

Accepted Manuscript

Title: More rapid reversal learning following overtraining in the rat is evidence that behavioural and cognitive flexibility are dissociable

Authors: Sandeep S. Dhawan, David S. Tait, Verity J. Brown

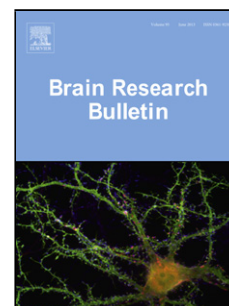
PII: S0166-4328(18)31568-7
DOI: <https://doi.org/10.1016/j.bbr.2019.01.055>
Reference: BBR 11779

To appear in: *Behavioural Brain Research*

Received date: 6 November 2018
Revised date: 29 January 2019
Accepted date: 29 January 2019

Please cite this article as: Dhawan SS, Tait DS, Brown VJ, More rapid reversal learning following overtraining in the rat is evidence that behavioural and cognitive flexibility are dissociable, *Behavioural Brain Research* (2019), <https://doi.org/10.1016/j.bbr.2019.01.055>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



More rapid reversal learning following overtraining in the rat is evidence that behavioural and cognitive flexibility are dissociable.

Dhawan, Sandeep S., Tait, David S. and Brown, Verity J.

School of Psychology & Neuroscience, University of St Andrews, St Mary's Quad, South Street, St Andrews, KY16 9JP, U.K.

Corresponding Author: Verity J Brown (<https://orcid.org/0000-0001-5762-1797>)

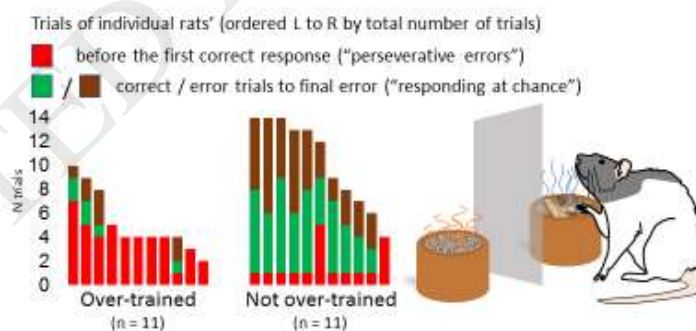
School of Psychology & Neuroscience, University of St Andrews

St Mary's Quad, South Street, St Andrews, KY16 9JP, U.K.

vjb@st-andrews.ac.uk

Tel: +44 (0) 7792 383 103

Graphical abstract



Highlights

- Over-trained rats perform reversal learning more rapidly than rats trained to criterion only.

- The same mechanism that can result in more rapid reversing slows attentional shifting.
- Greater alacrity of reversal learning does not necessarily indicate greater cognitive flexibility.

Abstract

Cognitive flexibility is a term used to describe the brain processes underlying the phenomenon of adaptive change in behaviour in response to changed contingencies in the internal or external environment. Cognitive flexibility is often assessed in complex tasks measuring perceptual attentional shifting or response or task switching, but, arguably, reversal learning is a simple assay of cognitive flexibility. Reversal learning requires the detection of a changed outcome, the cessation of a previously-rewarded response and the selection of an alternative, previously-unrewarded, response. This study addressed the issue of the relationship between reversal learning and cognitive flexibility.

In a single testing session, rats completed a series of 2-alternative forced-choice discriminations between digging bowls. The bowls differed according to both the medium within the bowl and the odor of the bowl. Having learned which cue (one of the odors or one of the digging media) indicated the food-baited bowl, half the rats were given additional trials of “over-training”. To test reversal learning, the meaning of the cues predictive of reward/non-reward was then switched.

There was a robust effect of over-training, with over-trained rats performing reversal learning in fewer trials than rats trained to criterion only. The pattern of errors supported the hypothesis that *more rapid* reversing results from the formation of an attentional set. This is the same attentional mechanism that results in *less rapid* shifting or switching. We conclude that the behavioural flexibility demonstrated in reversal learning does not provide a scale on which cognitive flexibility can be measured.

Keywords: overtraining; schizophrenia; discrimination learning; ASST;

1 Introduction

Just as “*Everyone knows what attention is*”[1], everyone knows what cognitive flexibility is. We have all experienced the annoyance of inflexible cognitions, such as when a song is ‘stuck in our head’ or we cannot find the solution to a problem and persist with a tactic that evidently does not work. Conversely, we know the joy of when it serves us well and delivers eureka moments, solutions ‘pop’ into mind, or we manage to ‘think outside the box’. For these reasons, we can also appreciate just how devastating are the consequences of impaired cognitive flexibility, which is common in so many psychiatric and neurological disorders[2].

Although we can introspect on our own cognitive flexibility, the reliable evidence of it in other people or animals is their behavioural adaptability. Indeed, behavioural flexibility has been referred to as the ‘gold standard’ of evidence for complex cognition [3]. The early psychology literature provides many examples of how it can be measured [4–7], but at its most simple, behavioural flexibility is the “*modifiability of learned behaviour*”[8]. Exemplary of this is discrimination reversal learning, which has been called a “*pre-eminent test of cognitive flexibility*”[9]. Reversal learning is considered a particularly important measure of flexibility because of its translational value: it is readily measured in many different species. Species appropriate stimuli, responses and rewards can be used and behavioural flexibility is measured as the number of trials required to change responding in response to a change in feedback, providing a simple to understand scale that can be directly compared within and between species. However, we must not be seduced by the apparent elegant simplicity of this paradigm: reversal learning is paradoxically complex.

We have previously argued [10] that reversal learning may represent a different kind of flexibility than is typically measured in tasks involving perceptual shifting or task switching. Specifically, in many shifting and switching tasks, the construct of interest is cognitive preparedness, also known as ‘mental set’ [11] (e.g., a perceptual set, response set or learning set). Mental set can be thought of as a cognitive structure that is a component of a model in model-based learning (model-based learning uses “*representations of the environment, expectations, and*

prospective calculations to make cognitive predictions of future value" [12] in contrast to model-free learning, which "*progressively acquires cached estimates of the long-run values of circumstances and actions from retrospective experience.*" [12]).

Mental set does not refer to a group (for example, a group of stimuli or features of a perceptual dimension, like 'colours' or 'shapes', or a group of problem-solving strategies) but rather it refers to the idea that the cognitive system is set, or prepared, to perform particular mental operations or process particular information. It is the same meaning as when used by the starter in a race: "*on your mark, get set, Go!*" 'Getting set', or prepared, benefits performance when the preparation is appropriate and impedes performance when it is not. Nevertheless, it is important to note that efficiency is not the same as capability: if the starter simply shouted 'Go!' the race would still begin, albeit more slowly and, probably, more chaotically. Conversely, if the starter unexpectedly shouted 'No!', rather than 'Go!', the runners would struggle to inhibit their prepared response. Tasks designed to assess mental set typically look for costs (e.g., failure to inhibit responses) and benefits (e.g., faster or more accurate responses) of being inappropriately *versus* appropriately prepared, or having the right cognitive structure or model. In the ID/ED task [5], new learning is assessed in the context of two different states of cognitive preparedness: at the ID stage, where there is an appropriate mental set, it confers a *learning benefit*; and at the ED stage, where the mental set is inconsistent with the new learning, it results in a *learning cost*. The state of mental set cannot be directly measured, but rather it is inferred from the difference between its benefit (trials to learn at the ID stage) and its cost (trials to learn at the ED stage).

Reversal learning is different in that behavioural flexibility is directly measured because it is an inherent performance requirement. This does not necessarily mean that reversal learning occurs in a model-free context (see [13,14]). Indeed, it has been suggested [9] that, in addition to reinforcement learning [15], at least two model-based processes are involved, namely: 'knowing that reversals can occur' [16,17]; and having 'a representation of the task demands' [18,19]. This means that either, or both, model-free and model-based processes may be activated and contribute to reversal learning in any particular task variant [9].

It seems intuitively plausible that the demands on cognitive flexibility would be rather different in the case of model-based, compared to model-free, reversal learning. This might account for some of the paradoxical effects that have been reported in the reversal learning literature, such as unexpected improvements where manipulations had been expected to result in impairment of performance [20–27]. An example of this is the over-training reversal effect (ORE) [26], which is the paradoxical finding that rats given additional, post-criterion, training on a visual discrimination require fewer trials to learn the reversal than rats trained to criterion only. This demonstrates that reversal learning involves more than simply inhibiting a pre-potent response and learning a new response (see also [28]), as model-free learning would require. Over-training may serve to increase the salience of the predictive cues and/or reduce or eliminate the number of ‘error factors’ (*i.e.*, cues that might be predictive of reward but which are, in fact, irrelevant) [20]. This modulation of salience enables the solution of the credit assignment problem by inhibiting the processing of, or down-regulating attention to, irrelevant information [29]. In other words, reversal learning is more rapid because over-training modulates the relative salience of cues, which is a change in the state of cognitive preparedness, or mental set.

Patterns of errors can be used to infer whether an animal is perseverating with a previous solution or is seeking to find or maintain a new solution (e.g., [30]). An attentional account of the ORE predicts that with over-training there will be an increase in the number of perseverative responses at the outset of reversal learning and a reduced probability of responding to irrelevant cues, hence fewer trials of ‘at chance’ responding [31]. It has been noted that an ORE is less likely to be observed if the task is ‘easy’ (such as spatial discrimination learning) compared to more ‘difficult’ (such as visual discrimination learning) [31]. Crucially, however, ‘problem difficulty’ is not determined by the absolute number of errors or trials to criterion, but rather by the initial probability of responding to the relevant aspect of the stimulus [31,32]. For example, allocentric, or absolute, spatial location is a salient feature for a rat in a maze with stable extra-maze spatial cues and is likely to be a default response. Therefore, an allocentric (one place vs another) spatial discrimination is ‘easy’ for a rat. A visual discrimination in the same maze is more ‘difficult’ because a rat might first explore

other possibilities (*e.g.*, allocentric location) before learning about the relevant visual stimulus.

In the experiment reported here, we examined reversal learning and the ORE in a bowl-digging paradigm, exploiting a number of advantages of this approach to explore the processes involved. Rats spontaneously forage and will readily dig in bowls filled with various scented media to find food. They quickly learn to discriminate between bowls which differ by at least two stimulus attributes (or dimensions), either odor (scent added to the bowl) or haptic (the digging media) cues, but only one of which is correlated with reward. We have previously undertaken an analysis of the performance of 375 intact rats (control groups tested in 22 different experiments, by different experimenters, over 18 years) and confirmed that the odor and haptic dimensions are both learned very rapidly [33]. Although rapid learning implies this is an 'easy' task for the rat, the initial probability of responding to the relevant stimulus attribute is chance. Our task, therefore, dissociates difficulty of learning (defined by the number of trials or errors) from problem difficulty. A second advantage is that because the rat learns the discrimination in a small number of trials, the number of over-training trials required for an ORE should also be small. We know from previous work [33] that an attentional set typically forms within 30 trials of performing successive discriminations, suggesting that, if the ORE is an attentional phenomenon, 30 additional trials should be sufficient to generate it in a bowl-digging context. We know that rats will continue to perform bowl-digging discriminations and consume the food bait (pieces of cereal) for well over 100 discrete trials in a single testing session, so this number of over-training trials, additional to the 30 or so trials of the test discriminations, is comfortably within the satiety limit. Therefore, in a single session, rats first learned that one of two bowls contained bait, signalled by its odor or by the digging material. Half then received an additional 30 trials of overtraining before the cues were reversed. Finally, they learned a new discrimination, with novel stimuli: if odor had been the relevant cue, it was now the digging media and vice versa.

The specific prediction was that 30 trials of over-training would generate an ORE in this bowl-digging task, reducing the number of trials to criterion for reversal learning. The improved reversal learning performance should not generalise to new learning of different stimulus attributes that were not over-trained.

2 Materials and methods

2.1 Animals

Twenty-two female Lister hooded rats (bred in-house; University of St Andrews, from Charles River stock), were housed in pairs in 40 x 23 x 19 cm plastic cages. This was a convenience sample of female rats, but previous studies have reported that there is not a sex difference in normal rats tested in the ID/ED bowl-digging paradigm, which includes reversal learning [33–36]. Therefore, there is good reason to suppose that these results would generalize to male rats.

Testing was conducted during the ‘lights-on’ phase in a 12-h light/dark cycle (lights on at 7:00 am). Although rats are more active during the dark cycle, they have sleep-wake periods throughout both the light and dark cycle. Furthermore, they have other circadian pacemakers, which can be entrained to food-cues (e.g., activity increases in anticipation of food-rewarded T-maze testing in the lights-on period [37]). Given that rats appear to be more adaptable in this regard than their human testers, all testing was conducted in normal light levels.

Rats were maintained with controlled access to food, receiving 15-20 g of lab chow each day in the latter half of the light period, with water available *ad libitum*. This level of controlled access did not result in weight loss and all rats gained weight during the experiment, tracking the normal body-weight curve supplied by Charles River. We monitored body condition [38] but we did not keep a rat with *ad libitum* food as a reference point. The weight range of the group was 180-250 g at the start of the experiment and 233-328 g at the end.

The work was conducted under UK Home Office licenced authority (Animal (Scientific Procedures) Act 1986) and local ethical and welfare oversight.

2.2 Apparatus and training

The apparatus, which has been described previously [39,40], was a plastic cage (70cm (length) × 40cm (width) × 18cm (height)), with acrylic dividers forming three sections: a large ‘waiting’ area (approximately two thirds of the length), and two smaller adjacent ‘choice’ compartments (the remaining third of the length, with each

being half of the width). Acrylic panels were used selectively to block access to one or both of the choice compartments.

A bowl (internal diameter of 7cm, with 4cm vertical walls) was placed in each of the two choice compartments. Each was filled, to a depth of 0.5 cm from the top, with a different digging substrate and with a different odor mixed in.

To encourage the rats to retrieve food bait (one half of a Honey Loop; Kellogg, Manchester, UK) from digging bowls, they were initially given a bowl, filled with sawdust into which had been mixed six pieces of the cereal food reward, in their home-cage. The bowl was given at the same time as the daily lab chow and left in the cage overnight for the rats to explore and consume the food. However, as the rats were housed in pairs, we could not be certain both rats had consumed the cereal bait. Therefore, before testing, they were introduced individually to the apparatus and presented with bowls filled with sawdust. Bait was initially placed on the surface of the sawdust but, over trials, was progressively buried deeper in the sawdust until rats were reliably and promptly retrieving it (typically within six presentations). In the same session, the rats were then given two discrimination problems, in which they had to learn which one of two bowls was baited. The procedure for this training is described below. The first discrimination was between two odors mixed in sawdust; the second was between two (similar smelling but with different haptic properties) media. On a subsequent day, rats were given another two-choice discrimination, but this time the two bowls differed from each other both in odor and in digging media (i.e., they were presented with a pair of bowls which were compound stimuli: one bowl contained medium 1, scented with either odor 1 or 2, and the other contained medium 2, scented with the other odor). The rat first had to learn which stimulus (e.g., odor 1) reliably indicated that the bowl was baited and then the learning was followed by a reversal of the discrimination, where the other stimulus (e.g., odor 2) now indicated the location of the bait. The purpose of the training was to ensure the rats had experience discriminating the bowls and with reversal learning. Stimuli used during any of these training stages were novel and not used again.

The procedure for testing the rats in these and all subsequent discrimination learning (including the test discriminations described below) was the same. Specifically, the bowls were introduced, simultaneously, to the 'choice chambers'

while the rat was in the 'waiting area'. The location of the baited bowl was pseudo-randomly determined, with no more than three consecutive trials having the rewarded bowl in the same location. The barrier was lifted and rat was allowed to approach either or both bowls freely and touch and whisk it, including the surface of the digging media. If the rat dug in the baited bowl, the trial was recorded as correct and terminated. If the rat dug in un-baited bowl, it was allowed to continue digging and searching until it moved away from the bowl, whereupon the trial was recorded as incorrect and terminated. When the rat started to dig, a barrier was lowered to block access to the bowl in the other choice chamber. The exception to this was the first four trials (comprised of two presentations of each pair of bowls) in each stage, including reversal learning, where access to the other bowl was not blocked. The rat, having first dug in the incorrect bowl, was thus permitted to dig in and retrieve the bait from the alternative bowl. This is equivalent to the 'self-correcting' procedure used by [26]. These trials were still recorded as incorrect but the 'self-correction' (*i.e.*, digging in the other bowl) gave information to the rat. Criterion performance was six consecutive correct responses ($P = 0.0156$), which could include one or more of the first four trials.

2.3 Design and testing

In a single testing session, rats completed a series of three, two-choice discriminations, an example of which is shown in Figure 1. The rats learned that one of the features of the bowl (either its odor or the digging medium) indicated which was the baited bowl, and the other feature was irrelevant. When they had reached criterion, half of the rats ($n=11$) were given 30 additional trials of over-training before proceeding to the next stage, while the remaining rats ($n=11$) proceeded immediately. The next stage was a reversal of the discrimination, in which the meaning of the cues predictive of reward / non-reward were switched, to test for the effect of over-training. Finally, to confirm any effect of over-training was specifically an ORE rather than a non-specific learning or performance improvement, rats were given a new discrimination with novel compound stimuli (NCD), based on the other, previously irrelevant, dimension.

2.4 Data Analysis

The number of trials to criterion was analysed by repeated-measures analysis of variance (ANOVA; SPSS v. 23), with a within-subjects factor of stage (three levels: CD; Rev; NCD), and a between-subjects factor of Group (two levels: over-trained and not over-trained) and Dimension (odor and medium).

In addition to the effect of over-training on the absolute number of trials, we also looked in detail at responding in the reversal stage. We divided responding into two learning phases, representing retention of the previous learning and ‘at chance’ responding, in a manner similar to previously reported [41]. Trials prior to the first correct response (all of which were errors, of course) were classified as ‘perseverative’. Responses including and subsequent to the first correct response, and up to and including the error trial preceding the final 6-correct trials were classified as ‘at chance’ responding. These data were analysed together, with a within-subjects factor of Response-type (two levels: perseverative and chance) and a between-subjects factor of Group (two levels: over-trained and not over-trained). For all analyses, interactions were further examined using restricted analyses (with appropriate corrections) to investigate Simple Main Effects [42].

We also examined the effect of over-training on the probability of an error on the early trials of the reversal stage as a function of whether the rat dug in the first bowl encountered or sampled both bowls before choosing to dig, and analysed the proportion of responses using chi-squared tests. As the rats do not know the location of the correct bowl until they approach and sample it, whether they approach the baited bowl first is by chance. They are free to dig in the first bowl they encounter, but they may also sample the alternative bowl before deciding where to dig. If a rat encounters the previously correct bowl first and immediately digs, this could be due to a failure to inhibit a pre-potent, well-trained response. However, it is more effortful to actively seek an alternative bowl. Therefore, if the rat did not dig in the first bowl encountered, but rather sought out the newly incorrect bowl, this would provide strong evidence that the response was perseverative. With the same reasoning, the converse should apply for correct trials: if the rat sought the correct bowl, this would

suggest that it had not merely dug in the first bowl and had a 'lucky encounter' (see Table 1).

3 Results

3.1 Trials to criterion

3.1.1 Discrimination learning (CD stage)

Rats in both groups attained criterion in the initial discrimination in around 11 trials (over-trained: $M = 11.5$, $SEM = 1.56$; not over-trained: $M = 10.7$, $SEM = 1.54$). Unexpectedly, we found that rats learned the medium discrimination in significantly fewer trials ($M = 9.33$, $SEM = 0.74$) than the odor discrimination ($M = 13.3$, $SEM = 1.9$) (omnibus analysis: Stage by Group by Dimension: $F_{(2,36)} = 4.68$, $p = 0.02$; $\eta_p^2 = 0.21$; simple main effect of Dimension, CD stage: $F_{(1,36)} = 7.09$, $p = 0.01$; $\eta_p^2 = 0.18$) (see Figure 2).

Of the 11 rats that were given 30 additional post-criterion trials, 9 made no errors in these overtraining trials; one rat made a single error (at the 27th trial of overtraining); and another made 2 errors (on the 5th and 17th trials of overtraining).

3.1.2 Reversal

Rats with over-training reached criterion at the reversal stage in significantly fewer trials ($M = 11.2$, $SEM = 0.78$) than rats trained to criterion only ($M = 16.5$, $SEM = 1.15$) confirming the expected ORE, with a large effect size (simple main effect of Group, Reversal stage: $F_{(1,36)} = 13.39$, $p < 0.001$; $\eta_p^2 = 0.595$; see Figure 2).

As with the initial acquisition, odor reversals required more trials than medium reversals (odor: $M = 15.9$ ($SEM = 1.38$); medium: $M = 12.1$ ($SEM = 0.69$)) (simple main effect of Dimension, Reversal stage: $F_{(1,36)} = 6.57$, $p < 0.015$; $\eta_p^2 = 0.42$). However, this effect of discrimination type was irrespective of whether the rats were over-trained (simple interaction of Group by Dimension, Reversal stage: $F_{(1,36)} = 1.71$, $p = 0.2$ (n.s.) $\eta_p^2 = 0.16$).

3.1.3 *New Learning*

The effect of the stimulus (odor or medium) on discrimination learning was also seen in the new discrimination, but only for the haptic stimuli (simple interaction of Group by Dimension, NCD learning stage: $F_{(1,36)} = 9.13$, $p = 0.005$; $\eta_p^2 = 0.19$; see Figure 2). The rats that had been over-trained on the odor discrimination – and consequently showed improved reversal learning – required significantly more trials than the rats not over-trained to learn the subsequent media discrimination ($M = 14.5$, $SEM = 2.46$ (over-trained), compared to $M = 9.4$, $SEM = 0.81$ (not over-trained)). This pattern confirms that the improved reversal learning performance after over-training does not generalize and improve performance of a subsequent acquisition.

3.2 *Response types in the reversal stage*

Because the reversal was un-cued, it was unsurprising that all of the rats responded on the first trial in a manner consistent with the reward history and the first trial of the reversal was an error (classified as perseverative). Some rats ($n=8$) made no further errors after they had made a correct response and so these rats had no ‘at chance’ responses. However, for those rats ($n = 14$) who did make errors after the first correct response, the proportion of trials that were errors was 0.51, suggesting that these responses were indeed ‘at chance’.

The over-trained rats made significantly more perseverative responses ($M = 3.9$, $SEM = 0.48$; range: 1 - 7) compared to the rats not over-trained ($M = 1.6$, $SEM = 0.42$; range: 1 - 3). The pattern was opposite for ‘at chance’ responding: the over-trained rats made significantly fewer responses ‘at chance’ ($M = 1.2$, $SEM = 0.54$; range: 0 - 3) compared to the rats not over-trained ($M = 8.7$, $SEM = 4.2$; range: 0 - 13). Of the 11 over-trained rats, 7 (64%) made no errors after the first correct response compared to only 1 (9%) of the 11 rats not over-trained (Figure 3). This pattern was confirmed by a significant interaction of Response type by Group: $F_{(1,20)} = 33.21$, $p < 0.01$; $\eta_p^2 = 0.62$; Bonferroni-corrected pairwise comparisons $p < 0.01$ for main effect of Group for each response type).

3.3 *Responses in the early trials of the reversal*

As noted above (Section 2.4), the rats do not know the location of the correct bowl until they approach it. They may dig in the first bowl approached or they may

sample the alternative bowl before deciding where to dig. On the very first trial of the reversal, all of the rats made an error and dug in the un-baited bowl. Almost half of them (9/22) encountered, by chance, the previously correct bowl first and they dug in it without sampling the other bowl. The remaining rats (13/22) encountered the newly correct bowl, but then sampled and elected to dig in the alternative (previously correct) bowl. This is entirely as expected: it demonstrates that, if the rat did not encounter by chance the bowl it expected to be baited, it would actively seek out that bowl.

The probability of a correct response on the subsequent three trials (*i.e.*, the 2nd, 3rd and 4th trials of the reversal) and whether the rat dug in the first bowl it encountered or whether it sampled both of the bowls before making a choice to dig were examined to see if the rats continued to actively seek out the previously reinforced bowl. As expected, about half of the first approaches were to the newly correct bowl on these trials (32 approaches in 66 trials; 48%). However, rats in the two groups were not equally likely to dig: the over-trained rats moved to the alternative (previously correct) bowl on 82% (14/17) of approaches, which is all the more remarkable given that we were using a self-correction procedure, but confirms the observations of [26], who also used a self-correction procedure. By contrast, the rats not over-trained only moved to the previously correct bowl on 30% (5/15) of approaches (Figure 4). This difference between the groups in the likelihood that they would persevere in seeking out the previously correct bowl was statistically reliable, with a moderate to large effect size ($\chi^2 = 7.94$, $df = 1$, $p = 0.005$, $\Phi = 0.498$).

Being more likely to 'accept' the newly-correct bowl when encountered first might imply that the rats not over-trained were learning more rapidly. However, they were as likely as the over-trained rats also to 'accept' the un-baited (previously correct) bowl when *that* was the first bowl they encountered ($\chi^2 = 1.7$, $df = 1$, *n.s.*, $\Phi = 0.225$).

4 Discussion

The aim of this study was to examine the nature of cognitive structures in the rat, using reversal learning as a case-study, and looking specifically at over-training

and the resultant ORE to see if it is possible to tease apart underlying processes and cognitive structures in reversal learning. We trained rats to discriminate bowls based on their odor or haptic properties, before giving half of them 30 additional trials (over-training).

Over-trained rats showed enhanced reversal learning, which is the over-training reversal effect, or ORE [26]. The effect of overtraining accounted for >59% of the variance and therefore, although amounting to a benefit in this task of a small number of trials (~5), overtraining had a large, as well as statistically significant, impact on performance. Obviously, no one has or would suggest that the absolute number of trials to criterion in any particular reversal learning task is indicative of flexibility *per se*: treatment groups are compared within tasks because different tasks or testing paradigms result in differing absolute number of trials to criterion. Even a seemingly small change (such as the inclusion of an additional stage or stimulus dimension, as in [39], or the difference between stimuli, as we observed here), can have an effect on the absolute number of trials for reversal learning. Nevertheless, the effect of overtraining was in spite of absolute differences in trials to criterion for odor and medium stimuli, thus demonstrating that the ORE is not influenced by ease of learning, but rather by problem difficulty [20].

The decreased trials to criterion at the reversals was seen in the context of a significant *increase* in perseverative responding (persistence of choosing the previously correct bowl) but a *greater decrease* in the number of trials in which responding was at chance. This pattern was predicted [20] and is consistent with the suggestion that the ORE reflects greater attentional focus on the relevant stimulus attributes and the down-regulation of attention to stimulus attributes that might have been relevant for predicting reward, but are in fact irrelevant (i.e., error factors).

One surprising aspect of the results was that discriminating between odors required more trials to learn and reverse than discriminating between media. Although we have previously reported [43] that *reversals* of odor discriminations require more trials, the effect size in that large sample was small and a power analysis confirmed a sample of >200 rats would be required. It is possible that a difference arose due to the particular odors and media used in this experiment. Nevertheless, importantly, the discrimination type (i.e., odor or medium) did not significantly change

the magnitude of the ORE. It did have an effect on learning a new discrimination that involved a shift of attention to a previously irrelevant feature. Rats that had been over-trained required more trials to learn the final novel discrimination compared to the initial discrimination, regardless of the discrimination type. This would be expected if the ORE arises from an attentional focus: the initially trained and then over-trained dimension means that the novel discrimination requires an extra-dimensional shift of attention. Rats that had not been over-trained showed no learning decrement for the medium discrimination, acquiring it in as few trials as required by rats performing a medium discrimination in the first stage.

It is intuitively plausible, that reversal learning comprises two parts (*e.g.*, see [20]). First, the previous learning must extinguish: in this initial stage, there is a strong tendency to respond according to prior reinforcement history, and error responses are therefore perseverative and these responses need to be inhibited. Second, new learning must occur: in this stage, responses are likely to be correct only by chance, prior to the formation of a new association. Because learning in the bowl-digging paradigm is so rapid, it is often not possible to identify the point at which responding ceases to be perseverative and is at chance. However, over-training had a differential effect on the pattern of these different kinds of responses. Specifically, over-training resulted in an increase in perseverative responses, with rats more likely to persist in seeking the previously correct bowl, and a decrease in responding at chance. More than half of the over-trained rats made no further errors after their first correct response (*i.e.*, there was no 'at chance' responding) and the rest made very few errors. By contrast, the rats not over-trained made very few perseverative errors, quickly adopting a new strategy of digging in the first bowl they encountered, consequently responding at chance, until they learned the new association.

As discussed above, reversal learning tasks can depend solely on model-free learning, but they may also involve model-based processes [9,44]. We suggested (Introduction) that mental set is an example of a model for model-based learning. Prior experience of one dimension enhances the learning of a new discrimination involving the same dimension and slows learning when another stimulus feature becomes relevant. Overtraining contributes to the formation of mental set – the model – and it is this model that results in more rapid reversal learning. Most

importantly, if reversal learning in this task were accomplished by model-free associative learning, by adding to the reinforcement history overtraining would increase the trials to criteria for a reversal.

One of the processes identified as a potential component in model-based reversal learning is that 'reversals can occur' [9]. This is a reasonable expectation for serial reversal learning: over successive reversals errors reduce because the animal appears to learn a win-stay/lose-shift rule [7]. The first error, or unexpected non-reward, is the cue that a reversal has occurred and the lose-shift rule is invoked. Clearly, this model is not relevant in the case of the ORE. We make this observation to re-emphasise the point that, in different circumstances, model-based reversal learning may involve different types of model.

At this point, it is as well to remind ourselves that "*when philosophers and psychologists speak of 'cognitive processes' and 'cognitive structures' it is not clear what they are talking about*" ([45]p.385). Even putting the term 'cognitive' aside, and thinking of very specific examples of cognition, the meaning of 'processes' and 'structures' is often still obscure [44]. One of the issues is that processes are often invoked in a manner that essentially describes the behaviour and, as such, they do not offer explanations. This is not to say that there is no value in a careful analysis and description of behaviour, but we must be cautious and not think that in naming something, we have explained it. It is also as well to remember that simple systems can result in surprisingly complex outputs: for example, it has been suggested that the 'reward prediction error' systems in the brain may be smarter than often assumed [44].

In this paper, we have used reversal learning as a specific example to examine cognitive processes that have been postulated to underlie cognitive flexibility. We have previously suggested that shifting and switching may be different forms of flexibility compared to reversal learning [10]. Nevertheless, notwithstanding the obvious differences between tasks, under some conditions performance of all of them may be modulated according to the regulation of attention to cues predictive of reward. Thus, a cognitive structure (mental set) is the product of an attentional process and this is common to all these tasks under particular experimental conditions. The real paradox of the ORE is that the *improvement* in the rate of reversal learning

arises from the same attentional mechanism or model that accounts for the *decrement* in performance in the ED stage of a shifting task or when a task requires a response switch. Of course, this is not so paradoxical if one considers that what affects rate of learning (including reversal learning) is the appropriateness of the mental set. The ORE is evidence of the benefit of a mental set and it is thus akin to the performance improvement seen at the ID stage of an ID/ED task. Nevertheless, it necessarily follows that the rate of reversal learning cannot, *per se*, be indicative of the capacity for cognitive flexibility.

Declarations of interest: none

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The research was undertaken by SSD, in partial fulfillment of the requirements of a PhD degree under the joint supervision of VJB and DST and supported by The University of St Andrews (QR Block Grant).

References

- [1] W. James, *The principles of psychology*, Vol I., 1890. doi:10.1037/10538-000.
- [2] L.A. Gunaydin, A.C. Kreitzer, Cortico–Basal Ganglia Circuit Function in Psychiatric Disease, *Annu. Rev. Physiol.* 78 (2016) 327–350. doi:10.1146/annurev-physiol-021115-105355.
- [3] I. Mikhalevich, R. Powell, C. Logan, Is behavioural flexibility evidence of cognitive complexity? How evolution can inform comparative cognition, *Interface Focus*. 7 (2017) 20160121. doi:10.1098/rsfs.2016.0121.
- [4] E.A. Berg, A simple objective technique for measuring flexibility in thinking, *J. Gen. Psychol.* 39 (1948) 15–22.
- [5] D.H. Lawrence, Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus, *J. Exp. Psychol.* 39 (1949) 770–784. doi:10.1037/h0058097.
- [6] A.T.A. Jersild, *Mental set and shift*, 1927. doi:not indicated.
- [7] H.F. Harlow, The formation of learning sets, *Psychol. Rev.* 56 (1949) 51–65. doi:10.1037/h0062474.
- [8] R.M. Yerkes, *The dancing mouse; a study in animal behavior*, The Macmillan company, New York, 1907. <https://www.biodiversitylibrary.org/item/15698>.
- [9] A. Izquierdo, J.L. Brigman, A.K. Radke, P.H. Rudebeck, A. Holmes, The neural basis of reversal learning: An updated perspective, *Neuroscience*. 345 (2017) 12–26. doi:10.1016/j.neuroscience.2016.03.021.
- [10] V.J. Brown, D.S. Tait, Behavioral Flexibility: Attentional Shifting, Rule Switching, and Response Reversal, *Encycl. Psychopharmacol.* (2015) 264–269. doi:10.1007/978-3-642-36172-2_340.
- [11] J.J. Gibson, A critical review of the concept of set in contemporary experimental psychology, *Psychol. Bull.* 38 (1941) 781–817. doi:10.1037/h0055307.
- [12] P. Dayan, K.C. Berridge, Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation, *Cogn. Affect. Behav. Neurosci.* 14 (2014) 473–492. doi:10.3758/s13415-014-0277-8.
- [13] G. Schoenbaum, M. Roesch, Orbitofrontal cortex, associative learning, and expectancies, *Neuron*. 47 (2005) 633–636. doi:10.1016/j.neuron.2005.07.018.
- [14] G. Schoenbaum, M.P. Saddoris, T.A. Stalnaker, Reconciling the roles of

- orbitofrontal cortex in reversal learning and the encoding of outcome expectancies, *Ann. N. Y. Acad. Sci.* 1121 (2007) 320–335. doi:10.1196/annals.1401.001.
- [15] T.A. Stalnaker, T.M. Franz, T. Singh, G. Schoenbaum, Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments, *Neuron*. 54 (2007) 51–58. doi:10.1016/j.neuron.2007.02.014.
- [16] V.D. Costa, V.L. Tran, J. Turchi, B.B. Averbeck, Reversal Learning and Dopamine : A Bayesian Perspective, *J. Neurosci.* 35 (2015) 2407–2416. doi:10.1523/JNEUROSCI.1989-14.2015.
- [17] A.I. Jang, V.D. Costa, P.H. Rudebeck, Y. Chudasama, E.A. Murray, B.B. Averbeck, The role of frontal cortical and medial-temporal lobe brain areas in learning a Bayesian prior belief on reversals, *J. Neurosci.* 35 (2015) 11751–11760. doi:10.1523/JNEUROSCI.1594-15.2015.
- [18] R.C. Wilson, Y.K. Takahashi, G. Schoenbaum, Y. Niv, Orbitofrontal cortex as a cognitive map of task space, *Neuron*. 81 (2014) 267–278. doi:10.1016/j.neuron.2013.11.005.
- [19] A. Saez, M. Rigotti, S. Ostojic, S. Fusi, C. Salzman, Abstract context representations in primate amygdala and prefrontal cortex, *Neuron*. 87 (2015) 869–881. doi:10.1016/j.neuron.2015.07.024.
- [20] N.S. Sutherland, N.J. Mackintosh, Mechanisms of animal discrimination learning, Academic Press Inc, New York, 1971.
- [21] S.M. Zola, H. Mahut, Paradoxical facilitation of object reversal learning after transection of the fornix in monkeys, *Neuropsychologia*. 11 (1973) 271–284. doi:10.1016/0028-3932(73)90038-9.
- [22] A. Krazem, D. Béracochéa, R. Jaffard, Effects of mammillary bodies and mediodorsal thalamic lesions on the acquisition and retention of a learning set in mice: paradoxical effect of the intersession interval, *Behav. Brain Res.* 67 (1995) 51–58. doi:10.1016/0166-4328(94)00103-M.
- [23] V. Boulougouris, J.W. Dalley, T.W. Robbins, Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat, *Behav. Brain Res.* 179 (2007) 219–228. doi:10.1016/j.bbr.2007.02.005.
- [24] C. Graybeal, M. Feyder, E. Schulman, L.M. Saksida, J. Bussey, Timothy, J.L.

- Brigman, A. Holmes, Paradoxical reversal learning enhancement by stress or prefrontal cortical damage : rescue with BDNF, *Nat. Neurosci.* 14 (2012) 1507–1509. doi:10.1038/nn.2954.
- [25] H. Eichenbaum, A. Fagan, N.J. Cohen, Normal olfactory discrimination learning set and facilitation of reversal learning after medial-temporal damage in rats: implications for an account of preserved learning abilities in amnesia., *J. Neurosci.* 6 (1986) 1876–1884.
- [26] L.S. Reid, The development of noncontinuity behavior through continuity learning., *J. Exp. Psychol.* 46 (1953) 107–112. doi:10.1037/h0062488.
- [27] J.O. van Horik, N.J. Emery, Serial reversal learning and cognitive flexibility in two species of Neotropical parrots (*Diopsittaca nobilis* and *Pionites melanocephala*), *Behav. Processes.* (2018) in press. doi:10.1016/j.beproc.2018.04.002.
- [28] S.R.O. Nilsson, J. Alsiö, E.M. Somerville, P.G. Clifton, The rat's not for turning: Dissociating the psychological components of cognitive inflexibility, *Neurosci. Biobehav. Rev.* 56 (2015) 1–14. doi:10.1016/j.neubiorev.2015.06.015.
- [29] R. Akaishi, N. Kolling, J.W. Brown, M. Rushworth, D. Op, Neural Mechanisms of Credit Assignment in a Multicue Environment, 36 (2016) 1096–1112. doi:10.1523/JNEUROSCI.3159-15.2016.
- [30] S.B. Floresco, Dissociable Roles for the Nucleus Accumbens Core and Shell in Regulating Set Shifting, *J. Neurosci.* 26 (2006) 2449–2457. doi:10.1523/JNEUROSCI.4431-05.2006.
- [31] N.J. Mackintosh, Selective attention in animal discrimination learning, *Psychol. Bull.* 64 (1965) 124–150. doi:10.1037/h0022347.
- [32] E. Lovejoy, Analysis of the overlearning reversal effect, *Psychol. Rev.* 73 (1966) 87–103.
- [33] D.S. Tait, E.M. Bowman, L.S. Neuwirth, V.J. Brown, Assessment of intradimensional/extradimensional attentional set-shifting in rats, *Neurosci. Biobehav. Rev.* 89 (2018) 72–84. doi:10.1016/j.neubiorev.2018.02.013.
- [34] W.M.Y. Mohamed, E.L. Unger, S.K. Kambhampati, B.C. Jones, Methylphenidate improves cognitive deficits produced by infantile iron deficiency in rats, *Behav. Brain Res.* 216 (2011) 146–152. doi:10.1016/j.bbr.2010.07.025.
- [35] K.L. Murphy, J. McGaughy, P.L. Croxson, M.G. Baxter, Exposure to sevoflurane

- anesthesia during development does not impair aspects of attention during adulthood in rats, *Neurotoxicol. Teratol.* 60 (2017) 87–94. doi:10.1016/j.ntt.2016.11.010.
- [36] S. Snigdha, J.C. Neill, S.L. McLean, G.K. Shemar, L. Cruise, M. Shahid, B. Henry, Phencyclidine (PCP)-induced disruption in cognitive performance is gender-specific and associated with a reduction in Brain-Derived Neurotrophic Factor (BDNF) in specific regions of the female rat brain, *J. Mol. Neurosci.* 43 (2011) 337–345. doi:10.1007/s12031-010-9447-5.
- [37] R.E. Mistlberger, M.H.M. De Groot, J.M. Bossert, E.G. Marchant, Discrimination of circadian phase in intact and suprachiasmatic nuclei-ablated rats, *Brain Res.* 739 (1996) 12–18. doi:10.1016/S0006-8993(96)00466-0.
- [38] D.L. Hickman, M. Swan, Use of a Body Condition Score Technique to Assess Health Status in a Rat Model of Polycystic Kidney Disease, *J. Am. Assoc. Lab. Anim. Sci.* 49 (2010) 155–159. doi:10.1021/ol200196m.
- [39] J.M. Birrell, V.J. Brown, Medial frontal cortex mediates perceptual attentional set shifting in the rat, *J. Neurosci.* 20 (2000) 4320–4324.
- [40] D.S.D.S. Tait, E.A.A. Chase, V.J.V.J. Brown, Attentional set-shifting in rodents: A review of behavioural methods and pharmacological results, *Curr. Pharm. Des.* 20 (2014) 5046–59. doi:10.2174/1381612819666131216115802.
- [41] B. Jones, M. Mishkin, Limbic lesions and the problem of stimulus—Reinforcement associations, *Exp. Neurol.* 36 (1972) 362–377. doi:10.1016/0014-4886(72)90030-1.
- [42] B. Winer, *Statistical principles in experimental design*, McGraw-Hill, New York, 1971.
- [43] D.S. Tait, E.M. Bowman, L.S. Neuwirth, V.J. Brown, Assessment of intradimensional/extradimensional attentional set-shifting in rats, *Neurosci. Biobehav. Rev.* 89 (2018) 72–84. doi:10.1016/j.neubiorev.2018.02.013.
- [44] B.B. Doll, D.A. Simon, N.D. Daw, The ubiquity of model-based reinforcement learning, *Curr. Opin. Neurobiol.* 22 (2012) 1075–1081. doi:10.1016/j.conb.2012.08.003.
- [45] N. Malcolm, The myth of cognitive processes and structures, in: T. Mischel (Ed.), *Cogn. Dev. Epistemol.*, 1st ed., Academic Press, 1971: pp. 385–392.

Figure legends

Figure 1: Schematic representation of the task. The rats were presented with a pair of bowls containing coarse or fine tea ('media') scented with ginger or cinnamon ('odor'). In the example, the bait (a piece of cereal, represented by the 'cut out' corner) is initially in the bowl with the odor of ginger (signified by upper case bold), but for half the rats it was associated with one of the media. In the reversal, the same pairs of bowls are used, but the bait is now associated with the alternative odor (or the alternative medium). For the novel compound discrimination, two new pairs of bowls are used, containing sand or grit scented with sage or paprika. Rats who had discriminated odors now had to discriminate between the media and vice versa. The criterion for learning was six consecutive correct trials.

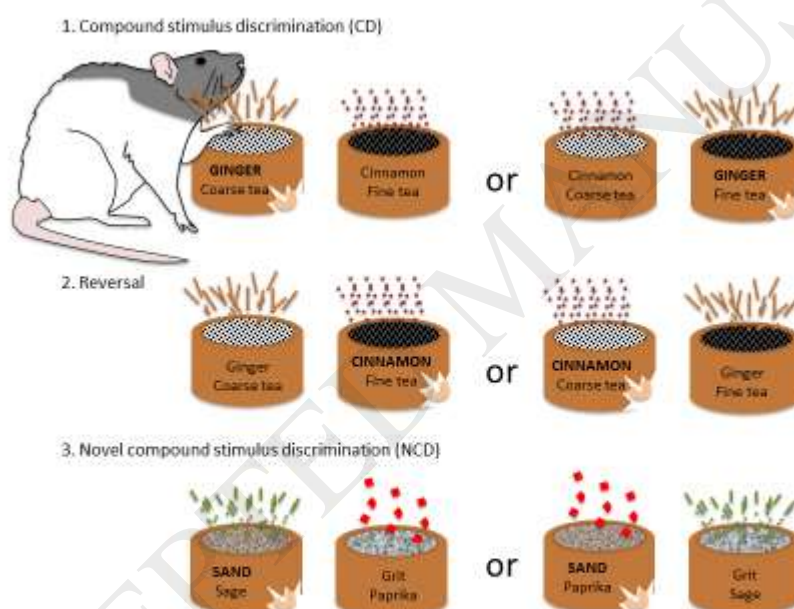


Figure 2: Mean trials to criterion (\pm SEM) for each stage of test, plotted by whether the initial discrimination (CD) was between odors or media. Thirty additional trials of over-training of the initial CD facilitated subsequent reversals of both odor and medium discriminations. Overall, odor discriminations required more trials to learn than media discriminations, except for rats over-trained on the odor discrimination: they failed to show better subsequent discrimination learning of the media. * $p < 0.05$.

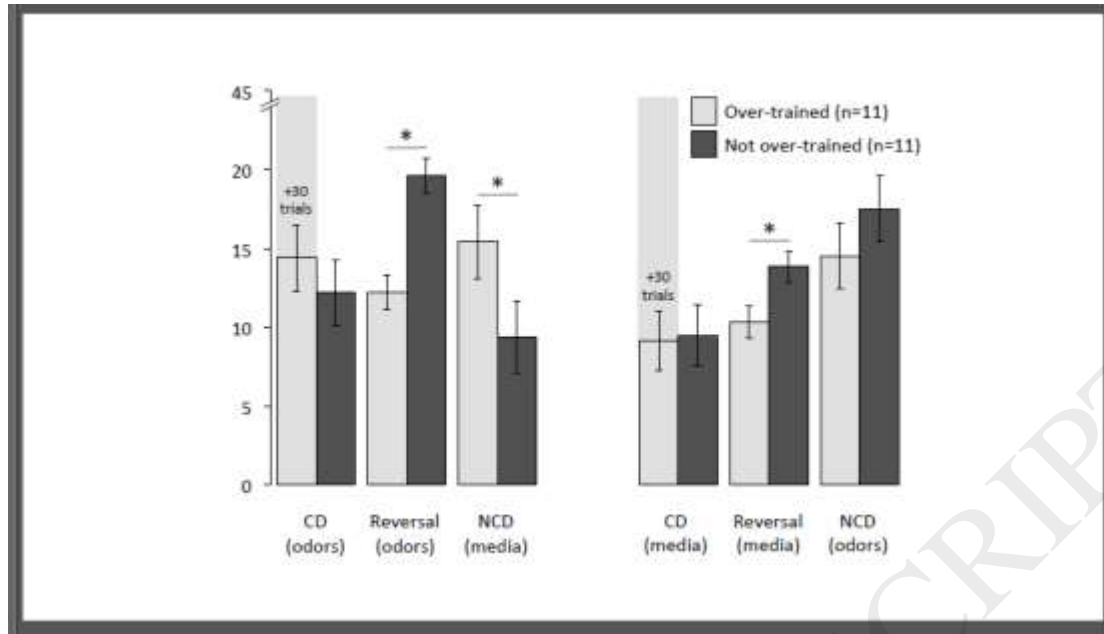


Figure 3: In the reversal learning stage, error responses prior to the first correct response were considered 'perseverative'. Responses from the first correct response to the last error response, inclusive, were considered to be 'at chance' responding. Over-trained rats made significantly more perseverative responses, but made significantly fewer responses 'at chance', than rats trained to criterion only. * $p < 0.05$.

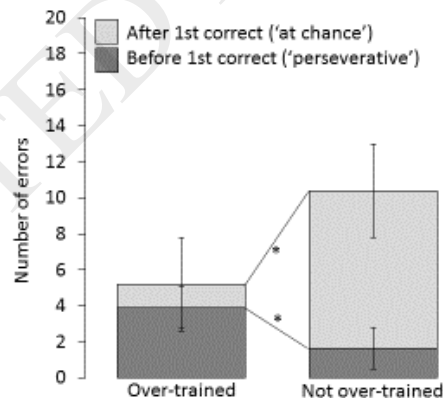


Figure 4: Responses on the 2nd, 3rd and 4th trials of the reversal stage for each group of 11 rats (i.e., 33 responses per group) were categorized according to whether the rat dug in the first bowl encountered without sampling the other bowl ('accept'; solid) or where it was sampled and rejected in favour of the alternative ('reject'; hashed).

Correct responses (i.e., accept the baited or reject the un-baited) are shown in white; incorrect responses (i.e., reject the baited or accept the un-baited) are shown in grey. The rats not over-trained were more likely to 'accept' the first bowl they encountered, while the over-trained rats sought out the previously rewarded bowl (i.e., perseverative responding).

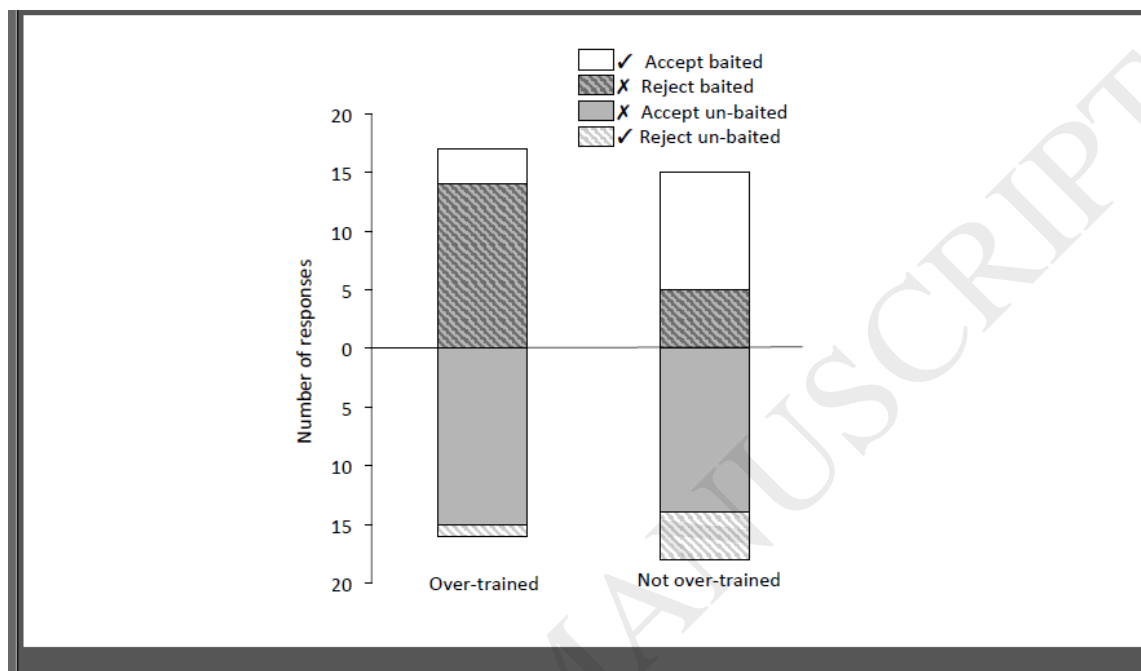


Table 1: A rat could dig in the first bowl encountered or could sample both bowls before deciding in which to dig. If the rat accepts the first bowl, the meaning of the choice is ambiguous. However, there is greater confidence in the interpretation if the rat samples both bowls: the seeking behavior suggests that it was either correctly rejecting the un-baited bowl or it was incorrectly rejecting and perseverating.

	<i>Dig in first bowl approached ("accept")</i>	<i>Sample both bowls before digging ("reject")</i>
<i>Dig in baited bowl ("correct")</i>	Made a lucky guess <u>or</u> knew it was correct?	Reject the unbaited bowl and seek out the baited bowl: correct response
<i>Dig in un-baited bowl ("incorrect")</i>	Made an unlucky guess <u>or</u> Failed to inhibit a pre-potent response <u>or</u> thought it was correct	Reject the baited bowl in favor of seeking out the previously-baited bowl: perseverative response