



Stereotypic route-tracing in experimentally caged songbirds correlates with general behavioural disinhibition

JOSEPH P. GARNER*, GEORGIA J. MASON† & REBECCA SMITH†

*Department of Animal Science, University of California, Davis

†Department of Zoology, University of Oxford

(Received 17 April 2002; initial acceptance 19 June 2002;
final acceptance 24 October 2002; MS. number: A9335R)

Repetitive, unvarying and apparently functionless behaviours called stereotypies are common in caged animals, but the mechanisms of cage stereotypy have remained elusive. We found that stereotypies correlate with a sign of altered brain functioning, the general disinhibition of behaviour, found in stereotyping human patients and animals treated with psychostimulants. We investigated route-tracing stereotypy in blue tits, *Parus caeruleus*, and marsh tits, *P. palustris*, caged in a behavioural research laboratory. In experiment 1, stereotypy correlated with disinhibition of responses, shown as persistent responding in extinction learning. In experiment 2a, stereotypy in blue tits correlated with repetitive sequence generation on a 'gambling task' that specifically measures disinhibited striatal functioning in humans and correlates with stereotypy in autistic and schizophrenic patients. In experiment 2b, the behavioural disinhibition correlated with stereotypy also affected food-storing behaviour in marsh tits, in particular their response to cache pilfering. In experiment 3, the sequencing of nonstereotypic home cage behaviour was correlated with stereotypy. Finally, changes in stereotypy induced by environmental enrichment correlated with changes in these measures of altered behavioural control. These results suggest that housing conditions that cause stereotypy thus alter many aspects of the behavioural control of caged subjects, mediated by altered striatal functioning. The implications for understanding cage stereotypies in laboratory, farm and zoo animals, and for laboratory-based behavioural experiments, are discussed. We suggest that improving housing conditions so that cage stereotypies do not develop would enhance the validity of laboratory-based behavioural research.

© 2003 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Stereotypies, which are abnormal repetitive, unvarying and apparently functionless behaviours, are perhaps the most common abnormal behaviour in captive animals. For example, stereotypic bar chewing is performed by at least 15 million intensively housed sows worldwide (calculated from data in: European Commission 1997), object pecking by over 56 million poultry on broiler breeder farms (calculated from data in: European Commission 2000), bar mouthing by approximately 10 million laboratory mice (calculated from data in: University of California Center for Animal Alternatives 1996; Würbel et al. 1996; European Commission 1999; Canadian Council on Animal Care 2000; National Animal Ethics Advisory Committee 2002) and wind-sucking, cribbing and weaving by 1.5 million horses in the U.S.A. alone (calculated from: McGreevy et al. 1995; J. Roser,

Correspondence: J. P. Garner, Department of Animal Science, University of California, One Shields Avenue, Davis, CA 95616, U.S.A. (email: jpgarner@ucdavis.edu). G. J. Mason is at the Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, U.K. R. Smith can be contacted c/o G. J. Mason.

personal communication). Stereotypies are mostly absent in the wild, relatively infrequent in large, environmentally enriched enclosures, and most common in barren and/or physically restricting cages (Mason 1991a). They are therefore often considered to indicate poor welfare (Berkson 1968; Mason 1991b, 2001; Wiedenmayer 1996, 1997).

In response, ethologists have focused on the motivational bases of cage stereotypies (e.g. Rushen et al. 1993), and the specific motivational origins of particular stereotypies are now well understood (e.g. Wiedenmayer 1997; Nevison et al. 1999; Nicol 1999). However, although focusing on individual species has led to some improvements in husbandry, a deeper mechanistic understanding of these important behaviours has been lacking. Furthermore, traditional approaches, which implicitly consider stereotypies the product of placing a normal animal in an abnormal environment, cannot account for all features of the behaviour (e.g. Mason 1991b, 1993). For example, individuals with cage stereotypies sometimes show other odd behavioural

traits, such as harming their own young (e.g. bank voles, *Clethrionomys glareolus*: Sørensen & Randrup 1986), making poor-quality nests (e.g. farmed mink, *Mustela vison*: Mason et al. 1995), or showing enhanced locomotor responsiveness to amphetamine (e.g. tethered sows: Terlow et al. 1992). Some cage stereotypies persist even when they harm the performer (e.g. pacing zoo carnivores, who may suffer sores and abrasions from repeatedly moving against their enclosure walls: Fentress 1976). Furthermore, with longer periods in captivity, some stereotypies are difficult to cure with environmental enrichments (e.g. bank voles: Cooper et al. 1996).

In the fields of experimental psychology, clinical psychology and psychiatry, in contrast, stereotypies in humans indicate well-understood types of behavioural dysfunction. Repetitive behaviours in general stem from a range of central nervous system deficits (e.g. Luria 1965; Frith & Done 1990; Turner 1997), and stereotypies seem specifically linked to dysfunction in the dorsal striatum and its efferent pathways (e.g. Lewis et al. 1996). Thus, drug- and isolation-induced stereotypies (e.g. Berkson 1968; Morgan et al. 1975; Hauber 1998), and many stereotypies, tics and choreas in human pathology (Owens et al. 1982; Frith & Done 1983; Albin et al. 1989; Crider 1997; Turner 1997), seem to reflect altered functioning of the cortico-striatal-pallido-thalamo motor loop (Alexander et al. 1990; Rolls 1994). Within this loop, the dorsal striatum mediates the control of behaviour via two efferent pathways. The first, 'direct pathway', activates behaviour chains and individual motor movements in a recursive manner. The second, 'indirect pathway' inhibits competing plans and movements, thus mediating the smooth transition between elements (Norman & Shallice 1986; Albin et al. 1989; Alexander et al. 1990; Rolls 1994). Suppression of the indirect pathway, which results in a general state of behavioural disinhibition, seems to occur in all stereotypies for which a neurological basis is known (Luria 1965; Joyce & Iversen 1984; Sandson & Albert 1984; Norman & Shallice 1986; Albin et al. 1989; Jones et al. 1991; Hauber 1998). Do such changes in the control of behaviour similarly underlie the cage stereotypies of captive animals? This hypothesis is feasible given the sensitivity of these pathways to stress (e.g. Scott et al. 1996) and to abnormal rearing conditions (e.g. Jones et al. 1991; Robbins et al. 1996), and is also consistent with the stranger features of cage stereotypies discussed above.

A noninvasive means of testing this hypothesis is to look for specific behavioural correlates of cage stereotypy, using approaches similar to those that have been taken with human subjects (e.g. Turner 1997). Reduced behavioural inhibition produces several behavioural effects that accompany clinical and drug- or lesion-induced stereotypies. For example, amphetamine enhances the rate at which behaviours are initiated and impairs performance in tasks where responses must be suppressed or slowed (Lyon & Robbins 1975; Evenden & Robbins 1983; Robbins et al. 1990; Robbins 1997). Human patients with lesions involving the dorsal striatum show a general tendency to inappropriately repeat previous responses and movements: a phenomenon known as 'recurrent

perseveration' (Luria 1965; Turner 1997). Deprivation-reared rats and primates also perform poorly in tasks sensitive to recurrent perseveration (Morgan et al. 1975), such as extinction learning (Beauchamp & Gluck 1988; Jones et al. 1991). Schizophrenic patients often inappropriately repeat words, and this tendency correlates with their levels of stereotypy (Manschreck et al. 1981; Crider 1997). Both autistic and schizophrenic patients also show recurrent perseveration in several diagnostic tasks, and again, this tendency correlates with their stereotypy (Frith & Done 1983; Turner 1997). Thus, if cage stereotypies similarly reflect changes in dorsal striatal function, then they too should correlate with other symptoms of decreased inhibition, such as the inappropriate repetition of previous responses in specific experimental tasks.

We tested the hypothesis in two wild-caught species of songbird, blue tits, *Parus caeruleus*, and marsh tits, *P. palustris*. Since the dorsal striatal motor circuits of birds are similar to those of mammals in terms of division into an indirect and direct pathway and connections with other motor areas (Medina & Reiner 1997; Reiner et al. 1998), we predicted that the stereotypies of caged birds would have similar correlates to those of mammals.

EXPERIMENT 1: PERSEVERATION IN STEREOTYPIC BLUE TITS AND MARSH TITS (EXTINCTION TASK)

Deprivation-reared rats and primates, with their developmental alterations in striatal function, perform poorly in tasks involving extinction learning (Beauchamp & Gluck 1988; Jones et al. 1991). Furthermore, general perseveration in extinction learning is related to cage stereotypy in rodents and in bears (Sharpe & Mason 2001, *in press*; Garner & Mason 2002). We tested whether stereotypic birds would show similar evidence of behavioural disinhibition, using a spatial discrimination and extinction task (cf. Jones et al. 1991), where the birds were first taught a spatial discrimination, and then required to suppress (extinguish) this response. We predicted that high stereotypy individuals would perseverate in this condition, and hence take longer to suppress the response. Extinction learning is a measure of perseveration that can either be used directly or refined, as we do here using checks and balances measured or imposed elsewhere in the task to control for alternative explanations of stereotypy, such as motivational differences (Terlow et al. 1991), differences in perceptual or learning abilities (Cooper & Nicol 1991) and hedonic differences (Rushen et al. 1990; Cooper & Nicol 1991).

Methods

Subjects and housing

Subjects were five blue tits and five marsh tits that had been caught in the wild 3 years earlier and were being used for other behavioural research. The birds were housed in wire-mesh cages (ca. 76 × 46 × 46 cm). Each cage contained a metal nestbox, two perches running across the cage and two wall-mounted storing blocks (a 2.5 × 20-cm wooden block with 20 drilled holes and six

small perches attached). Some cages had an additional branch running diagonally and additional storing blocks. Cage furniture was not standardized at the start of the experiment, because this could disrupt the performance of established stereotypies. Birds were fed Orlux birdfeed with additional hemp seeds, peanuts, sunflower seeds and three mealworms, *Tenebrio molitor*, per day. The birds were kept under a 14:10 h light:dark cycle with lights on at 0700 hours. Animal care and use was in accordance with Home Office regulations and institutional guidelines.

Video observations of stereotypy

Birds were videotaped from 1200 to 1600 hours, and 1 h of behaviour was analysed using a one-zero method, with sample times of 1 min; one-zero recording is well suited to intermittent behaviours (Martin & Bateson 1986) such as stereotypy (Garner 1999). We recorded the time spent active (i.e. all behaviours except sleeping) and performing stereotypy. Birds of both species displayed route-tracing stereotypies. Stereotypic routes were identified for all animals; birds differed in the number of points that they visited in a complete circuit. To avoid confounding variability and repetition, we did not set the criterion for scoring stereotypy within each time sample as the completion of a constant number of circuits: otherwise, with increasing route length, a bird would have had to be more invariable to reach criterion. Instead, birds had to visit the same number of points, rounded to the nearest complete circuit. Thus, if two points were visited, four consecutive circuits were required for criterion; if three, three circuits; and if four or more, two circuits. Stereotypy was expressed as the proportion of activity spent engaged in the behaviour.

The problem of feather wear

The assessment of stereotypy was complicated by the fact that many of the birds suffered from feather wear. Stereotypy probably contributed to this; for example, the wing feathers of one blue tit (BA) were worn only on the wing that brushed against the cage during stereotypy. In turn, feather wear seemed to affect stereotypies. Many birds with low feather wear incorporated airborne portions into their stereotypies, and some displayed bizarre aerobic manoeuvres. In contrast, birds with high feather wear were unable to do this because their flight was so impaired. Therefore, feather wear (which reflected in part the past performance of stereotypy) probably affected both the form and amount of current stereotypy, and thus needed to be controlled for. We therefore assessed feather wear on a five-point scale: (1) no damage, flight unimpaired, to (5) no viable flight feathers, bird unable to fly, with the midpoint (3) representing four or more abraded or snapped primaries and awkward flight. The impact of feather wear could then be investigated and dealt with statistically.

Operant training and extinction learning protocol

Birds were food-deprived at approximately 0800 hours. Both training and testing sessions began at 1000 hours.

Order of testing was randomized and balanced from day to day to avoid confounding feeding motivation with individual differences. Birds were transported to a separate cage for testing. They were taught to peck a computer-controlled touch screen for a reward of chopped peanut, having been shaped in prior experiments.

In 'attention training', birds were trained to peck on a rectangle presented in the centre of the screen. Three consecutive pecks were rewarded with a piece of chopped peanut. The stimulus timed out after 30 s. The intertrial interval (ITI) was 30 s. Training ended after 20 rewards had been collected. In the spatial discrimination task, birds pecked a central attention stimulus three times to bring up two choice stimuli. One side, chosen at random at the start of the task, was always correct, and choosing the correct side provided a reward of a small piece of chopped peanut. As before, trials timed out after 30 s, and the ITI was also 30 s. Sessions lasted 1 h. Criterion for this task was 18 out of 20 trials correct maintained for three consecutive trials. Spatial discrimination performance was scored as the number of responses (completed trials) to criterion. To measure motivational differences, we also calculated completion rate, defined as the proportion of trials in which animals responded.

In the extinction task, stimuli were presented as before, but no reward was dispensed. Rather than returning to a 50/50 choice (as is common in rodents; e.g. Jones et al. 1991; Garner 1999), the birds gradually ceased responding. Extinction learning performance was therefore measured as the number of trials required to drop to a completion rate, measured over 25 trials, that was 25% of the completion rate during initial learning, thereby controlling for individual differences in general learning ability, perception and motivation.

Statistical methods

In all analyses, transformations were applied where necessary to meet the assumptions of parametric methods (homogeneity of variance, normality of error and linearity). The general linear method (GLM) was used throughout. All analyses were performed in Minitab 12 for Windows.

Sexes of the birds were unknown. We used wing lengths at capture, and, in the marsh tits, response to recorded male song, to attempt to deduce sex. However, these methods did not agree, so sex was excluded from all analyses.

We first investigated the effects of feather wear and route length on stereotypy. Serious feather wear would have prevented any airborne elements in each route. Birds might then replace these long-distance flights with a series of hops along the cage side, in which case feather wear should positively correlate with the mean route length (in points visited) of each animal. To test for this, we correlated feather wear and mean route length; both continuous variables were logged to meet the assumption of linearity, and the analysis was blocked by species. We also checked that the criterion used for rating stereotypy did not bias the stereotypy score by correlating stereotypy

and mean route length. In this analysis, species and feather wear were partialled out (i.e. the analysis was blocked by these), and again all continuous variables were logged.

Testing the correlation between stereotypy and recurrent perseveration is potentially complicated if the relationship between these variables is affected by feather wear. We therefore included the interaction between stereotypy and feather wear to control for any such effects. Thus, extinction learning performance was correlated with feather wear, stereotypy and their interaction, blocking by species, and also partialling for spatial discrimination performance to exclude any common learning, motivational or hedonic differences. Extinction learning performance was inverse-transformed to yield homogeneity of variance. For purposes of data display and incorporation with later experiments, the resulting regression equation was rearranged to give stereotypy partialled for the other terms in the analysis, and the partialled stereotypy score was calculated for each individual using this equation and their original data (a method conceptually similar to the calculation of factor scores for subjects in factor analysis). The partialled stereotypy score was correlated with extinction learning performance, blocking by species. Spatial discrimination performance was also correlated with feather wear, stereotypy, and their interaction, partialling for species, and extinction learning performance. No transformations were required here.

Results

Stereotypy

Individual levels of stereotypy performance ranged from 7% to 90% of active behaviour, and no two birds traced the same route. All the birds traced at least two routes. Some routes shared common subelements with other routes performed by the same bird (e.g. MF; Fig. 1), and birds often switched between routes at the point at which the routes diverged (e.g. the point nearest the observer in Fig. 1). Some animals also appeared to vary in the sequencing of particular route elements. For instance, marsh tit ME showed a variety of repeated routes, which were often sequenced interchangeably, and shared common elements that were performed invariably once begun. Thus, each 'family' of routes sharing common elements could be interpreted as a single route with a variable element (Fig. 2). We also observed animals shift between sequences with shared elements via 'transitory stereotypies' (sensu Cronin & Wiepkema 1984) built from a cocktail of repeated elements from both stereotypies (Fig. 2). Oscillation between routes meant that the same routes often were not performed consecutively and hence did not reach criterion. This could potentially have confounded variability and repetition. However, mean route length and stereotypy were not significantly correlated ($F_{1,6}=0.59$, $P=0.471$, $r=-0.085$, significant correlations greater than $r \pm 0.233$ could be detected with our data set at a power of 0.8; Fig. 3), suggesting that the criteria used for scoring stereotypy was unlikely to have biased the

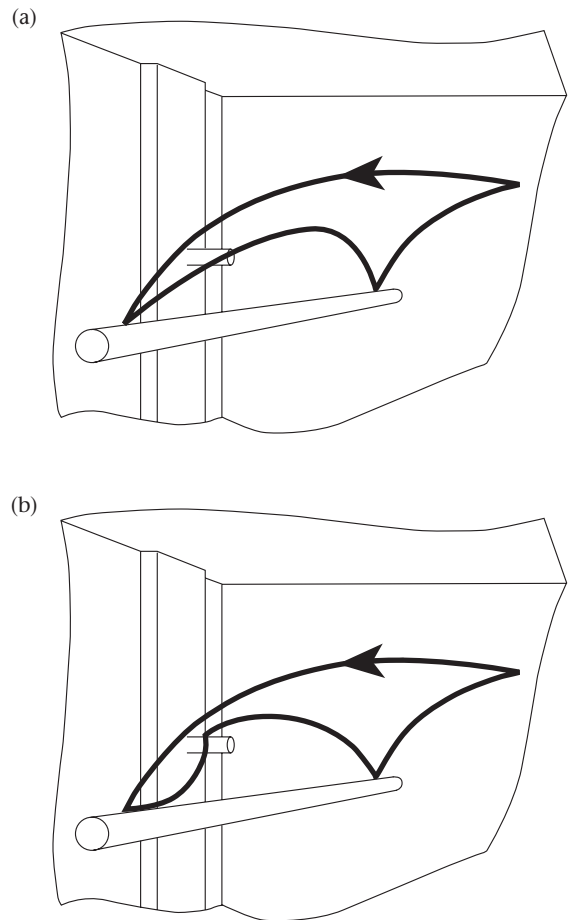


Figure 1. Two routes traced by marsh tit MF, which share common elements. Thus the bird always jumped from the far end of the perch, to the cage back, and then to the end of the perch close to the observer. (a) The bird jumped immediately to the start position. (b) The bird jumped via a perch on the storage block on the side wall of the cage.

results. As predicted; mean route length increased with feather wear ($F_{1,8}=7.76$, $P=0.024$).

Extinction learning

Extinction learning performance (trials to extinguish the learned response) was correlated with the stereotypy-by-feather wear interaction term ($F_{1,4}=190.85$, $P<0.0005$); stereotypy showed a positive correlation with extinction learning performance at low levels of feather wear, which flattened and reversed slightly at high levels of feather wear. Marsh tits also extinguished more rapidly than did blue tits ($F_{1,4}=132.91$, $P<0.0005$) (as preliminarily reported in: Garner et al. 1998). The partialled stereotypy score calculated for each individual was positively correlated with extinction learning performance ($F_{1,7}=373.71$, $P<0.0005$; Fig. 4). In contrast, stereotypy failed to correlate with spatial discrimination performance ($F_{1,4}=1.62$, $P=0.272$, $r=0.116$, significant correlations greater than $r \pm 0.202$ could be detected with our data set at a power of 0.8). Hence, stereotypy showed a selective correlation with extinction learning but not with the initial discrimination acquisition.

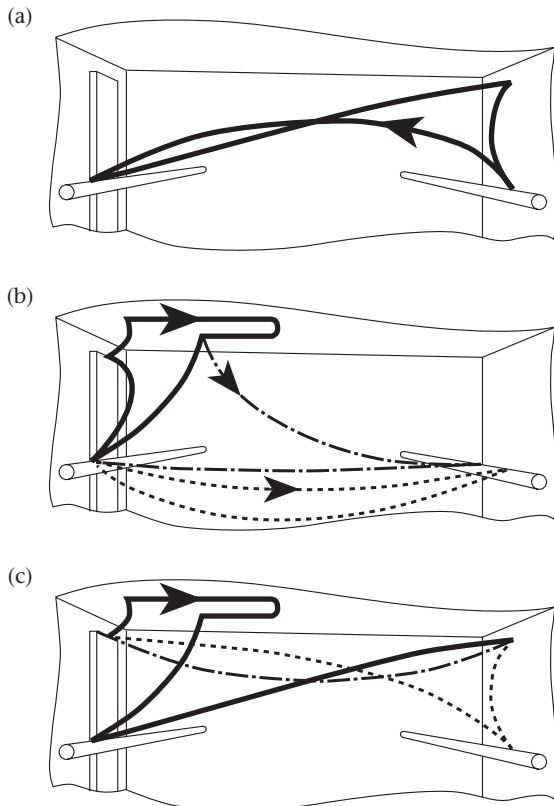


Figure 2. Routes traced by marsh tit ME. (a) Route comparable to a 'fixed routine' in pig oral stereotypies. (b) Route comparable to a 'loose stereotypy' in pigs. Solid line depicts the unvarying portions of the route. The bird jumped from the left perch to the top of the storing block, then to the roof, where it walked upside down, turned around, then jumped back to the perch. Occasionally the bird returned via the right perch (dash-dot line), or began the route by flying to the right perch and back again (dotted line). (c) Route equivalent to a 'transitory stereotypy' in pigs, containing elements from (b) and (a). The bird flew from the perch to the upper right side, then to the storing block, either directly (dash-dot line) or via the right perch (dotted line).

Discussion

The caged parids studied here showed hopping and airborne route-tracing stereotypies similar to those described for caged canaries, *Serinus canarius*, and other birds (Keiper 1969, 1970). However, compared with Keiper's descriptions, the stereotypies that we observed seemed complex and/or somewhat flexible. In this regard, they appeared more akin to the 'nonfixed' stereotypies of mink (Mason 1993) or the oral stereotypies of stalled sows (Cronin & Wiepkema 1984). Thus, stalled sows also perform stereotypies ('fixed complete routines') that are comprised of repeated elements ('fixed subroutines') common to more than one stereotypy (Cronin & Wiepkema 1984). The high incidence of 'loose' (Fig. 2b) and 'transitory' (Fig. 2c) stereotypies may also indicate that the birds' stereotypies had not yet become fixed and completely 'established' (Cronin & Wiepkema 1984), despite their 3 years in captivity.

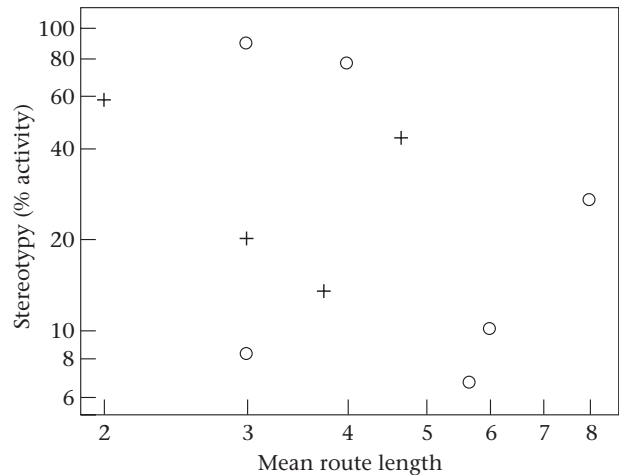


Figure 3. Stereotypy plotted against mean route length. Stereotypy is expressed as the percentage of activity. Circles: blue tits. Crosses: marsh tits. No significant relationship was found between these variables, indicating that the criteria used for scoring stereotypy was unlikely to have biased the results.

The results confirmed our prediction that stereotypy would be linked with perseveration as measured by impaired extinction learning. Extinction learning and simple discrimination involve many common processes, but they differ in one important regard, the inhibition of responses. Thus, the correlation between extinction learning and stereotypy when partialled for simple discrimination performance (i.e. when the effect of common processes is statistically eliminated), coupled with the lack of correlation between simple discrimination and stereotypy, suggests that stereotypy is associated with a general disinhibition of responses. The interaction of feather wear and stereotypy is also consistent with this interpretation. The high-feather-wear individuals that still showed higher stereotypy probably modified their route to replace flying portions with hops, as evidenced by the positive correlation between feather wear and route length. Such birds arguably behaved more flexibly than did other high-stereotypers, and so could be expected to show lower extinction scores.

Extinction learning may reveal behavioural disinhibition, but it does not implicate a unique brain region. For example, extinction is sensitive to recurrent perseveration caused by striatal changes, but it might also reflect a fixity of goals (i.e. the stuck-in-set perseveration characteristic of compulsive behaviours and prefrontal cortex dysfunction: Turner 1997; Garner 1999). Furthermore, our data were complicated by the birds' differing degrees of feather wear. Therefore, in experiment 2, we retested our hypothesis using a different task that specifically measured recurrent perseveration, after allowing the birds to moult new flight feathers.

EXPERIMENT 2A: PERSEVERATION IN STEREOTYPIC BLUE TITS (GAMBLING TASK)

Recurrent perseveration, the inappropriate repetition of previous specific responses or movements, is a cardinal

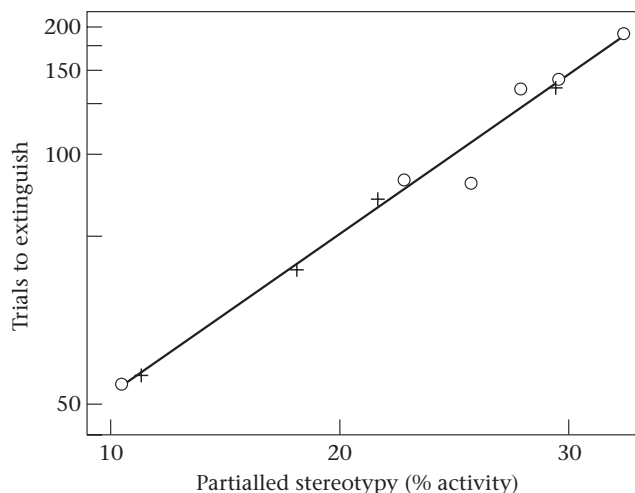


Figure 4. The relationship between perseveration measured by extinction learning and stereotypy. Stereotypy was partialled for species, feather wear and performance on the spatial discrimination task. Circles and solid line: blue tits. Crosses and dashed line: marsh tits.

symptom of dorsal striatal dysfunction (Luria 1965; Norman & Shallice 1986; Turner 1997). Frith & Done (1983) first demonstrated the correlation between recurrent perseveration and stereotypy in schizophrenia using a simple gambling task. The same correlation was subsequently reported in autistic patients (Turner 1997). In the gambling task, subjects are instructed to search for 'the rule' governing which of two squares on a computer screen hides a reward. In fact, no rule is present, and the correct side is selected at random by the computer. The instruction to 'search for a rule' steers a human subject away from developing a side bias. Healthy (control) subjects tended to produce semirandom sequences of choices, but impaired subjects produced sequences rich in repetition, or alternation (e.g. 'LRLLLLLLRRRRRRRR', or 'LRLRLRLRLRLRLRL', respectively). We therefore predicted that measures of recurrent perseveration obtained from this task would correlate with the spontaneous performance of cage stereotypy.

Methods

Subjects and housing

We used the same birds as in experiment 1. However, in an attempt to remove the confound of feather wear, we used the birds after they had been group-housed in an outside aviary for 2 months, allowing them to moult new flight feathers. The birds were then returned to their original cages. We equipped each cage with the original arrangement of cage furniture to avoid disrupting stereotypy. After completion of the gambling task, stereotypy and feather wear were assessed as before.

Operant training and gambling task protocol

Birds underwent 'attention training' as before, followed by 'gambling training', designed to emulate the verbal

instruction to 'search for the rule' in human experiments (Frith & Done 1983, 1990). As before, birds initiated a trial by pecking an attention stimulus. The bird was then presented with either a choice of both squares (choice trials) or a single square on the left (forced left), or right (forced right). Choice, forced-left and forced-right trials occurred at random with an equal probability. In choice trials, each square was rewarded with a probability of 50%. Following an 'incorrect' choice, a choice trial was presented again, where the same side was rewarded as before (correction trials). Forced trials ensured that the animal was continuously rewarded on both sides, and correction trials were intended to prevent superstitious conditioning.

'Gambling training' was planned to continue until 20 rewards were collected on each side in choice trials. In practice, none of the marsh tits achieved this criterion; they simply ceased responding to the attention stimulus. We later discovered that all the marsh tits cached food in locations hidden from day-to-day husbandry, so they were probably less food-motivated than the blue tits, and the unreliable reward in the gambling experiment may have been insufficient to maintain their interest. We therefore withdrew them from this experiment, but exploited their food-storing behaviour in experiment 2b.

The blue tits completed gambling training successfully and were presented with a gambling task. As before, birds had to peck three times on a central attention stimulus (time out: 30 s) to bring up two choice squares. Each square was correct 50% of the time. Stimuli timed out after 30 s. As before, three pecks indicated a choice. Gambling sessions were repeated until 306 responses were recorded for each animal, to allow the data to be suitably subdivided as described below.

Data processing

In human experiments (Frith & Done 1983; Turner 1997), the degree to which individuals pattern their responses was quantified using information theory (Attneave 1959; Frith & Done 1983). However this method suffers from the disadvantage that the information statistic is affected by side bias. Thus, even if two birds both choose the squares at random, if they differed in their side bias, their sequence of responses would be quantified differently (Attneave 1959). Side bias and perseveration are distinct phenomena: the sequence LLLLLRRRRRLLLL is more perseverant than LRLRLRLRLRLRL despite having the same global side bias (L:R=10:6). As we wanted to ensure that birds that chose both squares with an equal level of randomness but with a different global side bias would receive the same score, we analysed the data using Markov analysis.

Markov chains describe the probability of a behaviour occurring as a function of previous behaviours, and provide a means of quantifying repetition that controls for side bias. The total sequence of 306 choices was divided into six segments of 51 choices. This allowed us to perform the Markov analysis on a localized area of the total sequence using the side bias of that localized area.

Table 1. A sample transition matrix and worked example for a third-order Markov analysis

Preceding choices	Current choice					
	L			R		
	Observed	Expected	χ^2	Observed	Expected	χ^2
L L L	5	0.735	24.735	2	1.366	0.295
L L R	1	1.366	0.098	1	2.536	0.930
L R L	0	1.366	1.366	1	2.536	0.930
L R R	2	2.536	0.113	5	4.710	0.018
R L L	1	1.366	0.098	0	2.536	2.536
R L R	0	2.536	2.536	6	4.710	0.353
R R L	1	2.536	0.930	5	4.710	0.018
R R R	4	4.710	0.107	15	8.747	4.470
					$\Sigma\chi^2$	39.534
					P	1.546×10^{-6}

The actual side bias shown in this example is 0.35 (the proportion of choices to the left). The expected number of times each combination of preceding and current choices should be observed can therefore be calculated directly from this probability. For example: $P(\text{LLR-L}) = P(\text{L}) \times P(\text{L}) \times P(\text{R}) \times P(\text{L})$. The probability associated with $\Sigma\chi^2$ represents the probability that such a sequence would occur by chance alone if the bird was choosing randomly with the given side bias. In this case, the key contributions are from the perseverant sequences LLL-L and RRR-R.

We calculated the observed probability of choosing each side on a given trial for each possible combination of three previous choices (i.e. a third-order Markov chain). Third-order Markov analysis was chosen to maintain consistency with the human gambling experiments, which uses information theory to examine the relation between the current choice and the preceding three. We calculated the expected probability of each combination of three preceding choices directly from the observed side bias for that 51-trial sequence segment, and we calculated the sum chi-square from the observed and expected counts. The probability of each sum chi-square is the probability of sequential independence of the observed chain. A Markov chain showing sequential independence is one where the probability of each response is unaffected by previous responses (Martin & Bateson 1986). Thus, not only does dividing the sequence into smaller segments ensure that each segment is analysed with an appropriate side bias, but also meets the assumption of homogeneity of event probability (i.e. side bias) necessary for the chi-square analysis. When analysing Markov chains with multiple preceding behaviours, each chain can be expressed as a two-way contingency table, with two columns for the two present choices, and with the number of rows determined by the number of possible preceding chains (8 in this case) (Cane 1978). The degrees of freedom are therefore $(2 - 1) \times (2^{\text{chain order}} - 1)$, or seven for a third-order chain. This approach explicitly tests the relationship between the current choice given the previous three. It also appropriately calculates degrees of freedom given the fact that the expected values are calculated from the data rather than specified by the null hypothesis (as they would be in a one-way table: Hogg & Craig 1978, page 276; Christensen 1990, page 31). We performed the Markov analysis using a custom-written computer program (available from the authors on request; see Table 1). The mean third-order probability of sequential independence was calculated for

each bird by taking the mean of the six probabilities calculated for each of the six segments of the original sequence.

Statistical methods

As only the blue tits completed this task, the sample size was small. P values calculated from the normal distribution can be inaccurate at small sample sizes, since deviations from normality of error and homogeneity of variance are progressively harder to detect. We therefore performed a rerandomization test to confirm each P value, in which the dependent and independent variables were shuffled at random, and the proportion of results that yielded a test statistic (i.e. the correlation coefficient) equal to or greater than that observed was counted over 5000 iterations. This method directly calculates the probability of a type I error given the observed error distribution, rather than an assumed normal error distribution.

Results

Stereotypy and feather wear

Although the stereotypies observed were qualitatively similar to those in experiment 1, both the routes traced by individual birds and the intensity of stereotypy changed dramatically. The range of stereotypy performed was 0–50% of active behaviour. None of the birds showed significant feather wear.

Gambling task

As predicted, the mean probability of sequential independence in third-order Markov chains was negatively correlated with stereotypy (GLM: $F_{1,4} = 79.04$, $P = 0.003$). The rerandomization test confirmed this result ($P = 0.007$). Thus, as stereotypy levels increased, the randomness of birds' responses decreased (Fig. 5).

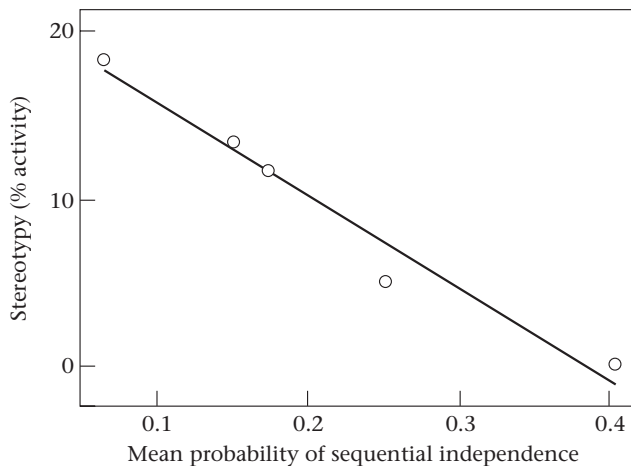


Figure 5. The relationship between perseveration measured by the mean probability of third-order sequential independence and stereotypy. Randomness increases along the *X* axis, indicating that high stereotypy birds showed higher levels of perseveration.

Discussion

The blue tits' performance confirmed our prediction that stereotypy would be linked with predictable responding in a gambling task. Thus, their stereotypy was associated with recurrent perseveration, measured using the same task as in schizophrenic and autistic humans. This result confirms the conclusion from experiment 1 that cage stereotypy in blue tits involves altered behavioural control and builds on it to implicate the dorsal striatum specifically.

We learned several lessons for future applications of the gambling task to nonhuman animals. First, the animal version of the task remains more susceptible to the development of side bias, despite careful pretraining, than does the human (verbally instructed) task. Information-based measures as used in humans are therefore inappropriate, and chi-square probabilities from high-order Markov chains are invaluable for separating recurrent perseveration from side bias. Relating this statistic to a known probability distribution is advantageous, because the meaning of an absolute probability is independent of side bias. Second, detailed post hoc analysis of the response sequences (for details see Garner 1999) revealed that longer repeated sequences are built up from smaller repeated units, and that testing for longer repeated sequences increases the sensitivity to perseveration. Thus, the need for long response sequences is emphasized. Given the problem of developing side bias, we recommend using the analysis of sequence segments with the mean as the test measure. Following these observations and the evolution of the human task (Turner 1997), we have recently refined the task (Garner et al., *in press*).

EXPERIMENT 2B: PERSEVERATION IN STEREOTYPIC MARSH TITS (FOOD-CACHING EXTINCTION)

In experiment 2a, the marsh tits apparently failed to engage in the task because of their food-storing behav-

our, which prevented the food deprivation necessary to engage in gambling. However, we still wanted to reassess perseveration in these subjects. In this experiment, we therefore exploited the marsh tits' storing behaviour to assess their perseveration in a naturalistic version of experiment 1, their extinction rates of unrewarded storing behaviours.

Food stores are often raided in the wild, and food-storing birds in the laboratory can learn to avoid storing food in sites that have been pilfered previously (Stevens 1984; Hampton & Sherry 1994). However, such avoidance does not always occur, nor does the predicted reduction in overall caching rates that might be expected with sustained pilfering (Baker & Anderson 1995; Lucas & Zielinski 1998). Functional explanations and/or constraints on learning may underlie such results; alternatively cache pilfering may represent a naturalistic extinction task, the response to which actually depends on the experimental animal's perseverative tendencies. We therefore tested the prediction that stereotypic marsh tits would show difficulty in suppressing unrewarded behaviour by examining their reduction in food storing in response to the daily pilfering of their caches. We predicted that high-stereotypy birds would persist with caching, whilst low-stereotypy birds would reduce the amount of food stored.

Methods

Extinction task

The five marsh tits were used from experiment 2a. We added five storing blocks to each cage: four wooden blocks, as described above, and an additional block consisting of 10 small rubber pockets mounted on a solid wooden backing. The birds were fed Orlux with chopped peanuts and sunflower seeds added in excess and eight peanut halves per day. The light cycle was set to LD 9:15 h to stimulate storing behaviour. We tested the birds for 4 weeks. In weeks 1 and 4, we collected baseline data; in weeks 2 and 3, we pilfered the birds' food stores and collected data.

At lights on (0800 hours), the birds were food-deprived for 90 min to increase the likelihood of storing later in the day (Clayton & Cristol 1996). At 0930 hours, the food cup and storage block were replaced in the cage. During baseline observations, the number and location of stored food was recorded every day at 1300 and 1615 hours. Stored items were marked with red food colouring for identification. The cages were also searched carefully for other caches, and food stored in locations other than the storing block was similarly recorded and marked. The storing blocks were removed 15 min before lights out (1645 hours). During the cache-pilfering treatment, at 1300 and 1615 hours each day, instead of marking the cached food, all cached food was removed and discarded.

Data processing

We calculated the total number of seeds stored during the baseline and cache-pilfering periods. The number of seeds stored during baseline was then divided by the

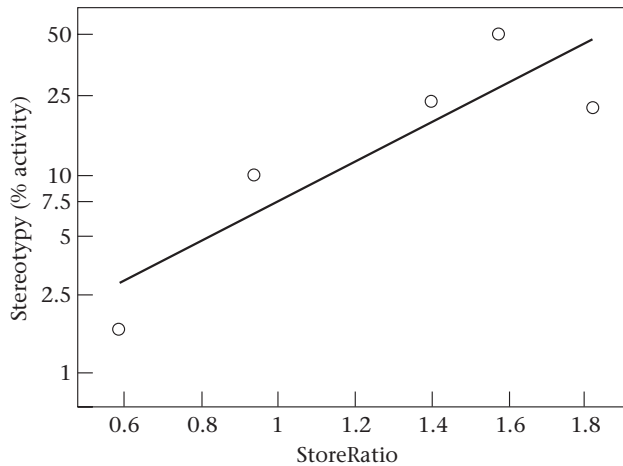


Figure 6. The relationship between stereotypy and the perseveration of food-storing behaviour when caches are pilfered. StoreRatio: number of seeds stored when caches were pilfered, divided by the number stored in the baseline condition. A StoreRatio greater than 1 indicates that the bird increased storing when caches were pilfered.

number stored during pilfering (StoreRatio). A StoreRatio less than one thus indicates a decrease in food-storing behaviour.

Statistical methods

Stereotypy measures taken in experiment 2a were regressed on StoreRatio. Stereotypy was log transformed to preserve homogeneity of variance. The results were confirmed by a rerandomization test as in experiment 2a.

Combining the results of experiments 2a and 2b

To compare the results from both experiments, increase the overall sample size and give an overview of the correlates of stereotypy, we pooled the findings from experiments 2a and 2b. Stereotypy scores were normalized for each species (i.e. the species mean was subtracted from each stereotypy score, and the result divided by the species standard deviation). For the blue tits, the mean probability of sequential independence from the gambling task (experiment 2a) was similarly normalized. This process was repeated for the StoreRatio score of the marsh tits (experiment 2b). The normalized mean probability of sequential independence was then multiplied by -1 , so that both of these variables could be combined on the same scale, giving a new variable that we termed 'perseveration'. Normalized stereotypy was then regressed on perseveration using the general linear model. Normalized stereotypy was log transformed to improve homogeneity of variance, and the analysis was blocked by species to avoid pseudoreplication.

Results

The StoreRatio of the marsh tits was positively correlated with stereotypy ($F_{1,3}=9.92$, $r=0.876$, $P=0.051$; Fig. 6). These results were confirmed by the rerandomization test ($P=0.044$). Thus, as predicted, the more stereotypic the bird, the more it persisted in storing food when its caches were pilfered. The three most stereotypic birds

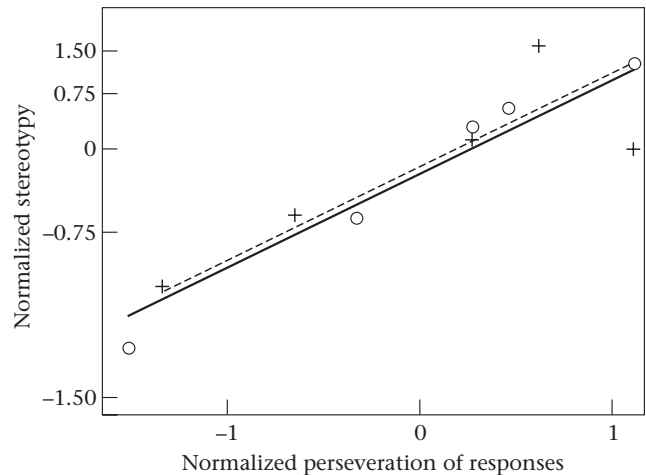


Figure 7. Combined data from experiments 2a and 2b measuring perseveration. Stereotypy, as well as the measure of perseveration taken in each experiment, was normalized for each species so that they could be combined on equivalent axes. The analysis is blocked by species. Normalized data are dimensionless, and expressed in units of standard deviation from the mean. Circles and solid line: blue tits. Crosses and dashed line: marsh tits.

even increased their storing rates during pilfering, while the two least stereotypic ones reduced them.

Both experiments 2a and 2b yielded similar correlates of stereotypy; thus the overall analysis of normalized stereotypy versus perseveration in both experiments was highly significant ($F_{1,7}=35.69$, $r=0.835$, $P=0.001$; Fig. 7). Thus, across both species, stereotypy was correlated with a difficulty in suppressing responses.

Discussion

To measure perseveration in the marsh tits we used a naturalistic extinction task, the pilfering of their daily food stores. As predicted, the high-stereotypy marsh tits persisted in inappropriate food storing. As the food stores were removed and replaced with the day's feed, pilfered birds would have been no more or less hungry than in the baseline condition. Furthermore, although it is possible that stereotypic birds used more energy, and that their greater rates of energy expenditure thus led them to store more (cf. effects of temperature decreases on storing; Clayton & Cristol 1996), the more stereotypic birds did not in fact store at greater rates than did less stereotypic birds ($r_3=-0.02$, $P=0.978$). This explanation is also difficult to reconcile with the results of experiments 1, 2a and 3, which suggest that general behavioural perseveration is the most parsimonious explanation for this insensitivity to store pilfering. Thus, as predicted, the most stereotypic animals appeared to have problems suppressing inappropriate storing.

EXAMINING EFFECTS OF ENRICHMENT: COMPARING EXPERIMENTS 1 AND 2

As described above, between experiments 1 and 2, the birds were housed in outdoor aviaries to moult new flight feathers. These aviaries were large ($h \times w \times d =$

2.4 × 1.2 × 2.7 m), allowing the birds to be group-housed, and were enriched with natural vegetation. As well as allowing the moult, this environment is exactly the type that we would expect to reduce or eliminate cage stereotypies (Keiper 1969; Mason 1991a), assuming that the birds' behaviour had not become 'established' and resistant to enrichment (e.g. Fentress 1976; Cooper & Nicol 1996; Cooper et al. 1996).

Although data could not be collected during this enrichment phase itself, there were reductions in stereotypy (see experiment 2a). Therefore it seemed likely that the birds would show individual differences in behaviour upon recaging. In human biological psychiatry, a methodological gold standard is to show that changes in symptom severity are correlated with changes in the underlying physiological process believed to be responsible for the symptom (e.g. for a fascinating example see Baxter et al. 1992). Time constraints prevented us from repeating the extinction protocol of experiment 1. However we were able to compare statistically data from experiments 1 (pre-enrichment) and 2 (postenrichment). We predicted that the birds' changes in stereotypy with enrichment would be correlated with changes in recurrent perseveration.

Statistical Methods

We used the normalized stereotypy score and the normalized perseveration score calculated in experiment 2b to represent stereotypy and perseveration postenrichment. We also normalized the partialled stereotypy score calculated in experiment 1 (Fig. 4) by species as described in experiment 2b to yield a pre-enrichment stereotypy score statistically adjusted for feather wear. We similarly normalized the number of trials taken to extinguish the learned response in experiment 1 to yield a pre-enrichment perseveration score. Thus, the data from all three experiments were represented on the same dimensionless normalized scale and controlled for confounding effects such as feather wear and species.

Pre- and postenrichment stereotypy scores were then correlated, partialling by species, to examine the nature of the change in stereotypy. Each individual's increase in stereotypy and perseveration was then calculated (i.e. postenrichment score minus pre-enrichment score), and the change in stereotypy correlated with the change in perseveration in a GLM partialled by species.

Results

Individual pre- and postenrichment stereotypy scores were uncorrelated (GLM: $F_{1,7}=0.18$, $r = -0.091$, $P=0.680$; a power of 0.8 requires $r \geq 0.182$ or $r \leq -0.182$), indicating that individuals differed in their responses to enrichment. However, as predicted, the change in stereotypy during enrichment was positively correlated with the change in recurrent perseveration (GLM: $F_{1,7}=30.93$, $P=0.001$; Fig. 8).

Discussion

This result demonstrates that the correlation of stereotypy and perseveration does not stem simply from

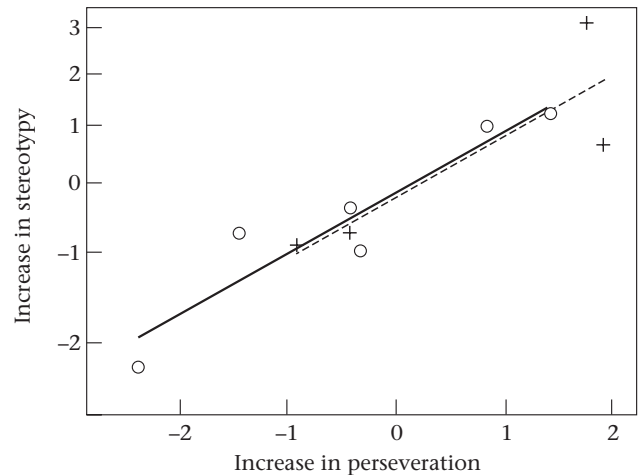


Figure 8. Changes in stereotypy with enrichment correlated with changes in perseveration. The analysis is partialled by sex. Axes are normalized. Circles and solid line: blue tits. Crosses and dashed line: marsh tits.

baseline individual variation in behavioural control. Instead, it shows that behavioural disinhibition is affected by housing, with environment-induced changes in stereotypy mirrored by environment-induced changes in perseveration. As in human methodologies, the fact that changes in measures of symptoms are correlated with changes in the putative underlying physiological process provides stronger evidence that recurrent perseveration is causally related to stereotypy.

EXPERIMENT 3: CAGE STEREOTYPY AND BASELINE RATES OF BEHAVIOURAL SWITCHING IN MARSH AND BLUE TITS

In our final experiment, we assessed baseline rates of behavioural switching as a measure of spontaneous, rather than experimentally elicited, signs of altered striatal functioning. Dorsal striatal dysfunction can have two paradoxical-seeming effects on baseline behaviour. First, the general disinhibition of behaviour can lead to an elevated rate of switching between behaviours and behavioural elements. Second, when disinhibition leads to the perseveration of a particular behaviour, an apparent increase in bout length may occur (although the behaviour may never be properly completed because it is being reinitiated before completion of the behaviour chain: Lyon & Robbins 1975). Thus, both increased behavioural switching rates and a tendency to repeat and persist with bouts of one action can occur with dorsal striatal dysfunction. For example, amphetamine enhances the rate at which behaviours are initiated, impairs performance in tasks where responses must be suppressed or slowed, and at higher doses leads to the constant repetition of a single set of movements (Lyon & Robbins 1975; Evenden & Robbins 1983; Robbins et al. 1990; Robbins 1997). Similarly, in institutionalized patients with Down's Syndrome, stereotypies are linked with higher frequencies of fewer behaviours (Francis 1966; Francis 1971), and in caged bank voles, enhanced rates of behavioural switching are

Table 2. Ethogram of behaviours recorded in experiment 3

Category	Description and subdivisions
Jump/Hop	A single jump from one cage location to another; the wings remain folded.
Turn	A single hop during which the animal rotates 180° on the perch.
Fly	Flight from one cage location to another, or a jump with open wings.
Climb	Walking or perching on the cage sides. Left, right, front, back and top sides were differentiated.
Perch	Bird grasped a perch with its feet. Very brief perches (e.g. between two jumps) were counted. Left, right, and third perches were differentiated, as was perching on left or right storage blocks.
Alert/Investigate	Bird was stationary, and made a series of rapid head movements. Short bouts resemble intention movements, but can be differentiated by orientation of the head towards different locations before movement, and a tendency to jump or fly to a location other than where the bird was looking. Intention movements were excluded.
Sing	The bill was opened and rapid movement of the throat and upper breast could be observed.
Feeding	The feeder was out of sight of the camera on the cage floor, but birds often returned to a perch to eat food that they had collected. The food was held in the feet and the bird pecked and pulled at the food item.
Bill wiping	While gripping the perch, the bird bent forward and wiped first one, and then the other side of its bill along the edge of the perch, sometimes repeating this sequence.
Grooming	Several grooming behaviours could be differentiated: preening the chest, shoulders (left and right), under the wings (left and right), legs and abdomen (left and right), and the chest, scratching with either left or right leg, and ruffling (bird erected its feathers and shook vigorously).
Pecking	Bird pecked at a location in the cage.
Out of sight	Time out of sight, either on cage floor or in transfer box.

correlated with both stereotypy and perseveration in extinction (Garner & Mason 2002). We therefore videotaped home-cage behaviour to determine whether the most stereotypic individuals showed elevated or reduced rates of switching between different activities.

Methods

Rate of behaviour switching

An ethogram was constructed (Table 2). For all 10 subjects, we chose four 1-min observation periods at random from the video recordings taken in experiment 2a; two contained stereotypic behaviour and two did not. We thus avoided the artefactual oversampling of stereotypic periods in high-stereotypy animals, and could make a fair comparison between individuals. A short observation period was used because of the very high frequency of behavioural transitions observed. We coded the sequence of behaviours observed as a series of events. Pauses in a behaviour followed by reinitiation were coded as two separate events, as were transitions between subdivisions of the same category (e.g. grooming the left wing followed by ruffling). We thus counted the total number of behaviours initiated by the bird during the observation period. Some cage areas were out of sight of the camera. We divided the total number of behaviours initiated across the 4 min by the time that the bird was visible, to give the mean rate of behaviour switching per second.

Statistical methods

For each individual, the difference in the mean rate of behaviour switching during the stereotypic and non-stereotypic observation sessions was regressed against stereotypy to check that stereotypic route tracing did not contribute artefactually to any result. This analysis was blocked by species, and stereotypy was log transformed.

If stereotypy contributed artefactually to the results, the difference between stereotypic and nonstereotypic observations would be positively correlated with stereotypy.

To test the hypothesis that stereotypic animals would show altered switching rates, stereotypy was regressed on the mean rate of behaviour switching using the general linear model. The analysis was blocked by species. However, initial model fit was poor, violating linearity, and so we then performed asymptotic and quadratic regressions. The asymptotic regression also fitted the data poorly. The quadratic regression provided the best fit (R^2 , the proportion of variation explained by the model, was far greater) and is reported here. Stereotypy was log transformed to maintain homogeneity of variance. Following the analysis of the relation between stereotypy and switching rate, stereotypy was then replaced by perseveration, and the mean rate of behaviour initiation normalized as in experiment 2b. We then repeated the analysis.

Results

The difference between switching rates in stereotypic and nonstereotypic periods was uncorrelated with stereotypy ($F_{1,7}=0.10$, $r=0.011$, $P=0.766$; a power of 0.8 requires $r \geq 0.167$ or $r \leq -0.167$), providing no evidence that route tracing artefactually inflated behavioural switching scores.

In the quadratic regression, stereotypy was correlated with the quadratic term ($F_{1,6}=30.62$, $P=0.001$). The overall model R^2 was 89.2% (Fig. 9). Perseveration was also correlated with the quadratic term ($F_{1,6}=18.14$, $P=0.005$), and the overall model R^2 here was 76.1%.

Discussion

There was no simple linear relationship between the rate of behaviour switching and cage stereotypy. The

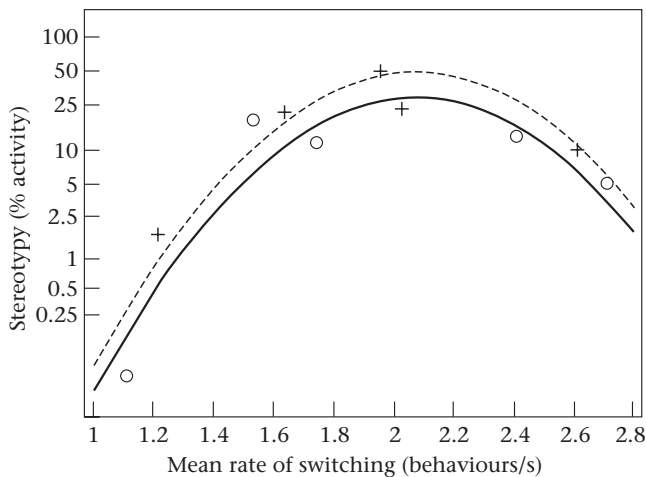


Figure 9. The relationship between stereotypy and the mean rate of switching between all behaviours. Circles and solid line: blue tits. Crosses and dashed line: marsh tits.

curve of best fit was a \sim -shaped quadratic, which had an order of magnitude higher significance, and an R^2 of 89.2%, compared with 54.8% in the asymptotic model. Thus, as switching rates increased from low to average levels, stereotypy levels increased, but as switching rates increased further, stereotypy levels then fell again. The relation between perseveration and switching rates was similar.

A \sim -shaped curve could conceivably reflect an artefact of data collection if extremely high rates of switching were erroneously recorded as lower rates. The rate of initiation observed in some birds did indeed push the limit of that observable on video (recorded at a standard frame speed), and thus birds may have been switching behaviours too quickly to observe. However, such an artefact would produce an asymptote, not a \sim -shaped curve. Another possibility is that extremely high levels of switching prevented birds from repeating routes the requisite number of times to be scored as stereotypic. However, if this were the case, a decrease in stereotypy complexity (i.e. shorter chain lengths) with increased levels of stereotypy would be predicted, as is seen in amphetamine stereotypy (Lyon & Robbins 1975). Such an effect was excluded in experiment 1.

Alternatively, this curve shape may reflect the summation of two processes described by a curve of decreasing return (i.e. ' \prime ') and a curve of increasing loss (i.e. ' \backslash '). Lesions or drugs up-regulating the direct pathway lead to a general activation of all behaviour, while lesions or drugs down-regulating the indirect pathway selectively induce stereotypy (Albin et al. 1989; Waddington et al. 1990). Elevated rates of switching between behaviours should result from both mechanisms (i.e. by the over-activation or disinhibition of competing behaviours, respectively). Therefore, if the highest rates of switching require both elevated direct pathway activity and depressed indirect pathway activity, and elevated direct pathway activity disrupts route-tracing stereotypy through the interfering activation of other behaviours, the interaction of these processes may explain the

observed curve. Data from bank voles provide some support for this hypothesis; bank voles show a linear relationship between bar-mouthing stereotypy and rates of switching (Garner & Mason 2002). Bar mouthing represents a far simpler motor pattern than route tracing, and is therefore likely to be more resistant to interruption and hence disruption (Lyon & Robbins 1975) in individuals with elevated direct pathway tone.

GENERAL DISCUSSION

The Stereotypes of Caged Parids

The individually caged blue and marsh tits in the present study spent 7–90% of their active time performing stereotypes, and in some, it was persistent enough to cause feather wear when their wings rubbed against the cage bars. Little work has been done on avian stereotypes, and the causal aspects of captive housing are not fully understood. However, restricted physical space may underlie similar route-tracing stereotypes in canaries (Keiper 1969, 1970), while the lack of naturalistic foraging may precipitate oral stereotypes (e.g. poultry: Savory et al. 1992). Pacing also seems to develop from escape attempts in caged poultry frustrated of access to food (Duncan & Wood-Gush 1972) and in captive emus, *Dromaius novaehollandiae*, subject to aggression (Glatz 2001). Finally, in hand-reared Hawaiian crows, *Corvus hawaiiensis*, showing the importance of early experience, birds raised in isolation are more stereotypic than birds raised with siblings. Despite the general lack of study, however, avian stereotypes have been widely reported in several species held in zoos (Casamitjana & Turner 2001), as well as in caged starlings, *Sturnus vulgaris*, other parids kept in behavioural research laboratories (e.g. S. Healy, personal communication; D. Wilson, personal communication) and in caged parrots, *Amazona amazonica* (Meehan et al. 2003; Garner et al., in press).

The stereotypes that we observed were quite variable, resembling the developing stereotypes of sows (Cronin & Wiepkema 1984) and mink (Mason 1993) more than the more extreme and rigid locomotor stereotypes attained by, say, some zoo-housed carnivores (Fentress 1976) or individually caged voles (Cooper & Nicol 1996; Cooper et al. 1996). Although such variability is perhaps surprising in birds caged for 3 years, this may reflect the fact that their caging was always interspersed with periods in a highly enriched aviary. A similar flexibility in long-term stereotypers has been observed in stabled horses, where even animals with a weaving stereotypy several years old can modulate this behaviour in response to environmental enrichment (e.g. Mills & Davenport 2002).

The Correlates of Caged Birds' Stereotypes

All four experiments showed that songbirds with high levels of cage stereotypy differed from those with low or negligible levels in several aspects of their behaviour.

They were more generally prone to repeat previous responses, as in the extinction task where they persisted in responding in the absence of reward, and in the gambling task where their responses were less random than those of low-stereotypers. They were also less inclined to reduce their levels of food storing when the repeated removal of these caches made the behaviour nonfunctional. Together, these results suggest that stereotypers are more perseverative (i.e. more prone to the inappropriate repetition of previous responses: Luria 1965). Their general home-cage behaviour also differed, linking progressively increasing rates of switching between normal home-cage activities like preening, perching, eating and flying, first with increased, and then decreased, levels of stereotypy. These altered switching rates further indicate that the birds with high cage stereotypies differed in how they sequenced their behaviours. Taken together, these results support our hypothesis that individual differences in the dorsal striatal control of behaviour underlie differences in cage stereotypy, just as they underlie drug-induced stereotypies (Lyon & Robbins 1975; Hauber 1998), the stereotypies of humans with autism and schizophrenia (Frith & Done 1983; Turner 1997), and the stereotypy-like behaviours caused by many neurodegenerative disorders (Albin et al. 1989) and brain lesions (Luria 1965; Norman & Shallice 1986). These findings in parids add to those in laboratory-reared bank voles (Garner & Mason 2002) and caged sun bears, *Helarctos malayanus*, and Asiatic black bears, *Ursus thibetanus* (Sharpe & Mason 2001, in press), whose cage stereotypies similarly correlated with poor abilities to extinguish learned responses and/or altered baseline rates of behavioural initiation. Future work could investigate these correlates of stereotypy in more detail, via experimental manipulations known to affect striatal function (e.g. isolation rearing), and more direct, invasive measures of brain function.

Comparative Issues in Avian and Mammalian Stereotypy

The similar correlates for bird and mammal stereotypies indicates that in birds, the homology of design in the dorsal striatal motor pathways (Reiner et al. 1998) reflects a homology of function in these pathways. As the association of stereotypy with dysfunction of the indirect pathway in mammals is a product of the particular design of this system and its particular role in behaviour, so the association of stereotypy in birds with signs of similar dysfunction suggests that the avian dorsal striatal motor pathways share the same peculiarities of design and function. This point is nicely underlined by the different kinds of stereotypies seen in the birds. The patterning of individual motor movements in pig oral stereotypies is beautifully described by Cronin & Wiepkema (1984). The fine analysis of the route elements in the birds' stereotypies revealed the same four distinct kinds of stereotypy as described in pigs (Figs 1, 2). Such similarities in the intricacies of an abnormal behaviour are further arguments for homologies of design and

function in the systems underlying bird and mammal behaviour.

Have Cage Stereotypers Been Altered by Captivity?

A priori individual differences provide one possible reason for our observed correlations between measures of dorsal striatal function and stereotypy: naturally occurring individual variation in behavioural control could differentially predispose animals to the development of cage stereotypy. Similar naturally occurring, baseline individual differences have been hypothesized to account for the variation that 'applied' and laboratory-housed animals show in their tendencies to develop behavioural routines (e.g. Koolhaas et al. 1999). If stereotypy is correlated with natural variation in dorsal striatal motor pathway function, it would be incorrect to claim that these pathways were dysfunctional or altered in high-stereotypy individuals.

However, an alternative explanation for our observed correlations is that captivity affects the behavioural control mechanisms of different animals to different extents, and that stereotypy levels then reflect the degree to which animals have been altered. This explanation is supported by the meta-analysis of the first three experiments. Changes in individual levels of stereotypy pre- and postenrichment were correlated with changes in the individual's perseverative tendency. Hence, individual differences were not stable. Instead, housing conditions induced changes in behavioural control mechanisms that were expressed as cage stereotypy. Such findings are consistent with evidence that dorsal striatal function is affected by abnormal rearing conditions (e.g. social isolation: Jones et al. 1991) and current stress levels (Cabib 1993; Scott et al. 1996), both common features of life in captivity. Whether this altered functioning is pathological remains an issue for future research. Overall, long-term CNS dysfunction seems likely in other cases of stereotypy (Garner & Mason 2002; Garner et al., in press), but more transient effects of stress may be implicated in other cases, such as those described here.

Implications For Animal Welfare

If stereotypy levels indicate that animals have been altered by being caged, what implications does this have for animal welfare, and also for the use of caged subjects in behavioural research?

There are several criteria for assessing welfare (Duncan & Fraser 1997; Sandøe et al. 1997; Mason 2001), of which two are particularly important. The first is based on the animal's mental or subjective experience: if it is suffering, welfare is compromised. The second considers the 'integrity' of the animal: regardless of subjective experience, if some function is compromised, welfare is compromised. Thus, if the animal had a painless but terminal illness, the first view would consider welfare to be acceptable, and the second would consider it to be poor (Mason & Mendl 1997; Mason 2001).

If stereotypy is a symptom of behavioural dysfunction, this implies the compromise of function and thus poor welfare according to the 'integrity' criterion. In terms of suffering, maladaptive behaviour (e.g. inappropriate food storing) and the disruption of complex behaviour chains may also induce both physical suffering and frustration of goal-directed behaviour. The general tendency to repeat previous responses associated with stereotypy may also lead to a novel form of frustration. In humans, stereotypies, stereotypy-like behaviours and recurrent perseveration can be experienced as frustrating, especially when they are performed against the individual's will (Luria 1965; Sandson & Albert 1984; Baron-Cohen et al. 1994; Turner 1997). If stereotypic nonhumans are similar, they may be frustrated by their difficulty in switching to highly motivated behaviours, or by their difficulty in completing complex behaviour chains, especially if such complex behaviours are critical for fitness (e.g. nest building in nursing mink: Mason et al. 1995). Thus, our results link stereotypy with poor welfare in the sense of altered function, and possibly with suffering in the form of enhanced frustration.

Implications for Scientific Validity

If housing conditions affect the control of behaviour, this has obvious implications for both the replicability and validity of laboratory-based behavioural work. From our results, we predict that, in housing conditions that induce high levels of stereotypy, experimental subjects will show alterations in response selection, extinction learning and behavioural patterning that would affect their responses in paradigms as diverse as the open-field test, activity levels and maze and operant-learning tasks. Subtle differences in husbandry, manifest as differences in stereotypy, could thus help explain some interlaboratory variability (e.g. in laboratory mouse behavioural tests: Crabbe et al. 1999; Wahlsten et al. 2003), and it could perhaps help to account for the differences in pilfering studies (Stevens 1984; Hampton & Sherry 1994; Baker & Anderson 1995; Lucas & Zielinski 1998).

The issue of scientific validity is particularly pressing to ethologists and behavioural ecologists, who assume that behaviour in the laboratory accurately reflects behavioural processes acting in the wild. If captivity makes animals generally more perseverative, this assumption is challenged. Stereotypic animals might be particularly inappropriate subjects for paradigms involving extinction learning or any other modification of a learned response, activity levels, the ability to switch between responses, behaviours or strategies, or involving complex behavioural sequences (which are more readily disrupted by the rapid switching between elements seen in stereotypic animals; Lyon & Robbins 1975). Perseveration may thus help to explain why animals in the laboratory sometimes show 'impulsivity', by responding for immediate small rewards when delaying responses for larger ones would be more productive (reviewed in Stephens & Krebs 1986) or leaving food patches prematurely (e.g. Stephens & Krebs 1986), why they may work for food, even when provided with it for free (e.g. Bean et al. 1999), why they can

delay patch-leaving in studies of optimal foraging (e.g. Stephens & Krebs 1986; Cassini et al. 1993), why they do not always modulate their risk sensitivity in response to changes in nutritional state (Abreu & Kacelnik 1999), and why they may persistently store inedible or repeatedly pilfered items (e.g. Baker & Anderson 1995; Clayton & Dickinson 1999). Rather than seeking only functional explanations for such paradoxes or attributing them to the natural constraints on optimality (see e.g. Baker & Anderson 1995; Lucas et al. 2001), it might be fruitful also to assess the extent to which altered behavioural control is a confounding factor in such experiments. We therefore suggest that either cage stereotypy be quantified and used as an explanatory variable (as is commonly done for sex, age and weight), or, preferably, that housing conditions be modified to prevent stereotypy in the first place. Overall, if stereotypies reflect captivity-induced behavioural disinhibition, then good welfare should also yield good science.

Acknowledgments

The authors acknowledge Hilary Broadbent for allowing us to work with her experimental animals and assisting with software development for the gambling task, Alan Woodington for animal care and invaluable advice on the behaviour and husbandry of the birds, Tony Price for help with the construction of apparatus, Mitch Watnick for statistical advice, Trevor Robbins, Nicky Clayton and Jonathan Newman for helpful discussions, and Cleide Falcone for assisting with video observations. J.P.G. was supported by a BBSRC Ph.D. scholarship and G.J.M. via a BBSRC Research Fellowship. This research was evaluated by the Animal Behavior Society's Animal Care Committee and approved on 5 December 2002.

References

- Abreu, F. B. E. & Kacelnik, A. 1999. Energy budgets and risk-sensitive foraging in starlings. *Behavioral Ecology*, **10**, 338–345.
- Albin, R. L., Young, A. B. & Penney, J. B. 1989. The functional anatomy of basal ganglia disorders. *Trends in Neurosciences*, **12**, 366–375.
- Alexander, G. E., Crutcher, M. D. & DeLong, M. R. 1990. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, 'prefrontal' and 'limbic' functions. *Progress in Brain Research*, **85**, 119–146.
- Attneave, F. 1959. *Applications of Information Theory to Psychology: a Summary of Basic Concepts, Methods and Results*. New York: H. Holt.
- Baker, M. C. & Anderson, P. 1995. Once-pilfered cache sites not avoided by black-capped chickadees. *Animal Behaviour*, **49**, 1599–1602.
- Baron-Cohen, S., Cross, P., Crowson, M. & Robertson, M. 1994. Can children with Gilles de la Tourette syndrome edit their intentions? *Psychological Medicine*, **24**, 29–40.
- Baxter, L. R., Schwartz, J. M., Bergman, K. S. & Szuba, M. P. 1992. Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Archives of General Psychiatry*, **49**, 681–689.
- Bean, D., Mason, G. J. & Bateson, M. 1999. Contrafreeloading in starlings: testing the information hypothesis. *Behaviour*, **136**, 1267–1282.

- Beauchamp, A. J. & Gluck, J. P.** 1988. Associative processes in differentially reared monkeys (*Macaca mulatta*): sensory preconditioning. *Developmental Psychobiology*, **21**, 355–364.
- Berkson, G.** 1968. Development of abnormal stereotyped behaviours. *Developmental Psychobiology*, **1**, 118–132.
- Cabib, S.** 1993. Neurobiological basis of stereotypies. In: *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* (Ed. by A. B. Lawrence & J. Rushen), pp. 119–145. Wallingford: CABI.
- Canadian Council on Animal Care** 2000. *CCAC Animal Use Survey, 2000* (<http://www.ccac.ca/english/facts/Facframeintro.htm>).
- Cane, V. R.** 1978. On fitting low-order Markov chains to behaviour sequences. *Animal Behaviour*, **26**, 332–338.
- Casamitjana, J. & Turner, D.** 2001. *The Official Zoo Health Check 2000: a Study of UK Zoological Collections*. Horsham: Zoo Check, Born Free Foundation.
- Cassini, M. H., Lichtenstein, G., Ongay, J. P. & Kacelnik, A.** 1993. Foraging behaviour in guinea pigs: further tests of the marginal value theorem. *Behavioural Processes*, **29**, 99–112.
- Christensen, R.** 1990. *Log-linear Models*. New York: Springer-Verlag.
- Clayton, N. S. & Cristol, D. A.** 1996. Effects of photoperiod on memory and food storing in captive marsh tits, *Parus palustris*. *Animal Behaviour*, **52**, 715–726.
- Clayton, N. S. & Dickinson, A.** 1999. Motivational control of caching behaviour in the scrub-jay, *Aphelocoma coerulescens*. *Animal Behaviour*, **57**, 435–444.
- Cooper, J. J. & Nicol, C. J.** 1991. Stereotypic behaviour affects environmental preference in bank voles, *Clethrionomys glareolus*. *Animal Behaviour*, **41**, 971–977.
- Cooper, J. J. & Nicol, C. J.** 1996. Stereotypic behaviour in wild caught and laboratory bred bank voles (*Clethrionomys glareolus*). *Animal Welfare*, **5**, 245–257.
- Cooper, J. J., Ödberg, F. & Nicol, C. J.** 1996. Limitations on the effectiveness of environmental improvement in reducing stereotypic behaviour in bank voles (*Clethrionomys glareolus*). *Applied Animal Behaviour Science*, **48**, 237–248.
- Crabbe, J. C., Wahlsten, D. & Dudek, B. C.** 1999. Genetics of mouse behavior: interactions with laboratory environment. *Science*, **284**, 1670–1672.
- Crider, A.** 1997. Perseveration in schizophrenia. *Schizophrenia Bulletin*, **23**, 63–74.
- Cronin, G. M. & Wiepkema, P. R.** 1984. An analysis of stereotyped behavior in tethered sows. *Annales de Recherches Veterinaires*, **15**, 263–270.
- Duncan, I. J. & Wood-Gush, D. G.** 1972. Thwarting of feeding behaviour in the domestic fowl. *Animal Behaviour*, **20**, 444–451.
- Duncan, I. J. H. & Fraser, D.** 1997. Understanding animal welfare. In: *Animal Welfare* (Ed. by M. C. Appleby & B. O. Hughes), pp. 19–31. Wallingford: CABI.
- European Commission** 1997. *The Welfare of Intensively Kept Pigs: Report of the Scientific Veterinary Committee*. XXIV/B3/ScVC/0005/1997 (http://europa.eu.int/comm/food/fs/sc/oldcomm4/out17_en.pdf).
- European Commission** 1999. *Second Report on the Statistics on the Number of Animals Used for Experimental and Other Scientific Purposes in the Member States of the European Union*. COM (99)191 (http://europa.eu.int/comm/environment/docum/99191_en.htm).
- European Commission** 2000. *The Welfare of Chickens Kept for Meat Production (Broilers): Report of the Scientific Committee on Animal Health and Animal Welfare*. SANCO.B.3/AH/R15/2000 (http://europa.eu.int/comm/food/fs/sc/scah/out39_en.pdf).
- Evenden, J. L. & Robbins, T. W.** 1983. Increased response switching, perseveration and perseverative switching following d-amphetamine in the rat. *Psychopharmacology*, **80**, 67–73.
- Fentress, J. C.** 1976. Dynamic boundaries of patterned behaviour: interaction and self-organisation. In: *Perspectives in Ethology* (Ed. by P. P. G. Bateson & R. A. Hinde), pp. 155–224. New York: Plenum.
- Francis, S.** 1966. An ethological study of mentally retarded individuals and normal individuals. Ph.D. thesis, University of Cambridge.
- Francis, S. H.** 1971. The effects of own-home and institution-rearing on the behavioural development of normal and mongol children. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, **12**, 173–190.
- Frith, C. D. & Done, D. J.** 1983. Stereotyped responding by schizophrenic-patients on a 2-choice guessing task. *Psychological Medicine*, **13**, 779–786.
- Frith, C. D. & Done, D. J.** 1990. Stereotyped behaviour in madness and in health. In: *Neurobiology of Stereotyped Behaviour* (Ed. by S. J. Cooper & C. T. Dourish), pp. 232–259. Oxford: Clarendon.
- Garner, J. P.** 1999. The aetiology of stereotypy in caged animals. D.Phil. thesis, University of Oxford.
- Garner, J. P. & Mason, G. J.** 2002. Evidence for a relationship between cage stereotypies and behavioural disinhibition in laboratory rodents. *Behavioural Brain Research*, **136**, 83–92.
- Garner, J. P., Mason, G. J. & Broadbent, H.** 1998. Cage stereotypy in blue tits (*Parus caeruleus*) and marsh tits (*Parus palustris*) is associated with a deficit in the inhibition of nonfunctional behaviour. In: *Proceedings of the 32nd International Congress of the International Society for Applied Ethology* (Ed. by I. Veissier & A. Boissy), page 56. Clermont-Ferrand, France: INRA.
- Garner, J. P., Meehan, C. L. & Mench, J. A.** In press. Stereotypies in caged parrots, schizophrenia and autism: evidence for a common mechanism. *Behavioural Brain Research*.
- Glatz, P. C.** 2001. Effect of declawing on behavior of farmed emus. *Asian-Australasian Journal of Animal Sciences*, **14**, 288–296.
- Hampton, R. R. & Sherry, D. F.** 1994. The effects of cache loss on choice of cache sites in black-capped chickadees. *Behavioral Ecology*, **5**, 44–50.
- Hauber, W.** 1998. Involvement of basal ganglia transmitter systems in movement initiation. *Progress in Neurobiology*, **56**, 507–540.
- Hogg, R. V. & Craig, A. T.** 1978. *Introduction to Mathematical Statistics*. New York: Macmillan.
- Jones, G. H., Marsden, C. A. & Robbins, T. W.** 1991. Behavioral rigidity and rule-learning deficits following isolation-rearing in the rat: neurochemical correlates. *Behavioural Brain Research*, **43**, 35–50.
- Joyce, E. M. & Iversen, S. D.** 1984. Dissociable effects of 6-OHDA-induced lesions of neostriatum on anorexia, locomotor-activity and stereotyp: the role of behavioral competition. *Psychopharmacology*, **83**, 363–366.
- Keiper, R. R.** 1969. Causal factors of stereotypies in caged birds. *Animal Behaviour*, **17**, 114–119.
- Keiper, R. R.** 1970. Studies of stereotype function in the canary (*Serinus canarius*). *Animal Behaviour*, **18**, 353–357.
- Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H., De Jong, I. C., Ruis, M. A. W. & Blokhuis, H. J.** 1999. Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience & Biobehavioral Reviews*, **23**, 925–935.
- Lewis, M. H., Gluck, J. P., Bodfish, J. W. & Beauchamp, A. J.** 1996. Neurobiological basis of stereotyped movement disorder. In: *Stereotyped Movements: Brain and Behavior Relationship* (Ed. by R. L. Sprague & K. M. Newell), pp. 37–67. Washington, D.C.: American Psychological Association.
- Lucas, J. R. & Zielinski, D. L.** 1998. Seasonal variation in the effect of cache pilferage on cache and body mass regulation in Carolina chickadees: what are the trade-offs? *Behavioral Ecology*, **9**, 193–200.
- Lucas, J. R., Pravosudov, V. V. & Zielinski, D. L.** 2001. A reevaluation of the logic of pilferage effects, predation risk, and environ-

- mental variability on avian energy regulation: the critical role of time budgets. *Behavioral Ecology*, **12**, 246–260.
- Luria, A. R.** 1965. Two kinds of motor perseveration in massive injury of the frontal lobes. *Brain*, **88**, 1–11.
- Lyon, M. & Robbins, T.** 1975. The action of central nervous system stimulant drugs: a general theory concerning amphetamine effects. *Current Developments in Psychopharmacology*, **2**, 79–163.
- McGreevy, P. D., French, N. P. & Nicol, C. J.** 1995. The Prevalence of abnormal behaviors in dressage, eventing and endurance horses in relation to stabling. *Veterinary Record*, **137**, 36–37.
- Manschreck, T. C., Maher, B. A. & Ader, D. N.** 1981. Formal thought disorder, the type-token ratio, and disturbed voluntary motor movement in schizophrenia. *British Journal of Psychiatry*, **139**, 7–15.
- Martin, P. & Bateson, P.** 1986. *Measuring Behaviour: an Introductory Guide*. Cambridge: Cambridge University Press.
- Mason, G. J.** 1991a. Stereotypies: a critical review. *Animal Behaviour*, **41**, 1015–1037.
- Mason, G. J.** 1991b. Stereotypies and suffering. *Behavioural Processes*, **25**, 103–115.
- Mason, G. J.** 1993. Forms of stereotypic behaviour. In: *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* (Ed. by A. B. Lawrence & J. Rushen), pp. 7–40. Wallingford: CABI.
- Mason, G. J.** 2001. Animal welfare. In: *Frontiers of Life. Vol. IV, Part One* (Ed. by P. G. Bateson & E. Alleva), pp. 29–43. San Diego: Academic Press.
- Mason, G. J. & Mendl, M.** 1997. Do the stereotypies of pigs, chickens and mink reflect adaptive species differences in the control of foraging? *Applied Animal Behaviour Science*, **53**, 45–58.
- Mason, G. J., Leipoldt, A. & de Jonge, G.** 1995. Why do female mink with high stereotypy levels have slow-growing offspring? In: *Proceedings of the 29th International Congress of the International Society for Applied Ethology* (Ed. by S. M. Rutter, J. Rushen, H. D. Randle & J. C. Eddison), pp. 133–134. Exeter: Exeter University.
- Medina, L. & Reiner, A.** 1997. The efferent projections of the dorsal and ventral pallidal parts of the pigeon basal ganglia, studied with biotinylated dextran amine. *Neuroscience*, **81**, 773–802.
- Meehan, C. L., Garner, J. P. & Mench, J. A.** 2003. Isosexual pair housing improves the welfare of young Amazon parrots. *Applied Animal Behaviour Science*, **80**, 73–88.
- Mills, D. S. & Davenport, K.** 2002. The effect of a neighbouring conspecific versus the use of a mirror for the control of stereotypic weaving behaviour in the stabled horse. *Animal Science*, **74**, 95–101.
- Morgan, M. J., Einon, D. F. & Nicholas, D.** 1975. The effects of isolation rearing on behavioural inhibition in the rat. *Quarterly Journal of Experimental Psychology*, **27**, 615–634.
- National Animal Ethics Advisory Committee** 2002. *National Animal Ethics Advisory Committee 2001 Annual Report* (<http://www.maf.govt.nz/biosecurity/animal-welfare/naeac/annual-report/naeac-ar-01.pdf>).
- Neivon, C. M., Hurst, J. L. & Barnard, C. J.** 1999. Why do male ICR(CD-1) mice perform bar-related (stereotypic) behaviour? *Behavioural Processes*, **47**, 95–111.
- Nicol, C. J.** 1999. Understanding equine stereotypies. *Equine Vet Journal, Supplement*, **28**, 20–25.
- Norman, D. A. & Shallice, T.** 1986. Attention to action: willed and automatic control of behaviour. In: *Consciousness and Self-regulation: Advances in Research and Theory* (Ed. by R. J. Davidson, G. E. Schwartz & D. Shapiro), pp. 1–18. New York: Plenum.
- Owens, D. G. C., Johnstone, E. C. & Frith, C. D.** 1982. Spontaneous involuntary disorders of movement: their prevalence, severity and distribution in chronic-schizophrenics with and without treatment with neuroleptics. *Archives of General Psychiatry*, **39**, 452–461.
- Reiner, A., Medina, L. & Veenman, C. L.** 1998. Structural and functional evolution of the basal ganglia in vertebrates. *Brain Research Reviews*, **28**, 235–285.
- Robbins, T. W.** 1997. Arousal systems and attentional processes. *Biological Psychology*, **45**, 57–71.
- Robbins, T. W., Jones, G. H. & Wilkinson, L. S.** 1996. Behavioural and neurochemical effects of early social deprivation in the rat. *Journal of Psychopharmacology*, **10**, 39–47.
- Robbins, T. W., Mittelman, G., O'Brien, J. & Winn, P.** 1990. The neuropsychological significance of stereotypy induced by stimulant drugs. In: *Neurobiology of Stereotyped Behaviour* (Ed. by S. J. Cooper & C. T. Dourish), pp. 25–63. Oxford: Clarendon.
- Rolls, E. T.** 1994. Neurophysiology and cognitive functions of the striatum. *Revue Neurologique*, **150**, 648–660.
- Rushen, J., dePassile, A. M. B. & Schouten, W.** 1990. Stereotypic behavior, endogenous opioids, and postfeeding hypoalgesia in pigs. *Physiology & Behavior*, **48**, 91–96.
- Rushen, J., Lawrence, A. B. & Terlouw, E. M. C.** 1993. The motivational basis of stereotypies. In: *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* (Ed. by A. B. Lawrence & J. Rushen), pp. 41–64. Wallingford: CABI.
- Sandøe, P., Crisp, R. & Holtug, N.** 1997. Ethics. In: *Animal Welfare* (Ed. by M. C. Appleby & B. O. Hughes), pp. 3–18. Wallingford: CABI.
- Sandson, J. & Albert, M. L.** 1984. Varieties of perseveration. *Neuropsychologia*, **22**, 715–732.
- Savory, C. J., Seawright, E. & Watson, A.** 1992. Stereotyped behaviour in broiler breeders in relation to husbandry and opioid receptor blockade. *Applied Animal Behaviour Science*, **32**, 349–360.
- Scott, P. A., Cierpial, M. A., Kilts, C. D. & Weiss, J. M.** 1996. Susceptibility and resistance of rats to stress-induced decreases in swim-test activity: a selective breeding study. *Brain Research*, **725**, 217–230.
- Sharpe, S. & Mason, G. J.** 2001. Investigating the severity of stereotypy in captive bears. In: *Proceedings of the 35th International Congress of the International Society for Applied Ethology* (Ed. by J. P. Garner, J. A. Mench & S. Heekin), page 139. Davis: Center for Animal Welfare, University of California.
- Sharpe, S. & Mason, G. J.** In press. Behavioural persistence in captive bears: implications for reintroduction. *Ursus*.
- Sørensen, G. & Randrup, A.** 1986. Possible protective value of severe psychopathology against lethal effects of an unfavourable milieu. *Stress Medicine*, **2**, 103–105.
- Stephens, D. W. & Krebs, J. R.** 1986. *Foraging Theory*. Princeton, New Jersey: Princeton University Press.
- Stevens, T. A.** 1984. Food-storing by marsh tits. D.Phil. thesis, Oxford University.
- Terlouw, E. C., Lawrence, A. B. & Illius, A. W.** 1992. Relationship between amphetamine and environmentally induced stereotypies in pigs. *Pharmacology, Biochemistry & Behavior*, **43**, 347–355.
- Terlouw, E. M., Lawrence, A. B. & Illius, A. W.** 1991. Influences of feeding level and physical restriction on development of stereotypies in sows. *Animal Behaviour*, **42**, 981–991.
- Turner, M.** 1997. Towards an executive dysfunction account of repetitive behaviour in autism. In: *Autism as an Executive Disorder* (Ed. by J. Russell), pp. 57–100. Oxford: Oxford University Press.
- University of California Center for Animal Alternatives** 1996. The Mouse in Science: Why Mice? (http://www.vetmed.ucdavis.edu/Animal_Alternatives/whymice.htm).
- Waddington, J. L., Molloy, A. G., O'Boyle, K. M. & Pugh, M. T.** 1990. Aspects of stereotyped and non-stereotyped behaviour in relation to dopamine receptor subtypes. In: *Neurobiology of Stereotyped Behaviour* (Ed. by S. J. Cooper & C. T. Dourish), pp. 64–90. Oxford: Clarendon.
- Wahlsten, D., Metten, P., Phillips, T. J., Boehm, S. L., Burkhardt-Kasch, S., Dorow, J., Doerksen, S., Downing, C., Fogarty, J., Rodd-Henricks, K., Hen, R., McKinnon, C. S., Merrill, C. M.,**

- Nolte, C., Schalomon, M., Schlumbohm, J. P., Sibert, J. R., Wenger, C. D., Dudek, B. C. & Crabbe, J. C. 2003. Different data from different labs: lessons from studies of gene-environment interaction. *Journal of Neurobiology*, **54**, 283–311.
- Wiedenmayer, C. 1996. Effect of cage size on the ontogeny of stereotyped behaviour in gerbils. *Applied Animal Behaviour Science*, **47**, 225–233.
- Wiedenmayer, C. 1997. Causation of the ontogenetic development of stereotypic digging in gerbils. *Animal Behaviour*, **53**, 461–470.
- Würbel, H., Stauffacher, M. & vonHolst, D. 1996. Stereotypies in laboratory mice: quantitative and qualitative description of the ontogeny of 'wire-gnawing' and 'jumping' in Zur:ICR and Zur:ICR nu. *Ethology*, **102**, 371–385.