



Middle-aged mice with enrichment-resistant stereotypic behaviour show reduced motivation for enrichment

Sarah-Lee C. Tilly^{a,1}, Jamie Dallaire^b, Georgia J. Mason^{b,*}

^a *Ethology & Welfare Centre (ECW), Department of Animals, Science and Society, Faculty of Veterinary Medicine, Utrecht University*

^b *Animal and Poultry Science, University of Guelph*

ARTICLE INFO

Article history:

Received 6 October 2009
Initial acceptance 9 November 2009
Final acceptance 4 June 2010
Available online 17 July 2010
MS. number: 09-00652R

Keywords:

age
animal welfare
enrichment
establishment
laboratory mice
motivation
Mus musculus
reward value
stereotypic behaviour

For captive animals, living in barren conditions leads to stereotypic behaviour that is hard to alleviate using environmental enrichment. This resistance to enrichment is often explained via mechanisms that decouple abnormal behaviour from current welfare, such as 'establishment': a hypothetical process whereby repetition increases behaviour's predictability and resistance to change. If such hypotheses are correct, then animals with enrichment-resistant stereotypic behaviour should still find enrichments rewarding. Alternatively, this behaviour could reflect a failure to improve welfare: plausible because age and chronic stress increase neophobia and anhedonia. If this hypothesis is correct, animals with enrichment-resistant stereotypic behaviour should value enrichments less than conspecifics. We tested these hypotheses using C57BL/6 mice, *Mus musculus*, aged 10–11 and 6–7 months, raised in barren laboratory cages. We observed their behaviour in both these and large enriched cages. Enrichment was more effective on the younger animals. However, contrary to ideas about establishment, the spontaneous predictability of stereotypic behaviour did not increase with age; nor was enrichment less effective on more predictable or time-consuming forms. We assessed the reward value of enriched cages by allowing access via progressively weighted doors (maximum weight pushed corresponding to peak motivation). In older mice, those individuals whose stereotypic behaviour was least reduced by enrichment were also the least motivated to gain access to enrichment. This suggests that the welfare of middle-aged-animals, as well as their stereotypic behaviour, is harder to improve using environmental enrichment.

© 2010 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Over 85 million captive animals worldwide display stereotypic behaviours, traditionally defined as repetitive, unvarying and without apparent function (Mason & Latham 2004). Examples include pacing in caged carnivores, rocking by isolated primates and sham chewing by hungry sows. These behaviours are ethically worrying because they are statistically associated with environments, treatments and feeding regimes that cause poor welfare (Mason & Latham 2004). They also reduce the appeal and educational value of zoos (Swaigood & Shepherdson 2006), and may indicate dysfunction of the forebrain's basal ganglia (Garner et al. 2003; Mason 2006; Lewis et al. 2007). Animal carers in zoos, stables and research facilities therefore often try to prevent or reduce them, most commonly by using environmental enrichment (Shepherdson et al. 1998; Young 2003). This involves providing rewarding stimuli, objects and substrates that trigger interaction

and naturalistic behaviour (e.g. Young 2003; Swaisgood & Shepherdson 2006). Compared to other means of preventing or reducing stereotypic behaviour (e.g. punishment, physical prevention or pharmacological treatment), enrichment is the preferred approach because of its potential to alleviate subjects' underlying poor welfare, not just their undesirable behaviour (Mason et al. 2007). Successful examples include the provision of naturalistic burrows and nest chambers to Mongolian gerbils, *Meriones unguiculatus*, which prevents the emergence of stereotypic digging (Weidenmayer 1997); and socially housing Amazon parrots, *Amazona amazonica*, which prevents development of the abnormal oral and cage-climbing behaviours common in isolated birds (Meehan et al. 2003).

However, despite such successes, some individuals' stereotypic behaviour seems very hard to reduce, even with improved housing that does reduce these behaviours in conspecifics. To illustrate, the earliest accounts of stereotypic behaviours included anecdotal reports of polar bears, *Ursus maritimus*, moved from small circus wagons to large zoo enclosures (Meyer-Holzappel 1968): these animals spent more time in stereotypic behaviour than typical zoo bears, and performed forms still delineated by their former tiny cages (see also Morris 1964 and Burghardt 2005 for similar

* Correspondence: G. J. Mason, Animal and Poultry Science, University of Guelph, Guelph, Ontario N1G 2W1, Canada.

E-mail address: gmason@uoguelph.ca (G.J. Mason).

¹ S.-L. C. Tilly is at the Ethology & Welfare Centre (ECW), Department of Animals, Science and Society, Faculty of Veterinary Medicine, Utrecht University, PO Box 80, 1663508 TD Utrecht, The Netherlands.

anecdotes from zoos). Similarly, Indian temple elephants, *Elephas maximus*, are shackled much of the time and so are more stereotypic than animals working in logging camps, but they reportedly remain so even when transferred from temple to camp environments (Kurt & Hartl 1995). These and similar observations led to suggestions that long-term exposure to impoverished environments, and/or prolonged performance of stereotypic behaviour, renders this behaviour particularly enrichment resistant (e.g. Mason 1991). Subsequent hypothesis-led experiments on rodents confirmed the importance of length of time in barren environments, and/or its correlates: sustained stereotyping and age.

Thus in one study, bank voles, *Clethrionomys glareolus*, showed a statistically significant effect of age (equivalent to length of time in standard cages) on enrichment efficacy (Cooper et al. 1996). Stereotypic behaviour was abolished by enrichment in voles aged 2 months (early adulthood) at the start of the experiment, falling from a mean of ca. 15% of observations to zero. However in voles aged 14 months at the start of the experiment (and elderly, 18 months, by the end), it fell by less than half from a mean of ca. 20% of observations to ca. 13%. In pilot work by the same researchers (Cooper & Ödberg 1991), the prevalence of stereotypic behaviour was also affected by interactions between age and enrichment: voles aged 2, 4 or 6 months at the start of this study all stopped stereotyping after 2 months of enriched housing, while voles aged 16 months all continued (although to what extent is unreported). Similarly, Hadley et al. (2006) demonstrated that age affects the impact of enrichment on stereotypic behaviour in deer mice, *Peromyscus maniculatus*. Here, deer mice were raised and housed in standard laboratory mouse cages. When aged either 3 months (young adults) or 11–14 months ('middle age' in this species, whose average life span is about double that of bank voles), they were either kept in such cages for 2 more months, or transferred for the same period to large dog kennels whose enrichments included ramps, a running wheel, plastic toys and scattered bird seed. In the young deer mice, 'counts/h' of stereotypic behaviour were approximately halved by enriched housing (ca. 900 for enriched subjects but ca. 1800 for controls); but in the middle-aged deer mice, both control and enriched animals showed a similar decline over the 2-month period (from ca. 1000 counts/h to ca. 600), with no effect of housing type.

All these cases seemingly involved improved housing, and so have typically been hypothesized to indicate a dissociation between stereotypic behaviour and current wellbeing. Stereotypic behaviours in older animals have thus been termed 'pure motoric automatism ...' (Kennes et al. 1988, page 20); independent from environmental and motivational variables' (Kennes et al. 1988, page 23); while Mason (1991, page 111) proposed that 'like a scar, a stereotypy tells us something about past events'; Ames (1994) stated that these behaviours tend to persist long after their initial causes have been removed; and Mason & Latham (2004, page S66) argued that 'enrichments that do not immediately reduce stereotypies should not be assumed failures with respect to welfare'. Some circumstantial evidence supports this view (for instance, isolated female rhesus monkeys, *Macaca mulatta*, given cagemates apparently benefited in terms of welfare, for example their infants thrived better, despite no decline in stereotypic behaviour; Eaton et al. 1994); and there are two plausible mechanisms that could feasibly dissociate the performance of stereotypic behaviour from current wellbeing (Mason & Latham 2004). The first is habit or routine formation, where repetition causes the control of a behaviour pattern to become less dependent on external circumstances (Fentress 1976, 1977; Benus et al. 1990; Golani et al. 1999). In a process sometimes called the 'establishment' of stereotypic behaviour (e.g. Ödberg 1978; Würbel 2006), repeated performance is thus suggested to render stereotypic behaviour both more

predictable (i.e. reiterations become increasingly similar to each other) and less influenced by changes in the external environment or affective state (see also Cronin 1985; Würbel 2006; Dodman et al. 2010). The second is perseveration, the tendency to repeat recently performed or otherwise prepotent behaviours inappropriately (e.g. Hotz & Helm-Estabrooks 1995). Perseveration is a known correlate of stereotypic behaviour inappropriately (e.g. Garner et al. 2003; Garner 2006), and can reflect changes in basal ganglia-mediated behavioural control. Stress and impoverished rearing environments both affect these forebrain systems, thence causing general impairments in abilities to suppress behavioural responses that are contextually inappropriate (reviewed in Mason 2006).

However, an alternative possible explanation for enrichment-resistant stereotypic behaviours is that affected animals find added 'enrichments' unrewarding or even aversive, like, for example, the zoo primates reportedly by Di Giovanni & Valente (2001) as frightened of new enrichments placed in their enclosures (Stevenson 1983 reported similar observations). Older individuals, in particular, often show little interest in environmental enrichments particularly if housed long term in impoverished conditions (e.g. Line et al. 1991; Novak et al. 1993; Baker 2000; Vickery 2003). Cooper & Nicol (1991) even found seemingly lowered preferences for naturalistic cover (e.g. hay) in stereotypic bank voles, as they aged between 2 and 5 months old (older subjects were not studied). Potential explanations for such aversion or indifference include preference for sameness and routine, as shown by autistic humans (for whom stereotypic behaviour is diagnostic, and perseveration typical; Hutt 1969; Turner 1999); neophobia, the fear of novelty; and/or anhedonia, a reduced capacity to feel pleasure (e.g. Snaith 1993). Neophobia and anhedonia increase with chronic stress (Papp et al. 1991; Baker 2000; Bondi et al. 2008), with age (Imhof et al. 1993; Herrera-Pérez et al. 2008), and with being raised in deprived housing conditions (Stevenson 1983; Novak et al. 1993), suggesting that they could well increase in animals held long term in barren conditions. Such hypotheses have not been tested, however; and nor has indifference or aversion to enrichments been investigated as a cause of enrichment-resistant stereotypic behaviour. If confirmed, this would have considerable implications for animal welfare. Here, we therefore investigated the reward value of enrichment to animals with enrichment-resistant stereotypic behaviours, testing the hypothesis that such individuals value enrichment less. We used laboratory mice, *Mus musculus*: model stereotypers as well as applied subjects in their own right. Reward value was assessed in a closed economy apparatus (Hursh 1980), via the maximum price subjects would pay ('breakpoint': e.g. Loh & Roberts 1990) to gain access to enrichment. The experiment additionally allowed us to test hypotheses about the potential establishment of stereotypic behaviour, which predicts positive associations between high degrees of repetition (greater daily time budgets devoted to stereotypic behaviour, and longer times in captivity), high spontaneous predictability and resistance to enrichment.

METHODS

Subjects were female C57BL/6 laboratory mice, purchased from the Jackson Laboratory, Maine, U.S.A. 'Young adults' (see Flurkey et al. 2007) were 6–7 months old at the start of the experiment, while 'Middle-aged adults' were 10–11 months. Throughout the experiment they were housed in pairs, since isolation is not recommended for female mice on welfare grounds (e.g. Sherwin 2003). Cagemates were the same age, and distinguished by hair-clipped marks redone every 4 weeks. Age was not visible on cage cards: thus the observer was always blind to mouse age throughout

data collection. They were kept in a standard animal facility room, on a reversed light cycle (lights-off at 1100 hours, on at 2300 hours), with supplementary red light available during the dark phase to allow observation. Their 'Standard' cages were of a type they had lived in all their lives: a typical 'shoebox' laboratory mouse cage (25 × 15 cm and 10 cm high) with clear plastic walls, woodchip bedding, a transparent red plastic shelter, a single 'nestlet' nesting square (replenished every week), and ad libitum food and water. The experiment involved two phases. In Phase 1, mice were assessed for stereotypic behaviour in both their original cages (Phase 1a) and in new, large enriched cages (Phase 1b). In Phase 2 they were re-paired with a new same-age cagemate that had shown a similar response to enrichment, and each new pair was tested for their strength of motivation to get access to the large enriched cage.

Phase 1: Impact of Environmental Enrichment

Pilot observations of our initial pool of 42 cages (84 mice) over three dark periods identified 24 cages in which both mice performed stereotypic behaviour (14 Young; 10 Middle-aged). Along with previous rodent studies (Cooper 1992; Cooper et al. 1996; Garner 1999; Latham 2004), these pilot data also allowed us to finalize an ethogram (Table 1) and choose an observation period for the experiment proper (a 4 h period starting 2 h after lights-off).

The 24 cages in which both cagemates showed stereotypic behaviour were then observed to quantify their behaviour in Standard conditions. Cages were scanned in a consistent order by a quiet observer once every 10 min for 4 h. As a row was scanned, instantaneous sampling was used to record patterns that were identifiable immediately and apparently normal (e.g. sleeping, standing, lying alert). However, this was unsuitable for putative stereotypic behaviours which are identified by their degree of repetition. In such instances the observer switched to focal observations, activity being continuously observed for up to 15 s or until a regular pattern emerged (whichever was the earlier). Behaviour was classified into one of four broad categories (Table 1). This observation regime was repeated for 8 consecutive days.

The 24 pairs were then transferred to new, large enriched cages. These contained bedding and ad libitum food and water (just as in standard cages), but measured 60 × 60 cm and were 25 cm high (24 times the volume of the previous cages), and contained a variety of enrichments selected on the basis of online recommendations for pet mice as well as published research (e.g. Roper 1975; Sherwin 1997; Van de Weerd et al. 1997, 1998a, b, 2002; Würbel et al. 1998; Van Loo et al. 2002; Augustsson et al. 2003; Pietropaolo et al. 2004). These included a piece of cardboard and cardboard tubes as shelters; nesting material, including shredded paper packed into both a washed yoghurt pot suspended with wire from the cage lid and into the cut-off foot of a sock; washed yoghurt

pots on the cage floor for shelter; autoclaved natural sticks and pinecones to chew; and the leg of a sock suspended by wire, tunnel- or hammock-like, from the cage lid. All cages had the same number and type of enrichments, all replenished at the same rate. Each enriched cage also contained one empty standard cage, with a mouse-sized access hole drilled in its side, so that any mouse motivated to continue stereotyping as before, using the original cage lid, could easily do so. After an acclimatization period of 24 h, behavioural data were collected as before, for 12 consecutive days, followed by a 2-day break (for data analysis), then a final 12 days.

For each mouse in each of Phases 1a and 1b, 'Unambiguous stereotypic behaviour' and 'Borderline stereotypic behaviour' values (Table 1) were summed to give 'Total stereotypic behaviour' (see Mason 1993). Similar to the 'rigidity index' calculations of Mason (1992, 1993), Clubb (2001) and Vickery (2003), 'predictability' was defined as the proportion of Total stereotypic behaviour comprising Unambiguous stereotypic behaviour. 'Predictability' was the preferred term for describing the behaviour's spontaneous morphological variability within a bout, since in this context the term 'rigidity' might erroneously imply resistance to enrichment. Note that 'Total stereotypic behaviour' is the value used in all subsequent analyses.

Phase 2: Motivation for Access to Enriched Cages

For each mouse, the effect of enrichment on the proportion of observations spent stereotyping was calculated in three different ways (see *Statistical analyses* below). Before the start of Phase 2, these three indices were used to guide the re-pairing of individuals, to re-cage each mouse with one whose stereotypic behaviour had been similarly reduced. This was because our motivation-assessing apparatus could not distinguish between cagemates, and we wanted to reduce the nonsystematic error potentially added by having cagemates within a single apparatus with very different levels of predicted motivation for enrichment. To create new pairs that were similarly impacted by enrichment, all 48 individuals were ranked by their scores in each of the three indices. New cagemates had to be within five ranks for at least two indices: an arbitrary, ad hoc rule allowing us to use all mice as best we could. Extreme divergences in the third of the three ranks were also avoided wherever possible. New cagemates also had to be in the same age group. So that all subjects had similar experiences at this time, all mice, even those resembling their original cagemates in the impact of enrichment, were re-caged with new cagemates.

To measure our subjects' motivations for the large enriched cage, each new pair was then placed in an apparatus consisting of a standard and an enriched cage connected by two clear plastic tunnels (length: 20 cm; diameter: 5 cm; Fig. 1a). One tunnel served as an entrance to the enriched cage, the other as an exit, with one-way passage being ensured via metal one-way doors (taken from

Table 1
Ethogram of mouse behaviours

Behaviour observed	Descriptions	Behavioural category
Sleeping	Lying or sitting curled up, with muzzle out-of-sight; or lying still out-of-sight in sock or nesting material	Inactive
Still & alert	Animal motionless, muzzle in sight and eyes open	Inactive
General activity	All locomotive behaviour performed on the lid and cage floor, except stereotypic behaviour	Normal activity
Maintenance	Nest building, grooming, feeding and drinking	Normal activity
Bar mouthing	A cage bar is held in diastema for 1 s or longer, while the mouse makes a series of sham-biting movements along the bar; or wipes the open mouth along the bar (sometimes while rotating or moving the head backwards and forwards)	Unambiguous stereotypic behaviour
Route tracing	Running on the cage floor along a fixed route, for three repetitions of the same path	Unambiguous stereotypic behaviour
Patterned climbing about cage lid	Climbing about the cage lid (hanging upside down by the paws) along a fixed route, for three repetitions of the same path	Unambiguous stereotypic behaviour
Other	Route tracing/patterned climbing involving only two repetitions; bar mouthing lasting under 1 s	Borderline stereotypic behaviour

commercial mouse traps designed to catch multiple live mice: Victor 'Poly Cat' and 'Tin Cat' models; Fig. 1b) that allowed just one mouse to pass at a time. Entering the enriched cage was made costly by way of progressively weighing the one-way entrance door with metal discs (Fig. 1a, b). Similar access costs have been used to measure motivation in mink, *Mustela vison* (e.g. Mason et al. 2001), poultry, *Gallus gallus* (Olsson & Keeling 2002) and laboratory mice (Latham 2004; Latham & Mason 2010). Food and water were available only in the standard cage; thus mice that entered the enriched cage had to return 'home' to meet basic homeostatic needs. This return to the standard cage was always free, the door being unweighted. Mice that chose not to gain access to enrichments therefore never had to leave the standard cage. Note that one

end of the entrance tunnel (that leading into the enriched cage) was covered daily by a piece of tissue with a small square hole in its centre. The mice had to rip this paper to enter the enriched cage, the rip then indicating to us that successful passage had occurred that day.

The first 3 days in this apparatus comprised 'training' whereby mice were gradually familiarized with the tunnels, one-way doors and tissue paper barrier. After 3 days, all subjects were successfully using both tunnels and both the Standard and Enriched parts of the apparatus. On the 1st day of Phase 2's data collection period, mice had unweighted ('free') access to the enriched cage. On the 2nd day, a screw and bolt (in total 3.5 g) were added to the one-way access door. On the 4th day, and every 48 h from then on, as long as the

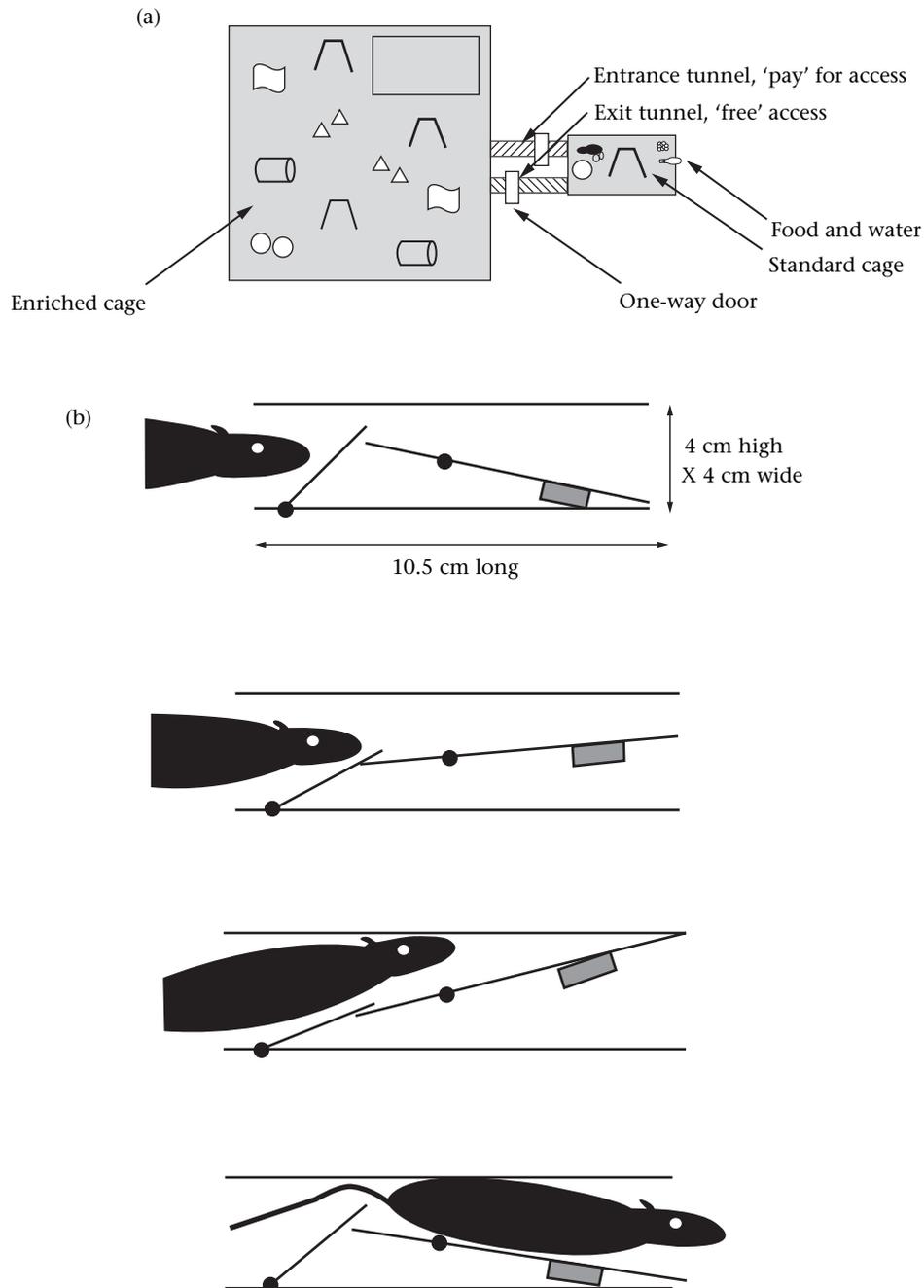


Figure 1. (a) Apparatus for assessing the motivation of mice to gain access to enrichments. (b) The weighted one-way entrance door (grey rectangle indicates the added weights). The dimensions were such that only one mouse could pass through at any one time.

preceding weight had been pushed and the enrichments accessed, a preweighed metal disc (a washer weighing ca. 4 g; range 3.8–4.2 g) was added. This method follows that validated by Latham (2004). Each time the weight was changed, a new, intact tissue paper screen was also placed over the entrance to the enriched cage. This procedure continued until neither cagemate gained access to enrichments for 2 consecutive days at the same weight; i.e. the tissue paper showed no sign of passage for 48 h. At this point, no further weight was added, and the cage was monitored for a further 48 h for signs of passage. If none occurred, the maximum weight pushed to get access to the enrichments (i.e. the exact weight pushed prior to this final weight) was noted as the 'breakpoint', and the experiment terminated by removing all weights so that the subjects again had free (unmonitored) enrichment access.

Ethical Note

To minimize experimental animal numbers, we used surplus mice from another research project; as a result our population had a less-than-ideal characteristic: ca. 40% of the mice were related (mothers and daughters) and ca. 60% were not. For how this was dealt with analytically see below. We pair-housed mice to avoid isolation stress, even though this made hypothesis testing in Phase 2 more challenging; and we avoided measuring the breakpoint for food to assess the maximum possible weight each pair could push, because the criterion is that animals fail to acquire food for a substantial period (24 h+). After the experiment, mice were euthanized with carbon dioxide. This is not an ideal method but was made less stressful by placing each pair's undisturbed home cage in a tank, which was then gradually filled with the gas. No animal showed severe signs of distress and all were unconscious within 2–3 min. The work met CCAC standards and was approved by the University of Guelph Animal Care Committee.

Statistical Analyses

Analyses tested the following hypotheses: that the time budget and predictability of stereotypic behaviour increase with age; that more time-consuming stereotypic behaviour is also more predictable; that time spent stereotyping is reduced by enrichment; that enrichment reduces stereotypic behaviour less effectively in older mice and/or mice with more predictable or higher pre-enrichment levels of stereotypic behaviour; that older mice show a slower rate of adjustment to the enriched cages; and that the effectiveness of enrichment in terms of reducing stereotypic behaviour correlates positively with the degree to which mice found the enriched cages rewarding. Prior to analyses, all data were checked for reliability by confirming that values collected on odd and even days of data collection covaried (see Martin & Bateson 2007).

Predictability in Phase 1a was calculated as the ratio of observations of Unambiguous stereotypic behaviour:Total stereotypic behaviour (where Total stereotypic behaviour = Unambiguous stereotypic behaviour + Borderline stereotypic behaviour, see Table 1).

In Phase 1, the effect of enrichment on the proportion of observations spent stereotyping was calculated in the following three ways: (1) Absolute reduction: 'Standard' stereotypic behaviour minus 'Enriched' stereotypic behaviour; (2) Relative reduction: $1 - (\text{'Enriched' stereotypic behaviour} / \text{'Standard' stereotypic behaviour})$; and (3) Residuals index: a mouse's individual residual from a simple linear regression of 'Enriched' stereotypic behaviour as dependent variable, versus 'Standard' stereotypic behaviour as independent variable (both arcsine square-root transformed).

We also assessed whether rates of adjustment to the new cages varied with age, first by running repeated measure analyses on the

stereotypic behaviour recorded over the 3.5 weeks of Phase 1b (to smooth the data, values from consecutive observation days were pooled into three 'periods' of 4–5 days each); and second by regressing observation day against stereotypic behaviour for each mouse, noting the slope, and then using a *t* test to compare the slopes of Young and Middle-aged mice.

Analyses on relationships between enrichment effectiveness and maximum weight pushed (Phase 2) were run using each of our three measures of change: this was useful because results could be visualized and plotted graphically. However, as a corroboration, because analysing residuals has been criticized (Darlington & Smulders 2001; García-Berthou 2001), we also ran a general linear model (GLM) with stereotypic behaviour in the enriched environment as the dependent variable, pre-enrichment stereotypic behaviour as the first term in the model, and maximum weight pushed and age as sequential factors in the model. This specific model order thus tests our hypothesis by investigating whether breakpoint relates to stereotypic behaviour in enriched conditions corrected for pre-enrichment levels (the prediction being that mice with relatively high levels of stereotypic behaviour, corrected for baseline levels, should show lower motivations for the enrichment). For all Phase 2 analyses, values for cagemates were first averaged, since each pair generated a single maximum price paid. However, as the maximum weight pushed could well reflect the breakpoint of just one mouse in each cage (if one cagemate gave up pushing at a lower weight, we would not be able to detect this), to check the robustness of any findings we re-ran these analyses using, first, the cagemate whose stereotypic behaviour had been most reduced by enrichment in Phase 1 (of the two cagemates this mouse should be the most motivated to push if reward value relates to the effectiveness of enrichment), and next, the cagemate whose stereotypic behaviour had been least reduced by enrichment.

All statistical analyses comprised GLMs in Minitab versions 13.0 and 15.0 (Minitab, State College, PA, U.S.A.). Where models were nonorthogonal (e.g. owing to covariates) they were re-run using the sequential instead of adjusted sums of squares (e.g. Grafen & Hails 2002). Where model order was not considered relevant to the hypothesis under test, such models were run and re-run with the main effects in all possible orders (robust results from such analyses are presented as '*F*>' or '*F*<' rather than '*F*='). Where only certain sequential analyses were deemed logically relevant to the hypothesis under test (as above), this is made clear. Proportional data were arcsine square-root transformed; no other transformations were needed to meet the assumptions of parametric tests. Models always included all possible interactions. For analyses of Phase 1 data, 'cage' was always treated as a random factor (thus the unit of replication), nested within age. Ideally 'family' would have also been included as a random factor. However, because only 20 of our 48 animals were related (10 Middle-aged females, each of which had a daughter in the Young group; see 'Ethical note'), it was impossible to control for 'family' while also using all data and including 'cage' in models. Therefore in follow-up analyses, these 20 related animals were considered separately, allowing the random effect of family to be factored out statistically.

This study thus involved multiple analyses. No correction was made for type I error rate because this study was novel and should be replicated (see Discussion). However, we did use two-tailed *P* values despite directional hypotheses, which should reduce risks of false positives. Note that for all figures shown, we averaged the values from cagemates (unless otherwise specified) to portray the correct number of units of replication. For bar graphs, we calculated the mean and SE of the transformed data points, and then back-transformed both the mean and the tips of the SE bars (error bars therefore are not always symmetrical).

RESULTS

Phase 1: Impact of Environmental Enrichment

In all quality checks on Phase 1's data, (see *Statistical Analyses*), odd and even day data showed strong, significant positive correlations ($P < 0.0001$ in all analyses), confirming the reliability of our sampling methods.

The transfer to enriched cages dramatically increased inactivity ($F_{1,22} = 43.12$, $P < 0.0001$) and tended to increase normal activity ($F_{1,22} = 4.19$, $P = 0.053$; Fig. 2), with no main or interactive effects of age. Resting, sleeping and being 'stationary but out-of-sight' (e.g. immobile within a sock or nesting material) increased approximately three-fold with enrichment ($F_{1,22} = 95.69$, $P < 0.0001$), but sitting still with the eyes open declined on average by about 50% ($F_{1,22} = 24.51$, $P < 0.0001$). The enriched cages also successfully decreased Total stereotypic behaviour ($F_{1,22} = 109.22$, $P < 0.0001$; Fig. 2). This effect seemed to occur rapidly after transfer to the new cage; thus 'Period' had no significant effect on stereotypic behaviour ($F_{2,44} = 0.62$, $P = 0.540$), and when stereotypic behaviour was regressed against 'day' to generate a slope for each replicate (cagemates averaged), the average slope was not significantly different from 0 ($t_{23} = 0.79$, $P = 0.438$). The reduction in stereotypic behaviour in Phase 1b also happened at similar rates in both age groups: the interactive effect of Age*Period was nonsignificant ($F_{2,44} = 0.96$, $P = 0.389$); and when stereotypic behaviour was regressed against 'day' to generate a slope for each replicate, slopes did not differ between Young and Middle-aged mice ($t_{23} = 0.38$, $P = 0.707$).

Age and housing type had no significant interactive effect on stereotypic behaviour, in analyses using all experimental animals ($F_{1,22} = 2.59$, $P = 0.11$). However, age and housing type did interact significantly for the mother–daughter subset (analyses here blocking by family: $F_{1,9} = 10.87$, $P = 0.009$). Splitting this data set by age revealed that this was because age had a nonsignificant tendency to affect Total stereotypic behaviour in Standard conditions, younger animals being more stereotypic (although see *Analyses Repeated Using Unambiguous Stereotypic Behaviour*, below): $F_{1,9} = 3.83$, $P = 0.082$. In contrast, in Enriched conditions, Young mice tended to be less stereotypic than their mothers ($F_{1,9} = 3.68$, $P = 0.087$). In all models run, stereotypic behaviour was also significantly affected by 'Cage ID'. This was because in both Standard ($F_{1,20} > 23$, $P < 0.0001$) and Enriched conditions ($F_{1,20} > 8$, $P < 0.01$), it significantly covaried between cagemates.

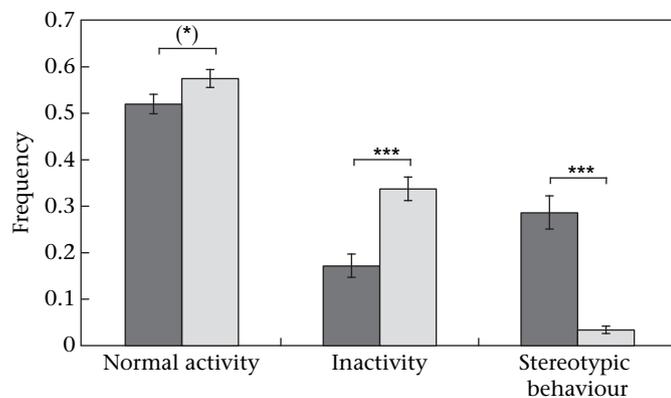


Figure 2. Frequency (proportion of scans) of normal activity, inactivity and stereotypic behaviour in standard (dark bar) and enriched (light bar) cages. (*) $P < 0.1$; *** $P < 0.001$. Means are shown with SE.

The amount of stereotypic behaviour performed in Enriched conditions was predicted by our subjects' levels of stereotypic behaviour in Standard conditions, interacting with age ($F_{1,22} = 6.24$, $P = 0.020$; Fig. 3). The model was nonorthogonal (see *Statistical Analyses*), and in all models with pre-enrichment stereotypic behaviour as the first term, age showed a nearly significant tendency to affect behaviour in Enriched conditions, older animals being more stereotypic ($F_{1,22} > 2.93$, $P < 0.101$; Fig. 3; also see *Analyses Repeated Using Unambiguous Stereotypic Behaviour*, below). Separating Young from Middle-aged mice showed that the interactive effect of age was because only in the Young mice did pre-enrichment stereotypic behaviour predict stereotypic behaviour with enrichment ($F_{1,13} > 7.62$, $P < 0.015$); Middle-aged mice, in contrast, showed no such effect ($F_{1,9} < 0.65$, $P > 0.440$). Repeating this analysis just using the mother–daughter subset of mice (thus controlling for familial effects) further suggested an age effect: when their pre-enrichment levels of stereotypic behaviour were controlled for statistically, Middle-aged mice tended to be more stereotypic than Young mice in the enriched cages ($F_{1,7} > 4.18$, $P < 0.08$).

The effectiveness of enrichment on stereotypic behaviour was further investigated using the three metrics calculated for each individual (as used for re-pairing in Phase 2; see *Statistical Analyses*). These three ways of assessing enrichment effectiveness all covaried significantly in the expected directions ($F_{1,22} > 4$, $P < 0.05$ in all cases). Absolute reduction covaried positively with pre-enrichment levels ($F_{1,22} > 76.2$, $P < 0.0001$), while the other two measures did not (Residuals index: $F_{1,22} < 0.09$, $P > 0.77$; Relative reduction: $F_{1,22} < 2.18$, $P > 0.15$). The relationship between pre-enrichment stereotypic behaviour and Absolute reduction was, however, not a simple main effect: there was an interaction with age ($F_{1,22} = 7.13$, $P = 0.014$). This model, which partialled out effects of pre-enrichment stereotypic behaviour from those of age, also revealed a nearly significant tendency for age to affect Absolute reduction, younger mice showing higher values ($F_{1,9} > 3.61$, $P < 0.072$). Splitting mice by age to investigate the interaction term showed that pre-enrichment stereotypic behaviour strongly predicted Absolute reduction in both groups (Young: $F_{1,13} > 52$, $P < 0.0001$; Middle-aged: $F_{1,9} > 48$, $P < 0.0001$); by inspection, the interaction reflected subtle differences in the slope, which was steeper and lower at the Y-intercept for Middle-aged mice. Absolute reduction in stereotypic behaviour was not affected by age in the population as a whole, when pre-enrichment levels of stereotypic behaviour were not controlled for ($F_{1,24} = 0.75$, $P = 0.40$). However, in our subset of related mice, daughters showed significantly bigger Absolute reductions than their mothers ($t = 2.53$, $N = 10$,

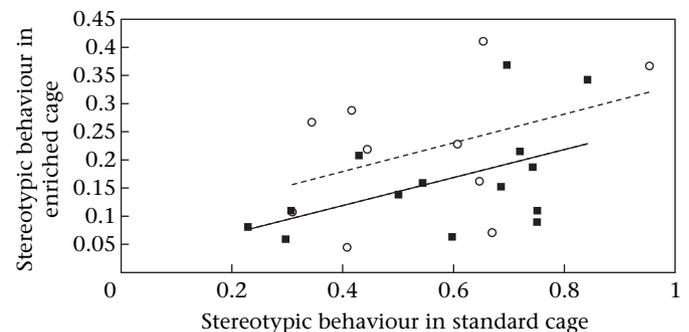


Figure 3. Relationships between pre-enrichment levels of stereotypic behaviour and stereotypic behaviour displayed in Enriched conditions, shown as proportion of scans, arcsine square-root transformed. Dark squares and solid line: Young adult mice (cagemates averaged); open circles and dashed line: Middle-aged mice (cagemates averaged).

$P = 0.032$; Fig. 4a). Young mice tended to have greater Relative reductions than Middle-aged mice ($F_{1,24} = 3.47$, $P = 0.076$), with daughters again showing significantly bigger Relative reductions than their mothers ($t = 2.35$, $N = 10$, $P = 0.043$; Fig. 4b). Young mice also showed a nearly significant tendency for greater reductions in stereotypic behaviour assessed by the Residuals index ($F_{1,24} = 2.98$, $P = 0.098$); and again this difference was significant when daughters were compared with their mothers ($t = 2.49$, $N = 10$, $P = 0.034$; Fig. 4c).

The 'predictability' of stereotypic behaviour was high, averaging at 94%; thus only 6% of Total stereotypic behaviour fell into the 'borderline' category across the whole population. The effectiveness of enrichment on stereotypic behaviour was not clearly related to the behaviour's predictability. For the Absolute and Relative reduction measures, there were nearly significant trends to the predicted negative relationships, but only in those models in which predictability was the last term ($F_{1,22} < 3.39$, $P < 0.08$), the relationship being nonsignificant in all other model orders suggesting no robust effect. The relationship was also nonsignificant for the Residuals measure ($F_{1,22} < 0.10$, $P > 0.80$). Furthermore, no significant effects were found for predictability or its interactions ($F_{1,20} < 1.00$, $P > 0.32$) in GLMs analysing whether stereotypic behaviour in enriched conditions was affected by predictability, controlling for pre-enrichment stereotypic behaviour and age. Finally, in a similar analysis to test the hypothesis that predictable stereotypic behaviours are more resistant to enrichment, the relative effectiveness of enrichment on Unambiguous and Borderline stereotypic behaviour was investigated by assessing whether the relationship between pre- and post-enrichment levels varied between these two subdivisions of Total stereotypic behaviour. They did not ($F_{1,45} < 33$, $P > 0.57$).

Age also did not affect the predictability of stereotypic behaviour in standard cages ($F_{1,24} = 0.33$, $P = 0.57$; mother–daughter comparisons: $F_{1,9} = 1.80$, $P = 0.20$). Examining relationships between the predictability and time budget of stereotypic behaviour, we found that mice with high levels of Unambiguous stereotypic behaviour performed more Borderline stereotypic behaviour ($F_{1,22} > 5.57$, $P < 0.028$), and perhaps because of this association, in regard to the relationship between Total stereotypic behaviour and its predictability there was a nearly significant tendency for a negative relationship in some orders of the GLM ($F_{1,22} > 3.38$, $P < 0.08$). However, this was nonsignificant in other model orders, and overall there was no sign of the positive relationship between predictability and time budget predicted by the 'establishment' hypothesis.

Phase 2: Motivation to get Access to Enriched Cages

Age did not predict breakpoint: the maximum weight pushed to get access to enrichment ($F_{1,22} = 0.01$, $P = 0.92$; mother–daughter comparisons: $F_{1,9} = 0.00$, $P = 0.97$).

Sequential GLMs first investigated how stereotypic behaviour in enriched conditions was predicted by age and breakpoint, after accounting for pre-enrichment stereotypic behaviour which was thus always the first term in each permutation of the model. Stereotypic behaviour in enriched conditions (cagemates averaged) was predicted by an interaction between maximum weight pushed and age ($F_{1,16} = 4.81$, $P = 0.043$). Splitting this data set by age showed that the interaction reflected a nearly significant tendency for the breakpoint to covary negatively with stereotypic behaviour in enriched conditions (corrected for baseline levels) among Middle-aged mice ($F_{1,6} = 5.65$, $P = 0.055$; also see [Analyses Repeated Using Unambiguous Stereotypic Behaviour](#), below), but not among Young mice ($F_{1,10} = 0.08$, $P = 0.789$).

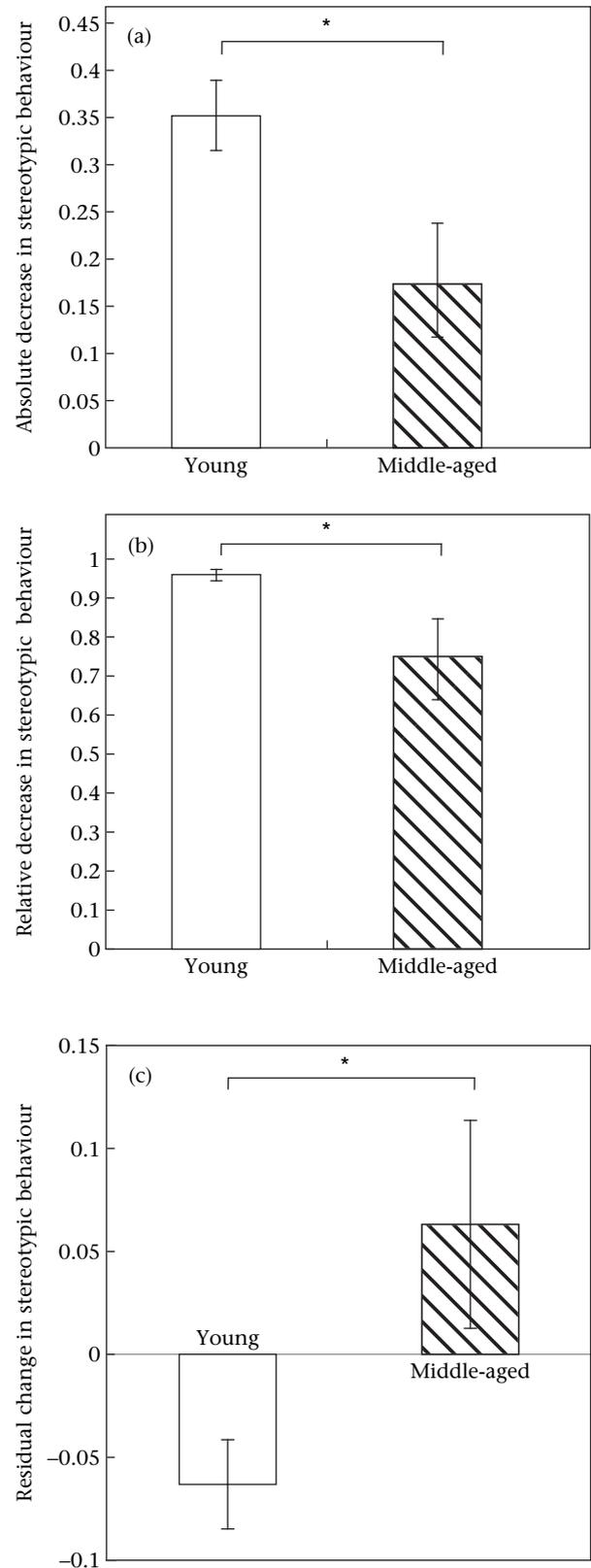


Figure 4. Relationships between age and three measures of enrichment effectiveness in Phase 1, in the subset of subjects where family effects could be controlled for. (a) Absolute, (b) relative and (c) residual change in stereotypic behaviour between mothers (Middle-aged) and their daughters (Young). Smaller, more negative residuals indicate less stereotypic behaviour in Enriched conditions than expected from pre-enrichment values.

These GLMs were then repeated focusing on the animal within each cage whose behaviour had benefited most from enrichment (see *Statistical Analyses*). Again, stereotypic behaviour in enriched conditions (corrected for baseline levels) was predicted by an interaction between the maximum weight pushed and age, for mice with the most extreme Absolute reductions ($F_{1,16} = 5.92$, $P = 0.027$), and mice with the largest Relative reductions in stereotypic behaviour ($F_{1,16} = 10.54$, $P = 0.005$). The same effect was also a nearly significant trend for individuals with the most extreme Residual index ($F_{1,16} = 3.87$, $P = 0.067$). Splitting this data set by age for the mice with the highest Absolute and Relative reductions showed that, just as above, breakpoint again covaried negatively with stereotypic behaviour in enriched conditions (corrected for baseline levels), or showed a nearly significant trend in that direction, in Middle-aged animals (Absolute reduction: $F_{1,6} = 5.28$, $P = 0.061$; Relative reduction: $F_{1,6} = 12.52$, $P = 0.012$), but not in Young mice (Absolute reduction: $F_{1,10} = 1.12$, $P = 0.287$; Relative reduction: $F_{1,10} = 0.67$, $P = 0.434$). The above GLMs were again repeated, this time focusing on the animal within each cage whose stereotypic behaviour was least reduced by enrichment, but in this case, there were no significant interactions between maximum weight pushed and age (Absolute reduction: $F_{1,16} = 2.61$, $P = 0.126$; Relative reduction: $F_{1,16} = 1.21$, $P = 0.288$; Residual index: $F_{1,16} = 2.85$, $P = 0.111$).

These findings suggested that the effectiveness of enrichment was positively related to the reward value of enrichment, but only in the Middle-aged mice. To confirm this and produce findings easier to show as figures, we then investigated the relationship between enrichment effectiveness and enrichment value using the three indices of enrichment effectiveness calculated for each subject. The relationship between these three measures and maximum weight pushed to get access to enrichment again varied with age, with the same pattern as above. Thus average Absolute reduction (cagemates' values averaged) tended to interact with age ($F_{1,20} = 3.61$, $P = 0.072$). This same interaction was significant for average Relative reduction ($F_{1,20} = 4.81$, $P = 0.040$), although not for the Residuals index ($F_{1,20} = 1.62$, $P = 0.22$). Focusing on the individual from each pair whose stereotypic behaviour had been most reduced by enrichment yielded similar effects. For these animals' Absolute reductions, there was no interaction with age ($F_{1,20} = 2.73$, $P = 0.11$; although see *Analyses Repeated Using Unambiguous Stereotypic Behaviour*, below). However, this interaction was significant for Relative reduction ($F_{1,20} = 7.32$, $P = 0.014$); and a weak trend could now be seen for the Residuals index ($F_{1,20} = 2.94$, $P = 0.10$). Focusing on those cagemates whose stereotypic behaviour had instead been least reduced by enrichment in Phase 1b, for Absolute reduction the trend for interaction with age returned ($F_{1,20} = 3.73$, $P = 0.068$), although it was not significant for Relative reduction ($F_{1,20} = 2.03$, $P = 0.16$) nor for the Residuals measure ($F_{1,20} = 0.71$, $P = 0.41$).

To understand the recurring interaction between enrichment effectiveness and age on maximum weight pushed, all data were therefore split according to age class. No measure of enrichment effectiveness predicted maximum weight pushed by Young mice (cagemates averaged: Absolute reduction: $F_{1,12} = 0.28$, $P = 0.60$; Relative reduction: $F_{1,12} = 1.08$, $P = 0.32$; Residuals index: $F_{1,12} = 0.204$, $P = 0.84$; extreme 'most affected' cagemate: Absolute reduction: $F_{1,12} = 0.18$, $P = 0.68$; Relative reduction: $F_{1,12} = 2.02$, $P = 0.18$; Residuals index: $F_{1,12} = 0.21$, $P = 0.65$; 'least affected' cagemate: Absolute reduction: $F_{1,12} = 0.39$, $P = 0.55$; Relative reduction: $F_{1,12} = 0.18$, $P = 0.67$; Residuals index: $F_{1,12} = 0.00$, $P = 0.97$). However, in the Middle-aged mice, subjects whose stereotypic behaviour had been little affected by enrichment in Phase 1 consistently worked less hard for the enrichment in Phase 2, displaying lower breakpoints. With the cagemates' mean values, these

relationships were significant for Absolute and Relative reductions, but just missed significance for the Residuals index (Absolute reduction: $F_{1,8} = 9.01$, $P = 0.017$; Relative reduction: $F_{1,8} = 8.86$, $P = 0.018$; Residuals index: $F_{1,8} = 4.98$, $P = 0.056$; Fig. 5). Focusing on the individual cagemate most affected by enrichment in Phase 1 gave very similar findings: maximum weight pushed covaried positively with their Absolute reduction in stereotypic behaviour ($F_{1,8} = 7.55$, $P = 0.025$) and their Relative reduction ($F_{1,8} = 10.06$, $P = 0.013$), while now significantly negatively covarying with their Residuals index ($F_{1,8} = 8.52$, $P = 0.019$). (The individual mice previously least affected by enrichment showed a similar pattern too, although their Residuals relationship was now not significant: Absolute reduction: $F_{1,8} = 7.35$, $P = 0.027$; Relative reduction: $F_{1,8} = 7.63$, $P = 0.025$; Residuals index: $F_{1,8} = 8.275$, $P = 0.14$.)

Analyses Repeated using Unambiguous Stereotypic Behaviour

As a last check for data robustness, we re-ran all analyses for both Phase 1 and Phase 2 using only Unambiguous stereotypic behaviour instead of the broader category Total stereotypic behaviour (Table 1). None of the previous results become nonsignificant using the stricter criteria for classifying behaviour as stereotypic: typically there were just minor changes in the exact

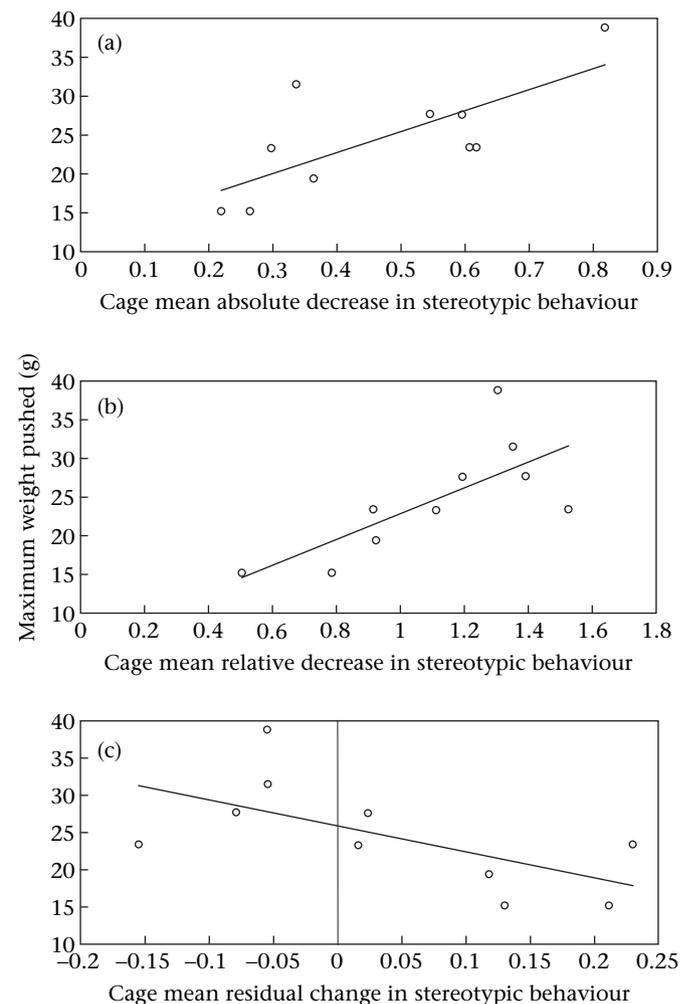


Figure 5. Relationships between the effectiveness of enrichment at reducing stereotypic behaviour, and its value as assessed by maximum weight pushed for access, in Middle-aged mice. X axis values in (a) and (b) are arcsine square-root transformed. (a) Absolute, (b) relative and (c) residual change in stereotypic behaviour. Values for cagemates are averaged.

values of F ratios and associated P values. However, four results did emerge more clearly using the stricter measure.

First, in analyses focusing on just mothers and their daughters, age now had no effect on Unambiguous stereotypic behaviour in Standard conditions, while in Enriched conditions, Young mice still tended to be less stereotypic than their mothers. Thus age and housing type still significantly interacted ($F_{1,9} = 9.95, P = 0.013$), but the interaction now more clearly reflected Young animals benefiting more from enrichment (rather than also having higher levels while still in Standard conditions, as was the case for Total stereotypic behaviour). Second, when we investigated the effect of age on stereotypic behaviour in Enriched conditions, controlling for stereotypic behaviour in Standard conditions, a nonsignificant trend now appeared when Unambiguous stereotypic behaviour was used instead of Total, with older mice being more stereotypic ($F_{1,22} > 3.02, P < 0.096$). The third change was seen in those analyses of Phase 2 data investigating how stereotypic behaviour in Enriched conditions was predicted by pre-enrichment behaviour and breakpoint. In Middle-aged mice, the relationship between stereotypic behaviour in Enriched conditions (controlling for baseline behaviour) and maximum weight pushed became significantly positive ($F_{1,6} = 6.41, P = 0.045$) instead of just a trend. Finally, focusing on those cagemates whose stereotypic behaviour had been most reduced by enrichment in Phase 1b, maximum weight pushed now tended to be predicted by Absolute reduction*age ($F_{1,20} = 2.99, P < 0.09$), a relationship that had just missed being a trend when Total stereotypic behaviour was used instead. These subtle effects of altering the criteria used to define stereotypic behaviour may reflect reduced measurement error when stricter criteria are applied.

DISCUSSION

Our main aim was to investigate whether enrichment-resistant stereotypic behaviours reflect a reduction in the reward value of the environmental enrichments used to tackle the behaviour. Our findings supported this hypothesis for Middle-aged mice, aged 10–11 months old, potentially raising concerns about the welfare of these animals and others like them. Our subsidiary aim was to test hypotheses about the repetition-dependent ‘establishment’ of stereotypic behaviour via routine-like processes; these hypotheses, in contrast, were not supported.

In Phase 1 of the study, being transferred to large enriched cages had substantial effects on the behaviour of our C57BL/6 mice: normal activity tended to increase; as reported in other rodent enrichment studies (e.g. Cooper et al. 1996; Würbel et al. 1998), inactivity, especially sleeping and resting, dramatically increased; and stereotypic behaviour was also successfully reduced in every subject. The impact of enrichment on stereotypic behaviour varied greatly in magnitude between individual mice, with one influencing factor being how stereotypic they were pre-enrichment. As we predicted, age was also important (although because of great individual variation, effects emerged only as a trend in analyses controlling for pre-enrichment levels of stereotypic behaviour, and were significant only if family effects were also controlled for). Thus despite being no more stereotypic to begin with, Middle-aged mice tended to stereotype more in enriched conditions than Young adults, and their behaviour was also less affected by enrichment as measured in terms of their absolute and proportionate changes in stereotypic behaviour, and their distances from the regression line of pre- versus postenrichment stereotypic behaviour. Our hypothesis that older mice would be less affected by enrichment was therefore supported, although effects were not dramatic, perhaps because our Young and Middle-aged animals diverged in age by only 4 months, not the 10–12 months of previous rodent studies reviewed in the Introduction.

Phase 1 gave us several further incidental findings, including some tests of the long-proposed hypotheses about establishment. One such incidental finding in Phase 1 was that in both types of environment, cagemates showed remarkably similar levels of stereotypic behaviour, either because they affected each other or because they were influenced alike by common local factors, emphasizing the importance of not treating mice housed together as statistically independent. A second was that cage-lid-directed stereotypic behaviour (bar mouthing and patterned climbing) after transfer to the enriched cage was directed at the new lid, not at the lid of the standard cage placed within the enriched cage: an observation supporting Nevison et al. (1999)'s hypothesis that this behaviour derives from escape attempts, but at odds with the idea that it represents fixed routines or habits. Further evidence that repetition did not cause stereotypic behaviour to become habit- or routine-like came from our findings that neither age nor time budget devoted to stereotyping predicted how predictable, that is, spontaneously unvarying, this behaviour was; that Unambiguous and Borderline stereotypic behaviours were similarly reduced by enrichment; and that the predictability of Total stereotypic behaviour did not determine how enrichment resistant an animal's behaviour was. Thus the increased behavioural predictability reported with age in some studies is not universal, and also seems to be independent of how easily stereotypic behaviour is alleviated by enrichment.

In Phase 2, the large enriched cages successfully proved rewarding, with all subjects pushing weights to get access to them, even though food, water and nesting material were all amply supplied in the standard cages. However, the maximum weights pushed before the ‘breakpoint’ (cessation of responding) varied greatly between animals. Our hypothesis was that this variation would reflect the degree to which enrichment reduced stereotypic behaviour, with enrichments being least valued by those mice with enrichment-resistant stereotypic behaviour. This hypothesis was strongly supported for Middle-aged mice. In this age group, mice with the most enrichment-resistant stereotypic behaviour in Phase 1 clearly had lower ‘breakpoints’ in Phase 2. Furthermore, if we assumed that within each cage, the individual mouse whose stereotypic behaviour had been most reduced by enrichment was responsible for pushing the maximum weight pushed, the results were clearer than if we assumed instead that the individual whose stereotypic behaviour had been least reduced by enrichment was responsible: a pattern further consistent with the effectiveness and reward value of enrichment being related. Overall, our findings thus show for the first time that when environmental enrichments have little impact on stereotypic behaviour, they may also not be very motivating, thence having little effect on welfare. This raises concern about the wellbeing of aged or long-captive animals, especially given that meta-analyses of zoo animal enrichment, for example, find not a single study in which stereotypic behaviour was successfully abolished (Swaigood & Shepherdson 2005, 2006; Shyne 2006). Our result is also practically important because in the zoo community, enrichment-resistant stereotypic behaviours are often attributed to past poor conditions rather than any current problems (Mason et al. 2007): an assumption that our results demonstrate is overoptimistic.

In the Introduction, we outlined some means by which enrichment could lose its reinforcing efficacy with age and/or with chronic stress, and future work should therefore now investigate the underlying mechanisms of the effect observed in Middle-aged mice. This should include quantifying their interactions with enrichment (see e.g. Cooper & Nicol 1991), and also screening mice in tests for anxiety and anhedonia (see e.g. Crawley 2007). Mice could be further screened for other depression-related changes such as learned helplessness (as assessed in Porsolt ‘forced swim’

tests, e.g. [Crawley 2007](#)). The co-occurrence of enrichment-resistant stereotypic behaviour and reduced motivation for enrichments in these older mice raises new questions about how exactly these two effects interrelate. In particular, we had assumed that finding enrichments unrewarding causes them to have little impact on stereotypic behaviour, but perhaps the causal relationship is in the other direction. They also raise new questions about how they relate to findings from neuroscience on the reduced effectiveness of enrichment on brain health and neurogenesis with age (e.g. [Saucier et al. 2007](#); [Mirochnic et al. 2009](#)). In addition, further work should record the identity, within a cage, of the individual willing to push the heaviest weights, to test directly our hypothesis that the most motivated cagemate is the one with the largest decrease in stereotypic behaviour, as well as also assessing the maximum weights all mice are willing or able to push to deal with a potential confound that could explain, as we discuss below, our most puzzling finding.

An unresolved puzzle from our results is why the Middle-aged adults did not seem, as a group, to show less motivation for enrichments than the Young adults. This would have been consistent with enrichment's effect on their stereotypic behaviour. We suspect the explanation is that Middle-aged mice were inherently more physically and/or psychologically persistent when faced with obstacles such as progressively weighted doors. Our Middle-aged mice were likely to have been 10% heavier than the Young adults (see data in [Sohal et al. 2009](#)), and, furthermore, older mice are more perseverative (e.g. [Ridderinkhof et al. 2002](#); A. N. Gross, J. P. Garner & H. Würbel, unpublished data), both of which could cause this group to achieve higher breakpoints more readily than Young adults. An oversight on our part was not to have weighed our individual mice. On the other hand, we suspect body weight alone was not a confound for two reasons. First, in a similar study by one of us ([Latham 2004](#); [Latham & Mason 2010](#)), when CD-1 mice were required to push weighted doors to reach their daily food allowance, the maximum weight pushed for food did not covary with body weight; and second, when we used the mean and variance in body weight for 7- and 11-month old ad libitum-fed C57BL/6J mice (data from [Sohal et al. 2009](#)) to create dummy body weight data for our two age groups, and added these dummy data as a control to models investigating the effect of age on maximum weight pushed, these analyses still did not reveal an age effect on the breakpoint, as if the small weight difference between the two age groups was not sufficient to be influential. The standard technique for controlling for nonmotivational factors in set-ups like Phase 2 is to assess how hard animals will work for food, using this as an index of peak 'push' and thus as a covariate in subsequent analyses. We had chosen not to do this on ethical grounds (see [Ethical Note](#)), hoping that such data would not be needed. However, it is now evident that follow-up studies cannot avoid running such checks. Our predictions in future work would thus be that when one controls for this potential confound, Middle-aged adults as a group do prove less motivated to reach enrichments than Young adults.

Overall, our findings highlight previously overlooked mechanisms which could cause enrichments to lose their effectiveness on old and/or long-caged animals: effects with potentially serious implications for these animals' welfare. Assumptions that persistent stereotypers simply have 'scar-like' habits that are dissociated from current welfare (e.g. [Mason 1991](#)) should thus never be assumed true without corroborating evidence.

Acknowledgments

Thanks to Claudia Vinke for co-supervision and sending S.-L.T. to Canada; Naomi Latham and Allison Bechard for invaluable advice on

running the experiment; Linda Grocock and Michelle Cieplak for excellent care of the animals; Frank Ödberg (University of Ghent) and Gabor Szalai (Peromyscus Genetic Stock Center, University of Southern Carolina) for advice on aging and longevity in bank voles and deer mice; Raj Sohal and Michael Forster (University of Southern California) for data on mouse body weights from their published study; NSERC for funding G.J.M. and J.D.; and two anonymous referees and the editor for very helpful comments on the manuscript.

References

- Ames, A.** 1994. *The Welfare and Management of Bears in Zoological Gardens*. Potters Bar: Universities Federation for Animal Welfare.
- Augustsson, H., Van de Weerd, H. A., Kruitwagen, C. L. J. J. & Baumans, V.** 2003. Effect of enrichment on variation and results in the light/dark test. *Laboratory Animals*, **37**, 328–340.
- Baker, K. C.** 2000. Advanced age influences chimpanzee behavior in small social groups. *Zoo Biology*, **19**, 111–119.
- Benus, R. F., den Daas, S., Koolhaas, J. M. & van Oortmerssen, G. A.** 1990. Routine-formation and flexibility in social and nonsocial behaviour of aggressive and non-aggressive male mice. *Behaviour*, **112**, 176–193.
- Bondi, C. O., Rodriguez, G., Gould, G. G., Frazer, A. & Morilak, D. A.** 2008. Chronic unpredictable stress induces a cognitive deficit and anxiety-like behavior in rats that is prevented by chronic antidepressant drug treatment. *Neuropsychopharmacology*, **33**, 320–331.
- Burghardt, G.** 2005. *The Genesis of Animal Play: Testing the Limits*. Cambridge, Massachusetts: MIT Press.
- Clubb, R.** 2001. Factors affecting the development of abnormal, stereotyped behaviour in zoo-housed carnivores. Ph.D. thesis, University of Oxford.
- Cooper, J. J.** 1992. Behavioural analysis of stereotypic behaviour in rodents. Ph.D. thesis, University of Bristol.
- 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. & Ödberg, F.** 1991. The emancipation of stereotypes with age. In: *Proceedings of the International Congress of the Society for Veterinary Ethology* (Ed. by M. C. Appleby, R. I. Horrell, J. C. Petherick & S. M. Rutter), p. 142. Edinburgh: Universities Federation for Animal Welfare.
- 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.
- Cronin, G. M.** 1985. The development and significance of abnormal stereotyped behaviours in tethered sows. Ph.D. thesis, Wageningen University.
- Crawley, J. N.** 2007. *What's Wrong With My Mouse? Behavioral Phenotyping of Transgenic and Knockout Mice*, 2nd edn. Wilmington: Wiley-Liss.
- Darlington, R. B. & Smulders, T. V.** 2001. Problems with residual analysis. *Animal Behaviour*, **62**, 599–602.
- Di Giovanni, M. & Valente, M.** 2001. Trial and error: attempting to change the lives of socially and environmentally deprived primates. In: *Proceedings of the 4th International Conference on Environmental Enrichment* (Ed. by V. J. Hare, K. E. Worley & K. Myers), pp. 358–363. San Diego: The Shape of Enrichment.
- Dodman, N. H., Karlsson, E. K., Moon-Fanelli, A., Galdzicka, M., Perloski, M., Shuster, L., Lindblad-Toh, K. & Ginns, E. I.** 2010. A canine chromosome 7 locus confers compulsive disorder susceptibility. *Molecular Psychiatry*, **15**, 8–10.
- Eaton, G. G., Kelley, S. T., Axthelm, M. K., Iliff-Sizemore, S. A. & Shiigi, S. M.** 1994. Psychological well-being in paired adult female rhesus (*Macaca mulatta*). *American Journal of Primatology*, **33**, 89–99.
- Fentress, J. C.** 1976. Dynamic boundaries of patterned behaviour: interaction and self-organisation. In: *Perspectives in Ethology. Vol. 1* (Ed. by P. P. G. Bateson & R. A. Hinde), pp. 155–224. New York: Plenum.
- Fentress, J. C.** 1977. The tonic hypothesis and the patterning of behavior. *Annals of the New York Academy of Sciences*, **290**, 370–395.
- Flurkey, K., Currer, J. M. & Harrison, D. E.** 2007. The mouse in aging research. In: *The Mouse in Biomedical Research. Vol. III. Normative Biology, Husbandry, and Models* (Ed. by J. G. Fox, S. W. Barthold, M. T. Davisson, C. E. Newcomer, F. W. Quimby & A. L. Smith), pp. 637–672, 2nd edn. Burlington: Academic Press.
- García-Berthou, E.** 2001. On the misuse of residuals in ecology: testing regression residuals vs. the analysis of covariance. *Journal of Animal Ecology*, **70**, 708–711.
- Garner, J. P.** 1999. The aetiology of stereotypy in caged animals. Ph.D. thesis, University of Oxford.
- Garner, J. P.** 2006. Perseveration and stereotypy: systems-level insights from clinical psychology. In: *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* (Ed. by G. Mason & J. Rushen), pp. 121–152, 2nd edn. Wallingford: CAB.
- Garner, J. P., Meehan, C. L. & Mench, J. A.** 2003. Stereotypies in caged parrots, schizophrenia and autism: evidence for a common mechanism. *Behavioural Brain Research*, **145**, 125–134.
- Golani, L., Kafkafi, N. & Draai, D.** 1999. Phenotyping stereotypic behaviour: collective variables, range of variation and predictability. *Applied Animal Behaviour Science*, **65**, 191–220.
- Grafen, A. & Hails, R.** 2002. *Modern Statistics for the Life Sciences*. Oxford: Oxford University Press.

- Hadley, C., Hadley, B., Ephraim, S., Yang, M. & Lewis, M. H. 2006. Spontaneous stereotypy and environmental enrichment in deer mice (*Peromyscus maniculatus*): reversibility of experience. *Applied Animal Behaviour Science*, **97**, 312–322.
- Herrera-Pérez, J. J., Martínez-Mota, L. & Fernández-Guasti, A. 2008. Aging increases the susceptibility to develop anhedonia in male rats. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, **32**, 1798–1803.
- Hotz, G. & Helm-Estabrooks, N. 1995. Perseveration. Part I: a review. *Brain Injury*, **9**, 151–159.
- Hursh, S. R. 1980. Economic concepts for the analysis of behavior. *Journal of the Experimental Analysis of Behavior*, **34**, 219–238.
- Hutt, C. 1969. Exploration, arousal and autism. *Psychologische Forschung*, **33**, 1–8.
- Imhof, J. T., Coelho, Z. M. I., Schmitt, M. L., Morato, G. S. & Carobrez, A. P. 1993. Influence of gender and age on performance of rats in the elevated plus maze apparatus. *Behavioural Brain Research*, **56**, 177–180.
- Kennes, D., Ödberg, F. O., Bouquet, Y. & De Rycke, P. H. 1988. Changes in naloxone and haloperidol effects during the development of captivity-induced jumping stereotypy in bank voles. *European Journal of Pharmacology*, **153**, 19–24.
- Kurt, F. & Hartl, C. B. 1995. Asian elephants (*Elephas maximus*) in captivity: a challenge for zoo biological research. In: *Research and Captive Propagation* (Ed. by U. Ganslosser, J. K. Hodges & W. Kaumanns), pp. 310–326. FÜRth: Filander Verlag.
- Latham, N. R. 2004. Refining the role of stereotypic behaviour in the assessment of welfare: stress, general motor persistence and early environment in the development of abnormal behaviours. Ph.D. thesis, University of Oxford.
- Latham, N. & Mason, G. 2010. Frustration and perseveration in stereotypic captive animals: is a taste of enrichment worse than none at all? *Behavioural Brain Research*, **211**, 96–104.
- Lewis, M. H., Tanimura, Y., Lee, L. W. & Bodfish, J. W. 2007. Animal models of restricted repetitive behavior in autism. *Behavioural Brain Research*, **176**, 66–74.
- Line, S. W., Morgan, K. N. & Markowitz, H. 1991. Simple toys do not alter the behavior of aged rhesus monkeys. *Zoo Biology*, **10**, 473–484.
- Loh, E. A. & Roberts, D. C. S. 1990. Break-points on a progressive ratio schedule reinforced by intravenous cocaine increase following depletion of forebrain serotonin. *Psychopharmacology*, **101**, 262–266.
- Martin, P. & Bateson, P. 2007. *Measuring Behaviour: An Introductory Guide*, 3rd edn. Cambridge: Cambridge University Press.
- Mason, G. J. 1991. Stereotypies and suffering. *Behavioural Processes*, **25**, 103–115.
- Mason, G. J. 1993. Age and context affect the stereotypies of caged mink. *Behaviour*, **127**, 191–229.
- Mason, G. J. 2006. Stereotypic behaviour in captive animals: fundamentals and implications for welfare and beyond. In: *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* (Ed. by G. Mason & J. Rushen), pp. 325–356, 2nd edn. Wallingford: CABI.
- Mason, G. J. & Latham, N. R. 2004. Can't stop, won't stop: is stereotypy a reliable animal welfare indicator? *Animal Welfare*, **13**, S57–S69.
- Mason, G. J., Cooper, J. & Clarebrough, C. 2001. Frustrations of fur-farmed mink. *Nature*, **410**, 35–36.
- Mason, G., Clubb, R., Latham, N. & Vickery, S. 2007. Why and how should we use environmental enrichment to tackle stereotypic behaviour? *Applied Animal Behaviour Science*, **102**, 163–188.
- Meehan, C. L., Garner, J. P. & Mench, J. A. 2003. Isosexual pair housing improves the welfare of young Amazon parrots. *Applied Animal Behaviour Science*, **81**, 73–88.
- Meyer-Holzappel, M. 1968. Abnormal behaviour in zoo animals. In: *Abnormal Behaviour in Animals* (Ed. by M. W. Fox), pp. 476–503. London: Saunders.
- Mirochnic, S., Wolf, S., Staufenbiel, M. & Kemperman, G. 2009. Age effects on the regulation of adult hippocampal neurogenesis by physical activity and environmental enrichment in the APP23 mouse model of Alzheimer disease. *Hippocampus*, **19**, 1008–1018.
- Morris, D. 1964. The response of animals to a restricted environment. *Symposia of the Zoological Society of London*, **13**, 99–118.
- Nevison, C. M., Hurst, J. L. & Barnard, C. J. 1999. Strain-specific effects of cage enrichment in male laboratory mice (*Mus musculus*). *Animal Welfare*, **8**, 361–379.
- Novak, M. A., Musante, A., Munroe, H., O'Neill, P. L., Price, C. & Suomi, S. J. 1993. Old, socially housed rhesus monkeys manipulate objects. *Zoo Biology*, **12**, 285–298.
- Ödberg, F. O. 1978. Introduction to round table on abnormal behaviours (stereotypies). In: *Proceedings of the 1st World Congress Ethology Applied to Zootechnics* (Editorial Garsi) pp. 475–480. Madrid: Industrias Gráficas España.
- Olsson, I. A. S. & Keeling, L. J. 2002. The push-door for measuring motivation in hens: laying hens are motivated to perch at night. *Animal Welfare*, **11**, 11–19.
- Papp, M., Willner, P. & Muscat, R. 1991. An animal model of anhedonia: attenuation of sucrose consumption and place preference conditioning by chronic unpredictable mild stress. *Psychopharmacology*, **104**, 255–259.
- Pietropaolo, S., Branchi, I., Chiarotti, F. & Alleva, E. 2004. Utilisation of a physically enriched environment by laboratory mice: age and gender differences. *Applied Animal Behaviour Science*, **88**, 149–162.
- Ridderinkhof, K. R., Span, M. M. & van der Molen, M. W. 2002. Perseverative behavior and adaptive control in older adults: performance monitoring, rule induction, and set shifting. *Brain and Cognition*, **49**, 382–401.
- Roper, T. J. 1975. Nest material and food as reinforcers for fixed-ratio responding in mice. *Learning and Motivation*, **6**, 327–343.
- Saucier, D. M., Yager, J. Y., Armstrong, E., Keller, A. & Shultz, S. 2007. Enriched environment and the effect of age on ischemic brain damage. *Brain Research*, **1170**, 31–36.
- Shepherdson, D., Mellen, J. & Hutchins, M. 1998. *Second Nature: Environmental Enrichment for Captive Animals*. Washington: Smithsonian Institution Press.
- Sherwin, C. M. 1997. Observations on the prevalence of nest-building in non-breeding TO strain mice and their use of two nesting materials. *Laboratory Animals*, **31**, 125–132.
- Sherwin, C. M. 2003. Social context affects the motivation of laboratory mice, *Mus musculus*, to gain access to resources. *Animal Behaviour*, **66**, 649–655.
- Shyne, A. 2006. Meta-analytic review of the effects of enrichment on stereotypic behavior in zoo mammals. *Zoo Biology*, **25**, 317–337.
- Snaith, P. 1993. Anhedonia: a neglected symptom of psychopathology. *Psychological Medicine*, **23**, 957–966.
- Sohal, R. S., Ferguson, M., Sohal, B. H. & Forster, M. J. 2009. Life span extension in mice by food restriction depends on an energy imbalance. *Journal of Nutrition*, **139**, 533–539.
- Stevenson, M. 1983. The captive environment: its effect on exploratory and related behavioural responses in wild animals. In: *Exploration in Animals and Humans* (Ed. by J. Archer & L. I. A. Birke), pp. 176–197. New York: Van Nostrand Reinhold.
- Swaigood, R. R. & Shepherdson, D. J. 2005. Scientific approaches to enrichment and stereotypies in zoo animals: what's been done and where should we go next? *Zoo Biology*, **24**, 499–518.
- Swaigood, R. R. & Shepherdson, D. 2006. Environmental enrichment as a strategy for mitigating stereotypies in zoo animals: a literature review and meta-analysis. In: *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* (Ed. by G. Mason & J. Rushen), pp. 256–285, 2nd edn. Wallingford: CABI.
- Turner, M. 1999. Repetitive behaviour in autism: a review of psychological research. *Journal of Child Psychology and Psychiatry*, **40**, 839–849.
- Van de Weerd, H. A., Van Loo, P. L. P., Van Zutphen, L. F. M., Koolhaas, J. M. & Baumans, V. 1997. Nesting material as environmental enrichment has no adverse effects on behavior and physiology of laboratory mice. *Physiology & Behavior*, **62**, 1019–1028.
- Van de Weerd, H. A., Van Loo, P. L. P., Van Zutphen, L. F. M., Koolhaas, J. M. & Baumans, V. 1998a. Strength of preference for nesting material as environmental enrichment for laboratory mice. *Applied Animal Behaviour Science*, **55**, 369–382.
- Van de Weerd, H. A., Van Loo, P. L. P., Van Zutphen, L. F. M., Koolhaas, J. M. & Baumans, V. 1998b. Preferences for nest boxes as environmental enrichment for laboratory mice. *Animal Welfare*, **7**, 11–25.
- Van de Weerd, H. A., Aarsen, E. L., Mulder, A., Kruitwagen, C. L. J. J., Hendriksen, C. F. M. & Baumans, V. 2002. Effects of environmental enrichment for mice: variation in experimental results. *Journal of Applied Animal Welfare Science*, **5**, 87–109.
- Van Loo, P. L. P., Kruitwagen, C. L. J. J., Koolhaas, J. M., Van de Weerd, H. A., Van Zutphen, L. F. M. & Baumans, V. 2002. Influence of cage enrichment on aggressive behaviour and physiological parameters in male mice. *Applied Animal Behaviour Science*, **76**, 65–81.
- Vickery, S. S. 2003. Stereotypic behaviour in caged bears: individual and husbandry factors. Ph.D. thesis, University of Oxford.
- Weidenmayer, C. 1997. Causation of the ontogenetic development of stereotypic digging in gerbils. *Animal Behaviour*, **53**, 461–470.
- Würbel, H. 2006. The motivational basis of caged rodents' stereotypies. In: *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare* (Ed. by G. Mason & J. Rushen), pp. 86–120, 2nd edn. Wallingford: CABI.
- Würbel, H., Chapman, R. & Rutland, C. 1998. Effect of feed and environmental enrichment on development of stereotypic wire-gnawing in laboratory mice. *Applied Animal Behaviour Science*, **60**, 69–81.
- Young, R. J. 2003. *Environmental Enrichment for Captive Animals*. Wheathampstead: Universities Federation for Animal Welfare.