

Stereotypies and Other Abnormal Repetitive Behaviors: Potential Impact on Validity, Reliability, and Replicability of Scientific Outcomes

Joseph P. Garner

Abstract

Normal behavior plays a key role in facilitating homeostasis, especially by allowing the animal to control and modify its environment. Captive environments may interfere with these behavioral responses, and the resulting stress may alter many physiological parameters. Abnormal behaviors indicate that an animal is unable to adjust behaviorally to the captive environment and, hence, may be expressing abnormal physiology. Therefore, captive environments may affect the following aspects of an experiment: validity, by introducing abnormal animals into experiments; reliability, by increasing interindividual variation through the introduction of such individuals; and replicability, by altering the number and type of such individuals between laboratories. Thus, far from increasing variability, enrichment may actually improve validity, reliability, and replicability by reducing the number of abnormal animals introduced into experiments. In this article, the specific example of abnormal repetitive behaviors (ARBs) is explored. ARBs in captive animals appear to involve the same mechanisms as ARBs in human psychiatry, which reflect underlying abnormalities of brain function. ARBs are also correlated with a wide range of behavioral changes that affect experimental outcomes. Thus, ARBs in laboratory animals may compromise validity, reliability, and replicability, especially in behavioral experiments; and enrichments that prevent ARB may enhance validity, reliability, and replicability. Although many links in this argument have been tested experimentally, key issues still remain in the interpretation of these data. In particular, it is currently unclear (1) whether or not the differences in brain function seen in animals performing ARB are abnormal, (2) which common behavioral paradigms are affected by ARB, and (3) whether enrichment does indeed improve the quality of behavioral data. Ongoing and future work addressing these issues is outlined.

Key Words: abnormal behavior; animal welfare; executive system; obsessive compulsive; perseveration; stereotypy

Potential for Abnormal Behavior to Affect Experimental Outcomes

Importance of Behavior in Any Experiment

It is useful to think of behavior as an organ, which is integrated with the biology of the whole animal. Like other organs, behavior ensures the survival and reproductive success of the animal. In so doing, behavior is intimately involved in homeostasis. For example, animals actively mix and match from a vast array of suboptimal foods to achieve the balanced diet they require to maximize growth (Raubenheimer and Simpson 1999). Behavior is also the means through which the animal can alter and control the environment to allow homeostasis—of which nest building is a particularly common example. For instance, laboratory mice are typically housed under conditions of cold stress. Mouse body temperature is in the 35 to 37°C range (Tankersley et al. 2002), yet the typical ambient room temperatures in vivaria are set at 20 to 22°C. For this reason, it is perhaps not surprising that laboratory mice consume less food and put on weight when they are provided with the opportunity to build a nest (Olsson and Dahlborn 2002).

A key component of stress is the failure of homeostasis (Moberg 2000). Thus, when an animal is housed in environments that render it unable to perform species-typical behaviors (e.g., nest building), which would normally allow it to control its environment and enhance homeostasis, it is likely to be under a state of stress (Olsson and Dahlborn 2002). Similarly, when animals are exposed to chronic aversive stimuli (e.g., cues leading to the perceived exposure to predators in an environment devoid of shelter), or an inability to perform behaviors normally essential to survival or reproduction in the wild (e.g., dispersal in mice; Latham and Mason 2004), they are likely to be stressed (Olsson and Dahlborn 2002). Indeed, the inability to predict or control stimuli that are themselves non-life-threatening can lead to chronic states of stress that can markedly decrease immunocompetence (e.g., social disruption in rhesus; Capitanio and Lerche 1998) or even lead to spontaneous mortality (e.g., foot shock in rats; Weiss 1970). One of the fundamental aims of enrichment, therefore, is to allow the animal to perform species-typical behaviors that return control over the environment to the animal and promote homeostasis.

Joseph P. Garner, Ph.D., is an Assistant Professor in the Department of Animals Sciences, Purdue University, West Lafayette, Indiana.

Abnormal Behaviors

Abnormal behavior often results when an animal is housed in an environment where it is exposed to chronic aversive stimuli, where it cannot perform behaviors normally essential to reproduction or survival in the wild, or where it cannot perform behaviors that would correct the homeostatic imbalance it is experiencing. Indeed, abnormal behavior serves not only as an indicator of poor welfare (i.e., negative internal experiences, impaired function, and/or denied natural behavior; Duncan and Fraser 1997), but often also as an indicator of stress and even failing homeostasis. Like abnormal behavior in human psychiatry (Davison and Neale 1998), abnormal behaviors in animals are identified on the basis of the following criteria, none of which is necessary or sufficient:

- Is the behavior seen only in captivity (e.g., stereotypies)?
- If seen in the wild as well as captivity, is the behavior performed in inappropriate circumstances (e.g., infanticide in mice), or performed excessively (e.g., screeching in parrots)?
- Does the behavior involve self-injury (e.g., self-biting in primates), affect social interactions (e.g., barbering in mice), or have deleterious consequences on growth or reproduction (e.g., stereotypies in mink)?
- Is the behavior peculiar to a subset of individuals (e.g., barbering in mice)?
- Does the behavior induce signs of distress in the animal or its companions?

These criteria parallel criteria in human psychiatry termed “violation of norms,” “unexpectedness,” “disability or dysfunction,” “statistical infrequency,” and “personal distress,” respectively (for discussion of these human criteria, see Davison and Neale 1998).

Abnormal behavior can be divided into two mechanistic categories, termed “maladaptive” and “malfunctional” behaviors (Mills 2003). **Maladaptive behaviors** reflect a normal animal in an abnormal environment: The animal is simply responding as well as it can with functionally intact behavioral mechanisms to the captive environment; the perversities of its behavior reflect those of the environment. For instance, infanticide in mice is a perfectly functional response to high levels of population density in the wild (Latham and Mason 2004), and high rates of infanticide in captivity may merely reflect the high level of population density perceived by mice in captivity. Maladaptive abnormal behaviors may nevertheless indicate a failure of behavior to regulate a stressor or properly correct a homeostatic imbalance, which over time could lead to altered physiology. **Malfunctional behaviors**, in contrast, are the product of abnormal psychology, brain development, or neurochemistry induced by features of the captive environment. For example, isolation rearing induces widespread changes in behavior, including inducing abnormal behaviors such as stereotypies as well as changes in brain development and

brain chemistry (e.g., Jones et al. 1991; Lewis et al. 1990; Würbel 2001).

General Hypothesis: Potential Effects of Abnormal Behavior, and Benefits of Enrichment, on Scientific Outcomes

To be of high scientific quality, a measure must be valid, reliable, and replicable (e.g., Martin and Bateson 1986). Although often confused, these three criteria are distinct.

Validity

A measure shows internal validity when it measures what it is intended to measure. When different measures of the same property do not correlate, they fail to show convergent validity (e.g., measures of fearfulness in quail; Miller 2003), and when they correlate with unrelated measures, they fail to show discriminant validity (e.g., Miller 2003). A measure in a model system shows external validity when it accurately models the applied system (Willner 1986) and when its generality is not constrained to the narrow experimental parameters of the model. For instance, some high-throughput behavioral measures are strongly affected by extraneous environmental variables (e.g., Chesler et al. 2002), which significantly limits the external validity of many experiments employing them.

Reliability

Reliable measures are measurements that provide the same result either when two measurement devices (or experimenters) return the same result on the same individual (inter-rater reliability), or when the same experimenter obtains the same result from the same individual with repeated observations (test-retest reliability). Good reliability does not imply good validity. For example, measures of fearfulness show superficial reliability simply because the individual’s environment remains constant, rather than because they are measuring a personality trait of the individual (Miller et al. 2005), indicating that the measures have poor validity.

Replicability

Replicable results are results that can be repeated independently in different laboratories. Again, some high-throughput behavioral measures show poor between-laboratory replicability despite good within-laboratory reliability (Crabbe et al. 1999). Although a replicable experiment does not ensure that the measures are valid, nonreplicable results may indicate poor external validity of the measures. These distinctions are described here because the general hypothesis of this article is that abnormal behavior could potentially affect the validity, reliability, and replicability of scientific outcomes (Garner and Mason 2002; Garner et al. 2003a; Würbel 2000, 2001).

Because both maladaptive and malfunctioning abnormal behavior potentially involve abnormal physiology, individuals performing abnormal behaviors can hardly be assumed to be good models of normal physiology (Würbel 2001). Thus, animals performing abnormal behavior could (and probably do) negatively affect experimental validity. One fascinating aspect of many examples of abnormal behavior is the fact that some individuals perform it, while others do not, even when they are of the same strain, sex, and age, experiencing the same housing, husbandry, and handling, and housed in the same cage. Moreover, of those animals that perform abnormal behavior, individuals differ significantly in the severity of the behavior.

The features of abnormal behavior described above correspond to the phenomena of “prevalence” and “symptom severity” in human psychiatry, respectively. This variation in abnormal behavior between individuals is likely to reflect differences in physiology and, hence, could add considerable between-individual noise to susceptible experimental outcomes, thereby reducing the reliability of experimental measures and of the experiments themselves. Furthermore, differences in husbandry conditions between laboratories, which appear innocuous to human caretakers, may have major effects on abnormal behavior and hence, on experimental outcomes. For instance, the severity but not the prevalence of barbering differs between rooms in a facility (Garner et al. 2004a,b, unpublished data). Thus, one potential component of the poor between-laboratory replicability of many high-throughput behavioral tests (Crabbe et al. 1999) may be between-laboratory variation in the severity or prevalence of abnormal behavior (Garner and Mason 2002).

Environmental enrichment, by definition, is the provision of biologically relevant resources or structuring to the cage that facilitate or allow highly motivated natural behaviors, particularly those that allow the animal to control its environment or homeostasis (Olsson and Dahlborn 2002). As a result, successful enrichments often reduce or prevent the occurrence of both malfunctioning and maladaptive abnormal behaviors. However, because malfunctioning behaviors may involve more permanent developmental changes in the brain, enrichment may have less effect on these behaviors once they have developed (which appears to be the case for many stereotypies; e.g., Cooper et al. 1996). Complex, structured environments and social contact (both of which can be forms of enrichment; Baumans 2005) are also known to alter brain development and many aspects of behavior (e.g., Jones et al. 1991; Rosenzweig et al. 1978). These alterations have led to the widespread worry that enriching laboratory animals will have one or more of the following effects: It will fundamentally alter their behavior; change the outcome of experiments (hence, decrease validity); increase between-individual variation (hence, decrease reliability); and increase between-laboratory variation (hence, decrease replicability if different facilities use different enrichments). However, this argument against enrichment rests on the implicit assumption that animals housed under

barren standardized conditions are more normal (i.e., more representative of wild conspecifics, or of humans) than enriched animals. If this assumption is false, then enrichment might even *benefit* scientific outcomes, as developed below.

In the foregoing text, the following arguments have been discussed: (1) Abnormal behaviors may be a marker of abnormal physiology; (2) this abnormal physiology may compromise the validity of animal experiments, and may add variation that causes poor reliability and replicability; and (3) the physiology of barren-housed animals cannot be considered normal simply because enrichment changes physiology, and, as a consequence, enrichments may prevent some abnormal behaviors by preventing abnormal physiology. If these arguments are accurate, then indeed enrichment may yield more “normal” animals and thereby increase the validity of animal experiments by normalizing abnormal physiology or by preventing abnormal brain development or neurochemistry. Similarly, enrichment could potentially reduce between-individual and between-laboratory variability by removing the range of variation introduced by abnormal physiology, thereby enhancing experimental reliability and replicability (Figure 1). Therefore, far from introducing another source of variability into experiments, enrichment could instead (1) help address the potential effects of abnormal behavior on experimental outcomes, (2) reduce variability, and (3) improve validity.

Abnormal behavior is not a necessary component of the argument that enriched animals may be more physiologically normal (i.e., representative of humans or wild conspecifics) and less physiologically variable, but it provides a concrete means to dissect the problem and test the hypothesis. Rather than a general review of enrichment and its effects on experimental outcomes or individual variability, a specific corner of this broader issue is explored in the ensuing discussion. This article focuses on **abnormal repetitive behavior** (ARB¹), which because of its close connection to brain function, provides a specific example where these hypotheses may in fact bear weight; where a complete testable argument can be formulated; where potential effects on experimental outcomes can be predicted; and where areas for further study are clear. For a broader review of the potential effects of enrichment on experimental outcomes, the reader is directed to other articles in this issue of *ILAR Journal* (e.g., Bayne 2005).

Abnormal Repetitive Behavior (ARBs)

ARBs are defined as behaviors that are inappropriate, repetitive and unvarying in either goal or motor pattern (Garner 2005; Turner 1997). ARBs may be subdivided into two basic categories based on the unvarying manner in which the behaviors are repeated: (1) **Stereotypies** involve the

¹Abbreviation used in this article: ARB, abnormal repetitive behavior.

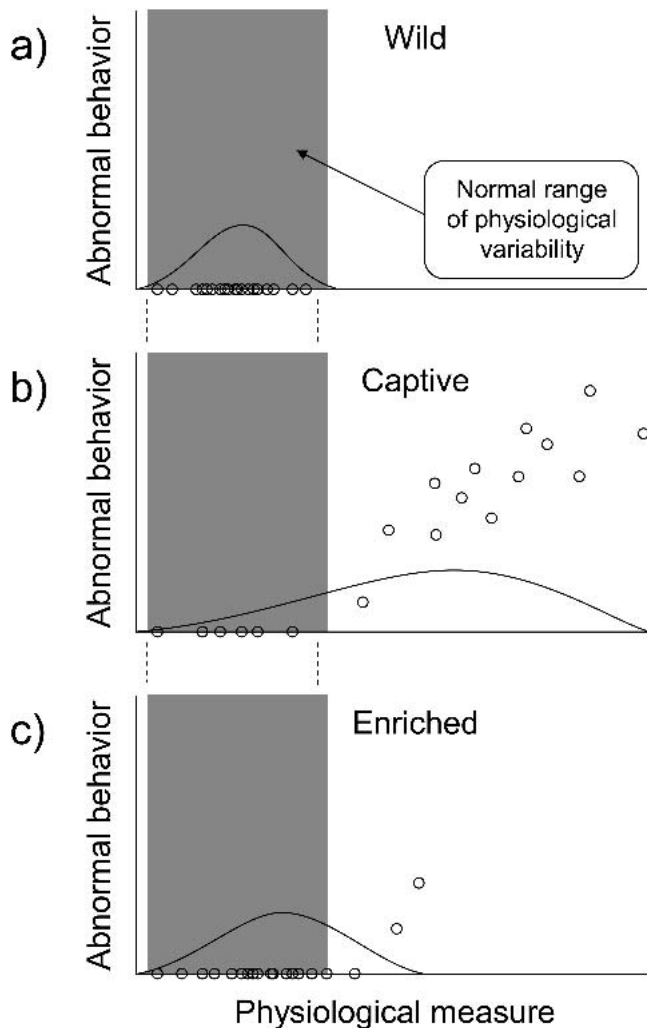


Figure 1 Hypothetical data set illustrating the general hypothesis of this article. a) Distribution of a physiological measure in a wild and hence normal population of animals. Note that none of the animals show any abnormal behavior. The gray box denotes the range of physiological variation that can be considered normal, and should be the target range for a valid study. b) Captive population of the same size, in which some aspect of the environment has induced abnormal behavior and concomitant abnormal physiology. Only a few individuals are within the normal range, hence, any experiment performed with this population will have poor validity. The increase in physiological variation induced by captivity will decrease the reliability of any experiment affected by the physiological measure, and will increase the number of individuals needed to achieve sufficient statistical power. Abnormal behavior therefore serves as a marker of abnormal physiology. c) Enrichments that return the population to physiological norms (i.e., similar to 1a) and reduce abnormal behavior will improve validity and reliability through reduced physiological variability.

unvarying inappropriate repetition of a particular set of movements and/or body postures that lack any goal or function. (2) **Impulsive/compulsive behaviors** involve the repetition of an inappropriate goal with variable flexible

goal-directed behavior (for detailed review, see Garner 2005). These categories are aptly illustrated by common ARBs seen in mice. Stereotypies in mice include bar-mouthing (when the mouse makes a series of repetitive functionless sham mouthing movements on a cage bar) and jumping (when the mouse rears on its hind legs and repeatedly jumps vertically on its hind legs, usually balancing on its tail, which is held rigid) (Garner and Mason 2002; Würbel et al. 1996). Other less common rodent stereotypies include “somersaulting,” “route-tracing,” and other forms of repetitive unvarying and functionless locomotion (e.g., Callard et al. 2000; Cooper and Nicol 1991; Nevison et al. 1999). The classic impulsive/compulsive behavior seen in mice is barbering (when one individual plucks similar idiosyncratic patterns of fur from its cagemates) (Garner et al. 2004a,b; Sarna et al. 2000), or from itself if singly housed (Garner et al. 2004b). Unlike a stereotypy, each episode of plucking behavior is flexible and goal directed (Sarna et al. 2000).

Prevalences of the behaviors described above can be extremely frequent under standard housing conditions. Approximately 98% of ICR (CD1) mice perform stereotypies (Würbel et al. 1996). In work with C57BL/6J mice, colleagues and I have observed stereotypies in approximately 80% of individuals (unpublished data). Barbering is less ubiquitous across mouse strains. Approximately 20% of C57BL/6J mice barber by 6 mo of age, whereas the behavior is rarely observed in CD1 mice (Garner et al. 2004b).

ARB and Brain Function in Humans

ARB in humans involves the brain systems responsible for selecting and sequencing behavior (Turner 1997), sometimes referred to as “executive systems.” The function of these systems and their role in ARB are outlined briefly below. For a detailed, recent review and to understand the relationship of ARB to experimental outcomes, the reader is referred to Garner (2005).

The control of behavior by the brain is divided between two systems. The first system selects and sequences individual responses on the basis of external (e.g., visual) and internal (e.g., blood sugar) stimuli that are immediately present. The second system selects and sequences goals, plans, and attention to salient features in the environment on the basis of internal abstract information that is not cued by stimuli (termed a “cognitive-attentional set”) (Fentress 1976; Garner 2005; Norman and Shallice 1986; Toates 2000). This goal selection system influences behavior by priming and editing the selections made by the response selection system. For example, when playing poker, I might bet very differently on the same hand depending on what other players have bet, on whether I am planning to bluff, and on what I know about the other players. The goal selection system is required to make these kinds of decisions, while the response selection system is responsible for initiating the actual movements involved in playing the game.

Similarly, the response selection system might divulge the information by initiating small unconscious movements (called “tells” by poker players) that are part of normal human body language. The goal selection system is involved in suppressing such automatic movements in particular circumstances (e.g., a good game of poker).

The brain is an expensive organ to maintain, and it is designed with very little redundancy. As a result, every area of the brain performs a different job, and if a particular area is damaged (i.e., physically lesioned) or dysfunctional (e.g., as a result of abnormal neurotransmission), there will be a job that the brain as a whole cannot accomplish. These failures are apparent as “software glitches” in behavior. The response selection system described above is distributed across a series of brain areas called the basal ganglia motor system. Damage or dysfunction in this system produces a software glitch called “recurrent perseveration,” or the inappropriate repetition of responses (Garner 2005; Luria 1965; Norman and Shallice 1986; Sandson and Albert 1984; Turner 1997). As an example, when a patient is asked the name of his or her father, the patient might respond correctly, but then when asked to name other members of the family, might continue to respond with the father’s name. In contrast, the goal selection system is distributed across the prefrontal corticostriatal loop, a series of brain areas. Damage or dysfunction in this system produces a software glitch called “stuck-in-set perseveration,” in which a cognitive-attentional set (i.e., goals, rules, plans, etc.) is inappropriately repeated while flexible goal-directed responses remain intact. In this instance, when a patient is asked to name the suit of each card turned face up from a deck of playing cards the patient responds correctly as each card is turned over (whereas with recurrent perseveration, the patient would have incorrectly repeated the suit named on previous cards). However, when the patient is then asked to stop naming the suit and to state the value of the card (e.g., ace, king, queen), the patient persists in naming the suit.

Psychiatrists quantify brain damage and brain function noninvasively, using “neuropsychological tasks” such as pencil-and-paper, card, or computer games that the subject can win only when a particular function (and hence a particular area of the brain) is intact. **Recurrent perseveration** can be quantified using a task called the “two choice gambling task,” and **stuck-in-set perseveration** can be quantified using one of several “set-shifting tasks” (Frith and Done 1983; Owen et al. 1993; Turner 1997). Several authors have argued that ARB in human patients is intimately related to perseveration, whereby ARB is the day-to-day expression of a deficit in behavioral control, and perseveration is the quantified experimental measure of the same deficit (e.g., Crider 1997; Garner 2005; Turner 1997). According to this hypothesis, the distinction between stereotypies and impulsive/compulsive ARBs mirrors the distinction between recurrent and stuck-in-set perseveration. Thus, stereotypies should be correlated with recurrent perseveration, but not stuck-in-set perseveration; and impulsive/compulsive ARBs should correlate with stuck-in-set perseveration, but

not recurrent perseveration. These predictions have been met in autism by correlating the severity of a patient’s ARB with his or her performance on neuropsychological tasks that measure perseveration (Turner 1997). The predictions have also been tested and confirmed for stereotypies in schizophrenia (Frith and Done 1983), and for either stereotypies or impulsive/compulsive behaviors in a series of other disorders (reviewed in Garner 2005).

ARB and Brain Function in Other Animals

Thus, ARB does indeed appear to be related to perseveration in humans, thereby raising the possibility that the same mechanisms underlie animal ARB. The involvement of executive (i.e., behavioral control) mechanisms in animal ARB is a particularly attractive hypothesis because the brain areas involved are highly conserved in vertebrates (Reiner et al. 1998), and it would explain (1) why animals persist in performing ARB despite self-injury or other deleterious consequences (Garner and Mason 2002); and (2) why some animal ARBs show remarkable epidemiological and phenomenological similarities to corresponding human disorders (e.g., barbering in mice; Garner et al. 2004b). Furthermore, the hypothesis is directly testable by extrapolation from human studies. Thus, animal stereotypies are predicted to correlate with measures of recurrent perseveration, and animal impulsive/compulsive behaviors are predicted to correlate with measures of stuck-in-set perseveration.

Stereotypies

To test the hypothesis described above, colleagues and I have assessed the correlation of stereotypy with recurrent perseveration. In earlier experiments, we used extinction learning as a measure of recurrent perseveration. We taught animals a simple spatial discrimination in a two choice maze, or in a touch-screen operant; and when this simple task was learned, we ceased to reward the correct response. Extinction learning is complete when the animal ceases responding to the previously rewarded side of the maze (or computer screen) and returns to choosing both options at random. Recurrent perseveration is apparent when the animal inappropriately persists in choosing the option that was previously rewarded. Using this simple paradigm, we demonstrated that stereotypies are indeed correlated with recurrent perseveration (and hence basal ganglia motor system function) in bank voles (Garner and Mason 2002), blue tits, and marsh tits (Garner et al. 2003a). This result has been replicated by other researchers with the same paradigm in Asiatic black bears and Malayan sun bears (Vickery and Mason 2003, 2005).

Extinction learning, however, is affected by processes other than extinction. Although we addressed these issues using internal controls in our earlier experiments, in an effort to improve the replication of the human data, we pro-

ceeded in later experiments to modify the two choice gambling task for use with animals. In this task, the subject is presented with a choice between two options and is randomly rewarded 50% of the time. The sequence of choices made by the subject is recorded and quantified for repetitiveness using information theory or Markov chain analysis (for introductions to these methods, see Attneave 1959; Martin and Bateson 1986). Using this paradigm, we again confirmed the correlation of stereotypy and recurrent perseveration in blue tits (Garner et al. 2003a). We further modified this task to avoid superstitious conditioning and developing side bias, so that the probability of being rewarded on each side of the maze (or computer screen) decreases the more that side is chosen. Using this bias-corrected gambling task, we have confirmed the correlation of recurrent perseveration and stereotypy in parrots (Garner et al. 2003b), and mice (unpublished data). Consistent with these findings, evidence is growing that a range of neurophysiological changes in the basal ganglia motor system may be involved in stereotypy (e.g., Martin et al. 1991; Turner and Lewis 2003; Turner et al. 2002).

Impulsive/Compulsive ARBs

We have begun working on the neuropsychological basis of barbering in mice. In autism, impulsive/compulsive ARBs are correlated with stuck-in-set perseveration that may be measured using the **IntraDimensional-ExtraDimensional set shifting task** (Turner 1997). This task has been adapted successfully for use in primates (Dias et al. 1996), rats (Birrell and Brown 2000), and mice (Garner et al. 2003c). The task and its use in animals are discussed in detail elsewhere (Garner 2005). Colleagues and I have found that barbers (i.e., hair-plucking individuals) performed poorly on the task compared with controls, and that the severity of barbering (measured as the amount of fur plucked from cage-mates) was correlated with stuck-in-set perseveration measured by the task (Garner et al. 2003c). In addition, we found that stereotypy and barbering, stereotypy and stuck-in-set perseveration, and barbering and recurrent perseveration were all uncorrelated, confirming the specificity of the association of particular forms of perseveration with particular forms of ARB. Thus, in summary, ARB in animals appears to involve the same neuropsychological mechanisms as ARB in humans. The question then arises regarding how this association between ARB and perseveration might affect experimental outcomes.

ARB, Brain Function, and Experimental Outcomes

In the foregoing text, perseveration is emphasized because it is an unequivocal indicator of the function of the systems controlling the selection and sequencing of behavior. However, disruption to these systems also affects many other

aspects of behavior. Manipulations such as isolation rearing (Jones et al. 1991) and amphetamine (Lyon and Robbins 1975), which induce stereotypy, perseveration, and alterations in behavioral control, provide useful clues to other potential effects on behavior. To confirm that the correlation of recurrent perseveration and stereotypy that we have observed does indeed reflect the function of the response selection system, we have also looked for the presence of these secondary indicators.

Animals treated with amphetamine show a rapid increase in the rate they switch between behaviors, which precedes the performance of amphetamine-induced stereotypy (Lyon and Robbins 1975). Initially, it might seem paradoxical that rapid switching between behaviors should be associated with stereotypy; however, this phenomenon parallels the distracted switching between goals that is associated with stuck-in-set perseveration (Norman and Shallice 1986). In both cases, perseveration and switching are consequences of a failure to inhibit behavior. Thus, recurrent perseveration and stereotypy can be seen as the failure to inhibit a response once it is completed, and switching inappropriately between behaviors can be seen as a failure to inhibit a new behavior before the current response is completed. Accordingly, recurrent perseveration and stereotypy should be correlated with the spontaneous rate at which animals switch between all behaviors in the home cage. We have confirmed this prediction in bank voles (Garner and Mason 2002). In blue tits, marsh tits (Garner et al. 2003a), and mice (unpublished data), we have identified more complex curvilinear relationships between these variables. Similarly, in schizophrenia, some authors (e.g., Crider 1997) have linked “disorganization” symptoms (which include pointless shifting between behaviors) to a disinhibition of behavioral control and to perseveration. However, in bears, Vickery and Mason (2005) failed to find a correlation between behavioral switching and stereotypy or perseveration. Thus, the commonality of this result is currently unclear.

The disinhibition of behavior that leads to high rates of behavioral switching in amphetamine-treated animals also induces hyperactivity (Lyon and Robbins 1975). Similarly, animals reared in isolation have been observed to be hyperactive (Einson and Morgan 1978). Accordingly, we found correlations between activity, behavioral switching, stereotypy, and perseveration in bank voles (Garner and Mason 2002). Other authors (e.g., Ödberg 1986) have similarly related stereotypies in captive animals to hyperactivity.

Amphetamine-treated animals not only make perseverative choices but also fail to suppress rapid responding to stimuli when it becomes inappropriate (Robbins 1997). When we examined the latency to respond to the two choices presented in an extinction task in bank voles, we found that stereotypy was correlated with persistent rapid responding (Garner and Mason 2002). We found stereotypy in parrots similarly correlated with perseveration and inappropriate rapid responding (Garner et al. 2003b).

Perseveration involves a failure of behavioral control, not of cognition, will, or knowledge. Thus, perseverative

human patients often report that they select incorrect perseverative responses despite knowing what the correct response should be and trying to make it (e.g., Luria 1965; Turner 1997)—the phenomenon termed “knowledge-action dissociation.” We examined the disparity between choices and timing of choices—both of which are indicators of knowledge of a learning task (Olton 1972)—to learn whether perseverative bank voles showed evidence of extinction task knowledge. Voles generally showed extinction of choices before they showed extinction of rapid responding, so we examined the degree to which each vole persisted in rapid responding at the point where it was choosing both sides of the maze equally. At this point, its choices revealed its “knowledge” of the answer to the problem, but its latency to choose did not necessarily agree. Accordingly, we found that stereotypy was correlated with a greater persistence of rapid responding at the point where choices indicated acquisition of the extinction task, which in turn indicated a greater knowledge-action dissociation. All of the data from bank voles were collected on the same individuals, therefore we were able to intercorrelate each of the measures discussed above.

If stereotypy is truly a product of a disinhibition of behavioral control, then each and every one of the measures described above should correlate with each and every other measure, as indeed we found (Garner and Mason 2002) (Table 1). In other experiments, however, these relationships have been less consistent than the relationship between stereotypy and perseveration (e.g., Garner et al. 2003a; Vickery and Mason 2005). Nevertheless, the general disinhibition of behavior that underlies stereotypy has the potential to affect a wide range of behavioral measures. Indeed, almost every widely used high-throughput behav-

ioral task is potentially affected by one of the correlates of stereotypy, as detailed in Table 2.

Thus, whether or not stereotypy is related to physiological differences (e.g., in stress physiology) that might affect other types of experiment, stereotypy is still likely to be a severely confounding factor in behavioral experiments. For example, in the food storing literature, a common paradigm is to pilfer the animal’s food stores, and to examine its behavioral response. The goal is to learn whether it gives up food storing (i.e., a naturalistic extinction response) or chooses to replenish its food stores. The chosen strategy, in theory, provides a great deal of information about the design and function of food storing behavior (Lucas and Zielinski 1998; Lucas et al. 2001). Unfortunately, however, when we examined the response of food storing marsh tits to pilfering, we found that the responses depended on their stereotypy level. Birds with high levels of stereotypy persisted with food storing, and birds with low levels of stereotypy gave up food storing (Garner et al. 2003a).

Potential for Concern

The disinhibition of response selection that appears to underlie stereotypy can and does affect experimental outcomes such as extinction learning, home cage activity, response latencies, and behavioral switching. Whether this phenomenon has a major impact on many of the other commonly used measures listed in Table 2 remains to be tested and is a subject of our ongoing research. Even if and when all of the effects hypothesized in Table 2 are observed, their meaning for scientific validity actually depends on a second issue—whether the level of brain function correlated with

Table 1 Intercorrelation of stereotypy and measures indicative of disinhibition of response selection in bank voles^{a,b}

| | Activity | Rate of switching of behaviors | Perseveration during extinction | Persistent rapid responding during extinction | Knowledge-action dissociation |
|-------------------|--------------------------------|--------------------------------|---------------------------------|---|--------------------------------|
| Stereotypy | $r = .857$ ($p = 0.0195$) | $r = .840$ ($p = 0.0035$) | $r = .642$ ($p = 0.0005$) | $r = .751$ ($p = 0.0095$) | $r = .790$ ($p = 0.0015$) |
| Activity | | $r = .871$ ($p = 0.020$) | $r = .791$ ($p = 0.030$) | $r = .878$ ($p = 0.007$) | $r = .883$ ($p = 0.020$) |
| Rate of switching | | | $r = .556$ ($p = 0.015$) | $r = .840$ ($p = 0.008$) | $r = .926$ ($p = 0.0005$) |
| Perseveration | | | | $r = .772$ ($p = 0.0145$) | $r = .830$ ($p = 0.0105$) |
| Rapid responding | | | | | $r = .820$ ($p = 0.017$) |

^aRedrawn with permission of Elsevier from Garner JP, Mason GJ. 2002. Evidence for a relationship between cage stereotypies and behavioural disinhibition in laboratory rodents. *Behav Brain Res* 136:83-92.

^bPartial correlation coefficients were controlled for the various internal controls in each measure. The p value of each is also given. This pattern of data indicates a single underlying causal factor consistent with disinhibition of response selection.

Table 2 Potential impact of demonstrated correlates of stereotypy on commonly used behavioral measures in ethology, experimental psychology, and high-throughput phenotyping

| Correlate of stereotypy | Species ^a | Behavioral tasks or measures that could be affected |
|--|--|--|
| Perseveration | Bank voles, parids, parrots, mice, bears | Any task where a response is switched or suppressed (e.g., many maze tasks, many operant tasks, patch switching, contrafreeloading, food store pilfering) |
| Knowledge-action dissociation | Bank voles | Any task in which behavior is used to detect the animal's knowledge, or the presence of cognitive process (e.g., most cognitive experiments) |
| Elevated activity levels | Bank voles | Activity, open field |
| Disinhibited switching between behaviors | Bank voles, parids, mice | Home cage behavior measures, and measures reliant on switching or the rate of responding (e.g., open field, elevated plus maze, social interaction test, emergence and runway paradigms) |
| Altered response timing | Bank voles, parrots | Startle, prepulse inhibition, resident intruder, runway paradigms |

^aThe species in which the correlate has been demonstrated are indicated. See text for details.

stereotypy is abnormal or simply a natural extreme that is expressed as an abnormal behavior in captivity. In other words, the question is whether stereotypy and barbering are malfunctional or maladaptive behaviors.

In the case of stereotypies induced by isolation rearing or amphetamine-induced stereotypies, we presume that the observed range of brain function is abnormal and that stereotypy is a malfunctional behavior. However, this presumption does not mean that all stereotypies, for example, indicate abnormal brain function. For example, obsessive and compulsive symptoms in obsessive compulsive disorder are correlated with abnormal brain metabolism (Baxter et al. 1992). Yet compulsive behavior cannot be taken blindly as an indicator of brain dysfunction because, as another example, subclinical levels of compulsive behavior remain correlated to neuropsychological measures of stuck-in-set-perseveration in healthy individuals within the normal range of brain function (Zohar et al. 1995).

It is critically important to disentangle this issue of ARB in laboratory animals for the following reasons: If stereotypies and barbering are maladaptive behaviors, then (1) animals showing ARB simply express a normal extreme of brain function strangely in captivity; (2) the range of variation seen in our behavioral tests can be considered normal, and *we do not need to worry* about the scientific consequences of ARB. If, however, ARBs are malfunctional, then (1) ARB is the product of abnormal extremes of brain function; (2) this abnormal brain function is a major source of excessive and uncontrolled variation, and *we really do need to worry* about scientific consequences. These possibilities are illustrated in Figure 2.

To date there is little direct evidence related to the distinction described above, but a very tentative picture does appear to be emerging in which many, if not most, examples of ARB appear to be malfunctional behaviors indicative of abnormal levels of perseveration. The following lines of

circumstantial evidence together argue that most stereotypies probably involve abnormal brain function (Garner and Mason 2002):

- Stereotypies often lead to self-injury (e.g., Novak et al. 2005; Odberg 1986; Veith 1986) and sometimes to impaired growth of offspring (e.g., Mason et al. 1995) or even increased mortality of offspring (e.g., Sørensen and Randrup 1986).
- Stereotypies appear not to develop in wild-caught adults, requiring instead an early critical period in captivity (Callard et al. 2000; Cooper and Nicol 1996).
- Stereotypies are thought to become increasingly resistant to enrichment as the animal ages, although only a few demonstrated examples have been reported (e.g., Cooper et al. 1996; Meyer-Holzappel 1968), and counter-examples exist (e.g., Garner et al. 2003a; Mills and Davenport 2002).

The only direct experimental evidence as to whether stereotypies are malfunctional or maladaptive comes from our work in songbirds. Midway through these experiments, we moved the birds to outside aviaries as part of routine husbandry for their yearly molt. When we returned them to their standard cages, we observed a reduction in stereotypy. Thus, this period of aviary housing served as a period of enrichment. For each bird, we maintained a stereotypy and a perseveration score from before and after the period of aviary housing. We then performed a meta-analysis wherein we normalized (i.e., z-scored) each of these two stereotypy and two perseveration measures for each individual from the pre- and post-“enrichment” experiments, and we calculated each individual’s change in stereotypy and perseveration using these normalized scores. The change in recurrent perseveration during aviary housing correlated with the change in stereotypy (Garner et al. 2003a). This result is consistent with the level of persevera-

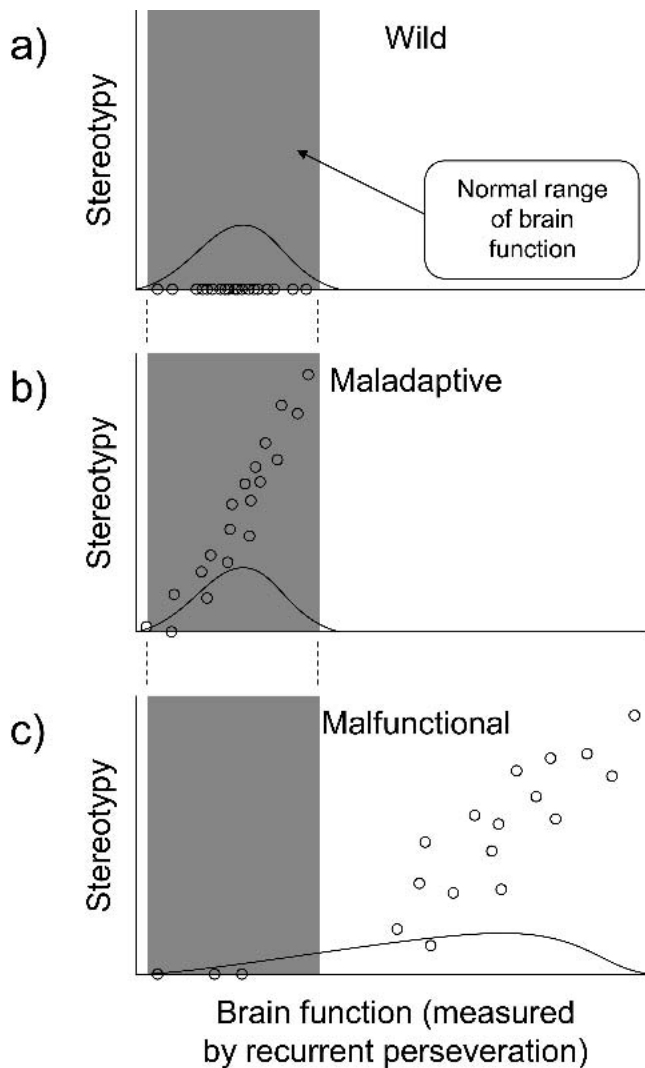


Figure 2 Two different explanations for the relationship between recurrent perseveration and stereotypy. a) Variation in brain function observed in a wild population of “normal” individuals. No stereotypy is observed. b) Captive animals show the same range of brain function as normal wild animals; however, this perfectly normal range becomes expressed as stereotypy in captivity. In this case, stereotypy would be an example of a maladaptive behavior. c) Captivity alters brain function, and the range of brain function observed in stereotypic animals is abnormal. In this case, stereotypy is a malfunctional behavior that indicates abnormal brain function. Nonstereotypic individuals should show levels of perseveration within the normal range, whereas stereotypic individuals would show a completely different range of perseveration, separated from the normal range.

tion seen before aviary housing (i.e., possibly before enrichment) being abnormal. However, these data should be interpreted cautiously because the experiment was essentially unintentional.

Although the case can be made that the level of perseveration associated with stereotypy may be abnormal, this

conclusion is extremely preliminary. Different answers to this central question almost certainly depend on the species, time in captivity, severity and type of stressor to which the animal is exposed, and type of stereotypy. The ultimate determination will be made in future work, but it is an empirical issue that can be assessed readily. We also are addressing this central question alongside current work of investigating the relationships between ARB, perseveration, and commonly used behavioral tests. To that end, we are investigating the effects of enrichment on stereotypy, perseveration, and experimental outcomes. If the level of perseveration observed in stereotypic animals is in fact normal, then enrichments that reduce stereotypy should not affect perseveration and should not reduce variability in experimental outcomes. Alternatively, if the level of perseveration seen in stereotypic animals is abnormal, then enrichment should reduce perseveration and reduce variability in experimental outcomes. Most importantly, the level of perseveration seen in nonstereotypic individuals should be the same in both enriched and unenriched animals.

Thus, by using enrichment, perseveration, and stereotypy in concert, it is possible to define the “normal” range of perseveration for animals such as laboratory mice. One cannot turn to wild house mice to define a normal range because laboratory mice are complex hybrids of multiple subspecies, so no “normal” wild population exists (Silver 1995).

Mouse Enrichment: Brief Commentary

In this article, a discussion of the mouse enrichment literature has been avoided for three reasons: (1) Other articles in this *ILAR Journal* issue address rodent enrichment directly (e.g., Baumans 2005; Smith and Corrow 2005); (2) the literature on mouse enrichment is relatively sparse, and few commercially available enrichments have been properly assessed or even demonstrated to be beneficial to mice; and (3) three recent reviews provide excellent overviews of the available literature (Latham and Mason 2004; Olsson and Dahlborn 2002; Sherwin 2002). In summary, for an enrichment to be enriching, it must be “biologically relevant” (i.e., it must address a behavioral need and facilitate the performance of highly motivated [often homeostatic] behaviors) (see the introductory text in this article, and Olsson and Dahlborn 2002). For instance, although toys and abstract stimulation may be biologically relevant for primates, the same stimulation is not necessarily appropriate for mice (Olsson and Dahlborn 2002). To date, the only device that does appear to be ubiquitously enriching to mice of different sexes, group sizes, and strains is nesting material (Olsson and Dahlborn 2002). Although shelters and complex cages may benefit some mice, these items may induce territoriality and hence become deleterious to others (especially group-housed males) (Olsson and Dahlborn 2002). However, stud-

ies explicitly testing whether nesting material reduces ARB are lacking, although cardboard tubes (which are often shredded and incorporated into nests) reduce stereotypy in mice (Würbel et al. 1998), and nesting material in the form of hay reduces stereotypy in bank voles (Cooper et al. 1996).

Conclusions and Closing Remarks

The following key conclusions are based on the information reviewed and discussed above:

1. Clear evidence exists that the brain mechanisms that produce ARBs also affect measures in behavioral experiments.
2. Enrichment might improve the validity, reliability, and replicability of behavioral experiment.
3. Conclusive empirical work has yet to be completed. In other words, in the particular example of the effects of ARB and enrichment on behavioral experiments, good welfare probably is good science; however, a good welfare scientist would wait for more data before arguing conclusively for the scientific benefits of enrichment.

Conclusion 3 can be further divided into three key unresolved issues: (a) whether or not the differences in brain function seen in animals performing ARB are abnormal; (b) which common behavioral paradigms are affected by ARB; and (c) whether or not enrichment does improve the quality of behavioral data. Despite the wide literature on the effects of enrichment on the mean outcome of various measures, few studies have tested specifically for effects on the variability of these measures (though see, e.g., Augustsson et al. 2003; Marashi et al. 2004; Wolfer et al. 2004). In addition, few have tested for relationships between ARB and behavioral experimental outcomes. Finally, the effects of enrichment on between-laboratory variability (or for that matter the relative importance of within- and between-laboratory effects on variability) are unclear—although recent work is beginning to explore this issue (Wolfer et al. 2004). Nevertheless, each of these issues can be answered empirically, and given the relationship of ARB to behavioral mechanisms discussed in detail above, its effects on other behavioral measures can be predicted directly. Thus, the interrelationships between ARB, brain function, enrichment, and (potentially) other behavioral measures provide a powerful system for exploring the argument that good welfare is good science.

Conclusion 3 is probably frustrating to many readers who are looking for scientific reasons that support or refute the enrichment of laboratory animals. In the meantime, as we await the results of additional studies, considering the following pragmatic questions may perhaps provide some perspective on the enrichment debate. (1) Is enrichment really such a dangerously noisy variable? Objecting to enrichment because of its possible negative effects on scientific

outcomes involves an implicit assumption—that enrichment has a greater potential negative effect on experimental outcomes than other sources of variability in the housing of laboratory animals. Therefore, (2) Is it reasonable to argue that providing a mouse with nesting material will perturb its physiology more than many other common laboratory provisions and decisions (e.g., switching from pine shaving to paper-product beddings, switching from open-top cages to ventilated cages, switching from the use of CRT computer monitors that produce ultrasound to LCD monitors that do not, or the hiring of new personnel who have different handling techniques)? Under most circumstances, the answer to both of these questions would appear to be “no.” Thus, in the face of these and other changes in husbandry and testing conditions that the average laboratory staff will experience over a number of years, the unknown negative impacts of introducing enrichment may be of little relative consequence.

Acknowledgments

Many of the ideas in this article grew from conversations with Drs. Kathryn Bayne, Georgia Mason, Joy Mench, Katherine Miller, Danny Mills, Chris Sherwin, and Hanno Würbel. The article also benefited from the comments of three anonymous referees.

Work on this manuscript was supported by Purdue Agricultural Experiment Station grant 2005-17554.

References

- Atneave F. 1959. Applications of Information Theory to Psychology: A Summary of Basic Concepts, Methods, and Results. New York: Henry Holt.
- Augustsson H, van de Weerd HA, Kruitwagen C, Baumans V. 2003. Effect of enrichment on variation and results in the light/dark test. *Lab Anim* 37:328-340.
- Baumans V. 2005. Environmental enrichment for laboratory rodents and rabbits: Rodent, Rabbit, and Research Requirements. *ILAR Journal* 46:162-170.
- Baxter LR, Schwartz JM, Bergman KS, Szuba MP. 1992. Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Arch Gen Psychiatry* 49:681-689.
- Bayne K. 2005. Potential for unintended consequences of environmental enrichment for laboratory animals and research results. *ILAR J* 46:129-139.
- Birrell JM, Brown VJ. 2000. Medial frontal cortex mediates perceptual attentional set shifting in the rat. *J Neurosci* 20:4320-4324.
- Callard MD, Bursten SN, Price EO. 2000. Repetitive backflipping behaviour in captive roof rats (*Rattus rattus*) and the effects of cage enrichment. *Anim Welf* 9:139-152.
- Capitanio JP, Lerche NW. 1998. Social separation, housing relocation, and survival in simian AIDS: A retrospective analysis. *Psychosom Med* 60:235-244.
- Chesler EJ, Wilson SG, Lariviere WR, Rodriguez-Zas SL, Mogil JS. 2002. Identification and ranking of genetic and laboratory environment factors influencing a behavioral trait, thermal nociception, via computational analysis of a large data archive. *Neurosci Biobehav R* 26:907-923.
- Cooper JJ, Nicol CJ. 1991. Stereotypic behaviour affects environmental preference in bank voles, *Clethrionomys glareolus*. *Anim Behav* 41:971-977.

- Cooper JJ, Nicol CJ. 1996. Stereotypic behaviour in wild caught and laboratory bred bank voles (*Clethrionomys glareolus*). *Anim Welf* 5:245-257.
- Cooper JJ, Ödberg F, Nicol CJ. 1996. Limitations on the effectiveness of environmental improvement in reducing stereotypic behaviour in bank voles (*Clethrionomys glareolus*). *Appl Anim Behav Sci* 48:237-248.
- Crabbe JC, Wahlsten D, Dudek BC. 1999. Genetics of mouse behavior: Interactions with laboratory environment. *Science* 284:1670-1672.
- Crider A. 1997. Perseveration in schizophrenia. *Schizophr Bull* 23:63-74.
- Davison GC, Neale JM. 1998. *Abnormal Psychology*. New York: John Wiley & Sons, Inc.
- Dias R, Robbins TW, Roberts AC. 1996. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380:69-72.
- Duncan IJH, Fraser D. 1997. Understanding animal welfare. In: MC Appleby, BO Hughes, eds. *Animal Welfare*. Wallingford UK: CAB International. p xiii, 316.
- Einon DF, Morgan MJ. 1978. Early isolation produces enduring hyperactivity in the rat, but no effect upon spontaneous alternation. *Q J Exp Psychol* 30:151-156.
- Fentress JC. 1976. Dynamic boundaries of patterned behaviour: Interaction and self-organisation. In: PPG Bateson, RA Hinde, eds. *Perspectives in Ethology*. Vol 1. New York: Plenum Press. p 155-224.
- Frith CD, Done DJ. 1983. Stereotyped responding by schizophrenic-patients on a 2-choice guessing task. *Psychol Med* 13:779-786.
- Garner JP. 2005. Perseveration and stereotypy—Systems-level insights from clinical psychology. In: J Rushen, G Mason, eds. *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare*. 2nd ed. Wallingford UK: CAB International (In Press).
- Garner JP, Dufour B, Gregg LE, Weisker SM, Mench JA. 2004a. Social and husbandry factors affecting the prevalence and severity of barbering (“whisker trimming”) in laboratory mice. *Appl Anim Behav Sci* 89:263-282.
- Garner JP, Mason GJ. 2002. Evidence for a relationship between cage stereotypies and behavioural disinhibition in laboratory rodents. *Behav Brain Res* 136:83-92.
- Garner JP, Mason G, Smith R. 2003a. Stereotypic route-tracing in experimentally-caged songbirds correlates with general behavioural disinhibition. *Anim Behav* 66:711-727.
- Garner JP, Meehan CL, Mench JA. 2003b. Stereotypies in caged parrots, schizophrenia and autism: Evidence for a common mechanism. *Behav Brain Res* 145:125-134.
- Garner JP, Wayne CM, Würbel H, Mench JA. 2003c. Barbering (whisker trimming) in laboratory mice involves the same brain systems as compulsive behaviors in trichotillomania, autism and other obsessive-compulsive spectrum disorders. *Proceedings of the 37th International Congress of The International Society for Applied Ethology*, Abano Terme, Italy: Fondazione Iniziative Zooprofilattiche e Zootecniche. p 75.
- Garner JP, Weisker SM, Dufour B, Mench JA. 2004b. Barbering (fur and whisker trimming) by laboratory mice as a model of human trichotillomania and obsessive-compulsive spectrum disorders. *Comp Med* 54: 216-224.
- Jones GH, Marsden CA, Robbins TW. 1991. Behavioral rigidity and rule-learning deficits following isolation-rearing in the rat—Neurochemical correlates. *Behav Brain Res* 43:35-50.
- Latham N, Mason G. 2004. From house mouse to mouse house: The behavioural biology of free-living *Mus musculus* and its implications for laboratory housing. *Appl Anim Behav Sci* 86:261-289.
- Lewis MH, Gluck JP, Beauchamp AJ, Keresztury MF. 1990. Long-term effects of early social isolation in *Macaca mulatta*: Changes in dopamine receptor function following apomorphine challenge. *Brain Res* 513:67-73.
- Lucas JR, Pravosudov VV, Zielinski DL. 2001. A reevaluation of the logic of pilferage effects, predation risk, and environmental variability on avian energy regulation: The critical role of time budgets. *Behav Ecol* 12:246-260.
- Lucas JR, Zielinski DL. 1998. Seasonal variation in the effect of cache pilferage on cache and body mass regulation in Carolina chickadees: What are the trade-offs? *Behav Ecol* 9:193-200.
- Luria AR. 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. *Curr Dev Psychopharmacol* 2:79-163.
- Marashi V, Barnekow A, Sachser N. 2004. Effects of environmental enrichment on males of a docile inbred strain of mice. *Physiol Behav* 82:765-776.
- Martin LJ, Spicer DM, Lewis MH, Gluck JP, Cork LC. 1991. Social deprivation of infant rhesus monkeys alters the chemoarchitecture of the brain: I. Subcortical regions. *J Neurosci* 11:3344-3358.
- Martin P, Bateson P. 1986. *Measuring Behaviour: An Introductory Guide*. Cambridge: Cambridge University Press.
- Mason GJ, Leipoldt A, de Jonge G. 1995. Why do female mink with high stereotypy levels have slow-growing offspring? *Proceedings of the 29th International Congress of the International Society for Applied Ethology*, Exeter, UK: Exeter University. p 133-134.
- Meyer-Holzappel M. 1968. Abnormal behaviour in zoo animals. In: MW Fox, ed. *Abnormal Behaviour in Animals*. London: Saunders. p 476-503.
- Miller KA. 2003. Behavioral measures of personality in Japanese quail (*Coturnix coturnix japonica*): Reliability, validity, and the effects of four types of environmental enrichment. PhD thesis. Department of Animal Science, University of California, Davis, CA.
- Miller KA, Garner JP, Mench JA. 2005. The test-retest reliability of four behavioural tests of fearfulness for quail: A critical evaluation. *Appl Anim Behav Sci* (In Press).
- Mills DS. 2003. Medical paradigms for the study of problem behaviour: A critical review. *Appl Anim Behav Sci* 81:265-277.
- Mills DS, 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. *Anim Sci* 74:95-101.
- Moberg GP. 2000. Biological response to stress: Implications for animal welfare. In: Moberg GP, Mench JA, eds. *The Biology of Animal Stress: Basic Principles and Implications for Animal Welfare*. Wallingford UK: CAB International. p 1-22.
- Nevison CM, Hurst JL, Barnard CJ. 1999. Strain-specific effects of cage enrichment in male laboratory mice (*Mus musculus*). *Anim Welf* 8: 361-379.
- Norman DA, Shallice T. 1986. Attention to action: Willed and automatic control of behaviour. In: Davidson RJ, Schwartz GE, Shapiro D, eds. *Consciousness and Self-regulation: Advances in Research and Theory*. New York: Plenum Press. Vol 4. p 1-18.
- Novak M, Meyer J, Tiefenbacher S, Lutz C. 2005. Deprivation rearing and stereotypies: Insights from primatology. In: Rushen J, Mason G, eds. *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare*. 2nd ed. Wallingford UK: CAB International (In Press).
- Ödberg FO. 1986. The jumping stereotypy in the bank vole (*Clethrionomys glareolus*). *Biol Behav* 11:130-143.
- Olsson A, Dahlborn K. 2002. Improving housing conditions for laboratory mice: A review of “environmental enrichment.” *Lab Anim* 36:243-270.
- Olton DS. 1972. Behavioral and neuroanatomical differentiation of response-suppression and response-shift mechanisms in the rat. *J Comp Physiol Psych* 78:450-456.
- Owen AM, Roberts AC, Hodges JR, Summers BA, Polkey CE, Robbins TW. 1993. Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinsons disease. *Brain* 116: 1159-1175.
- Raubenheimer D, Simpson SJ. 1999. Integrating nutrition: A geometrical approach. *Entomologia Experimentalis Et Applicata* 91:67-82.
- Reiner A, Medina L, Veenman CL. 1998. Structural and functional evolution of the basal ganglia in vertebrates. *Brain Res Rev* 28:235-285.
- Robbins TW. 1997. Arousal systems and attentional processes. *Biol Psychol* 45:57-71.

- Rosenzweig MR, Bennett EL, Hebert M, Morimoto H. 1978. Social grouping cannot account for cerebral effects of enriched environments. *Brain Res* 153:563-576.
- Sandson J, Albert ML. 1984. Varieties of perseveration. *Neuropsychologia* 22:715-732.
- Sarna JR, Dyck RH, Whishaw IQ. 2000. The Dalila effect: C57BL6 mice barber whiskers by plucking. *Behav Brain Res* 108:39-45.
- Sherwin C. 2002. Comfortable quarters for mice in research institutions. In: Reinhardt V, Reinhardt A, eds. *Comfortable Quarters for Laboratory Animals*. 9th ed. Washington DC: Animal Welfare Institute.
- Silver LM. 1995. *Mouse Genetics: Concepts and Applications*. New York: Oxford University Press.
- Smith AL, Corrow DJ. 2005. Modifications to husbandry and housing conditions of laboratory rodents for improved well-being. *ILAR J* 46: 140-147.
- Sørensen G, Randrup A. 1986. Possible protective value of severe psychopathology against lethal effects of an unfavourable milieu. *Stress Med* 2:103-105.
- Tankersley CG, Irizarry R, Flanders S, Rabold R. 2002. Functional genomics of sleep and circadian rhythm: Selected contribution: Circadian rhythm variation in activity, body temperature, and heart rate between C3H/HeJ and C57BL/6J inbred strains. *J Appl Physiol* 92:870-877.
- Toates F. 2000. Multiple factors controlling behaviour: Implications for stress and welfare. In: Moberg GP, Mench JA, eds. *The Biology of Animal Stress: Basic Principles and Implications for Animal Welfare*. Wallingford UK: CAB International. p 199-226.
- Turner CA, Lewis MH. 2003. Environmental enrichment: Effects on stereotyped behavior and neurotrophin levels. *Physiol Behav* 80:259-266.
- Turner CA, Yang MC, Lewis MH. 2002. Environmental enrichment: Effects on stereotyped behavior and regional neuronal metabolic activity. *Brain Res* 938:15-21.
- Turner M. 1997. Towards an executive dysfunction account of repetitive behaviour in autism. In: Russell J, ed. *Autism As an Executive Disorder*. New York: Oxford University Press. p 57-100.
- Veith L. 1986. Acral lick dermatitis in the dog. *Canine Pract* 13:15-22.
- Vickery SS, Mason GJ. 2003. Behavioural persistence in captive bears: Implications for reintroduction. *Ursus* 14:35-43.
- Vickery SS, Mason GJ. 2005. Stereotypy in caged bears correlates with perseverative responding on an extinction task. *Appl Anim Behav Sci* (In Press).
- Weiss JM. 1970. Somatic effects of predictable and unpredictable shock. *Psychosom Med* 32:397-408.
- Willner P. 1986. Validation criteria for animal models of human mental disorders: Learned helplessness as a paradigm case. *Prog Neuro-Psychopharmacol Biol Psychiatry* 10:677-690.
- Wolfer DP, Litvin O, Morf S, Nitsch RM, Lipp H-P, Würbel H. 2004. Cage enrichment and mouse behaviour. *Nature* 432:821-822.
- Würbel H. 2000. Behaviour and the standardization fallacy. *Nat Genet* 26:263.
- Würbel H. 2001. Ideal homes? Housing effects on rodent brain and behaviour. *Trends Neurosci* 24:207-211.
- Würbel H, Chapman R, Rutland C. 1998. Effect of feed and environmental enrichment on development of stereotypic wire-gnawing in laboratory mice. *Appl Anim Behav Sci* 60:69-81.
- 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.
- Zohar AH, LaBuda M, Moschel-Ravid O. 1995. Obsessive-compulsive behaviors and cognitive functioning: A study of compulsivity, frame shifting and type A activity patterns in a normal population. *Neuropsychiatry Neuropsychol Behav Neurol* 8:163-167.