Guiding Principles for Behavioural Laboratory Animal Science

Edition One: November 2013

Acknowledgements:

We thank **LASA** *(Laboratory Animal Science Association*), particularly members of the Education, Training and Ethics Section, the **BAP** *(British Association for Psychopharmacology)*, **BNA** *(British Neuroscience Association)* and the **ESSWAP Foundation** *(European Courses in Whole Animal Pharmacology)* for their financial support and professional input during the preparation and writing of these Guiding Principles.

Contents

Preface

These Guidelines are designed to help with the process of making informed decisions about the best way to carry out studies of animal behaviour in biomedical experiments. Although the topics concentrate on laboratory research, some apply to ethological studies in the natural environment as well. Even investigators who need to comply with regulatory requirements (and so cannot modify either the choice of procedure or the design of their studies) need to be aware of the principles described in these Guidelines.

New recruits to this field are the main target audience but anyone with a professional interest in behavioural laboratory animal science should find these Guidelines useful, even if only as a 'refresher'. Lay members of panels that deal with animal welfare and ethical review should find them helpful, also.

Each section includes questions that are intended to stimulate thoughts on how best to approach and carry out the experiments. This appraisal process will help to make objective decisions on topics such as: the justification for the choice of procedure; opportunities for optimising the welfare of the animals; implementation of the 3Rs[1]; and key ethical questions. It will also help to ensure that the experiments adopt the best design and avoid common pitfalls when interpreting the results.

Obviously, it would be impossible to provide point-by-point instructions on individual procedures because so many are used, in different experimental contexts, to study a wide range of species at various stages of development. Instead, the Guidelines focus mainly on principles that apply to all species. When necessary, they deal specifically with rats and mice because these are the species that are used most often in laboratory animal science.

The points that have been included were all flagged as fundamental at a workshop held in London in April 2012. This was attended by more than fifty experts, from the UK and beyond, who represented all aspects of behavioural laboratory animal science: breeders, veterinary surgeons, animal technicians, statisticians, regulators, animal welfare charities, industry, biotechnology centres, contract research organisations and universities. A second workshop, held in Harrogate in 2013, sought feedback on a draft copy before the final version was completed.

These Guidelines are based on the collective views and ideas of everyone who contributed to their development. The strong bias towards neurological and pharmacological procedures reflects the balance of interests of those who were energetic about passing on the benefits of their experience. However, we acknowledge that there is scope for expanding these Guidelines, in terms of detail and fields of research. For that reason, we regard this version merely as a first Edition. We hope that they will form a template for a document that evolves in future years.

[1] See Section 1.1

Summary and Overall Checklist

These Guidelines highlight important questions which should be resolved before, during and after studying the behaviour of laboratory animals.

1. They start by offering advice on **how to ensure** that the work will meet the highest standards in terms of **3Rs compliance, ethical considerations and animal welfare** (**Section 1**).

2. This is followed by an outline of how to **assess scientific validity**, which should be considered in order to optimise scope for translation into humans (**Section 2**).

3. The next section incorporates a list of questions that aim to prompt a **critical appraisal of the proposed work**. In particular, whether the use of *in vitro* alternatives has been considered and whether there are options for refining the procedure and/or optimising the experimental design (**Section 3**).

4. Ways of **ensuring that the experimental work is carried out by competent investigators** is covered in the next section. The legislative requirements for training and competence, as defined in the EU Directive (2010/63/EU), are covered first but it is pointed out that these merely focus on generic skills that are used in any field of research. Evaluation of animal behaviour often requires more specialised technical expertise: here, we offer some suggestions on how to ensure competence in these more specific experimental procedure(s) (**Section 4**).

5. Differences in the behaviour of animals, both within a group of individuals and between different groups, can influence experimental findings and even the main conclusions. This section gives some examples of the **many sources of such variation that should be considered** when designing the study and carrying out the work (**Section 5**).

6. This section gives examples of **environmental factors that can affect animals' behaviour.** The investigator must decide whether or not it is better to control for any of these factors, which would eliminate them as possible confounders, or to carry out the experiment in a more naturalistic environment, which could have greater translational relevance (**Section 6**).

7. The final section re-emphasises the **need to plan the statistical analysis of the data and the experimental design simultaneously (*i.e.* before starting the experiment**). There are also suggestions on how to avoid subjective or systematic bias in the data and an evaluation of the *pros* and *cons* of using the same subjects in more than one procedure. This section ends with a warning not to make unjustified assumptions about the cause(s) and consequence(s) of changes in animal behaviour, or the direction of causation of any such change (**Section 7**).

GLOSSARY

[In the context of these Guiding Principles]

Procedure:	A single behavioural protocol (*e.g*. the elevated plus-maze; Open Field)
Experiment:	One or more behavioural procedures, which collectively provide the data needed to meet a specific scientific objective
Project/Program:	A series of experiments, some of which will incorporate behavioural procedures, with a common aim

The 3Rs and Ethical Evaluation

1 THE 3Rs AND ETHICAL EVALUATION

This section outlines factors that should be considered at the start of the experiment so as to ensure that the work meets the highest welfare and scientific standards.

As in all biomedical research, it is essential to have **a clear experimental design and to ensure that this matches the planned statistical analysis of the data.** That is the only way to be certain that the data that emerge from the experiment will be amenable to valid statistical analyses.

This section considers the first step in this process, which is how to ensure 3Rs compliance and optimisation of animal welfare, and offers advice on who should be involved in that process.

It is important to acknowledge that animals are not merely a 'research resource' and that their use in an experimental procedure crosses an ethical boundary

There are two key principles that must be addressed in all research projects involving animals:

- the **3Rs principles of humane science;** and
- the **harm-benefit assessment** and **'ethical evaluation'** of the work.

1.1 The 3Rs

The 3Rs principles of humane science were published by Russell and Burch (1959) and are now internationally recognised as important both for good animal welfare and good science. For every project, and every procedure within each project, it should be evident that the 3Rs have been considered and that, wherever feasible, they have been adopted within the experimental design.

- ***Replacement***: Use of an alternative to a living animal (*e.g*. an *in vitro* alternative, computer simulation or human volunteer).
- ***Reduction:*** The use of the fewest number of animals needed to reach a clear conclusion.
- ***Refinement***: Procedural changes that reduce pain, suffering, distress and lasting harm for the animals and improve their welfare. Since animal welfare can have a direct effect on the validity and reproducibility of scientific data, this also benefits science. Refinement should be applied throughout the lifetime experience of the animal, including housing and care, transport, handling and restraint, identification, procedures and their long-term effects, and euthanasia.

Before designing an experiment, always consider who should be consulted to discuss the experimental procedure, its severity, humane end-points and whether the animals' welfare could be improved without undermining the experimental objectives. There may already be an established 'team' of animal care staff and other investigators dealing with such questions. If not, consider whether it would be helpful to assemble such a '3Rs group' or to discuss the work in depth with the local ethics and welfare body.

1.2 Ethical evaluation

In countries with a well-developed system for regulation of experiments that use animals, there is a requirement for **ethical evaluation of research projects**. This process may differ from one country (or even one establishment) to another. Nevertheless, the principles are common to all. Do not assume that the principles of the 3Rs and a favourable benefits/costs balance can be ignored if the procedure is to be carried out on animals that are not protected (*e.g.* most invertebrates under the Animals (Scientific Procedures) Act [A(SP)A] in the UK).

Within the EU, project authorisation is required for any study that could cause the animal pain, suffering, distress or lasting harm. Simple observational studies of animal behaviour are unlikely to cause any distress. However, authorisation by Member States is likely to be required for procedures that involve interventions, such as social isolation of rats or mice, or where the animal is subjected to a stimulus that evokes an escape or avoidance response.

In all cases, every effort must be made to ensure that the 3Rs have been considered adequately when approving an application to carry out a specific piece of research and that the objectives:

- ***maximise the likely benefits*** *of the project: e.g. by ensuring that the work optimizes its relevance to the human condition and through systematic dissemination of the results; and*
- ***minimise the predicted harms*** *through application of the 3Rs.*

To achieve these objectives, everyone involved in this process needs to be aware of the ethical implications of the research and to understand its complexities and constraints and to ensure that it does not violate accepted norms. This means that the first step, for investigators who intend to carry out the research, is to **identify the ethical and animal welfare issues** at an early stage when planning the work.

In addition to national regulations, individual research institutions should operate within a set of **local ethical principles.** These are usually implemented through bodies such as an *Animal Welfare and Ethics Review Body*[2] or an institutional *Animal Care and Use*

[2] AWERB in the UK

Committee[3]. The composition of these bodies can vary, but everyone (including the lay members and secretariat) must be aware of, and understand, the principles of the 3Rs.

These groups may go further than required by legislation: for example, they may limit the use of certain species, specific experimental procedures, or levels of suffering.

There are likely to be disparate views on:

- *whether the specific research objectives justify the use of animals;*
- *what constitutes a 'harm' to an animal; and*
- *the extent of harm that is acceptable and permissible in meeting the research objectives.*

As a consequence**, it is not possible for individual investigators to carry out a truly *independent* or *objective* ethical assessment of their own work**. Comprehensive assessment and weighing of harms and benefits needs informed input from a wide range of lay persons and regulators as well as professionals who include: *animal technologists, animal care and welfare professionals, laboratory animal veterinarians, scientific colleagues and collaborators*.

All these people will look at the ethical evaluation from **different perspectives.** By working together, as a coherent team, they will ensure robust decisions on the experimental procedure, its severity, humane end-points and whether the animals' welfare could be improved without undermining the experimental objectives.

Bear in mind that **animals, as well as humans, could benefit** from the work and this possibility needs to be taken into account when appraising harm *versus* benefit.

On the other hand, some experimental procedures might be **hard to justify** because of their harm *versus* benefit, or because they have little (or no) equivalence to any event likely to be experienced by humans ('translational validity'). Such procedures will need particularly **stringent ethical evaluation,** which might even decide that they should not be used at all.

Finally, it is unethical to waste animals by abandoning an experiment before its completion, without scientific justification (*e.g.* through lack of adequate funding).

1.2.1 Assessing benefits

Some questions to consider when assessing the benefits of a program of work and the procedures that will be used are:

- *Is it certain that the questions cannot be answered without using animals?*

[3] ACUC in the USA

- *Do the aims and objectives of the research fulfil a clear unmet medical/ veterinary/animal welfare need, and/or...*
- *.....Do they address an important scientific question?*
- *Will the findings from the series of individual experiments help to meet the aims and objectives of the overall program of work?*
- *Is the work original, relevant, realistic and timely?*
- *Is the science of high quality and demonstrably robust?*
- *Will the findings help us to understand the explanation for a behavioural disorder, or is the work aimed at trying to find a new treatment? If the latter, it might be necessary to use animals at an early stage of a progressive disorder and the ethical bars might differ.*
- *Will the benefits (both positive and negative findings) be disseminated as widely as possible and are they likely to be taken into account in future work?*
- *Will any contribution to the 3Rs be disseminated and applied in future work?*

1.2.2 Assessing harms to animals

It is important to have a comprehensive understanding of all the types and sources of harms for animals in order to carry out a harm-benefit analysis and to apply the 3Rs. Any pain, suffering, distress or lasting harm arising from factors such as **transport, housing and husbandry, identification, handling and restraint** (if applied) should be identified as well as harms from the main procedures. This 'contingent suffering' is important from a scientific point of view, as well as an ethical one, because it can affect the quality of the experimental results.

Laboratory animal investigators need to know whether the experiment is adversely affecting the animals' welfare. To achieve this, a good system of **welfare assessment** of the animals, targeted at the particular experiment, is an essential pre-requisite (see: Hawkins *et al*., 2011). But, in order to recognise changes following an experimental intervention, it is essential for the team to be familiar with the profile of behaviour that is typical ('baseline') for the species or strain being studied (see: Section 4 (Training); Berdoy 2002).

A **useful, preliminary assessment of the ethical burden of harm** imposed during a procedure is to decide whether the researcher would willingly use the equivalent procedure on another human being. However, it is important to bear in mind that perception of, or actual, harm in humans cannot be assumed to have the same effect in other animals. Some stimuli (*e.g.* smell) can have a greater influence on animals than humans and *vice versa* (*e.g.* the spoken word). Furthermore, there may be stimuli that humans cannot detect at all and yet can be disturbing or even harmful to animals: *e.g*. ultrasonic sound or exposure of a nocturnal species to bright light.

1.3 Record-keeping

Rigorous record-keeping is important for welfare and scientific purposes as well as to ensure compliance with legal requirements. The records should include documentation of unexpected or adverse effects, which are needed for **retrospective reporting of harm** within the EU.

Use **data-recording sheets**, to provide a clear, dated (and signed, if necessary) record of the results. Record precisely the conditions of the experiment. All details relevant to the particular procedure (*e.g.* ambient light intensity, food consumption), should be noted as a matter of routine. There should be sufficient detail to enable any other competent investigator to audit and replicate your study. The details on record should comply with the **ARRIVE Guidelines**, at least (Kilkenny *et al*. 2010).

The score sheets need to be **carefully designed**. They should include potential adverse effects as well as changes in behaviour (which may not be the same). Categories of behaviour to be included on the score sheets will be different for each procedure because different procedures will affect different behaviours in different ways.

The form should be refined, if any unexpected adverse effects or behavioural changes come to light during the course of the experiments. **The input of everyone** involved with the animals (animal care professionals, veterinary clinicians and regulators, as well as the investigators) will be helpful in this process.

An awareness of the spirit of **Good Laboratory Practice** (GLP) is helpful because this strives to reach the highest standards. However, the implementation of GLP for all research facilities is not a legal requirement and usually not feasible for small research groups who have limited administrative support for their work.

For more information about the process of ethical review, see:

- Guiding Principles on Good Practice for Ethical Review Processes http://www.lasa.co.uk/publications.html or www.rspca.org.uk/researchanimals.
- Animal Procedures Committee (UK: 2003) Review of cost-benefit assessment in the use of animals in research. www.gov.uk/government/uploads/system/uploads/attachment_data/file/119027/cost-benefit-assessment.pdf.
- IACUC: www.research.psu.edu/arp/experimental-guidelines/rodent-behavioral-tests-1
- www.nap.edu/openbook.php?record_id=10732&page=7
- http://grants.nih.gov/grants/olaw/guidebook.pdf
- www.animalethics.org.au/animal-ethics-committees

1.4 Retrospective Assessment

1.4.1 The project

There is a legal requirement within the EU to carry out a **retrospective assessment** of some projects to check:

- *whether the **objectives of the project have been achieved**;*
- *the **actual harms** that were suffered by each individual animal, including the severity of the procedures (see: Section 1.4.2);*
- *any contributions to the **further implementation of the 3Rs.***

A full retrospective assessment, as outlined above, is mandatory only for projects involving 'severe' procedures (using any species) or projects using primates. Nevertheless, it is good practice to conduct a full retrospective assessment for all projects. As well as helping to ensure that the science is on track, this process will bring additional benefits through highlighting any matters concerning training, resource and project management that need attention (see: Jennings *et al*., 2007).

1.4.2 Assessment of actual harms

Legislation within the EU requires actual harms to be assessed for all projects and this information must be submitted for statistical reporting purposes. This legislation (A(SP)A within the UK) further stipulates that a suitably qualified person must classify the actual severity of each procedure as 'non-recovery', 'mild', 'moderate', or 'severe' after the project has ended. This is done by reviewing day-to-day ('cageside') records of welfare assessments.

It is essential to have a good welfare assessment protocol that facilitates effective recognition and assessment of harms, both of which depend on a reliable record-keeping system (see: Section 1.3) and an appropriate monitoring regime. These are best achieved through **effective teamwork** between investigators, animal care-workers, veterinary staff and the regulators[4]. It might also be helpful to have input from the local ethical or animal care and use committee, such as the *Animal Welfare and Ethical Review Body* (AWERB). Collaboration between all these professionals will enable appropriate welfare indicators to be tailored to each procedure.

The process of assessment must consider the lifetime experience of the animal and identify different types and sources of suffering. These must consider the effects of identification, handling, restraint, any husbandry restrictions and euthanasia in addition to direct effects of the scientific procedures. It should also be borne in mind that these procedures can impose anxiety and distress, as well as physical pain.

[4] *e.g.* the Home Office Inspectorate (in the UK).

A sound understanding of normal animal behaviour is essential in order to recognise when there is a health or welfare problem within a study. This is especially important when assessing the welfare of animals used in behavioural research (and when interpreting experimental results). The records of welfare assessments, such as score sheets (see: Section 1.3), are used to assess actual harms for reporting to the regulator. They should also be used to review and reflect on each animal's lifetime experience and to help identify any scope for further refinement. For example, if some animals lost weight after surgery it might be necessary to adjust the pain relief protocol. Or, if animals displayed stereotyped behaviour after periods of restraint, the duration of restraint might need to be reduced, or steps taken to ensure that the animals are habituated to this procedure (but see: Section 6.2).

The European Commission has published guidelines on severity assessment, including how this can be assessed, with some worked examples (see: European Commission, 2012 & 2013). For more information on welfare assessment, see: Hawkins *et al.*, (2011).

1.5 Conflicts

Ensure that any contractual relationships relating to the research, including sponsorship, do not compromise the ethics of the work. If there is the possibility of **personal benefit** from the research outcome, this should be declared openly. In some cases, a higher benefit(s)/harm balance might be needed to be assured of the ethical basis for the research.

> ***Happy animals make good science***
> *T Poole (1997)*

References

Berdoy M. (2002) The Laboratory Rat: A Natural History. Film 27 minutes. www.ratlife.org .

Hawkins P, Morton DB, Burman O, Dennison N, Honess P, Jennings M, Lane S, Middleton V, Roughan JV, Wells S, Westwood K. (2011) A guide to defining and implementing protocols for the welfare assessment of laboratory animals: eleventh report of the BVAAWF/FRAME/RSPCA/UFAW UK Joint Working Group on Refinement. Lab Anim. 45:1-13.

European Commission Working document on a severity assessment framework (2012)

Federation of Laboratory Animal Science Associations (2005) Principles and practice in ethical review of animal experiments across Europe. Available at: www.felasa.eu

LASA (2007) The Value of Looking Back: Improving Science and Welfare through Retrospective Review. Poster produced by M. Jennings (RSPCA), B. Howard (LASA) & M. Berdoy (Oxford University) www.lasa.co.uk/publications.html.

Poole T. (1997) Happy animals make good science. Lab Anim. 31:116-24.

Russell WMS & Burch RL. (1959) The Principles of Humane Experimental Technique (2nd Edition, 1992 UFAW). Methuen, London.

Further reading:

Guillen J. (2010) The use of performance standards by AAALAC International to evaluate ethical review in European institutions. Lab Anim (NY). 39:49-53.

National Competent Authorities for the implementation of Directive 2010/63/EU on the protection of animals used for scientific purposes. Working document on a severity assessment framework. ec.europa.eu.
http://ec.europa.eu/environment/chemicals/lab_animals/interpretation_en.htm

Resource book for lay members of Ethical Review Processes: www.rspca.org.uk

Smith JA, van den Broek FA, Martorell JC, Hackbarth H, Ruksenas O, Zeller W. (2007) Principles and practice in ethical review of animal experiments across Europe: summary of the report of a FELASA working group on ethical evaluation of animal experiments. Lab Anim. 41:143-60.

Justifying Studies of Laboratory Animal Behaviour

2 JUSTIFYING STUDIES OF LABORATORY ANIMAL BEHAVIOUR

This section highlights criteria that must be satisfied in order to ensure the valid use of animals in behavioural studies and to optimise the likelihood that the findings will translate into humans.

The different elements of scientific validity are considered first. This is followed by an outline of the main factors that could limit or undermine the translational relevance of a specific behavioural procedure.

Always consider the translational relevance of procedures used to evaluate behavioural responses in animals

An ability to evaluate different aspects of animal behaviour is essential for improving our understanding of how those behaviours are regulated, the causes of illnesses that disrupt them and also for the discovery and development of new medical treatments. However, such studies are justified only if their validity is assured.

It would be helpful to liaise with a clinician who specialises in the behaviour/disorder that is the focus of the research in animals. It would be equally helpful to meet patients suffering from the disorder. These contacts will enable the strengths and limitations of the animal research to be evaluated and placed in context: *i.e*. what key behavioural signs of the disorder that is being studied would clinicians and their patients expect to see in animals?

2.1 Validity

The **validity** of all behavioural studies using laboratory animals depends on the research context and objective(s). The criteria that need to be met in order to satisfy the different types of experimental validity (*e.g.* translational validity, predictive validity, construct validity, face validity) have been debated for many years. It has become increasingly recognised that, while it is impossible to model human mental disorders in animals, intermediate behavioural phenotypes (sometimes called 'endophenotypes') that underlie, or form part of, such mental disorders can be identified. In some cases, directly homologous behaviours can be studied in animals. The details of these discussions are beyond the scope of these Guidelines but can be found elsewhere, (*e.g.* Willner, 1984; 1986; Belzung & Lemoine, 2011; Stephens *et al*., 2013).

2.1.1 Translational validity

The only firm **confirmation of validity for any behavioural study is through translation of any conclusions into humans** from other animals. This is not necessarily a single step, or a one-way process. Progressing research across a range of species can enable the construction of an effective translational pathway. For instance, back-translation from humans to monkey to mouse and then translation from mouse to monkey and then to humans can provide the validation that is needed

to make decisions that turn out to be robust and reliable. Yet, they are based on primary data that might only be determined ethically in a mouse.

Overall, translational validity implies that:

- *a novel treatment (e.g. drug candidate, other industrial chemical or surgical intervention) has no harmful or unwanted effects on the behaviour of humans that are not evident in animals*
- *the novel treatment (e.g. drug or surgical intervention) prevents the behavioural disorder(s) in both animals and humans*

The first step in testing for translational validity is to evaluate the **safety and efficacy** of the experimental intervention. This rests on a **systematic study** of its effects on the animals' spontaneous behaviour and gross physiological signs. This is an **open-ended interrogation**, which involves monitoring animals in a series of different environments, starting with the home-cage, followed by an assessment when held in the hand and then in a novel environment (*e.g.* open field).

A recommended step-by-step guide to the systematic behavioural assessment of small mammals, incorporating features of the Irwin and FOB (Functional Observational Battery) tests, was published in 2012 by the ESSWAP foundation.
Copies can be purchased through e-mail to: email@esswap.org

2.1.2 Predictive validity

This category of validity assumes that a specific behavioural response following an experimental procedure enables us to **predict** the human response. This assumption is based on an empirical finding that manipulation of a biological target (*e.g.* through a drug challenge, genetic alteration or surgical lesion) in an animal that produces a distinct behavioural response signals that it will also cause a specific response in humans. For example, self-administration of a drug by animals is taken as an indication that the drug is likely to be addictive in humans.

The behavioural responses in animals and humans do not have to be qualitatively similar. Also, it must not be assumed that the behaviour being evaluated in animals is equivalent to the human condition. For instance, **'immobility' in the *forced swim test*** is often use as a screen for antidepressant drugs (*i.e.* when animals stop swimming and adopt a floating posture) **but does not necessarily indicate that the animals' emotional state is the same as that experienced by depressed humans.**

When using these procedures, it is important to be aware of the risk of '**false positives'** (a positive response, which does not translate into humans). For instance, a drug that increases arousal and/or motor activity could reduce immobility in the

Forced Swim Test and yet turn out to be ineffective as an antidepressant (false positive): *e.g.* NK1 receptor antagonists.

There is also a risk of **false negatives** (a negative response in the test but a positive one in humans). For instance, increased exploration of the open-arms of an elevated plus-maze predicts an anti-anxiety action in humans. Yet, although some drugs do not produce this response in rodents they do relieve anxiety in humans (*e.g.* buspirone).

In fact, buspirone is an interesting example because it highlights the cautionary principle that some procedures can produce positive results only when testing **drugs with a similar mechanism of action.** For instance, the elevated plus-maze reliably produces positive results for anti-anxiety drugs that act at the same biological target (*e.g.* benzodiazepines and other drugs that bind to the GABAA-receptor), but not when testing drugs with a different mechanism action even though they might also have anti-anxiety properties in humans (*e.g.* buspirone, which modifies the function of serotonin-releasing neurones).

Behavioural procedures in which a specific response to an experimental challenge in animals is consistently paired with a specific response in humans acquire 'predictive validity'. However, the behaviour of the animals, before or after the treatment, is not necessarily a 'model' of any human behaviour or disorder

2.1.3 Construct validity

Further reasons for studying the behaviour of laboratory animals are:

- *to increase our understanding of the biological processes thought to govern behaviour in health and illness;*
- *to characterise the behavioural phenotype of an experimental intervention (e.g. genetic alteration, surgical intervention or drug treatment) that is thought to explain, exacerbate or prevent a human illness.*

The rationale for these studies is that greater insight into these biological processes will lead to better treatments for, or prevention of, human illnesses.

These studies **assume that the biological processes, which influence behaviour in health and illness, generalise across different species**. Whereas this seems to be the case for most aspects of behaviour, it is important to acknowledge that this might not always be so. For instance, some elements of cognitive processing in humans might be evident in other animals (*e.g.* spatial mapping) whereas others might not (*e.g.* the ability to solve complex mathematical problems).

The validity of any claim that a biological process is causally linked with behaviour ('construct validity') depends on a process of **continual reappraisal.** This must also

involve accumulating supporting evidence from a wide range of research fields (*e.g.* gene microarray, histopathology, endocrinology), not just behavioural studies.

2.1.4 Face validity

Face validity implies that there are overt similarities between behavioural responses and their context in animals and humans. However, this criterion is often ignored, not least because some procedures that modify animals' behaviour are **not relevant to day-to-day human experiences** (*e.g.* repeated inescapable electric shocks to induce a behavioural change, known as 'learned helplessness', in animals or the experiential conditions of the Morris Water-Maze test of cognition).

Also, a behaviour induced in animals might not have any obvious human analogue (*e.g.* fear-induced foot-thumping in gerbils). Such species-specific behaviours provide a window onto underlying neurobiological systems but it is worth considering the extent to which such limitations in face validity undermine the translational validity of the experimental procedure.

2.2 Complicating factors

2.2.1 More than one procedure in each animal?

Consider at the outset whether the best experimental design would involve carrying out a **series (battery) of procedures** using the same animals**.** An **advantage** of this approach is that it could strengthen the conclusions and reduce the number of animals needed for the study.

A **disadvantage** is that the sequence of the procedures could influence the findings (see: Section 7.4). Using a battery of procedures in the same animals can also compromise welfare because it could increase the **cumulative harm** experienced by each animal. Discuss with animal care, veterinary staff and regulators, the advantages and disadvantages of this approach and set limits on what should be done.

2.2.2 More than one symptom/sign in humans

Obviously, it is easier to evaluate a single behaviour in animals (*e.g.* locomotor activity) than a cluster of different behaviours. For instance, abnormal behaviours in rodents that are presumed (often with little, or no, justification) to be analogues of certain features of depression in humans include: a reduction in rodents' preference for sweetened water in the sucrose preference test (*anhedonia*) or an increase in submissive behaviour (*low self-esteem*).

However, psychiatric and neurological disorders in humans rarely affect only one aspect of our behaviour. As a consequence, a prediction that the treatment will have a therapeutic effect in humans cannot be justified merely on the basis of evidence that it prevents a single behavioural abnormality in animals.

Although the use of one well-founded test is worth more than a panoply of meaningless ones, it follows that the **translational validity of behavioural tests in laboratory animals can be strengthened by studying as many different aspects of their behaviour as possible**, both within a single procedure and across a battery of procedures. The use of multiple tests also helps to ensure that the behavioural response of interest does not have an extraneous explanation that is not relevant to the focus of research.

Consider the extent to which the experimental procedure is relevant to complex human behaviour(s), mood(s) and emotion(s), especially if only one aspect of behaviour is being evaluated

2.2.3 More than one illness

Humans often express symptoms of more than one disorder concurrently (‘**co-morbidity’**). As a consequence, it is essential to consider whether a change in behaviour, following an experimental intervention, could have more than one explanation. For instance, depression is common in patients with Parkinson’s Disease and both illnesses slow movement, but for different reasons. It follows that a reduction in motor activity in animal studies of Parkinsonism could be explained by an effect on mood (*e.g.* a loss of motivation to move) or impaired motor function (*e.g.* bradykinesia of neurological origin), or both.

Be clear about the aspect(s) of the human disorder(s) that are being studied in animals following an experimental intervention

2.2.4 More than one cause of an illness

Most, if not all, disorders that affect the behaviour of humans are **multifactorial**: *i.e.* there is no single genetic, or other biological, causal factor. However, several factors can increase vulnerability to the disorder, such as in schizophrenia. Also, it is now recognised that the behaviour of adults is strongly influenced by interactions between the genome and its environment (‘**epigenetic’** changes): *e.g.* harmful experiences in childhood.

These factors make it unlikely that any single targeted intervention in animals will replicate the full spectrum of the human condition. At best, they may only replicate limited aspects of the disorder.

Be wary of claims that a specific experimental intervention, which produces abnormal behaviour in animals, is a ‘model’ of a complex human disorder

2.3 Overcoming problems through back-translation from humans to animals

Recent research exploits the discovery of associations between inherited (genetic) traits and specific behavioural abnormalities in humans: these associations are defined as '**endophenotypes**'. The aim is to replicate this aspect of the human disorder through genetic alteration or drug treatment, for instance. This approach is helping to circumvent the problem that some key (diagnostic) elements of human illnesses (*e.g.* hallucinations in schizophrenia or melancholia in depression) **cannot be evaluated** in either humans or other animals.

A good example is a deficit in a behavioural trait, known as 'latent inhibition'. Normally subjects (humans or other animals), who have learnt that a buzzer signals nothing ('neutral stimulus'), find it hard to adapt to a change in that contingency (*e.g.* the same buzzer later signals a 'startling stimulus'). Latent inhibition is impaired in schizophrenics with specific genetic mutations ('endophenotype') and in rodents after treatment with a hallucinogenic drug (*e.g.* phencyclidine). In both cases, the impairment is prevented in humans and rodents by treatment with antipsychotic drugs. On the basis of such findings, measures of latent inhibition are used in studies of the neurobiology of schizophrenia and its treatment.

References

Belzung C & Lemoine M. (2011) Criteria of validity for animal models of psychiatric disorders: focus on anxiety disorders and depression. Biol Mood Anxiety Disord. 1:9. doi: 10.1186/2045-5380-1-9.

ESSWAP (2012): Behavioural observation: A guide to the systematic behavioural assessment of small mammals. Complied for ESSWAP and published by the University of Utrecht. Can be purchased from the ESSWAP office by contacting email@esswap.org .

Stephens DN, Crombag HS, Duka T. (2013) The challenge of studying parallel behaviours in human and animal models. Curr Top Behav Neurosci. 13:611-45.

Willner P. (1984) The validity of animal models of depression. Psychopharmacology (Berl). 1984;83(1):1-16.

Willner P. (1986) Validation criteria for animal models of human mental disorders: Learned helplessness as a paradigm case. Progress in Neuro-Psychopharmacology and Biological Psychiatry 10: 677–690

Further reading:

Amann LC, Gandal MJ, Halene TB, Ehrlichman RS, White SL, McCarren HS, Siegel SJ. (2010) Mouse behavioral endophenotypes for schizophrenia. Brain Res Bull. 83:147-61.

Choosing the Procedure

3 CHOOSING THE PROCEDURE

This section suggests factors that should be considered when deciding which behavioural procedure is the most appropriate for the proposed research. These factors are drawn up as a series of questions, which also seek confirmation that the use of animals is unavoidable and that both the procedures and the experimental design are 3Rs compliant.

Most of the points covered in this Section apply to all fields of research, even those that do not use animals. However, because 'behaviour' is such a complex response, which can be difficult to interpret correctly, there is a need to be especially clear about the nature and purpose of the procedure and to be satisfied that it is appropriate for achieving the experimental objectives.

If in doubt, consult others who have more experience in the field of behavioural research in question. Joining on-line discussion groups could also be helpful.

Ensure that all decisions (from choosing procedures to designing the whole experiment) are based on a thorough and critical appraisal of the literature, rather than historical or current tradition. It is essential to **read the literature** that describes the development and validation of the procedure you intend to use, **even though many of these landmark studies are unavailable online**. It is sensible to use a procedure that has produced consistent findings in more than one laboratory. Even so, consider whether the use of a newly developed procedure would be ethically *and* scientifically more appropriate than the use of a long-established one. A new approach could also provide valuable information on factors that can influence behaviour but have not been studied before in detail (*e.g.* see Section 7).

Ideally, an experimental challenge should be **ethologically relevant and use naturalistic stimuli** so as to produce responses that are typical of the animal's behavioural repertoire. These procedures are most likely to relate to analogous experiences in humans (*e.g.* exposure to a novel environment or reward).

Fulfilling this objective is not straightforward partly because **laboratory animals are not being studied in their natural environment**, which could influence the findings. For instance, wild rats and mice have complex, but quite different social systems. Whereas rats prefer social groups, male mice are often highly territorial and aggressive. Nevertheless, both species are group housed in the laboratory and their behaviour is usually monitored when housed individually (see: Berdoy and Drickammer, 2007).

3.1 How to decide on the procedure

The best experiments will ensure that all options for developing the 3Rs within the protocol have been explored and adopted into the design, whenever possible. Be prepared to **justify procedures that are regarded as contentious**: such as the water maze, fear-conditioning (using electric shock) and oral gavage. Specific procedures are

often regarded as the accepted benchmark but always consider the choice of procedure afresh, along with the basic scientific principles and ethics. Some important examples are given in Tables 1 - 4, which highlight key questions that should be resolved before embarking on an experiment to monitor the behaviour of laboratory animals.

Questions to ask and answer before starting the experiment

Table 1: General and technical points

GENERAL	
• ***Will the findings help to relieve suffering in humans or other animals?***	If not, how important and necessary is the study and over what time-scale?
• ***Is the choice of procedure influenced by a need to comply with published work?***	If so, consider whether there is scope for modifying the procedure, in line with the 3Rs (below) to develop a new approach
• ***Is this the best time to carry out this experiment?***	Would it be better to wait for an imminent technical advance or for results from other ongoing work? Would further in vitro testing be useful?
• ***Is the procedure used already in your laboratory?***	If not, who should be approached for training and advice during the early stages of the project? Ideally, carry out a pilot study to ensure that the findings replicate those from other studies/laboratories
• ***Are the facilities (equipment, rooms, expertise of technical staff) state of the art for the work?***	If not, consider carrying out the work in a laboratory that does have appropriate facilities
• ***Could the procedure be adapted to enable the use of a non-rodent species instead?***	Consider developing and validating the procedure to enable its use in a less sentient species
• ***Could any extraneous factors affect the results?***	Ensure that the test environment and the equipment are appropriate for studying animal behaviour

Table 2: Is there scope for Replacement?

REPLACEMENT	
• ***Is it certain that there is no <u>in vitro</u> alternative?***	Check for further information on this, *e.g.*: – NC3Rs (National Centre for the 3Rs (UK)) – FRAME (Fund for the Replacement of Animals in Research (UK) – ECOPA (<u>European Consensus-Platform for Alternatives</u> (EU)
• ***Is the choice of procedure based on a strong scientific rationale?***	If not, what is the justification for using the procedure?
• ***Is there a more direct route to reaching human studies (if this is the goal)?***	– If so, why is the proposed experiment necessary and how can it be justified? – Would an additional in vitro study help to interpret the behavioural results?
• ***Does the experiment test a specific hypothesis?***	Is this the <u>most</u> appropriate method to test that hypothesis?
• ***If there is no hypothesis: i.e. there is no specific prediction about the response to an experimental intervention.....***	What information will be gained from the experiment and what will it add to our understanding of the regulation of normal or abnormal behaviour, or therapeutics?
• ***Will the experiment enable you to decide what to do next?***	If not, in what other way(s) will the conclusion be useful?

Table 3: Is there scope for Refinement?

REFINEMENT	
• ***Of all the procedures that could be used to meet the research objectives, is this the most refined?*** • ***Does the choice of procedure provide the greatest amount of information that could help to answer the scientific question?***	If not, has the procedure been chosen merely because it is: – the most technically straightforward – the one that is used most frequently in other laboratories? – the most widely cited – the cheapest (no need to buy specialist equipment) – regulatory restrictions If any of these is the case, consider revising the choice of procedure to resolve any such conflict(s)
• ***Is there a biomarker that could be measured without causing additional harm***	Is so, this could 'add value' to the findings and increase their translational validity
• ***Is the procedure the most appropriate for studying the behaviour, or phase of a progressive neurological disorder, that is of greatest interest and/or relevance to therapeutic interventions in humans?***	If not, could this lead to misleading conclusions, especially regarding potential for translation?
• ***Are there any test conditions (e.g. husbandry, environment, handling protocol) that can be improved without affecting the conclusions?***	If so, consider changing factors that would reduce any associated suffering and/or improve animal welfare

Table 4: Is there scope for Reduction?

REDUCTION	
• ***Has the experiment been done before?***	If so, is the repetition necessary and do the objectives justify the use of more animals? For instance, is the proposed experimental procedure needed to ensure that findings in one laboratory can be replicated in another or to serve as an active control?
• ***Have any similar experiments been performed previously?***	If so, how will the results increase our understanding of the field and is this a sufficient advance on previous work?
• ***Does the experiment use the best (factorial) design that enables the maximum amount of information and the minimum use of animals?***	Could inclusion of another test factor at this stage save having to do another experiment later?
• ***Has the experimental design been checked by a statistician to ensure that the planned statistical analysis is valid?*** • ***Have all factors that could influence the results been considered in the experimental design?***	An NC3Rs interactive website is currently under development and is scheduled for launch early in 2014. This package is designed to facilitate and optimise these processes
• ***Would a pilot study help (particularly if the experiment has not been done before in your laboratory)?*** • ***Is it feasible to carry out a power analysis to estimate the appropriate sample size?***	Small-scale experiments (pilot studies) can provide essential information such as: – optimal stimulus intensity or dose range – when to carry out different parts of the study (*e.g.* timing of behavioural observations post-dose) – likely (and likelihood of) side-effects – sample variance (needed for power analysis) – unforeseen problems with the procedure But.... bear in mind that there is no way of knowing whether the mean and variance of data from a small, pilot experiment are typical of those from a larger sample

Reference

Berdoy M & Drickammer L. (2007) Comparative Social Organisation and life history of Rattus and Mus pp 380 – 392, in Rodent Societies: An Ecological and Evolutionary Perspective. J. Wolff and P. Sherman (ed.) University of Chicago Press. pp610

Training

4 TRAINING

This section deals with different aspects of training. The legislative requirements, specified in the EU Directive, are covered first. However, these are merely the generic requirements that apply to all aspects of laboratory animal science that are essential, as an introduction, to provide a general background to the subject. Further training will be required to ensure that researchers are competent to carry out each specific procedure in their experiments.

Successful studies of laboratory animal behaviour need research teams with relevant training and competencies

It is critically important to ensure that experiments are well designed with full implementation of all 3Rs. This will ensure that the work produces good quality data that can be interpreted properly and is sensitive to issues of animal welfare and public concern. This requires **competence** on the part of everyone in the research team. As in any other field of research, it takes time to acquire competence, a process that needs effective training and adequate supervision.

There are many excellent text-books that describe how to evaluate the behaviour of laboratory animals when using specific procedures, but these cannot substitute for advice from, and discussion with, experienced staff. Ideally, skills in specific methods and procedures should be available within your own establishment, but if this is not the case then you may need to **visit a laboratory where the behavioural procedure of interest is already well-established.**

Different laboratories may have different (but equally valid) views on which (and how) procedure(s) should be used. In such cases, it will be necessary to make on objective decision on which advice takes precedence (and why) in the context of the planned work. Always be prepared to seek further advice and revisit laboratories, if necessary, because questions/queries often come to light only when trying to set up the procedure in a new laboratory setting.

4.1 Legal requirements

The achievement and maintenance of competence, including the importance of supervision in this respect, is central to the EU Directive and UK law. In the context of animal work carried out under A(SP)A, there are clear **responsibilities on the part of both the researchers and the animal care staf**f to comply with the terms of the work authorised and any associated licences (*e.g.* project and personal licences).

In the UK, formal training of prospective personal and project licensees is delivered initially through mandatory attendance of accredited modular training courses (currently **Modules 1-5**, although these may change, see: EC Guidelines[5] (in prep)). However, these

[5] http://ec.europa.eu/environment/chemicals/lab_animals/pdf/Endorsed_consensus.pdf

training modules are brief and provide only an introduction to the ethical, legal and practical issues. **They are not designed to produce researchers who are competent in the required practical and specialist skills of behavioural science – or any other field of research**. After completion of the courses, personnel must therefore go on to acquire competence through on-the-job training, under supervision, together with other forms of continuous professional development (CPD).

One or more individuals within an establishment[6] will have a responsibility to ensure that all staff in the establishment are adequately educated and trained and that they are supervised until they are competent. For more details on these requirements and responsibilities see: LASA: Guiding Principles for Supervision and Assessment of Competence as Required under EU and UK Legislation.

There needs to be a clear training protocol for everyone involved in each procedure with clear expectations of how competence is defined

4.2 Specific training for behavioural studies

Provision of advice on how to carry out specific procedures is beyond the scope of these Guidelines, but some general principles apply to them all.

In addition to proficiency in practical procedures, researchers need to have a sound understanding of:

- *different types of validity and translatability;*
- *how to select the most appropriate procedure to address the scientific questions; and*
- *the many factors (species, environment etc. as set out in this report) that can affect the experimental data.*

Probably the most important principle is that **correct interpretation of changes in the behaviour** of animals, following an experimental intervention, **rests on a clear understanding of what counts as 'normal' and 'abnormal' behaviour** in each species or strain (not to mention reflecting on why these behaviours exist at all: *i.e.* why they have evolved). This means that it is essential to become familiar with the spontaneous behaviour of experimentally-naïve animals.

The best way to achieve this is first to study recordings of the behaviour of the animals in both their home cage and the experimental setting. This is also an opportunity to develop a system for scoring different elements of the animals' behaviour. Ideally, this process would be done manually and involve producing a score-sheet as a record (there is no

[6] e.g. Named Training and Competence Officer (in the UK)

substitute for watching the animals). Also, this process should be carried out using records (*e.g.* video or digital) of past work, rather than rehearsing on animals *in vivo*.

The next step is to become familiar with the effects of an experimental procedure that has well-documented effects on the animals' behaviour. As part of this process, **ensure that there is a clear training protocol for everyone involved in the procedure.** Only then, is it appropriate to go on to investigate changes in behaviour following an experimental intervention.

Seek advice from other researchers, who are experienced in using the procedure, as well as care staff, vets, statisticians

Reference

LASA 2013 Guiding Principles for Supervision and Assessment of Competence as required under EU and UK Legislation. A report by the LASA Education, Training and Ethics Section. www.lasa.co.uk/publications.html.

National Competent Authorities for the implementation of Directive 2010/63/EU on the protection of animals used for scientific purposes. Working document on Education and Training.. http://ec.europa.eu/environment/chemicals/lab_animals/interpretation_en.htm

The Animal

5 THE ANIMAL

Many factors can affect the behavioural response to a specific experimental intervention in animals and can **undermine translational validity.** This section considers factors that relate to the animal itself: *i.e.* regards the 'animal' as an experimental factor. The topics draw attention to many of the variables that could account for individual differences in behaviour. These variables could also help to explain why findings can differ from one laboratory to another, even when they all use the same procedure in a standardised environment.

Animals are not models of humans
Mice are not miniature rats

When starting a programme of work, it is essential to **select the most appropriate animal** (species/strain/sex/age) to achieve the scientific objectives. This is particularly important in behavioural studies because each species, strain (and even individual animal) can have its own behavioural (as well as physical) phenotype.

5.1 The use of outbred *versus* inbred strains

It is important to make an informed decision on whether to study an inbred or outbred strain because they could express different behavioural responses to a specific experimental challenge.

Studies of inbred *versus* outbred strains each have advantages and disadvantages. When a genetic trait (*e.g.* a specific behavioural response) is the focus of the research it is probably advisable to use an **inbred strain,** which expresses the trait prominently (Festing, 2010). This is not least because this approach will reduce sample variance and increase statistical power and so reduce the number of animals used in the study.

When investigating the extent to which strain could influence behaviour, consider starting the study by comparing **several inbred strains**, with 'strain' included as a fixed factor in a multifactorial design (see also: Section 5.2 and Section 7).

Note, however, that inbred strains that have been selected to display a particular behaviour (*e.g.* alcohol drinking) that is influenced by several genes (*e.g.* genes A – H), may actually have been selected for only a few (A-C) of those genes. Other strains, with an apparently similar phenotype may have been selected because of variations in different genes (*e.g.* C-F). **Results obtained with the two inbred strains may then give quite different conclusions as to the efficacy of a specific drug treatment.**

Bear in mind that experimental interventions that have **inconsistent effects on different inbred strains are unlikely to have strong potential for translation** into another species (especially humans) unless prompted by a reliable biomarker.

By contrast, variation in the behaviour of **outbred strains** could result in the behaviour of interest being masked or confounded by other behaviours (Festing, 2010). Nevertheless, a clear behavioural response to an experimental intervention in an outbred strain is a strong indication that the experimental procedure could have **translational potential.**

A database of information on the phenotypes of inbred mouse strains (sourced internationally) can be obtained from the ***Jackson Laboratory at www.jax.org/***
The rat genome database http://rgd.mcw.edu (rats) *gives equivalent information for different rat strains*

5.2 Choice of species or strain

This is an important factor because i**t cannot be assumed that a behavioural trait in one species or strain will generalise to others** (*e.g.* from rats to either humans or mice. See: Section 5.1). For instance, spontaneous motor activity of the C57BL/6J inbred strain of mice is greater than that of the sv129ev inbred strain. The latter also have a low 'ceiling' for motor activity. The C57BL/6J strain also shows vigorous tail-climbing and so is not suited to the Tail Suspension Test, which is used as a preliminary screen for antidepressant drugs. It follows that the choice of background strain can confound the effects of experimental interventions that affect motor activity or arousal.

For similar reasons, it should be borne in mind that the **background strain of genetically-altered animals** can influence the behaviour of interest and even interact with the effects of the experimental intervention. For instance, when mice are placed in the Open Field or elevated plus maze, the anxiety-like behaviour of mice with genetic ablation of the NK1 gene, depends on whether mice are derived from a C57 BL/6J or sv129ev background strain. Note, too, that mice from different suppliers may not be genetically identical. For instance, C57BL/6 mice supplied by Harlan have undergone a chromosomal deletion that has resulted in the loss of two genes.

Choice of species/strain is also important because **the structure of the biological target (*e.g.* an enzyme or a neurotransmitter receptor) in humans can differ substantially from the equivalent target in some species/strains, but not others.** These differences can influence the magnitude of the behavioural response to drug treatment and its translational relevance.

Given the potential for 'strain' to determine the behavioural response to an experimental treatment, it is worth considering whether it would be advisable **to assess routinely the effect of the test treatment in more than one strain**.

Examples of other factors to consider are:

- In some cases, the choice of species will be **constrained by a regulatory requirement**.

- Choice might be restricted by **practical criteria**. For instance:
  - It is not advisable to use albino animals, which are usually visually impaired, in tasks that rely on visual acuity: *e.g.* complex learning tasks that require animals to respond to visual stimuli in their environment.
  - Similarly, auditory impairment will impair Prepulse Inhibition or Latent Inhibition. Age-related impairment will also affect performance in these tests.
  - Measurement of pupil diameter is difficult in pigmented strains.
- The fate of a drug in the body (from its adsorption to elimination: **'pharmacokinetics')** can depend on species and strain. In turn, a difference in pharmacokinetics across different species/strains will influence the behavioural response to a drug challenge (*e.g.* Martignoni *et al*. 2006).

Space requirements and budget can be relevant considerations as well but it would be wrong, both scientifically and ethically, to allow such factors to determine the choice of the most appropriate species/strain.

Any decision to study a particular species or inbred strain must be supported by a rigorous scientific rationale, which makes it clear that the scientific objectives are likely to be met

5.3 Other sources of variation

5.3.1 Different suppliers

The behaviour of animals from different suppliers is not always the same, even when using the same inbred strain. The behaviour of different batches of animals from the same supplier can also differ. The source of this variability can be hard to trace not least because it can arise for a number of reasons. For instance, a given supplier does not always acquire the animals from the same **breeding colony**. Suppliers are also likely to combine animals from **different litters** or social colonies in a single batch for delivery. **Spontaneous genetic mutation** (genetic drift) is another possibility (see review by: Cassellas, 2011) but **epigenetic changes** (*i.e.* long-term (life-long?)) changes in gene expression, caused by certain environmental stimuli), are more likely. Such factors could include: maternal behaviour, housing or husbandry. For instance, different colonies will experience different handling protocols, or the litters might be weaned at different ages. Another example is whether or not animals are housed in individually-ventilated cages, which isolate groups of mice and deprive them of olfactory or auditory stimuli (see: Section 6.1.2).

Animals' experience during transportation is another factor which could affect their behaviour, irrespective of the time allowed for them to habituate to their new environment.

5.3.2 Age

A great deal of behavioural research focuses on sexually mature, young adults. This approach neglects the possibility that young and/or elderly subjects might respond to the experimental intervention in different ways. Age is an important variable to take into account, whether in the context of using animals at a particular stage of life (*e.g.* aged animals, as in research on dementia) or as a potential source of variability of the data. For instance:

- Many studies use rats from PND 35-60 until their weight reaches 300 g. The exact period of 'adolescence' is poorly defined in rodents but is likely to fall within this window. However, many developmental changes occur during adolescence and they can affect behaviour and drug responses.
- Longitudinal studies sometimes involve a continual series of measurements of behaviour over a period of several months. It is important to be certain that the same elements of behaviour are being measured at all time points.
- Some aspects of **motivation** change with age *(e.g.* old rats are less easily motivated by hunger or thirst). This can be a problem if gustatory stimuli are used to motivate the animals because a reduction in appetite will reduce animals' motivation to perform the task.
- The **sensitivity to pharmacological agents** can depend on age. In some cases, the response declines as the animals grow older (*e.g.* the effect of cannabinoid receptor antagonists on alcohol preference and food intake of mice).
- The **response to a drug can become more vigorous** and prolonged as the subjects grow older. This is usually because elimination of drugs from the body relies on the liver and/or kidneys and their efficiency deteriorates with age.
- The integrity of the **blood-brain barrier** is not fully developed in juvenile animals and can deteriorate in elderly subjects. In both cases, there will be greater penetration of the brain by drugs that are normally prevented from reaching the brain by this physiological 'filter'.

Also, it should be recognised that when using procedures that involve many months of animal training, they may be near the end of their life-span at the time of testing. It is important to consider whether conclusions from such studies might not be valid for younger animals.

5.3.3 Size

Body mass (and volume) can be an important variable for many reasons. For example:

- Body volume can affect the likelihood of intercepting light-beams in activity cages.

- Regional blood flow and the size of the fat store will affect the rate of redistribution of lipid-soluble drugs between the plasma and brain and, as a consequence, the behavioural response.

The **body mass** of genetically-altered mice sometimes differs from that of their wildtype. In such cases, it is important to decide whether drugs should be tested in animals of the same age (*i.e.* adjusting the dose to correct for the difference in weight) or of the same weight (*i.e.* using the same dose, but testing mice of different ages).

Whichever of these two factors is the most important will depend on the objectives of the experiment, but mass is normally regarded as the fixed factor.

5.3.4 Sex

The behavioural repertoire, physiology, pharmacology and biochemistry of male and females differ in all species.

The **variability** arising from cyclic changes in the behaviour of females (during the oestrous cycle) is often used to justify studying only male subjects. This approach reduces the number of animals that need to be studied in order to reach a firm conclusion but there are a number of factors to consider when deciding whether it is justified to study only males:

- Given that treatments for the majority of human disorders are intended to be effective in both sexes, it is worth considering whether the behavioural changes in laboratory animals should be monitored in both males and females (see: Section 2).
- Several disorders are more prevalent in one sex than the other (*e.g.* Attention Deficit Hyperactivity Disorder (ADHD) is more common in males but depression and anxiety are more common in females). These differences could give important clues to their causation and/or prevention. Yet, most experiments use sexually mature, adolescent males (see: Section 5.3.2).
- Studying both sexes reduces the waste of (unstudied) animals and cost of maintaining the breeding colonies.
- Take account of the possibility that using the same apparatus to study both males and females could affect the behaviour of both sexes.

More information on the scientific implications of sex differences in behavioural studies, together with common misconceptions in this field, has been reviewed by McCarthy *et al* (2012).

Consider the extent to which the animals to be studied can be regarded as analogous to the target human population in terms of age, sex, stage of development.....

References

Casellas J. (2011) Inbred mouse strains and genetic stability: a review. Animal. 5:1-7.

Festing MF. (2010) Inbred strains should replace outbred stocks in toxicology, safety testing, and drug development. Toxicol Pathol. 38:681-90. doi: 10.1177/0192623310373776. Epub 2010 Jun 18.

Martignoni M, Groothuis GM, de Kanter R. (2006) Species differences between mouse, rat, dog, monkey and human CYP-mediated drug metabolism, inhibition and induction. Expert Opin. Drug Metab. Toxicol. 2:875-94.

McCarthy MM, Arnold AP, Ball GF, Blaustein JD, De Vries GJ. (2012) Sex differences in the brain: the not so inconvenient truth. J Neuroscience 32(7): 2241-2247.

Further reading

Crawley JN. (2010) What's Wrong With My Mouse?: Behavioral Phenotyping of Transgenic and Knockout Mice, 2nd Edition. Wiley & Sons. ISBN: 978-0-471-47192-

Guidelines for the Care and Use of Mammals in Neuroscience and Behavioral Research. Washington, DC: The National Academies Press, 2003.

Iivonen H, Nurminen L, Harri M, Tanila H, Puolivali J. (2003) Hypothermia in mice tested in Morris Water Maze. Behav Brain Res 141: 207-213.

Terry AV. Jr (2009) Spatial Navigation (Water Maze) Tasks. In: Buccafusco JJ, editor. Methods of Behavior Analysis in Neuroscience. 2nd edition. Boca Raton (FL): CRC Press; 2009. Chapter 13. Available from: www.ncbi.nlm.nih.gov/books/NBK5217/

Wahlsten D. (2011) Mouse Behavioral Testing: How to use Mice in Behavioral Neuroscience. Amsterdam: Elsevier.

The Environment

6 THE ENVIRONMENT

The ways in which differences in the environment of experimental animals can affect their behaviour are unlimited. This section describes some well-documented examples. It would be impossible to control for all such sources of variation, but the researcher needs to decide which, if any, should be standardised across all experiments so as to eliminate it as a factor that could confound interpretation of the results.

The experiment starts at the moment the animal is conceived and the researcher is part of the environment

In behavioural tests, the environment is a source of many experimental variables and so it is important to be aware of its impact on the animal's lifetime experience - not just at the point of collecting the data.

As in humans, factors that affect adult behaviour range from subject-related factors, such as **early life experience** (*e.g.* a change of littermates) to a change in the **laboratory environment** (*e.g.* building noise, or the smell of perfume or paint).

Consider whether standardising the environment would improve the quality of the data. On the other hand, consider whether behaviours that are evident only when the environment is stringently controlled are likely to be interesting or important in a translational context. The advantages of carrying out an experiment in a stringently controlled environment (*e.g.* a sound-proofed box) must be compared with those arising from the alternative design in which there is negligible standardisation of environment. It could be argued that only behaviours expressed under the latter condition (a 'naturalistic environment') are likely to have strong translational potential.

Do not assume that collaboration with another research group justifies animals travelling from one establishment to another. Ideally, the research should be performed on one site. Certification of animals as fit to travel does not by itself imply that the travel is ethically acceptable. If travel is justified, then it should not be a permanent arrangement or involve a large numbers of animals.

If the experiment involves adjusting the environment appreciably, for example by imposing 24 h darkness or a reduced ambient temperature, it is advisable to consider the long-term consequences of such a change before carrying out further experimental procedures that are intended to affect behaviour.

Examples of common environmental confounders are described below. In all cases, objective and systematic welfare assessments should be developed to ensure that any behavioural differences, due to extraneous environmental factors, are identified and eliminated at an early stage.

But....it is important to consider whether behavioural changes that are evident only when animals are tested in a stringently controlled environment are likely to have translational relevance

6.1 Housing

Housing can have a major influence on animal behaviour and welfare. After weaning, the **social hierarchies** that develop in different cages can determine animal behaviour and cognition (*e.g.* Fitchett *et al.*, 2005). The extent of this influence will depend on the species and strain of animal. For instance, the effects of social isolation on the behaviour of adult animals that are highly territorial (*e.g.* male mice) will differ from those that live in social groups (*e.g.* rats).

Differences in the behaviour of genetically-altered and wildtype mice, derived from separate, inbred homozygote colonies, could arise from differences in the **maternal behaviour** of the two strains. Conversely, a difference in interactions with littermates could well lead to differences in the behaviour of homozygote adults bred from heterozygote littermates *versus* homozygote inbred strains.

In fact, the **early experiences** of laboratory animals, especially mice and rats, are now being manipulated deliberately in the laboratory. These procedures usually involve imposing some form of stress (*e.g.* maternal separation). Such stimuli cause long-term disruption of the behavioural and hormonal phenotype that could help to explain behavioural disorders in human adults (*e.g.* impaired cognition or vulnerability to psychiatric disorders associated with stress, such as anxiety and depression).

Be aware of the provenance of each subject and regard 'litter' (which affects the influence of genetic and early-life experiences on behaviour) as an experimental factor

6.1.1 Stocking density

Litter size and composition affects nutritional status and interaction with the mother. This variability can be resolved by culling pups to standardise the number of offspring in each litter. If the intention is to study only males, then the females are culled after weaning as a matter of routine. However, this practice needs to take into account the ethical burden of **overbreeding.**

Some experiments need subjects to be housed **singly**, despite evidence that separation of rats from their conspecifics, even if only for a short period, is stressful and will affect their behaviour. As a consequence, **social isolation** is a welfare issue that should always be questioned and avoided, if possible.

6.1.2 Cage environment

- **Environmental enrichment**
  Although enriching the animals environment (cage/pen) benefits their welfare, there is concern that a lack of standardisation of **environmental enrichment increases the variance** of a wide range of physiological and behavioural measures (Simpson & Kelly, 2012).

  The type of enrichment needs careful consideration. For instance, a running wheel would not be appropriate for an animal equipped for EEG monitoring, which will make them vulnerable to snagging and physical harm.

  Given the importance of environmental enrichment to animal welfare, these concerns about variability are best addressed by **ensuring enrichment items are standardised** across all groups of subjects (and experiments), rather than by withholding enrichment altogether.

- **The position of the cage in the rack**
  This will affect light exposure, which can influence behaviour (*e.g.* in studies of the effects of light entrainment of circadian rhythms). Consider whether the position of the home-cage should be rotated through the rack, or remain constant, especially in studies of behaviours that are influenced by light intensity.

- **Bedding material**
  Appropriate bedding offers a form of enrichment that is easy to instigate (*e.g.* using bedding that mice can use to make nests). For reasons that are not understood, mice and rats express a preference for wood shavings, over wood chips (Kirchner *et al.*, 2012), and corncob bedding inhibits reproductive behaviour and increases aggression in rats. Also, animals' behaviour is more consistent when they have adapted to their new bedding after their cages have been cleaned: *i.e.* when territorial scents have been re-established.

- **Individually ventilated cages (IVCs)**
  IVCs are being used increasingly commonly for the breeding and maintenance of laboratory rodents because each cage is microbiologically isolated. This minimises the risk of infection in the colony and the exposure of humans to laboratory animal allergens (LAAs). Also, cage-change frequency can be reduced. This is thought to have beneficial effects for the well-being of rodents (Burn and Mason, 2008), with no measurable effects on micro-environmental conditions, health or behaviour of mice (Hawkins *et al*, 2003; Rosenbaum *et al*, 2009).

Mice need time to adapt to the environment in an IVC rack but breeding performance is either the same as, or better than, in open-cage systems. This is also true for animals moved to open-cage systems and could be due to factors such as a change in background noise (Tsai, *et al*, 2002). Little is known about the effects of housing in IVCs on behaviour, but this should be considered as a possibility. For instance, most rodent behavioural testing is performed outside the home-cage, usually in an open (non-IVC) environment: it cannot be assumed that the behaviour of the animals will be unaffected by acclimatization to their new surroundings.

6.1.3 Facility environment

Many environmental factors in the holding unit (*e.g.* temperature, humidity) are stringently controlled, but others are not (*e.g.* noise, olfactory stimuli and personnel who come into contact with the animals). **Any change in the animals' environment can affect their behaviour.**

This source of variation should be taken into account when **moving animals** between holding and procedure rooms, or when there are changes in husbandry routine or animal care staff/personnel carrying out procedures. The influence of such extraneous factors on behaviour might need to be assessed separately.

- **Olfactory stimuli** (*e.g.* from a predator species)
  These are an obvious source of variability, especially in multi-use procedure rooms. The disturbance to olfaction arising from cage cleaning is another factor to consider. These stimuli can emanate from staff themselves: *e.g.* through soaps, perfumes or contact with other animals (especially predators) at home and can be carried on clothes.

- **Noise**
  Noise in the environment affects animals' behaviour, which can be a problem (*e.g.* File and Fernandes, 1994). Carrying out experiments in **individually-ventilated, sound-proofed enclosures** helps to ensure that animals are not stressed or distracted. However, this approach risks studying behaviours that are evident only in a low-stress, soundproof environment and so are not relevant to human day-to-day experiences.

  Some electrical equipment (including electronic cameras) can emit at **ultrasonic frequencies,** which cannot be detected by humans. Ultrasonic vocalisation is common in rodents and the frequency of emission depends on whether the animal is experiencing pleasurable or aversive stimuli (Lahvis *et al.*, 2011). However, emissions in the region of 20 kHz are within the range for an alarm signal for mice and rats.

- **Light**
  Light intensity is a major factor that can affect behaviour and, if too great, can even provoke retinal degeneration. This is particularly important for mice, which (like rats) are most active at dusk and dawn. Be aware that it

needs only a 2-3 min burst of light during the dark-phase to reset the circadian clock.

6.2 Habituation *versus* sensitisation to the test environment

Some experiments start by repeatedly exposing the animals to the test environment so as to produce **habituation.** The aim is to prevent an acute behavioural response, following transfer to the novel environment, which could confound the results of the study. However, it is important to consider whether the animals could become **sensitized** to the environmental stimuli instead. For instance, the behavioural response can increase (through sensitization) rather than diminish (through habituation) after repeated exposure to a stressful environment.

Factors that can determine whether animals habituate or become sensitized include: the type of stimulus; its intensity; and the inter-stimulus interval. Whether or not **habituation or sensitization** has developed cannot be assumed and must be determined empirically (*e.g.* by measuring secretion of stress hormones).

Be aware of, and control for, environmental stimuli that could affect animals' spontaneous behaviour, especially if these stimuli could modify the response to the experimental challenge.

- ***Ensure that everyone who works with the animals is aware of environmental factors that could influence animal behaviour, in all relevant settings***
- ***Consider whether the behavioural changes that are evident only when animals are tested in a stringently controlled environment are likely to have translational relevance***
- ***Do not assume that animals develop behavioural habituation when they are exposed repeatedly to the test environment***

References

Burn C & Mason GJ. (2008) Effects of cage-cleaning frequency on laboratory rat reproduction, cannibalism and welfare. Applied Animal Behaviour Science. 114: 235-247.

File SE & Fernandes C. (1994) Noise stress and the development of benzodiazepine dependence. Anxiety 1: 8-12.

Fitchett AE, Collins SA, Barnard CJ, Cassaday HJ. (2005) Subordinate male mice show long-lasting differences in spatial learning that persist when housed alone. Neurobiol Learn Mem. 84:247-51. Epub 2005 Oct 5.

Hawkins P, Anderson D, Applebee K, Key D, Wallace J, Milite G, MacArthur-Clark J, Hubrecht R, Jennings M. (2003) Individually ventilated cages and rodent welfare: Report of the 2002 RSPCA/UFAW rodent welfare group meeting. Animal Technology and Welfare, 2: 23-34.

Kirchner J, Hackbarth H, Stelzer HD, Tsai PP. (2012) Preferences of group-housed female mice regarding structure of softwood bedding. Lab Anim. 46:95-100. doi: 10.1258/la.2011.010173. Epub 2012 Jan 31.

Lahvis GP, Alleva E, Scattoni ML. (2011) Translating mouse vocalizations: prosody and frequency modulation. Genes Brain Behav. 10:4-16. doi: 10.1111/j.1601-183X.2010.00603.x.

Rosenblaum MD, VandeWoude S, Johnson TE. (2009) Effects of cage-change frequency and bedding volumes on mice and their microenvironment. Journal of the American Association for Laboratory Animal Science. 48: 763-773.

Simpson J & Kelly JP. (2011) The impact of environmental enrichment in laboratory rats--behavioural and neurochemical aspects. Behav Brain Res. 222:246-264. doi: 10.1016/j.bbr.2011.04.002. Epub 2011 Apr 12.

Tsai P-P, Oppermann D, Stelzer HD, Mähler M, Hackbarth H.(2003) The effects of different rack systems on the breeding performance of DBA/2 mice. Laboratory Animals. 37: 44-53.

Further reading

Hogg S. (1996) A review of the validity and variability of the elevated plus-maze as an animal model of anxiety. Pharmacol Biochem Behav. 54:21-30.

Kallnik M, Elvert R, Ehrhardt N, Kissling D. *et al*. (2007) Impact of IVC housing on emotionality and fear learning in male C3HeB/FeJ and C57BL/6J mice. Mammalian Genome, 18: 173-186.

York JM, McDaniel AW, Blevins NA. *et al*. (2012) Individually ventilated cages cause chronic low-grade hypoxia impacting mice hematologically and behaviorally. Brain Behav Immun, 26: 951-958.

Casterlhano-Carlos MJ & Baumans V (2009) The impact of light, noise, cage cleaning and in-house transport on welfare and stress of laboratory rats. Lab Anim., 43:311-27.

The Experiment & Analysis of the Data

7 THE EXPERIMENT AND ANALYSIS OF THE DATA

This final section suggests precautions that can be built into an experimental design so as to prevent any subjective or systematic bias in the data. Overlapping with this, is the need to ensure that the experimental design is suitable for the planned statistical comparisons, including ways of dealing with problems that could arise from repeated testing of the same subjects. Finally, there is a warning that a specific change in behaviour could have many underlying causes and that it is essential to be cautious when interpreting the cause(s) and consequence(s) of any change in animal behaviour.

The only truly 'negative' result is one that emerges from an experiment that was flawed, either through its design or execution

7.1 The first step

Seek advice from a statistician about the design of the experiment: *i.e.,* **before carrying out the work**. This will ensure that there is a clear plan for the analysis of the data and that the results that emerge from the study are amenable to the planned statistical analysis. Part of this process is to establish that the number of animals to be used is adequate to reach a firm conclusion, but does not exceed that needed to reveal statistical significance. Essential elements of the design of experiments using animals and options for analysing the data are described in Festing *et al*., (2002) and Bate & Clark (in press).

7.2 Control group(s)

All experiments need at least one control group, which serves as a 'baseline' to help interpret the effects of the experimental factor(s). However, what counts as a control group (*e.g.* genetic wildtype, handled, vehicle-injected, sham-operated,) **depends on the purpose and context** of the experiment. Even an injection of saline, or anaesthesia, changes the behaviour of rodents in ways that can interact with, and even mask, the response to the test treatment. The influence of such effects can also depend on expertise of the researcher, who could even be considered as a variable factor in the experiment.

Sometimes, more than one control group is needed. For instance, it might be necessary to establish whether or not the stress of an injection has modified the behavioural response to the injection of a test drug. This can be done by including a procedural control, which is a group of subjects that has not experienced any experimental intervention (*e.g.* no vehicle injection: '**naïve control'**).

In general, repetition of experiments should be avoided. However, if it is uncertain that there will be a response to the experimental intervention (treatment A), it is essential to include a group that serves as a **positive control** (treatment B, which is similar to treatment A, but is known already to induce the behavioural response of interest). This will help to confirm that any lack of a response to treatment A is not explained by a fault in the design or execution of the experiment.

Do not assume that, because Treatment B has always/never produced a response in the past there is no need to include this group as a matter of routine: **contemporaneous control group(s) are always essential.**

Although increasing the number of groups to control for different experimental factors increases the number of animals used in the experiment, this approach enables more detailed interpretation of the findings and so could reduce the need for further experiments later on.

7.3 Basic practicalities that help to ensure a successful experiment

- Start with a relevant, **well-established protocol** and confirm that this can be replicated in your laboratory.
- Knowing the experimental group to which an animal has been assigned can influence an observers' score. To avoid this bias, ensure that the experiment is carried out **'blind':** *i.e.* the researcher is unaware of the treatment (*e.g.* genotype, vehicle/drug dose, surgery/anaesthesia) given to each of the animals, especially if the scoring is to be carried out manually (by observation).
- Look out for atypical results that seem to emerge only when **certain researchers** are carrying out the experiment. The cause of this anomaly might be untraceable but could be due to any number of factors, such as olfactory, visual, or aural stimuli that affect the animals' behaviour. In such cases, it will be necessary to decide who would be the best person to complete the work.
- **Check the consistency of the behavioural measures,** especially when scoring is done manually. Test this by ensuring that scores by individuals can be replicated by other members of the team and also that a given individual produces consistent scores. This will ensure that everyone involved in the experiment is using the same criteria: (*e.g.* the proportion of the animals' body has to enter an open arm of an elevated plus-maze to count as an 'entry'). A standard error of less than 5% of the mean score is normally regarded as the accepted criterion for 'reliable'.
- **Check the consistency of scores whether collected mechanically or electronically.** Unless the output is checked against a manual score, it is easy to be unaware that an automated device is counting additional (irrelevant) behaviours, or not counting important ones. This can happen because:
  - *the apparatus was not designed for the purpose for which it is being used;*
  - *the design does not exclude irrelevant movements;*
  - *the designers decided not to include them; or*
  - *the researchers are not clear about what is being measured.*
- Note that automated behavioural scoring should merely **aid the process of collecting data. It should not substitute for watching the animals.** Observing

the animals' behaviour during the experiment will provide assurance about their welfare and could reveal behaviours of potential scientific interest that would otherwise be missed.

- Consider all the factors that could influence the behaviour being measured, especially if the results differ from those obtained by other laboratories. In particular, watch out for **floor effects** (*i.e.* the behaviour is already close to the minimum before the experimental intervention) **or ceiling effects** (*i.e.* the behaviour is already close to the maximum before the experimental intervention) by ensuring that bidirectional changes in behaviour are possible.
- Consider the possibility that **the animal might become bored** during prolonged or repeated procedures and take account of any consequences of that change in mental state on the behaviour of interest.
- To ensure valid evaluation of changes in behaviour after an experimental intervention, ensure that **the behaviour of interest is the same at baseline (pre-intervention) across all groups**.
- **Validate routinely the consistency** of any experimental intervention (*e.g.* extent and location of surgery/lesion) that is intended to affect the behaviour of interest.
- **Do not assume that the behaviour of interest is the only one to be affected by the experimental intervention**. Monitor as many aspects of the behaviour as possible to ensure that other changes are not confounding the results. For example, in the Open Field, a reduction in locomotor activity can arise because there is an increase in another, mutually exclusive behaviour: *e.g.* grooming or rearing.
- Bear in mind that some procedures, which are apparently extremely simple and straightforward, produce **results that are ambiguous and difficult to interpret.** Again, the Open Field is a good example because a change in locomotor activity could be attributed to a change in motor behaviour, anxiety-like behaviour (explosive running or freezing), sedation, muscle relaxation.....
- If it is uncertain what counts as biologically significant behaviour, **ask an expert** in the field.
- **Be flexible:** if the original hypothesis is not upheld, it will be necessary to abandon or redesign the project – but that might take you to more interesting scientific territory.

Many ground-breaking findings have emerged from experiments that did not produce the predicted results

7.4 The sequence of test treatments and controls

Normally, all fixed factors in the experiment *(i.e.* the factors that are predetermined, such as genotype or drug dose) must be **fully randomised across all the subjects.** This is necessary to ensure that there is no systematic bias in the results arising from, for instance, a change in the researcher, animal husbandry, batch of test drug....

There are some circumstances in which a **different approach might be adopted,** especially if it is not possible to test all the subjects simultaneously. For instance, when comparing the effects of a range of drug doses on the behaviour of a mutant mouse and its wildtype, it might be helpful to evaluate the effects of a single treatment on the behaviour of pairs of the two genotypes, simultaneously. This will reduce the influence of nuisance factors when comparing the two genotypes. The other experimental factors (*e.g.* drug dose) should then be either randomised or counter-balanced according to a predetermined sequence, with the different treatments being given consecutively.

***Ad hoc* selection** of animals from their home-cage for each test treatment will not produce a randomisation assignment to different treatment groups. This is because mice displaying high avoidance behaviour are more likely to escape handling by the researcher, whereas others (with low avoidance behaviour) will be caught more easily.

There are several **approaches to randomisation.**

- Use the RAND function in Excel

  (see: http://spreadsheets.about.com)
- Assign each mouse a number as it appears in sequence in the number, π (3.142....)
- Use the random number sequence generator to be found at www.random.org/sequences/

7.5 Retesting the animals

When studying the behaviour of animals over a prolonged period, ensure that the **method used to identify** individual subjects will last long enough (*i.e.* do not use marker-pens for long-term experiments). Also, consider (and check, if necessary) whether the chosen identification procedure could affect behaviour (*e.g.* distal phalanx removal).

7.5.1 Retesting in the same behavioural procedure

- This approach can produce **high quality data** because each subject acts as its own control. Repeated testing of the same animals' behavioural responses to different experimental challenges has the added advantage of reducing the number of animals needed for the study.

- Bear in mind that **habituation** or **sensitization** to the test environment (see: Section 7.4) and **learning** can modify the animals' behavioural response on repeated experience of the procedure. The inclusion of a group of naïve (untreated) subjects is essential to control for these changes.
- Taking a series of measurements from the same animal, or increasing the duration of a study, could increase the **cumulative 'harm'.** Justification for this approach has to balance the harm to the animals against the increase in the number of animals that would be needed if each was tested only once.
- If the animals have been trained or tested under one set of conditions (*e.g.* whilst under the influence of a test drug) it is important to establish whether **state-dependence** has developed: *i.e.* animals' performance is maintained only when the drug is present, regardless of its pharmacological effects. This can be ascertained by counterbalancing control and test groups during the training and testing phases of the experiment (see: Table 5). State-dependence can develop with any environmental stimulus, not just drugs. For instance, in humans the presence and type of background music affects recall in a memory task (Balch *et al.*, 1993).

Table 5: Controlling for state-dependency

Subject Group	Training	Testing
1	Vehicle	Vehicle
2	Vehicle	Drug
3	Drug	Vehicle
4	Drug	Drug

7.5.2 Testing animals in a battery of procedures

If the animals are to be tested in a series of different procedures, it is important to consider whether this could increase cumulative harm (see: Section 2). Bear in mind that this might not always be the case because some procedures might not be distressing or harmful in any way. For instance, some involve rewarding the animals, or placing them in an environment that they find more interesting than their home cage.

There could also be **carry-over effects** from one procedure to another that cause long-lasting changes in the animals' behaviour. If this is the case, then the sequence of tests will influence the results.

It is generally agreed that it is best to carry out the test that causes the **least discomfort first** in order to minimise the influence of any carry-over effects on behaviour. However, if all the procedures are regarded as equal severity, a counterbalanced, blocked design might be more appropriate.

7.6 When to test the animals

Procedural variables can affect animals' behaviour and influence their expectations - especially those that depend on time of day: *e.g.* the regime for cage-cleaning and general husbandry.

These can be especially problematic when the **clock is adjusted** for daylight-saving, or during public holidays and weekends, when there are changes in the routine for feeding and cleaning, as well as general background activity.

If it is decided to carry out the measurements at a **fixed time** each day, it could be helpful to carry out a pilot experiment to check whether the behaviour is the same when collected at a different time of day. If not, then **time-of-day should be regarded as another experimental fixed factor.**

Such a difference could be explained by any of a number of factors, such as:

- The time since the animals last ate, which will be especially important if the behaviour of interest is driven by appetitive reward.
- Sleep deprivation (if the procedure is carried out in the light phase).
- Circadian fluctuations in behaviour (see below).

7.7 Reversed lighting schedule for rats and mice?

Changes in behaviour that depend on **time of day** are well-documented: notably, locomotor activity. This experimental factor is especially important because rodents are nocturnal and so are most active during the human dark-phase. For this reason, some laboratories monitor the behaviour of rodents that have been housed on a **reversed lighting schedule.** Measurements are taken during their dark phase (the human light phase), when they are most aroused.

This approach has the additional advantage of preventing the animals from becoming **sleep-deprived**, which can affect their behaviour. Nevertheless, it can be difficult, if not impossible, to ensure that the animals' exposure to the reversed light / dark cycle is maintained throughout all stages of the experiment. Even a brief pulse of light (c 3 min) during the animals' dark phase will disrupt their circadian rhythms.

In any case, **extraneous environmental noise** during the human day (the animals' subjective night), such as heating or ventilation systems switching on/off or staff arriving at, or leaving, the unit can still act as a *zeitgeber* and disrupt the experiment.

Another problem is that a reversed lighting schedule requires the experimental procedure, observation and recording of the animals' behaviour to be carried out in **red light.** Whereas some strains of mice are blind to short-wavelength red light, there are photoreceptors that respond to red light of longer wavelengths. These activate neuronal pathways, which project to brain regions that influence behaviour (Delwig *et al.*, 2012).

Finally, consider the possibility that an experimental procedure that could disrupt animals' circadian rhythms, could also cause secondary changes in behaviour that affect the controls/test groups of animals in different ways.

7.8 Statistical analysis of the data

Advice on the statistical tests that should be used to analyse data from experiments of different designs is beyond the scope of these Guidelines but there are many textbooks on this subject, some of which are listed at the end of this section.

There are also several **software packages** that offer a wide range of statistical tests. A comparison of the functionality of several statistical software packages has been published recently (Clark *et al.*, 2012)**.**

7.9 Forming conclusions: some final precautions

Check that other useful findings have been considered too - or whether the conclusions are based only the results of direct relevance to the hypothesis.

Appraise the evidence to **ensure that ALL the behavioural changes support the hypothesis**, and not merely those of interest.

Consider the possibility that a change in the behaviour that is of greatest interest is a **secondary consequence of the experimental intervention**. If so, this process might need further experimental investigation. Some well-known examples include:

- The genetic mutation has **physiological consequences** that indirectly disrupts behaviour (*e.g.* impairs motor function or memory).
- Some procedures can cause **pain, amnesia or sedation**. All these actions will affect other behaviours, especially those that involve movement. Such factors need to be eliminated as confounders by incorporating appropriate controls into the experimental design.
- Drugs that have been administered systemically will affect peripheral systems first (*e.g.* by causing a change in blood pressure). Check whether a peripheral response could contribute to the behavioural response.
- Some aspects of animals' behaviour are **mutually exclusive**, such as locomotor activity and vegetative behaviours (*e.g.* grooming). An increase in one of these behaviours could be an indirect consequence of a reduction in the other. (See also: Section 7.4).
- Animals' motivation to respond in operant procedures that are shaped by gustatory reward will be affected by drug-induced changes in appetite (as in the use of food or sucrose solution in procedures for evaluating cognitive performance).
- Some forms of drug **vehicle** can **affect behaviour** (*e.g.* polyethylene glycol).

It is also important to consider whether the magnitude of the response is large enough to have an appreciable effect in humans? If not, is the finding likely to have translational relevance? Consider using statistical ('**effect size'**) analyses that provide an objective evaluation of whether the magnitude of the response is large enough to be 'real'.

But, bear in mind that the importance of a response of given magnitude depends on context. For instance, a reduction of 5% in a behavioural measure such as food intake might be **biologically interesting but is functionally negligible**. By contrast, in terms of accuracy or attention whilst driving, a reduction of this magnitude would be worrying.

Do not assume that a **parallel change** in two measurements implies that they are causally linked. It is possible that they are both induced, independently, by a third, common factor. A statistically significant correlation between two measures offers stronger evidence for (but still not proof of) a causal link. However, this type of analysis needs a large number of data points (>50), which would be hard to justify. Small samples should not be used in this way because individual (extreme) points distort the analysis, and suggest a statistically significant correlation when none exists. (See: Salmon and Stanford 1992; Stanford and Salmon, 1993;

Finally, consider whether anything useful can be learned from the **results that do not support the hypothesis.** Although often regarded as 'negative' results, these findings still enhance the scientific knowledge-base and merit publication. This is not least because it is important to prevent others from repeating an experiment that has turned out to be less interesting than expected. But failure to find an outcome predicted from an established hypothesis challenges that hypothesis and can inspire a new one. That is what science is about!

Beware of an interpretation of behavioural changes that relies on anthropomorphic assumptions

References

Balch WR, Bowman K, Mohler L. (1992) Music-dependent memory in immediate and delayed word recall. Mem Cognit. 20:21-8.

Clark RA, Shoaib M, Hewitt KN, Stanford SC, Bate ST. (2012) A comparison of InVivoStat with other statistical software packages for analysis of data generated from animal experiments. J Psychopharmacol. 26:1136-42.

Delwig A, Logan AM, Copenhagen DR, Ahn AH. (2012) Light evokes melanopsin-dependent vocalization and neural activation associated with aversive experience in neonatal mice. PLoS One. 7:e43787. doi: 10.1371/journal.pone.0043787. Epub 2012 Sep 13.

Festing M, Overend P, Gaines Das R, Cortina Borja M, Berdoy M. (2002) The Design of Animal Experiments: Reducing the use of animals through better experimental design. Laboratory Animal Handbook Number 14. Pub. The Royal Society of Medecine Press.

Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. (2010) Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol. 2010 Jun 29;8(6):e1000412.

Salmon P & Stanford SC. (1992) Critique: Research strategies for decoding the neurochemical basis of resistance to stress. J Psychopharmacol. 6:1-7.

Stanford SC & Salmon P. (1993) Decoding the neurochemical basis of resistance to stress: a reply. J Psychopharmacol. 7:89-90.

Further reading:

Howell DC (2008) Fundamental Statistics for the Behavioral Sciences. 6e. Thomson Wadsworth
Bate ST and Clarke RA. (2013) The Design and Statistical Analysis of Animal Experiments. Cambridge University Press (in press)

Tabachnick BG & Fidell LS. (2007) Using Multivariate Statistics, 3e. Pearson Education.

Steering Panel

(Responsible for assembly, production and editing of the Guidelines)

Anne-Marie Farmer*
A(SP)A Inspector
Home Office observer (UK)

Alma Gower
Affiliated to ESSWAP

Joop de Graaf
ESSWAP Foundation
(Courses in Whole Animal Pharmacology)

Maggy Jennings*
Royal Society for the Prevention of Cruelty to Animals (RSPCA)

Hugh Marston
TPP Global Development

Paula Moran
University of Nottingham

John Rodgers
University of Leeds

Clare Stanford*
University College London

LASA ETES
Education, Training and Ethics Section (Members)*

These Guidelines have been approved by the Councils of LASA, BAP, BNA and ESSWAP, on behalf of their members.

We thank the following people for contributing points for inclusion in the Guidelines but wish to point out that the list of names does not imply that any of these individuals endorse fully all elements of their contents.

David Bannerman (*University of Oxford*)
Duncan Banks (*Open University/BNA*)
Manuel Berdoy* (University of Oxford)
Verity Brown (*University of St Andrews*)
Steve Brown (*MRC, Harwell)*
James Bussell (*Sanger Institute*)
Vicky Chapman (*University of Nottingham*)
Peter Clifton (*University of Sussex*)
Robert Deacon (*University of Oxford*)
Adrian Deeny (*University College London*)
Dawn Eagle (*University of Cambridge*)
Fran Ebling (*University of Nottingham*)
Cathy Fernandes (*Institute of Psychiatry, London*)
Kevin Fone (*University of Nottingham*)
Paul Gard (*University of Brighton*)
Maggie Green (*University of Cambridge*)
Ashley Grimmé (*PhD student, UCL*)
Jim Hagan (*Kings College, University of London*)
Penny Hawkins (*RSPCA*)
Katie Hewitt (*University of Oxford*)
Robert Hubrecht (*UFAW*)
Trevor Humby (*University of Cardiff*)
Anne Jackson (*University of Brighton*)
Angela Kerton* (*Imperial College, London*)
John Kelly (*NUI, Galway*)
Emma Lane (*University of Cardiff)*
Darren Logan (*Sanger Institute)*
Beverley Law* (*University of Leeds*)
Wendy Matthews (*AWERB (lay member) UCL*)
John McCann (*AWERB (lay member) MRC/ University of Cambridge*)
Ken Meecham (*Huntingdon Life Sciences*)
Paul Mitchell (*University of Bath*)
Paul Moser *(Laboratoires Pierre Fabre Ltd)*
Judy Pratt (*University of Stirling*)
Will Redfern (*AstraZeneca, UK*)
Trevor Robbins (*University of Cambridge*)
Sally Robinson (*AstraZeneca, UK*)
Jennifer Salisbury (*Sanger Institute*)
Patrick Sinnett-Smith* (*Pfizer UK*)
Jacqui White (*Sanger Institute*)
Clea Warburton (*University of Bristol*)
Sara Wells (MRC, Harwell)
Jacqui White (*Sanger Institute*)
Lucy Whitfield* (*Royal Veterinary College, London*)

*** Members of LASA Education, Training and Ethics Section**

It is the Steering Panel's intention that these Guiding Principles will be reviewed and revised periodically. If you have any comments or suggestions to make please e-mail info@lasa.co.uk with "Guiding Principles for Behavioural LAS" in the message subject.