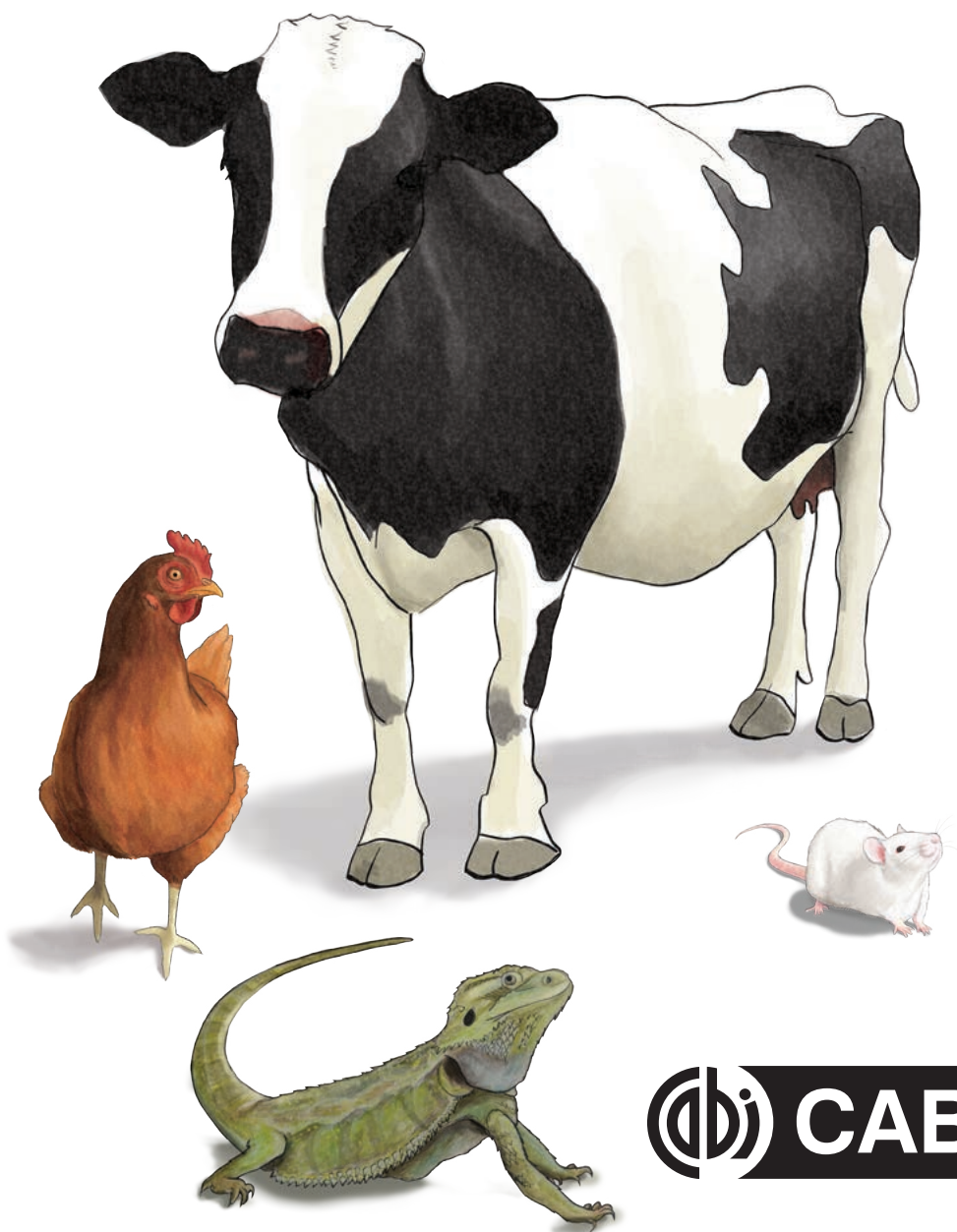


ASKING ANIMALS

AN INTRODUCTION TO ANIMAL BEHAVIOUR TESTING

Birte L. Nielsen



Asking Animals

An Introduction to Animal Behaviour Testing

Asking Animals

An Introduction to Animal Behaviour Testing

Birte L. Nielsen

*UMR MoSAR, INRAE
France*

Illustrations by Elinor L. Friggens



CABI is a trading name of CAB International

CABI

Nosworthy Way
Wallingford
Oxfordshire OX10 8DE
UK
Tel: +44 (0)1491 832111
Fax: +44 (0)1491 833508
E-mail: info@cabi.org
Website: www.cabi.org

CABI

WeWork, One Lincoln St
24th Floor
Boston, MA 02111
USA
T: +1 (617)682-9015
E-mail: cabi-nao@cabi.org

© Birte L. Nielsen 2020. All rights reserved. No part of this publication may be reproduced in any form or by any means, electronically, mechanically, by photocopying, recording or otherwise, without the prior permission of the copyright owners.

References to Internet websites (URLs) were accurate at the time of writing.

A catalogue record for this book is available from the British Library, London, UK.

Library of Congress Cataloging-in-Publication Data

Names: Nielsen, Birte Lindstrøm, author.

Title: Asking animals : an introduction to animal behaviour testing / Birte

L. Nielsen, UMR MoSAR, INRAE, France.

Description: Wallingford, Oxfordshire, UK ; Boston : CABI, [2020] |

Includes bibliographical references and index. | Summary: "Contemporary, thought-provoking yet utterly practical, this book gives an introduction to the use and misuse of behaviour tests applied to animals. By including illustrative examples from a variety of species, the book aims to inspire the animal scientist to think about what a given behavioural test can be used for and how the results can be interpreted.

It is valuable to students, established researchers, teachers and practitioners of applied ethology, animal welfare science, and veterinary science"— Provided by publisher.

Identifiers: LCCN 2019042207 (print) | LCCN 2019042208 (ebook) |

ISBN 9781789240603 (hardback) | ISBN 9781789240610 (paperback) |

ISBN 9781789240627 (ebook) | ISBN 9781789240634 (epub)

Subjects: LCSH: Animal behavior—Testing.

Classification: LCC QL751 .N526 2020 (print) | LCC QL751 (ebook) | DDC 591.5—dc23

LC record available at <https://lcn.loc.gov/2019042207>

LC ebook record available at <https://lcn.loc.gov/2019042208>

ISBN-13: 9781789240603 (hardback)

9781789240610 (paperback)

9781789240627 (ePDF)

9781789240634 (ePub)

Commissioning Editor: Caroline Makepeace

Editorial Assistant: Emma McCann

Production Editor: Tim Kapp

Typeset by Exeter Premedia Services Pvt Ltd, Chennai, India

Printed and bound in the UK by Severn, Gloucester

Contents

Preface	vii
Acknowledgements	ix
Part I Setting the Scene	
1 Five Things This Book is Not	1
2 Non-test Observations	7
3 How to Choose a Test	19
Part II Types of Tests	
4 Tests to Characterize the Animal	29
5 Choice, Preference and Motivation	50
6 Ability to Detect and Distinguish	65
7 Effects of Age and Treatment	78
8 Reinforcement and Punishment	93
9 Learning Capacity, Memory and Cognitive Ability	107
10 Genetic Components of Behaviour	124
Part III Additional Aspects	
11 Other Test Considerations	137
12 Legislation, Guidelines and Ethical Considerations	154
13 Future Methodologies and Technological Advances	166
Index	173

Preface

Most of us have heard about Pavlov's dogs and the Skinner box, and many are aware that applied behavioural science has come a long way since then. Behavioural testing of animals is used in many different scientific disciplines, from laboratory-based studies in neuroscience to fieldwork in behavioural ecology. But why do we use the tests we do? What can they tell us and – not least – what are their limitations?

This book will give an introduction to the use (and perhaps misuse) of behaviour tests applied to animals. Through illustrative examples from a variety of species, the aim is to inspire the animal experimenter to think about what a given behavioural test can be used for and how the results can be interpreted. It is *not* meant as a dictionary or list of tests from which a researcher can choose, but as an inspiration on what to do (and not to do) when developing and executing tests of animal behaviour.

I could have chosen to delve into the history of behavioural experimentation with a detailed presentation and discussion of the tests most commonly used. Instead, I have opted for a lighter tone (and tome), hoping that you may actually read it to the end. This has, of course, some drawbacks. There will be omissions and the purists among you may scoff at some of the simplifications used to describe the hows and whys in the different chapters. However, I believe this to be justified if it makes more people read about this subject and, perhaps, as a consequence, develop an interest in the practical use of behavioural testing to ask animals questions.

Birte L. Nielsen
September 2019

Acknowledgements

Writing a textbook has been very interesting, highly educational for the author and great fun. But it can be somewhat frustrating at times, and without the encouragement of people around me, this undertaking would never have come to fruition.

Friends and colleagues, both from my former and my current research unit, have provided great support during the writing process. I would also like to thank Ophélie Dhumez, Turid Burvik, Cecilie M. Mejdell, Alexandra Courty, Keelin O'Driscoll, Margit Bak Jensen, Lene Munksgaard, Lene Juul Pedersen and Maria V. Rørvang for kindly allowing me to use their photos. The wonderful drawings for many of the figures were created by Elinor L. Friggens, for which I am extremely grateful.

A special thanks goes to Jes Lynne Harfeld and Janne W. Christensen, for reading through earlier versions of different chapters. I would also like to thank Justin Varholick from the University of Bern for guiding me to the article by David Lahti, and Jeremy Marchant-Forde for information on the US regulations for animal use.

Writing a textbook about animal behaviour testing makes you appreciate human social interactions even more. Thank you to all who believed this could be possible, especially Caroline Makepeace and Tim Kapp from CABI, who were always there with advice and reassurance. A heartfelt thank you and several bear hugs also go to Nina, Elinor and Nic for their moral support and continued encouragement.

Five Things This Book is Not

You may have looked at your dog, horse or goldfish, and wondered what they were thinking. Does Rover like his new dog-house? Is my pony feeling cold in this weather? Wouldn't it be nice if we could just ask them? Well, as the title of this book indicates, we can. The concept of using well-designed behavioural tests of animals as a way of asking them questions has been known about and used for a long time. Initially, behavioural tests were mainly carried out on laboratory rodents. Konrad Lorenz's demonstrations of imprinting in greylag geese in the 1930s were a form of behavioural testing, but it wasn't until later that domestic livestock species were included: Hughes and Black (1973) and Dawkins (1977) were among the first to apply behavioural tests to farm animals, in their case the domestic hen, by studying the responses of the birds in behavioural tests of preference for cage size and floor type. Since then, a plethora of tests have been developed to ask animals questions by monitoring their behaviour in different situations.

The subject of behavioural testing of animals is complex, rich and potentially controversial (see Chapter 12). And in an era where almost everything can be found online, do we really need another book on this subject? Yes, because many of the existing books are quite specialized in their approach or do not give much practical advice. These books, together with those on animal behaviour in general in which various behavioural tests are inevitably mentioned, tend to be focused on specific groupings of animals. This includes laboratory rodents and primates (Whishaw and Kolb, 2005; Crawley, 2007; Buccafusco, 2009), domestic animals (Fraser, 2010; Broom and Fraser, 2015; Jensen, 2017; Appleby *et al.*, 2018) or wildlife species (Manning and Dawkins, 2012). These textbooks rarely dwell on the experimental test design, and – because it is not the purpose of these books – do not always consider the pros and cons of a given testing paradigm.

Having worked for most of my career in applied ethology of farm animal species and their welfare, I have also spent 9 years in a neuroscience

research unit, carrying out behavioural experiments on olfactory responses of rodents, mainly rats. This has provided me with the privilege of seeing two very different sides of animal behaviour testing, and made me realize how rarely methods and behavioural knowledge are transferred between scientific disciplines. This book is an attempt to start bridging that gap.

Before you delve into the different chapters of this book, here is some important information to prevent confusion and disappointment, and to put you in the right frame of mind to make the most of the next 170 pages. You should be aware of the following:

This Book is Not Complete

It almost goes without saying that this book is but a snapshot of some of the tests developed to study animal behaviour. Each chapter heading could be a book in itself, and not all behavioural tests are included, nor are they described in depth. Space restriction is among the reasons why the book is not even trying to be more exhaustive. In order to have enough space to include a broad variety of behavioural tests, it has been necessary to exclude some tests to allow a more in-depth description and discussion of others. As happened to me when researching this book, this is likely to introduce you to test types or formats that you have not come across before. This, in turn, may provide new inspiration for your own scientific work, not only as a student but perhaps also as an experienced silverback in applied ethology. If you want to know more about specific tests, there are other more dedicated textbooks (e.g. Lehner, 1998; Wyatt, 2014). There are also fascinating articles describing how knowledge is obtained from animal research in terms of reproducibility of results and the limitations of our chosen model (e.g. Garner *et al.*, 2017). Finally, Bueno-Guerra and Amici (2018) cover field and laboratory methods in animal cognition of more exotic species, including tortoises, sharks and bats.

This Book is Not Representative

Unlike a review article, the chapters are telling a story about different types of tests, and yes, cherry-picking has occurred. This has been done intentionally to introduce the reader to some of the more interesting examples of animal behaviour testing within each category. The book also does not give the history of animal behaviour testing, nor the origin of most of the tests mentioned, as many tests have already been refined and further developed since their first use. Descriptions of earlier incarnations of a given behavioural test are therefore only included if they are relevant for the understanding of the tests in question. Having worked for many years in olfaction, this sensory modality tends to crop up more often than it should by chance in this book, and I apologize in advance for this slight selection

bias. However, if it piques your interest in the importance of olfaction for animal behaviour and welfare, I can (humbly) recommend a book written by distinguished colleagues in the field and edited by me (Nielsen, 2017).

This Book is Not About One Scientific Discipline

Although the main scientific discipline of animal behaviour testing is applied ethology, the subject does embrace a number of scientific disciplines, such as neuroscience, behavioural ecology and animal behaviour science in general, as well as genetics and nutrition, just to mention a few. This has also made it possible to cover a wide range of species (but see below), and I have found myself marvelling at tests done in animal models largely unknown to me, such as zebrafish and chimpanzees. I hope that by including examples from species not usually seen in the neuroscience or pharmacological labs, such as dairy goats and laying hens, this book can evoke the same sense of discovery that I experienced when researching it. The importance of this is more wide-ranging: when reading about the same type of behavioural test carried out, say, in mice by neuroscientists and in piglets by animal welfare scientists, it becomes clear that the approach to the test is very different. This is perhaps not surprising, because the goals of the study, and the scientific questions asked, are very different. It is, however, something we should all be aware of when using results arising from different scientific studies and disciplines.

This Book is Not About Insects

Apologies to all the insect aficionados out there, but there is a – hopefully good – reason to exclude them: I wanted to put the emphasis on sentient species and animals managed by humans, especially those covered by legislation on the use of animals for scientific purposes, such as the EU Directive (2010). The main species you will come across in the following chapters are therefore vertebrates. However, I cannot exclude the possibility that a single bee trial may have sneaked in without my noticing. If you are interested in insects, and specifically the neuroscience aspects of their behaviour, you may find the book by Menini (2009) of interest.

This Book is Not About Statistics

It would have been relevant and useful to have a section on statistical analysis of results from various behaviour tests and how to interpret the statistical results correctly. Researchers are sometimes unclear about what the replicate in their study is (e.g. individual or group), and what to do if the residuals of their analyses are not normally distributed. But, alas, I am no statistician. It is, nevertheless, an extremely important aspect of animal

behaviour testing, and assistance should be sought from statistical experts in the field (Kaps and Lamberson, 2017). The first place to look for guidance in this specific area of biology is Martin and Bateson (2007), a text book that focuses on statistical issues when analysing behavioural data. In Chapter 11, different test considerations to take into account (or at least be aware of) are discussed, as a lot of statistical grief can be prevented by careful planning.

So the structure of this book is as follows: In the first chapters, I try to set the scene, describing how non-test observations provide information that is often the basis on which many behavioural tests rest. This leads on to a chapter on how to choose a test, both in theory but also very much in terms of practical considerations. The core of the book, Chapters 4–10, covers the main types of behavioural testing themes, such as tests characterizing an individual, standard tests of treatment effects, choice and preference tests, and ways to assess learning ability, as well as genetic aspects of behaviour. Each chapter covers only some of the available tests within each theme, and for each test type, I have chosen one or two examples from the current literature to illustrate the practical use of the test in question (Fig. 1.1). These examples are meant to demonstrate the breadth as well as the limitations of the tests, while covering a variety of species. The

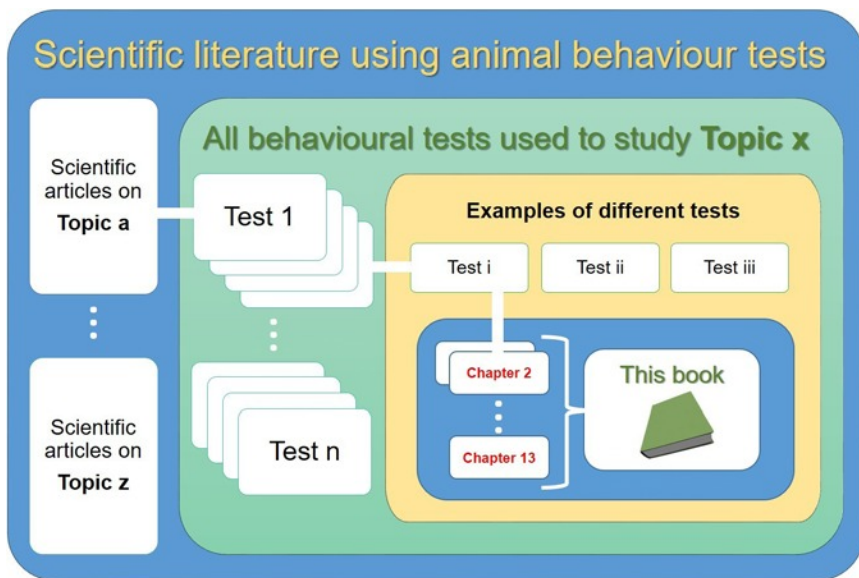


Fig. 1.1. Each core chapter is based on a few select examples of the practical use of a limited set of tests within the chapter topic. These test examples have been chosen so as to cover a variety of vertebrate species across the whole book, as well as to highlight specific details in the tests and methods used.

examples are often also included because they were the most interesting and fun to read.

I only have to glance at all the half-read ... aargh, who am I kidding? – unread textbooks on my shelves to realize that although we may have the best intentions to read up on, say, the behaviour of cattle or the neurobiology of olfaction, when push comes to shove, there are only so many hours in the day. Most people working in science are already struggling to keep up-to-date with the scientific papers in their subject area. How should they find the time to read whole books, in particular one that deals with more methodological aspects and spans several scientific disciplines? The only chance that anybody (other than the technical editor) will read this book, is if I make it as easy to read as possible. I have therefore endeavoured to the best of my ability to write a relatively short book, which includes the more interesting examples of animal behaviour testing, written in language that is easily digestible and printed in a format that can be read while you are lying down. I hope I have succeeded.

References

- Appleby, M.C., Olsson, A. and Galindo, F. (eds) (2018) *Animal Welfare*, 3rd edn. CAB International, Wallingford, UK.
- Broom, D.M. and Fraser, A.F. (eds) (2015) *Domestic Animal Behaviour and Welfare*, 5th edn. CAB International, Wallingford, UK.
- Buccafusco, J.J. (2009) *Methods of Behavior Analysis in Neuroscience*. CRC Press, Taylor & Francis Group, Boca Raton, Florida.
- Bueno-Guerra, N. and Amici, F. (eds) (2018) *Field and Laboratory Methods in Animal Cognition – A Comparative Guide*. Cambridge University Press, Cambridge, UK.
- Crawley, J.N. (2007) *What's Wrong with My Mouse? Behavioral Phenotyping of Transgenic and Knockout Mice*. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Dawkins, M. (1977) Do hens suffer in battery cages? environmental preferences and welfare. *Animal Behaviour* 25, 1034–1046. DOI: 10.1016/0003-3472(77)90054-9.
- EU Directive (2010) Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Official Journal of the European Union* 20.10.2010, L276, 33–79. Available at: <http://data.europa.eu/eli/dir/2010/63/oj>
- Fraser, A.F. (2010) *Behaviour and Welfare of the Horse*, 2nd edn. CAB International, Wallingford, UK.
- Garner, J.P., Gaskill, B.N., Weber, E.M., Ahloy-Dallaire, J. and Pritchett-Corning, K.R. (2017) Introducing Therioepistemology: the study of how knowledge is gained from animal research. *Lab Animal* 46(4), 103–113. DOI: 10.1038/labon.1224.
- Hughes, B.O. and Black, A.J. (1973) The preference of domestic hens for different types of battery cage floor. *British Poultry Science* 14(6), 615–619. DOI: 10.1080/00071667308416071.
- Jensen, P. (2017) *The Ethology of Domestic Animals – An Introductory Text*. CAB International, Wallingford, UK.

- Kaps, M. and Lamberson, W.R. (2017) *Biostatistics for Animal Science*. CAB International, Wallingford, UK.
- Lehner, P.N. (1998) *Handbook of Ethological Methods*. Cambridge University Press, Cambridge, UK.
- Manning, A. and Dawkins, M.S. (2012) *An Introduction to Animal Behaviour*, 6th edn. Cambridge University Press, Cambridge, UK.
- Martin, P. and Bateson, P. (2007) *Measuring Behaviour – An Introductory Guide*. Cambridge University Press, Cambridge, UK.
- Menini, A. (2009) *The Neurobiology of Olfaction*. CRC Press, Taylor & Francis Group, Boca Raton, Florida.
- Nielsen, B.L. (ed.) (2017) *Olfaction in Animal Behaviour and Welfare*. CAB International, Wallingford, UK.
- Whishaw, I.Q. and Kolb, B. (eds) (2005) *The Behavior of the Laboratory Rat – A Handbook with Tests*. Oxford University Press, Inc., New York.
- Wyatt, T.D. (2014) *Pheromones and Animal Behaviour: Chemical Signals and Signatures*, 2nd edn. Cambridge University Press, Cambridge, UK.

Non-test Observations

For many people over the age of 40, their first encounter with the scientific study of animal behaviour was when Sigourney Weaver played the role of Dian Fossey in the film *Gorillas in the Mist*, released in 1988. Although this book is about behavioural tests, knowledge of the behaviour of animals is largely based on observations of a given species in its natural surroundings. For many ethologists, the study of animal behaviour thus consists of hours and hours (and hours...) of field work, where the species studied is observed in its natural environment. These data form the basis of the so-called ethograms used also in applied ethology, where the complete behavioural repertoire of the species studied is listed and described in a mutually exclusive way according to the posture and activity of the animal within a given environment.

The behaviour of numerous species of animals has been studied in natural settings, and that includes a variety of domestic species. An example of this is the Edinburgh Pig Park, where the behaviour and interactions of domestic pigs were observed while they were kept in a large (2.3 ha) enclosure with varied topography and vegetation (woodland, streams and pasture) in Scotland (Newberry and Wood-Gush, 1985, 1986; Stolba and Wood-Gush, 1989). This was the first study to demonstrate that, despite having been domesticated for millennia, individual sows engaged in nest-building prior to farrowing. These individual sows had never previously experienced the outdoors and, in this case, never had access to material with which to nest-build, such as straw. Nevertheless, the sows began to construct intricate nests of branches and greenery for farrowing, in a way similar to that seen in sows of the wild boar.

Results from studies such as the one mentioned above, should – at least in theory – enable us to take into account the physiological and behavioural needs in the housing of animals managed by humans. Physiological needs include access to food and water, and examples of behavioural needs could be access to perches in birds (Olsson and Keeling, 2002), nest-building materials in sows and rats (Arey *et al.*, 1991; Patterson-Kane, 2004), and

swimming water for ducks and mink (Rodenburg *et al.*, 2005; Kornum *et al.*, 2017). The housing environment is nevertheless likely to be limiting in some form or other, not least because of space constraints relative to free range living. Having said this, some positive aspects of housing animals exist, such as protection from adverse climate conditions and protection from natural predators. In the following, examples of behavioural studies in a housed setting are given, emphasizing the non-test situation and what we can learn from this for use in behavioural tests.

Time Budgets and Behavioural Development Over Time

To know what has changed you need to know what is normal. One way to measure this is to observe the animal in the environment in which it is kept, and quantify the occurrence and duration of different behaviours. This may range from continuous observations (often done via video recordings) of the complete behavioural repertoire of the animal based on their ethogram, to registrations of a subset of these behaviours, such as whether the animal is active or passive. Estimates of the time budget of animals kept in groups may be based on scan-sampling of the group at regular intervals. Depending on the species and the behaviour of interest, this could consist of observations every 5 min, where the number of animals in the process of doing predetermined behaviours is counted (for more details of this and other observation methods, please see Martin and Bateson, 2007).

When behaviour is scored by an observer from a video recording, it is essential that the individual animals can be easily identified. Using a marker pen, rodents may be given different combinations of stripes and dots on their tails, and cattle can have numbers dyed or bleached on to their coats. The spray-marking of moving objects is rather difficult, and numbers can be difficult to see on videos unless they are put on all sides of the animals. One method to identify individual pigs on video recordings, is to use a coding system of stripes, which are both quicker and easier to apply than digits (Nielsen, 1995). This system was developed from the (now abandoned) ear notching system used to identify pigs before the advent of ear tags. In this modified marking system, each number can be expressed through combinations of stripes on the rear, middle and/or shoulder of the pig. Each stripe on the rear represents the value 1, each stripe on the middle represents the value 9 and each stripe on the shoulder represents the value 3. A pig allocated the number 7 would therefore have two spray lines across the shoulder and one across the rear. [Figure 2.1](#) shows three pigs marked according to this spraying system. Each stripe is visible on the top and on both sides of the animal, even when the pig is lying on its side, allowing easy identification on the video recordings. This system covers all integers up to 26 if a maximum of two stripes are used on each body section.

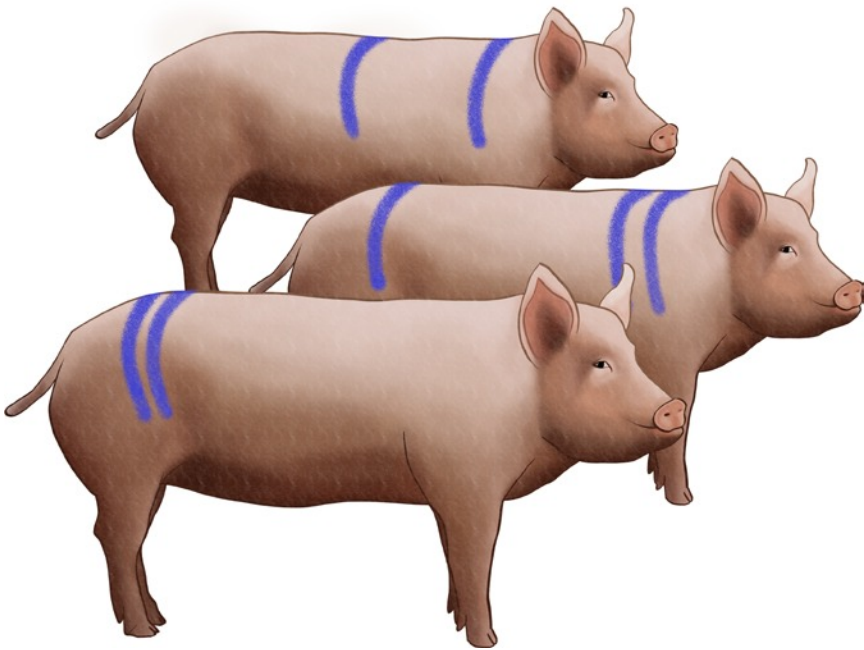


Fig. 2.1. Method used for spray-marking individuals prior to video recording. These pigs, in order from front to back, have the numbers 2, 7 and 12 (see text for details or try to guess the system).

Animals are often observed only during specific periods, either due to time constraints for the observer or due to low visibility during the evening and night. This is specifically problematic when working with animals, such as rodents, that are nocturnally active. For this reason, rodent houses can have an inverse lighting schedule, allowing research to be carried out within normal working hours while studying the animals during their active time. However, many labs do not consistently employ such lighting, often because it involves performing cage cleaning and behavioural observations under red light, which is straining for the human eye. Another method is available, as demonstrated by McLennan and Taylor-Jeffs (2004): low-pressure sodium bulbs provide sufficient light for humans to see to read and write, but this type of light has a very narrow wavelength (589 nm) in which rodents are unable to see (Fig. 2.2). The animals thus behave as if it were dark, allowing the researchers and caretakers to perform maintenance activities and observe behaviour during the naturally active phase of the animals.

Once we have our time budget measured on healthy individuals, we know what to expect as normal – in the broadest sense of the term. By extension, this can be used to detect abnormalities, such as leg and hoof problems in cattle. Even moderate lameness in dairy cows can lead to

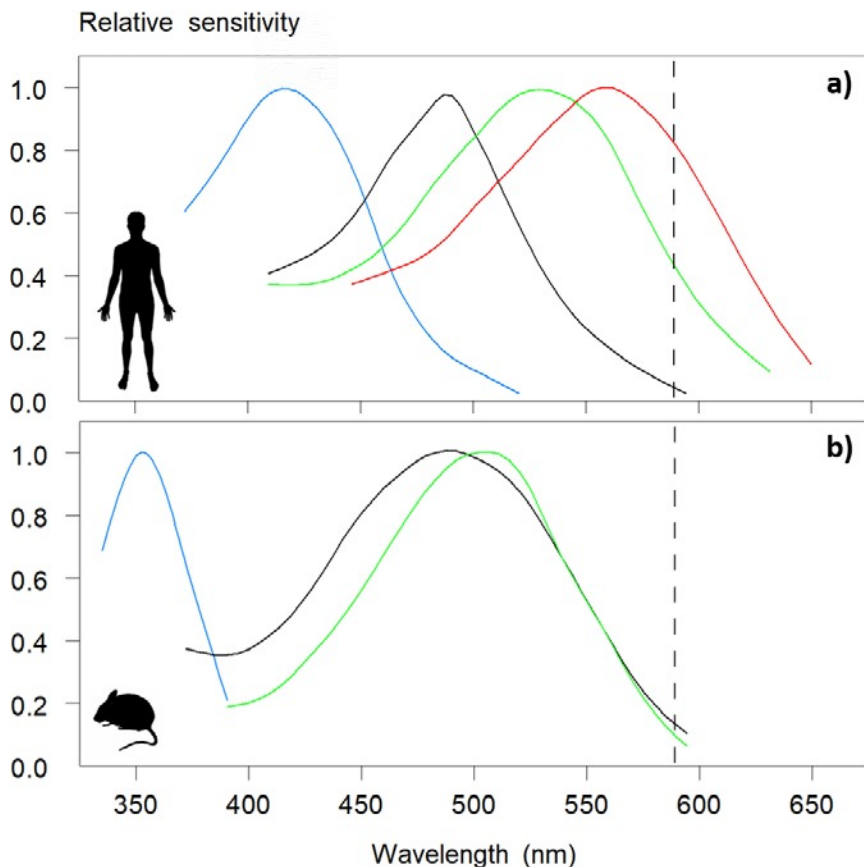


Fig. 2.2. Relative sensitivity to different colour wavelengths for (a) humans and (b) mice. The wavelength (589 nm) emitted by low-pressure sodium light bulbs is indicated by the vertical dashed line (adapted from McLennan and Taylor-Jeffs, 2004).

detectable differences in their time budget compared with non-lame conspecifics (Weigele *et al.*, 2018). Recently, Mandel *et al.* (2018) showed that a certain degree of lameness in cows could be detected indirectly through differences between cows in their use of cow brushes, devices installed in cow sheds to allow the animals to scratch difficult-to-reach places on their body. However, as highlighted in the review by Van Nuffel *et al.* (2015), in many studies mildly lame cows are lumped together with the non-lame cows, making it difficult to use the method to detect early signs of lameness. This is where development over time comes in. When an animal is used as its own control, even very subtle changes in its behaviour are more easily detected, and this is also the case for the development of lameness in dairy cows (de Mol *et al.*, 2013).

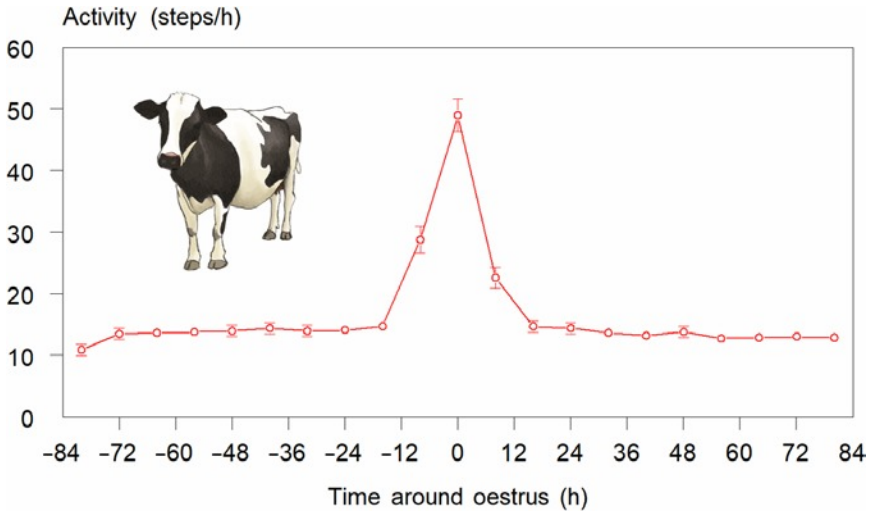


Fig. 2.3. Number of steps measured every 8 h around oestrus in 49 dairy cows (adapted from Arney *et al.*, 1994).

Time- and situation-specific observations

Across many species, a lot of behavioural elements are linked to the circadian rhythm of the animal. As mentioned earlier, mice and rats are nocturnal species, and if we are interested in their behaviour and general activity levels, it is appropriate to study them during their natural period of activity. Some behaviours are situation specific, such as flight responses when a predator is encountered. Others are cyclic across periods longer than 24 h; this includes the state of being in oestrus for mature and non-pregnant female mammals. In rats, this occurs roughly every 4 days, where the female will display receptive behaviour, such as lordosis, where the rat takes a prone position with an exaggerated inward curvature of the spine thereby allowing easier access for copulation. In sows, oestrus gives rise to an increased likelihood of continued standing when light pressure is put on the rump of the sow, a signal that she will accept mounting by the boar. In some species, oestrus increases the locomotion of the female, and cows in heat will walk up to four times the distance measured at other times during the on-average 21-day oestrus cycle (Fig. 2.3; Arney *et al.*, 1994).

Some behaviour patterns are seen only at specific times, such as dust-bathing in hens. This behaviour consists of the bird transferring a friable substrate, such as sand, in between its feathers through a sequence of different behavioural elements (including scratching, bill raking, wing shaking and head rubbing) lasting several minutes, and finishing with a whole body shake (Nicol, 2015). It serves as a grooming behaviour for cleaning and maintenance of integrity of the feathers (Vestergaard, 1981), and is more likely to take place around midday (Mishra *et al.*, 2005; Campbell

et al., 2016; Mutibvu *et al.*, 2017). However, hens do not necessarily dust-bathe every day, and registration and quantification of this and similarly infrequent behaviour should take this into account.

Locomotor Activity

Activity in the form of locomotion is one of the simplest but also most important measures of animal behaviour. This can take the form of tracking individuals on video recordings filmed from above the animal enclosure (e.g. Meunier and Nielsen, 2014), providing the animals with running wheels (e.g. Bartling *et al.*, 2017), or monitoring movement of the legs by means of pedometers or accelerometers (e.g. Thorup *et al.*, 2016). However, when animals are kept in groups and we want to measure general activity, other methods may be more appropriate.

At some point during a research project, I needed to be able to measure the activity levels of groups of day-old chickens. At the time, my children were still quite young, and out of curiosity I asked them how they would measure how much a group of chickens moved. Having thought about it for a surprisingly short time, they came up with the idea of a pen with a floor of compacted mud, on which they would simply count the number of chicken footprints. Perhaps the feasibility of this suggestion was less than ideal, but still not bad for a couple of 7 and 8 year-olds. We ended up instead using passive infrared detectors (PIDs), which are mostly known as the sensors that make your porch-lamp light up automatically when you come home late at night. These sensors are activated by temperature differences that move, which is why the lights also come on when a (warm) cat passes the (colder) driveway. We used versions of PIDs that registered and stored files of the monitored movement in volts relative to time (Pedersen and Pedersen, 1995). This allowed us to estimate overall movements, achieving similar results to those obtained by the (at the time more laborious) logging of pixel changes between consecutive frames of a video recording (Nielsen, 2003; Nielsen *et al.*, 2004). I have included this example, because the PID curves we obtained from these newly hatched chicks appeared to indicate that the groups showed rhythmic bouts of activity (Nielsen *et al.*, 2008). These rhythms (Fig. 2.4a) were not synchronized among pens, thereby excluding the possibility of some external time-keeper like the turning on and off of the ventilation. However, simulation of the behaviour of individual chicks allowed us to reproduce the curves obtained from the PIDs (Fig. 2.4b). The apparent rhythmicity turned out to be an artefact of the superposition of individual activity cycles, occurring when the periods of inactivity of individual chicks are interspersed with shorter bouts of activity (Fig. 2.4c). Summation of these data gave rise to an undulating curve (Fig. 2.4b), which does not reflect the behaviour of the individuals in the group, but is a result of the so-called beats effect, when two or more oscillations of different frequencies interfere. We had been very

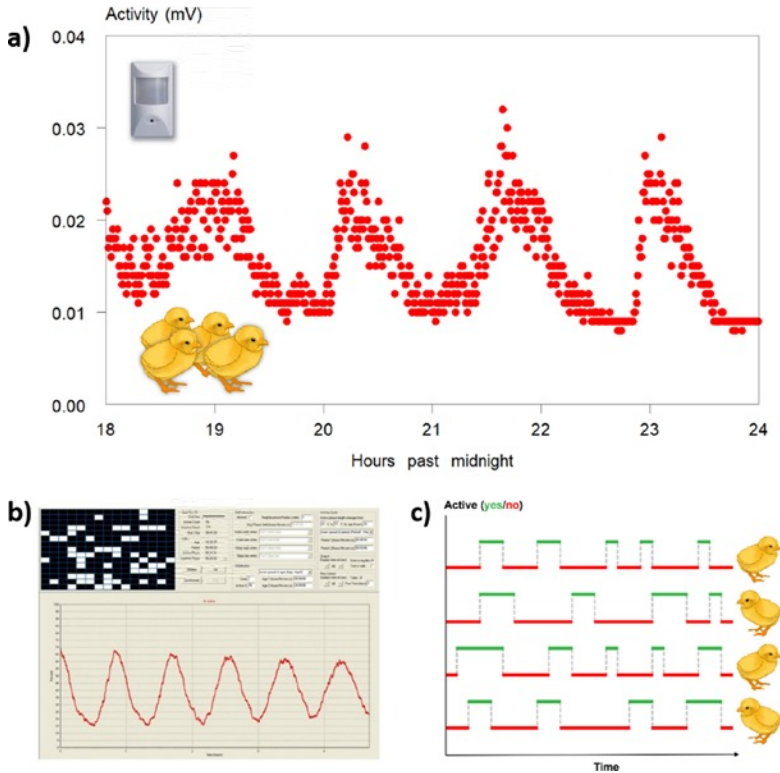


Fig. 2.4. (a) Activity rhythms measured by a passive infrared detector (PID; photo inset) in a group ($n = 225$) of day-old broiler chicks. (b) Computer simulations revealed that these rhythms were an artefact of the sum of movements by individual chicks being either active or inactive (c), as long as the activity bouts were shorter than the inactive periods (adapted from Nielsen *et al.*, 2008).

excited when we first saw the rhythmic activity in the PID data, thinking (incorrectly) that groups of young chicks were able to maintain a synchronized activity rhythm in the absence of a mother hen, at least for the first few days after hatch. However, this turned out not to be the case. Whenever automatic registration of activity is carried out, it is essential to ascertain that the data obtained are true representations of the behaviour of the animals. This is also briefly discussed in Chapter 13.

Feeding Behaviour

A behavioural activity of great interest across a number of scientific disciplines is feeding behaviour. In many studies this is measured only as daily feed intake (DFI) by daily subtractions of the weights of feed delivered and feed left over. However, feeding behaviour is obviously much more detailed

than the – albeit important – measure of total intake in a day. Individuals differ in the way they feed, with some eating little but often, while others consume few but large meals. We also know that feeding behaviour is affected by the social environment, so that pigs housed individually have been found to feed more than twice as frequently as pigs housed in groups (de Haer and de Vries, 1993). When feeding behaviour is not the main subject of a study, registration of behaviour around feeding may nevertheless add information that can be useful for the interpretation of other behavioural measurements. I have previously argued (Nielsen, 1999) that changes in the speed with which an animal eats can reflect two things: its degree of hunger, and the constraints imposed by the social environment. In other words, if you are hungry, or if access to feed is somewhat limited or easily interrupted, you will eat faster (Nielsen *et al.*, 1995).

Unless we are dealing with adult mayflies, most studies of animal behaviour over a certain length of time are bound to include feeding behaviour. In the experimental situation this most often involves feeding on one type of highly homogenized feed, such as a pelleted or pre-mixed diet. The reason for this is to standardize feeds across the treatment of interest in order to reduce the variation between animals in feeding-related measures. But, as can be seen in the hypothetical example in Fig. 2.5, even when individuals show identical daily feed intake and eat at the same speed and for the same amount of time, they may still differ greatly in terms of their meal pattern. From the three variables describing meals (i.e. meal size, meal frequency and meal duration) other feeding behaviour characteristics can be calculated, such as daily feed intake, feeding rate and time spent feeding. However, the reverse calculation is not possible, as different meal patterns can give rise to the same feeding behaviour characteristics (Fig. 2.5). This should be kept in mind when designing experiments where feeding behaviour can differ, and – if measurements are possible – meal patterns can potentially be used to account for inter-individual variation in other, non-feed-related variables.

This gives me an opportunity to draw your attention to another consequence of the inverse relationship between meal size and meal frequency. As mentioned above, calculation of feed intake is most often done by weighing the amount of feed left over and subtracting this from the weight of the feed delivered 24h earlier. This can be done for individuals or on a group basis, and for the three hypothetical goats in Fig. 2.5, we would arrive at an average intake of 5.5 kg for the day shown. Within a day, the intake of individual goats can also be calculated as the meal frequency multiplied by the mean meal size, e.g. goat B eats 5 meals of 1.1 kg per day, giving rise to a DFI of 5.5 kg. When we have measures of individual meal patterns, we might be tempted to calculate the daily intake of the group as the product of mean meal frequency and mean meal size of the group. This is where things go wrong. The three goats in the example have a mean meal frequency of $(2+5+10)/3=5.7$ meals/day. They also have a mean meal size

Daily feed intake : **5.5 kg/day**
 Feeding rate: **23 g/min**
 Time spent feeding: **4 h/day**



Animal:	A	B	C
Meal frequency (meals/day)	2	5	10
Meal duration (min/meal)	120	48	24
Meal size (g/meal)	2 750	1 100	550

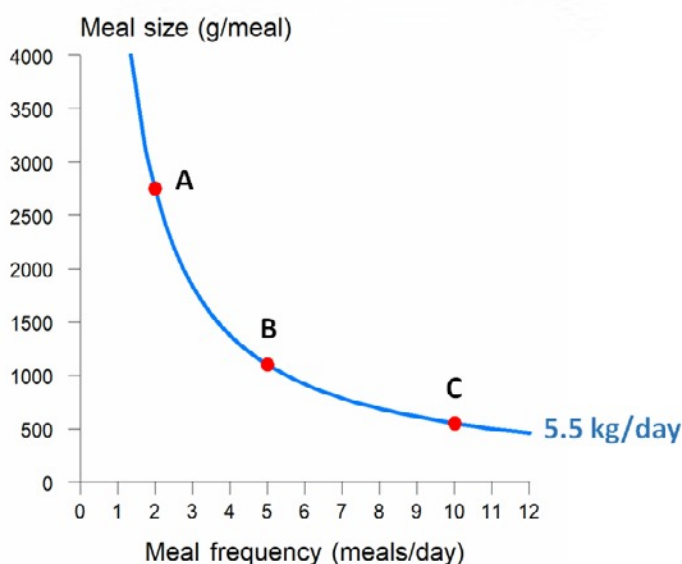


Fig. 2.5. Three animals with identical daily feed intake, feeding rate and time spent feeding may vary greatly in terms of their meal patterns. The graph illustrates the relationship between meal frequency and meal size, where all combinations along the curve (isocline) give rise to a daily intake of 5.5 kg, with the meal pattern of the three goats indicated. Data are simulated for clarity (adapted from Nielsen, 1999; photo: Ophélie Dhumez).

of $(2.75 + 1.10 + 0.55) / 3 = 1.467$ kg/meal. However, when we multiply these two numbers together, we get $5.7 \text{ meals/day} \times 1.467 \text{ kg/meal} = 8.3 \text{ kg/day}$. This is over 50% more than the 5.5 kg we know they eat on a daily basis. Why is this? The overestimation arises because we have calculated the product of two variables that are inversely correlated. If in Fig. 2.5 you draw a

vertical line through 5.7 meals/day on the x -axis, and a horizontal line through 1.467 kg/meal on the y -axis, the lines will cross above the isocline of 5.5 kg/day. This is because meal size (S) does not correlate directly with meal frequency (F), but with $1/F$. So if we recalculate using $1/F$, we get $(1/2 + 1/5 + 1/10)/3 = 0.27$ day/meal (the unit is inversed), and the calculated mean daily intake for the group is 1.467 kg/meal divided by 0.27 day/meal, which gives 5.5 kg/day. Apologies for this mathematical digression, which is but loosely related to the subject of this book.

Concluding Remarks

In this chapter, I have tried to highlight the importance of knowing the behavioural repertoire of the species you work with. This will help to identify potential constraining effects of the housing or testing environment, which may serve as causal factors when certain behaviours are not observed or changed in their way of being expressed, both qualitatively and quantitatively. An example of this is restlessness seen in cows prior to calving. This has been suggested to reflect the motivation of the parturient cow to isolate herself from the herd to give birth, but indoor housing does not allow the cow to increase her distance to the other group members (Rørvang *et al.*, 2018). Knowledge of the likely time of occurrence of different behaviours is also important as shown by the dustbathing example from hens mentioned earlier. Changes in behaviour may also be caused by human activity, both within a housing system and in the wild. Recently, Gaynor *et al.* (2018) showed that many wild-living species were becoming more nocturnal, on average an increase of 36% in nocturnality, to avoid human disturbance. This included human activities such as hunting and urban development, but also less dramatic activities, such as hiking. In Chapter 11, examples will be given of behavioural tests carried out in a natural setting.

References

- Arey, D.S., Petchey, A.M. and Fowler, V.R. (1991) The preparturient behaviour of sows in enriched pens and the effect of pre-formed nests. *Applied Animal Behaviour Science* 31(1–2), 61–68. DOI: 10.1016/0168-1591(91)90153-O.
- Arney, D.R., Kitwood, S.E. and Phillips, C.J.C. (1994) The increase in activity during oestrus in dairy cows. *Applied Animal Behaviour Science* 40(3–4), 211–218. DOI: 10.1016/0168-1591(94)90062-0.
- Bartling, B., Al-Robaiy, S., Lehnich, H., Binder, L., Hiebl, B. *et al.* (2017) Sex-related differences in the wheel-running activity of mice decline with increasing age. *Experimental Gerontology* 87, 139–147. DOI: 10.1016/j.exger.2016.04.011.
- Campbell, D.L.M., Makagon, M.M., Swanson, J.C. and Siegford, J.M. (2016) Litter use by laying hens in a commercial aviary: dust bathing and piling. *Poultry Science* 95(1), 164–175. DOI: 10.3382/ps/pev183.

- de Haer, L.C.M. and de Vries, A.G. (1993) Feed intake patterns of and feed digestibility in growing pigs housed individually or in groups. *Livestock Production Science* 33(3–4), 277–292. DOI: 10.1016/0301-6226(93)90008-6.
- de Mol, R.M., André, G., Bleumer, E.J.B., van der Werf, J.T.N., de Haas, Y. *et al.* (2013) Applicability of day-to-day variation in behavior for the automated detection of lameness in dairy cows. *Journal of Dairy Science* 96(6), 3703–3712. DOI: 10.3168/jds.2012-6305.
- Gaynor, K.M., Hojnowski, C.E., Carter, N.H. and Brashares, J.S. (2018) The influence of human disturbance on wildlife nocturnality. *Science* 360(6394), 1232–1235. DOI: 10.1126/science.aar7121.
- Kornum, A.L., Röcklinsberg, H. and Gjerris, M. (2017) The concept of behavioural needs in contemporary fur science: do we know what American mink (*Mustela vison*) really need? *Animal Welfare* 26(2), 151–164. DOI: 10.7120/09627286.26.2.151.
- Mandel, R., Harazy, H., Gygas, L., Nicol, C.J., Ben-David, A. *et al.* (2018) Detection of lameness in dairy cows using a grooming device. *Journal of Dairy Science* 101, 1511–1517.
- Martin, P. and Bateson, P. (2007) *Measuring Behaviour – An Introductory Guide*. Cambridge University Press, Cambridge, UK.
- McLennan, I.S. and Taylor-Jeffs, J. (2004) The use of sodium lamps to brightly illuminate mouse houses during their dark phases. *Laboratory Animals* 38(4), 384–392. DOI: 10.1258/0023677041958927.
- Meunier, N. and Nielsen, B.L. (2014) Using scalar products to refine the interpretative value of an orientation choice test. *Chemical Senses* 39(6), 551–560. DOI: 10.1093/chemse/bju024.
- Mishra, A., Koene, P., Schouten, W., Spruijt, B., van Beek, P. *et al.* (2005) Temporal and sequential structure of behavior and facility usage of laying hens in an enriched environment. *Poultry Science* 84(7), 979–991. DOI: 10.1093/ps/84.7.979.
- Mutibvu, T., Chimonyo, M. and Halimani, T.E. (2017) Effects of strain and sex on the behaviour of free-range slow-growing chickens raised in a hot environment. *Journal of Applied Animal Research* 46(1), 224–231. DOI: 10.1080/09712119.2017.1287079.
- Newberry, R.C. and Wood-Gush, D.G.M. (1985) The suckling behaviour of domestic pigs in a semi-natural environment. *Behaviour* 95(1–2), 11–25. DOI: 10.1163/156853985X00028.
- Newberry, R.C. and Wood-Gush, D.G.M. (1986) Social relationships of piglets in a semi-natural environment. *Animal Behaviour* 34(5), 1311–1318. DOI: 10.1016/S0003-3472(86)80202-0.
- Nicol, C.J. (2015) *The Behavioural Biology of Chickens*. CAB International, Wallingford, UK.
- Nielsen, B.L. (1995) *Feeding Behaviour of Growing Pigs: Effects of the Social and Physical Environment*. PhD thesis. The University of Edinburgh, UK.
- Nielsen, B.L. (1999) On the interpretation of feeding behaviour measures and the use of feeding rate as an indicator of social constraint. *Applied Animal Behaviour Science* 63(1), 79–91. DOI: 10.1016/S0168-1591(99)00003-9.
- Nielsen, B.L. (2003) Use of passive infrared detectors (PIDs) to monitor bursts of activity in groups of broilers. *Archiv für Geflügelkunde* 67, 42–44.

- Nielsen, B.L., Lawrence, A.B. and Whitemore, C.T. (1995) Effects of single-space feeder design on feeding behaviour and performance of growing pigs. *Animal Science* 61(3), 575–579. DOI: 10.1017/S1357729800014168.
- Nielsen, B.L., Kjær, J.B. and Friggens, N.C. (2004) Temporal changes in activity measured by passive infra-red detection (PID) of broiler strains growing at different rates. *Archiv für Geflügelkunde* 68, 106–110.
- Nielsen, B.L., Erhard, H.W., Friggens, N.C. and McLeod, J.E. (2008) Ultradian activity rhythms in large groups of newly hatched chicks (*Gallus gallus domesticus*). *Behavioural Processes* 78(3), 408–415. DOI: 10.1016/j.beproc.2008.02.010.
- Olsson, I.A.S. and Keeling, L.J. (2002) The push-door for measuring motivation in hens: laying hens are motivated to perch at night. *Animal Welfare* 11, 11–19.
- Patterson-Kane, E.G. (2004) Enrichment of laboratory caging for rats: a review. *Animal Welfare* 13, S209–S214.
- Pedersen, S. and Pedersen, C.B. (1995) Animal activity measured by infrared detectors. *Journal of Agricultural Engineering Research* 61(4), 239–246. DOI: 10.1006/jaer.1995.1051.
- Rodenburg, T.B., Bracke, M.B.M., Berk, J., Cooper, J., Faure, J.M. *et al.* (2005) Welfare of ducks in European duck husbandry systems. *World's Poultry Science Journal* 61(4), 633–646. DOI: 10.1079/WPS200575.
- Rørvang, M.V., Nielsen, B.L., Herskin, M.S. and Jensen, M.B. (2018) Prepartum maternal behavior of domesticated cattle: a comparison with managed, feral, and wild ungulates. *Frontiers in Veterinary Science* 5, 45. DOI: 10.3389/fvets.2018.00045.
- Stolba, A. and Wood-Gush, D.G.M. (1989) The behavior of pigs in a semi-natural environment. *Animal Production* 48, 419–425.
- Thorup, V.M., Nielsen, B.L., Robert, P.-E., Giger-Reverdin, S., Konka, J. *et al.* (2016) Characterising individual cow feeding and rumination behavior from sensor data to detect lameness. *Frontiers in Veterinary Science* 3, 37.
- Van Nuffel, A., Zwervaegher, I., Van Weyenberg, S., Pastell, M., Thorup, V.M. *et al.* (2015) Lameness detection in dairy cows: Part 2. Use of sensors to automatically register changes in locomotion or behavior. *Animals* 5(3), 861–885. DOI: 10.3390/ani5030388.
- Vestergaard, K. (1981) Dust-bathing behaviour in the domestic hen. *Applied Animal Ethology* 7(4), 386–387. DOI: 10.1016/0304-3762(81)90070-5.
- Weigele, H.C., Gygas, L., Steiner, A., Wechsler, B. and Burla, J.-B. (2018) Moderate lameness leads to marked behavioral changes in dairy cows. *Journal of Dairy Science* 101(3), 2370–2382. DOI: 10.3168/jds.2017-13120.

How to Choose a Test

In order to select the most appropriate behavioural test, we first need to consider why we use behavioural tests at all. What is the purpose of putting an animal in a situation that will always be somewhat artificial, no matter how hard we try to create test environments suitable for the species in question? As described in the previous chapter, a lot of information about an animal and its behaviour can be gained from observing the animal in its home environment, even when this is non-natural as is the case for pets kept at home, housed livestock and zoo enclosures, to mention but a few. So a behavioural test needs to add something else, something we cannot easily know by simply observing the species.

In that sense, behavioural tests are no different from any other test: we are seeking to answer a specific question in the most optimal way. And an optimum is always a balance between a number of competing issues: we want to use as few animals as possible for as little time as possible, while still obtaining results that are unbiased, interpretable and – hopefully – significant. Many behavioural tests can thus be regarded as a proxy for showing effects that we otherwise could only unveil by observing the animals for long enough and in every thinkable situation. An example of this is the use of operant techniques where an animal is asked to press a lever to obtain access to a resource. These are described in more detail in Chapters 5 and 8, and are a good illustration of how we can ask an animal a specific question, such as ‘How much are you willing to work for access to a rotating brush to scratch yourself with if you are a dairy cow?’ (McConnachie *et al.*, 2018). It turns out that cows are willing to work as hard for access to a scratching opportunity as they would for access to fresh feed. To establish this through behavioural surveillance alone would require many days of observation and give rise to experimental design challenges, such as how to compare the comfort of cows with access to brushes with those that do not have this opportunity.

Behavioural tests can often provide faster and more accurate information than direct behavioural observation over time, as the following

example shows. A group of Norwegian researchers trained horses to indicate through a choice of different symbols whether they wanted to wear a blanket or not, or, if already wearing one, whether they wanted it off (Mejdell *et al.*, 2016; Fig. 3.1). All the horses ($n = 23$) learnt this within 2 weeks of training, and were able to use it afterwards, as their choices mirrored the prevailing weather; thus they wanted the blankets off when the weather was warm, and on when the weather was colder, wetter and windier. It would have taken longer than 2 weeks of observation to establish such preferences, putting blankets on and taking blankets off horses in different kinds of weather while noting down their behaviour in order to establish if they were more or less thermally comfortable. Different horses showed slightly different preferences, so perhaps some horses would like to be without a blanket on colder days, where we would put one on them by default (in the same way that parents tell their child to put on a sweater when they are cold themselves). This testing paradigm was set up to assess if learning to indicate preferences via symbols was possible in horses and, if so, within a feasible learning period. As the results demonstrate, the horses learnt the task, with some of them at times being very eager to communicate their preference. The authors describe occasions where horses were allowed to indicate their choice before testing was due to start, with the horses indicating the ‘blanket off’ symbol. When the blanket was removed, the horses were found to be sweaty underneath.



Fig. 3.1. Horses are able to indicate whether they want to wear a blanket or not by touching square boards with different symbols. When wearing a blanket, the horse is given a choice between no change (white square) or blanket off (vertical line). The two horses shown here chose differently under similar weather conditions. When not wearing a blanket, the horses can choose between the symbols for blanket on (horizontal line) and no change (Mejdell *et al.*, 2016; photos: Turid Burvik and Cecilie M. Mejdell).

Once we have established that we cannot easily answer our behavioural question by observing the animal in its home environment, we need to identify the objective before deciding on which behavioural test to use: what is the goal we would like to achieve? It may be a good idea to consider this goal at two levels: the *general goal* in terms of the results of the test telling us something fundamental about the behaviour of the species or type of animal tested (in other words, to what extent can we extrapolate the results and interpret them in a wider context?), and the more *specific goal* of the test in question (what are the measurements necessary to be able to make the more general interpretation?). The generalizability of our results is closely linked to the way we ask our scientific question. 'Does a given dose of drug A give rise to increased locomotion in mice?' is a different question from 'Can activity be affected by drugs?'.

It goes without saying that different types of tests are employed dependent on what we are interested in finding out. Do we want to know more about the behaviour of the species itself? Do we want to know more about behavioural mechanisms in general? Or do we want to investigate the effects of different treatments, whatever form they take? One test can often provide answers to more than one question, but the detail of the design may allow for more or less wide-ranging interpretations. Sometimes confounding effects need to be carefully considered to ensure that the measurements made are a reflection of the question asked. One example is the question of hunger in feed-restricted animals, such as the parents of broiler chickens. Broiler chickens have been selected for decades for rapid and efficient lean growth; in other words they put on a lot of muscle fast with a minimum of feed. Commercial broilers are usually slaughtered around 35 days of age, having grown from the weight of the egg (~60 g) to around 2 kg in this period. This also means that their parents, having to reach maturity in order to reproduce, have to be severely feed restricted so as not to become excessively heavy with cardiovascular, locomotor and reproductive problems as a consequence. This has been called the 'broiler breeder dilemma' (Mench, 2002), because both ad libitum and restrictive feeding cause problems, the latter resulting in prolonged hunger as female broiler breeders can be fed at 30% of their ad libitum intake. One question is how to compare the degree of hunger in birds that are used to being feed restricted with conspecifics that are ad libitum fed? The confounding effect here is that the two situations are comparing birds at greatly different live weights, thus we are not comparing like with like (D'Eath *et al.*, 2009). Another problem in the comparison is that, having been fed restrictively for most of their lives, commercial broiler breeders are likely to be able to eat more and faster in one sitting, which makes direct tests on feed intake difficult. Together with colleagues, we tried to overcome these issues by adapting existing tests of hunger previously used in broiler breeders (Sandilands *et al.*, 2005).

Groups of broiler breeders were raised on three different pelleted feed types, and when the birds were 11-14 weeks of age, a subset was tested in pairs to estimate their hunger (Nielsen *et al.*, 2011). The speed with which you eat can reflect your level of hunger (Nielsen, 1999), and the test was carried out at the time of day when these restrictively fed birds were given their daily ration, as they had been once a day throughout their lives. After feeding, the two birds to be tested were placed in a small pen in which they had been group housed when younger. It contained a circular feeder containing a known amount of their normal feed, but covered by a lid to prevent the birds eating. The following day at their usual feeding time, the birds were given access to the feed for 2 min, after which time the feed trough was covered again and the remaining feed weighed. Immediately following the test, the birds were given *ad libitum* access to the feed as well as water. On the fifth day, the lid was put back on the feed trough for 24 h,

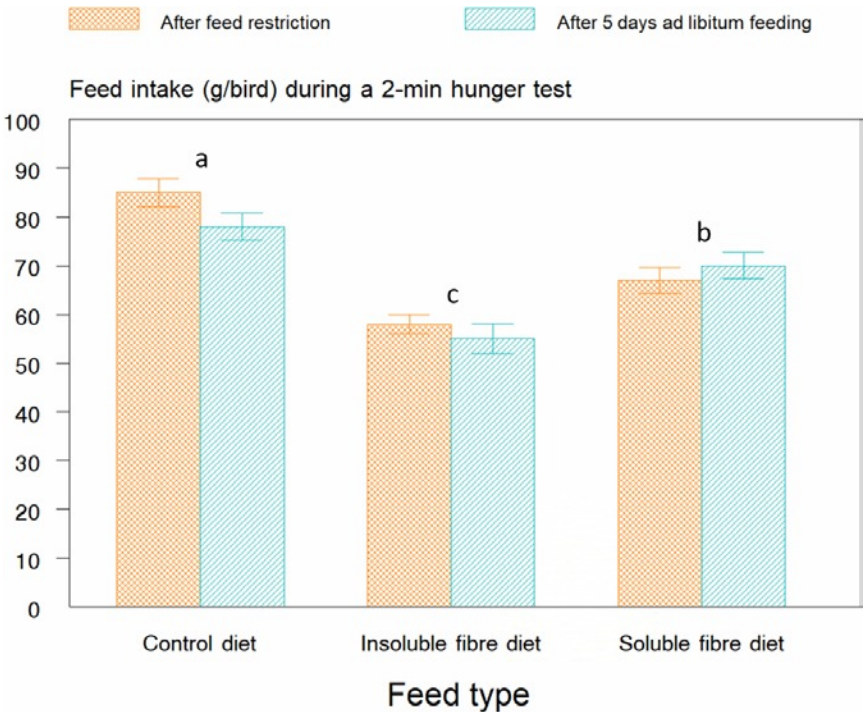


Fig. 3.2. Feed intake (g/bird) of three diets differing in proportion of insoluble fibre by female broiler breeders of around 12 weeks of age after fasting for 24 h. Birds were given access to the pelleted feed for 2 min, either when previously feed restricted or following 5 days of *ad libitum* feeding on the same diets. No differences were found between the two tests within each feed type, but a higher fibre inclusion led to slower eating, which was more pronounced when the feed contained more insoluble fibre (data from Nielsen *et al.*, 2011).

at which point the birds were again given access to the feed for 2 min, and the remaining feed weighed.

The idea was to see if birds that were accustomed to 24-h feed withdrawals (because they had been restrictively fed once a day for several weeks) would be less hungry and therefore eat slower than birds that had been *ad libitum* fed before the 24-h feed withdrawal. Three different feed types were tested in this trial, and the results are shown in [Fig. 3.2](#). Birds fed the control diet, which contained the most energy of the three feeds, ate more during the 2-min test, than birds fed one of the two fibre-rich feeds with lower energy content. A similar result was obtained when comparing volume rather than weight of feeds. Throughout their lives, the broiler breeders on the fibre diets had been given larger quantities of feed at the daily feeding than the birds on the control feed so as to give all birds equivalent daily energy intake. As the fibre diets led to a lower feeding rate during the test, it could indicate that they gave rise to greater satiety. Having said that, as no overall differences were found in feeding rate between previously restricted and *ad libitum* fed birds, would this not mean that the test does not reflect level of hunger? The two fibre diets differed in their proportion of insoluble fibre, with a high content of oat hulls (insoluble fibre) in one and lots of beet and potato pulp (soluble fibre) in the other. It could be that the birds were simply not able to eat faster on the fibre diets, especially when the feed contained a lot of insoluble fibre. However, all three diets were in pelleted form, and differed only slightly in terms of density (mass to volume ratio). Perhaps 2 min was not sufficient time to fully reveal the differences in hunger, or perhaps 5 days of *ad libitum* feeding did not make the birds unlearn that feed is not always there? Another way to interpret these results is that 24h of feed withdrawal makes birds hungry, whether or not they are used to it, and the restricted feeding of broiler breeders is an animal welfare problem with no easy solution. Sometimes behavioural tests raise as many questions as they provide answers.

On a very pragmatic level, the behavioural tests used in a variety of laboratories are often chosen on much less lofty criteria than the ones mentioned at the beginning of this chapter. Within a research team, the subject area remains relatively constant over time, and the same behavioural tests are often used in different projects. This is amplified by the available experimental equipment, so that the presence of an elevated plus maze (see Chapter 4) may lead to it being used without consideration for its suitability or for potential alternatives. In a world where funding is limited, it sometimes boils down to what is possible within the constraints of a given project, in terms of time, staff availability and money. Many behavioural tests are relatively cheap to carry out, which can be a blessing in disguise: to perform a forced swim test, which is a severely stressful test for rodents, you only need access to a glass beaker, some lukewarm water, a stopwatch and a towel (see Chapter 7 for a critical discussion of this test). Other tests require more sophisticated apparatus; an incomplete list of test types and their equipment is shown in [Table 3.1](#).

Table 3.1. Examples of equipment for use in behavioural tests. Lists such as these can be found online, often published by companies producing the equipment. The same equipment can sometimes be used for different tests, and the same test may occasionally be used to investigate different aspects of an animal's behavioural responses.

Research area	Behavioural test	Examples of equipment
Activity and exploration	Locomotor activity/rearing	Video camera, (infrared) activity meter
	Open field test	Open field arena
	Hole-board test	Hole-board, video camera
	Response to novelty	Open field arena, objects
	Voluntary exercise	Activity wheel (rodents)
Anxiety	Open field test	Open field arena
	Locomotor activity/rearing	(Infrared) activity meter, video camera
	Elevated plus maze test	Elevated plus maze
	Dark/light test	Black and white box
Depression	Forced swim test	Cylindrical glass beaker
	Tail suspension test	Sticky tape
Learning and memory	Passive or active avoidance	Passive box or shuttle box
	Aron test	Aron box
	Morris water maze	Circular pool
	Radial maze test	Radial maze
	T-maze test	T-maze
	Object recognition test	Open field arena, objects
	Operant procedures	Operant box, Skinner box
	5/9-hole test	5/9-hole box
Reward	Place preference test	Place preference box
	Operant procedures	Operant box
Sensory and motor skills	Coordination and equilibrium	Rotating rod (rodents)
	Grip strength	Grip strength meter
	Exercise training	Treadmill
	Nociception	von Frey filaments
	Startle response	Sound equipment, umbrella
Social behaviour	Social interaction	Open field arena, home cage
	Social reinstatement	Open field arena

The popularity of a behavioural test often spreads between research labs as more papers using the test are published. A test used by many will often lead to it being used even more. This may be the case for the aforementioned forced swim test, as the use of this test has increased over time (Fig. 3.3). Many behavioural tests were first developed in rodents, mainly

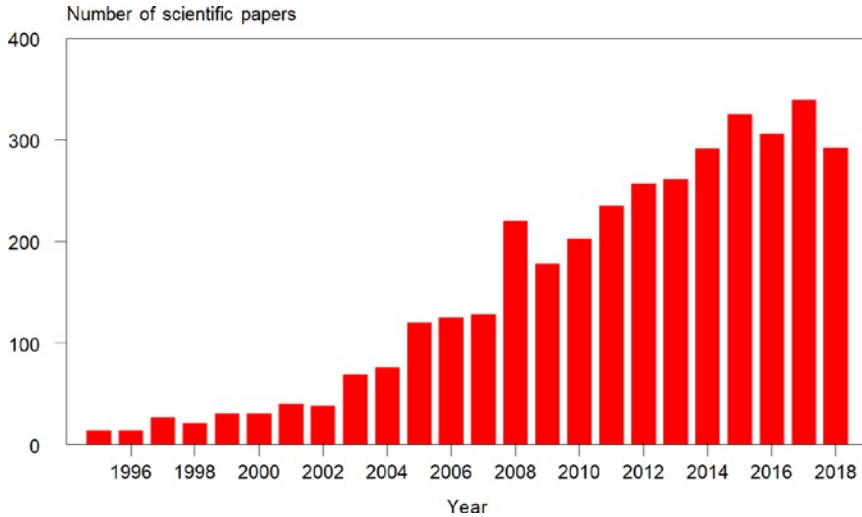


Fig. 3.3. Number of scientific articles published annually since 1980 in which 'forced swim test' or 'forced swimming test' are included in the title, abstract or keywords (data from Web of Science™).

mice and rats, and were later applied and adapted for other species and situations. The habituation/dishabituation test, described in more detail in Chapter 6, has mostly been applied in mice, but has recently been scaled up from a rodent cage to a dairy barn to investigate the olfactory capacities of cows (Rørvang *et al.*, 2017).

Returning to the subject of test choice, I cannot emphasize enough that the validity of any given test is of course of utmost importance. Does the test measure what we think it does? This is subtly different from the questions asked earlier (What is the purpose of the test? What question do I want the answer to?), as we may erroneously employ a test to answer one thing when it actually reflects something else. Such issues will be covered in Chapter 11, as they are easier to relate to once a large number of different test types have been presented. Instead, I will present an example of how we can build validity checks into a test. The open field test is used for a variety of purposes (Table 3.1), and a detailed discussion of this widely used test can be found in Chapter 4. In the following example, the open field test was used to estimate the stressfulness of different housing environments.

Pigs are sometimes housed in individual metabolism crates for research purposes. This can be for extended periods depending on the experimental treatment, resulting in the animals being socially isolated, which is a known stressor for pigs. Herskin and Jensen (2000) wanted to know if the stress associated with this housing could be reduced if the metabolism crates were pushed together so that the pigs at least had some degree of contact with a conspecific, albeit limited. One of the measures collected

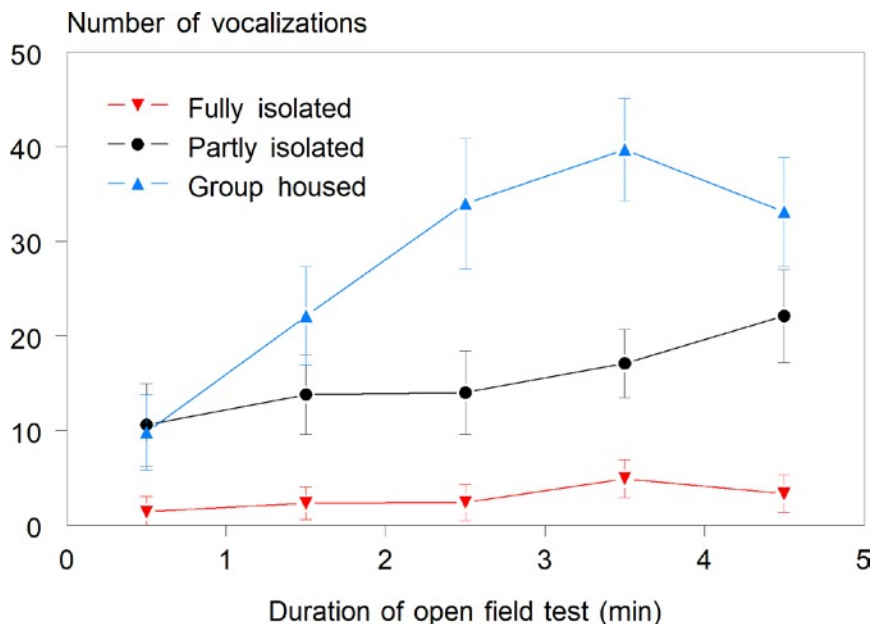


Fig. 3.4. Vocalizations (means $\pm$ SEM) by pigs during an open field test. The pigs were housed either in groups, fully isolated in individual metabolism crates or partly isolated with restricted physical contact with other pigs (data from Herskin and Jensen, 2000).

to investigate the effect of this partial isolation was vocalizations during an open field test. This was compared to two control situations, where pigs were either housed in groups or fully isolated in the metabolism crates. We would expect pigs on these two treatments to differ in their responses to the social isolation that an open field test imposes. This turned out to be the case, as shown in Fig. 3.4, with pigs coming from group housing vocalizing significantly more than pigs that were used to social isolation. The finding that partly isolated pigs were intermediate between the two indicates that the pushing together of the metabolism crates had a positive effect on the welfare of these pigs, as they vocalized more than the fully isolated conspecifics. It clearly did not fully compensate for a social environment as the partly isolated pigs did not respond with the same magnitude as the group housed ones. By including two control treatments in the experimental set-up, thereby identifying the two extremes of the behavioural response, the interpretability of these open field test results was made simple.

In order to know which test to use, it would be nice to make a library or filing system that would allow us to choose which test is appropriate for what situation, such as the list given in Table 3.1. However, any categorization comes with constraints and caveats. It is clearly not very useful to divide behavioural tests into groups according to the type of animal tested,

as many tests can be used for a range of species as diverse as lizards, fish and chickens. Instead, behavioural tests could be categorized according to their purpose or goals, e.g. tests to determine ability to smell, tests to quantify to what extent animals can generalize stimuli and tests to measure fear. However, sometimes the same behavioural test measures several aspects of the animal's current state. For example, the open field test can reflect aspects of fear, animal coping strategies and desire to be socially reinstated, dependent on the experimental treatment (see also Chapter 4).

Another way to label behavioural tests is to focus on the mechanics of the test. Is it done individually or in a group? What equipment is needed? How long does it last? Alternatively, tests can be grouped according to which behavioural system or domain is involved, such as learning, memory, hunger and affective state. These groupings are either not useful or too diverse to bring any real benefit for the choice of behavioural tests. It soon becomes clear that whichever method we chose for categorization, some overlap is inevitable. Behavioural tests are like a cloud of soap bubbles and, looked at from different angles, any one bubble will overlap with others to varying extents. In the following chapters, I have nevertheless clustered different behavioural tests into some umbrella chapter headings, within which some – but far from all – of these different tests are described, exemplified and discussed. Please remember the caveats from Chapter 1, and – even more importantly – use your common sense when using and executing a behavioural test.

References

- D'Eath, R.B., Tolkamp, B.J., Kyriazakis, I. and Lawrence, A.B. (2009) 'Freedom from hunger' and preventing obesity: the animal welfare implications of reducing food quantity or quality. *Animal Behaviour* 77(2), 275–288. DOI: 10.1016/j.anbehav.2008.10.028.
- Herskin, M.S. and Jensen, K.H. (2000) Effects of different degrees of social isolation on the behaviour of weaned piglets kept for experimental purposes. *Animal Welfare* 9, 237–249.
- McConnachie, E., Smid, A.M.C., Thompson, A.J., Weary, D.M., Gaworski, M.A. *et al.* (2018) Cows are highly motivated to access a grooming substrate. *Biology Letters* 14(8), 20180303. DOI: 10.1098/rsbl.2018.0303.
- Mejdell, C.M., Buvik, T., Jørgensen, G.H.M. and Bøe, K.E. (2016) Horses can learn to use symbols to communicate their preferences. *Applied Animal Behaviour Science* 184, 66–73. DOI: 10.1016/j.applanim.2016.07.014.
- Mench, J.A. (2002) Broiler breeders: feed restriction and welfare. *World's Poultry Science Journal* 58(1), 23–29. DOI: 10.1079/WPS20020004.
- Nielsen, B.L. (1999) On the interpretation of feeding behaviour measures and the use of feeding rate as an indicator of social constraint. *Applied Animal Behaviour Science* 63(1), 79–91. DOI: 10.1016/S0168-1591(99)00003-9.
- Nielsen, B.L., Thodberg, K., Malmkvist, J. and Steinfeldt, S. (2011) Proportion of insoluble fibre in the diet affects behaviour and hunger in broiler

- breeders growing at similar rates. *Animal* 5(8), 1247–1258. DOI: 10.1017/S1751731111000218.
- Rørvang, M.V., Jensen, M.B. and Nielsen, B.L. (2017) Development of test for determining olfactory investigation of complex odours in cattle. *Applied Animal Behaviour Science* 196, 84–90. DOI: 10.1016/j.applanim.2017.07.008.
- Sandilands, V., Tolkamp, B. and Kyriazakis, I. (2005) Behaviour of food restricted broilers during rearing and lay—effects of an alternative feeding method. *Physiology & Behavior* 85(2), 115–123. DOI: 10.1016/j.physbeh.2005.03.001.

Tests to Characterize the Animal

Whatever our ultimate scientific goal, we often (but not always) aim to use a cohort of animals that are as similar as possible. This is to minimize existing differences between our animals, so that the effects found can be attributed to the treatment applied, be it a drug, a new type of housing or a change in feeding schedule (but see the discussion on heterogeneity of treatment groups in Chapter 11). The common use of inbred strains of rodents is an example of trying to reduce the inherent variability between individuals. But it is not always possible to have access to a sufficiently large group of genetically similar animals, and even within the common strains of laboratory rodents individual differences can be observed. It can therefore be very useful to be able to group animals according to certain characteristics. When these are visually obvious (e.g. coat colour), or routinely measured (e.g. live weight), this is relatively easy to do. However, we may want to categorize our test animals based on some behavioural trait, and for this we want to make sure that what we measure is a true reflection of the animal's character. Ideally, we want some kind of measure that tells us something about the animal's personality.

No single test exists that can quantify the complete personality or temperament of a non-human animal. An array of tests have been designed to identify specific aspects of an animal's character, and most of these originate from work on laboratory rodents. These tests vary dependent on the species tested, they cover a wide range of behavioural phenotyping and many of them are often used also to investigate changes caused by an experimental treatment. Whole books have been written about personality testing of animals (e.g. MacKay, 2018), and in this chapter only some of the most common examples of tests of individual behavioural characteristics are included. Personality determines a propensity to react in a similar manner across a range of situations, and one of the basic features of individual temperament is fearfulness (Boissy, 1995). For this reason, a large number of tests are centred around aspects of fear and anxiety, but the examples presented here from a variety of species deal with the assessment of many

different characteristics of an individual, such as optimism, anxiety, fearfulness, boldness and dominance.

Open Field Test

In one of the first published examples of an open field test, a measure of defecation was used to estimate fearfulness in rats (Hall, 1934). The open field test has since become one of the most used behavioural tests across a large number of animal species and scientific disciplines. The open field test (sometimes called an arena test) can be simply described as placing an individual animal for a short period of time in an empty arena surrounded by walls to prevent the animal from escaping, while recording its behaviour. But within this simple description, there is a lot of room for variation. In their highly cited (>1300 times) review of this method, Walsh and Cummins (1976) lamented the lack of standardization of the open field apparatus, which was found to differ between studies in shape, size, colour, floor type, wall height and location of starting point, as well as variation in the surrounding light, sound and odour-scape. Indeed, the authors state that ‘it is hard to think of any facet which has not been modified’. I am less worried by this lack of standardization. First, the open field used for rodents is by default different from that used by goats, which again differs (luckily!) from that used for fish. Open field tests have even been carried out in wild caught squirrels (Mazzamuto *et al.*, 2019) and great tits (Dingemanse *et al.*, 2012), where movements in all three dimensions are not only possible, but much more likely. Secondly, in order to be useful, any test needs to be sufficiently robust to withstand minor differences in experimental layout. And finally, the comparisons of interests are those made within an experiment where the same open field arena is used across different treatments or age groups. It should, of course, be ensured that the walls are sufficiently high to prevent the animal escaping, that the floor is not slippery and that the arena is evenly lit with no shadows. Also, the test is best performed under video surveillance in the absence of the experimenter, who may provide unintended olfactory or auditory cues.

So, whether the arena is circular or square should make little difference, even if there can be no hiding in the corners in the former. Incidentally, if automatic video tracking is not possible, the movement around the open field arena is often scored by placing a grid of squares on the floor, either in real life or on the video image. This is easier to do in a square arena, although care needs to be taken when trying to assess the degree of thigmotaxis, where the animal is staying close to the wall and avoiding the open space towards the centre of the arena. This has been found to reflect anxiety in mice (Simon *et al.*, 1994), and is often – at least in the older literature – defined as spending time in the grid squares running along the wall. However, if, say, the arena used is of the dimensions 1 m by 1 m with a

grid of 5×5 squares on the floor, the likelihood of being in the 16 squares along the edges is 64% (16 divided by 25) compared with 36% for the nine central squares. This should be taken into account when calculating the propensity to stay near the wall. In a 1-m² arena, there is a 50% likelihood that, by chance, the test animal will be in a 14.6-cm-wide virtual corridor along the walls, as this is equivalent to an area of 0.5 m², or half the arena.

The most common behavioural measures performed when using an open field test are listed in Table 4.1. Some of these variables are clearly correlated, such as latency to move and time spent active. This should be taken into account when analysing the data, and one of the recommended ways to amalgamate the full spectrum of open field behaviour is to carry out a principal component analysis (PCA) or similar form of factor analysis, which will allow the many interdependent variables to be expressed in fewer dimensions. This has already been recommended by Walsh and Cummins (1976), who found the most common feature could be summarized by variables describing emotionality, and the second feature would contain variables reflecting exploration level. These are similar to the dimensions of valence and arousal proposed by Mendl *et al.* (2010b) for the study of animal emotion.

One of the problems when transferring testing methods, such as the open field test, from rats to other species like livestock, is that the evolutionary trajectory differs among species (Forkman *et al.*, 2007). For the nocturnal rodent, a brightly lit open space will induce thigmotaxis, whereas for animals that have evolved to forage in the open during the day, this may not be the case. Although described mainly as a fear test by Forkman *et al.*

Table 4.1. Behavioural variables commonly scored in open field tests in terrestrial mammals and non-flying birds. The grid and squares refer to markings on the arena floor. Adapted from Walsh and Cummins (1976) and Forkman *et al.* (2007).

Behaviours related to...	Measure
Activity	Latency to move
	Distance covered (e.g. number of grid-lines crossed)
	Time spent active (vs immobile)
Exploration	Use of arena (e.g. squares visited)
	Degree of thigmotaxis (staying close to the wall)
	Sniffing and licking
Emotionality	Freezing and alertness
	Rearing
	Escape attempts (i.e. jumps at the wall)
	Grooming (potential displacement behaviour)
	Vocalization
	Defecation and urination

(2007), they also concur with the observation that other, sometimes overlapping emotional aspects are at play. These include motivation to become reinstated with conspecifics, as the open field test socially isolates the animal, which is a known stressor for social species such as most farm animals. Increased freezing may thus reflect a high level of fearfulness, whereas escape attempts may indicate motivation for social reinstatement. Other behaviours can be more difficult to interpret, as immobility and a long latency to move could indicate both alertness and relaxedness, dependent on the details of the behaviour displayed. This highlights the importance of considering different variables from the test simultaneously.

As well as comparing animals from two extremes of housing environment, as in the example mentioned in Chapter 3 (Herskin and Jensen, 2000; Fig. 3.4), another way to examine the validity of various open field measures is to test animals that vary in specific behaviour traits achieved through genetic selection (see also Chapter 10). Jones *et al.* (1992) tested Japanese quail chicks in an open field. The birds came from breeding stock divergently selected for and against a stress response when being immobilized, as measured by differences in plasma corticosterone, which is found to correlate with fearfulness in birds. The researchers found a reduction in freezing behaviour and shorter latency to start moving in the open field test in chicks from the low stress line, indicating that these measures could be used as indicators of fear.

Novel Object Test

A test that is often performed immediately after and in the same arena as an open field test is the novel object test. In its simplest form, it consists of exposing a test animal to an unknown object while recording the behavioural responses of the animal for a set period of time, often 5-10 min. The object used is often brightly coloured and conspicuously shaped, both to ensure its novelty as well as making the animal notice the object immediately. For example, Sneddon *et al.* (2003) constructed different novel objects out of LEGO bricks when testing rainbow trout. A novel object test creates a motivational conflict in the animal: because it has not experienced the object before, the animal is simultaneously scared of this novel thing and curious as to what it is. Latency to approach and touch the object, as well as frequency and duration of interacting with the object, are common measures used to quantify the relative degree of fear and exploratory motivation.

The novel object test has been found to be a valid method for assessing fear in a number of species (Forkman *et al.*, 2007), and the test animals are usually tested individually. It is obviously difficult to repeat a novel object test as the animal will be familiar with not only the object but also the test situation (Boissy *et al.*, 2018), as is also the case for the open field test. This can to a certain extent be overcome by changing the test arena and using

different novel objects (Nawroth *et al.*, 2017). Dalmau *et al.* (2017) used a novel object test to assess fear in growing pigs in commercial conditions and looked into repeatability using a different approach. The researchers argued that by performing the same novel object test on a large number of pigs from many different pens both within and between farms, they would expect the variability to be lower within farms than between farms. This, according to the authors, would indicate that the test is a reliable way to assess the prevailing fear levels among pigs on a farm, where they are likely to be treated in a similar manner, and therefore have comparable levels of fear. For practical reasons, they tested pigs in groups, using the home pens of the animals as the arena, and three balloons of different colours as the novel objects. The balloons were filled with helium, hence buoyant, and tied to a weight with one piece of string, thereby keeping the balloons together and at the eye level of a standing pig. The weight with the balloons attached was placed in the centre of the pen at the start of the test, and the latency for the first animal to touch one of the balloons was recorded, together with the proportion of animals watching or touching the balloons within 2 min of the start of the test. Contact latency was found to be a reliable measure for assessing the general fearfulness of pigs among the farms (Fig. 4.1), and it was also the easiest to score, showing a high inter-observer repeatability.

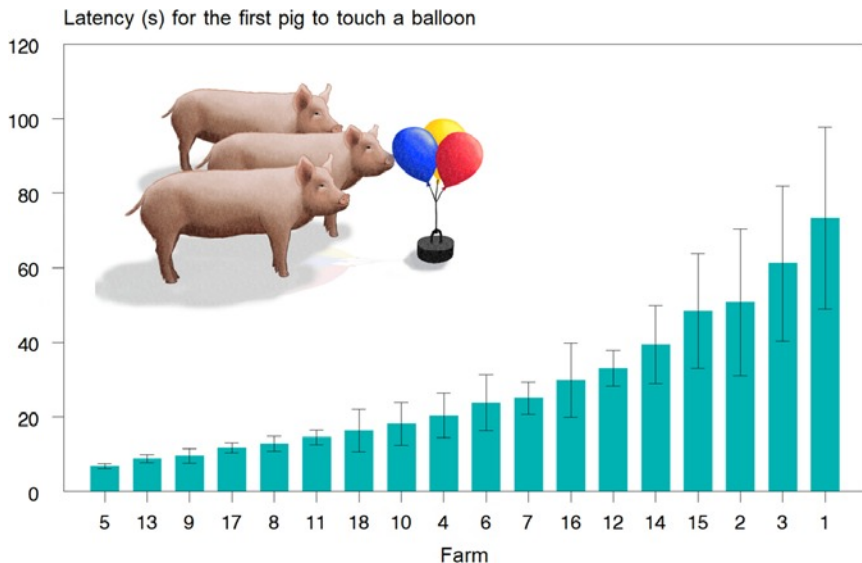


Fig. 4.1. Latency ($\pm$ SE) in seconds for the first pig in a pen to touch the novel object: a bunch of helium-filled balloons of different colours floating at pig eye-level when tied with string to a weight placed centrally in the home pen. The variation in this measure is greater among farms ($n = 17$) than within farms (data from Dalmau *et al.*, 2017).

In the novel object test, the animal may react upon discovering the object with behaviours indicative of fear, such as retreat away from the object, vocalization and alert immobility (freezing). However, if the animal does not display overt signs of being frightened, but still does not interact with or approach the object, it is difficult to know if the animal tested has a low level of fearfulness towards novelty or whether it is just not motivated to explore its surroundings. This is partly why the novel object test is often carried out in a barren arena, such as that used for an open field test, because it makes the novel object the dominant and only feature in the enclosure.

Startle Test

On an evolutionary basis, some of the key features of a predator attack are novelty, unpredictability and suddenness. In order to measure the response of an animal to these different aspects of predation within a test situation, the presentation of the unknown object can be made more startling. This can be done by dropping the novel object without prior warning into the arena, but it can be difficult to predict where the object will end up, which can make inter-test comparisons problematic: unless the object is in the centre (or at least in the same place for all animals tested) the possible distance to the novel object will differ among tests. One way to control the position of the novel object in the test arena is to lower it down rapidly from the ceiling. In this set-up, the animal may position itself underneath the trajectory of the object, and being hit from above by, say, an orange traffic cone is not part of the test procedure. Usually this can be prevented by delaying the entry of the object slightly if the animal is not in an appropriate place at the allotted time.

One approach is to make the presentation of a startling stimulus provoke a behavioural response in a relatively standardized manner. An alternative to lowering traffic cones or other objects rapidly from the ceiling is to use something that suddenly changes its size. An umbrella is ideal for this purpose, as it can be presented through a hole in the arena wall, first closed and functioning as a novel object in the same way as the LEGO bricks, the traffic cone and the balloons in the previous examples. Erhard *et al.* (2004) used an orange umbrella to test the reactivity to suddenness in male ($n = 12$) and female ($n = 20$) sheep. When the test animal was 20 cm away from the umbrella, the umbrella was opened, thereby transforming into a completely different shape in a split-second. They found that female sheep not only had higher levels of locomotion following the surprise, but also vocalized considerably more than their male conspecifics for the 3-min duration of the test (5.5 ± 1.6 vs 0.7 ± 0.3 vocalizations per min). This result indicates that female sheep have a higher emotional reactivity than do males.

Another, non-physical stimulus used to test for emotional reactivity is a sudden sound. This is known as the acoustic startle response, which

basically measures the jumpiness of an animal when exposed to an unpredictable, short, loud noise. The test is often used in rats, where an individual is restrained within a narrow box (a restrainer) and placed on a load cell within a soundproof chamber (Russo and Parsons, 2017). Speakers in the ceiling play constant white noise to mask other sounds except the acoustic stimulus, in this example a 50-ms burst of 95 dB white noise. The load cell automatically registers the displacement (maximal jump force) of the restrainer containing the test animal. Figure 4.2 shows that this measure appears to be a relatively stable characteristic of an animal, with a relatively large variation in jump force among rats tested. The acoustic startle response has been found to be exaggerated in human patients suffering

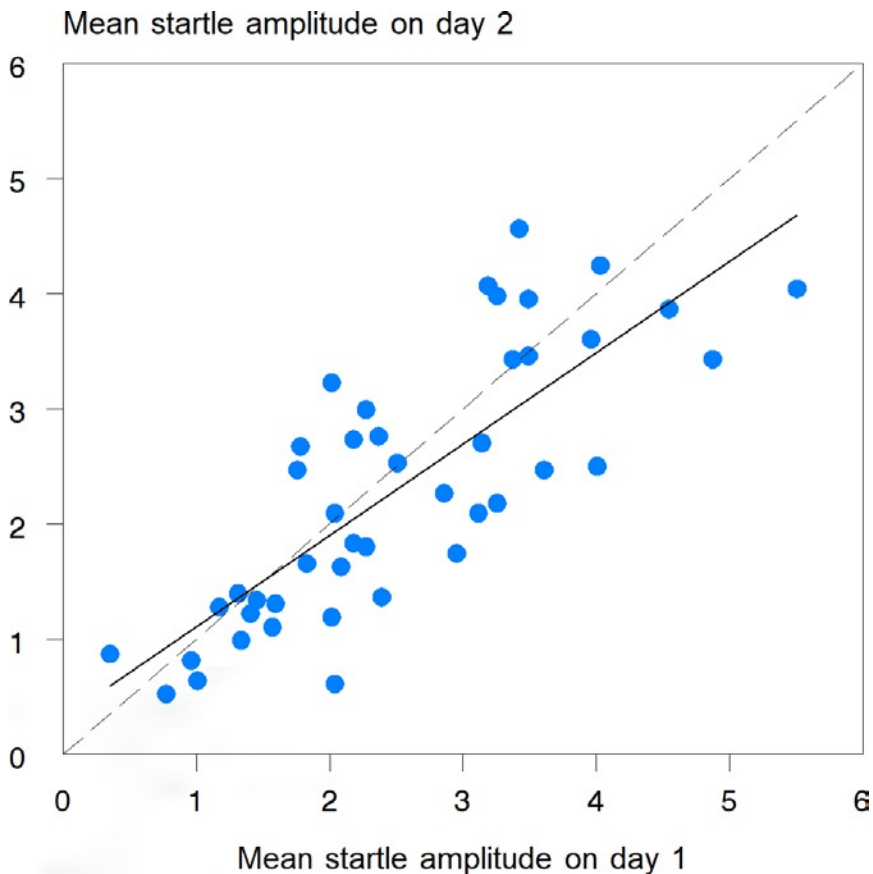


Fig. 4.2. Acoustic startle response (amplitude of jump following a 50-ms burst of 95 dB white noise; mean of 30 bursts per test) of individual rats ($n = 48$) tested on two consecutive days. The dashed line indicates where $y = x$ (i.e. slope is 1), and the slope of the best fit regression (solid line) is 0.8 (data from Russo and Parsons, 2017).

from post-traumatic stress disorder (PTSD), and the test is therefore often used in animal models of PTSD (e.g. Shalev *et al.*, 1997).

Human Approach Test

Just like the acoustic startle test, the human approach test is designed to provoke a reaction. The main measure is the minimum distance achieved between an approaching human and the test animal, i.e. how close can you get before the animal retreats. It is a test mostly used outside the rodent laboratory, and is suitable for animals housed in enclosures large enough for the human to begin their approach from a certain distance, and with enough space to allow the animal to move away from the approaching human. It differs from tests using a stationary human, somewhat similar to the novel object test, but where a human observer positions themselves passively inside the test arena, and the time taken by the animal to approach the human is measured. For this reason, the human approach test where the person moves is sometimes referred to as the forced human approach test or the withdrawal response test.

Whenever humans are involved as part of a behavioural test, the previous experience that the test animal has had with humans will influence the outcome (Waiblinger *et al.*, 2006). To an animal, an approaching unknown human can be seen as a threat or as something worth investigating. For this reason, among others, it has been proposed that the human approach test is a measure of willingness to take risks, rather than an estimate of fearfulness (Marchant-Forde, 2002). This may allow us to phenotype animals into behavioural types (also suggested for the back test presented below), which in turn may correlate with other traits. Chapter 10 includes an example of genetic selection of mink based on their responses to humans.

A combination of a stationary and approaching human has been found to predict aggressive behaviour in sows. Two months prior to farrowing, gilts ($n = 62$) were tested in a human approach test as follows (Marchant-Forde, 2002): Individual gilts were briefly habituated to a square test arena, after which an unfamiliar human entered the arena, and stood passively against the arena wall. After 3 min, the person approached and touched the snout of the gilt. An average shy/boldness score was calculated for each gilt, based on their relative ranking on five different variables relating to latency to approach and frequency of contact with the human, activity and vocalization as well as heart rate change when touched on the snout. Following the subsequent parturition, the degree of aggressiveness shown by the sows towards the stock person was scored on a 5-point scale on three occasions when the piglets were weighed. Five of the gilts savaged their piglets, and five other gilts showed high levels of aggression towards the stock person. This propensity for aggression was found to be highly correlated with the shy/boldness score from the human approach test: gilts

that were bolder in the human approach test were more likely to be aggressive towards the stock person, whereas the sows savaging their piglets were the shyest when tested as gilts. This suggests that extreme aggression in pigs towards humans and towards offspring, respectively, is predictable to a certain extent and stems from different behavioural characteristics. Results from a human approach test have also been found to correlate to behavioural reactivity in cattle, with cows that accepted being touched by the approaching human, being 3.5 times more likely to kick during milking (Rousing *et al.*, 2004).

Elevated Plus Maze

One method used to characterize the anxiety level of rodents is the elevated plus maze (EPM) test (Walf and Frye, 2007). The test is based on the propensity for nocturnal mice and rats to avoid brightly lit areas, and the equipment consist of a plus-shaped arena with four arms. Two of the arms are enclosed by high walls, whereas the remaining two arms are left completely open. The opposing arms are of the same type, and the whole test device is elevated from the ground to discourage the test animal from jumping off the open arms during testing. The EPM has different dimensions for tests of mice and rats, respectively, but the protocol is the same. In a fully lit room, the test animal is placed in the centre of the EPM where the four arms meet, and is left to explore for 5 min. Usually, entries into and time spent in each arm are registered from video recordings, but more detailed observations of the behaviour of the mouse or rat can be made. A version of this test has also been adapted for use in mink (Malmkvist and Hansen, 2002).

The behaviour of the animal in the EPM test, more specifically the proportion of time spent in the open arms, indicates its degree of anxiety, with less anxious animals spending the highest proportion of time in the open arms. In one of the early uses of the EPM, Lister (1987) found that mice treated with anxiolytics, i.e. drugs used to relieve anxiety, spent a higher proportion of time in the open arms than untreated conspecifics, whereas the opposite was found for mice treated with anxiogenics, i.e. drugs that induce anxiety. However, the test can also be used to characterize the animals before testing (e.g. Bombail *et al.*, 2018) to ensure that animals are allocated to subsequent treatments in a balanced way based on their degree of anxiousness. For a list of behavioural tests used to assess anxiety levels, mainly in rodents, see Cryan and Sweeney (2011).

Tonic Immobility

As will become evident throughout this book, animal behaviour tests most often take their outset, not surprisingly, in the natural behaviour of the

animal tested. This was evident from the EPM, which would not be particularly useful in, say, pigs: not because of the impracticality of building a maze of sufficient sturdiness, but because pigs are diurnal and therefore not innately afraid of visiting the open arms. However, what pigs do share with a number of other species, most notably birds but also sharks (Watsky and Gruber, 1990), is the ability to enter a state of tonic immobility (TI). As the name indicates, this is a state of immobility, sometimes referred to as death feigning, and it is seen in nature when birds are faced with imminent danger: the animal enters what resembles a catatonic state, with no or little movement and a relaxed muscle tone. Although dated, a good review on tonic immobility in the domestic fowl has been written by Bryan Jones (1986).

In the laboratory, the test is carried out using a table (for birds) or a V-shaped crate (for piglets; Fig. 4.3). The experimenter turns the animal on its back and places a light weight on the animal's chest, either by placing a hand there, or using a small sand-filled bag. Sometimes several attempts to induce TI are needed, and a good experimental protocol will have a maximum limit for the number of induction attempts. Once the animal enters into TI, the duration until the animal attempts to right itself from the prone position is measured. In the case of chickens, if the bird does not attempt to right itself within 5 min, the experimenter will gently turn the bird over and ensure that it regains full mobility. This maximum period of 600s is most likely chosen because it gives a sufficient spread of variation in the TI duration, without causing potential discomfort to the animal tested. Chickens are able to remain in TI for very long periods, with Gallup (1977) reporting a bird staying in TI for more than 5 h. The propensity to enter TI is closely linked to the duration of TI, as Japanese quail genetically selected for short duration of TI also require a higher number of attempts to induce



Fig. 4.3. Example of a V-shaped crate used to induce tonic immobility (TI) in piglets (photo: Alexandra Courty).

TI (Mills and Faure, 1991). In piglets, the test is often referred to as the back test, where either TI is induced (Erhard *et al.*, 1999) or the piglet is gently restrained on its back for 1 min and the number of escape attempts counted (Bolhuis *et al.*, 2005).

One of the major points of discussion associated with the use of TI tests in piglets has been whether or not animals could be categorized according to their response when submitted to this test. This idea arose from the finding that piglets which readily responded to TI induction, i.e. few induction attempts needed and long TI duration, were different in their overall personality to piglets where TI was not easily induced (Erhard *et al.*, 1999; Erhard and Mendl, 1999). Pigs show different types of behavioural strategies (Mendl *et al.*, 1992) and these were related to the ability to cope in different situations (Hessing *et al.*, 1993), with pigs being either active or passive copers (Hessing *et al.*, 1994). Proof of coping strategies has been found in rodents (Jensen, 1995), but not pigs (Forkman *et al.*, 1995), whereas Mendl and Deag (1995) took a more pragmatic view but acknowledged the limitations of the concept. In their recent review of personality testing in pigs, O'Malley *et al.* (2019) report that the back test has been used in two out of three studies so far, although it is unclear what aspects of personality are actually being measured.

This is not the place for a prolonged review of coping strategies in pigs, and the published aspects of the discussion appear to have subsided, but the subject is nevertheless fascinating. If the TI test cannot characterize the animal and predict with some degree of certainty how the animal is likely to behave in other situations, why would we use it? Like many behavioural tests, sometimes TI has a relationship to many other behavioural traits (Bolhuis *et al.*, 2005), sometimes to only a few (Boersma *et al.*, 2017). It is not always possible to disentangle the genetic and environmental causes. Courty and O'Driscoll (2018) found that piglets born from gilts characterized as fearful, neutral or friendly based on their response to an approaching human also tended to differ divergently in their response to various behavioural tests, including a back test, an open field test and a human approach test. However, to what extent these differences are due to inherited traits or correlated differences in the maternal behaviour of the gilts is unknown. In a recent review of 83 articles on personality testing in pigs, O'Malley *et al.* (2019) lament that, although personality may be as heritable as some other porcine traits used in commercial breeding, there is a lack of standardization, reliability and validity of measures used to assess personality in pigs. Adult personality of junglefowl, for example, can only be predicted to a degree from behavioural responses early in life (Favati *et al.*, 2016; Zidar *et al.*, 2017). In carp, however, MacKenzie *et al.* (2009) found that gene expression for cortisol receptor and other genes only differ between carp exposed to either a challenge or a control treatment if the coping style of the fish is taken into account (Fig. 4.4). In this study, a carp was characterized as proactive or bold if it was among the first

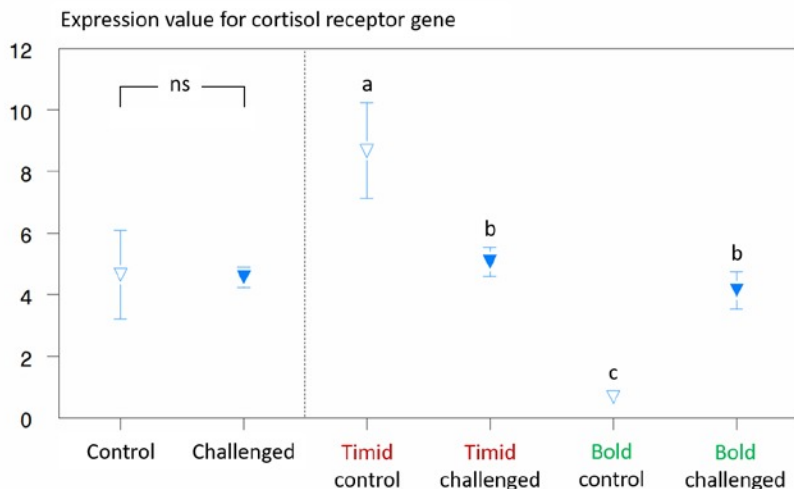


Fig. 4.4. When carp are exposed to a challenge with an inflammatory agent (lipopolysaccharide *E. coli*), no differences in expression (mean $\pm$ SE) were found for cortisol receptor gene (left-hand panel). When the same data were analysed taking into account the coping style of the fish in terms of risk-taking in a novel environment, significant differences were found in gene expression in the control group, with timid fish showing higher expression values than the bold ones (right-hand panel). This revealed significantly different responses to the challenge between the two types of fish, with timid carp showing a reduction and bold carp showing an increase in cortisol receptor gene expression when challenged (adapted from MacKenzie *et al.*, 2009).

three fish (out of a randomly selected group of ten carp) to emerge into a novel environment, whereas the last three carp left behind were regarded as reactive or timid.

Tests at Group Level and Social Encounters

Resident–intruder test

One aspect of an animal's personality is its propensity to display aggression towards conspecifics. Erhard and Mendl (1997) investigated if aggressiveness of pigs could be ascertained by testing them in a resident–intruder test. This test was first developed for use in rodents (Thomas, 1973). In the version used for pigs, it consisted of the animal to be tested (the resident) housed in one half of its home pen, into which a smaller conspecific (the intruder) was introduced. The measurement used was the time taken from first contact between the pigs until the resident pig attacked the intruder, after which point the test was immediately stopped, and the intruder pig removed from the pen. An attack was defined as at least a bite, so for example

chewing the intruder was not considered an attack. If no attack occurred, the test was terminated after 3.5 min. Erhard and Mendl (1997) wanted to know if the latency to attack was a consistent characteristic of the test animal (repeatable) and independent of other factors such as age and sex. They used data from 436 tests using 218 resident pigs from 23 different litters. In only 13 tests (3%) did the intruder initiate the attack. Although pigs generally attacked faster in the second test, the attack latencies were highly correlated between tests, indicating a promising degree of stability of this measure within individuals. No effects of age or sex were found on aggressiveness of the resident pigs either, but there were indications that the likelihood of attack was reduced if the intruder was very small (less than half the weight of the resident). Overall, the distribution of attack latencies was bimodal, with fast attackers and non-attackers divided by a much smaller fraction of intermediate types. The authors concluded that this measure could be used to assess the level of aggressiveness of a pig without compromising the welfare of the animals used in the test.

Others have found that an aggressiveness score based on two resident-intruder tests is correlated with the persistency of aggression in groups of pigs after mixing, confirming that attack latency can be used as a proxy for aggressiveness (D'Eath, 2002). However, the simplicity of the test has been questioned, as latency to first attack may not accurately reflect the level of aggressiveness of the animal tested, and Camerlink *et al.* (2016) recommend more detailed recording of the behaviours seen before an attack to remedy this. Mostly applied to rodents, a version of the test, where the encounter is not interrupted at first attack, is used to establish chronic social defeat, inducing depression in the defeated animal (Berton *et al.*, 2006). The ethical implications of this are discussed in Chapter 12.

Tests of hierarchy

In groups of animals, a stable social hierarchy minimizes aggression as access to resources, such as food and reproductive females, is dominated by the high-ranking individuals. When unfamiliar domestic animals are housed in groups, they will initially fight to establish a dominance hierarchy. This will remain relatively stable with little need to ascertain rank order among the group members unless major changes occur, such as removal of the most dominant animal, or introduction of new group members. It is important to keep in mind that dominance is always a relative measure, and not an individual characteristic (Drews, 1993). When we want to assess the social hierarchy in a group of animals, we often base this on observations of pairwise (dyadic) interactions. These can be provoked by offering a high-value resource, such as feed, in a limited quantity, leading to competition for access to the resource and displacement of one individual by another, thereby indicating their relative rank (Parent *et al.*, 2012).

In rodents, a tube test can be used to establish the direction of social rank between two individuals. Using a narrow transparent tube, two mice enter the tube from opposite ends, and the mouse that is able to force the other mouse backwards out of the tube is deemed to be the dominant, whereas a retreating mouse is subordinate to its opponent. This has been found to be a very stable way of assessing dyadic dominance, with 86% of pairs maintaining their relative rank when tested on two consecutive days (Fan *et al.*, 2019). In addition, these experimenters found that if mouse A forces mouse B out of the tube, and mouse B wins over mouse C, then there is a 95% chance that mouse C will retreat when facing mouse A. Linear hierarchies, such as this, are not always applicable in social groups, as circular dominance relationships may exist among some group members. Linear or near-linear hierarchies also have a high probability of appearing by chance if results between some dyads are missing (Appleby, 1983). Indeed, if the hierarchy assessment is based on spontaneous fights, it is unlikely that all animals in a group have been engaged in the same number of agonistic interactions, which may skew the assessment of certain rankings (Shev *et al.*, 2014).

Cognitive Bias

We all have among our acquaintances someone who's a super optimist, or someone who always appears to find the negative side of any issue. Such propensity to reveal your inner mood by seeing the glass as either half-full or half-empty has been exploited in human behavioural tests. Test persons are presented with unambiguous (clearly positive or clearly negative) sentences as well as ambiguous ones, and clinically anxious people are more likely to interpret the ambiguous sentences in a negative way (Eysenck *et al.*, 1991). Using this concept of cognitive bias, Mike Mendl and Liz Paul from the University of Bristol developed a test paradigm for use in animals (Harding *et al.*, 2004; Mendl and Paul, 2004; Paul *et al.*, 2005). Presenting an animal with two clearly distinguishable cues and letting the animal learn to associate each cue with a positive and negative stimulus, respectively, makes it possible to test the response of the animal when ambiguous cues are presented. These ambiguous cues are always mixtures of the two initial cues, and it is this stroke of genius that makes this judgement bias paradigm so adaptable. It not only allows the use of many different cue types and species, such as auditory notes in mink (Svendsen *et al.*, 2012), colour plates in laying hens (Hernandez *et al.*, 2015) and left/right positioning in dogs (Mendl *et al.*, 2010a), it also makes it possible to quantify the amount of bias by mixing the two cues to varying degrees (Paul *et al.*, 2005).

In the following example, Doyle *et al.* (2011) tested if sheep given injections of a serotonin inhibitor would respond differently in a judgement

bias test, as low serotonin levels are associated with mood disorders and irritability in humans (Young and Leyton, 2002). All the sheep were trained to associate a bucket with a positive reinforcer (a feed reward) when the bucket was placed in one corner of the test arena and to associate the bucket with a negative reinforcer (exposure to a dog) when placed in the alternate corner (Fig. 4.5). This meant that the sheep would readily approach the bucket on one side, but would not (or only reluctantly) approach the bucket when it was placed in the other corner. Subsequently, half of the sheep were injected twice daily with a serotonin inhibitor (pCPA; p-chlorophenylalanine) in a dose that had previously been shown not to affect locomotion or feeding motivation of the sheep compared to a control group. The sheep were then tested with the bucket placed at intermediate positions between the two corners (Fig. 4.5). After 5 days of pCPA treatment, the sheep were significantly less likely than control sheep to approach the bucket when it was placed in the 50 and 60% positive position. This result illustrates that judgement bias tests can be a means to demonstrate mood differences, even in a livestock species such as sheep. In this and many other cases, the pre-test condition of a negative or positive affective state is created as part of the experimental design. Others have used judgement bias tests to show that dogs with separation-related behavioural problems, such as destroying things when left alone, were of a more negative disposition (Mendl *et al.*, 2010a).

As always, it is important to ensure that the test conditions do not influence the outcome, giving rise to unexpected results. Freymond *et al.* (2014) trained horses using either positive or negative reinforcement (see Chapter 8) to see if the training methods gave rise to different affective states. When the horses were tested in a judgement bias test with a baited and an empty feed bucket at the two extremes, the horses trained with negative reinforcement (i.e. the removal of an unpleasant stimulus when the desired behaviour is displayed) showed the most optimistic judgement, which was not the result expected by the authors. They suggest that the use of food as positive reinforcement during training may have created a difference in feeding motivation between the two groups of horses, giving rise to the surprising results in the judgement bias test. In a different trial, Asher *et al.* (2016) found that the housing conditions of pigs interacted with their personality to influence how the pigs responded in a judgement bias test. Other examples like these exist (Monk *et al.*, 2019), and one can only assume that many more tests with unexpected results have been left unpublished. Care therefore needs to be taken when designing the details of these types of tests. This is to prevent other biases than the one tested from affecting the judgement made by the test animal.

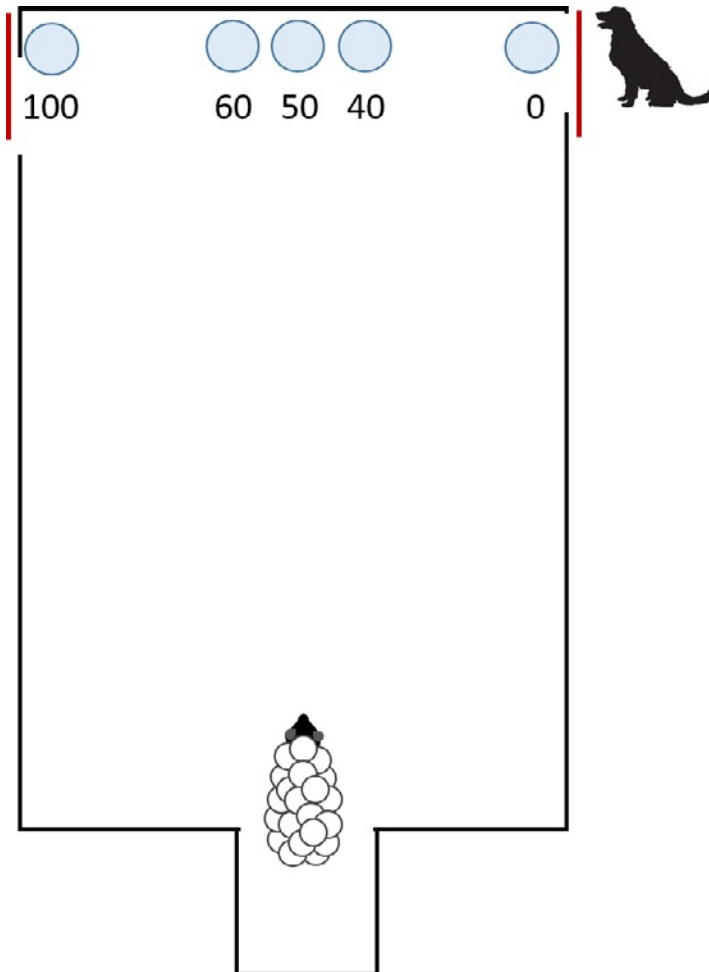


Fig. 4.5. Diagram of arena (2 m × 3 m) used for judgement bias test of sheep. The blue circles represent placement of a bucket, and only one bucket was present during a test. The red lines indicate sliding panels. The animals first learnt to associate the placement of the bucket in position 100 with a positive stimulus (a feed reward), whereas a bucket placed in position 0 was associated with a negative stimulus (presence of a dog). Half of the sheep were always presented with the dog on the left side, and the other half had the dog on the right. The sheep learnt to approach the bucket when in position 100, and not to approach when in position 0. For the judgement bias test, the bucket was placed in intermediate positions between the 100% positive and 0% positive, corresponding to 40, 50 and 60% positive. Sheep treated for 5 days with a serotonin inhibitor known to affect mood in humans were less likely than untreated sheep to approach the bucket when it was in positions 50 or 60 (adapted from Doyle *et al.*, 2011).

References

- Appleby, M.C. (1983) The probability of linearity in hierarchies. *Animal Behaviour* 31(2), 600–608. DOI: 10.1016/S0003-3472(83)80084-0.
- Asher, L., Friel, M., Griffin, K. and Collins, L.M. (2016) Mood and personality interact to determine cognitive biases in pigs. *Biology Letters* 12(11), 20160402. DOI: 10.1098/rsbl.2016.0402.
- Berton, O., McClung, C.A., Dileone, R.J., Krishnan, V., Renthal, W. *et al.* (2006) Essential role of BDNF in the mesolimbic dopamine pathway in social defeat stress. *Science* 311(5762), 864–868. DOI: 10.1126/science.1120972.
- Boersma, G.J., Smeltzer, M.D., Scott, K.A., Scheurink, A.J., Tamashiro, K.L. *et al.* (2017) Stress coping style does not determine social status, but influences the consequences of social subordination stress. *Physiology & Behavior* 178, 126–133. DOI: 10.1016/j.physbeh.2016.12.041.
- Boissy, A. (1995) Fear and fearfulness in animals. *The Quarterly Review of Biology* 70(2), 165–191. DOI: 10.1086/418981.
- Boissy, A., Dwyer, C.M. and Jones, R.B. (2018) Fear and other negative emotions. In: Appleby, M.C., Olsson, I.A.S. and Galindo, F. (eds) *Animal Welfare*, 3rd edn. CAB International, Wallingford, UK, pp. 92–113.
- Bolhuis, J.E., Schouten, W.G.P., Schrama, J.W. and Wiegant, V.M. (2005) Individual coping characteristics, aggressiveness and fighting strategies in pigs. *Animal Behaviour* 69(5), 1085–1091. DOI: 10.1016/j.anbehav.2004.09.013.
- Bombail, V., Barret, B., Raynaud, A., Jérôme, N., Saint-Albin, A. *et al.* (2018) In search of stress odours across species: behavioural responses of rats to faeces from chickens and rats subjected to various types of stressful events. *Applied Animal Behaviour Science* 205, 216–226. DOI: 10.1016/j.applanim.2017.10.013.
- Camerlink, I., Arnott, G., Farish, M. and Turner, S.P. (2016) Complex contests and the influence of aggressiveness in pigs. *Animal Behaviour* 121, 71–78. DOI: 10.1016/j.anbehav.2016.08.021.
- Courty, A. and O'Driscoll, K. (2018) Like mother like piglet: a pilot investigation into the relationship between maternal and piglet traits. *Proceedings of the ISAE UK & Ireland Regional Meeting*, 7 February 2018, Bristol, UK.
- Cryan, J.F. and Sweeney, F.F. (2011) The age of anxiety: role of animal models of anxiolytic action in drug discovery. *British Journal of Pharmacology* 164(4), 1129–1161. DOI: 10.1111/j.1476-5381.2011.01362.x.
- D'Eath, R.B. (2002) Individual aggressiveness measured in a resident-intruder test predicts the persistence of aggressive behaviour and weight gain of young pigs after mixing. *Applied Animal Behaviour Science* 77(4), 267–283. DOI: 10.1016/S0168-1591(02)00077-1.
- Dalmau, A., Mainau, E. and Velarde, A. (2017) Reliability of fear assessment in growing pigs exposed to a novel object test in commercial conditions. *Journal of Applied Animal Welfare Science* 20(3), 280–288. DOI: 10.1080/10888705.2017.1310043.
- Dingemanse, N.J., Bouwman, K.M., van de Pol, M., van Overveld, T., Patrick, S.C. *et al.* (2012) Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*. *Journal of Animal Ecology* 81(1), 116–126. DOI: 10.1111/j.1365-2656.2011.01877.x.
- Doyle, R.E., Hinch, G.N., Fisher, A.D., Boissy, A., Henshall, J.M. *et al.* (2011) Administration of serotonin inhibitor p-chlorophenylalanine induces

- pessimistic-like judgement bias in sheep. *Psychoneuroendocrinology* 36(2), 279–288. DOI: 10.1016/j.psyneuen.2010.07.018.
- Drews, C. (1993) The concept and definition of dominance in animal behaviour. *Behaviour* 125(3–4), 283–313. DOI: 10.1163/156853993X00290.
- Erhard, H.W. and Mendl, M. (1997) Measuring aggressiveness in growing pigs in a resident-intruder situation. *Applied Animal Behaviour Science* 54(2–3), 123–136. DOI: 10.1016/S0168-1591(97)00069-5.
- Erhard, H.W. and Mendl, M. (1999) Tonic immobility and emergence time in pigs—more evidence for behavioural strategies. *Applied Animal Behaviour Science* 61(3), 227–237. DOI: 10.1016/S0168-1591(98)00196-8.
- Erhard, H.W., Mendl, M. and Christiansen, S.B. (1999) Individual differences in tonic immobility may reflect behavioural strategies. *Applied Animal Behaviour Science* 64(1), 31–46. DOI: 10.1016/S0168-1591(99)00028-3.
- Erhard, H.W., Boissy, A., Rae, M.T. and Rhind, S.M. (2004) Effects of prenatal undernutrition on emotional reactivity and cognitive flexibility in adult sheep. *Behavioural Brain Research* 151(1–2), 25–35. DOI: 10.1016/j.bbr.2003.08.003.
- Eysenck, M.W., Mogg, K., May, J., Richards, A. and Mathews, A. (1991) Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology* 100(2), 144–150. DOI: 10.1037/0021-843X.100.2.144.
- Fan, Z.X., Zhu, H., Zhou, T.T., Wang, S., Wu, Y. *et al.* (2019) Using the tube test to measure social hierarchy in mice. *Nature Protocols* 14(3), 819–831. DOI: 10.1038/s41596-018-0116-4.
- Favati, A., Zidar, J., Thorpe, H., Jensen, P. and Løvlie, H. (2016) The ontogeny of personality traits in the red junglefowl, *Gallus gallus*. *Behavioral Ecology* 27(2), 484–493. DOI: 10.1093/beheco/arv177.
- Forkman, B., Furuhaug, I.L. and Jensen, P. (1995) Personality, coping patterns, and aggression in piglets. *Applied Animal Behaviour Science* 45(1–2), 31–42. DOI: 10.1016/0168-1591(95)00601-N.
- Forkman, B., Boissy, A., Meunier-Salaün, M.-C., Canali, E. and Jones, R.B. (2007) A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiology & Behavior* 92(3), 340–374. DOI: 10.1016/j.physbeh.2007.03.016.
- Freymond, S.B., Briefer, E.F., Zollinger, A., Gindrat-von Allmen, Y., Wyss, C. *et al.* (2014) Behaviour of horses in a judgment bias test associated with positive or negative reinforcement. *Applied Animal Behaviour Science* 158, 34–45. DOI: 10.1016/j.applanim.2014.06.006.
- Gallup, G.G. Jr. (1977) Tonic immobility: the role of fear and predation. *The Psychological Record* 27(S1), 41–61. DOI: 10.1007/BF03394432.
- Hall, C.S. (1934) Emotional behavior in the rat. I. defecation and urination as measures of individual differences in emotionality. *Journal of Comparative Psychology* 18(3), 385–403. DOI: 10.1037/h0071444.
- Harding, E.J., Paul, E.S. and Mendl, M. (2004) Cognitive bias and affective state. *Nature* 427(6972), 312. DOI: 10.1038/427312a.
- Hernandez, C.E., Hinch, G., Lea, J., Ferguson, D. and Lee, C. (2015) Acute stress enhances sensitivity to a highly attractive food reward without affecting judgement bias in laying hens. *Applied Animal Behaviour Science* 163, 135–143. DOI: 10.1016/j.applanim.2014.12.002.
- Herskin, M.S. and Jensen, K.H. (2000) Effects of different degrees of social isolation on the behaviour of weaned piglets kept for experimental purposes. *Animal Welfare* 9, 237–249.

- Hessing, M.J.C., Hagelsø, A.M., van Beek, J.A.M., Wiepkema, R.P., Schouten, W.G.P. *et al.* (1993) Individual behavioural characteristics in pigs. *Applied Animal Behaviour Science* 37(4), 285–295. DOI: 10.1016/0168-1591(93)90118-9.
- Hessing, M.J.C., Hagelsø, A.M., Schouten, W.G.P., Wiepkema, P.R. and Van Beek, J.A.M. (1994) Individual behavioral and physiological strategies in pigs. *Physiology & Behavior* 55(1), 39–46. DOI: 10.1016/0031-9384(94)90007-8.
- Jensen, P. (1995) Individual variation in the behaviour of pigs—noise or functional coping strategies? *Applied Animal Behaviour Science* 44(2–4), 245–255. DOI: 10.1016/0168-1591(95)00617-2.
- Jones, R.B. (1986) The tonic immobility reaction of the domestic fowl: a review. *World's Poultry Science Journal* 42(1), 82–96. DOI: 10.1079/WPS19860008.
- Jones, R.B., Satterlee, D.G. and Ryder, F.H. (1992) Research note: open-field behavior of Japanese quail chicks genetically selected for low or high plasma corticosterone response to immobilization stress. *Poultry Science* 71(8), 1403–1407. DOI: 10.3382/ps.0711403.
- Lister, R. (1987) The use of a plus-maze to measure anxiety in the mouse. *Psychopharmacology* 92(2), 180–185. DOI: 10.1007/BF00177912.
- MacKay, J.R.D. (2018) *Animal Personality: The Science Behind Individual Variation*. 5M Publishing Ltd, Sheffield, UK.
- MacKenzie, S., Ribas, L., Pilarczyk, M., Capdevila, D.M., Kadri, S. *et al.* (2009) Screening for coping style increases the power of gene expression studies. *PLoS ONE* 4(4), e5314. DOI: 10.1371/journal.pone.0005314.
- Malmkvist, J. and Hansen, S.W. (2002) Generalization of fear in farm mink, *Mustela vison*, genetically selected for behaviour towards humans. *Animal Behaviour* 64(3), 487–501. DOI: 10.1006/anbe.2002.3058.
- Marchant-Forde, J.N. (2002) Piglet- and stockperson-directed sow aggression after farrowing and the relationship with a pre-farrowing, human approach test. *Applied Animal Behaviour Science* 75(2), 115–132. DOI: 10.1016/S0168-1591(01)00170-8.
- Mazzamuto, M.V., Cremonesi, G., Santicchia, F., Preatoni, D., Martinoli, A. *et al.* (2019) Rodents in the arena: a critical evaluation of methods measuring personality traits. *Ethology Ecology & Evolution* 31(1), 38–58. DOI: 10.1080/03949370.2018.1488768.
- Mendl, M. and Deag, J.M. (1995) How useful are the concepts of alternative strategy and coping strategy in applied studies of social behaviour? *Applied Animal Behaviour Science* 44(2–4), 119–137. DOI: 10.1016/0168-1591(95)00609-V.
- Mendl, M. and Paul, E.S. (2004) Consciousness, emotion and animal welfare: insights from cognitive science. *Animal Welfare* 13, 17–25.
- Mendl, M., Zanella, A.J. and Broom, D.M. (1992) Physiological and reproductive correlates of behavioural strategies in female domestic pigs. *Animal Behaviour* 44(6), 1107–1121. DOI: 10.1016/S0003-3472(05)80323-9.
- Mendl, M., Brooks, J., Basse, C., Burman, O., Paul, E. *et al.* (2010a) Dogs showing separation-related behaviour exhibit a ‘pessimistic’ cognitive bias. *Current Biology* 20(19), R839–R840. DOI: 10.1016/j.cub.2010.08.030.
- Mendl, M., Burman, O.H.P. and Paul, E.S. (2010b) An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences* 277(1696), 2895–2904. DOI: 10.1098/rspb.2010.0303.

- Mills, A.D. and Faure, J.-M. (1991) Divergent selection for duration of tonic immobility and social reinstatement behavior in Japanese quail (*Coturnix coturnix japonica*) chicks. *Journal of Comparative Psychology* 105(1), 25–38. DOI: 10.1037/0735-7036.105.1.25.
- Monk, J.E., Lee, C., Belson, S., Colditz, I.G. and Campbell, D.L.M. (2019) The influence of pharmacologically-induced affective states on attention bias in sheep. *PeerJ* 7(1), e7033. DOI: 10.7717/peerj.7033.
- Nawroth, C., Prentice, P.M. and McElligott, A.G. (2017) Individual personality differences in goats predict their performance in visual learning and non-associative cognitive tasks. *Behavioural Processes* 134, 43–53. DOI: 10.1016/j.beproc.2016.08.001.
- O'Malley, C.I., Turner, S.P., D'Eath, R.B., Steibel, J.P., Bates, R.O. *et al.* (2019) Animal personality in the management and welfare of pigs. *Applied Animal Behaviour Science* 218, 104821. DOI: 10.1016/j.applanim.2019.06.002.
- Parent, J.-P., Meunier-Salaün, M.-C., Vasseur, E. and Bergeron, R. (2012) Stability of social hierarchy in growing female pigs and pregnant sows. *Applied Animal Behaviour Science* 142(1–2), 1–10. DOI: 10.1016/j.applanim.2012.09.011.
- Paul, E.S., Harding, E.J. and Mendl, M. (2005) Measuring emotional processes in animals: the utility of a cognitive approach. *Neuroscience & Biobehavioral Reviews* 29(3), 469–491. DOI: 10.1016/j.neubiorev.2005.01.002.
- Rousing, T., Bonde, M., Badsberg, J.H. and Sørensen, J.T. (2004) Stepping and kicking behaviour during milking in relation to response in human–animal interaction test and clinical health in loose housed dairy cows. *Livestock Production Science* 88(1–2), 1–8. DOI: 10.1016/j.livprodsci.2003.12.001.
- Russo, A.S. and Parsons, R.G. (2017) Acoustic startle response in rats predicts inter-individual variation in fear extinction. *Neurobiology of Learning and Memory* 139, 157–164. DOI: 10.1016/j.nlm.2017.01.008.
- Shalev, A.Y., Peri, T., Orr, S.P., Bonne, O. and Pitman, R.K. (1997) Auditory startle responses in help-seeking trauma survivors. *Psychiatry Research* 69(1), 1–7. DOI: 10.1016/S0165-1781(96)03001-6.
- Shev, A., Fujii, K., Hsieh, F. and McCowan, B. (2014) Systemic testing on Bradley-Terry model against nonlinear ranking hierarchy. *PLoS ONE* 9(12), e115367. DOI: 10.1371/journal.pone.0115367.
- Simon, P., Dupuis, R. and Costentin, J. (1994) Thigmotaxis as an index of anxiety in mice. Influence of dopaminergic transmissions. *Behavioural Brain Research* 61(1), 59–64. DOI: 10.1016/0166-4328(94)90008-6.
- Sneddon, L.U., Braithwaite, V.A. and Gentle, M.J. (2003) Novel object test: examining nociception and fear in the rainbow trout. *The Journal of Pain* 4(8), 431–440. DOI: 10.1067/S1526-5900(03)00717-X.
- Svendsen, P.M., Malmkvist, J., Halekoh, U. and Mendl, M. (2012) Responses of mink to auditory stimuli: prerequisites for applying the 'cognitive bias' approach. *Behavioural Processes* 91(3), 291–297. DOI: 10.1016/j.beproc.2012.09.010.
- Thomas, T.R. (1973) Use of a resident-intruder paradigm to measure the influence of environmental adaptation on mounting behavior. *The Psychological Record* 23(3), 405–416. DOI: 10.1007/BF03394185.
- Waiblinger, S., Boivin, X., Pedersen, V., Tosi, M.-V., Janczak, A.M. *et al.* (2006) Assessing the human–animal relationship in farmed species: a critical review. *Applied Animal Behaviour Science* 101(3–4), 185–242. DOI: 10.1016/j.applanim.2006.02.001.

- Walf, A.A. and Frye, C.A. (2007) The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. *Nature Protocols* 2(2), 322–328. DOI: 10.1038/nprot.2007.44.
- Walsh, R.N. and Cummins, R.A. (1976) Open-field test – critical-review. *Psychological Bulletin* 83, 482–504.
- Watsky, M.A. and Gruber, S.H. (1990) Induction and duration of tonic immobility in the lemon shark, *Negaprion brevirostris*. *Fish Physiology and Biochemistry* 8(3), 207–210. DOI: 10.1007/BF00004459.
- Young, S.N. and Leyton, M. (2002) The role of serotonin in human mood and social interaction: insight from altered tryptophan levels. *Pharmacology Biochemistry and Behavior* 71, 857–865.
- Zidar, J., Sorato, E., Malmqvist, A.-M., Jansson, E., Rosher, C. *et al.* (2017) Early experience affects adult personality in the red junglefowl: a role for cognitive stimulation? *Behavioural Processes* 134, 78–86. DOI: 10.1016/j.beproc.2016.06.003.

Choice, Preference and Motivation

Choice Tests

At first glance, the concept of a choice test appears deceptively simple: you present the animal with two different options and observe which option the animal chooses. What could possibly go wrong? Well, on closer inspection, a number of potential problems emerge. First, we need to make sure that we are comparing like with like. The basis of a choice test is to reveal the animal's motivation to seek one resource above another, and if the paired options differ greatly, the choice made loses any interpretative value (Correa *et al.*, 2016). For example, if a male rat is given the choice between a highly palatable food source and a female rat in oestrus, he will most likely choose the latter. However, this will be highly dependent on his state of hunger, as well as his previous experience with female rats, and choosing sex over food in the test situation does not mean that this is a preference applicable across situations.

Even when comparing the choice of two feed sources to test for preference, the choices should differ in only one dimension – if they have different nutritional value, the form and size of the feed should be the same. If different colour feeds are used during training, these should be balanced across both types of feed. Wadhera *et al.* (2018) found that rats chose feed presented as 30 smaller pellets over one, large pellet with the same total weight and nutritive value (Fig. 5.1). As this choice differed in at least three dimensions (numerosity, surface area and density), these authors used two additional comparisons to disentangle the importance of these different aspects. To test for the effect of surface area, they compared the rats' preferences for the same amount of mashed potato presented either in a pancake shape (large surface area) or as a sphere (small surface area). The rats preferred the flattened mash, thus choosing the largest surface area. To investigate the relative importance of numerosity and density, which are negatively correlated, the rats were given a choice of 15 pellets either spread out or presented in a cluster. The rats showed no preference between these

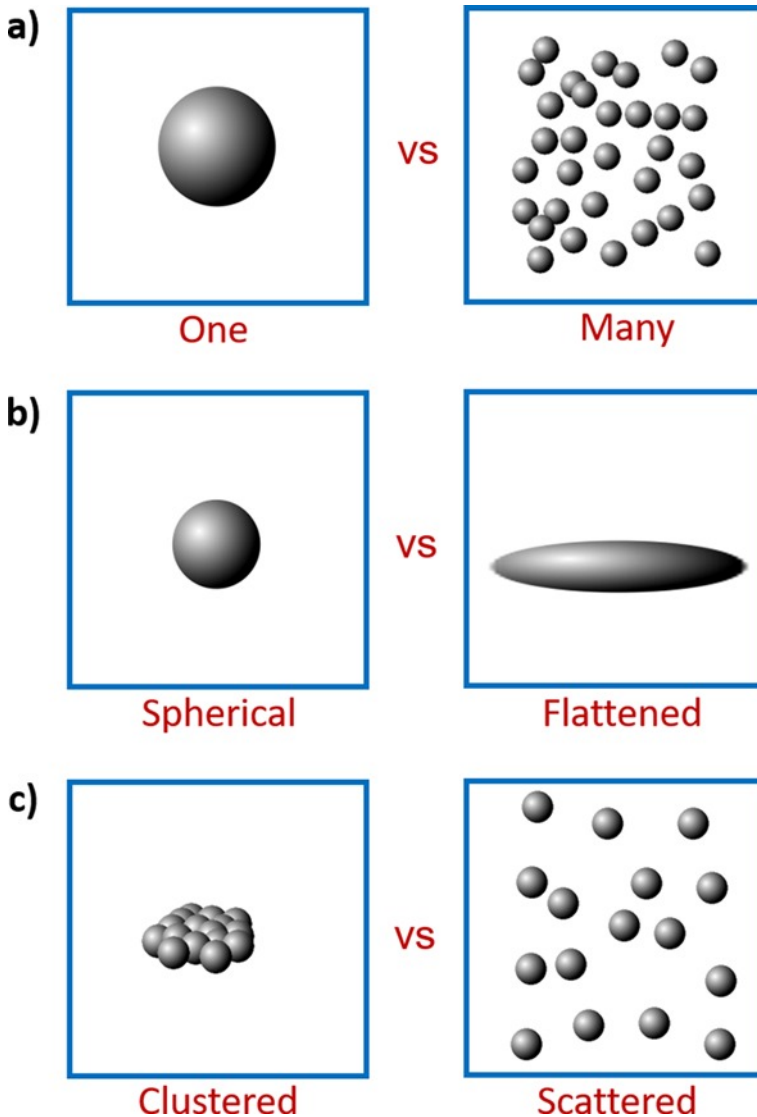


Fig. 5.1. Rats were given a choice of (a) one, large pellet with the same total weight and nutritive value as 30 smaller pellets (one vs many); (b) the same amount of mashed potato presented either as a sphere or a pancake shape (spherical vs flattened); and (c) 15 pellets either presented in a cluster or spread out (clustered vs scattered) (adapted from Wadhera *et al.*, 2018).

two options, indicating that they were able to take food density into account, so that the preferences for high numerosity and large surface area are diminished when the space between food items is large.

Other factors need to be considered when asking an animal to choose between different options. Many choice tests are carried out in Y-mazes or T-mazes, so called because their shape from above resembles the letters Y and T. The animal is placed in the lower part of the letter and, in the Y-maze, will be able to see the contents of the two arms before setting off to make a choice. In the T-maze, the choice to go left or right is made at the intersection, and if the arms of the maze have angled extensions, no visual information is available prior to choosing. The choice is then based on olfactory and perhaps auditory clues, as well as previous experience. A well-known example of the latter is the experiments by Marian Dawkins (1976, 1977), in which she showed that hens raised in battery cages were more likely to choose a battery cage over an open hen run than were hens with previous experience with the larger enclosure.

An important aspect when using Y- and T-tests is laterality, as some animals have a propensity to always go to one side, independent of the option offered (Adámková *et al.*, 2017). In some species, such as cats, a handedness can even be detected, with males showing a preference for using their left paw and the opposite preference found in female cats (McDowell *et al.*, 2018). To account for pre-existing laterality, the animals can be pre-tested to quantify their individual degree of laterality, and the placement of the choice options in the arms should be interchanged, preferably in a non-systematic but balanced manner. The duration of access to the different options also needs to be taken into consideration, as the choice may differ if the animal is stuck with the chosen option for a long time (Hughes, 1976) or, inversely, if there is not enough time to benefit from the choice.

Preference or Avoidance?

A choice between two things can be a display of preference or attractiveness of the chosen option, but may also reflect avoidance of the alternative if this option is aversive to some degree. This can be the outcome when animals are tested in pairs or groups, if dominant individuals prevent subordinates from accessing the preferred option. One such example comes from zebrafish, which were given a choice – either in pairs or as a group – between two tanks containing different types and degrees of enrichment, ranging from barren tanks to floating plants and gravel (Schroeder *et al.*, 2014). When tested in pairs, the dominant zebrafish spent more time in and prevented the other, subordinate fish from entering the compartment preferred when groups were tested. Gravel was a popular choice over tanks that were barren or contained sand; even a picture of gravel on the bottom sufficed to make a compartment attractive.

A paradigm often used to test for preference and avoidance is the conditioned place preference (CPP) and conditioned place avoidance (CPA) tests. As the names indicate, a degree of learning is needed, where the

animal is taught (conditioned) to associate a certain place (or half of an arena) with either something positive/attractive or something negative/aversive. Following this training, when put into the arena without the associated stimulus, the animal is expected to show a preference for the arena half previously associated with either presence of a positive (in the case of CPP) or absence of a negative (in the case of CPA) stimulus. This type of test is based on the classical conditioning paradigm (also known as Pavlovian conditioning) where a neutral stimulus (such as the sound of a bell) is paired with a stimulus with either positive (e.g. food) or negative (e.g. puff of air) valence (Fig. 5.2). These two stimuli become associated to the extent that the sound of the bell elicits the same responses previously only seen when the positive or negative stimulus was presented, such as salivation or escape attempts, respectively. In the CPP/CPA tests, the previously neutral stimulus is simply substituted for a previously neutral place.

CPP tests have been extensively employed in studies of cocaine or alcohol dependence with rodents as the model animal (e.g. Katebi *et al.*, 2018;



Fig. 5.2. The principle of classical (or Pavlovian) conditioning. A stimulus (x) gives rise to a behavioural response. Before conditioning, another stimulus (y) does not elicit the same response. During conditioning, stimulus y is presented just before or together with stimulus x, which means that after conditioning has taken place, stimulus y now gives rise to the same behavioural response as stimulus x. Stimulus x is known as the unconditioned stimulus (US), and y as the conditioned stimulus (CS). Below the box are two examples of conditioning: the example from Pavlov, where dogs began to salivate at the sound of a bell, and another example of inadvertent conditioning, as an animal may begin to associate a visit to the veterinarian with an aversive experience, such as pain.

review by Tzschentke, 2007). The CPA concept has also been successfully applied to laying hens, where potentially aversive stimuli were tested in a four-chamber experimental set-up, but with only one aversive stimulus presented (Paul *et al.*, 2018). This was to prevent any pre-test chamber preferences from influencing the results, with the four chambers decorated with different colour and pattern combinations. The authors found that laying hens showed CPA to air puffs and water spray, but not visual exposure to a moving snake model, nor alarm calls from conspecifics. Others have tried to use a two-chamber set-up to test both aversive (air puffs, social isolation) and attractive (increasing amounts of feed) stimuli in feed-restricted broiler breeders (Dixon *et al.*, 2013). As described in more detail in Chapter 3, these birds are highly motivated to feed, being raised on a feed allowance well below their *ad libitum* intake. However, these experiments were consistently unsuccessful in inducing CPPs or aversions, as the birds appeared to be too hungry not to investigate any chamber for potential feed sources. It is rare that negative results are published, so this is an important example of the significant influence of feed deprivation on the use of preference/avoidance paradigms.

Multiple choices

Choice tests are also carried out using other experimental set-ups where a choice among more than two options is presented. These can take the form of radial mazes with many arms, often 4 or 8, but sometimes 12 or more. Together with hole-board tests, this type of equipment is also used in tests of memory (as in 'Where is the food reward?'), which is dealt with in Chapter 9. The use of a test arena with multiple compartments not only allows us to compare the animals' choices among different options simultaneously, but also represents a more robust experimental set-up in terms of interpretation of the results. In the hypothetical example mentioned in the introduction to this chapter, a male rat was asked to choose between food and copulation. However, if the animal was thirsty and thus more motivated to drink, any choice taken by the rat may not indicate a preference; indeed, if the choice is made fast it could indicate that the rat has learnt that the test will finish sooner if it chooses quickly.

A multiple choice test was used to investigate what type and brightness of light broilers prefer (Kristensen *et al.*, 2007). A central, square box with no light, no feed and no water was fitted with lightproof doors on each side, giving access to four different compartments with four different light sources. Each of the compartments had feed and water, they were kept at the same ambient temperature and humidity, and cleaned daily. Small groups of broilers ($n = 6$) were tested in this set-up at two different ages (1 week and 6 weeks of age), and half of the tests were done with dim light in all compartments (5 lux), with the other half offering bright lights (100 lux). The birds were kept in the test arena for a week,

with the first couple of days being the habituation period before usable data were collected. The broilers spent 62% of their time feeding and foraging, and showed no clear preference for a particular light source when 1 week old. However, the older birds spent most of their time in compartments with light resembling daylight (biolux) or with warm-white light, irrespective of the light intensity. In this study, the behaviour of the individual birds was registered using scan sampling every 15 min. It should be noted that when social animals are tested in groups, they are likely to influence each other's choices, mostly in the sense of spending more time in the same compartment as conspecifics. However, this can be taken into account by testing several groups, and observing the group dynamics in more detail. The influence of social factors should also be weighed against the negative aspects of testing social animals in isolation.

Another type of multiple choice test can be found in experimental set-ups using closed-economy designs. Here, the animal is kept in a housing system, where different chambers offer different environments with the animal free to move between chambers and remain in any chamber for any length of time. This is an excellent set-up for solitary species, such as mink, and Cooper and Mason (2000) investigated the choices made by mink when given free access to seven resource compartments accessible via a corridor from their home cage, the latter providing feed, water and a nest box. One of the compartments was empty, and the others contained one of the following: a box of hay, a water bath, a raised platform, a wire cylinder, a novel object and small toys. The mink spent most of their time in the home cage (31%) or in the hay compartment (38%), with 10% of their time spent in the corridor. The two most popular resources after hay were the bath (7%) and the novel object (6%), with the empty compartment being visited the least (<1%). The researchers then put weighted doors on the entry to the compartments as a way to impose an entrance fee for the mink. No cost was put on exiting a compartment, nor on entry to the home cage. By increasing the weight of the door, the mink had to use more effort to enter a given compartment, and this changed the time budget of the mink: heavier doors predictably meant fewer but longer visits to a compartment, and for all of the compartments the latency to interact with the resource fell as door weight increased. For some of the compartments, the mink also interacted more with the resource during the visits and, together with the reduced latency, this meant that even at the heaviest door weight, the time spent in the hay box, manipulating the novel objects, as well as swimming remained constant. As a consequence, the interaction with the remaining resources declined, indicating a shift in the behavioural priorities of the mink when asked to pay for access to different resources.

Contrafreeloading

If you have never heard of contrafreeloading, you may (like me) find the word strange, even borderline annoying. It covers a situation where animals will make an effort to gain access to feed even though the same feed is freely available. The term was coined following a study by Glen D. Jensen (1963), who gave rats the option of eating pellets from a dish or pressing a bar a pre-set number of times to obtain pellets. Out of 200 rats tested, 199 obtained at least some of their food following bar-presses, thus not 'freeloading' by simply eating the freely available food. Contrafreeloading has been demonstrated in a variety of species, including starlings (Bean *et al.*, 1999), pigs (de Jonge *et al.*, 2008), wolves (Vasconcellos *et al.*, 2012) and cattle (Van Os *et al.*, 2018). A wonderful test of this phenomenon comes from a study in grizzly bears (McGowan *et al.*, 2010). The bears were offered salmon and apples frozen into large blocks of ice, while at the same time having free access to salmon and apples placed on the floor of their home pen. They were also offered a block of plain ice to control for the potential attractiveness of manipulating an ice-block. Of the four bears taking part, only two managed to get food out of the ice-blocks, but all bears spent more time manipulating the ice-blocks containing food than they did the plain ice. The authors list various criteria used in the literature for defining the occurrence of contrafreeloading (from just spending time trying to gain access to the non-free food to working for more than 50% of their food intake). The ice-blocks turned out to make it very difficult for the bears to gain access to the food, and a subsequent experiment using apples in boxes, free apples and empty boxes proved more successful. What I like about the ice-block trial, however, is the attention to detail applied in the experimental design: the inclusion of the plain ice-block, the freezing of all ice-blocks in two halves to allow the (initially floating) food items to freeze in the middle of the ice-block, and the piercing of the apples to mimic the flow of juices from the salmon into the water as it froze.

Contrafreeloading is thought to provide an animal with information about its foraging environment to be used in the event that the free food is no longer available. It has been demonstrated that the degree of contrafreeloading is dependent on the success rate of obtaining the reward that needs to be worked for (Inglis *et al.*, 1997), as demonstrated in the grizzly bear example. Differences in feeding motivation also play a role, with fast-growing broilers showing a lesser degree of contrafreeloading than laying hens, which have lower nutritional requirements (Lindqvist *et al.*, 2006).

Tests of Motivation

When we stand in the queue at the bakers looking at the display of baked goods, we weigh up a number of factors before arriving at our choice. The choice may appear spontaneous to us if the queue is short or non-existent, as we quickly move a pointing finger in front of the glass panel

before saying ‘That one, please’. But if the queue is a little longer, we may take a few moments to consider the various options, mentally tasting different items, and even changing our minds from rye bread to a wheat loaf because it leaves room for a cake. The choice is also dependent on the money available. What it boils down to in the end is what are we motivated to eat within the limits of our available funds and the items on offer? This paradigm is very pertinent also when we want to study the choices of animals. How do they make their choices? What is their motivation and how do we measure it?

Operant conditioning

The strength of an animal’s motivation to obtain a resource can be estimated in a test where we ask the animal to pay a price to gain access to the resource. This was briefly mentioned earlier in the closed-economy example from mink, but here I would like to present in a little more detail, the concept of operant methods with which to obtain so-called demand curves. These stem from economic theory in which the demand for goods can be assessed by functions (demand functions) estimating how much of a given resource will be consumed dependent on the price. Again, this is dependent on the money available, and the question most often asked is ‘if you have a given sum, say €100, how much do you buy, respectively, of different items?’ Staple foods, such as potatoes and rice, may have gently sloping demand curves, as they form an essential part of human nutrition, and hence will – subject to certain caveats – continue to be bought even when their price increases. Life-dependent substances, such as insulin if you are diabetic, will have virtually inelastic, flat demand curves. Other items, such as sherry or cream, may be consumed less or even disappear from the shopping list if their prices go up. They are thus seen as luxury goods, because their demand curves are much more elastic. Demand functions estimate consumption relative to price, and the slope of the demand curve indicates the degree of elasticity. A recent study of university students showed that the amount of alcohol the student said they would buy declined if the price was high, and French students were more willing than American students to substitute an expensive alcoholic drink with a cheaper non-alcoholic alternative (Martinetti *et al.*, 2019).

How can we ask animals the same type of question? Operant conditioning techniques, where a test animal has to press a lever to obtain access to a resource, allow us to change the price of the resource by increasing and decreasing the number of lever presses necessary to obtain one unit of said resource (Dawkins, 1983; Matthews and Ladewig, 1994). This will tell us something about the value an animal puts on the resource, and how this changes as a function of price. It also makes it possible to compare the relative values of two different resources.

An example of the use of this concept was carried out to estimate the amount of time a cow needed to lie down every day. This question is relevant as a short lying time may indicate a sub-optimal environment, such as insufficient or unsuitable lying space, whereas long lying times may be caused by lameness or other health issues. Jensen *et al.* (2005) habituated young cows to wear a harness around their girth, and this could be attached to a wire in the ceiling above the cow, thus preventing it from lying down. The wire could be released by the cow pressing a panel (Fig. 5.3a) a set number of times, a so-called fixed ratio (FR) schedule, and this number could be increased to see how much the cow was willing to work for access to recumbent rest. In each 24-h period, the cows were prevented from lying down for 9 h, were free to lie down for 9 h and in the remaining 6 h had the possibility to work for access to lie down. The workload varied from 10 to 50 panel presses (FR10 to FR50) to obtain the reward of being able to lie down, and different reward lengths were investigated, from 20 up to 80 min of access to lie down for each completed FR. Figure 5.3b shows the duration of lying time obtained as a function of the FR value. When only 20 min lying was possible, the demand was quite elastic, as indicated by the slope of the demand curve. If instead 50- or 80-min periods became available, the demand was almost inelastic, with cows willing to work at the highest FR

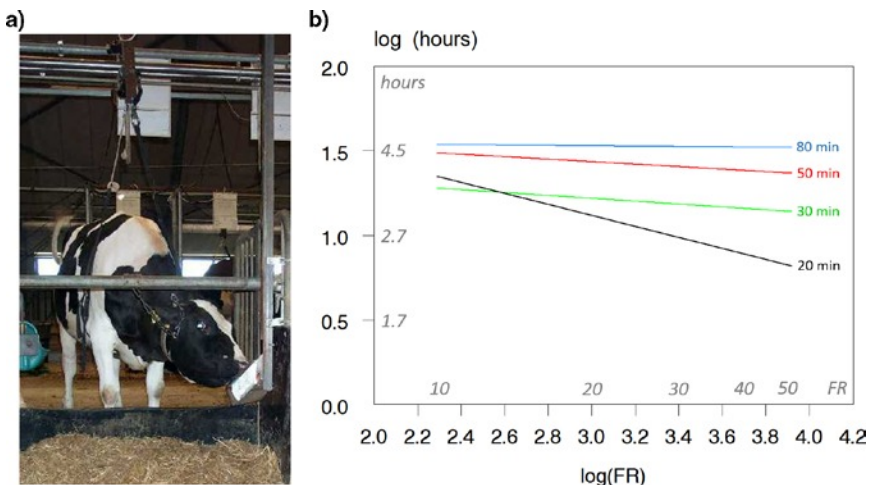


Fig. 5.3. (a) Cow pressing a panel to release a harness permitting the cow to lie down; (b) demand curves for different fixed ratio (FR) values, i.e. the number of panel presses needed to obtain a reward, and for different reward durations. If the reward time was only 20 min of lying down, the demand dropped with increasing FR. For reward durations of 50 and 80 min, the demand was inelastic at just over 4 h (data from Jensen *et al.*, 2005; photo: Margit Bak Jensen and Lene Munksgaard).

to be able to lie down. Using these reward lengths was therefore a better way to estimate the demand for lying, and cows obtained approximately 4 h of lying time, which – together with the 9 h they spent lying when it was freely available – indicates that cows have a need for recumbent rest for around 13 h per day. This finding has been confirmed by other experiments, where cows were found to defend their lying time, even to the extent that it affected their daily feed intake, and the cows lost weight (Munksgaard *et al.*, 2005).

Operant techniques can also be used to estimate the preference for one resource over another. This method was employed to ask pigs which rooting materials they preferred, and how motivated they were to get access to it; in other words, the strength of their preference (Pedersen *et al.*, 2005). In this set-up, the pigs were first trained to associate the pressing of a panel with 3 min of access to rooting material. The price of access to the rooting material could be raised by increasing the number of panel presses required to release the rooting material (Fig. 5.4). A second panel gave access to another type of rooting material, and by changing the relative number of panel presses between the two panels, the pigs were asked to vote with their snout. A companion animal in an



Fig. 5.4. Pigs are trained to press a panel (the terracotta-coloured square on each of the two vertical cylinders) to obtain a portion of rooting material in which the pig can root for up to 3 min. The test is performed adjacent to a companion pig that does not press a panel but simply gets the same type and amount of rooting material as the test pig. By giving access to two different panels, different types of rooting material can be tested and their relative attractiveness compared by altering the price (number of panel presses necessary) of one relative to the other (Pedersen *et al.*, 2005). In the photo, the test pig has chosen peat over long straw (photo: Margit Bak Jensen and Lene Juul Pedersen).

adjacent pen was there to prevent social isolation of the test animal, and the choice made by the test pig released the same rooting material into the pen of the companion pig.

Unchopped (long) straw was always used as one of the two rooting materials. This was first compared also with long straw to ensure that there was no side preference and the pigs would always choose the panel with the cheapest access, i.e. fewest presses. The price on the two panels was set according to different FR values, i.e. the number of panel presses necessary to obtain a reward, in this case access to a rooting material for 3 min. The different schedules meant that when the price on one panel was low the other was high. The five schedules used went from 8 presses necessary on one panel and 40 on the other (FR8/FR40) over FR16/FR32 and equal price at FR24/FR24 up to FR32/FR16 and FR40/FR8. Before each FR combination was tested over four test days, the pigs were left alternately with only one of the panels working so that they were acquainted with the workforce necessary to obtain the reward. Also, on each test day the five combinations of FR pairings were tested in random order.

The strength of preference for a rooting material (long straw, chopped straw, branches or peat) compared with long straw could then be assessed by seeing at which price the pigs obtained more of one than the other. For example, when long straw was available from both dispensers, we would expect the pigs to choose more from the side with the lowest price, and for the crossover between the two demand curves (consumption plotted against FR on a log scale) to be when the price was the same at FR24/FR24. Indeed, the crossover point was at FR29, which was not significantly different from FR24. For the three other rooting materials, chopped straw was not preferred over long straw (crossover at FR24), but this was significantly lower for branches (FR18) and even more so for peat (FR9), indicating that the pigs were willing to pay more for access to these rooting materials, and hence had a stronger preference for these compared with long straw. The authors concluded that pigs preferred peat over branches, which they in turn preferred over chopped and long straw.

I find the use of operant techniques to ask animals what they want fascinating. And as shown by the few examples summarized above, the method can be applied for a range of species and a wide range of resources or reward types. But of course we need to be careful that the answer we get is applicable to the question we were asking. Often the devil is in the detail, and what I haven't covered here are the protocols for training the animals prior to testing, both in the use of the operant apparatus and the associative learning involved. If you want to use this technique in your own studies, it is of paramount importance that you have enough time to habituate and train your test animals. Several issues need to be taken into account, but among the main ones are reward type and operant equipment.

As demonstrated in the example of lying time in dairy cows, the reward characteristics can affect the result. There was a difference in demand curve shape when the maximum duration of lying bouts possible was only 20 min, which clearly was not worth working for although the cows would have been rest-deprived to some degree. Rats pressing a bar to obtain heat are less responsive when the duration of the heat reward is short (Refinetti and Carlisle, 1987), and male rats only develop a specific odour preference for a female if they are allowed enough time to copulate and perform post-ejaculatory behaviour (Kippen and Pfaus, 2001). In contrast, mink are still willing to work for access to a swim bath, even when the access time is short (Hansen and Jensen, 2006). The time taken to complete the necessary number of lever presses may also reduce the number of rewards obtained (Jowaisas *et al.*, 1971). Another characteristic of reward is the quality vs quantity dimension. We may be interested in knowing what type of rooting material pigs prefer, but how can we present these in a manner that makes them comparable, so that the choice is about the type of rooting material and not the amount? In the example above, the authors did their best to estimate how many branches and of what length are comparable to a given amount of long straw, which in turn is much lighter than peat on a weight-to-volume basis.

Different types of operant equipment can be used. For example, the assessment of lying time in cows has also been estimated by placing a push door between the cow and the lying area (Tucker *et al.*, 2018). However, the cows found the use of the door difficult, and the method was found to underestimate the motivation of the cows to lie down. The same type of apparatus has been used successfully to assess cows' motivation to gain access to a rotating brush with which to be groomed (McConnachie *et al.*, 2018), which they were highly motivated to do. Laying hens and cichlid fish will also push through weighted swing doors to gain access, respectively, to perches at night (Olsson and Keeling, 2002) and social partners (Galhardo *et al.*, 2011). It has also been found that pulling and pushing is not the same to a mink: when required to pull a chain to obtain food the demand curve was more elastic (steeper) than if mink were asked to press a lever for the same reward (Hansen *et al.*, 2002). Indeed, in some animal species the design of the operant task needs to correspond to the behaviour usually required to obtain the reward type in question, such as pecking or pushing a lever to obtain a food reward as opposed to flapping a wing for pigeons (Smith and Keller, 1970; Domjan, 1983), and scratching themselves for rats (Pearce *et al.*, 1978). Also, the movement made by the animal the first time a reward appears will often be repeated (Skinner, 1948).

In conclusion, use the different methods mentioned in this chapter to gain insight into the preferences, aversions and motivations of your animal subjects. But use the techniques with care and consideration for the question you want answered.

References

- Adámková, J., Svoboda, J., Benediktová, K., Martini, S., Nováková, P. *et al.* (2017) Directional preference in dogs: laterality and 'pull of the north'. *PLoS ONE* 12(9), e0185243. DOI: 10.1371/journal.pone.0185243.
- Bean, D., Mason, G.J. and Bateson, M. (1999) Contrafreeloading in starlings: testing the information hypothesis. *Behaviour* 136(10–11), 1267–1282. DOI: 10.1163/156853999500712.
- Cooper, J.J. and Mason, G.J. (2000) Increasing costs of access to resources cause re-scheduling of behaviour in American mink (*Mustela vison*): implications for the assessment of behavioural priorities. *Applied Animal Behaviour Science* 66(1–2), 135–151. DOI: 10.1016/S0168-1591(99)00069-6.
- Correa, M., Pardo, M., Bayarri, P., López-Cruz, L., San Miguel, N. *et al.* (2016) Choosing voluntary exercise over sucrose consumption depends upon dopamine transmission: effects of haloperidol in wild type and adenosine A2AKO mice. *Psychopharmacology* 233(3), 393–404. DOI: 10.1007/s00213-015-4127-3.
- Dawkins, M. (1976) Towards an objective method of assessing welfare in domestic fowl. *Applied Animal Ethology* 2(3), 245–254. DOI: 10.1016/0304-3762(76)90056-0.
- Dawkins, M. (1977) Do hens suffer in battery cages? Environmental preferences and welfare. *Animal Behaviour* 25, 1034–1046. DOI: 10.1016/0003-3472(77)90054-9.
- Dawkins, M.S. (1983) Battery hens name their price: consumer demand theory and the measurement of ethological 'needs'. *Animal Behaviour* 31(4), 1195–1205. DOI: 10.1016/S0003-3472(83)80026-8.
- de Jonge, F.H., Tilly, S.-L., Baars, A.M. and Spruijt, B.M. (2008) On the rewarding nature of appetitive feeding behaviour in pigs (*Sus scrofa*): do domesticated pigs contrafreeload? *Applied Animal Behaviour Science* 114(3–4), 359–372. DOI: 10.1016/j.applanim.2008.03.006.
- Dixon, L.M., Sandilands, V., Bateson, M., Brocklehurst, S., Tolkamp, B.J. *et al.* (2013) Conditioned place preference or aversion as animal welfare assessment tools: limitations in their application. *Applied Animal Behaviour Science* 148(1–2), 164–176. DOI: 10.1016/j.applanim.2013.07.012.
- Domjan, M. (1983) Biological constraints on instrumental and classical-conditioning – implications for general process theory. *Psychology of Learning and Motivation – Advances in Research and Theory* 17, 215–277.
- Galhardo, L., Almeida, O. and Oliveira, R.F. (2011) Measuring motivation in a cichlid fish: an adaptation of the push-door paradigm. *Applied Animal Behaviour Science* 130(1–2), 60–70. DOI: 10.1016/j.applanim.2010.12.008.
- Hansen, S.W. and Jensen, M.B. (2006) Demand for swimming water and running wheel with 1 min of access per reward. *Applied Animal Behaviour Science* 98(1–2), 145–154. DOI: 10.1016/j.applanim.2005.08.016.
- Hansen, S.W., Jensen, M.B., Pedersen, L.J., Munksgaard, L., Ladewig, J. *et al.* (2002) The type of operant response affects the slope of the demand curve for food in mink. *Applied Animal Behaviour Science* 76(4), 327–338. DOI: 10.1016/S0168-1591(02)00008-4.
- Hughes, B.O. (1976) Preference decisions of domestic hens for wire or litter floors. *Applied Animal Ethology* 2(2), 155–165. DOI: 10.1016/0304-3762(76)90043-2.
- Inglis, I.R., Forkman, B. and Lazarus, J. (1997) Free food or earned food? A review and fuzzy model of contrafreeloading. *Animal Behaviour* 53(6), 1171–1191. DOI: 10.1006/anbe.1996.0320.

- Jensen, G.D. (1963) Preference for bar pressing over 'freeloading' as a function of number of rewarded presses. *Journal of Experimental Psychology* 65(5), 451–454. DOI: 10.1037/h0049174.
- Jensen, M.B., Pedersen, L.J. and Munksgaard, L. (2005) The effect of reward duration on demand functions for rest in dairy heifers and lying requirements as measured by demand functions. *Applied Animal Behaviour Science* 90(3–4), 207–217. DOI: 10.1016/j.applanim.2004.08.006.
- Jowaisas, D., Taylor, J., Dewsbury, D.A. and Malagodi, E.F. (1971) Copulatory behavior of male rats under an imposed operant requirement. *Psychonomic Science* 25(5), 287–290. DOI: 10.3758/BF03335879.
- Katebi, N., Farahimanesh, S., Fatahi, Z., Zarrabian, S. and Haghparast, A. (2018) Involvement of D1- and D2-like dopamine receptors in the dentate gyrus in the acquisition, expression, and extinction of the morphine-induced conditioned place preference in rats. *Behavioural Brain Research* 353, 185–193. DOI: 10.1016/j.bbr.2018.07.018.
- Kippin, T.E. and Pfaus, J.G. (2001) The development of olfactory conditioned ejaculatory preferences in the male rat I. Nature of the unconditioned stimulus. *Physiology & Behavior* 73, 457–469.
- Kristensen, H.H., Prescott, N.B., Perry, G.C., Ladewig, J., Ersbøll, A.K. *et al.* (2007) The behaviour of broiler chickens in different light sources and illuminances. *Applied Animal Behaviour Science* 103(1–2), 75–89. DOI: 10.1016/j.applanim.2006.04.017.
- Lindqvist, C., Zimmerman, P. and Jensen, P. (2006) A note on contrafreeloading in broilers compared to layer chicks. *Applied Animal Behaviour Science* 101(1–2), 161–166. DOI: 10.1016/j.applanim.2006.01.006.
- Martinetti, M.P., Caughron, R.L., Berman, H.L., André, J., Sokolowski, M.B.C. *et al.* (2019) The behavioral economics of alcohol demand in French and American university students. *Alcoholism: Clinical and Experimental Research* 43(3), 531–544. DOI: 10.1111/acer.13954.
- Matthews, L.R. and Ladewig, J. (1994) Environmental requirements of pigs measured by behavioural demand functions. *Animal Behaviour* 47(3), 713–719. DOI: 10.1006/anbe.1994.1096.
- McConnachie, E., Smid, A.M.C., Thompson, A.J., Weary, D.M., Gaworski, M.A. *et al.* (2018) Cows are highly motivated to access a grooming substrate. *Biology Letters* 14(8), 20180303. DOI: 10.1098/rsbl.2018.0303.
- McDowell, L.J., Wells, D.L. and Hepper, P.G. (2018) Lateralization of spontaneous behaviours in the domestic cat, *Felis silvestris*. *Animal Behaviour* 135, 37–43. DOI: 10.1016/j.anbehav.2017.11.002.
- McGowan, R.T.S., Robbins, C.T., Alldredge, J.R. and Newberry, R.C. (2010) Contrafreeloading in grizzly bears: implications for captive foraging enrichment. *Zoo Biology* 29, 484–502.
- Munksgaard, L., Jensen, M.B., Pedersen, L.J., Hansen, S.W. and Matthews, L. (2005) Quantifying behavioural priorities—effects of time constraints on behaviour of dairy cows, *Bos taurus*. *Applied Animal Behaviour Science* 92(1–2), 3–14. DOI: 10.1016/j.applanim.2004.11.005.
- Olsson, I.A.S. and Keeling, L.J. (2002) The push-door for measuring motivation in hens: laying hens are motivated to perch at night. *Animal Welfare* 11, 11–19.
- Paul, E.S., Edgar, J.L., Caplen, G. and Nicol, C.J. (2018) Examining affective structure in chickens: valence, intensity, persistence and generalization measured

- using a conditioned place preference test. *Applied Animal Behaviour Science* 207, 39–48. DOI: 10.1016/j.applanim.2018.07.007.
- Pearce, J.M., Colwill, R.M. and Hall, G. (1978) Instrumental conditioning of scratching in the laboratory rat. *Learning and Motivation* 9(3), 255–271. DOI: 10.1016/0023-9690(78)90009-7.
- Pedersen, L.J., Holm, L., Jensen, M.B. and Jørgensen, E. (2005) The strength of pigs' preferences for different rooting materials measured using concurrent schedules of reinforcement. *Applied Animal Behaviour Science* 94(1–2), 31–48. DOI: 10.1016/j.applanim.2004.11.023.
- Refinetti, R. and Carlisle, H.J. (1987) Behavioral heat intake as a function of reward duration in rats. *Animal Learning & Behavior* 15(2), 228–237. DOI: 10.3758/BF03204966.
- Schroeder, P., Jones, S., Young, I.S. and Sneddon, L.U. (2014) What do zebrafish want? Impact of social grouping, dominance and gender on preference for enrichment. *Laboratory Animals* 48(4), 328–337. DOI: 10.1177/0023677214538239.
- Skinner, B.F. (1948) 'Superstition' in the pigeon. *Journal of Experimental Psychology* 38(2), 168–172. DOI: 10.1037/h0055873.
- Smith, R.F. and Keller, F.R. (1970) Free-operant avoidance in pigeon using a treadle response. *Journal of the Experimental Analysis of Behavior* 13(2), 211–214. DOI: 10.1901/jeab.1970.13.211.
- Tucker, C.B., Munksgaard, L., Mintline, E.M. and Jensen, M.B. (2018) Use of a pneumatic push gate to measure dairy cattle motivation to lie down in a deep-bedded area. *Applied Animal Behaviour Science* 201, 15–24. DOI: 10.1016/j.applanim.2017.12.018.
- Tzschentke, T.M. (2007) Measuring reward with the conditioned place preference (CPP) paradigm: update of the last decade. *Addiction Biology* 12(3–4), 227–462. DOI: 10.1111/j.1369-1600.2007.00070.x.
- Van Os, J.M.C., Mintline, E.M., DeVries, T.J. and Tucker, C.B. (2018) Domestic cattle (*Bos taurus taurus*) are motivated to obtain forage and demonstrate contrafreeloading. *PLoS ONE* 13(3), e0193109. DOI: 10.1371/journal.pone.0193109.
- Vasconcellos, A.S., Harumi Adania, C. and Ades, C. (2012) Contrafreeloading in maned wolves: implications for their management and welfare. *Applied Animal Behaviour Science* 140(1–2), 85–91. DOI: 10.1016/j.applanim.2012.04.012.
- Wadhera, D., Wilkie, L.M. and Capaldi-Phillips, E.D. (2018) The rewarding effects of number and surface area of food in rats. *Learning & Behavior* 46(3), 242–255. DOI: 10.3758/s13420-017-0305-y.

Ability to Detect and Distinguish

In many of the behavioural tests used to ask animals questions, we present the animal with different stimuli. These can take many forms and require different sensory modalities to be engaged. We may use pictures of different geometrical shapes, three-dimensional objects in various colours, audible tones at different frequencies or different odorous compounds. Sometimes the animal is asked to choose between two stimuli that differ in some aspect, and the animal has to estimate if one choice is better than another. But how do we know that the animal is able to distinguish between two options if these differ only in, say, colour such as a red and a blue triangle? In other words, how do we know which colours an animal of a given species is capable of seeing? Or which odours it is able to detect?

Categorization of Stimuli

In Chapter 4 we heard about tests of cognitive bias, where animals were trained to associated two stimuli differing in one dimension (e.g. a high and low frequency tone) with a positive and negative experience, respectively, and their responses to intermediate tones were then observed. If instead an animal is exposed to the same positive experience with both stimuli, how does it react to intermediate versions of these stimuli? Jones *et al.* (2001) used this type of test to investigate how young chicks categorized colours. Male poultry chicks of 1 week of age were trained in pairs to obtain food rewards from paper cones printed with squares of one of two colours on a grey background. In the example chosen here, the two colours were blue and green. One-third of the chick pairs was rewarded with the colour blue, i.e. only cones with blue squares contained a food reward, the other cones (either grey or printed with green squares) were empty. Similarly, a third of the chick pairs was rewarded in green cones only, while the remainder of chicks were rewarded in both blue and green cones. All the animals were then tested once with empty cones printed with green, blue or an intermediate colour (turquoise) and the attractiveness of each of these was scored based on the number of

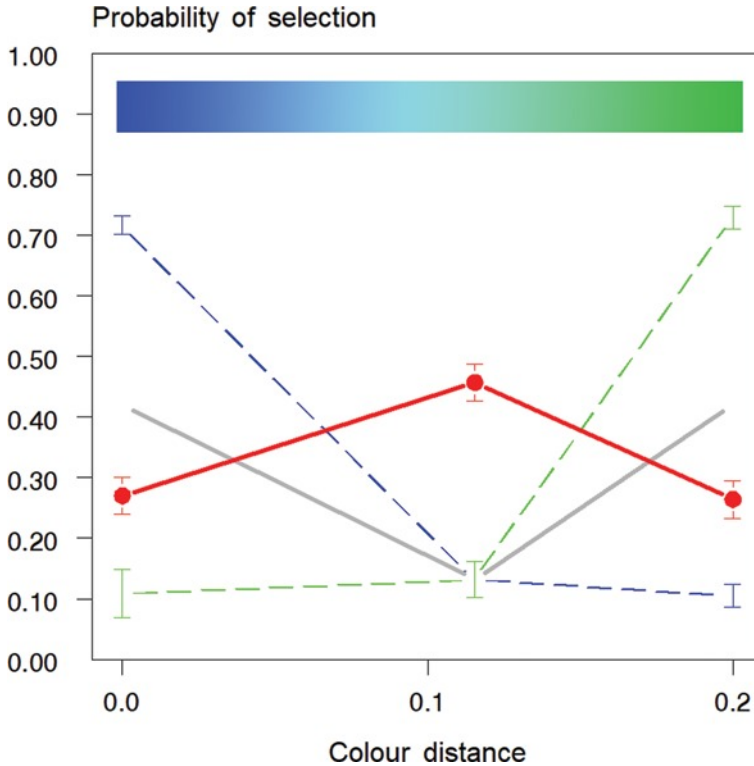


Fig. 6.1. Likelihood of chicks pecking a blue, green or turquoise cone after having previously found rewards only in blue (blue dashed line), or green (green dashed line), or both blue and green cones (red line with circular markers). The solid grey line indicates the response predicted for the chicks trained with both colours, if their preference for the intermediate colour (turquoise) had been the mean of the two other treatments (blue and green) (data from Jones *et al.*, 2001).

pecks they received from the pair of chicks. The results are shown in Fig. 6.1. Chicks that had previously been rewarded with both blue and green cones were significantly more attracted to the turquoise cones than the blue and green cones. It thus appears that the chicks interpolate the two colours, and expect the colour mixture *more* likely to be associated with a reward. This is not just due to the colour being novel, as the researchers also tested colours outside this spectrum and the chicks did not extrapolate to colours outside the blue–green continuum.

Estimation of Quantity

As humans, we usually think of ourselves as pretty good at judging quantities, especially when it comes to food. Cut a pizza into triangular segments,

and we can quickly point to the largest piece, probably based on the length of the crust. If two children are asked to share a cake evenly, the best solution will always be that one cuts, the other chooses (this method even has a dedicated Wikipedia page with the heading 'Divide and choose'). When it comes to liquids we may be less adept, especially if the containers are of different shapes. When asked to pour a shot (44.3 ml) of alcohol by eye, even experienced bartenders have been found to pour 20% more into short, wide glasses than tall, slender ones (Wansink and van Ittersum, 2005).

Chimpanzees, on the other hand, are very good at judging liquid quantities. To test whether this also holds true when only limited visual information is available, Beran (2010) presented chimps with different quantities of fruit juice with various amounts of visible evidence. First, they asked the chimps to choose between two different quantities (from 1 to 6 units) of juice presented in identical transparent glasses. This was to ensure that the chimps would choose the largest amount when given very clear information about the relative quantity. All 30 pair-wise combinations were presented in random order to each of the three chimpanzees taking part, and on only one occasion was the smallest amount chosen (5 vs 6 units). Next, the researcher tested 1 vs 4 units of juice presented in opaque cups. This was to make sure that the surface of the liquid was not visible to the animals, who, indeed, chose between these opaque cups at random, indicating that they were unable to ascertain the quantities before choosing. It should be added that the choice was indicated by the chimp reaching out and touching one of the cups, after which the experimenter emptied the chosen cup into the mouth of the animal. Also, the researcher placed the cups at each end of a box, pushed the box within reach of the chimp and immediately looked down to prevent themselves from giving any unintentional cues as to which cup contained the most juice.

Having thus established the baseline for the experiment, the chimps were now given a choice between a transparent cup containing a quantity of juice that was visible to the animal while the researcher filled an opaque cup from an opaque syringe, held above the cup and emptied at a constant rate (Fig. 6.2). The chimpanzees thus had to estimate if the amount dispensed from the syringe into the cup was more or less than the amount visible in the adjacent glass. The chimpanzees were still able to identify the cup containing the largest amount on average 83% of the time, with their first choice being correct for 12–13 of the 15 combinations presented. Now, this could be because the chimps were able to gauge the time it took to empty the syringe, so in the next trial, two identical opaque cups were filled, one after the other, from opaque syringes held above the cup and emptied at the same rate, but the second cup was always filled from twice the height. This gave both a greater auditory feedback when the liquid hit the bottom of the cup, and the liquid was visible for longer during the filling. However, this did not lead to a different result from before, with the animals identifying the largest amount of liquid on average 81% of

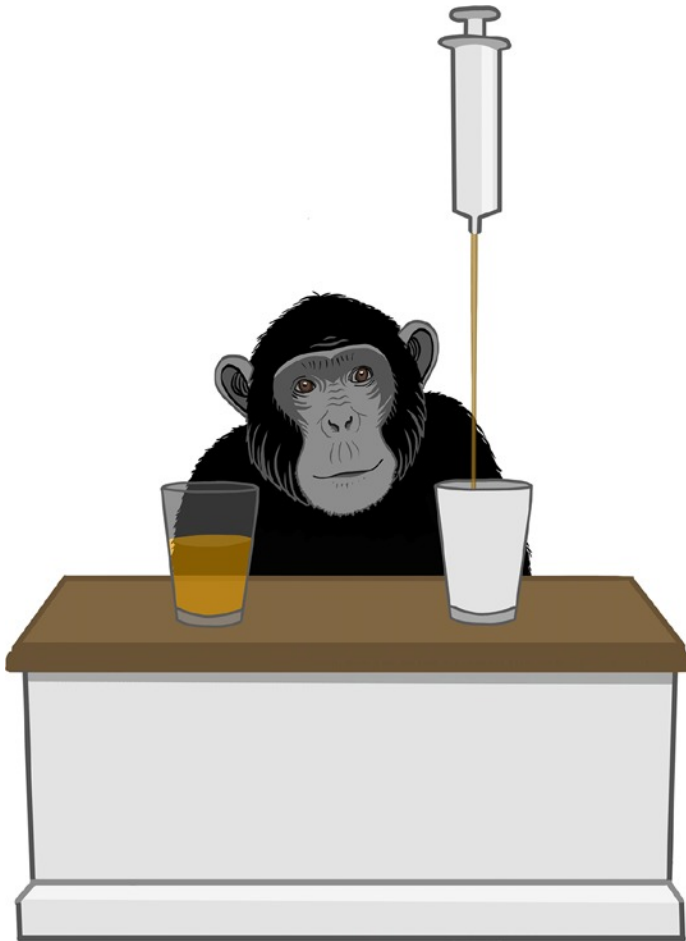


Fig. 6.2. Chimpanzee given a choice of two cups: one transparent cup containing a quantity of fruit juice, and one opaque cup being filled with a different quantity of juice from an opaque syringe. The animals tested were able to choose the cup with the greatest quantity of juice the vast majority of times (from Beran, 2010).

the time, and with 12–13 first correct choices out of 15 possible. This is an example of using simple means with which to present an animal with different options while still changing only a few parameters each time. This enables the experimenter to tease out which aspects are used by the animal in making the decision. In effect, the chimpanzees were able to continue to estimate the quantities, even when exposed to various manipulations and forms of presentation, involving greater auditory and visual feedback. The author suggests one way to continue this investigation could be to use syringes with different gauges to dispense different amounts of liquid in the same amount of time. Other species, such as grey parrots are also able to

grasp the concept of liquid conservation – that the amount doesn't change when the liquid is transferred to a container of a different size (Pepperberg *et al.*, 2017).

Being able to estimate the size of a food source relative to another is a valuable skill to have. It is one of the cornerstones of optimal foraging, which is when animals weigh up the pros and cons of food items in terms of availability and effort needed to procure them. In a classic experiment by Krebs *et al.* (1977), four wild-caught great tits (*Parus major*) were presented with large and small meal worms passing on a conveyor belt, visible and accessible through a small opening for 0.5s only. The size of the meal worm passing was alternately large and small, and when only one or two meal worms passed per 40s, the birds were able to pick and eat them all. However, when the frequency of passing prey was increased to six meal worms per 40s, the birds started to choose the larger ones (86% of worms) in preference to the smaller. The great tits were thus able to tell the difference and deliberately await the next large prey passing. The attraction of large over small when it comes to food items is very difficult for animals to ignore. Dogs can learn to touch a plate representing a large amount of food if the dog afterwards receives the large amount of food. However, in a so-called reverse-reward contingency task, where the dogs are required to select the smaller option in order to receive the large portion, they are unable to do so (Fernand *et al.*, 2018).

Discrimination Between Stimuli

Habituation–dishabituation test

A commonly used paradigm to test if animals are able to distinguish between two similar stimuli is the habituation–dishabituation test. In its simplest form, an animal is repeatedly presented with a stimulus, say, an odour, and the animal will usually show less and less interest in the odour in the form of reduced sniffing and investigation of the odour source. After three to four such presentations, the animal has habituated to the odour. The odour is then replaced by another compound with a different smell, and the behaviour of the animal is again observed. If the animal continues to show a lack of interest in the novel odour, it is concluded that the animal cannot distinguish between the two odours. If, however, the animal increases its investigation of the odour source upon being presented with the second, novel odour, the animal is demonstrating that it has detected the change, and hence is able to distinguish between the two odours. This change in behaviour following the odour change is what has been termed 'dishabituation'. However, the term is used incorrectly, as dishabituation involves a behavioural change to the same, previously habituated stimulus (see Rankin *et al.*, 2009), but for ease of reference, I will continue to refer to this test as habituation–dishabituation.

Habituation–dishabituation tests have been used to determine if human infants are able to distinguish between two visual stimuli (Kavšek, 2004), but in animals they are often applied in an olfactory context using rodents as subjects. Sometimes the ability to distinguish between two odours that are very similar is dependent on prior experience with both odours; in other words, the animal appears to learn to distinguish through exposure to the odours. An example of this comes from Moreno *et al.* (2009), using the compound limonene in its two isomer forms, +limonene and –limonene. In this trial, mice with no prior experience of limonene were exposed to one of the two types of limonene four times in a row, resulting in a decrease in interest (habituation; red bars in Fig. 6.3a). When the other limonene type was presented on the fifth occasion, no increase in investigation occurred, indicating that the mice could not distinguish between the two odours. Subsequently, the mice were exposed to both +limonene and –limonene in their home cage for 1 h daily for 10 days. The odours were presented inside two metal tea balls hung from the cage lid, each containing one of the two types of limonene. Having thus gained experience with the two odours, the habituation–dishabituation test was repeated, and the mice were now able to tell the two odours apart, as indicated by the rekindling of interest when the odour changed on the fifth presentation (dishabituation; blue bar in Fig. 6.3b). Incidentally, the test can be adapted to other species than rodents, as shown by Rørvang *et al.* (2017), who investigated if the method could be used to test dairy cows' ability to distinguish

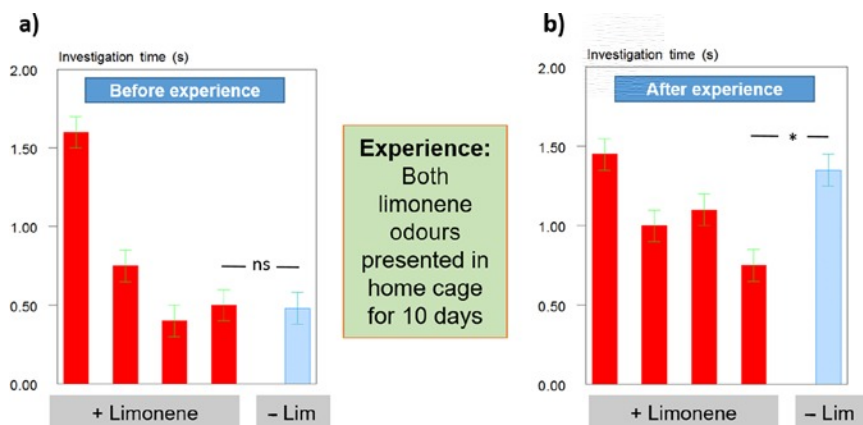


Fig. 6.3. Experience with odours improves the ability to distinguish them. (a) Mice with no prior experience of two types of limonene (+ and –) cannot initially distinguish between them, but after exposure to both odours for 1 h per day in their home cage for 10 days, (b) the mice are able to tell the two odours apart, as indicated by the increased investigation time when the odour is changed after repeated exposure to the other odour during the habituation–dishabituation test (data from Moreno *et al.*, 2009).



Fig. 6.4. Dairy cow in a habituation–dishabituation test with different odours (photo: Maria V. Rørvang).

between coffee and orange juice (and yes, they could, and they preferred the smell of coffee; [Fig. 6.4](#)).

Just to include an example of the habituation–dishabituation test being used for something other than odours, Saito *et al.* (2019) tested if cats could recognize their own names using a version of this paradigm. Briefly, the response of the cats was observed while recordings were played of different words spoken by the same person, either a stranger or the owner. The words were either nouns that were similar in duration, pitch and intensity to the cat’s name, or names of other cats living in the same household. The habituation in this study was therefore not the same stimulus repeated, but different sounds. After being exposed to four different nouns or names, the cat’s own name was played. The responses registered were mainly ear and head movement, but also tail movement and vocalizations. Overall, 60% of the cats (67 out of 112) responded with less and less magnitude as the four sounds were played, in a similar manner to that shown in [Fig. 6.3](#), independent of sound type or person speaking. When the name of the test cat was played as the fifth sound, 44 of the 67 habituated cats (66%) showed a marked increase in response magnitude to the sound of their own name, indicating that they were able to distinguish it from the previous words. However, there is still a long way from your cat knowing its own name to your being able to make it come to you when called.

Training to Discriminate

Another form of test used to evaluate discrimination is the ability to distinguish between handlers wearing different coloured clothing. In an experiment with cattle (Rushen *et al.*, 1999), dairy cows were treated three to six times per day by two handlers, with one handler treating the cow gently and the other handler treating the cow aversively. The gentle treatment consisted of talking in a gentle voice while brushing the cow and offering tasty food. The aversive treatment included slapping four times with a flat hand, shouting at the cow, using a cattle prod once on the flank of the cow and hitting it with a plastic paddle. The treatments varied in length from 1 to 6 min to prevent the cows from predicting when the treatment stopped. The handlers wore either red or yellow overalls, and the treatments were balanced among cows for colour of overall as well as sex of the handler, and all cows were treated by both handlers each day, but in an unpredictable order. After 5 days of treatment, the handlers (wearing their red or yellow overalls) took turns standing for 60s with their hands in their pockets in front of the cows, and the behaviour of the cow was scored every 5s. This was converted into a distance score, and no differences were found between scores of response to the handlers before treatment, but with a significant increase in distance score after treatment when the aversive handler was present. No differences in reaction to handlers were found if the handlers swapped their overalls, or if they stood in front of the cows wearing green overalls. Also, after the 5 days of treatment, if the handlers stood close to the cows during milking, the heart rate of the cows increased significantly more with the aversive compared with the gentle handler present, whereas before treatment the heart rate was slightly reduced during milking. This was even more pronounced among the cows that were best at discriminating between the handlers. Cows are thus able to use the colour of clothing to discriminate between people, but as the cows did not reverse their response to the handlers when they swapped clothes, the cows also use other cues. Similar results have been found with dairy calves, and there are indications that cattle are also able to distinguish people by their faces or by body height (Rybarczyk *et al.*, 2001, 2003).

The importance of knowing whether an animal is able to detect a stimulus and distinguish it from another is clear from the examples given in this chapter. What also becomes apparent is the necessity to obtain this knowledge in order to ensure the validity of other behavioural tests we may want to carry out. If an animal is unable to sense a stimuli used in a test, the results have little value. Luckily, the sensory capacities of most of the common test species are already well known, and we do not need to start from scratch every time. It is worth noting, however, that some sensory modalities, not least olfaction, are susceptible to external disturbances and the stress level of the animal may affect its responsiveness as well as its sensitivity (Raynaud *et al.*, 2015; Bombail, 2019).

An experimental set-up frequently used in olfactory research with rats is a nose-poke device. Various types exist, but in general the rat is presented with a wall containing circular holes, into which the rat can insert its nose. The central hole is fitted with a constant flow of clean air, and by poking its nose into the central hole, the rat triggers the release of an odour into the airstream. A reward can be obtained from the adjacent holes, often one on each side of the odourized hole. Training the rat on two different odours, a reward is given in the hole on the right when odour A is released and on the left with odour B. The rewards can be small food rewards or drops of water, dependent on the motivation of the rat, which is manipulated by restricting access to feed or water before the test session. Imminent release of an odour is indicated by a light or a sound to alert the rat to the possibility of obtaining a reward. The nose-poke apparatus may have more holes, with five and nine being common, but with some of these blocked to provide the desired number of choices (e.g. three out of nine holes available; Dowd and Dunnett, 2005). The sensitivity to odours can be tested by lowering the concentration of the odours released, and relative discrimination between two odours can also be assessed by gradually diluting each of the odours with the other one. Changes in the response time, i.e. duration from odour release until a choice is made, as well as changes in the number of errors made compared to trials with strong and pure dilutions, are used to assess how different treatments affect the odour processing ability of the test animals.

In recent years, a number of scientific papers have emerged on the ability of animals to detect diseases in humans based on odours. This has included a range of species used to detect a variety of diseases, most commonly cancer but also conditions such as epilepsy, where pet dogs have been found to display seizure-specific behaviour, with some of them indicating an epileptic seizure in their owner before it happens (Kirton *et al.*, 2004). African pouched rats have been trained to detect tuberculosis in saliva samples (Weetjens *et al.*, 2009) and even humans are in on the act, as in the case of Joy Milne, who in 2015 demonstrated that based on smell alone she could detect Parkinson's disease in patients before other clinical symptoms became apparent. The ability of dogs to detect different types of cancer via smell is now well established (e.g. McCulloch *et al.*, 2012; Taverna *et al.*, 2015), although, unlike Joy Milne, they cannot reliably identify the disease prior to other clinical symptoms (Jezierski, 2017). And the training of dogs to distinguish between biological samples from patients and non-patients can turn up some remarkable results, but not always for the right reason.

Elliker *et al.* (2014) did a trial with 10 dogs in which they trained these animals to distinguish between urine samples from people with or without prostate cancer. In a large arena, the dogs were presented with a 3-m-long plastic rack with four holes (2 cm in diameter) spaced 75 cm apart. Underneath these scent-holes, odour samples could be placed without the

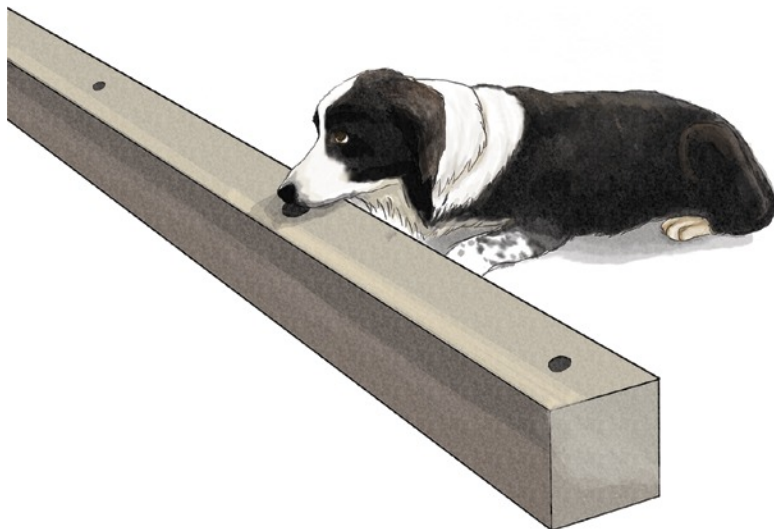


Fig. 6.5. Dog indicating which scent-hole out of four contains a urine sample from a cancer patient (based on Elliker *et al.*, 2014).

dogs being able to see or access the open-top containers holding the samples. Initially, a random scent-hole was baited with food, and when the dog placed its nose on the baited hole (Fig. 6.5), the dogs were praised and rewarded with food as well as clicker training (see Chapter 8 for more details on positive reinforcement schedules). This initial training ensured that the dogs were interested in the scent-holes, and the testing procedure now progressed to consist of one hole baited with a urine sample from a prostate cancer patient (CaP), while the other three holes were left empty. Again, the dogs were positively reinforced to indicate the baited hole. Once the dogs mastered this, the four holes were baited with one CaP and three non-CaP urine samples, with the placement of samples in the four scent-holes being random. A handler was in the arena with the dog, but was unaware of the placement of the CaP sample. The handler rewarded the dogs for a correct odour indication only when told to do so by a person following the training via a video feed from outside the test room. This was to ensure that the handler could not unintentionally reveal the correct scent-hole to the dog through micro-movements or facial expressions. Unfortunately, the supply of CaP samples was limited, and samples from the same patients were presented more than once. The experimenters also used pooled samples from up to three prostate cancer patients, to vary the odour profiles in order to train the dogs on the general smell of prostate cancer. After 5 months of training, two of the dogs had displayed a high level of correct responses in training. These two dogs were now tested on two novel sets of urine samples, one with 15 CaP and 45 control samples (making 15 arrays

of 4 samples), and another with 16 CaP and 48 control samples. All of the samples were from different donors and unknown to both the dogs. The test was similar to that used during training, so that only 1 of the 4 urine samples was a CaP sample, and the position in the rack was randomized, with the handler being blind to the position of the CaP sample. Only one dog was tested with the first 15 arrays of samples, and it identified only 2 out of the 15 CaP samples. Using the second set of samples, both dogs were tested and each identified 2 and 4 of the 16 samples, respectively. The performance of the dogs in finding the right sample was no different from a random choice. Why?

The experimenters suggest that during the many months of training, the dogs must have learnt to distinguish among the individual urine samples from different donors and memorized which specific samples were rewarded. This in itself is no small feat as many samples were used during training, but the re-use of samples during training would have played a major role in this. Even pooling of samples was not able to overcome this. The experimenters did not foresee this problem as they thought the sample size was larger than the capacity of the dogs' olfactory memory. It should be mentioned that during the tests with the novel sets of samples, none of the experimenters knew which sample was the CaP sample. The dogs were rewarded when they indicated a sample, and the performance was only analysed after the tests when the sample identity was revealed. This meant that the dogs were rewarded for indicating the wrong samples during these tests of their capacity, a factor that may have contributed to the disappointing result. The authors list various recommendations for future studies to improve the success rate of cancer detection by dogs through odours, including using urine samples from several forms of cancer; training the dogs with both disease and control samples from the beginning; avoiding rewarding wrong choice in the final tests; and exploring the possibility of using other test paradigms, such as the habituation–dishabituation test mentioned earlier.

From the examples included in this chapter, it is clear that animals from a range of different species are able to assess their surroundings with a high degree of accuracy. A number of tests developed to investigate these animal characteristics have been described above. Some types of tests have been left out, as they also cover additional behavioural and cognitive abilities of animals, such as tests of object permanence, which will be covered in Chapter 9.

References

- Beran, M.J. (2010) Chimpanzees (*Pan troglodytes*) accurately compare poured liquid quantities. *Animal Cognition* 13(4), 641–649. DOI: 10.1007/s10071-010-0314-0.

- Bombail, V. (2019) Perception and emotions: on the relationships between stress and olfaction. *Applied Animal Behaviour Science* 212, 98–108. DOI: 10.1016/j.applanim.2018.12.013.
- Dowd, E. and Dunnett, S.B. (2005) Comparison of 6-hydroxydopamine-induced medial forebrain bundle and nigrostriatal terminal lesions in a lateralised nose-poking task in rats. *Behavioural Brain Research* 159(1), 153–161. DOI: 10.1016/j.bbr.2004.10.010.
- Elliker, K.R., Sommerville, B.A., Broom, D.M., Neal, D.E., Armstrong, S. *et al.* (2014) Key considerations for the experimental training and evaluation of cancer odour detection dogs: lessons learnt from a double-blind, controlled trial of prostate cancer detection. *BMC Urology* 14(1), 22–30. DOI: 10.1186/1471-2490-14-22.
- Fernand, J.K., Amanieh, H., Cox, D.J. and Dorey, N.R. (2018) Any reward will do: effects of a reverse-reward contingency on size preference with pet dogs (*Canis lupus familiaris*). *Learning & Behavior* 46(4), 472–478. DOI: 10.3758/s13420-018-0343-0.
- Jezierski, T. (2017) The role of olfaction in disease detection and prevention. In: Nielsen, B.L. (ed.) *Olfaction in Animal Behaviour and Welfare*. CAB International, Wallingford, UK, pp. 123–135.
- Jones, C.D., Osorio, D. and Baddeley, R.J. (2001) Colour categorization by domestic chicks. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268(1481), 2077–2084. DOI: 10.1098/rspb.2001.1734.
- Kavšek, M. (2004) Predicting later IQ from infant visual habituation and dishabituation: a meta-analysis. *Journal of Applied Developmental Psychology* 25(3), 369–393. DOI: 10.1016/j.appdev.2004.04.006.
- Kirton, A., Wirrell, E., Zhang, J. and Hamiwka, L. (2004) Seizure-alerting and -response behaviors in dogs living with epileptic children. *Neurology* 62(12), 2303–2305. DOI: 10.1212/WNL.62.12.2303.
- Krebs, J.R., Erichsen, J.T., Webber, M.I. and Charnov, E.L. (1977) Optimal prey selection in the great tit (*Parus major*). *Animal Behaviour* 25, 30–38. DOI: 10.1016/0003-3472(77)90064-1.
- McCulloch, M., Turner, K. and Broffman, M. (2012) Lung cancer detection by canine scent: will there be a lab in the lab? *European Respiratory Journal* 39(3), 511–512. DOI: 10.1183/09031936.00215511.
- Moreno, M.M., Linster, C., Escanilla, O., Sacquet, J., Didier, A. *et al.* (2009) Olfactory perceptual learning requires adult neurogenesis. *Proceedings of the National Academy of Sciences* 106(42), 17980–17985. DOI: 10.1073/pnas.0907063106.
- Pepperberg, I.M., Gray, S.L., Lesser, J.S. and Hartsfield, L.A. (2017) Piagetian liquid conservation in grey parrots (*Psittacus erithacus*). *Journal of Comparative Psychology* 131(4), 370–383. DOI: 10.1037/com0000085.
- Rankin, C.H., Abrams, T., Barry, R.J., Bhatnagar, S., Clayton, D.F. *et al.* (2009) Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory* 92(2), 135–138. DOI: 10.1016/j.nlm.2008.09.012.
- Raynaud, A., Meunier, N., Acquistapace, A. and Bombail, V. (2015) Chronic variable stress exposure in male Wistar rats affects the first step of olfactory detection. *Behavioural Brain Research* 291, 36–45. DOI: 10.1016/j.bbr.2015.05.013.

- Rørvang, M.V., Jensen, M.B. and Nielsen, B.L. (2017) Development of test for determining olfactory investigation of complex odours in cattle. *Applied Animal Behaviour Science* 196, 84–90. DOI: 10.1016/j.applanim.2017.07.008.
- Rushen, J., de Passillé, A.M.B. and Munksgaard, L. (1999) Fear of people by cows and effects on milk yield, behavior, and heart rate at milking. *Journal of Dairy Science* 82(4), 720–727. DOI: 10.3168/jds.S0022-0302(99)75289-6.
- Rybarczyk, P., Koba, Y., Rushen, J., Tanida, H. and de Passillé, A.M. (2001) Can cows discriminate people by their faces? *Applied Animal Behaviour Science* 74(3), 175–189. DOI: 10.1016/S0168-1591(01)00162-9.
- Rybarczyk, P., Rushen, J. and de Passillé, A.M. (2003) Recognition of people by dairy calves using colour of clothing. *Applied Animal Behaviour Science* 81(4), 307–319. DOI: 10.1016/S0168-1591(02)00273-3.
- Saito, A., Shinozuka, K., Ito, Y. and Hasegawa, T. (2019) Domestic cats (*Felis catus*) discriminate their names from other words. *Scientific Reports* 9(1), 5394. DOI: 10.1038/s41598-019-40616-4.
- Taverna, G., Tidu, L., Grizzi, F., Torri, V., Mandressi, A. *et al.* (2015) Olfactory system of highly trained dogs detects prostate cancer in urine samples. *Journal of Urology* 193(4), 1382–1387. DOI: 10.1016/j.juro.2014.09.099.
- Wansink, B. and van Ittersum, K. (2005) Shape of glass and amount of alcohol poured: comparative study of effect of practice and concentration. *BMJ* 331(7531), 1512–1514. DOI: 10.1136/bmj.331.7531.1512.
- Weetjens, B.J.C., Mgoode, G.F., Machang'u, R.S., Kazwala, R., Mfinanga, G. *et al.* (2009) African pouched rats for the detection of pulmonary tuberculosis in sputum samples. *The International Journal of Tuberculosis and Lung Disease* 13(6), 737–743.

Effects of Age and Treatment

Age and Time Effects

Changes over time in the behaviour of animals can be caused by different factors. Differences in behavioural responses between age groups when comparing young, adolescent and adult animals are often due to differences in development, such as physiological, physical or even morphological changes with time. An example of the latter would be the absence and presence of horns in some ruminant males, which affects the ability to attain dominance, even if mainly through posturing. Another time effect is increased experience, both in the short and long term. Older animals – all other things being equal – will have had more time to acquire certain skills and will have been exposed to a greater variety of situations. However, the reverse effect of ageing also occurs, where older animals may react differently to their younger conspecifics due to a deterioration in some form of sensory, physical or mental capacity.

Many behavioural tests will be suitable for testing any of the three categories mentioned above, but disentangling the relative contribution of age and experience can sometimes be difficult. In the following, examples will be given of tests employed to investigate behavioural effects of ageing, development and experience, respectively. It should be noted that the behavioural protocols chosen in these examples may also be used in other contexts.

Behavioural aspects of ageing

Biological ageing or senescence can be defined as a gradual deterioration of functional characteristics with time. This is an area of research of increasing importance for human health and welfare, especially in terms of neurodegenerative disorders, such as Alzheimer's disease. As a result, more and more experiments are carried out to simulate the progression of senescence, with animal models being developed in both rodents and pigs (e.g.

Søndergaard *et al.*, 2012; Park *et al.*, 2019). In general, however, we know relatively little about the effects of ageing in most of the animal species managed by humans. This is mainly due to their being put down, either before they reach maturity, as is the case with most meat-producing livestock, or before severe deterioration sets in, which is often (but unfortunately not always) the case with companion animals.

Compared with larger domestic species, the lifespan of rats and mice is relatively short, and rodents are often used in research related to the effects of ageing. One method is to assess differences in learning ability between young and old rats in a Morris water maze where the rats are trained to find a hidden platform (Fig. 7.1a). In the example below, the circular tank used was 150 cm in diameter with a water depth of 60 cm. In order to hide the platform, the water is made opaque by adding water-based colouring, often odourless paint. Villarreal *et al.* (2002) painted the inside of the tank black, and added black paint to the water to prevent the rats from seeing the platform before they reached it. The water was heated to 25°C, and a square platform (12 cm by 12 cm) was placed in one of the four quadrants of the tank 2 cm below the surface of the blackened water. The platform remained in the same place throughout the training, but the rats were released from different places around the edge of the tank. Visual cues were placed in the room to allow the rats to navigate towards the platform. Each rat was trained three times a day for 8 days, and the maximum duration of a swim was 90 s. If the rat did not locate the platform within this time, it was gently caught and placed on the platform for 15 s.

The researchers tested male Fisher rats at two different ages: young rats of 2–4 months of age, and old rats of 20–22 months of age. The latter is the mean life expectancy of this breed (Chesky and Rockstein, 1976), and

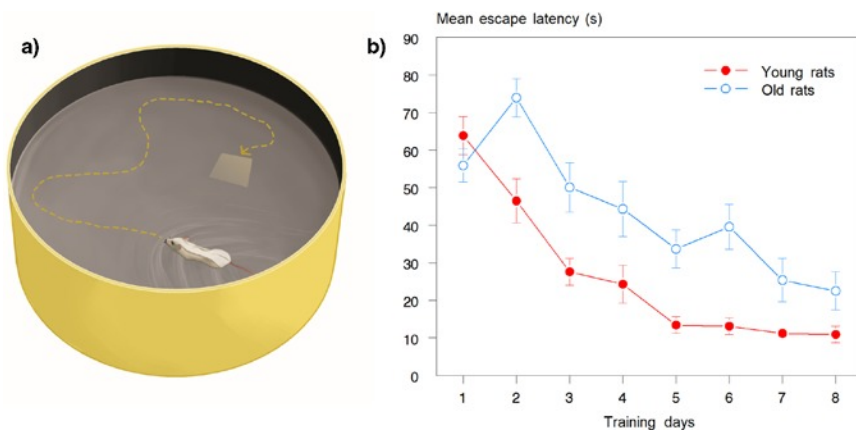


Fig. 7.1. (a) Young and old rats trained in a Morris water maze to locate a submerged platform 2 cm below the surface of the opaque water; (b) older rats take longer to learn the position of the platform (data from Villarreal *et al.*, 2002).

the equivalent in humans would be 79 years, which is the current average life expectancy of men in Western Europe (Statista, 2019). Indeed, there were signs that the older rats had impaired hearing, although this was unlikely to affect their performance in the test situation.

Rats are motivated to escape from the tank, and will actively search for the platform. [Figure 7.1b](#) shows the mean escape latency for young and old rats. The latter took much longer to find the platform than the young rats, which appeared to have reached their minimum latency possible already at training day 5. Interestingly, swimming speed did not differ between the two age groups, so the increased latency was not due to differences in velocity, which remained at roughly 20 cm/s. This indicates that the older rats had impaired learning ability compared with their younger conspecifics, and were still trying to learn the task on the final training day.

Other parameters can be measured when using a Morris water maze. Also used in the open field test (Chapter 4), degree of thigmotaxis, or propensity to stay close to the edge of the tank, is one measurement found in articles applying this testing method. There are also several variables measured during so-called probe tests, which are when the animal is tested but the platform is not present. This is to quantify the search near the position where the platform should have been, thereby showing the experimenter that the animal has indeed learnt the position of the platform: Moser's zone, Whishaw's corridor and Gallagher's proximity all form part of this palette of measures. Moser's zone is time spent in the area surrounding the (now absent) platform, measured at various distances from its centre (Moser *et al.*, 1993). This is often expressed as a percentage of the total surface area, so in the example above, a Moser's zone of 20 cm radius around the platform is equivalent to 7.3% of the total water surface. Spending significantly more than this percentage of time in the zone indicates that the animal has learnt where the platform should be. Whishaw's corridor measures time and length of the swim path in a strip leading directly from the release point to the platform location (Whishaw, 1985). Gallagher's proximity is the average distance of the animal from the centre of the platform location across the duration of the probe test (Gallagher *et al.*, 1993). This last measure has been found to be the most sensitive for detection of group differences (Maei *et al.*, 2009). The validity of the probe test has been questioned as some of the differences seen between groups are heavily influenced by degree of thigmotaxis, and passive floating (Wolfer *et al.*, 1998), which should be taken into account in the analysis of the recorded data. In addition, the Morris water maze test is likely to be an aversive experience for the rats, because the rats are being forced to swim and the escape platform is not immediately obvious. The stressfulness of the test is highest for rats that have not yet learnt to find the platform (Villarréal *et al.*, 2002) and the probe test mentioned above would also be very stressful as the platform has disappeared. Such animal welfare aspects will be

discussed below in connection with examples of tests of treatments, and further developed in Chapter 12.

Behavioural aspects of development

In the previous section, the behavioural changes seen in older animals were a result of a reduced capacity to learn as a consequence of senescence. However, differences between generations can also be the result of developmental changes. Some of these are physiological, morphological and cognitive changes resulting from ontogeny, the growth and development of the individual animal. Many of these changes are a result of improvements in the physical capacity of the animal as it develops, giving rise to a wider behavioural repertoire and with behaviour expressed in a more competent manner. However, it is worth noting that certain types of behaviour, such as the occurrence of play, are more common in young animals (Boissy *et al.*, 2007; Palagi, 2018).

One of the classic examples of behavioural development with age is the display of dustbathing behaviour in poultry species. Dustbathing, described in Chapter 2, consists of a series of behavioural components performed in a specific order. When the chicks are hatched, they are not yet able to perform a complete series of dustbathing movements; the individual components develop over time and – for some of them – always in the same order. Borchelt (1977) tested a group of newly hatched bobwhite quail by giving them access for 30 min to a tray of sifted, dry earth every day at 14:00 h. Different behaviours were noted for each chick, and Fig. 7.2 shows the cumulative percentage of birds performing the different components of dustbathing, with dust tossing seen in all chicks before head rubbing occurred, which in turn was followed by the development of side rubbing.

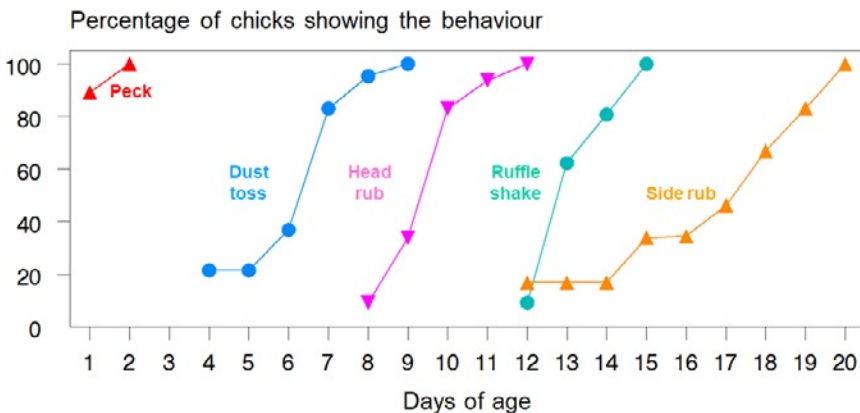


Fig. 7.2. Development over time of different components of dustbathing behaviour in bobwhite quail chicks (adapted from Borchelt, 1977).

This is a nice demonstration of behavioural ontogeny in the early stages of an animal's life.

Behavioural aspects of experience

Experience is part of the learning process, and in animals can be defined as the process of gaining knowledge or skills from doing or sensing something. The simplest form is habituation, where repeated exposure to a stimulus, which does not result in reward or punishment, diminishes the initial response seen at first exposure. Examples of this were shown in the habituation–dishabituation tests described in Chapter 6. Habituation differs from our usual concept of learning as it does not involve the animal acquiring new behaviours, but rather loss of existing responses. It can be difficult to distinguish from fatigue or sensory adaptation, which are when the animal can no longer express the behaviour or sense the stimulus. If the habituation is carried out in stages, the process is sometimes referred to as desensitization. Changes in behaviour following repeated exposure are often easy to monitor and quantify, and because habituation is often relatively quick to apply, there is not a confounding effect of age to take into account. In the following, I will describe two examples of behavioural aspects of experience: one is a desensitization of horses to a variety of objects in their environment, and the other demonstrates how long-term effects of experience can be ascertained, using an example of social experience with different phenotypes in chickens.

Habituation is usually thought to be stimulus specific. In other words, the animal gets used to the particular neutral object or gesture to which it has been repeatedly exposed until habituated, so that the animal ceases to respond to the stimulus. Christensen *et al.* (2011) wanted to know if horses were able to generalize across physical stimuli that differed in shape, size and colour. First, a large group of horses were habituated to being fed on their own from two feed buckets in a large test arena (10 m × 10 m). Then half of the horses ($n = 15$; Hab group) were habituated to eat from the two buckets in the presence of five different objects, which were previously unknown to the horses. These objects were an open, white umbrella; a yellow box; a purple ball; a blue cylinder; and a green frustum (a cone with the top cut off). The dimensions of the objects were between 40 and 100 cm, and they were all present during the habituation. The other half of the horses (Control group) were also fed individually from the two buckets, but with no objects present. The time taken to start eating was measured for both groups on the last day before testing (pre-test; Fig. 7.3) to ensure that the groups did not differ. On the test day, all horses entered the test arena individually as usual, but for all the horses one of the five objects was present. This time, the object was placed in front of the feed buckets, and latency to start eating was measured. The final test was identical, except that the object was behind the feed buckets and novel for all horses, being a red and

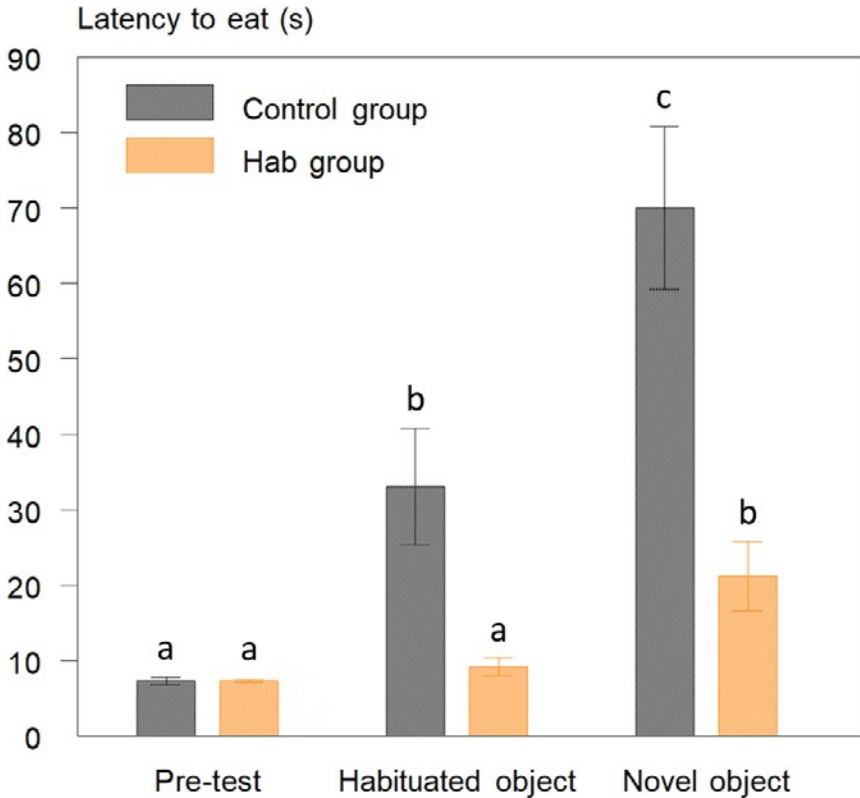


Fig. 7.3. When horses have experience with physical objects they are quicker to begin eating in their presence. Horses were habituated to feed either in the presence of five objects (Hab group) or not (Control group). This bar chart shows the latency to start eating (mean $\pm$ SEM) for both groups of horses following this initial habituation (pre-test). The same measure was recorded during two subsequent tests (i) in the presence of only one of the objects to which the Hab group had been habituated but placed in a novel position (habituated object), and (ii) with an object present that was novel to both groups (novel object). Different letters indicate significant differences at $P < 0.05$ (adapted from Christensen *et al.*, 2011).

white striped cone. It is clear from Fig. 7.3 that the horses that had been habituated to the presence of the five objects (Hab group) did not change their behaviour when only one of the objects was present in a novel location. The control horses, however, greatly increased their latency to eat in this situation. Habituated horses also reacted less profoundly to the novel object than did control horses, and only control horses had an increase in heart rate during the two tests. This indicates that horses are able to generalize across different stimuli, applying experience gained from exposure to other stimuli in similar situations. If you are particularly interested in this

species, an excellent reasoning for the application of learning theory in the training of horses can be found in the article by McLean and Christensen (2017).

Previous experience can also have long-term effects on animal behaviour. An example of this is the higher risk of behavioural problems in dogs that have been mistreated by a previous owner (Vitulová *et al.*, 2018). Experience can also alter the behaviour of wild animals, with examples of female elk changing their behaviour to avoid hunters (Thurfjell *et al.*, 2017), or the likelihood of capture in wild birds diminishing with their age and number of previous trappings, thereby changing the representativeness of the samples (Camacho *et al.*, 2017). The behavioural test chosen as an example of long-term effects of previous experience comes from chickens, and only a subset of the results (group size 40) is presented to illustrate the method. Marin *et al.* (2014) wanted to know if previous experience with conspecifics of different phenotypic appearance would have long-term effects on how the birds reacted to diversity within their social group. Although breeds with different plumage colours could have been used, this would have introduced an additional genetic effect, as chicken breeds are known to differ in their behavioural expression (Hughes and Duncan, 1972). Instead, the researchers changed the phenotypic appearance of some of the birds by dyeing the feathers on the back of their neck. This also allowed the markings to be removed, making a test for phenotypic changes in both directions possible.

Figure 7.4 shows the marking protocol over time in the experimental groups used. The chicks were housed after hatch in groups of 40, and markings were applied (and maintained if fading) from day 1 until the birds were 34 weeks of age. During this period, the groups had either 0, 30, 50, 70 or 100% of the birds marked. When the birds were 34 weeks of age, birds in the groups from the two extreme treatments (0 and 100% marked) were re-marked so that 30% of the birds were now either marked or unmarked. This was repeated when the birds were 38 and 44 weeks of age, with an increasing number of birds being marked/unmarked (Fig. 7.4). When the first changes were made at 34 weeks of age, the body weight of the phenotypically changed birds and the egg production of their groups were reduced compared to the groups with the same percentage of birds marked since hatching. In the groups going from 0 to 30% marked birds, this could be due to the novel stress of being handled for marking in these chickens, as the marked birds in the other groups had been accustomed to this throughout their life. However, the de-marked birds in the groups going from 100 to 70% marked individuals also lost weight, and the authors suggest that the novelty of birds with different appearance led to aversive social interactions towards them by their group mates. Also, when additional birds had their neck plumage changed in weeks 38 and 44, no differences in growth or production were found. These results indicate that changes in the markings of the birds when adult initially affect them

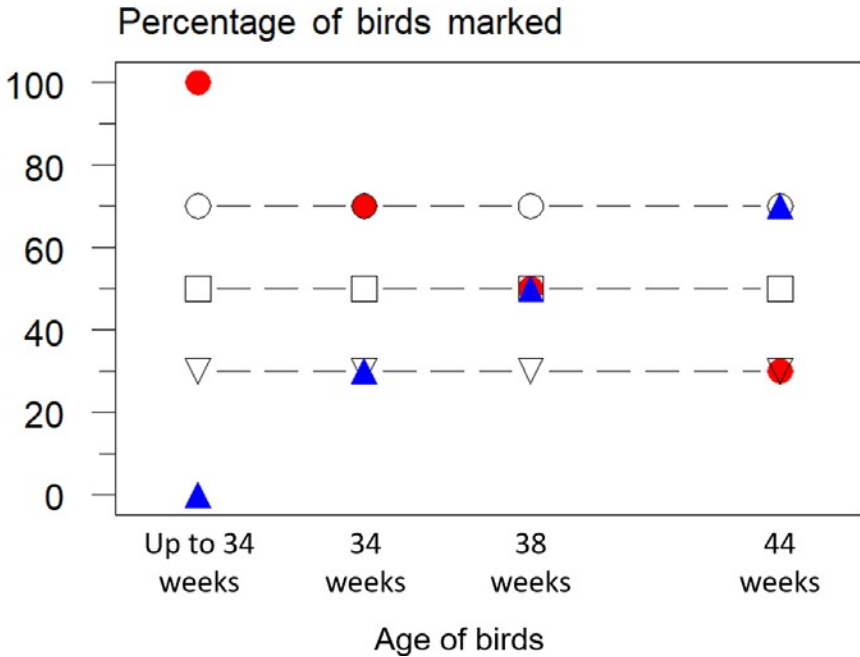


Fig. 7.4. Percentage of birds marked in each group over time. The markings were black dye applied to or removed from the back of the neck of the birds. Open symbols indicate the groups where the percentage of marked birds was not changed, whereas filled symbols denote groups where the percentage of marked birds was increased (▲) or decreased (●) at different time points (adapted from Marin *et al.*, 2014).

negatively compared with the birds housed with a mixture of phenotypes since hatching. This effect appears to diminish with time, as further changes in the appearance of other birds in the group did not have the same affect. The authors suggest that being used to conspecifics with different appearances from an early age prevents the negative effects of changes.

Tests of Effects of Treatment

One could argue that most behavioural tests of animals can be put into the category of tests of effects of treatment. It is after all one of the cornerstones in applied ethology to test various situations against each other and measure how the animal responds. In this context, treatments could be something we do or don't do to the animal, such as giving a drug or a placebo to a mouse, and trimming (or not) the hooves of a cow. A treatment can also be external to the animal, such as raising dogs in barren or enriched environments, and exposing rats to being tickled or not. In the following, examples of tests relating to each of these aspects, internal or

external to the animal, will be shown, although many of them can be applied in both situations.

Internal treatment

One of many tests developed to investigate effects of drugs on anxiety-like behaviour in rats is the test of novelty-suppressed feeding (NSF; Bodnoff *et al.*, 1988). The idea behind the test is that when animals motivated to eat are given the opportunity to do so, but in an unknown environment, they will be more reluctant to begin feeding and eat less than expected. An important aspect of the test is to ensure that the drug treatment is not inducing some form of nausea or other appetite-reducing effects. It is therefore necessary also to observe the feeding behaviour in the home environment to ensure that reluctance to approach the feed in the test situation is not due to a lack of motivation to eat. In the original test, Bodnoff *et al.* (1988) presented each rat with 12 pellets in a novel environment with a maximum test duration of 6 min, and measured only latency to start eating. Others have used a test duration of 10 min, and only one pre-weighed pellet to include also a measure of feed intake (e.g. Miragaia *et al.*, 2018). In both studies, the rats had been food deprived for 48 h prior to testing, and it could, and should, be questioned whether this length of fasting is necessary to obtain valid results (see Chapter 12). Rats treated for 2 weeks with an anxiolytic, i.e. drug used to relieve anxiety, started eating the pellets much sooner than the non-treated group (after 80 s compared with 271 s on average). When their latency to eat after fasting was tested in the home cage, rats in either group started eating after 48 s (Bodnoff *et al.*, 1988). In other words, administration of an anxiety-reducing drug reduced the latency to start eating in a novel environment, which serves as validation that the test is assessing the degree of anxiety experienced by the rat during the test. An evaluation of behavioural tests used in pharmacological research can be found in Hånell and Marklund (2014).

Forced swim test and tail suspension test

Some of the behavioural tests used to assess the efficacy of pharmacological compounds aimed to reduce negative states, such as anxiety and depression, are even less pleasant than the example above. One such test, developed in rats by Porsolt *et al.* (1977), is the forced swim test. It is based on the observation that a rat or a mouse, when forced to swim with no apparent means of escape, after a while will remain immobile, only moving enough to keep its head above water. This is thought to reflect a state of despair, and has been used to test the efficacy of potential anti-depressant drugs. The test is usually carried out over 2 days, using a large beaker half-filled with lukewarm water. On the first day, the rat or mouse is put into the water and left to swim or float for 15 min (Slattery and Cryan, 2012). This

is a considerable environmental stressor, and the animal will usually make several escape attempts before resigning itself to floating. The animal is then lifted out, dried and returned to its home cage. On the second day, the test is repeated, but the animal is only left in the beaker for 5 min. Without any drug treatment, most rats will float for about 75% of the time, which according to Porsolt *et al.* (1977) serves as a suitable baseline allowing treated rats to reduce but also increase their amount of floating. The variables recorded are most often number of escape attempts, latency to start floating and duration of floating, but other behavioural responses may be included, such as head bobbing (Paré, 1994). The test has been criticized for a number of reasons, not least because of its welfare implications, but also because its validity has been questioned (see below).

The forced swim test was also mentioned briefly in Chapter 3, where Fig. 3.3 showed the increasing use of this test over the past decades. Use of the tail suspension test follows the same pattern over time. This is probably because the two tests are often used in the same experiment, although the tail suspension test is unsuitable for rats, as suspension by the tail is painful in this species. As the name indicates, a mouse is suspended using a length (15–20 cm) of sticky tape, with one end attached to the end of the animal's tail and the other end stuck on to a solid fixture, such as a shelf, allowing enough space for the animal to be suspended without touching any vertical walls nearby. To prevent the mouse from climbing up its own tail to the tape, a small piece (4 cm) of transparent solid tube is sometimes placed over the base of the tail (Can *et al.*, 2012). The mouse is suspended for 6 min, and the one variable scored is the time spent immobile.

Both the forced swim test and the tail suspension test are based on the premise that more depressed animals will give up the struggle to escape sooner and for longer than less depressed animals. Injecting a test animal with an anti-depressant drug should therefore increase the duration of swimming and number of escape attempts in the forced swim test, and reduce immobility in the tail suspension test compared with animals treated with saline (e.g. Ahmed *et al.*, 2009). However, the validity of these two tests continues to be questioned (Reardon, 2019), and I will summarize here the concerns raised by Dzirasa and Covington (2012). The first issue they raise is the timing, as many anti-depressants do not have acute effects when used in humans, but take time to work. This does not tally with the use of tests carried out in animals soon after injection. Sometimes the rodents used have been genetically modified to induce depression, and this may affect their motivation and ability to move. A third concern is that the test does not allow the animal to escape, whatever it does. Dzirasa and Covington (2012) therefore argue that it is more adaptive for the animal to stop trying and conserve energy. Their final argument against these tests is that hopelessness is not part of the diagnosis for major depressive disorders in humans, and as this is the essential measurement in both tests, one can question their suitability as a test model for anti-depressive drugs.

In a recent meta-analysis of the use of forced swim tests in mice to assess effects of anti-depressants, Kara *et al.* (2018) concluded that the test did not show consistent dose-response correlations, and comparisons across experiments should be avoided.

External treatment

Unless we are interested in immediate reactions to sudden changes in the animal's surroundings, we need to expose the animal to different external treatments, such as barren and enriched environments, for some time. But how can we compare being raised in barren and enriched environments, when they differ in so many ways for the animal to express its behaviour, in terms of social interactions, feed choice, space allowance and equipment with which to interact? One way to overcome this, at least in part, is to carry out a suitable behavioural test in a standardized manner after some time, and with the test being novel to animals raised in either environment.

Many aspects of behaviour can be affected by the complexity of the environment in which an animal is raised. One of these relates to changes in cognitive ability, with animals raised in barren surroundings performing less well than their enriched conspecifics in tests of learning and memory (Brantsæter *et al.*, 2016). A hole-board test was carried out using pigs kept in groups of ten (Grimberg-Henrici *et al.*, 2016). Compared with the barren environment, the enriched group had double the floor area and it was covered in straw. They also had access to peat for rooting and a variety of enrichment objects (wooden sticks, balls, as well as jute bags and ropes). After 5 weeks in these environments, the pigs were tested in a spatial hole-board discrimination test. This consisted of an arena (5.4 m × 5.4 m) with 4 × 4 holes in the floor under which food bowls were attached. The 16 holes were 95 cm apart, and covered with opaque semi-domes (red balls cut in half) to prevent the pig from seeing the content. For each test, four bowls were baited with one sugar-coated chocolate sweet each and in one of the four patterns shown in Fig. 7.5. In reality, all bowls contained sweets, but these were placed under a false bottom, and thus inaccessible

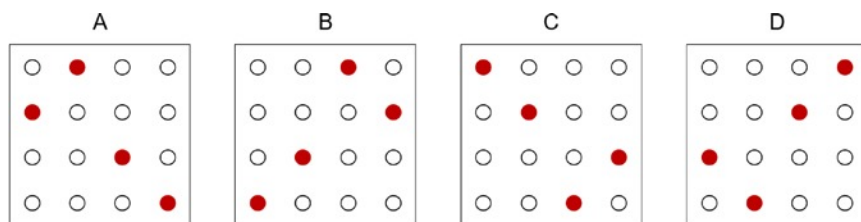


Fig. 7.5. Configuration of baited (●) and unbaited (○) holes in a hole-board test used to assess cognitive performance of pigs raised in either an enriched or barren environment (from Grimberg-Henrici *et al.*, 2016).

to the pig, to prevent the pigs from using odour cues to locate the sweets. The pigs were fed only part of their daily ration the morning of testing to motivate them to search for food. It has to be added that, before the tests began, the pigs went through a lot of habituation to the human handler, the test arena, the holes, the ball coverings, as well as the sweets themselves. For the real test sessions, the test pig entered the arena through one of four doors in a random pattern, only one of which was open for each session. First, the pigs were tested on one specific pattern of baited holes (acquisition phase, 40 sessions) and then exposed to the diagonally mirrored pattern (A vs C, and B vs D or *vice versa* in Fig. 7.5). This reversal phase lasted 20 sessions per pig. For each of these sessions, a number of parameters were recorded, including the order of visits to holes, whether the visits were rewarded (baited holes) or not, how many sweets were consumed, as well as the latency and duration of the search. Each session lasted a maximum of 5 min, but was ended sooner once all four baited holes had been visited. During the acquisition phase, enriched pigs had fewer visits to unbaited holes than did the pigs raised in barren pens, and they also had fewer revisits to holes during the reversal phase. In both phases, the enriched pigs also had a shorter latency to search for the sweets compared with the barren pigs. Environmental enrichment thus improved the cognitive performance of the pigs in this spatial learning task, albeit only slightly. The experimenters suggest that the long habituation and handling necessary to perform this type of test could have acted as a form of enrichment for the barren pigs, thereby reducing the magnitude of the differences found.

Tests, such as the one described above, are devised to compare animals exposed to treatments that give rise to so many differences in their situation that observations are extremely difficult if not impossible to compare in an unbiased manner. In the example with the enriched vs barren environment, activity in the home pen is likely to be greater in the enriched group simply because there is more room in which to manoeuvre. If the barren pigs were tested in a larger pen, the mere novelty of the increased space would lead to increases in locomotion, play and perhaps even fighting. When working on a project in our laboratory, where we investigated if rats could learn to associate an odour with being tickled (yes, they could), we were unable to compare the behaviour of the rats during tickling, which mimics the rough-and-tumble play seen in adolescent rats (LaFollette *et al.*, 2017; Hammond *et al.*, 2019), with that of untickled rats. Instead, we compared their behaviour in the pauses between tickling bouts with behaviour during the same time slots for the untickled conspecifics (Bombail *et al.*, 2019). We found that tickled rats spend a larger proportion of their time focusing on the immobile hand of the human tickler and performed more play-jumping than the untickled rats, which in turn were more active. In this manner we created a form of ‘test within treatment’, comparing like with like to the greatest extent possible.

Concluding Remarks

As stated at the beginning of this chapter, the examples given here are but a fraction of possible behavioural tests that can be used to test effects of ageing, experience and various treatments. When testing the behavioural responses of animals, we should aim for tests that impose a minimum of discomfort and a maximum of joy whenever possible. The criticism of the forced swim test mentioned above should be taken into consideration when planning experiments, and development of better, less stressful behavioural tests to replace these methods is needed. In Chapter 12, aspects of animal welfare and ethical considerations are discussed together with the existing legislation in Europe and the USA for use of animals in scientific experiments. It should not be forgotten, however, that animal behaviour testing is not restricted to scientific use, and all companion animal owners are encouraged to engage their pets in behavioural training and tests, which, if done well and without aversive treatments, lead to animals better adapted to domestic life.

References

- Ahmed, A.A.E., Al-Rasheed, N.M. and Al-Rasheed, N.M. (2009) Antidepressant-Like effects of rosiglitazone, a PPAR γ agonist, in the rat forced swim and mouse tail suspension tests. *Behavioural Pharmacology* 20(7), 635–642. DOI: 10.1097/FBP.0b013e328331b9bf.
- Bodnoff, S.R., Suranyi-Cadotte, B., Aitken, D.H., Quirion, R. and Meaney, M.J. (1988) The effects of chronic antidepressant treatment in an animal model of anxiety. *Psychopharmacology* 95(3), 298–302. DOI: 10.1007/BF00181937.
- Boissy, A., Manteuffel, G., Jensen, M.B., Moe, R.O., Spruijt, B. *et al.* (2007) Assessment of positive emotions in animals to improve their welfare. *Physiology & Behavior* 92(3), 375–397. DOI: 10.1016/j.physbeh.2007.02.003.
- Bombail, V., Jérôme, N., Lam, H., Muszlak, S., Meddle, S.L. *et al.* (2019) Odour conditioning of positive affective states: rats can learn to associate an odour with being tickled. *PLoS ONE* 14(6), e0212829. DOI: 10.1371/journal.pone.0212829.
- Borchelt, P.L. (1977) Development of dustbathing components in bobwhite and Japanese quail. *Developmental Psychobiology* 10(2), 97–103. DOI: 10.1002/dev.420100202.
- Brantsæter, M., Nordgreen, J., Rodenburg, T.B., Tahamtani, F.M., Popova, A. *et al.* (2016) Exposure to increased environmental complexity during rearing reduces fearfulness and increases use of three-dimensional space in laying hens (*Gallus gallus domesticus*). *Frontiers in Veterinary Science* 3(2), 14. DOI: 10.3389/fvets.2016.00014.
- Camacho, C., Canal, D. and Potti, J. (2017) Lifelong effects of trapping experience lead to age-biased sampling: lessons from a wild bird population. *Animal Behaviour* 130, 133–139. DOI: 10.1016/j.anbehav.2017.06.018.
- Can, A., Dao, D.T., Terrillion, C.E., Piantadosi, S.C., Bhat, S. *et al.* (2012) The tail suspension test. *Journal of Visual Experiments* 59, 3769.

- Chesky, J.A. and Rockstein, M. (1976) Life span characteristics in the male Fischer rat. *Experimental Aging Research* 2(5), 399–407. DOI: 10.1080/03610737608257998.
- Christensen, J.W., Zharkikh, T. and Chovaux, E. (2011) Object recognition and generalisation during habituation in horses. *Applied Animal Behaviour Science* 129(2–4), 83–91. DOI: 10.1016/j.applanim.2010.11.016.
- Dzirasa, K. and Covington, H.E. (2012) Increasing the validity of experimental models for depression. *Annals of the New York Academy of Sciences* 1265(1), 36–45. DOI: 10.1111/j.1749-6632.2012.06669.x.
- Gallagher, M., Burwell, R. and Burchinal, M.R. (1993) Severity of spatial learning impairment in aging: development of a learning index for performance in the Morris water maze. *Behavioral Neuroscience* 107(4), 618–626. DOI: 10.1037/0735-7044.107.4.618.
- Grimberg-Henrici, C.G.E., Vermaak, P., Bolhuis, J.E., Nordquist, R.E. and van der Staay, F.J. (2016) Effects of environmental enrichment on cognitive performance of pigs in a spatial holeboard discrimination task. *Animal Cognition* 19(2), 271–283. DOI: 10.1007/s10071-015-0932-7.
- Hammond, T., Bombail, V., Nielsen, B.L., Meddle, S.L., Lawrence, A.B. *et al.* (2019) Relationships between play and responses to tickling in male juvenile rats. *Applied Animal Behaviour Science* 221, 104879. DOI: 10.1016/j.applanim.2019.104879.
- Hånell, A. and Marklund, N. (2014) Structured evaluation of rodent behavioral tests used in drug discovery research. *Frontiers in Behavioral Neuroscience* 8, 252.
- Hughes, B.O. and Duncan, I.J.H. (1972) The influence of strain and environmental factors upon feather pecking and cannibalism in fowls. *British Poultry Science* 13(6), 525–547. DOI: 10.1080/00071667208415981.
- Kara, N.Z., Stukalin, Y. and Einat, H. (2018) Revisiting the validity of the mouse forced swim test: systematic review and meta-analysis of the effects of prototypic antidepressants. *Neuroscience & Biobehavioral Reviews* 84, 1–11. DOI: 10.1016/j.neubiorev.2017.11.003.
- LaFollette, M.R., O’Haire, M.E., Cloutier, S., Blankenberger, W.B. and Gaskill, B.N. (2017) Rat tickling: a systematic review of applications, outcomes, and moderators. *PLoS ONE* 12(4), e0175320. DOI: 10.1371/journal.pone.0175320.
- Maei, H.R., Zaslavsky, K., Teixeira, C.M. and Frankland, P.W. (2009) What is the most sensitive measure of water maze probe test performance? *Frontiers in Integrative Neuroscience* 3, 4. DOI: 10.3389/neuro.07.004.2009.
- Marin, R.H., Liste, M.G., Campderrich, I. and Estevez, I. (2014) The impact of phenotypic appearance on body weight and egg production in laying hens: a group-size- and experience-dependent phenomenon. *Poultry Science* 93(7), 1623–1635. DOI: 10.3382/ps.2013-03705.
- McLean, A.N. and Christensen, J.W. (2017) The application of learning theory in horse training. *Applied Animal Behaviour Science* 190, 18–27. DOI: 10.1016/j.applanim.2017.02.020.
- Miragaia, A.S., de Oliveira Wertheimer, G.S., Consoli, A.C., Cabbia, R., Longo, B.M. *et al.* (2018) Maternal deprivation increases anxiety- and depressive-like behaviors in an age-dependent fashion and reduces neuropeptide Y expression in the amygdala and hippocampus of male and female young adult rats. *Frontiers in Behavioral Neuroscience* 12, 159. DOI: 10.3389/fnbeh.2018.00159.
- Moser, E., Moser, M.B. and Andersen, P. (1993) Spatial learning impairment parallels the magnitude of dorsal hippocampal lesions, but is hardly present

- following ventral lesions. *The Journal of Neuroscience* 13(9), 3916–3925. DOI: 10.1523/JNEUROSCI.13-09-03916.1993.
- Palagi, E. (2018) Not just for fun! Social play as a springboard for adult social competence in human and non-human primates. *Behavioral Ecology and Sociobiology* 72(6). DOI: 10.1007/s00265-018-2506-6.
- Paré, W.P. (1994) Open field, learned helplessness, conditioned defensive burying, and forced-swim tests in WKY rats. *Physiology & Behavior* 55(3), 433–439. DOI: 10.1016/0031-9384(94)90097-3.
- Park, H., Kang, S., Nam, E., Suh, Y.-H. and Chang, K.-A. (2019) The protective effects of PSM-04 against beta amyloid-induced neurotoxicity in primary cortical neurons and an animal model of Alzheimer's disease. *Frontiers in Pharmacology* 10. DOI: 10.3389/fphar.2019.00002.
- Porsolt, R.D., Lepichon, M. and Jalfre, M. (1977) Depression: a new animal model sensitive to antidepressant treatments. *Nature* 266(5604), 730–732. DOI: 10.1038/266730a0.
- Reardon, S. (2019) Depression researchers rethink popular mouse swim tests. *Nature* 571(7766), 456–457. DOI: 10.1038/d41586-019-02133-2.
- Slattery, D.A. and Cryan, J.F. (2012) Using the rat forced swim test to assess antidepressant-like activity in rodents. *Nature Protocols* 7(6), 1009–1014. DOI: 10.1038/nprot.2012.044.
- Statista (2019) The Statistics Portal. Available at: www.statista.com/statistics/274514/life-expectancy-in-europe (accessed 14 February 2019).
- Søndergaard, L.V., Ladewig, J., Dagnæs-Hansen, F., Herskin, M.S. and Holm, I.E. (2012) Object recognition as a measure of memory in 1–2 years old transgenic minipigs carrying the APPsw mutation for Alzheimer's disease. *Transgenic Research* 21(6), 1341–1348. DOI: 10.1007/s11248-012-9620-4.
- Thurfjell, H., Ciuti, S. and Boyce, M.S. (2017) Learning from the mistakes of others: how female elk (*Cervus elaphus*) adjust behaviour with age to avoid hunters. *PLoS ONE* 12(6), e0178082. DOI: 10.1371/journal.pone.0178082.
- Villarréal, J.S., Gonzalez-Lima, F., Berndt, J. and Barea-Rodriguez, E.J. (2002) Water maze training in aged rats: effects on brain metabolic capacity and behavior. *Brain Research* 939(1–2), 43–51. DOI: 10.1016/S0006-8993(02)02545-3.
- Vitulová, S., Voslařová, E., Večerek, V. and Bedáňová, I. (2018) Behaviour of dogs adopted from an animal shelter. *Acta Veterinaria Brno* 87(2), 155–163. DOI: 10.2754/avb201887020155.
- Whishaw, I.Q. (1985) Cholinergic receptor blockade in the rat impairs locale but not taxon strategies for place navigation in a swimming pool. *Behavioral Neuroscience* 99(5), 979–1005. DOI: 10.1037/0735-7044.99.5.979.
- Wolfer, D.P., Stagljär-Bozicevic, M., Errington, M.L. and Lipp, H.-P. (1998) Spatial memory and learning in transgenic mice: fact or artifact? *Physiology* 13(3), 118–123. DOI: 10.1152/physiologyonline.1998.13.3.118.

Reinforcement and Punishment

The learning paradigms that are essential for animal training are also inherent in many of the behavioural tests we use, not least the operant conditioning examples given in Chapter 5. This book would therefore not be complete without a chapter on this type of associative learning. I shall admit that it was always going to be the most difficult chapter to write. Not because there are fewer examples to choose from, quite the opposite, but because the concepts of positive and negative reinforcement and punishment are difficult to get your head around. I have always had trouble with the notion of ‘positive punishment’, but we’ll get to that later.

The focus of this chapter will be on the concept of reinforcement, and how it is different from punishment. I will use the opportunity to include more examples from wild animals under the management of humans. Horses and calves will also feature, and clicker training will be briefly described, as it is based on the learning principle of positive reinforcement.

Positive and Negative Reinforcement

Reinforcement is used in the training of animals, and the goal is to encourage specific behaviours. Making a behaviour more likely to occur is the essence of reinforcement, in stark contrast to the use of punishment, which makes a behaviour less likely to occur. In [Fig. 8.1](#), I have tried to make a simple diagram to illustrate these differences, and to highlight what is meant when we use the words positive and negative. Positive refers to a stimulus or reinforcer being added or applied, whereas negative means a stimulus being removed or disallowed. Positive reinforcement is therefore to provide an attractive stimulus, such as food, to encourage the behaviour shown by the animal. Negative reinforcement, on the other hand, is to remove an aversive stimulus when the animal shows the behaviour we want to promote. Aversive just means something the animal wants to avoid, and it does not have to involve a painful or frightening stimulus (Innes and McBride, 2008). An example of this is a rider pulling the reins of a

	<i>Punishment:</i> ↓ Behaviour	<i>Reinforcement:</i> ↑ Behaviour
Positive: Add (+)	Aversive stimulus	Attractive stimulus
Negative: Remove (–)	Attractive stimulus	Aversive stimulus

Fig. 8.1. The (sometimes tricky) relationship between positive and negative reinforcement and punishment. *Reinforcement* makes a behaviour more likely to happen, *punishment* makes a behaviour less likely. *Positive* refers to a stimulus or reinforcer being added, such as a food reward when the animal shows a behaviour we want to encourage. *Negative* refers to a stimulus being removed or denied, i.e. negative reinforcement means removing an aversive stimulus as soon as the desired behaviour is shown by the animal. Please note the inverse relationship between the valence of the stimulus/reinforcer (i.e. nasty vs nice) for the two types of training.

horse, which is mildly unpleasant for the animal, to motivate the horse to stop moving. As soon as the horse stops (i.e. the desired behaviour that we want to encourage), the reins are immediately made slack to remove the aversive stimulus (the unpleasantness of the pulling), thereby rewarding the horse for displaying the correct behaviour, i.e. stopping or slowing. In other words, we are removing the nasty stimulus, and training the horse using negative reinforcement.

Let us start with positive reinforcement, as this is possibly the most animal welfare-friendly way to train animals, involving no aversive stimuli if carried out correctly. So, when the animal displays the desired behaviour, this is rewarded by giving the animal an attractive stimulus, such as food. The animal will learn to associate the behaviour it has just performed with a pleasant experience (the reward), and the behaviour will increase in frequency or be performed more readily. This is the principle used in operant conditioning, where an animal presses a lever (or performs some other form of operant response), which releases a reward. Examples of this technique were given in Chapter 5 where the technique was used to ask what price an animal was willing to pay for access to a given resource. Positive reinforcement was first described in an almost incomprehensible article

by Premack (1959). The technique has been extensively used to train wild animals in captivity to allow clinical examination and blood sampling with a minimum of stress, including marine mammals (Brando, 2010) and grizzly bears (Joyce-Zuniga *et al.*, 2016), the latter example is described later in this chapter. It can, however, be quite difficult to train animals using *only* positive reinforcement. When training your dog to sit on command, you give it a titbit only when the desired behaviour, i.e. sitting, occurs in response to your chosen command-word, usually 'sit' if your dog speaks English. However, the dog may do lots of other behaviours, including nuzzling you because you are a source of food rewards. This behaviour should be ignored, and by doing so, you are using negative punishment, which is the withholding of a positive stimulus to make a behaviour, the nuzzling, disappear. The concept of negative punishment is described in more detail below.

Another problem encountered during positive reinforcement is how to reward behaviours that are not optimal, but not completely wrong either. One method is to break the desired behaviour into sections, thereby gradually shaping the behaviour of the animal, which is required to perform more and more complex behaviour to obtain the reward. This technique is often used when training an animal for an operant conditioning trial. Initially, the animal is rewarded for just sniffing the lever. Then the animal has to touch the lever to obtain a reward, then to press it once, and finally press it several times before the reward is given. Fischer and Wegener (2018) suggest that we should not reward in a binary fashion (i.e. either presence or absence of reward dependent on the behaviour shown), but instead make the rewards more graded, allowing the animal to learn from the amount of reward which behaviours are more desired than others. The experimenters used varying amounts of diluted red grape juice to train macaque monkeys to keep their gaze fixed on a central dot on a screen, while detecting a peripheral feature change without making any eye movements. As that sentence indicates, this was a complex learning task, which nevertheless was mastered by the three monkeys trained. Such graded rewarding may, however, not be a suitable method for species with less cognitive ability, whereas shaping, as described above, has been successfully applied to a number of species, including training goldfish to play basketball.

Negative reinforcement is often confused (e.g. Marranzino, 2013) with positive punishment, which is an unfavourable stimulus delivered to decrease the occurrence of a behaviour. As shown in [Fig. 8.1](#), the negative refers to the removal of something unpleasant, and reinforcement is always aimed at increasing a given behaviour. A lot of horse training is based on negative reinforcement, illustrated by the example given at the start of the chapter on pulling a horse's reins. As the aversive stimulus is often some kind of pressure applied to the horse, Ahrendt *et al.* (2015) investigated if the pressure needed to elicit a correct response would decrease as the

horse learnt to ‘turn off’ the negative stimulus by responding appropriately. They applied increasing pressure to the middle of the hindquarter of each horse using an algometer and measured the force necessary to make the horse move away from the pressure, at which point the pressure was removed immediately to reward the appropriate response, i.e. the movement performed by the horse. It is important to note that if the maximum pressure that was measurable on the algometer was reached (30 newtons), extra force was applied with the experimenter’s free hand to make the horse move. Had they instead removed the pressure, the horse would have been wrongly rewarded for not moving. The pressure test was carried out several times for each horse, and the pressure needed for a test horse to move decreased significantly over the first day of testing (Fig. 8.2), showing that the horses learnt the task from the negative reinforcement. On the subsequent days no decrease in pressure was seen, but the horses responded at the lower level of pressure, thus indicating that they remembered the required behaviour. Christensen *et al.* (2017) noted that the diminishing response seen on day 1 (Fig. 8.2) is sometimes interpreted as an increase in pain sensitivity in studies of nociceptive thresholds, as they often employ the same experimental protocol using repeated stimulation. As the stimulation

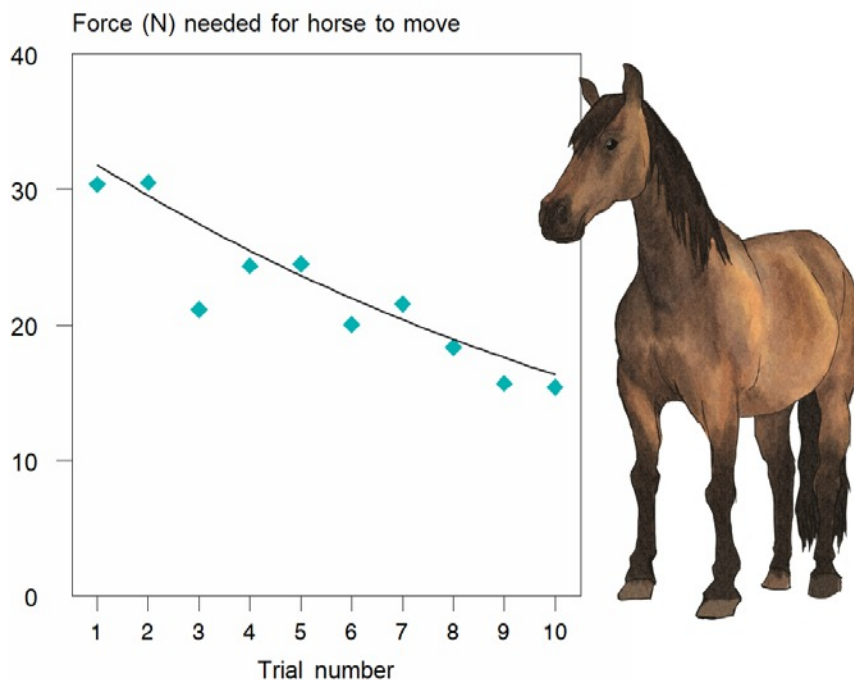


Fig. 8.2. Development over trials within the first test day of mean force (N) needed to make a horse move (adapted from Ahrendt *et al.*, 2015).

is removed when the animal shows avoidance behaviour, this behaviour is negatively reinforced, and the learning paradigm will interfere with any changes in sensitivity.

The use of only positive reinforcement training can take longer than when using a combination of positive and negative reinforcement. Wergård *et al.* (2015) compared these two approaches in female macaques in order to train them to move to a shelf in the cage and accept being briefly enclosed. Upon arrival in the experimental station, the monkeys were not used to humans. They were housed in pairs in 2-m-high cages provided with enrichment material, and the monkeys had tactile, visual, auditory and olfactory access to neighbouring cages. The trainer wore clothing that differed in colour from that of husbandry staff, and the two treatment groups were housed in separate rooms to prevent the training of one group influencing that of the other. Before the different training protocols were applied, all monkeys were first habituated to the cages, and then desensitized to the presence of the trainer by repeated exposure, gradually decreasing her distance to the cage. The positive reinforcement schedule consisted of rewarding the desired behaviour with edible treats in the form of nuts, fruit and pasta. The negative reinforcers were three novel objects. This type of stimuli was chosen because novel objects are seen as aversive (i.e. something to avoid) when first presented. In this experiment, they consisted of a bucket, a chain and a glove. Each item was attached to the end of a long stick and remained outside the cage when presented, which consisted of slowly raising the object from floor level, moving it upwards and towards the cage. The appearance and approach of the novel objects were signalled using a tone, which differed for each pair trained. Four pairs ($n = 8$) were trained using positive reinforcement only, whereas six pairs ($n = 12$) were trained with combined positive and negative reinforcement. The approach of the novel objects and, over time, the sound of the tone, made the animals move towards the desired part of the cage. This behaviour was then reinforced with edible treats. Over the 30 training sessions, none of the monkeys trained only with positive reinforcement succeeded in accepting the gate being closed, whereas 10 out of 12 monkeys trained with the combined schedule managed at least once. It should be mentioned that the negative reinforcers were only necessary in less than a third of the training sessions. The authors propose that the negative reinforcement is perceived by the animal as being less aversive when it is used in combination with positive reinforcement. The use of stimuli that are only mildly aversive, as compared with electric shocks or loud noises, and the pre-signalling of the aversive stimulus allowing avoidance as opposed to an escape response, will have contributed to the success rate achieved in this study.

Just to muddy the waters a little, the clean-cut arrangement of positive and negative reinforcement presented in [Fig. 8.1](#) is not always apparent. In

an article by Michael (1975), later discussed by Baron and Galizio (2005), the question is raised of how we can distinguish between the two types of reinforcements. The examples given are rats pushing a lever to turn on a heat lamp when they feel cold: are the rats performing the behaviour (lever pressing) because they are given positive reinforcement (the pleasure of receiving warmth) or negative reinforcement (the removal of the unpleasantness of being cold)? A similar comparison could be made when food is used as reinforcement in food-deprived animals – is it the addition of something attractive (positive reinforcement) or the removal of the unpleasant feeling of hunger (negative reinforcement)? An interesting discussion of this conundrum, and the relative effectiveness of positive and negative reinforcement – and why we should never perform the ultimate test of this – can be found in Nevin and Mandell (2017).

Positive and Negative Punishment

Whereas reinforcement makes a behaviour more likely to occur, punishment leads to a reduction in the behaviour that is being punished. Although timing is always important in animal training, and also when using reinforcement, it is even more important when using a punishment to reduce a behaviour. Indeed, it is easy to achieve the opposite effect of what was intended: it is never good to whip your dog, especially if you do it when it returns to you, and your aim was to punish it for running away. As indicated in [Fig. 8.1](#), punishment also comes in two forms: one, where you apply some aversive treatment to the animal when it displays a certain behaviour, and another, where you remove or deny it an attractive treatment when a given behaviour is shown. The former is referred to a positive punishment, which is somewhat of a confusing misnomer (remember that positive means addition). An example of positive punishment is the shock received when touching an electric fence, which reduces the likelihood of touching, or even approaching, the fence again ([Fig. 8.3](#)). Negative punishment, where a desired stimulus is withdrawn or denied, can be achieved by ignoring an animal when it is performing an undesired behaviour, such as when a dog is begging for food at the table or barking for attention. By withholding the object of the behaviour, i.e. food and attention in these examples, the behaviour performed by the dog is not rewarded. Having said that, this is no easy task for the trainer, as timing is crucial: taking away a favourite toy may work for misbehaving children, but it is doubtful that a dog will associate the behaviour it performed with the withdrawal of, say, a chew toy – especially if the bad behaviour had nothing to do with the toy in the first place. If you put your dog outside whenever it pees on your carpet, the dog may quickly learn that if it wants to go outside, it just needs to pee on your carpet.



Fig. 8.3. An example of how positive punishment (i.e. adding an aversive stimulus) can reduce a behaviour: touching an electric fence makes you – and cows – less likely to touch it again (photo: Maria V. Rørvang).

Timing is indeed pivotal, especially for punishment, which needs to be immediate and consistently following the behaviour the trainer wants to reduce. Electric shock collars used in dog training are remotely controlled to deliver an aversive stimulus immediately when an undesirable behaviour is shown. However, the delivery of this positive punishment is still dependent on the skill of the trainer, and shock collars have been found to induce anxiety, stress, fear and pain in dogs (Steiss *et al.*, 2007; Cooper *et al.*, 2014). These collars have been found to be ineffective in many training situations (Makowska, 2018). Even when they do reduce an unwanted behaviour, such as nervous barking, the dogs remain anxious indicating that only the symptoms but not the cause of the problem have been removed (Moffat *et al.*, 2003). One study indicated that for some dogs the mere presence of their owner had become associated with getting shocked (Schilder and van der Borg, 2004). It is therefore not surprising that the use of shock collars is banned in at least nine European countries (Makowska, 2018), but the sale of them is often legal, and they are widely available online. The importance of timing makes punishment much more difficult to master than reinforcement, especially as punishment only indicates what the animal should *not* do, and not what it *should* do. Some dog trainers try to avoid any form of punishment by instead signalling a desired behaviour that cannot be performed together with the unwanted behaviour. Reinforcement

is much easier to apply and, if combined with clicker training (see below), can promote quite specific behaviours.

Another use of shock collars is to establish a virtual fence for livestock on pasture. Each of the animals to be contained within a certain boundary wears a device, often a collar. When the animal approaches the virtual fence, i.e. not a visible boundary but a specific geolocation, a warning signal is given, e.g. a sound or a vibration from the collar. If the animal tries to cross the invisible boundary, an electric shock is triggered from the collar. Most of these systems now work via GPS location of the animal, but the reliability and accuracy of these are variable. Unlike the old-fashioned electric fence, where the source of the shock is a physical feature, virtual fencing is an invisible line in the landscape. The idea of being able to contain animals, such as farm livestock, in a terrain where construction and maintenance of physical fences are impossible, is a good one. However, like the shock collars used for dog training, virtual fencing likely comes with a cost to the animal.

Bøe and Eftang (2019) found that small groups of goats ($n = 5\text{--}16$) monitored over the first 5 days after being introduced to the virtual fence, received between 0.2 and 5.5 shocks per animal per day on average, with some animals receiving 11, 13 and even 29 shocks in a day. In cows, values of 1.0 to 6.5 daily shocks per animal have been reported (Lomax *et al.*, 2019). Marini *et al.* (2018) found that sheep responded to the audio cue and avoided the electrical stimulus with a likelihood of only 52%. Campbell *et al.* (2019) observed large individual differences in cows in terms of ability to associate a sound with the proximity of a virtual fence with a 25% likelihood that a cow would receive an electric shock after an audio cue. In contrast, an investigation of privately owned cats kept either free-roaming or with a virtual fence found no long-term negative effects on the behaviour and welfare of the cats contained with the virtual fencing system (Kasbaoui *et al.*, 2016). Some systems have tried to replace the electric shock with an irritating and persistent sound. For a thorough review of the pros and cons of using virtual fencing with livestock, also in terms of animal welfare, please see Umstatter (2011), although more studies have been published since (e.g. Markus *et al.*, 2014; Brunberg *et al.*, 2017; Lee *et al.*, 2018; Kearton *et al.*, 2019). Unless the functionality and precision of virtual fencing can be improved, and given the negative and stressful effects found when training dogs with electric collars, the use of these devices to replace visible fencing does not appear to be conducive to good animal welfare.

An example using positive reinforcement and negative punishment

One of the most entertaining trials I have come across is an attempt to teach female calves to urinate in a particular place (Vaughan *et al.*, 2014). Not only is the subject area fascinating, but the authors have gone to great lengths

to ensure that the experimental design is statistically valid. The test arena consisted of a stall in which a test calf was placed. If the calf urinated within 15 min of entering the stall, it was given access to a milk feeder and received a milk reward (positive reinforcement). If the calf didn't pee, it was placed in a so-called time-out zone, where it was left for 5 min and did not receive any milk (negative punishment). Before you get to hear about the results, I would like to draw your attention to some of the details of the protocol. The calves had their milk ration gradually reduced to half before the trial started to ensure that the calves were motivated to obtain a milk reward. In addition, the calves had no access to milk for an hour before being trained or tested. Before testing began, the calves were habituated to the experimental set-up by being walked through the test arena 18 times over a period of 3 days. During this habituation, they either received a milk reward paired with the sound of a bell, or were left in the time-out zone for 5 min. This was done in a particular order (Gellermann, 1933) to ensure that the calves could not predict what was about to happen based on previous walk-throughs. Training or testing was then carried out for 17 days. Calves were allocated to pairs, where one calf was the test calf being trained to urinate, whereas the paired control calf was subjected to exactly the same conditions on the following day independent of its own urination pattern. A training day consisted of the test calf receiving an injection with a diuretic to induce urination. When the calf urinated, the bell was rung and the calf rewarded with milk. One injection sufficed for the calf to urinate several times, allowing the researchers to train the calf three times. The following day was a test day, where no diuretic was used. It was checked that calf had not urinated in the home pen within 30 min of the start of the test. Once in the test stall, if the calf urinated spontaneously, it was rewarded with milk, but if it didn't urinate within 15 min, it was moved to the time-out zone for 5 min as a form of punishment, and did not receive any milk. Calves that urinated spontaneously continued on the test schedule the following day, whereas calves that failed to pee were returned to the training schedule. The authors also checked that urination by the previous calf tested did not influence the likelihood of urination in the test arena by the subsequent calf.

So did it work? I'm happy to say that it did, although there was a certain variation in the ability of calves to make the association between their urinating and being rewarded with milk. Of the six pairs tested, one pair did not differ in their urination frequency in the stall, mainly because the test calf rarely peed when in there. In contrast, following only one training day with a diuretic, another test calf always urinated within 3 min of entering the stall (a video of this can be found with the online version of this reference). Overall, the test calves were more than twice as likely to urinate (on average 5.3 ± 0.95 urinations) than control calves (2.3 ± 0.52 urinations) across the test days. Vaughan *et al.* (2014) conclude that it may be feasible to train cattle to urinate in specific places using operant conditioning.

Clicker Training

Although reinforcement is better and easier to apply than punishment, at least two aspects of reinforcement can be difficult to manage. One is the use of rewards that eventually lose their reinforcing properties, which is often the case when food is used: the animal becomes less and less motivated to obtain the reward as it becomes satiated. Another problem encountered when using reinforcement as a training method is that the timing of the reward relative to the expression of the behaviour is important. The shorter the delay between the display of the desired behaviour and delivery of the reward, the easier it is for the animal to make the connection between the two, and the quicker the animal will learn. It can be difficult to achieve short delays if the trainer is not in the immediate vicinity of the animal, which is often the case with marine mammals and zoo species (e.g. Bloomsmith *et al.*, 2015).

Clicker training is a way to overcome these problems, making training less dependent on the environment in general, as well as being easy to perform. It consists of an acoustic secondary reinforcer, often the sound of a bell, whistle or click, used as a bridging stimulus between the desired behaviour and the reward. In other words, we want the animal to make the link between performing the desired behaviour, the sound of the click and the subsequent reward. We are temporarily replacing the reward with the click. This is achieved by first pairing the primary reward with the sound of the click through the process of classical conditioning (see Fig. 5.2 in Chapter 5): the animal is given a small food reward at the same time as it hears the click. Subsequently, the click occurs only when a correct behavioural response is performed, and over time the animal learns that this will eventually lead to a reward, i.e. reward delivery is removed in time from the behaviour/click, and not necessarily given at every click. Using a clicker – a small, handheld device with a button, emitting a short and audible sound – allows the trainer to deliver a click sound, which is identical every time and immediately contingent upon the desired behaviour. There is a plethora of videos available online showing how to clicker train your dog and your rat, as well as horses, fish and even tortoises.

Clicker training has been used to facilitate routine procedures, such as when elephants are trained through positive reinforcement to have their trunks washed voluntarily for tuberculosis testing (Fagen *et al.*, 2014). Joyce-Zuniga *et al.* (2016) trained captive grizzly bears to enter a holding crate and present their leg through the bars of the crate to have a blood sample taken. Initially, this was done using small amounts of diluted honey as the reward, delivered through a straw into the mouth of the bear, but this was paired with the sound of a click as a secondary reinforcer. The experimenters found serum cortisol to be undetectable in the trained bears, whereas bears that were chemically immobilized via a blow-dart injection prior to blood sampling showed elevated cortisol levels, especially if they

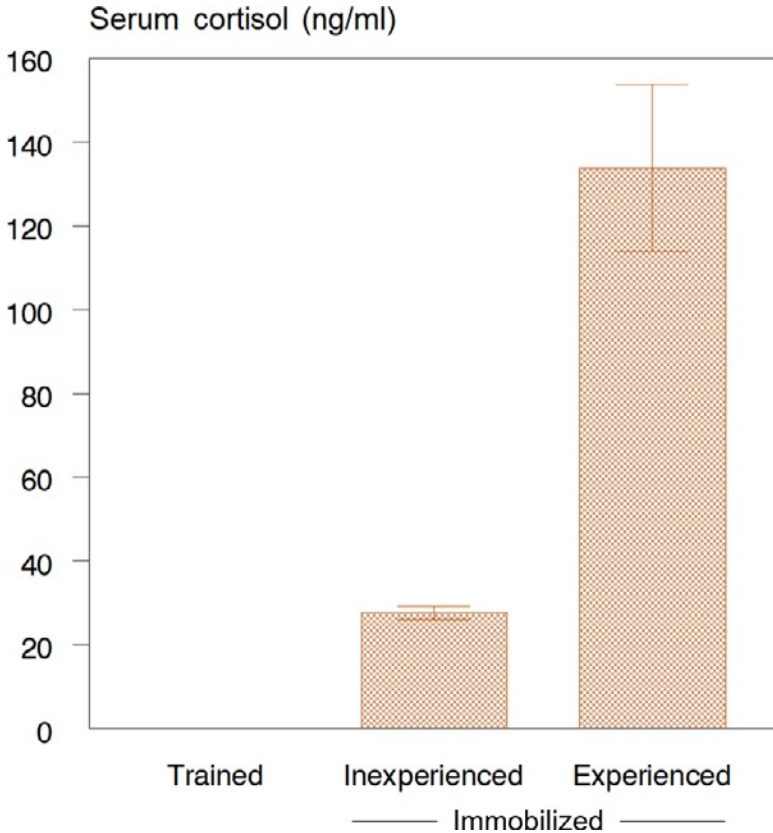


Fig. 8.4. Serum cortisol (ng/ml) in blood samples from grizzly bears, either trained ($n = 4$) to present their leg voluntarily, or immobilized using dissociative anaesthetics. The immobilized bears had either experienced ($n = 2$) or not ($n = 2$) a similar immobilization before (data from Joyce-Zuniga *et al.*, 2016).

had recently experienced similar immobilization (Fig. 8.4). Although the sample size is small, this demonstrates nicely the beneficial effects of positive reinforcement to reduce anxiety in animals during routine procedures.

Concluding Remarks

Reinforcement is an effective way to promote a variety of behaviours, and the use of positive reinforcement avoids the use of any form of aversive stimuli. Combining this with a secondary reinforcer, such as the use of a clicker, may not only accelerate the learning trajectory, but allow behaviours that are more subtle to be rewarded. As shown in the example with macaque monkeys, the addition of negative reinforcement can sometimes provide faster and more reliable results.

Punishment leads to a reduction in the behaviour penalized, but the use of aversive stimuli, such as electric shocks, should be avoided. The removal of an attractive stimulus (negative punishment) should always be preferred to positive punishment, and it should also be remembered that aversive does not imply experiencing pain. The confusion caused by the terms positive and negative in connection with both reinforcement and punishment can lead to unintentional misuse and misunderstandings. Thus, McLean and Christensen (2017) have wisely proposed the use of the words addition and subtraction instead of positive and negative.

Timing is pivotal in all types of learning, as is the linkage between the reinforcement or punishment and the behaviour. A story emerging from a university in the US tells the tale of a class of students deciding to use positive reinforcement to train one of their male lecturers: when the teacher positioned himself on the left side of the room, the students would pretend to be bored and uninterested. However, when the lecturer moved to the right-hand side of the classroom, they would look at him, smile and nod, as if taking in every word. Rumour has it that by the end of the semester, the teacher was performing all his lectures within touching distance of the right-hand wall. Whether true or not, this anecdote may serve as an amusing reminder of the difference between addition and subtraction reinforcement.

References

- Ahrendt, L.P., Labouriau, R., Malmkvist, J., Nicol, C.J. and Christensen, J.W. (2015) Development of a standard test to assess negative reinforcement learning in horses. *Applied Animal Behaviour Science* 169, 38–42. DOI: 10.1016/j.applanim.2015.05.005.
- Baron, A. and Galizio, M. (2005) Positive and negative reinforcement: should the distinction be preserved? *The Behavior Analyst* 28(2), 85–98. DOI: 10.1007/BF03392107.
- Bloomsmith, M., Neu, K., Franklin, A., Griffis, C. and McMillan, J. (2015) Positive reinforcement methods to train chimpanzees to cooperate with urine collection. *Journal of the American Association for Laboratory Animal Science* 54, 66–69.
- Bøe, K.E. and Eftang, S. (2019) Virtual fences for goats on commercial farms, an animal welfare concern. *Proceedings of the 53rd International Congress of the International Society of Applied Ethology*, Bergen, Norway, 5–9 August, p. 303.
- Brando, S.I.C.A. (2010) Advances in husbandry training in marine mammal care programs. *International Journal of Comparative Psychology* 23, 777–791.
- Brunberg, E.I., Bergslid, I.K., Bøe, K.E. and Sørheim, K.M. (2017) The ability of ewes with lambs to learn a virtual fencing system. *Animal* 11(11), 2045–2050. DOI: 10.1017/S1751731117000891.
- Campbell, D., Haynes, S., Lea, J., Farrer, W. and Lee, C. (2019) Temporary exclusion of cattle from a riparian zone using virtual fencing technology. *Animals* 9(1), 5. DOI: 10.3390/ani9010005.

- Christensen, J.W., McLean, A. and Ladewig, J. (2017) From Pavlov to Bekhterev – and the pros and cons of positive and negative reinforcement. *Proceedings of the 51st Congress of the International Society for Applied Ethology (ISAE)*, Aarhus, Denmark, 7–10 August, p. 158.
- Cooper, J.J., Cracknell, N., Hardiman, J., Wright, H. and Mills, D. (2014) The welfare consequences and efficacy of training pet dogs with remote electronic training collars in comparison to reward based training. *PLoS ONE* 9(9), e102722. DOI: 10.1371/journal.pone.0102722.
- Fagen, A., Acharya, N. and Kaufman, G.E. (2014) Positive reinforcement training for a trunk wash in Nepal's working elephants: demonstrating alternatives to traditional elephant training techniques. *Journal of Applied Animal Welfare Science* 17(2), 83–97. DOI: 10.1080/10888705.2014.856258.
- Fischer, B. and Wegener, D. (2018) Emphasizing the 'positive' in positive reinforcement: using nonbinary rewarding for training monkeys on cognitive tasks. *Journal of Neurophysiology* 120(1), 115–128. DOI: 10.1152/jn.00572.2017.
- Gellermann, L.W. (1933) Chance orders of alternating stimuli in visual discrimination experiments. *The Pedagogical Seminary and Journal of Genetic Psychology* 42(1), 206–208. DOI: 10.1080/08856559.1933.10534237.
- Innes, L. and McBride, S. (2008) Negative versus positive reinforcement: an evaluation of training strategies for rehabilitated horses. *Applied Animal Behaviour Science* 112(3–4), 357–368. DOI: 10.1016/j.applanim.2007.08.011.
- Joyce-Zuniga, N.M., Newberry, R.C., Robbins, C.T., Ware, J.V., Jansen, H.T. *et al.* (2016) Positive reinforcement training for blood collection in grizzly bears (*Ursus arctos horribilis*) results in undetectable elevations in serum cortisol levels: a preliminary investigation. *Journal of Applied Animal Welfare Science* 19(2), 210–215. DOI: 10.1080/10888705.2015.1126523.
- Kasbaoui, N., Cooper, J., Mills, D.S. and Burman, O. (2016) Effects of long-term exposure to an electronic containment system on the behaviour and welfare of domestic cats. *PLoS ONE* 11(9), e0162073. DOI: 10.1371/journal.pone.0162073.
- Kearton, T., Marini, D., Cowley, F., Belson, S. and Lee, C. (2019) The effect of virtual fencing stimuli on stress responses and behavior in sheep. *Animals* 9(1), 30. DOI: 10.3390/ani9010030.
- Lee, C., Colditz, I.G. and Campbell, D.L.M. (2018) A framework to assess the impact of new animal management technologies on welfare: a case study of virtual fencing. *Frontiers in Veterinary Science* 5, 187. DOI: 10.3389/fvets.2018.00187.
- Lomax, S., Colusso, P. and Clark, C.E.F. (2019) Does virtual fencing work for grazing dairy cattle? *Animals* 9(7), 429. DOI: 10.3390/ani9070429.
- Makowska, I.J. (2018) *Review of Dog Training Methods: Welfare, Learning Ability and Current Standards*. Published by the British Columbia Society for the Prevention of Cruelty to Animals (PC SPCA), Vancouver, Canada, p. 48.
- Marini, D., Meuleman, M.D., Belson, S., Rodenburg, T.B., Llewellyn, R. *et al.* (2018) Developing an ethically acceptable virtual fencing system for sheep. *Animals* 8(3), 33. DOI: 10.3390/ani8030033.
- Markus, S.B., Bailey, D.W. and Jensen, D. (2014) Comparison of electric fence and a simulated fenceless control system on cattle movements. *Livestock Science* 170, 203–209. DOI: 10.1016/j.livsci.2014.10.011.

- Marranzino, A. (2013) The use of positive reinforcement in training zebra sharks (*Stegostoma fasciatum*). *Journal of Applied Animal Welfare Science* 16(3), 239–253. DOI: 10.1080/10888705.2013.798555.
- McLean, A.N. and Christensen, J.W. (2017) The application of learning theory in horse training. *Applied Animal Behaviour Science* 190, 18–27. DOI: 10.1016/j.applanim.2017.02.020.
- Michael, J. (1975) Positive and negative reinforcement: a distinction that is no longer necessary; or a better way to talk about bad things. *Behaviorism* 3, 33–44.
- Moffat, K.S., Landsberg, G.M. and Beaudet, R. (2003) Effectiveness and comparison of citronella and scentless spray bark collars for the control of barking in a veterinary hospital setting. *Journal of the American Animal Hospital Association* 39(4), 343–348. DOI: 10.5326/0390343.
- Nevin, J.A. and Mandell, C. (2017) Comparing positive and negative reinforcement: a fantasy experiment. *Journal of the Experimental Analysis of Behavior* 107(1), 34–38. DOI: 10.1002/jeab.237.
- Premack, D. (1959) Toward empirical behavior laws: I. positive reinforcement. *Psychological Review* 66(4), 219–233. DOI: 10.1037/h0040891.
- Schilder, M.B.H. and van der Borg, J.A.M. (2004) Training dogs with help of the shock collar: short and long term behavioural effects. *Applied Animal Behaviour Science* 85(3–4), 319–334. DOI: 10.1016/j.applanim.2003.10.004.
- Steiss, J.E., Schaffer, C., Ahmad, H.A. and Voith, V.L. (2007) Evaluation of plasma cortisol levels and behavior in dogs wearing bark control collars. *Applied Animal Behaviour Science* 106(1–3), 96–106. DOI: 10.1016/j.applanim.2006.06.018.
- Umstatter, C. (2011) The evolution of virtual fences: a review. *Computers and Electronics in Agriculture* 75(1), 10–22. DOI: 10.1016/j.compag.2010.10.005.
- Vaughan, A., de Passillé, A.M., Stookey, J. and Rushen, J. (2014) Operant conditioning of urination by calves. *Applied Animal Behaviour Science* 158, 8–15. DOI: 10.1016/j.applanim.2014.06.009.
- Wergård, E.-M., Temrin, H., Forkman, B., Spångberg, M., Fredlund, H. *et al.* (2015) Training pair-housed Rhesus macaques (*Macaca mulatta*) using a combination of negative and positive reinforcement. *Behavioural Processes* 113, 51–59. DOI: 10.1016/j.beproc.2014.12.008.

Learning Capacity, Memory and Cognitive Ability

The ability of an animal to learn and remember, together with its capacity to demonstrate a certain cognitive level, are often associated with estimates of intelligence. Although primates are usually considered highly intelligent (Roth and Dicke, 2012), the concept is often debatable and has a tendency to serve merely as a ranking method of different species, with little functional merit. Does our training of dogs show that they are more intelligent than cats, or are cats just too clever to submit to being trained by humans? A recent review of the literature concluded (controversially) that dogs were not cognitively exceptional (Lea and Osthaus, 2018). So, although I will no doubt be criticized for putting tests of learning capacity and cognition into the same chapter, so be it. My argument will hopefully be evident in the subsequent sections: it is not always straightforward to determine what we are actually testing, because of overlaps and dependencies among these skills. Learning is to a large extent dependent on the ability to remember. The behavioural tests described below are but a handful of possible testing paradigms for use in assessing learning, memory and cognitive skills in a range of species.

Tests of Cognition

Object permanence

The notion that something continues to exist, even when you can no longer see it, is one of the earliest developmental skills used when assessing how human babies construct their reality (Piaget, 1954). Object permanence is learnt as we develop, and it forms the basis of the universal game of peek-a-boo, i.e. covering your face with your hands, then suddenly removing them while saying ‘Peek-a-boo!’, which can bring about a variety of emotions of expectation, surprise and laughter in a child: it is funny because the child knew you were there all along. Humans less than 5 months of age cannot grasp the concept of object permanence. In animals, tests have been

developed using this notion to assess the developmental skill of different species. These tests vary from simply covering an object to see if the animal will look for it in the right place, to more complex tests, such as moving and hiding an object, not only once but several times. Impossible scenarios may also be included, where the test animal is shown an object being hidden, but when the cover is removed the object is no longer there. An example of this, and of dogs mastering object permanence can be seen in the work by Finnish magician Jose Ahonen, who performs ‘taikuutta koirille’ – magic for dogs, by making a titbit in his hand disappear in front of a dog, with the dog clearly searching for the missing treat. The YouTube videos have been watched more than 18 million times at this point in time.

Caicoya *et al.* (2019) studied the object permanence ability of giraffes. According to the authors, this is a sparsely studied species with complex social relationships and a highly varied diet, both of which indicate a certain degree of memory capacity and cognitive ability. In this trial, giraffes ($n = 6$) were habituated to being shown a small, opaque box with the lid open and containing a food reward. The lid was subsequently closed, and the box held in front of the giraffe. When the animal pushed the box with its nose, the box was opened and the food reward given to the giraffe. Having tested that the giraffes were unable to locate the reward by olfactory cues alone, the animals were now tested for their ability to identify the one box out of two that contained a reward, both having been shown to the giraffe before the lids were closed (Fig. 9.1).

The giraffes consistently performed above chance level in choosing the baited box, demonstrating that they grasped the concept of an object remaining in place, even though it was no longer visible. When the experimenter prolonged the time between closing the lids and letting the giraffe choose, the animals could still make the right choice after a delay of 30 s, but not if a minute or more elapsed. This test was done to assess the short-term memory of the giraffes. Finally, instead of showing the content of the two boxes to the test animal, the experimenter shook one of the closed boxes before presenting both to the animal. When the baited box was shaken, and therefore made a noise, the giraffes were able to infer that it contained the food reward. However, if the empty box was shaken and thus made no noise, the giraffes were unable to make the inverse (exclusion) inference, i.e. if there is no reward in there, then the other box must be the right one. Interestingly, this rather complex ability of using indirect information has been mastered by goats, a species also able to keep track of invisible displacements (Nawroth *et al.*, 2014, 2015), which is a skill more often reported in great apes and humans (Barth and Call, 2006). Dogs are able to find a rubber ball shown to be hidden in one of four boxes, even when the view of the boxes is subsequently blocked by an opaque screen for up to 4 min (which was the maximum delay tested; Fiset *et al.*, 2002). Although still succeeding above chance level, the dogs did make more errors when the start of the search was delayed (Fig. 9.2). The mistakes were



Fig. 9.1. Test of ability to understand object permanence in giraffes. The animal is presented with open opaque boxes, one of which is baited with a food reward. Upon closing the lids, the boxes are moved towards the test animal, and the chosen box is the one the giraffe touches first with its lips or tongue (from Caicoya *et al.*, 2019).

mainly due to the dogs choosing boxes adjacent to the one containing the ball, indicating difficulties in remembering the exact spatial position of the baited box.

Detour test

A blocked view is also used in the so-called detour test (see Kabadayi *et al.*, 2018 for a recent review). An animal is shown a path to a given goal, and then an obstacle is placed on the route so that the animal will have to move around it to arrive at the goal. Sovrano *et al.* (2018) tested different species of fish in a detour test (Fig. 9.3). The animal being tested was temporarily confined within a small area of the tank, where it could see a group of conspecifics (the goal) through a small window. Upon being released from the smaller enclosure, the fish would have to turn its back to the goal and

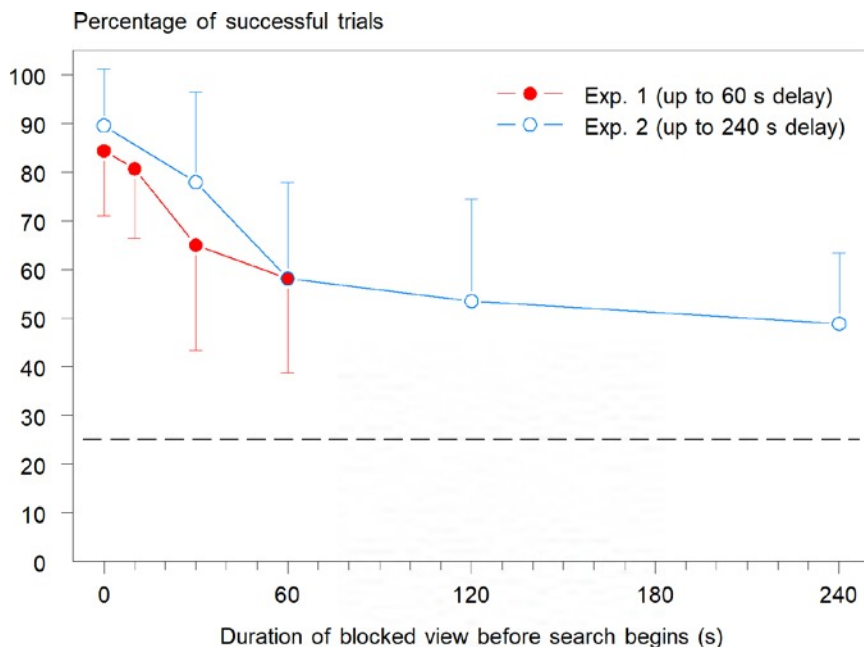


Fig. 9.2. Successfulness of dogs in locating a ball hidden in one of four boxes as a function of the duration of time since the hiding place was shown to the dogs (Exp. 1: $n = 11$ dogs; Exp. 2: $n = 8$ dogs). The percentage of successful trials decreases as the start of the search is delayed, but the dogs still find the ball above chance level (dashed line) on average (adapted from Fiset *et al.*, 2002).

swim in the opposite direction, before making a choice of turning into one of four boxes, of which two were in the correct direction towards the goal. The water in the two tanks was not connected, and the four boxes to choose from were mirror versions of the same shape and size. Of the four species tested, goldfish, angelfish and redtail splitfins were able to choose one of the two correct compartments in terms of moving towards the goal (compartments A and B in Fig. 9.3) in a proportion significantly different from chance (in total 80 out of 108 fish tested; 74%). The only species that did not convincingly succeed on the first choice was zebrafish (24 out of 41; 59%).

Cylinder test

Another, almost inverse version of the detour test described above is the cylinder test. I particularly liked the finding that certain bird species equal great apes in this particular cognitive skill (Kabadayi *et al.*, 2016). The test subjects, in this particular example three different crow species, were habituated to an opaque, hollow cylinder with open ends from which they

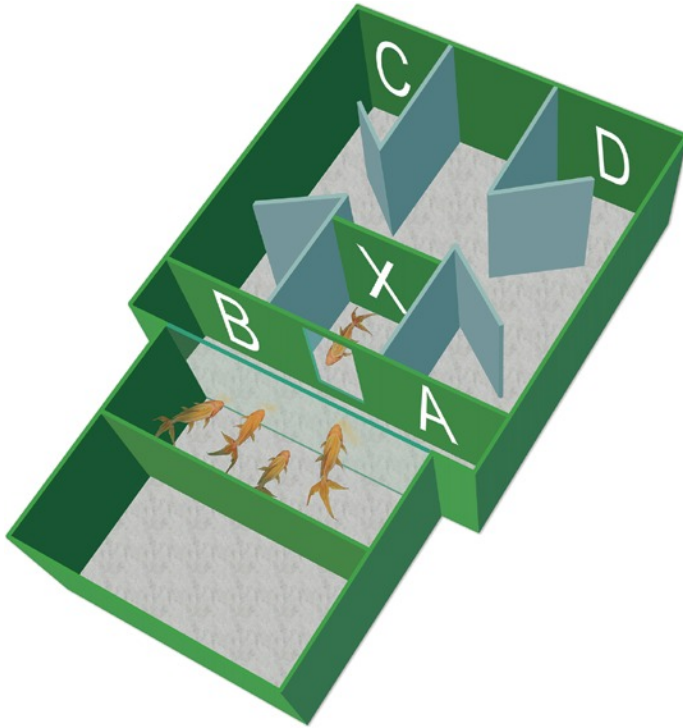


Fig. 9.3. Example of a set-up for a detour test in fish. Through a small window, the fish being tested is able to see a group of conspecifics acting as a social attractant (goal) in an adjacent tank. When panel X is removed, in order to try and join the social group, the test fish has to turn its back to the goal and choose among four different compartments (A–D) of which two are in the right direction (A and B). Out of four fish species, three were able to do so for 74% of the individuals tested, which is significantly different from chance (from Sovrano *et al.*, 2018).

learnt to retrieve a food reward placed inside at the centre. For the test itself, the opaque cylinder was replaced with a transparent one. If the bird tried to access the food reward, which was now visible, by attempting to reach it through the side of the transparent cylinder, the test was scored as a fail. If, on the other hand, the bird approached one of the ends of the cylinder, as it had done with the opaque version, in order to retrieve the reward, the test was scored as a success. Kabadayi *et al.* (2016) found that ravens succeeded in all trials (100%) and jackdaws averaged 97% success. This is comparable with the success rates found in great apes (MacLean *et al.*, 2014), including chimpanzees (100%), bonobos (95%) and gorillas (94%).

The impressive cognitive abilities of corvids are well known. Not only are they capable of shaping and using tools, they also appear to be able

to plan their actions based on the available means, thus form mental representations of solutions to problems. However, when one problem has one direct solution (e.g. the food dangling on the string from my perch is difficult to reach, so I use my beak and a foot to gradually pull it up), the behavioural display can be assigned to other factors than reasoning, such as perceptual motor feedback (Taylor *et al.*, 2010). Bugnyar (2019) gives an excellent summary of how we can test for these very complex cognitive skills. In a series of trials, New Caledonian crows were given different tasks to gain access to a food reward. These included retrieving a stick from one apparatus to be used to release a stone from a tube, and the stone was then used to release food from another structure while the crow at the same time ignored a distractor stimulus containing another stone. The details of the experimental set-up are too intricate to be included here, but the complexity of some of the tasks, and the finding that some crows are able to solve them, is mind-blowing. It makes you wonder if you would find the solution as fast as the crow – or even at all?

Tests of Memory

The hole-board discrimination test described in Chapter 7 can be used to assess different aspects of spatial memory (Tahamtani *et al.*, 2015). Number of holes used may differ between tests, and the holes are positioned equidistantly in a square of 3×3 , 4×4 or 5×5 holes, of which the square root, i.e. 3, 4 and 5 holes, respectively, are baited with a small, highly palatable food reward. The time taken to find and eat all the baits (trial duration), as well as the number of times a baited hole is revisited reflects negatively on the *working memory* of the animal tested, as does revisits to holes in general. The number of visits necessary in total before all baits have been found is a reflection of the *reference memory* of the test animal, and the extent to which it is able to discriminate between holes.

Delayed match-to-sample

One of the most widely used tests for working or short-term memory is the delayed match-to-sample test (Nielsen *et al.*, 2009; Bimonte-Nelson, 2015a). Just like the cognitive bias test described in Chapter 4, this test can be carried out using spatial, visual, auditory or olfactory cues, and the design is relatively simple: following habituation to cues and the test environment, the working memory of the animal is tested by rewarding one of several cues. Lee *et al.* (2018) trained rats to retrieve small food rewards hidden in cups containing scented sand. Four different odours were used, and in the first part of the test (learning the match-to-sample principle), the rat was presented with a sample cup of scented sand containing half of a small food reward. Once the reward has been retrieved, the rat is presented with

four cups containing sand with different smells, one of which matches the odour of the sample cup. The rat may sniff all the cups, but only dig in one, and only the cup matching the sample odour contains a whole food reward (the other cups have crumpled, and thus inaccessible food below the surface to make all cups have the added smell of food). When the rat starts to learn the connection between the odours of the sample cup and the rewarded cup, a delay is introduced before being given access to choose one of the four test cups. By increasing or decreasing the duration of this delay, the ability of the animal to remember can be quantified. Sometimes the test is performed as the inverse, i.e. a non-match-to-sample test (Callaghan *et al.*, 2012). An added twist to the delayed match-to-sample test is when the animal is given the option of not choosing (Lee *et al.*, 2018). This can be done by adding a neutral choice that gives rise to a smaller reward than a correct choice, but does not lead to no reward like a wrong choice would. Such an experimental set-up allows the test animal to show that it can assess the strength of its memory. The cognitive ability to know whether or not you are able to remember something is referred to as metamemory, which has been demonstrated in a number of species, including rats (Templer *et al.*, 2017).

I warned you in Chapter 1 that an example of behavioural tests using honey bees may be mentioned. You can think of them as being included in a very broad definition of animals managed by humans. We have no doubt all been fascinated by the ability of honey bees to communicate through their waggle dance (von Frisch, 1927), and these insects do not disappoint when it comes to learning. Although not the first, nor the most recent study on this subject, the article by Gross *et al.* (2009) is a nice and relatively simple example of the very complex tasks honey bees are able to perform. At the entrance to a Y-maze, bees were presented with a pattern before flying through a 1-m-long tunnel leading to a chamber with a choice of two patterns (Fig. 9.4a). One of these matches the pattern at the entrance to the tunnel, thus indicating the hole leading to a sugar reward. Bees were first trained on relatively simple patterns, with two or three identical blue dots presented in the same configuration in the sample pattern and the choice pattern. Subsequently, the bees were tested on different designs, but always so that the quantity of symbols on the sample pattern indicated the quantity of symbols on the pattern leading to the rewarded choice.

The bees were first tested with the configuration of dots no longer matching, so that the pattern but not the number of dots differed between the sample and rewarded choice pattern. The bees were able to choose correctly in 79% of cases with two dots and 70% of cases when three dots were the sample pattern. In a subsequent test, which the bees also mastered, the blue dots were replaced with yellow stars, thus changing the colour and shape of the symbols used to make the pattern. Even when the rewarded choice pattern differed from the sample pattern in the type and mixture of symbols used, as well as when the two choice patterns covered

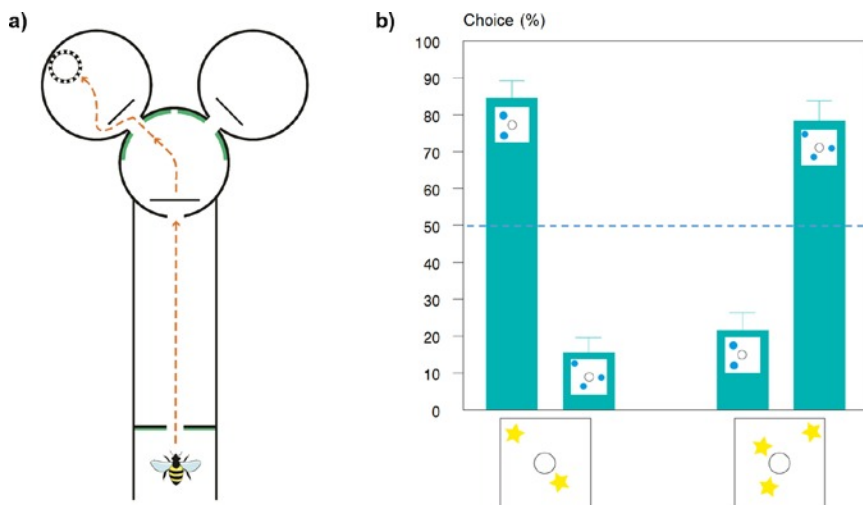


Fig. 9.4. Example of a delayed match-to-sample test performed with honey bees. (a) The bee enters the maze through the centre hole of a marker with a sample pattern, and flies through a 1-m tunnel into another chamber, where the bee has a choice of two marker patterns, each surrounding the entry to a chamber, where only the chamber with the same number of symbols as on the sample pattern contains a sugar reward; positions of markers are indicated in green. (b) Results of two tests with different sample patterns with stars shown below the bars, and the choice patterns indicated on the bars of the histogram. Bees are able to choose the correct chamber entry, even when the shape, colour, configuration and size of the symbols differ between the sample and the choice patterns. The dashed line indicates the threshold for a random choice (adapted from Gross *et al.*, 2009).

the same total area (Fig. 9.4b) or had the same total edge length, the bees were able to choose correctly in more than 70% of cases. Keep in mind that bees had to remember the sample pattern encountered at the tunnel entrance until they arrived at the choice patterns. Only when there were four or more symbols in the pattern did the bees not consistently make the right choice. Gross *et al.* (2009) clearly state that they have not shown that bees can count or put numbers in order of magnitude, but simply that this insect has an ability to assess numerosity and can learn to distinguish between small quantities. Bees have also recently been found able to match characters to small quantities and vice versa (Howard *et al.*, 2019), in the same way we associate the symbol 3 with the quantity of three. However, the bees were not able to reverse the association spontaneously upon first presentation, which the authors suggest is due to the numerical nature of the task.

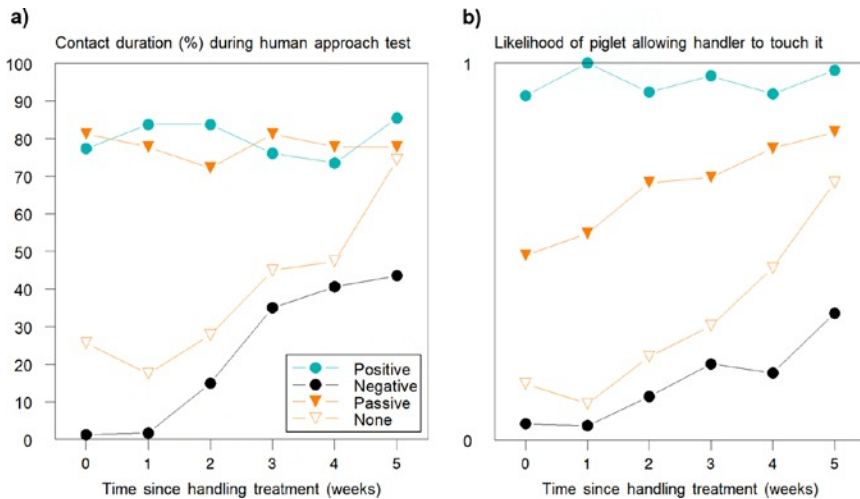


Fig. 9.5. (a) Mean percentage of time piglets (three piglets per group; six groups per treatment) spent in contact with the human handler, who was sitting motionless on a chair in the corner of the home pen. The test was repeated immediately after the end of the handling treatment (0 weeks) and weekly thereafter. The handling treatments (18 sessions over 5 days) were positive (gentle stroking), negative (rough handling), passive (human presence with no handling) and none (no human presence). (b) Likelihood of piglets allowing the human handler to touch them (adapted from Brajon *et al.*, 2015).

Human approach test revisited

The human approach test described in Chapter 4 can also be used to test the memory of previous experience with humans. Brajon *et al.* (2015) applied one of eight different treatments to groups of three piglets, with group composition balanced for live weight of the piglets and their response in a novel object test (see Chapter 4). Treatments consisted of different types of handling, four of which are presented here: positive (gentle stroking), negative (rough handling) and passive (no handling, but human present), as well as a control treatment where no human was present. The piglets were subjected to 18 sessions of these 5-min treatments across 5 consecutive days. In order to investigate to what extent the piglets were able to remember and associate the human handler with the treatment, a passive human approach test was carried out immediately after the end of the treatments and weekly for the next 5 weeks. Figure 9.5a shows the percentage of time the piglets spent in contact with the human at each test for each of the four treatments. Although there is a clear and persistent difference between pigs on the positive and negative handling treatments, it is also clear that the mere habituation to the presence of a passive human is sufficient for the pigs to approach and interact with the person. However, when the handler tries to touch the piglet (Fig. 9.5b), the positive

treatment differs significantly from the passive treatment for the first two tests, demonstrating that these piglets had learnt to associate the handler with a positive experience, and that this association persisted, even though the non-handled piglets caught up over time.

Passive and active avoidance

One way to assess memory in laboratory rodents is the passive avoidance test. It is based on the innate preference for dark spaces by nocturnal rats and mice. The test box consists of two connected compartments, with one being brightly lit and the other being dark. Once the animal is habituated to the test box, it is exposed to one trial only where it receives a mild electric foot shock when it retreats into the dark half of the test box. Following this one acquisition trial, the animal is tested again with no shock given. However, if the animal remembers the previous experience, in order to avoid the aversive shock stimulus, the animal has to remain in the (also aversive) brightly lit compartment, i.e. not respond as it would otherwise do. In other words, to avoid the shock requires the animal not to move, hence display a passive avoidance response. Latency to escape from the lit compartment is a reflection of memory: the longer the rat stays in the lit compartment, the better it recalls the aversive shock treatment. The test can be used to evaluate the effects of different drugs on learning and memory, but many years ago, Ader *et al.* (1972) wanted to see how the latency and hence memory of rats was affected by the duration and intensity of this one shock. The researchers tested this at different intervals after the initial shock and, although no shocks were given during these subsequent tests, the sound of the electric floor grid was audible to the rats. [Figure 9.6](#) shows the result for the mildest shock intensity used (0.125 mA): it is clear that the longer the shock was applied during the single acquisition trial, the more reluctant the rat was to enter the dark compartment. The memory of the aversive shock faded over time, with faster entries after a day or two. When more intense currents (0.25 and 0.5 mA) were used, the latencies increased sharply (data not shown).

Active avoidance tests use a similar set-up (a two-compartment shuttle box) as previously described for passive avoidance, but the electric shock is now signalled with an innocuous stimulus, such as a light, which the animal is conditioned to associate with the onset of the current. The animal can then move to the other compartment to avoid the shock, hence active avoidance. If the animal does not move, it will receive a foot shock, but can escape by moving to the other compartment. The test can thus distinguish between a conditioned response (avoidance) and an unconditioned response (escape). It should always be taken into account that these tests often (but not always) use electric shocks as the aversive treatment, and therefore should not be used without very good reason. Indeed, other less aversive stimuli have been tried, such as air puffs (Moriarty *et al.*, 2012).

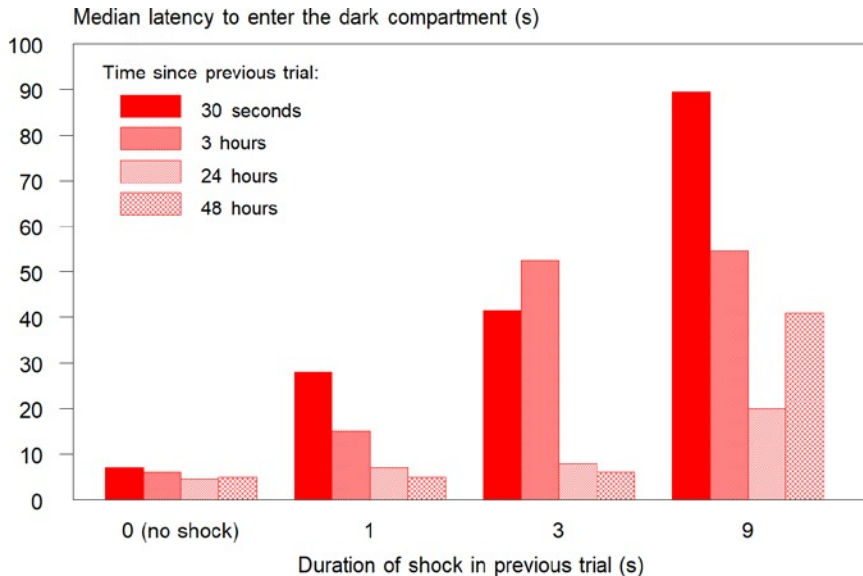


Fig. 9.6. Median latency (s) to enter a dark compartment previously associated with a mild foot shock (0.125 mA) as a function of the duration of the shock and the time since this was experienced. Each bar is the median of 10 rats (except no shock: $n = 6$) (data are from Ader *et al.*, 1972).

Tests of Learning

Go/No-go test

Many of the tests described in this book involve a degree of learning. The nose-poke paradigm described in Chapter 6 depends on the rodent learning to associate each of two odours with a certain response (go left or right). However, it can also be used in a slightly more complex manner, where the animal has to respond to one odour and not respond when another odour is presented. This is known as a Go/No-go task (e.g. Frederick *et al.*, 2011), where only one reward hole is used, and if the animal responds to the wrong (i.e. the No-go) odour, a mild punishment is introduced before the animal can try again. This is done by delaying the onset of the light or sound indicating the imminent odour release, so that the animal has to wait a short while for the next opportunity to be rewarded. This is the same type of negative punishment used in the example of training calves to urinate described in Chapter 8.

It is usually much more difficult to learn *not* to do something than it is to learn to do something, and the Go/No-go protocol can be used to assess impulsiveness. The test can of course be carried out using experimental setups different from the nose-poke apparatus and not necessarily investigating olfactory stimuli. The test is, for example, used on human subjects to assess

response speed and accuracy, but also response inhibition. I thought I should give it a try, and performed an online version of a Go/No-go test, where I was asked to press the space bar within 2s if a green Go was shown on my screen, but not to press the bar and wait for 2s if a red No-go was shown. Simple, eh? Well, I did not do too badly (Fig. 9.7) but I did make a mistake, even though I was hell-bent on demonstrating (not least to myself) that this was an easy task once you knew the concept and were prepared and ready. Before I did the test, I was not planning to show these data here, but perhaps we should all be better at trying different testing paradigms to fully appreciate what we are asking our animal subjects to do?

Daros *et al.* (2014) trained dairy calves using a Go/No-go set-up with monitors displaying either a red or a white screen. The calves started from one corner of a rectangular arena with the monitor placed farthest away from the start position (Fig. 9.8). When the screen was white (positive cue), the calf was rewarded with milk if it approached the screen to within 20 cm. When the screen was red (negative cue), the calf should refrain from approaching. If a calf approached the red screen, a whistle was blown, and the calf was punished by delaying the onset of the next screen display, and hence the opportunity to obtain a milk reward, for 1 min. The calves were first trained with only the positive cue (Go response) and, upon reaching a 90% correct response rate, the negative cue was gradually introduced (from 2 up to 20), interspersed among the 20 positive cues. For the negative cue, a 100% learning criterion was applied, whereas the minimum of

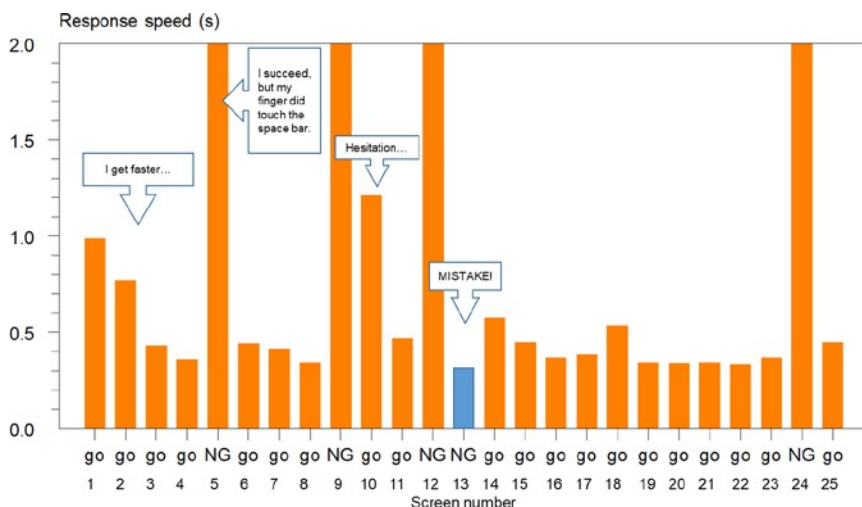


Fig. 9.7. Data from my trying an online Go/No-go test, where I had to press the space bar within 2s when a green 'go' sign appeared on my computer screen, but refrain from responding for 2s if a red 'No-go' (NG) appeared. For each test, 25 screens appear with on average one in five screens displaying the No-go sign. The test can be found at www.psytoolkit.org.

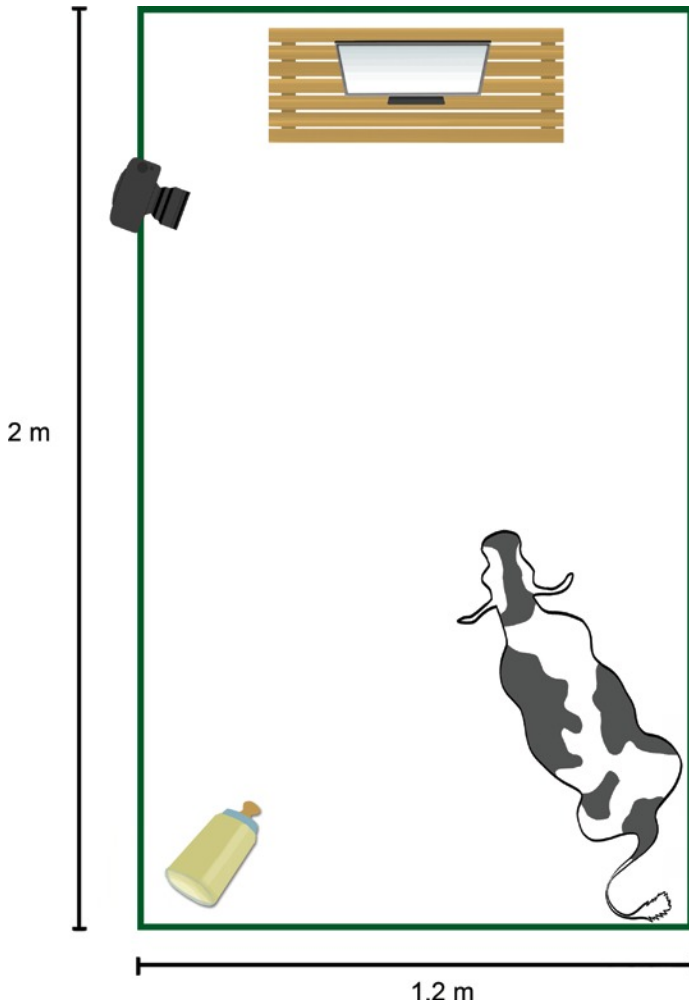


Fig. 9.8. Arena used in a Go/No-go test of dairy calves. A monitor at the far end of the arena displays either a white or a red screen. If the monitor display is white, the calf is trained to approach the monitor to within 20 cm, a response that is rewarded by access to the milk bottle in the opposite corner to the start position. If the monitor displays a red screen, the calf should not approach. However, if the calf approaches the screen when red, no milk is given and the calf is punished by delaying the display of the next screen by 1 min (adapted from Neave *et al.*, 2013 and Daros *et al.*, 2014).

90% correct responses was kept for the positive cue. All of the 13 calves tested were able to learn the Go/No-go task with 20 screens of each colour and at least 90 and 100% correct responses for the Go and No-go parts, respectively, using on average 32 sessions across all training.

Learning from conspecifics

Behavioural tests using demonstrator animals can be used to investigate if animals are able to learn from each other by observation alone. In horses, it has been found that fear responses are reduced if the horse is paired with an experienced calm companion (Christensen *et al.*, 2008). However, Rørvang *et al.* (2018) has questioned whether horses are able to learn from social observation, and proposed that the transfer of behavioural responses in social situations is not learning, but a social transmission caused by other mechanisms, such as social facilitation and stimulus enhancement. Social learning tests have most often been used in mammals and birds, but I would like to include an example from a more exotic category of companion animals, the bearded dragon, which – somewhat disappointingly – is but a large lizard. Kis *et al.* (2015) wanted to know if bearded dragons could learn to open a sliding door by watching a video of a conspecific trained to open said door. The animals had been habituated to eating meal worms from a Petri dish and, by placing a dish of meal worms behind a wire door that could slide open to the right and to the left, the researchers created an incentive for the bearded dragons to open the door. The protocol used the following clever trick to ensure that the learning was indeed a result of observing a conspecific: One group of bearded dragons ($n = 4$) watched a video of the demonstrator lizard opening the door by sliding it to the right with rapid head movements. Another group ($n = 4$) watched the same video in a flipped version, so that the door was now opened to the left, whereas the control group ($n = 4$) watched the door opening by itself to the right while a conspecific was standing passively in front of it. [Figure 9.9](#) shows the results: only the bearded dragons that had seen the videos with the door being opened by a conspecific were able to open the door and, in their first successful trial, they all opened it to the same side they had seen the demonstrator lizard do. Also, none of the control lizards made sideways head movement, whereas all the other bearded dragons did perform this movement, emphasizing that it had been learnt from observing the video. This is a neat demonstration of social learning in a species not usually associated with this type of apprenticeship.

Concluding Remarks

The variety of tests developed for assessing cognition, learning and memory are plentiful, and only a handful have been presented here. I hope, however, that the demonstration of these different abilities and tests using examples from mammals, reptiles and even insects, has provided inspiration to investigate these types of behavioural tests in detail. More information on some of the tests described above can be found in Bimonte-Nelson (2015b). For a quick overview with examples of tests on cognition and learning, see also Nawroth and Langbein (2019). If you are looking for

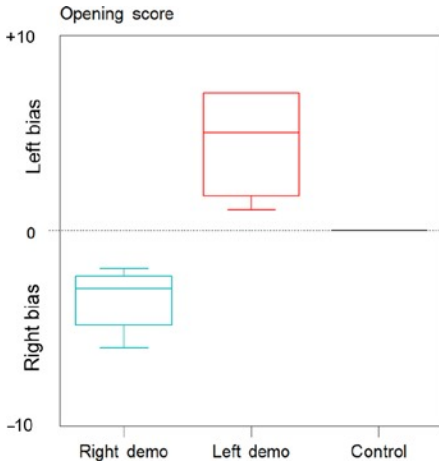
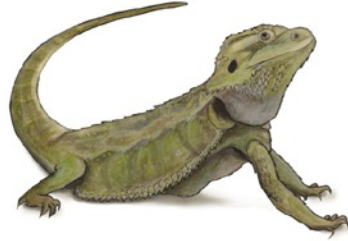


Fig. 9.9. Propensity for bearded dragons to open a sliding door to the left (score +1), to the right (-1) or not at all (0) out of 10 trials, as a function of having seen a conspecific opening the door to the right (Right demo), or to the left (Left demo) or the door opening to the right by itself while a passive bearded dragon is standing by it (Control) (adapted from Kis *et al.*, 2015).

species-specific information, I can recommend Marino (2017) for a review covering cognition in the domestic chicken. Finally, a comprehensive introduction to the subject of *Animal Learning and Cognition* can be found in the book of that name by John M. Pearce (2008).

References

- Ader, R., Weijnen, J.A.W.M. and Moleman, P. (1972) Retention of a passive avoidance response as a function of the intensity and duration of electric shock. *Psychonomic Science* 26(3), 125–128. DOI: 10.3758/BF03335453.
- Barth, J. and Call, J. (2006) Tracking the displacement of objects: a series of tasks with great apes (*Pan troglodytes*, *Pan paniscus*, *Gorilla gorilla*, and *Pongo pygmaeus*) and young children (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes* 32(3), 239–252. DOI: 10.1037/0097-7403.32.3.239.
- Bimonte-Nelson, H.A. (2015a) The delayed match to sample protocol for rodents. In: Bimonte-Nelson, H.A. (ed.) *Maze Book: Theories, Practice, and Protocols for Testing Rodent Cognition*. Book Series: Neuromethods, vol. 94. Humana Press Inc., Totowa, New Jersey, pp. 433–440.
- Bimonte-Nelson, H.A. (ed.) (2015b) *Maze Book: Theories, Practice, and Protocols for Testing Rodent Cognition*. Book Series: Neuromethods, vol. 94. Humana Press Inc., Totowa, New Jersey.
- Brajon, S., Laforest, J.-P., Bergeron, R., Tallet, C., Hötzel, M.-J. *et al.* (2015) Persistency of the piglet's reactivity to the handler following a previous positive or negative experience. *Applied Animal Behaviour Science* 162, 9–19. DOI: 10.1016/j.applanim.2014.11.009.

- Bugnyar, T. (2019) Tool use: New Caledonian crows engage in mental planning. *Current Biology* 29(6), R200–R202. DOI: 10.1016/j.cub.2019.01.059.
- Caicoya, A.L., Amici, F., Ensenyat, C. and Colell, M. (2019) Object permanence in *Giraffa camelopardalis*: first steps in giraffes' physical cognition. *Journal of Comparative Psychology* 133(2), 207–214. DOI: 10.1037/com0000142.
- Callaghan, C.K., Hok, V., Della-Chiesa, A., Virley, D.J., Upton, N. *et al.* (2012) Age-related declines in delayed non-match-to-sample performance (DNMS) are reversed by the novel 5HT6 receptor antagonist SB742457. *Neuropharmacology* 63(5), 890–897. DOI: 10.1016/j.neuropharm.2012.06.034.
- Christensen, J.W., Malmkvist, J., Nielsen, B.L. and Keeling, L.J. (2008) Effects of a calm companion on fear reactions in naive test horses. *Equine Veterinary Journal* 40(1), 46–50. DOI: 10.2746/042516408X245171.
- Daros, R.R., Costa, J.H.C., von Keyserlingk, M.A.G., Hötzel, M.J. and Weary, D.M. (2014) Separation from the dam causes negative judgement bias in dairy calves. *PLoS ONE* 9(5), e98429. DOI: 10.1371/journal.pone.0098429.
- Fiset, S., Beaulieu, C. and Landry, F. (2002) Duration of dogs' (*Canis familiaris*) working memory in search for disappearing objects. *Animal Cognition* 6(1), 1–10. DOI: 10.1007/s10071-002-0157-4.
- Frederick, D.E., Rojas-Libano, D., Scott, M. and Kay, L.M. (2011) Rat behavior in go/no-go and two-alternative choice odor discrimination: differences and similarities. *Behavioral Neuroscience* 125(4), 588–603. DOI: 10.1037/a0024371.
- Gross, H.J., Pahl, M., Si, A., Zhu, H., Tautz, J. *et al.* (2009) Number-based visual generalisation in the honeybee. *PLoS ONE* 4(1), e4263. DOI: 10.1371/journal.pone.0004263.
- Howard, S.R., Avarguès-Weber, A., Garcia, J.E., Greentree, A.D. and Dyer, A.G. (2019) Symbolic representation of numerosity by honeybees (*Apis mellifera*): matching characters to small quantities. *Proceedings of the Royal Society B: Biological Sciences* 286(1904), 20190238. DOI: 10.1098/rspb.2019.0238.
- Kabadayi, C., Taylor, L.A., von Bayern, A.M.P. and Osvath, M. (2016) Ravens, New Caledonian crows and jackdaws parallel great apes in motor self-regulation despite smaller brains. *Royal Society Open Science* 3(4), 160104. DOI: 10.1098/rsos.160104.
- Kabadayi, C., Bobrowicz, K. and Osvath, M. (2018) The detour paradigm in animal cognition. *Animal Cognition* 21(1), 21–35. DOI: 10.1007/s10071-017-1152-0.
- Kis, A., Huber, L. and Wilkinson, A. (2015) Social learning by imitation in a reptile (*Pogona vitticeps*). *Animal Cognition* 18(1), 325–331. DOI: 10.1007/s10071-014-0803-7.
- Lea, S.E.G. and Osthau, B. (2018) In what sense are dogs special? Canine cognition in comparative context. *Learning & Behavior* 46(4), 335–363. DOI: 10.3758/s13420-018-0349-7.
- Lee, K.A., Preston, A.J., Wise, T.B. and Templer, V.L. (2018) Testing for metacognitive responding using an odor-based delayed match-to-sample test in rats. *Journal of Visualized Experiments* (136), e57489. DOI: 10.3791/57489.
- MacLean, E.L., Hare, B., Nunn, C.L., Addessi, E., Amici, F. *et al.* (2014) The evolution of self-control. *Proceedings of the National Academy of Sciences* 111(20), E2140–E2148. DOI: 10.1073/pnas.1323533111.
- Marino, L. (2017) Thinking chickens: a review of cognition, emotion, and behavior in the domestic chicken. *Animal Cognition* 20(2), 127–147. DOI: 10.1007/s10071-016-1064-4.

- Moriarty, O., Roche, M., McGuire, B.E. and Finn, D.P. (2012) Validation of an air-puff passive-avoidance paradigm for assessment of aversive learning and memory in rat models of chronic pain. *Journal of Neuroscience Methods* 204(1), 1–8. DOI: 10.1016/j.jneumeth.2011.10.024.
- Nawroth, C. and Langbein, J. (2019) Editorial: advances and perspectives in farm animal learning and cognition. *Frontiers in Veterinary Science* 6, 172. DOI: 10.3389/fvets.2019.00172.
- Nawroth, C., von Borell, E. and Langbein, J. (2014) Exclusion performance in dwarf goats (*Capra aegagrus hircus*) and sheep (*Ovis orientalis aries*). *PLoS ONE* 9(4), e93534. DOI: 10.1371/journal.pone.0093534.
- Nawroth, C., von Borell, E. and Langbein, J. (2015) Object permanence in the dwarf goat (*Capra aegagrus hircus*): perseveration errors and the tracking of complex movements of hidden objects. *Applied Animal Behaviour Science* 167, 20–26. DOI: 10.1016/j.applanim.2015.03.010.
- Neave, H.W., Daros, R.R., Costa, J.H.C., von Keyserlingk, M.A.G. and Weary, D.M. (2013) Pain and pessimism: dairy calves exhibit negative judgement bias following hot-iron disbudding. *PLoS ONE* 8(12), e80556. DOI: 10.1371/journal.pone.0080556.
- Nielsen, T.R., Kornum, B.R., Moustgaard, A., Gade, A., Lind, N.M. *et al.* (2009) A novel spatial delayed non-match to sample (DNMS) task in the Göttingen minipig. *Behavioural Brain Research* 196(1), 93–98. DOI: 10.1016/j.bbr.2008.07.019.
- Pearce, J.M. (2008) *Animal Learning and Cognition – An Introduction*, 3rd edn. Psychology Press, London.
- Piaget, J. (1954) *The Construction of Reality in the Child*, Translated by Margaret Cook from *La construction du réel chez l'enfant* (1950). Routledge and Kegan Paul, London.
- Roth, G. and Dicke, U. (2012) Evolution of the brain and intelligence in primates. *Progress in Brain Research* 195, 413–430.
- Rørvang, M.V., Christensen, J.W., Ladewig, J. and McLean, A. (2018) Social learning in horses – fact or fiction? *Frontiers in Veterinary Science* 5, 212. DOI: 10.3389/fvets.2018.00212.
- Sovrano, V.A., Baratti, G. and Potrich, D. (2018) A detour task in four species of fishes. *Frontiers in Psychology* 9, 2341. DOI: 10.3389/fpsyg.2018.02341.
- Tahamtani, F.M., Nordgreen, J., Nordquist, R.E. and Janczak, A.M. (2015) Early life in a barren environment adversely affects spatial cognition in laying hens (*Gallus gallus domesticus*). *Frontiers in Veterinary Science* 2, 3. DOI: 10.3389/fvets.2015.00003.
- Taylor, A.H., Medina, F.S., Holzhaider, J.C., Hearne, L.J., Hunt, G.R. *et al.* (2010) An investigation into the cognition behind spontaneous string pulling in New Caledonian crows. *PLoS ONE* 5(2), e9345. DOI: 10.1371/journal.pone.0009345.
- Templer, V.L., Lee, K.A. and Preston, A.J. (2017) Rats know when they remember: transfer of metacognitive responding across odor-based delayed match-to-sample tests. *Animal Cognition* 20(5), 891–906. DOI: 10.1007/s10071-017-1109-3.
- von Frisch, K. (1927) *Aus dem Leben der Bienen*, Springer verlag. Translated 1953: *The Dancing Bees: An Account of the Life and Senses of the Honey Bee*. Harvest Books, New York.

Genetic Components of Behaviour

Like the majority of animal traits, such as milk yield, fur colour and longevity, behaviour is subject to genetic mechanisms. This will vary among different types of behaviour, and the environment in which an animal is raised affects the expression of different behavioural components. However, as described in Chapter 2, even though domestic sows have been selected for many decades in an indoor environment with, until recently, limited space and quite barren features, these animals have still retained the ability to build a nest (Stolba and Wood-Gush, 1989). Similarly, laying hens have been bred to produce a large amount of eggs without becoming broody (i.e. cessation of egg laying and high motivation to lie on and incubate the eggs), so this behavioural trait (broodiness) has almost completely disappeared, yet can be revived if the selection pressure against it is eased (Sharp, 2009). The genetic component of behaviour makes it possible to select for behavioural traits, such as maternal ability in mice, and against others, like aggression in dogs. In order to carry out behavioural selection without unwanted side-effects, we need to know more about the links between various traits. Unless we have an estimate of the genetic correlations among wanted and unwanted behavioural characteristics, we cannot ensure that the individuals used for breeding are selected in an optimal way. One way to elucidate some of these heritable links is via behavioural tests.

Selection For and Against Behavioural Traits

One of the major problems in modern egg production is the occurrence of feather pecking, where birds peck and pull out the feathers of other birds. Despite the often gruesome consequences, this is a non-aggressive behaviour, which is difficult to study because it often emerges quite suddenly, and not all birds peck. In order to be able to study the behaviour in a more systematic way, and to confirm the suggestion that feather pecking could be selected against (Kjær and Sørensen, 1997), Kjær *et al.* (2001) carried out a long-term selection trial. Hens, which had been categorized

according to their feather pecking pedigree using a genetic model, were used as breeding stock for the creation of divergent lines: a high pecking (HP) and a low pecking (LP) line. In each generation, the birds were housed in groups of 20 individuals with an even mix of both lines in each group. The groups were observed for 3 h in order to measure the number of feather pecking bouts per hour of each bird, as this was the breeding value criterion estimated by the genetic model. In the third generation, this value was 4.6 bouts per hour for the HP hens and 0.6 for the LP hens, compared with the generation zero value of 1.9 bouts per hour. The behavioural observations thus confirmed that the genetic calculations used had indeed produced changes between the two lines in their propensity to feather peck. Interestingly, the HP hens were heavier than hens from the LP line by the third generation, indicating that these traits may be linked. One of the reasons for using genetic models for behavioural selection is that it becomes increasingly difficult to observe difference between individuals in the LP line as the feather pecking diminishes and it is no longer possible to rank the birds. In order to continue selection based on behavioural observations (which are also time consuming), it would be necessary to create environments that are more likely to provoke feather pecking. This is obviously not a desirable solution.

Divergent selection of mice for and against activity in the open field test (see Chapter 4) was carried out for 30 generations by DeFries *et al.* (1978). This resulted in the active line being 30 times more active than the inactive line, with the control line intermediate between the two (Fig. 10.1a). The mice had been selected based on 3 min testing in an open field arena on 2 consecutive days when they were around 40 days of age, with number of

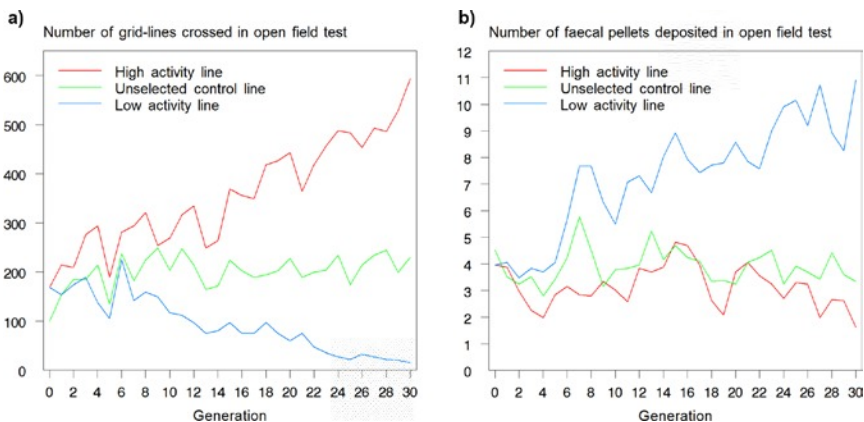


Fig. 10.1. Lines of mice selected for 30 generations for high and low activity in an open field test, as well as an unselected control line. (a) Number of grid-lines crossed in the open field test across generations; (b) number of faecal pellets deposited during the test (adapted from DeFries *et al.*, 1978).

grid-lines crossed logged automatically by photo-beams. From Fig. 10.1a it is evident that the same type of floor effect as in the previous feather pecking example appears to be reached for the line with low activity, as this variable cannot descend below zero. However, the low activity line also deposits three times as many faecal pellets during the test (Fig. 10.1b). This high defecation rate is an indirect consequence of the selection against activity, indicating an increased level of fear in this line of mice (Forkman *et al.*, 2007). So, selection against activity in an open field test is associated (unintentionally) with selection for fearfulness. It is often possible to prevent this from happening as such genetic correlations are rarely 100% linked. In other words, one could include a measure of fearfulness in the selection criteria, and choose to breed for mice with both low activity and low fear. It would result in a less rapid divergence between the lines than that seen in Fig. 10.1a.

Correlated Genetic Changes

Non-behavioural traits can also be indirectly co-selected when animals are chosen for breeding based on their responses in behavioural tests. Quail divergently selected for long and short duration of tonic immobility (TI) in a TI test (see Chapter 4) produced eggs with significantly different yolk steroid levels, so that higher levels of progesterone and androstenedione were found in the birds with low TI duration, a behaviour that is a reflection of low inherent fearfulness (Bertin *et al.*, 2009). The eggs produced by the more fearful quail were larger and with less heavy eggshells. Another aspect of genetic correlations is when different behavioural traits show correlations over time, so that one behavioural characteristic can be used to predict another when the animal is older. Response of female pigs to handling at 5 weeks of age (Horback and Parsons, 2018) has been found to predict their propensity to be aggressive when they become sows (but see also the discussion of personality in Chapter 4).

Epigenetics and Behaviour

Natural variation in behavioural characteristics, such as maternal behaviour, can be used in the selection of breeding dams to promote their ability to raise young. Champagne *et al.* (2003) found that the amount of licking and grooming of pups in their first week of life differed among female Long—Evans rats, ranging from 5 to 20% of the time spent in this behaviour, and being normally distributed (i.e. following a bell curve) within the cohort studied. It was a stable characteristic for individual dams, and was independent of the amount of time they spent in contact with their pups. The authors also found that they passed this behavioural trait to their

female offspring, although they obviously showed the same natural variation as their mothers. But then another effect arrived on the scene.

It has long been known that the environment affects the expression of different genes. Domestic pigs, for example, are usually not very hairy, but if they are reared in unheated housing during a Scottish winter, they become incredibly furry (personal observation). This is how many genes work, but it was not until more recently that some of these environmental effects were discovered to have become permanent and heritable: the emergence of epigenetics. This is one of the more talked about topics in genetics in recent years, and I mention it here because behaviour has been found to affect epigenetic programming. In another study of maternal behaviour in rats, Weaver *et al.* (2004) found that rat mothers who licked and groomed their pups a lot and arched their back when suckling their young effectively altered the epigenome of their offspring. To check that this was not just a phenotypic correlation, the researchers cross-fostered rat pups between mothers with high and low grooming tendency, and found that the epigenomic changes were linked to the rearing dam and not the biological mother of the rat pups. The pups reared by high grooming dams showed reduced hypothalamic–pituitary–adrenal (HPA) responses to stress as adults. At the time of writing, this article (Weaver *et al.*, 2004) has been cited over 3500 times.

Per Jensen (2013, 2014 and 2015) was among the first to consider the importance of epigenetics for behavioural genetics in general, and the domestication process in particular. He and his colleagues compared two breeds of chickens: the domesticated strain White Leghorn and the ancestor to all chickens, the red junglefowl (Lindqvist *et al.*, 2007). As opposed to rats, chickens can be hatched and reared without the presence of a mother hen, eliminating any effects of maternal care. Birds from both breeds were housed and raised with unpredictable light schedules, which is stressful for them and which led to reduced spatial learning in both breeds. However, this epigenetic change was only transferred to the offspring in the domesticated breed. Jensen (2014) thinks that this is because the behavioural selection that happened during domestication has led to an increased ability to respond to epigenetic changes (Fig. 10.2). Indeed, the following year, he suggested that ‘the interaction between the genome and the environment is far more dynamic and complex than previously thought and our view on domestication and evolution may have to be rather drastically revised’ (Jensen, 2015, p. 35).

Genotype and Environment Interactions

One way to disentangle the respective effects of genetics and environment on behaviour is through cross-fostering experiments, as in the example above, or – even better – using embryo transfer. Dwyer and Lawrence

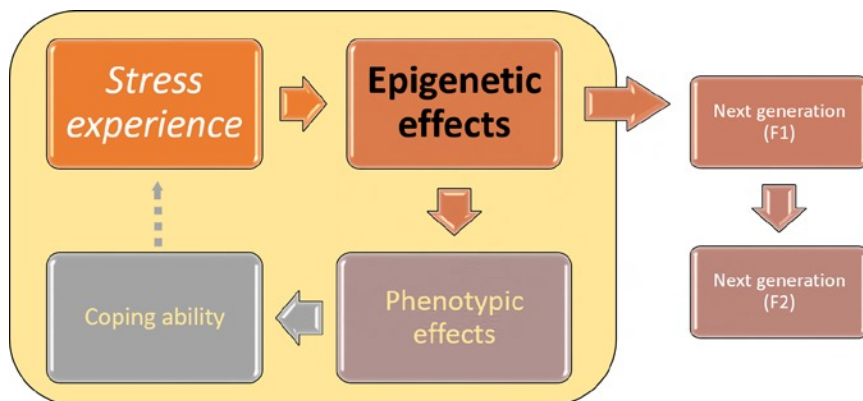


Fig. 10.2. Diagram showing the association between stress experience and epigenetic change. Stressful events can lead to modifications in gene expression (epigenetic changes), which affects the animal itself (phenotypic effects) leading to altered ability to cope with future stress. These epigenetic changes may transfer to the next generation (*in ovo* or *in utero*) or further generations by epigenetic germline modifications (adapted from Jensen, 2015).

(2000) investigated the offspring of two breeds of sheep, Suffolk and Scottish Blackface, known to differ in their foraging and social behaviour, as well as level of activity. Embryos of Suffolk and Blackface were implanted in 30 ewes from each breed, half of them giving birth to lambs of their own breed and the other half to lambs of the other breed in a 2×2 experimental design. The lambs were thus all reared by their birth mother, but differed in their genetic origin. The behaviour of the lambs for the first few days of life depended on the breed of the lamb, independent of the ewe that raised it. However, subsequently, the behaviour of the lambs was significantly influenced by the breed of the ewe, with lambs reared by Blackface mothers suckling less often, and being more active, both before and after weaning. When at grass, Blackface ewes and their lambs stayed closer together, independent of the breed of the lamb, compared with Suffolk ewes. This study was carried out before epigenetics became the new black, and it would have been interesting to follow the subsequent maternal behaviour of the lambs when mature.

Selection for stable behavioural characteristics, such as the maternal behaviour in the example above, can be done based on behavioural tests. When mink are raised for fur production, individuals differ in their behavioural characteristics, with some mink showing high levels of fear. Although fur production is a highly controversial subject in terms of animal welfare, it would – all other things being equal – be an improvement for the welfare of farmed mink if they were less fearful. For this to happen, we need to know if this characteristic is sufficiently heritable, and if any potential test of fearfulness is picking up the trait in general; in other words, if we select

mink based on their response to humans, will their fearfulness in general be reduced? Hansen (1996) divergently selected mink based on their reaction when a human put a wooden spatula through the wire cage and held it there for up to 60 s. The mink that retreated to the back of the cage were chosen for breeding a fearful selection line, whereas the mink that approached, sniffed and bit the spatula were selected for the confident line. After ten generations of selection, the mink were tested in six different behavioural tests to find out if the difference in fearfulness towards humans between the two lines was consistent across other situations (Malmkvist and Hansen, 2002). The tests included both social and non-social stimuli, including the spatula test (voluntary human contact), attempts to handle the mink by a gloved human hand (forced human contact), exposure to a novel object (a wooden cube), a novel feed (cat food), a novel environment and an unknown conspecific. Mink from the two selection lines differed in the expected manner, with the confident mink showing less fearful responses across all tests compared with mink from the divergently selected line, indicating a general difference in fearfulness across several social and non-social situations. These results show that it is possible to devise simple behavioural tests for use as selection criterion and still achieve general genetic changes in animal personality.

Strain and Breed Differences in Behaviour

From a behavioural point of view, one of the factors giving rise to differences in behavioural traits is the breed or strain of the animal model used. Dog owners will be well aware that different breeds of dogs may react differently to a given situation, although this can be difficult to disentangle from the environmental effects caused by the training of the dog and the competence of the owner. Strain differences are abundant in rodents, with certain mice strains being more aggressive than others (e.g. Schicknick *et al.*, 1993). Behavioural differences are often observed between pigmented and albino rat strains, with the latter having a behavioural expression similar but not identical to that of offspring of wild-caught rats (Stryjek *et al.*, 2012, 2013). When pigmented Long–Evans rats and albino Wistar rats were tested for their ability to perceive motion detection using a random pattern of moving dots, the pigmented strain was found to have a well-developed perception of motion, whereas the albino strain showed severe impairment (Hupfeld and Hoffmann, 2006). Similar impairment of albino rat strains has been found in the sexual response of males to the odour of female rats in heat (Sachs, 1996). It is also worth noting that albinism became fixed at generation eight in the mice selected for low activity in the open field test depicted in [Fig. 10.1](#) (DeFries *et al.*, 1978).

So what is the link between pigmentation and behaviour? When Russian scientists Dmitry Belyaev and Lyudmila Trut started selecting farmed foxes

based on their level of tameness, they mimicked the domestication process (Belyaev *et al.*, 1979; Belyaev and Trut, 1986). Behavioural changes in the foxes continued over generations, but the animals also became less uniform in their fur colour, had curlier tails and more floppy ears. The increase in fur pigmentation is thus the opposite of the increased albinism seen in the more fearful line of mice in Fig. 10.1. Recently it has been suggested that the morphological changes seen in the Russian fox breeding experiment have been more directly favoured by humans in our domestication of companion animals, because we have a tendency to infer calmness from cuteness (Lane, 2018).

It is not only laboratory and companion species that show behavioural differences among breeds; this has also been found in livestock, such as poultry. Perches are environmental structures that are used a lot by laying hens (Brendler and Schrader, 2016; Campbell *et al.*, 2016). When different strains of slow-growing broiler chickens are raised with access to perches, they do not use the perches to the same extent (Nielsen, 2004; Fig. 10.3). Both strains show similar growth trajectories, so differences in perching behaviour are not a consequence of size differences. They may be linked to differences in the form of the keel bone, as the low-perching strain in this example (Labresse) had more breast blisters, potentially caused by more sharply angled keel bones, which may make perching painful.

When we use animal models selected for specific non-behavioural characteristics such as certain diseases in inbred strains of rodents, there may be associated behavioural changes. Such changes in behaviour can be undesirable, in particular if they affect the results of the experiments where these models are used. The extent to which this is widespread among the many different animal models available today is unknown. In some cases, it is difficult to distinguish between genetic correlations and physiological constraints of a selection: mice selected for inherent obesity show reduced aggression (Yamashita *et al.*, 1989), and genetically obese rats are less active, less exploratory and have a higher level of anxiety-like behaviour (Vogel *et al.*, 2017). The reduced activity in obese rats could be a phenotypic rather than a genetic consequence of being obese. However, when the animals were compared within strain, differences in body weight did not correlate with behavioural differences, indicating that the behavioural changes are of genetic origin.

Domestication *per se* is found to lead to a reduction in stress responses, as calmer animals are selected. It is generally thought that a less reactive animal may have better welfare as it is less fearful and copes better, both in general and with novelty. However, Rauw *et al.* (2017) suggest that in some production species the coping individuals can still have a high stress response, but may just not show this behaviourally, which raises a number of ethical questions. Narrow and specific selection for production traits is bound to lead to behavioural changes, e.g. selection for productivity in laying hens is found to be associated with an increase

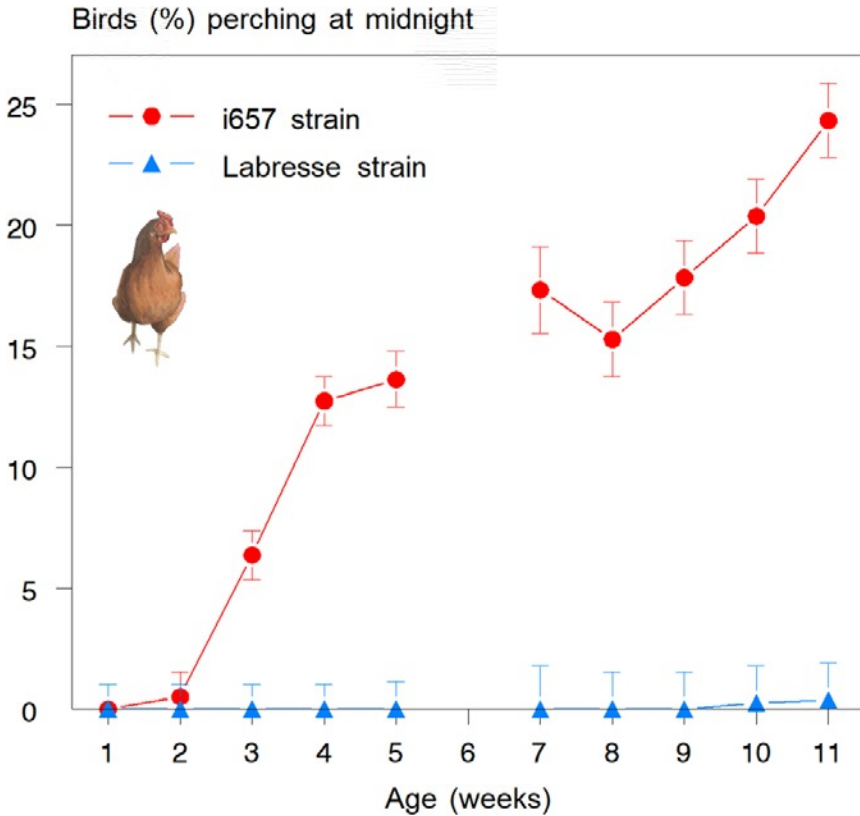


Fig. 10.3. Use of perches by two slow-growing broiler strains. Until 5 weeks of age, the birds were raised indoors, and subsequently moved in week six to housing in tents with access to an outdoor area (data from Nielsen, 2004).

in aggressiveness (Muir, 2003). Such genetic correlations can be difficult to identify, especially in production species, where some behavioural traits may be selected for or against almost involuntarily. This can happen when, for example, a high-producing dairy cow is difficult to milk due to kicking, or when a prolific sow is too aggressive to handle. In such cases, animals may be culled and taken out of the breeding stock for non-productive reasons, but these causes are rarely systematically registered. There are also significant influences of the selection environment, which is beyond the scope of this book. An example is the feeding behaviour of brown trout, which was thought to be affected by the genetic selection for growth. It turned out to be differences in the social environment during the growth period that led to a correlated response in feeding behaviour (Boujard *et al.*, 2007).

Innate and Learned Behaviour

I will finish this chapter with some examples of behavioural tests that have been used to distinguish between innate and learned behaviour. They are unashamedly all from the world of olfaction, and concern laboratory rodents. The first is the response of rats to the smell of a predator in the form of faeces collected from farmed foxes. When rats are exposed to the odour of fox faeces, they show an increase in freezing behaviour, i.e. immobility, often sudden, with raised ears and open eyes (Rampin *et al.*, 2006). As the rats used in these tests were bred in the laboratory, they had never before been exposed to fox odour, and any behavioural response is therefore considered to be innate. A compound (trimethylthiazoline or TMT) was found in fox faeces by Vernet-Maury (1980), and she identified this as the odorant that induced fear responses in rats. TMT has since been used as a predator odour proxy in many studies (Rosen *et al.*, 2015). However, in a recent analysis we failed to find TMT in fox faeces, and have encouraged others to repeat this analysis using more refined methods (Rampin *et al.*, 2018). Another potentially innate odour response is found in male rats that are mature but sexually naïve; that is, with no prior experience of mating behaviour. They show increased sexual responses, in the form of penile erections, when exposed to faeces from a female rat in oestrus (i.e. in heat) compared with when presented with a neutral herbal odour (Nielsen *et al.*, 2016). This would indicate that the response to this odour is innate, although subsequent sexual experience significantly increases the response to the odour of receptive female rats, indicating a degree of associative learning, with the male rats becoming more sensitized to oestrous odours through experience.

Seemingly innate behavioural responses can sometimes turn out not to be innate. It has long been known that mice born without a sense of smell (anosmic mice) have great difficulty in initiating suckling, and most of them die as a consequence (Wong *et al.*, 2000). In rabbits, a suckling pheromone has been identified, which initiates the suckling response in newborn rabbits (Schaal *et al.*, 2003), so it was an obvious idea to try and look for a similar mechanism in mice. Logan *et al.* (2012) studied 2-h-old normal (as in a functioning sense of smell), wild-type strains of laboratory mice, born by caesarean section. The authors found that all pups initiated suckling less than 60 s after being presented with the nipple of their mother. Cleaning the nipple with water, on the other hand, resulted in 64% of pups failing to suckle within 15 min, which was the duration of the test (Fig. 10.4). Different odours, such as vanillin and garlic, were then applied to the washed nipple, but only amniotic fluid elicited the suckling response, with over 90% of the pups beginning to suckle within the 2 min test duration when amniotic fluid was used as the odorant. However, if the – previously ineffective – odours were brushed on to the washed nipples of the dam as well as on to the pups themselves within 5 min of their being

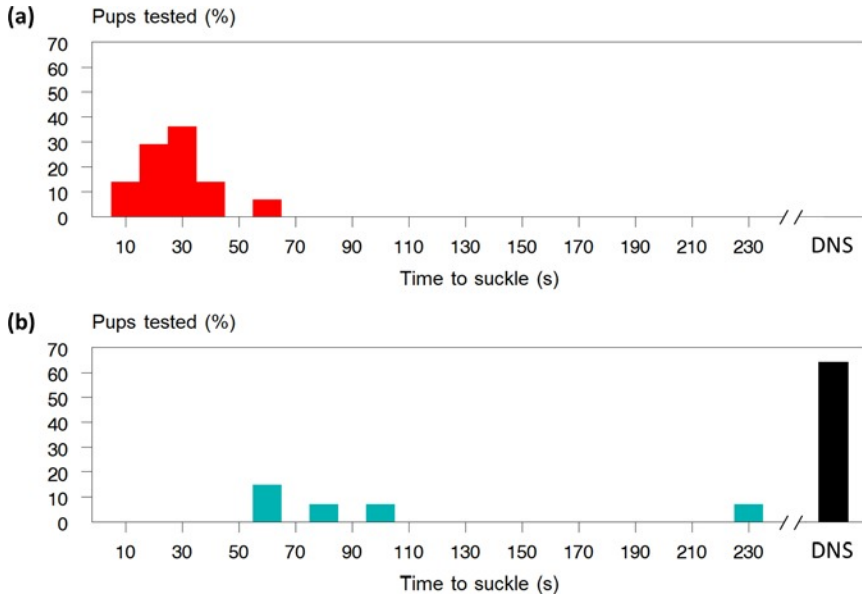


Fig. 10.4. Time to first suckle (s) in 2-h-old mice pups born via caesarean section. (a) If the nipples of the dam were left unwashed, all pups suckled within 1 min. (b) However, if the nipples had been washed clean, i.e. all odours removed, a large proportion of pups (64%) did not suckle (DNS) within the 15 min duration of the test (data from Logan *et al.*, 2012).

born, they responded as the pups tested with amniotic fluid. Finally, the researchers fed garlic- or vanillin-flavoured food to pregnant dams, and these odours were now able to initiate suckling when applied to the nipple. This shows us that, under normal circumstances, suckling initiated by an odour is something the pup learns *in utero*, and is not an innate behavioural response.

Concluding Remarks

It is not surprising in itself that a lot of behaviour seen in animals is heritable. This is, after all, how most traits are passed on from generation to generation, and is the foundation of the process of domestication. What is important to keep in mind when using behavioural tests for animals is the correlations that exist among different traits, whether they are behavioural or not. We can deliberately exacerbate differences between groups of animals in their behaviours by divergently selecting for certain behavioural characteristics, often measured via specific behaviour tests. This can allow us to study the consequences of these differences in more detail, but we

should always keep in mind that the simple measure used for selection may carry with it other, sometimes undesirable, traits.

References

- Belyaev, D.K. and Trut, L.N. (1986) Genetic interrelations of specific changes in standard coat color of silver foxes (singes and star) arising during domestication. *Genetika* 22, 119–128.
- Belyaev, D.K., Ruvinsky, A.O. and Trut, L.N. (1979) Significance of inherited gene activation-inactivation in the domestication of animals. *Genetika* 15, 2033–2050.
- Bertin, A., Richard-Yris, M.A., Houdelier, C., Richard, S., Lumineau, S. *et al.* (2009) Divergent selection for inherent fearfulness leads to divergent yolk steroid levels in quail. *Behaviour* 146, 757–770.
- Boujard, T., Ramezi, J., Vandeputte, M., Labbé, L. and Mambrini, M. (2007) Group feeding behavior of brown trout is a correlated response to selection for growth shaped by the environment. *Behavior Genetics* 37(3), 525–534. DOI: 10.1007/s10519-007-9148-1.
- Brendler, C. and Schrader, L. (2016) Perch use by laying hens in aviary systems. *Applied Animal Behaviour Science* 182, 9–14. DOI: 10.1016/j.applanim.2016.06.002.
- Campbell, D.L.M., Makagon, M.M., Swanson, J.C. and Siegford, J.M. (2016) Perch use by laying hens in a commercial aviary. *Poultry Science* 95(8), 1736–1742. DOI: 10.3382/ps/pew111.
- Champagne, F.A., Francis, D.D., Mar, A. and Meaney, M.J. (2003) Variations in maternal care in the rat as a mediating influence for the effects of environment on development. *Physiology & Behavior* 79(3), 359–371. DOI: 10.1016/S0031-9384(03)00149-5.
- DeFries, J.C., Gervais, M.C. and Thomas, E.A. (1978) Response to 30 generations of selection for open-field activity in laboratory mice. *Behavior Genetics* 8(1), 3–13. DOI: 10.1007/BF01067700.
- Dwyer, C.M. and Lawrence, A.B. (2000) Effects of maternal genotype and behaviour on the behavioural development of their offspring in sheep. *Behaviour* 137(12), 1629–1654. DOI: 10.1163/156853900502754.
- Forkman, B., Boissy, A., Meunier-Salaün, M.-C., Canali, E. and Jones, R.B. (2007) A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiology & Behavior* 92(3), 340–374. DOI: 10.1016/j.physbeh.2007.03.016.
- Hansen, S.W. (1996) Selection for behavioural traits in farm mink. *Applied Animal Behaviour Science* 49(2), 137–148. DOI: 10.1016/0168-1591(96)01045-3.
- Horback, K.M. and Parsons, T.D. (2018) Ontogeny of behavioral traits in commercial sows. *Animal* 12(11), 2365–2372. DOI: 10.1017/S1751731118000149.
- Hupfeld, D. and Hoffmann, K. (2006) Motion perception in rats (*Rattus norvegicus* sp.): deficits in albino Wistar rats compared to pigmented Long-Evans rats. *Behavioural Brain Research* 170(1), 29–33. DOI: 10.1016/j.bbr.2006.01.022.
- Jensen, P. (2013) Transgenerational epigenetic effects on animal behaviour. *Progress in Biophysics and Molecular Biology* 113(3), 447–454. DOI: 10.1016/j.pbiomolbio.2013.01.001.
- Jensen, P. (2014) Behavior genetics and the domestication of animals. *Annual Review of Animal Biosciences* 2(1), 85–104. DOI: 10.1146/annurev-animal-022513-114135.

- Jensen, P. (2015) Adding 'epi-' to behaviour genetics: implications for animal domestication. *Journal of Experimental Biology* 218(1), 32–40. DOI: 10.1242/jeb.106799.
- Kjær, J.B. and Sørensen, P. (1997) Feather pecking in White Leghorn chickens – a genetic study. *British Poultry Science* 38, 333–341.
- Kjær, J.B., Sørensen, P. and Su, G. (2001) Divergent selection on feather pecking behaviour in laying hens (*Gallus gallus domesticus*). *Applied Animal Behaviour Science* 71(3), 229–239. DOI: 10.1016/S0168-1591(00)00184-2.
- Lane, J.D. (2018) Revisiting the famous farm foxes: a psychological perspective. *Learning & Behavior* 46(4), 586–590. DOI: 10.3758/s13420-018-0333-2.
- Lindqvist, C., Janczak, A.M., Nätt, D., Baranowska, I., Lindqvist, N. *et al.* (2007) Transmission of stress-induced learning impairment and associated brain gene expression from parents to offspring in chickens. *PLoS ONE* 2(4), e364. DOI: 10.1371/journal.pone.0000364.
- Logan, D.W., Brunet, L.J., Webb, W.R., Cutforth, T., Ngai, J. *et al.* (2012) Learned recognition of maternal signature odors mediates the first suckling episode in mice. *Current Biology* 22(21), 1–10. DOI: 10.1016/j.cub.2012.08.041.
- Malmkvist, J. and Hansen, S.W. (2002) Generalization of fear in farm mink, *Mustela vison*, genetically selected for behaviour towards humans. *Animal Behaviour* 64(3), 487–501. DOI: 10.1006/anbe.2002.3058.
- Muir, W. (2003) Indirect selection for improvement of animal well-being. In: Muir, W.M. and Aggrey, S.E. (eds) *Poultry Genetics, Breeding, and Biotechnology*. CAB International, Wallingford, UK, pp. 67–82.
- Nielsen, B.L. (2004) Breast blisters in groups of slow-growing broilers in relation to strain and the availability and use of perches. *British Poultry Science* 45(3), 306–315. DOI: 10.1080/00071660410001730798.
- Nielsen, B.L., Jérôme, N., Saint-Albin, A., Ouali, C., Rochut, S. *et al.* (2016) Oestrus odours from rats and mares: behavioural responses of sexually naive and experienced rats to natural odours and odorants. *Applied Animal Behaviour Science* 176, 128–135. DOI: 10.1016/j.applanim.2016.01.014.
- Rampin, O., Jérôme, N., Briant, C., Boué, F. and Maurin, Y. (2006) Are oestrus odours species specific? *Behavioural Brain Research* 172(1), 169–172. DOI: 10.1016/j.bbr.2006.04.005.
- Rampin, O., Jérôme, N., Saint-Albin, A., Ouali, C., Boué, F. *et al.* (2018) Where is the TMT? GC-MS analyses of fox feces and behavioral responses of rats to fear-inducing odors. *Chemical Senses* 43(2), 105–115. DOI: 10.1093/chemse/bjx075.
- Rauw, W.M., Johnson, A.K., Gomez-Raya, L. and Dekkers, J.C.M. (2017) A hypothesis and review of the relationship between selection for improved production efficiency, coping behavior, and domestication. *Frontiers in Genetics* 8, 134. DOI: 10.3389/fgene.2017.00134.
- Rosen, J.B., Asok, A. and Chakraborty, T. (2015) The smell of fear: innate threat of 2,5-dihydro-2,4,5-trimethylthiazoline, a single molecule component of a predator odor. *Frontiers in Neuroscience* 9(72), 292. DOI: 10.3389/fnins.2015.00292.
- Sachs, B.D. (1996) Penile erection in response to remote cues from females: albino rats severely impaired relative to pigmented strains. *Physiology & Behavior* 60(3), 803–808. DOI: 10.1016/0031-9384(96)00158-8.
- Schaal, B., Coureaud, G., Langlois, D., Giniès, C., Sémon, E. *et al.* (2003) Chemical and behavioural characterization of the rabbit mammary pheromone. *Nature* 424(6944), 68–72. DOI: 10.1038/nature01739.

- Schicknick, H., Hoffmann, H.-J., Schneider, R. and Crusio, W.E. (1993) Genetic analysis of isolation-induced aggression in the mouse. III. Classical cross-breeding analysis of differences between two closely related inbred strains. *Behavioral and Neural Biology* 59(3), 242–248. DOI: 10.1016/0163-1047(93)91010-K.
- Sharp, P.J. (2009) Broodiness and broody control. In: Hocking, P.M. (ed.) *Biology of Breeding Poultry. Poultry Science Symposium Series*. 29. CAB International, Wallingford, UK, pp. 181–205.
- Stolba, A. and Wood-Gush, D.G.M. (1989) The behaviour of pigs in a semi-natural environment. *Animal Science* 48(2), 419–425. DOI: 10.1017/S0003356100040411.
- Stryjek, R., Modlińska, K. and Pisula, W. (2012) Species specific behavioural patterns (digging and swimming) and reaction to novel objects in wild type, Wistar, Sprague-Dawley and Brown Norway rats. *PLoS ONE* 7(7), e40642. DOI: 10.1371/journal.pone.0040642.
- Stryjek, R., Modlińska, K., Turlejski, K. and Pisula, W. (2013) Circadian rhythm of outside-nest activity in wild (WWCPS), albino and pigmented laboratory rats. *PLoS ONE* 8(6), e66055. DOI: 10.1371/journal.pone.0066055.
- Vernet-Maury, E. (1980) Trimethyl-thiazoline in fox feces: a natural alarming substance for the rat. In: van der Starre, H. (ed.) *Olfaction and Taste VII. Proceedings of the 7th International Symposium on Olfaction and Taste and of the 4th Congress of the European Chemoreception Research Organization*, 22–25 July 1980. Noordwijkerhout, the Netherlands, p. 407.
- Vogel, H., Kraemer, M., Rabasa, C., Askevik, K., Adan, R.A.H. *et al.* (2017) Genetic predisposition to obesity affects behavioural traits including food reward and anxiety-like behaviour in rats. *Behavioural Brain Research* 328, 95–104. DOI: 10.1016/j.bbr.2017.02.037.
- Weaver, I.C.G., Cervoni, N., Champagne, F.A., D'Alessio, A.C., Sharma, S. *et al.* (2004) Epigenetic programming by maternal behavior. *Nature Neuroscience* 7(8), 847–854. DOI: 10.1038/nn1276.
- Wong, S.T., Trinh, K., Hacker, B., Chan, G.C.K., Lowe, G. *et al.* (2000) Disruption of the type III adenylyl cyclase gene leads to peripheral and behavioral anosmia in transgenic mice. *Neuron* 27(3), 487–497. DOI: 10.1016/S0896-6273(00)00060-X.
- Yamashita, J., Hayashi, S. and Hirata, Y. (1989) Reduced size of preputial glands and absence of aggressive-behavior in the genetically-obese (ob/ob) mouse. *Zoological Science* 6, 1033–1036.

Other Test Considerations

Behavioural testing of animals covers a vast range of methods, from simple tests, such as presenting your cat with two types of feed, to quite complex protocols, like training monkeys to fix their gaze on a dot while they detect a peripheral feature change (Fischer and Wegener, 2018; Chapter 8). In this chapter, I have collected a variety of issues to consider, things to remember before, during and after testing, and pitfalls to avoid. They are in no particular order, and there are likely to be some – even many – subjects missing. The idea is to highlight certain overall matters that are important to keep in mind, independent of the complexity of your protocol and the question you want to ask.

Do We Measure What We Think We Do?

Tiny copper rods move when put on a gently vibrating plate (Ravilious, 2017). When at low densities, the rods move in a random fashion, but at high densities, they appear to move in a more organized way, producing ‘flocking characteristics’. As copper rods are inanimate objects, we know that these organized movements are not intentional on the part of the rods. However, how can we know when observed behavioural changes originate from the animal tested, and when they are artefacts of external, physical factors? In Chapter 2, the example of clear, rhythmic undulations of activity in groups of day-old chicks (Nielsen *et al.*, 2008; Fig. 2.4) is an illustration of this type of deceptive phenomenon. The activity patterns could not be attributed to an obvious ‘zeitgeber’ (a time-keeper; a rhythmically occurring external cue that regulates the observed patterns) such as the onset of the ventilation fans or something equivalent. We originally thought it could be proof of an innate rhythmicity in young chicks that disappeared in the absence of a mother hen to keep it going. Alas, it turned out to be an artefact of adding together the binomial states of individuals being either passive or active. This gives rise to seemingly regular bursts of activity of the

whole group as long as the passive phase is longer than the active phase for each individual.

Sometimes we use experimental stimuli or markings that do not work as intended. David Lahti (2015) has written an excellent review on the influence of egg characteristics on egg rejection in certain bird species. He discusses the consequences of using eggs of exaggerated and artificial sizes, colours and materials to study this phenomenon. We can all learn a lot from Lahti's description of how he carefully constructed painted, wooden eggs to see which colour the birds would reject, only to discover later that the birds in question would use their beak to puncture any egg they wanted to remove – a behaviour not possible with the wooden replicates. In the author's own words 'What is it that distinguishes a good, biologically relevant and scientifically productive use of artificial stimuli, from one that demonstrates nothing but an idiosyncratic behavioral response to an oddity?' (Lahti, 2015, p. 530). He highlights the importance of not varying more than one feature at a time, and – when using supernormal stimuli – not to differ so much from what is naturally possible that the context of the stimulus is lost. This is often made most obvious to us when we inadvertently make mistakes. Another, well-known example is the use of differently coloured leg-bands to identify birds, where Nancy Burley unintentionally made some male birds more attractive to the females, as the latter preferred some colours over others (Burley, 1988).

Many issues need to be considered when using a behavioural test, and one of these is whether the possible outcomes are always interpretable. Although this may be stating the obvious, tests exist where care needs to be taken to prevent undecipherable results. One such example is the buried food test. This is a test that has been widely used in the assessment of olfaction in rodents, and I have to admit that I do not like this test. On the surface, the test paradigm is very simple: the test animal is placed in an arena where a small food reward is buried in the litter. The experimenter simply records the time it takes for the animal to find the food reward, and this is supposed to measure an animal's general ability to smell (Yang and Crawley, 2009). My scepticism regarding the usefulness of this simple test is twofold. First, it does not take into account the general activity levels of the animal tested. If the test is used to detect effects of a certain drug on the capacity to smell, it should at least measure the latency to move and the activity of the test animal from the instant it is placed in the test arena. Some drugs can render an animal lethargic while others may induce hyperactivity, and the animal may therefore explore the arena with more or less efficiency. If lethargic, the animal may spend time motionless, which is not a reflection of its olfactory capacity. Likewise, if the animal is hyperactive, it may not simultaneously be searching its surroundings for titbits. Second, factors other than a reduced sense of smell can affect the time taken to retrieve the food reward. Feeding motivation can vary, and it should be ensured that the test animals are at similar levels of satiety when tested. The

types of food reward used, typically cookies or chocolate cereals, may also affect the outcome. If the drug tested leaves the animal feeling nauseous, the motivation to search for sweet treats may be reduced, again for other reasons than olfactory deficits.

Ethograms and Test Development

It can never be emphasized enough that in order to be able to record the behaviour of an animal – whether in a test situation or not – we need to know the behavioural repertoire of the animal in question. For most of the animals commonly used in behavioural research, ethograms have been widely published. However, construction of an ethogram may be necessary if little or no research has been done on the species studied. An ethogram can be more or less detailed, depending on the test situation, but it should also be kept in mind that behavioural components which could be but are *not* displayed during a test may indicate certain limitations of the testing protocol.

Spending time observing your animals is rarely wasted. It can give inspiration for new questions to seek answers to, and one may see behaviours that are not usually displayed in connection with routine maintenance and handling. One of the ways to attack a novel experimental protocol is to put yourself in the place of the animal. We, as humans, have a tendency to focus on our own sensory modalities, and may end up using visual or auditory stimuli when olfactory ones would have been a better choice. Look at (and listen to, and smell) things from different angles, as they may change how you and not least your test animal perceive them (Fig. 11.1). I have previously emphasized the importance of taking animal sense into account,



Fig. 11.1. Stained glass window seen from the outside and the inside of the church in Saint-Roman de Tousque, Gard, France (photo: Birte L. Nielsen).

not only in their housing and management, but also during experimental testing as it will ensure the best and most valid results (Nielsen *et al.*, 2015; Nielsen, 2018).

In Chapter 3, ways to choose a test were described, and examples given of how tests could be adapted to suit specific situations or species. Even if using a validated and commonly used test, it is always a good idea to run a few pilot tests. Pilot testing is, unfortunately, not always possible. This may be due to the available time window for testing being very short, such as in tests of oestrus behaviour of seasonable breeders like vixens, or because the equipment to be used is not available for prior testing. On the other hand, pilot tests can be referred to in scientific articles, sometimes leaving the impression that these were meant to be part of the main experiment, but didn't quite work. In my humble view that is okay, because at least we got to know about them. Even if the treatment showed no significant differences, the protocol may reveal possible reasons for this result, and an experiment is rarely, if ever, result-less (in the broad, non-scientific sense of the word). One emerging way to encourage and allow more 'result-less' experiments to be published is through registered reports, where experimental protocols are logged prior to the experiment, including which statistical methods will be employed. These are peer-reviewed by colleagues and, if found valid, the process allows the results to be published no matter the outcome of the tests, provided the protocol has been followed.

Habituation, Training and Test Environment

It has been underlined several times already, but habituating experimental animals to being handled is key to a successful behavioural test. This is also true even if the test is an open field test, which should be new to the animal at least on their first testing. Animals that are unhandled or scared of humans not only have their welfare compromised when being subjected to behavioural test arenas, but also cannot possibly provide us with valid results if all we are measuring is the anxiety displayed by the test animal (Raundal *et al.*, 2015).

It would be nice to know to what extent our treatment and handling affect the animals tested. An example of the relative stressfulness of open field tests and blood sampling comes from pigs (Herskin and Jensen, 2002). Individuals within a group of piglets were tested in a 10-min open field test on different days around weaning at 4 weeks of age, with each pig tested only once. Another group of piglets were not subjected to the open field test but left in their home pen for the same amount of time as the duration of the test. Blood samples were taken from all pigs before and after the 10-min test period, thus all pigs were subjected to handling and blood sampling; the only difference between the two treatments was the open field test. [Figure 11.2](#) shows the cortisol response of the pigs before and after the

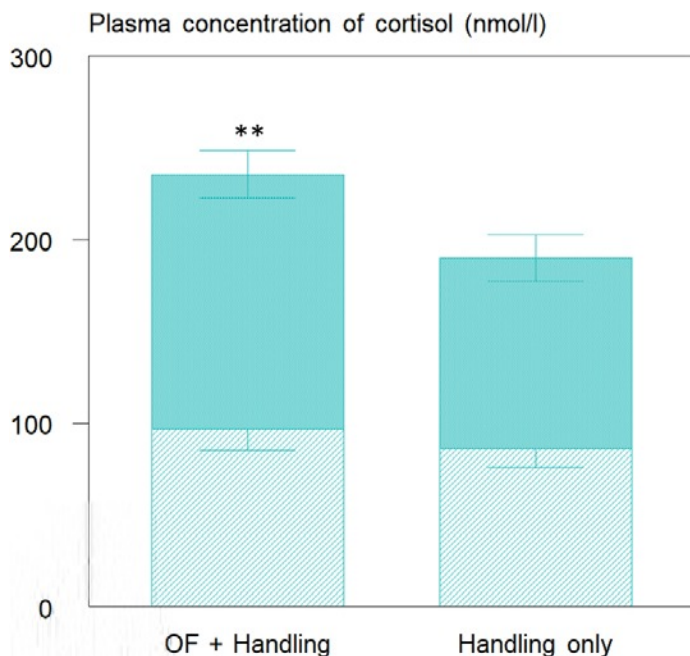


Fig. 11.2. Blood plasma concentration of cortisol (nmol/l) in pigs ($n = 32\text{--}36$ per treatment) before (lower part of the bars) being subjected to one of two treatments: an open field (OF) test and the handling involved during blood sampling, or the handling only while being returned to the home pen instead of an open field test for the 10-min test period. The solid upper part of the bars is the increase in cortisol after the treatment, which was significantly higher for piglets exposed to an open field test (data from Herskin and Jensen, 2002).

10-min test period. No difference was seen in cortisol levels before the test, but the increase after the test was significantly higher for the pigs tested in the open field, and the blood sampling procedure in itself accounted for 74% of this increase.

One way to carry out behavioural tests with a minimum of handling is to make the home environment the test arena. The closed economy set-up described for mink in Chapter 5 is one such approach (Cooper and Mason, 2000). Another method employed is the use of large group-housing facilities for mice, where more than 20 individuals live together without interruption for months at a time (Winter and Schaefer, 2011; Schaefer and Claridge-Chang, 2012). Behavioural data are collected automatically from sensors connected to a water source, accessible through gated tunnels, where the RFID (radio-frequency identification) tagged mice are given tasks involving discrimination of different stimuli in order to obtain drops of water. Thus, no disturbance or transport of the mice is necessary in order to apply different experimental protocols.

Behavioural tests can also be carried out with animals in a natural setting. Such experiments are often put in place to test specific theories on what triggers a given behaviour observed in the wild. One such study investigated if odours from known and unknown predators would induce threat-sensitive foraging in deer (Chamaillé-Jammes *et al.*, 2014). The researchers chose a remote island off the west coast of Canada where a population of black-tailed (sitka) deer had been living for over 90 years. Wolves had never been present in this group of islands, whereas brown bears lived in the area. In the study, 98 bait stations were set up, in which a circle (1 m diameter) of flat ground was sprayed with one of four different odours, including urine from bears and wolves, with water being used as the control treatment. Pieces of apple were then spread on top of the odorized area and small open-top containers holding the odour were pushed into the ground to ensure a continuous perfusion of the smell. Following odorization and baiting, the first visit to each station by a deer was monitored using motion-detector cameras recording videos at 1 frame/s. [Figure 11.3](#) shows the percentage of sniffing observed at these visits and the proportion of apple pieces eaten by the deer. It is clear that the response of the deer to the wolf odour is greater than that seen with water and bear urine. Two additional odours (cologne and petrol) were tested, but did not yield results different from those obtained with water and bear urine, indicating that it is not the novelty of wolf odour that gives rise to the observed behaviour. One explanation for the stronger response is that wolves are a more dangerous threat than black bears to the deer and, despite the long absence of wolves, their odour has remained innately fear-inducing for the deer. The authors add that the different responses to the two predator odours are not related to risk of encounter as wolves range more than black bears, and thus are less likely to be found close to the urination sites.

Although slightly off topic, I have to present an attempt to remake a famous experiment conducted in 1937 by Konrad Lorenz and Nikolaas Tinbergen, in which they flew different cardboard silhouettes over groups of young poultry of various species. One silhouette resembled a goose when flown with the long protrusion forward ([Fig. 11.4a](#)) and a hawk when flown with the short protrusion forward ([Fig. 11.4c](#)). According to Schleidt *et al.* (2011), the two grand masters of ethology differed in their interpretation of the results, with Tinbergen stating that alarm responses were elicited only when the silhouette was flown to resemble a hawk (short neck, long tail; Tinbergen, 1939), whereas Lorenz (1939) concluded that the shape of the silhouette was not important, and novelty and speed were the factors that influenced the response of the birds. Schleidt (1961) did a remake of the experiment on five turkey poults using five different shapes ([Fig. 11.4a-e](#)), all having the same surface area and flown at the same speed and height (2.3 m above the pen). The size of the buzzard silhouette flown at this height corresponds to a real buzzard flying at 27 m height.

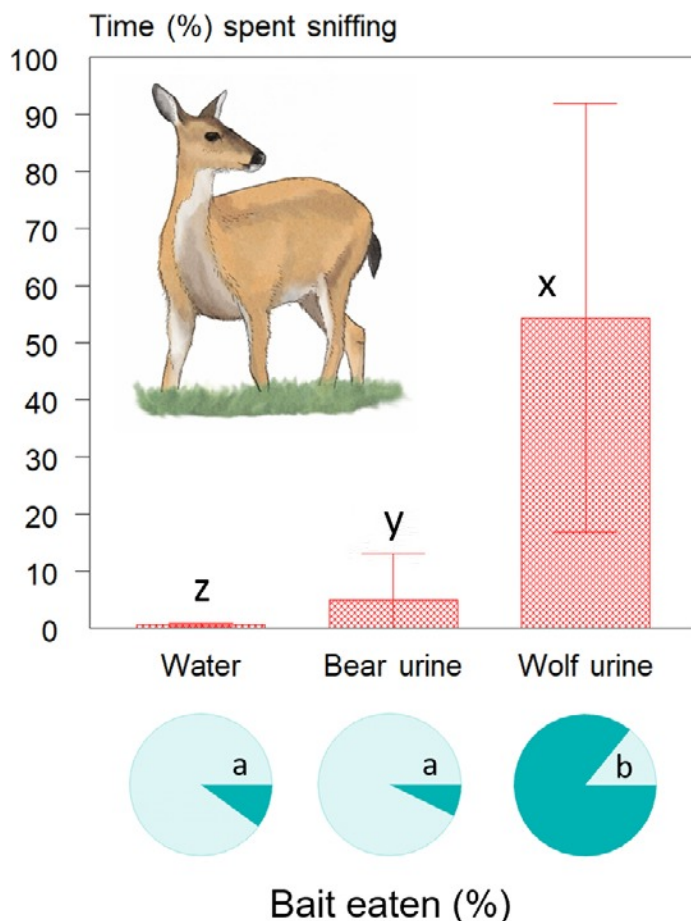


Fig. 11.3. Behavioural response of black-tailed deer when visiting feeding stations sprayed with water (control), bear urine or wolf urine, and baited with apple pieces. The bar chart shows the percentage of time the deer spent sniffing, and the pie charts show the percentage of bait eaten (in light shading) for the three odour treatments, respectively. For each type of chart, different letters indicate significant differences (data from Chamailé-Jammes *et al.*, 2014).

The results are shown in the graph in Fig. 11.4. On day 1, four of the shapes were used once, and this was repeated on day 2. The hawk silhouette was flown last, which may have given rise to the lower alarm call response observed for this shape on day 1. The following day, the birds were already showing a much reduced response. Over the subsequent 22 days, the turkeys were tested on various days, always with ten presentations of the goose shape, with the hawk shape inserted once at a random time along the test. The turkeys responded significantly more to the hawk shape across the test days, but with the greatest response seen on days 3 and 4. This could

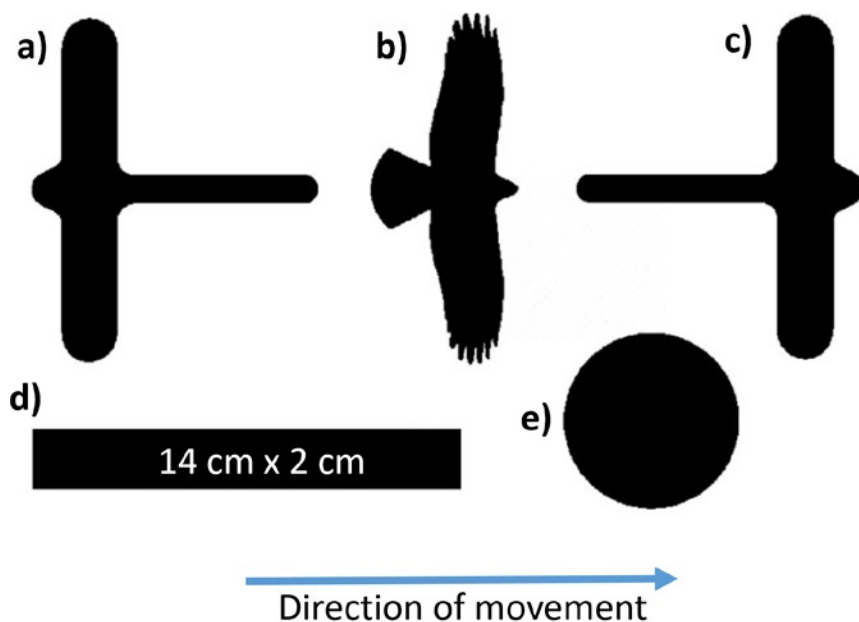
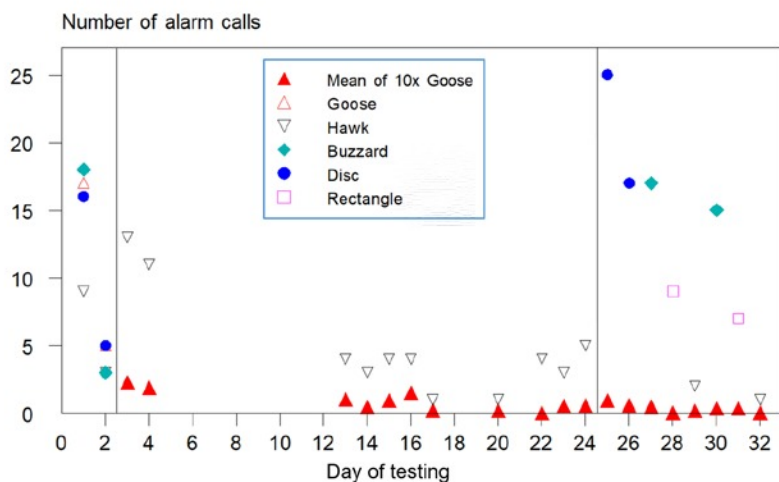


Fig. 11.4. A re-test of the famous experiment by Lorenz and Tinbergen, investigating the alarm response of five young turkeys to different silhouettes gliding above their pen. Five different silhouettes were used: (a) goose; (b) buzzard; (c) hawk; (d) rectangle; and (e) disc, all of the same surface area (28 cm^2) and moving at the same speed (0.25 m/s). Please note that shape (c) is shape (a) reversed. See text for details (based on original data from Schleidt, 1961, reproduced in Schleidt *et al.*, 2011).

indicate that Tinbergen was right, but as the reverse experiment was not performed (i.e. ten times a hawk with the goose inserted at random), the results cannot differentiate between novelty and shape effects. Therefore, from day 25 onwards, other shapes were included as the odd one out during the ten presentations of the goose shape. The results indicate, that Lorenz was right: it is indeed the novelty of the shape that is important for the alarm response, and not the short neck of the flying object (Schleidt *et al.*, 2011).

The experiment by Lorenz and Tinbergen has been used extensively in ethological teaching as an example of innate responses. Indeed, the raptor silhouettes used on windows to prevent birds flying into the pane of glass is based on the Lorenz–Tinbergen raptor icon, although it quickly becomes ineffective as the surrounding birds habituate to the shape as well as its fixed location (Shalter, 1975).

How to Ensure the Best Data Possible?

A multitude of factors may influence the outcome of a behavioural test, and plenty of forethought needs to go into the experimental planning, even when quite simple tests are employed. The importance of baseline observations is perhaps obvious, but it cannot be emphasized enough, especially if the testing period is of extended duration. One variable that is sometimes forgotten in behavioural studies until too late is the recording of body weight. Pre-test weighing is a relatively simple measure, which is often relevant for use as a covariate in the statistical analysis. Regular live weight recording across the testing period may also reveal unexpected fluctuations, the cause of which needs to be looked into.

As stated in Chapter 1, this book is not about statistics. That said, it probably never hurts to remind everyone that for a behavioural test to be valid, the chosen sample size needs to be appropriate for the question asked. Clearly, if the question is ‘Can pigs learn to tap dance?’, and you are able to teach one pig to do so, then a sample size of one is sufficient. Joking aside, surprisingly small sample sizes can be statistically valid if the effect of treatment is large and the variation among test subjects is small. In other words, small studies can only detect large effects (Button *et al.*, 2013). Conversely, large studies can detect quite small effects, some of which may not be biologically relevant. True effect size is difficult to estimate beforehand, but we should always aim for a sample size that is predicted to ensure a statistically valid comparison, without testing more animals than necessary. The statistical power of a test is the probability that it will *correctly* reject the null hypothesis (i.e. when the null hypothesis is false). Clever experimental designs can prevent a lot of heartache when interpreting the results. The use of a Latin square, for example, may allow you to adjust for baseline differences between your test animals as well as differences

i) Sequential testing							ii) Counterbalanced design						
Animal	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Animal	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
ID#1	A	F	E	D	C	B	ID#1	A	B	F	C	E	D
ID#2	B	A	F	E	D	C	ID#2	B	C	A	D	F	E
ID#3	C	B	A	F	E	D	ID#3	C	D	B	E	A	F
ID#4	D	C	B	A	F	E	ID#4	D	E	C	F	B	A
ID#5	E	D	C	B	A	F	ID#5	E	F	D	A	C	B
ID#6	F	E	D	C	B	A	ID#6	F	A	E	B	D	C

Fig. 11.5. Examples of two 6 × 6 Latin square designs where (i) the treatments are always in the same sequence, e.g. treatment F always follows treatment A, and (ii) the treatments are never in the same sequence. The latter is only possible when there is an even number of treatments (from Bradley, 1958).

between test days. For even numbers of treatments, it may be worth considering the Bradley solution to balance for sequential effects (Bradley, 1958; Fig. 11.5). This can prevent systematic carry-over effects, if treatments are likely to affect each other across time. Garamszegi (2016) provides a statistical guide for when sample size is limited or skewed.

One of the factors we use to choose, say, a particular strain of mice for tests in the laboratory is their behavioural characteristics, as different strains are known to differ in their type and level of activity. Crabbe *et al.* (1999) tested groups of eight different strains of mice under near-identical circumstances in a battery of behavioural tests carried out in three different laboratories at the same time. The results from activity in the open field test for four of the strains are shown in Fig. 11.6. As expected, differences among strains are clearly visible, but both the absolute and the relative results vary among the laboratories. The authors call for prudence when interpreting behaviours that differ little, such as is often the case when using gene knock-out lines of rodents, as the observed results may not be the same from one laboratory to the next. For this exact reason, some authors suggest to increase the heterogeneity of the treatment groups to ensure that results are reproducible across laboratories and situations, and not only valid for very narrow and specific scenarios (Richter *et al.*, 2010).

Many behavioural tests still require the recorded or live test scenarios to be transformed into analysable data by observation. Although some can be logged automatically (see Chapter 13), a lot is still dependent on human labour. Video recordings are quite labour intensive to transcribe, and several people are often engaged in observing videos of tests and recording data from the same experiment. In order to ensure that any two observers score the behaviour in the same way, both need to score the same subset of the test videos, thus allowing us to calculate the degree of agreement on a scale from none to perfect (e.g. a Cohen's Kappa coefficient). This can – and should – also be used to check for intra-observer variation, especially when a lot of video recordings are being watched over an extended period by the same person. As time passes, we may gradually and inadvertently change the way we score certain behaviours. This is also the reason why

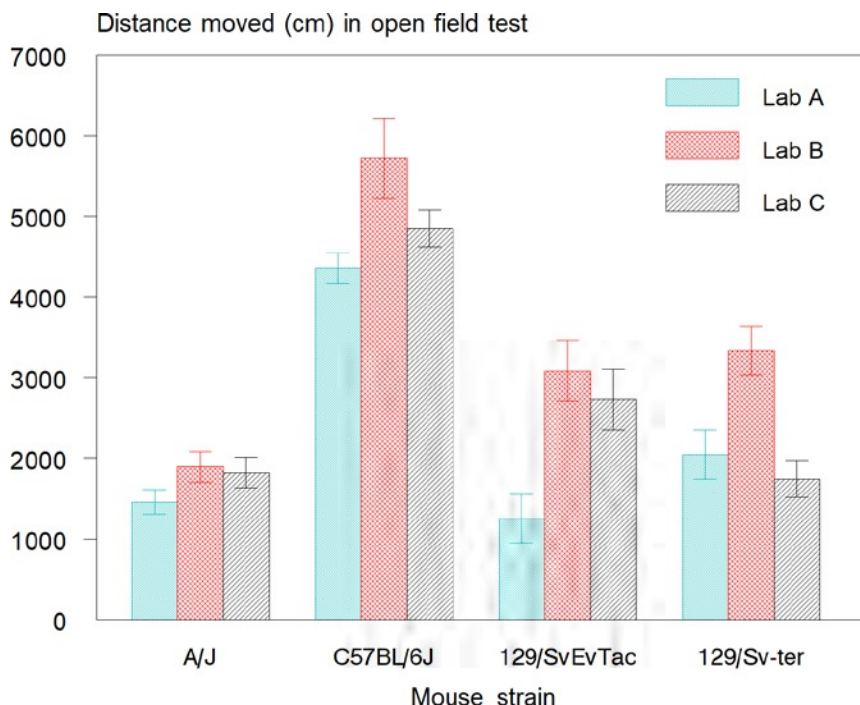


Fig. 11.6. Activity in a 15-min open field test of four different mice strains tested in three different laboratories under near-identical experimental conditions (adapted from Crabbe *et al.*, 1999).

videos should not be observed in chronological order (nor for hours on end! If you are a technician or student observing videos, show this page to your boss or supervisor).

Re-testing the Same Animal

Just a short remark on re-testing an animal using the same behavioural test. It can be useful, as the animal can be used as its own control, but it is often necessary to include animals for the second test with no previous experience of the test, i.e. a non-tested control group. It should also be kept in mind when re-testing individuals that the test is no longer novel to the animal. Indeed, the second test may be influenced by how the first test went, and data from these are clearly not independent measures. This is, in a sense, what is being exploited in the habituation–dishabituation test, where repeated exposure to the same stimulus reduces the response of the animal, triggering an increase when, finally, a different stimulus is introduced (see Fig. 6.3 in Chapter 6).

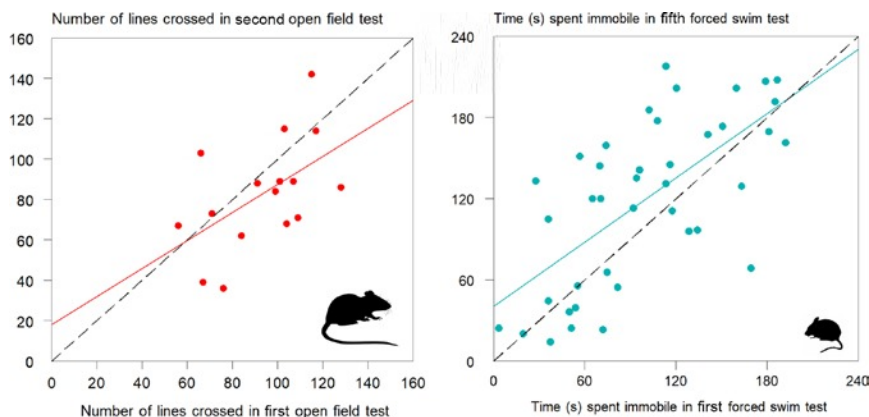


Fig. 11.7. Re-testing animals in behavioural tests: (a) activity of the same rats in two open field tests performed 17 days apart (data from Bombail *et al.*, 2018); (b) immobility of mice in a forced swim test for the first and fifth exposure to the test (data from Kazavchinsky *et al.*, 2019). In both graphs, each point is an animal, the dashed black line indicates equivalence between the two tests ($y=x$) and the solid line is the best fit regression line of the data.

As an example of re-testing, I dug out some raw data from a recent experiment (Bombail *et al.*, 2018), where we had tested 6-month-old male Wistar rats in a 5-min open field test twice with a 17-day interval. In-between these tests, the rats had been exposed individually to various odours and their behaviour recorded. The results from the two open field tests are shown in Fig. 11.7a. There was a general reduction in activity during the second test of 70% (slope of the regressions line), but the correlation of 0.54 between the two tests was significant. Another example is from Kazavchinsky *et al.* (2019), who tested female and male mice five times in a forced swim test, each separated by 2 days. Duration of immobility varied among mice, but stayed relatively constant within individuals across the tests, with an overall increase in immobility from 99 ± 8 s during the first test to 119 ± 10 s in the fifth test. The slope of the regression line in Fig. 11.7b is 0.79 and the correlation of 0.67 is significant. The validity of the forced swim test is discussed in Chapter 7.

Males vs Females

An important factor to consider in behavioural testing is how to take the sex of the animal into account. Clearly, there is a need to use only males for some tests and observations, like studies of aggression in group-housed bulls, just like mature females are needed to study maternal behaviour in any species. However, for many behavioural tests, sex is not an influencing factor. Nevertheless, females are often excluded because they are thought

to show more individual variation than males due to hormonal fluctuations. But males have hormones too. In a recent article in the journal *Science*, Rebecca Shansky (2019) notes (and laments) that hormonal differences in male rodents arising from different positions in the dominance hierarchy of group-housed males are treated as random variation and most often not considered in the experimental design, whereas females are often excluded from study based on their having an oestrous cycle. Meta-analyses of 293 articles on mice (Prendergast *et al.*, 2014) and 6000 data points from rats (Becker *et al.*, 2016) found no differences in the variability between males and females. Shansky (2019) takes the argument further because not only are females predominantly excluded in neuroscience and biomedical research, but when they *are* tested it is to a large extent after an effect has been found in males, thus treating the male brain and biology as the standard from which females may or may not deviate.

It is not that no behavioural differences between the sexes exist. They do. Male goat kids have been found to perform better than females at different stages of an object permanence test (Vas *et al.*, 2019). Studying aged (1-year-old) rats, Domonkos *et al.* (2017) found less anxious behaviour displayed by females in a number of behavioural tests, including the open field and the elevated plus maze. Booher *et al.* (2019) wanted to see if access to exercise in the form of a running wheel would reduce voluntary alcohol consumption in mice. When the authors checked the baseline differences between males and females of the strain investigated (129/SvEvTac) for these two variables, they found that female mice consumed more alcohol and ran further than their male conspecifics, leading to different results for the two sexes (Fig. 11.8).

In passive avoidance tests (see Chapter 9), female rats are much more likely to re-enter the box in which they have previously received an electric shock than are male rats, even when the shock intensity was high (Van Oyen *et al.*, 1979). This was originally thought to be caused by females being less able to remember that they had received the shock. However, by offering the rats three instead of two options this was found not to be the case (van Haaren and van de Poll, 1984). When the rats had the choice between staying in the light compartment, or entering either of two dark compartments, in one of which the rat had previously been shocked, males and females differed in their response. Almost half of the male rats tested (8 out of 17; 47%) remained in the light compartment, despite being offered an alternative dark compartment where no shock had been experienced, which was the choice of the nine other males. In contrast, only 2 out of 17 females (12%) did not leave the light compartment, and of the 15 that entered a dark compartment, only one rat chose the one in which they had been previously shocked. The authors conclude that female rats appear to find the light compartment more aversive than do male rats.

Sometimes the choice of sex of your experimental subjects is more pragmatic: one of my colleagues always recommends using male pigs for

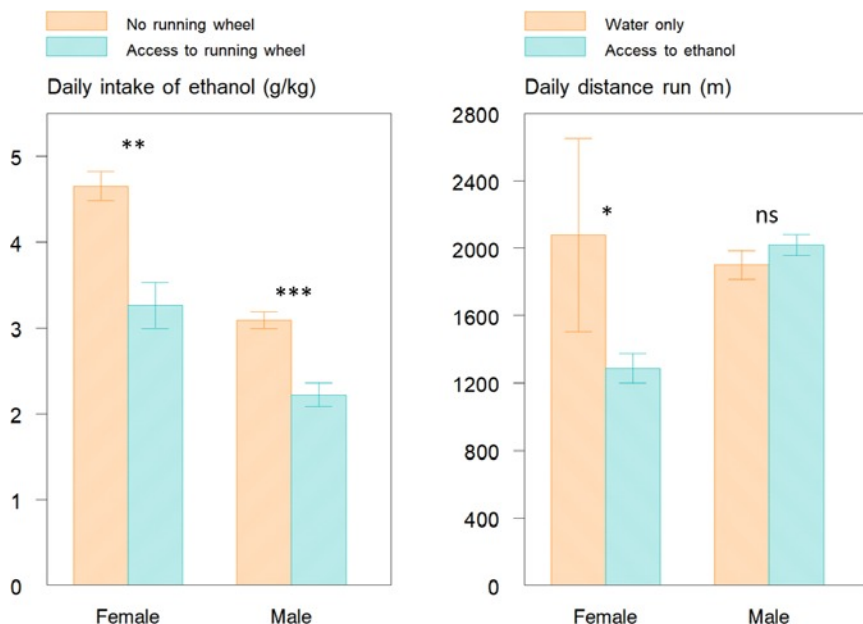


Fig. 11.8. Difference between sexes in alcohol consumption and daily running activity in mice. (a) Compared with males, female mice have a higher intake of alcohol, and this is reduced in both sexes when given access to a running wheel. (b) Only females showed a reduction in running when given access to alcohol as well as water (from Booher *et al.*, 2019).

feeding trials for the simple reason that they are less likely than females to pee in the feed trough. It is also worth noting that the words *gender* and *sex* used to be interchangeable. However, the former is now associated with (human) identity, and gender should no longer be applied to the biological sex differences seen in animals, including when referring to castrated males, such as steers and geldings (castrated bulls and stallions, respectively).

Concluding Remarks

This chapter describes only a handful of the factors to consider when planning and executing behavioural tests for a scientific purpose. It quickly becomes clear that the experimental planning is pivotal for a successful outcome, whether it involves simple tests of choice or intensive training over long periods. Some of the examples missing here are scattered in the previous chapters. In the final two chapters of this book, I will try to tackle issues relating to ethical considerations as well as current and future technological advances in the field of animal behaviour testing.

References

- Becker, J.B., Prendergast, B.J. and Liang, J.W. (2016) Female rats are not more variable than male rats: a meta-analysis of neuroscience studies. *Biology of Sex Differences* 7(1), 34. DOI: 10.1186/s13293-016-0087-5.
- Bombail, V., Barret, B., Raynaud, A., Jérôme, N., Saint-Albin, A. *et al.* (2018) In search of stress odours across species: behavioural responses of rats to faeces from chickens and rats subjected to various types of stressful events. *Applied Animal Behaviour Science* 205, 216–226. DOI: 10.1016/j.applanim.2017.10.013.
- Booher, W.C., Hoft, N.R. and Ehringer, M.A. (2019) The effect of voluntary wheel running on 129/SvEvTac and C3H/Ibg alcohol consumption. *Alcohol* 77, 91–99. DOI: 10.1016/j.alcohol.2018.10.007.
- Bradley, J.V. (1958) Complete counterbalancing of immediate sequential effects in a Latin square design. *Journal of the American Statistical Association* 53(282), 525–528. DOI: 10.1080/01621459.1958.10501456.
- Burley, N. (1988) Wild zebra finches have band-colour preferences. *Animal Behaviour* 36(4), 1235–1237. DOI: 10.1016/S0003-3472(88)80085-X.
- Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J. *et al.* (2013) Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience* 14(5), 365–376. DOI: 10.1038/nrn3475.
- Chamaillé-Jammes, S., Malcuit, H., Le Saout, S. and Martin, J.-L. (2014) Innate threat-sensitive foraging: black-tailed deer remain more fearful of wolf than of the less dangerous black bear even after 100 years of wolf absence. *Oecologia* 174(4), 1151–1158. DOI: 10.1007/s00442-013-2843-0.
- Cooper, J.J. and Mason, G.J. (2000) Increasing costs of access to resources cause re-scheduling of behaviour in American mink (*Mustela vison*): implications for the assessment of behavioural priorities. *Applied Animal Behaviour Science* 66(1–2), 135–151. DOI: 10.1016/S0168-1591(99)00069-6.
- Crabbe, J.C., Wahlsten, D. and Dudek, B.C. (1999) Genetics of mouse behavior: interactions with laboratory environment. *Science* 284(5420), 1670–1672. DOI: 10.1126/science.284.5420.1670.
- Domonkos, E., Borbélyová, V., Csongová, M., Bosý, M., Kačmárová, M. *et al.* (2017) Sex differences and sex hormones in anxiety-like behavior of aging rats. *Hormones and Behavior* 93, 159–165. DOI: 10.1016/j.yhbeh.2017.05.019.
- Fischer, B. and Wegener, D. (2018) Emphasizing the ‘positive’ in positive reinforcement: using nonbinary rewarding for training monkeys on cognitive tasks. *Journal of Neurophysiology* 120(1), 115–128. DOI: 10.1152/jn.00572.2017.
- Garamszegi, L.Z. (2016) A simple statistical guide for the analysis of behaviour when data are constrained due to practical or ethical reasons. *Animal Behaviour* 120, 223–234. DOI: 10.1016/j.anbehav.2015.11.009.
- Herskin, M.S. and Jensen, K.H. (2002) Effects of open field testing and associated handling *v.* handling alone on the adrenocortical reactivity of piglets around weaning. *Animal Science* 74(3), 485–491. DOI: 10.1017/S1357729800052644.
- Kazavchinsky, L., Dafna, A. and Einat, H. (2019) Individual variability in female and male mice in a test-retest protocol of the forced swim test. *Journal of Pharmacological and Toxicological Methods* 95, 12–15. DOI: 10.1016/j.vascn.2018.11.007.
- Lahti, D.C. (2015) The limits of artificial stimuli in behavioral research: the umwelt gamble. *Ethology* 121(6), 529–537. DOI: 10.1111/eth.12361.

- Lorenz, K. (1939) Vergleichende verhaltensforschung [Comparative behavioral research]. *Verhandlungen der Deutschen Zoologischen Gesellschaft Zoologischer Anzeiger, Supplementband* 12, 69–102.
- Nielsen, B.L. (2018) Making sense of it all: the importance of taking into account the sensory abilities of animals in their housing and management. *Applied Animal Behaviour Science* 205, 175–180. DOI: 10.1016/j.applanim.2018.04.013.
- Nielsen, B.L., Erhard, H.W., Friggens, N.C. and McLeod, J.E. (2008) Ultradian activity rhythms in large groups of newly hatched chicks (*Gallus gallus domesticus*). *Behavioural Processes* 78(3), 408–415. DOI: 10.1016/j.beproc.2008.02.010.
- Nielsen, B.L., Jezierski, T., Bolhuis, J.E., Amo, L., Rosell, F. *et al.* (2015) Olfaction: an overlooked sensory modality in applied ethology and animal welfare. *Frontiers in Veterinary Science* 2(Suppl. 1), 69. DOI: 10.3389/fvets.2015.00069.
- Prendergast, B.J., Onishi, K.G. and Zucker, I. (2014) Female mice liberated for inclusion in neuroscience and biomedical research. *Neuroscience & Biobehavioral Reviews* 40, 1–5. DOI: 10.1016/j.neubiorev.2014.01.001.
- Raundal, P.M., Andersen, P.H., Toft, N., Herskin, M.S., Forkman, B. *et al.* (2015) Pre-test habituation improves the reliability of a handheld test of mechanical nociceptive threshold in dairy cows. *Research in Veterinary Science* 102, 189–195. DOI: 10.1016/j.rvsc.2015.07.020.
- Ravilious, K. (2017) Find the flow. *New Scientist* 235(3132), 32–35. DOI: 10.1016/S0262-4079(17)31269-1.
- Richter, S.H., Garner, J.P., Auer, C., Kunert, J. and Würbel, H. (2010) Systematic variation improves reproducibility of animal experiments. *Nature Methods* 7(3), 167–168. DOI: 10.1038/nmeth0310-167.
- Schaefer, A.T. and Claridge-Chang, A. (2012) The surveillance state of behavioral automation. *Current Opinion in Neurobiology* 22(1), 170–176. DOI: 10.1016/j.conb.2011.11.004.
- Schleidt, W.M. (1961) Über die Auslösung der Flucht vor Raubvögeln bei Truthühnern [On the triggering of flight from birds of prey in turkeys]. *Die Naturwissenschaften* 48, 141–142.
- Schleidt, W., Shalter, M.D. and Moura-Neto, H. (2011) The hawk/goose story: the classical ethological experiments of Lorenz and Tinbergen, revisited. *Journal of Comparative Psychology* 125(2), 121–133. DOI: 10.1037/a0022068.
- Shalter, M.D. (1975) Lack of spatial generalization in habituation tests of fowl. *Journal of Comparative and Physiological Psychology* 89(3), 258–262. DOI: 10.1037/h0076817.
- Shansky, R.M. (2019) Are hormones a ‘female problem’ for animal research? *Science* 364(6443), 825–826. DOI: 10.1126/science.aaw7570.
- Tinbergen, N. (1939) Why do birds behave as they do? (II) *Bird-Lore* 41, 23–30.
- van Haaren, F. and van de Poll, N.E. (1984) The effect of a choice alternative on sex differences in passive avoidance behavior. *Physiology & Behavior* 32(2), 211–215. DOI: 10.1016/0031-9384(84)90131-8.
- Van Oyen, H.G., Van de Poll, N.E. and De Bruin, J.P.C. (1979) Sex, age and shock-intensity as factors in passive avoidance. *Physiology & Behavior* 23(5), 915–918. DOI: 10.1016/0031-9384(79)90200-2.
- Vas, J., Chojnacki, R.M. and Andersen, I.L. (2019) Search behavior in goat (*Capra hircus*) kids from mothers kept at different animal densities throughout pregnancy. *Frontiers in Veterinary Science* 6, 21. DOI: 10.3389/fvets.2019.00021.

- Winter, Y. and Schaefer, A.T.U. (2011) A sorting system with automated gates permits individual operant experiments with mice from a social home cage. *Journal of Neuroscience Methods* 196(2), 276–280. DOI: 10.1016/j.jneumeth.2011.01.017.
- Yang, M. and Crawley, J.N. (2009) Simple behavioral assessment of mouse olfaction. *Current Protocols in Neuroscience* 48, 8.24.1–8.24.8.

Legislation, Guidelines and Ethical Considerations

As stated in Chapter 1, any book on animal behaviour tests is bound to be incomplete. I have tried to ensure that the tests featured are those that attempt to ask the animal a question by placing it in a (test) situation, allowing the researcher to interpret the behaviour observed and measured. But even when a test tries to mimic a naturally occurring situation without any form of sampling of bodily fluids from the test subject, it is nevertheless using live animals for scientific purposes. And this has raised – and continues to raise – concerns among the public, and is an issue that cannot go unmentioned in a book such as this. In the following sections, I will attempt to make a pithy summary of the ethical considerations associated with the scientific use of animals in behavioural tests, starting with the legislation put in place to ensure that animals are treated in an ethically defensible manner. It is unavoidable that this in places will be coloured by my personal opinions on the subject. I am clearly not against animal testing, but there are tests that we should no longer use. You may not agree with some (or all) of my points of view, and there may be issues I have not taken into consideration. That said, I hope the topics raised here can initiate thoughts and discussions to bring us forward towards acceptable solutions instead of polarizing the debate. This chapter is an attempt to provide a common basis from which to start this dialogue.

Legislation and Guidelines

Within Europe, all use of animals for scientific and educational purposes is subject to the EU Directive (2010). This document stipulates rules and limits for experiments carried out on non-human vertebrate animals and cephalopods (e.g. octopus, squid and cuttlefish). It also covers foetal forms of mammals in the last trimester of normal development, but does not include birds still in the egg. One statement in the preamble (Section 10) declares that ‘this Directive represents an important step towards achieving the final goal of full replacement of procedures on live

animals for scientific and educational purposes as soon as it is scientifically possible to do so' (EU Directive, 2010, p. 34). This may come as a surprise to many European animal scientists, and I fail to see how we can investigate certain aspects of animal behaviour of relevance for the welfare of a given species without asking the animal through behavioural testing. The aim of this statement is, of course, to speed up the replacement of animals in neuroscience and pharmacological research, where procedures may cause discomfort, harm and pain. Indeed, the EU Directive embraces the replacement, reduction and refinement of animal use in procedures, also known as the 3Rs and originally developed with humans in mind by Russell and Burch (1959). The reduction refers to a decrease in number of animals, so as to obtain the same information from fewer animals, and using only the minimum required to obtain statistically significant results. Refinement applies to both the breeding and the procedures used in animal science to alleviate or minimize potential pain, suffering and distress. An excellent source for guidelines, information and demonstration of the 3Rs in principle and in practice is the National Centre for the Replacement, Refinement and Reduction of Animals in Research (www.nc3rs.org.uk).

In the USA, the regulation of the use of animals in research is somewhat complex, and littered with acronyms. The main regulation is a federal law called the Animal Welfare Act. It is administered by the United States Department of Agriculture (USDA) Animal Care unit within the Animal and Plant Health Inspection Service (APHIS). However, not all animals are covered by the Animal Welfare Act; most notable is the exclusion of mice and rats (Table 12.1). These species are protected under the Public Health Service Policy on Humane Care and Use of Laboratory Animals (PHS Policy), though only in research facilities that receive federal funding. Scientists must comply with the Guide for the Care and Use of Laboratory Animals, and each research facility must have an Institutional Animal Care and Use Committee (IACUC; Table 12.1).

Complementing the existing legislation are guidelines for use of animals in research, and these continue to be published and improved, including the ARRIVE guidelines (Kilkenny *et al.*, 2010). One such publication, which is freely available online, is the Guiding Principles for Behavioural Laboratory Animal Science (2013) published by a group of scientific associations in animal science. It is heavily rodent oriented, focusing on ethical assessment and compliance with the 3Rs. Another example is the ethical guidelines developed by the International Society for Applied Ethology (ISAE), for which an updated (2017) version of the original published article (Sherwin *et al.*, 2003) can be found online. These documents also contain lists of international and national legislation and regulations regarding animal use and procurement of animals.

Table 12.1. Summary of regulations, guides and oversight bodies for the use of animals in research in the USA. Acronyms used are explained elsewhere in the table.

Regulation and oversight body	Short description
Animal Welfare Act (AWA)	Protects all warm-blooded animals <i>except</i> rats, mice, birds bred for research, farm animals used for food or fibre (fur, hide, etc.), cold-blooded species (amphibians and reptiles), horses not used for research purposes, fish and invertebrates (crustaceans, insects, etc.) Included are zoos, circuses, research labs, hospitals, businesses, federal agencies, dealers, breeders, etc.
Institutional Animal Care and Use Committee (IACUC)	A committee, organized at every research facility subject to the AWA, PHS Policy or AAALAC accreditation, which must review and approve or reject every proposed animal protocol IACUC members must inspect their research facility twice a year
United States Department of Agriculture (USDA) Animal Care unit	The USDA licenses research facilities and conducts annual, unannounced inspections Violations are punished with fines, cease-and-desist orders, and licence suspension or revocation
The Guide for the Care and Use of Laboratory Animals (the <i>Guide</i>)	Guide published by the National Research Council and the Institute for Laboratory Animal Research The central part of Public Health Service Policy on the humane care and use of laboratory animals and basis for AAALAC International accreditation
The Public Health Service (PHS) Policy	Protects all vertebrate animals (including fish, reptiles, rats, mice and birds) used in research funded by the PHS Violations or loss of AAALAC accreditation can result in loss of PHS funding
Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International	Non-profit organization that accredits research facilities on a voluntary basis for compliance with the <i>Guide</i> Announced site visits are conducted every 3 years

Ethical Considerations

What should we avoid?

In Chapter 8, the use of punishment as well as negative reinforcement to train animals were described, and it is clear that some behavioural tests involve aversive stimuli. These do not necessarily involve painful treatment,

as aversive stimuli are simply those that the animal wants to avoid, such as an unpleasant smell. Positive reinforcement, where the animal is rewarded for showing the desired behaviour, should be used as much as possible. Of the four categories presented in Fig. 8.1, those involving attractive stimuli are to be preferred, even when this involves removing or withholding the reward as is the case with negative (subtraction) punishment. Remember that negative reinforcement is when an aversive stimulus is removed as soon as the desired behaviour is shown. This method often results in fast learning provided that the aversive stimulus is not too severe, and can be employed together with positive reinforcement to improve efficiency of the latter. An example of this is the gentle and gradually increasing pressure placed on the flank of a horse to make it move, with the pressure completely removed as soon as the horse takes a step sideways, leading to less and less force necessary to make this happen (Fig. 8.2). The training method to be avoided is positive (addition) punishment, where an aversive stimulus is applied to reduce the frequency of a behaviour. For this to be effective, the aversive stimulus has to be moderate to severe. Unfortunately, this often involves applying electric shocks, which are highly aversive to most animals, including humans. Application of this punishment needs to be immediately contingent upon the behaviour being shown, such as when livestock touches an electric fence. As the animals usually learn from a single shock not to touch the fence again, this is one of the special cases where positive punishment can justifiably be used. However, this is not the case when shock collars are used for training dogs or as virtual fences (see Chapter 8 for a discussion of the negative welfare impact of these collars).

Tests exist that in my opinion should not be used given our current knowledge on animal welfare and the availability of alternatives. One such test is the Vogel conflict test, which is a test of anxiety involving the application of electric shocks when the animal attempts to feed or drink. Obviously, this reduces the feeding and drinking behaviour of the animal, whereas administration of anxiolytic drugs can make the animal approach the feeder and drinker although it will still receive shocks. When searching through Web of Science, I was somewhat disappointed to find that more than seven publications per year have been using this method in the past 25 years. The Geller Seifter test is a slightly milder variation of the same concept, as the animal can choose to feed (albeit infrequently) without being shocked.

In Chapter 5, examples were given of tests where the animal's willingness to pay for a given resource was assessed by increasing the number of lever presses necessary or increasing the weight of the door to be pushed open to gain access. Walker and Mason (2018) recently suggested that mildly electrified grids could be used to impose a cost on access to resources in mice. The authors note that as long as the intensity of the current is low ($\leq 0.6\text{mA}$ in their example) and the mouse has access to all necessary resources (water, food, bedding and nesting material) on both sides of the

grid, the set-up does not force the animal to ever cross the grid. Walker and Mason (2018) argue that using mild electric shocks in this manner can be used to estimate the motivation of mice to gain access to enrichment.

Some tests include feed or water deprivation prior to testing in order to motivate the animal to obtain feed or water as rewards for correct behavioural responses (see Chapters 5 and 8). Efforts are made to improve the testing conditions to ensure better animal welfare. One such development is the use of water rewards to train rats, but without restricting access to water. Reinagel (2018) found that by slightly acidifying the ad libitum water source in the home pen by adding 2% citric acid, rats would still perform trials to obtain pure water but without being water deprived to do so. The rats performed trials at a level of 74% of those performed when all water had to be obtained in this way. It should be noted that this study was carried out with female rats, which is a nice break with the norm of preferring male test subjects.

If feed restriction is necessary to perform a given test, it is important to keep this to a minimum, i.e. only fast the animals for the shortest possible duration for the test to work. For food rewards to be attractive, they need only to be highly palatable titbits that the animal finds appetitive with minimal, if any, fasting depending on the species. As long as these rewards are given in very small quantities, their desirability can remain high enough to function as reinforcement for a period of time (see also Chapter 8). A number of scientific papers on rats have been published where feed restriction of 48 h has been used, and weight loss is detectable throughout this period of fasting. However, it has been known for at least a decade that fasting-induced changes in metabolic function and most blood parameters occur within or at 16 h of feed deprivation, with only very small subsequent changes (Fig. 12.1; Kale *et al.*, 2009). The use of prolonged fasting in scientific studies of these rodents should be of durations not exceeding 16 h unless it can be shown that longer deprivation is necessary – and justifiable – for the scientific question asked.

At the risk of stating the obvious, it is important to distinguish between behavioural tests and behavioural treatments. Behavioural tests of animals, which this book is about, are ways of asking the animal questions. In order to get a sensible answer, the animal needs to be in a suitable state and, as far as possible, it should be ensured that the animal is not stressed. Handling our animals regularly, and keeping them in suitable housing, as stipulated by the legislation (EU Directive, 2010) contributes to the validity of the results we want to obtain. Behavioural treatments, on the other hand, are applied in order to change the psychological state of the animal. When enrichment is applied, it is supposed to improve the affective state of the animal, but far too often it consists of merely adding a few stimuli to an otherwise barren environment, and the term enrichment, although relative, appears inappropriate (Newberry, 1995; Olsson and Dahlborn, 2002; van de Weerd and Day, 2009). Frequently, behavioural treatments are applied

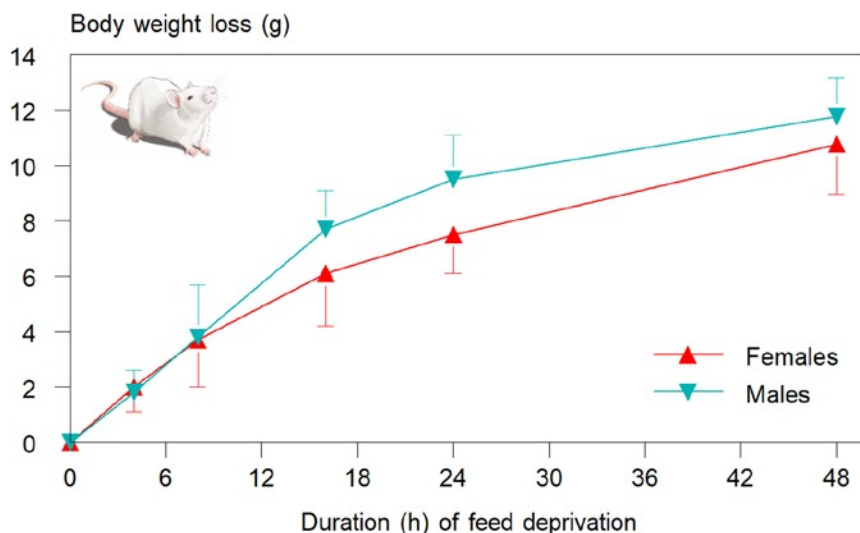


Fig. 12.1. Body weight loss in female and male Wistar rats as a function of duration of feed restriction. Changes in most blood constituents are evident and stabilized at 16 h of fasting (data from Kale *et al.*, 2009).

to induce negative affective states in the animals in order to test pharmaceutical remedies. These may include severe behavioural restrictions to induce stress, anxiety or depression, such as chronic social defeat (e.g. Browne *et al.*, 2018), prolonged deprivation of REM (rapid eye movement) sleep (Jouvet *et al.*, 1964; Machado *et al.*, 2004) and chronic unpredictable stress (Descalzi *et al.*, 2017). The latter consists of exposing mice or rats to mildly unpleasant environmental conditions at frequent but unpredictable intervals, such as tilting the cage, removing or wetting the bedding, temporary removal of food, unexpected light exposure and short-term restraint. It is important to keep in mind that this is not what is meant by animal behaviour testing.

Animal welfare – are we talking about the same thing?

Many definitions of animal welfare have been proposed over time, and this is not the place for a thorough discussion of their respective merits and shortcomings. It is worth mentioning that the importance of an animal's affective state has been increasingly acknowledged over time, and is now fundamental in most of the commonly used animal welfare definitions (Fraser and Duncan, 1998; Dawkins, 2006). I would like to present, however, one of the latest attempts to define animal welfare. This work was carried out by a working group in the French Agency for Food, Environmental and Occupational Health and Safety (Anses), and it is further presented and discussed in an article by Mormede *et al.* (2018):

The welfare of an animal is the positive mental and physical state linked to the satisfaction of its physiological and behavioural needs, as well as to its expectations. This state varies according to the perception of the situation by the animal.

(Anses, 2018, p. 16)

It was met with some controversy in France when first published, mainly to do with the difficulties of applying it in practice. However, what I like about it is that it not only combines physiological and psychological dimensions of welfare, it also includes the dynamic aspect of time as well as the expectations of the animal. To me, this encapsulates the notion that animal welfare is not a static state and differs among individuals, dependent on their experiences, and within individuals depending on the prevailing situation. Housing a group of pigs with *ad libitum* food and water in an appropriate environmental temperature may be sufficient in the short term, but if this situation never changes, these animals (that will spend the majority of their time foraging and exploring when kept in free range conditions) are likely to lack environmental stimulation, leading to an increased risk of ear- and tail-biting behaviour developing. So even if the definition above comes across as being more convoluted than some other animal welfare definitions, it makes us remember the many facets encompassed by the concept of animal welfare: mental and physical state, experience of the animal and changes over time.

The more we include the feelings or emotional state of the animal in our definition of welfare, the more important it is to know what we mean by these different concepts. A great discussion of these can be found in a recent paper by Webb *et al.* (2019), who dare to ask ‘What is happiness?’. The authors include a figure, a version of which is reproduced in [Fig. 12.2](#), which is intuitively very clear, showing the relationship between the terms emotion, mood and affective state in a beautiful way.

It is relatively common to find newspaper articles or blogs that speak about animal welfare and animal ethics as if these were interchangeable concepts. They are not. Many people find it difficult to put their finger on the differences, and because welfare and ethics are interlinked it is not surprising to see the confusion between the two. Animal ethics are moral principles that govern human behaviour. Animal welfare science is the study of how animals ‘fare’, either as a consequence of their genetic makeup, under different housing and management systems or in specific situations, so as to assess these objectively, based on their impact on the welfare of the animals. Animal ethics should be (but is not always) based on animal welfare science. To give an example of the difference between animal ethics and welfare: a rat subjected to a chronic intermittent and unpredictable stress protocol, as described earlier, will have its welfare severely and negatively impacted. Some may argue that the application of this treatment can be justified by the goal, which is to develop effective treatments for stress and anxiety in humans. The latter is the ethical judgement we need to make

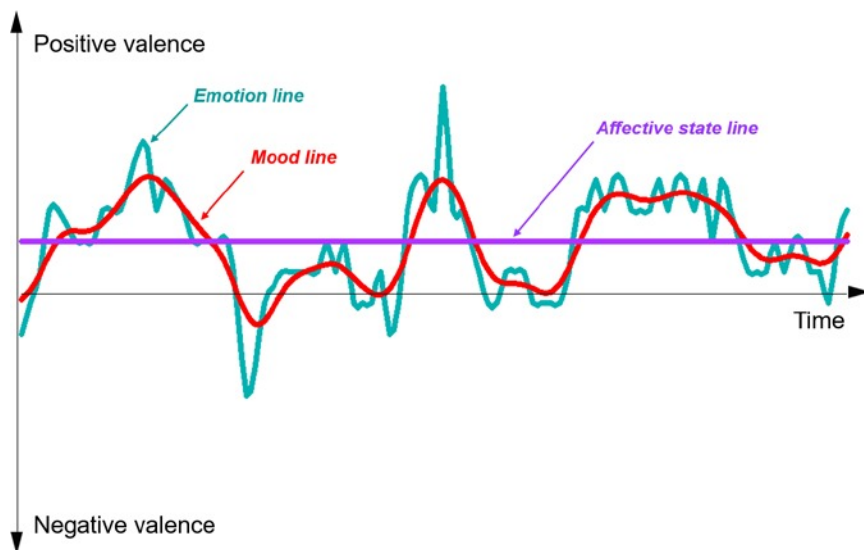


Fig. 12.2. Schematic representation of the differences between emotion (in cyan), mood (in red), and affective state (in violet) over time. Emotions and moods are transient affective experiences, whereas affective state is stable under nearly stable conditions (adapted from Webb *et al.*, 2019).

before applying aversive treatments to our animals. However, the rat itself does not care about ethics – if it receives the aversive treatment, no matter for what reason, its welfare is compromised.

As an animal behaviour scientist, I should endeavour in my research to use only the number of animals necessary, and perform behavioural tests that have positive or at least neutral effects on the welfare of my subjects of study, and that are not aversive. If I do use aversive tests or treatments, these should be of the lowest severity and shortest duration possible. And one could argue that I have a moral obligation to publish my results so that it is not necessary for others to repeat the same experiments. But the latter does not impact on the welfare of the animals I have already tested.

Ethical considerations also apply when developing genetically modified lines of animals, most commonly rats and mice, to create models for various human diseases and mental conditions. Many of these modifications are changes to immune capacity or physiological aspects, but some alter the behaviour of the animal in significant ways. An example of this are rats bred for helplessness, used as models for depression. This phenotypic expression of being powerless in the face of adversity has in the past been a state that animals could develop if exposed to prolonged and inescapable harsh conditions, so-called learned helplessness. In the equivalent genetic model, rats are bred to display these phenotypical traits. Schulz *et al.* (2016) found that these rats were impaired compared with wild-type rats when

learning a positive reinforcement schedule. Although an anti-depressant (deprenyl) reduced immobility of these rats when subjected to a forced swim test, the medication did not improve their ability to learn the positive reinforcement schedule. This challenges the usefulness of such an animal model, and emphasizes the questionable efficacy of the forced swim test as a measure for effectiveness of anxiolytic drugs, as discussed in more detail in Chapter 7.

In Chapter 3, I mentioned briefly that our choice of species for a behavioural test depends, not surprisingly, on what we are trying to achieve. Are we trying to increase the knowledge of the behavioural biology of the species tested, or do we want to know more about behavioural mechanisms in general? Rats and mice are frequently used as models for human applications, e.g. in pharmacological testing, but one could argue that rodents may not be the optimal animal model for *Homo sapiens*. Often we work with the animals that are available, so dog owners will train their dogs, and scientific laboratories with animal housing often have only one or two species available, mostly rodents and zebrafish judging by the numbers used in scientific studies. A prominent critic of some of the current animal models is Dr Joseph Garner, who asks the question why scientific results achieved with rodents translate so poorly into the equivalent human conditions. Together with Professor Hanno Würbel, he has been a proponent of introducing more heterogeneity into our animal models – named ‘the standardization fallacy’ – to mimic better real-life conditions in the human world (Richter *et al.*, 2009). Garner has questioned the validity of genetic animal models of conditions such as autism and Tourette’s, as the changes giving rise to the tics in the Tourette’s syndrome mouse model are not similar to those found in the brains in human Tourette’s syndrome cases (Rutkin, 2016). His laboratory also hosts the website mousebehavior.org, which describes various methods and protocols, including ethograms, for observing mouse behaviour in their home cages.

Concluding Remarks

It is important to keep in mind that scientists are under strict regulations in terms of what they can do to animals. The EU Directive (2010) for example, has detailed standards for the housing and management of a large variety of species, including not only mice, rats, rabbits, cats and dogs, but also ferrets, farm animals, birds, amphibians, reptiles and fish. Special restrictions for use are in place for non-human primates, such as marmosets, macaques and baboons. Experiments using animal models need to follow scientific protocols that adhere to the current legislation on the use of animals for scientific purposes and have been approved by ethical committees. Thus, some of the studies found in the older literature would not be permitted today, including raising kittens in darkness to investigate links

between movement and visually guided behaviour (Held and Hein, 1963), and tests on rats where the endpoint was drowning (Richter, 1957). The EU Directive (2010) contains a statement (p. 33), that it is ‘necessary to improve the welfare of animals used in scientific procedures by raising the minimum standards for their protection in line with the latest scientific developments’. It also states that animals have intrinsic value, and that they should always be treated as sentient creatures. The vast majority of behavioural tests for use in animals are non-aversive, and many tests even appear to improve the affective state of the animals; in other words provide the test animal with a pleasurable, even joyful experience (e.g. Cloutier *et al.*, 2018). During interactions with animals in our care, we can all do our best to improve their welfare, and to apply training and test methods that confer a maximum of contentment, restricting aversive treatment to the bare minimum. This applies equally to scientific studies, to animal trainers and pet owners, as well as zoo-keepers and farmers.

References

- Anses (2018) Avis de l'Anses relatif au 'Bien-etre animal : contexte, definition et evaluation' [Anses opinion on animal welfare: context, definition and evaluation]. Available at: www.anses.fr/fr/system/files/SABA2016SA0288.pdf (accessed 8 July 2019).
- Browne, C.A., Falcon, E., Robinson, S.A., Berton, O. and Lucki, I. (2018) Reversal of stress-induced social interaction deficits by buprenorphine. *International Journal of Neuropsychopharmacology* 21(2), 164–174. DOI: 10.1093/ijnp/pyx079.
- Cloutier, S., LaFollette, M.R., Gaskill, B.N., Panksepp, J. and Newberry, R.C. (2018) Tickling, a technique for inducing positive affect when handling rats. *Journal of Visualized Experiments* 135(135), e57190. DOI: 10.3791/57190.
- Dawkins, M. (2006) A user's guide to animal welfare science. *Trends in Ecology & Evolution* 21(2), 77–82. DOI: 10.1016/j.tree.2005.10.017.
- Descalzi, G., Mitsi, V., Purushothaman, I., Gaspari, S., Avrampou, K. *et al.* (2017) Neuropathic pain promotes adaptive changes in gene expression in brain networks involved in stress and depression. *Science Signaling* 10(471), eaaj1549. DOI: 10.1126/scisignal.aaj1549.
- EU Directive (2010) Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Official Journal of the European Union* L276, 33–79.
- Fraser, D. and Duncan, I.J.H. (1998) 'Pleasures', 'pains' and animal welfare: toward a natural history of affect. *Animal Welfare* 7, 383–396.
- Guiding Principles for Behavioural Laboratory Animal Science (2013) Financed by Laboratory Animal Science Association (LASA), British Association for Psychopharmacology (BAP), British Neuroscience Association (BNA) and the European Courses in Whole Animal Pharmacology (ESSWAP) Foundation. Available at: www.lasa.co.uk/PDF/LASA_BAP_BNA_ESSWAP_GP_Behavioural_LAS_Nov13.pdf (accessed 12 June 2019).

- Held, R. and Hein, A. (1963) Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology* 56(5), 872–876. DOI: 10.1037/h0040546.
- ISAE (2017) Updated version of the *Guidelines for Ethical Treatment of Animals in Applied Animal Behaviour and Welfare Research*, developed by the International Society for Applied Ethology. Available at: www.applied-ethology.org/Ethical_Guidelines.html (accessed 9 July 2019).
- Jouvet, D., Vimont, P. and Delorme, F. (1964) Etude de la privation sélective de la phase paradoxale du sommeil chez le chat [Study of selective deprivation of the paradoxal phase of sleep in the cat]. *Journal of Physiology* 56(381) (meeting abstract).
- Kale, V.P., Joshi, G.S., Gohil, P.B. and Jain, M.R. (2009) Effect of fasting duration on clinical pathology results in Wistar rats. *Veterinary Clinical Pathology* 38(3), 361–366. DOI: 10.1111/j.1939-165X.2009.00143.x.
- Kilkenny, C., Browne, W.J., Cuthill, I.C., Emerson, M. and Altman, D.G. (2010) Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biology* 8(6), e1000412. DOI: 10.1371/journal.pbio.1000412.
- Machado, R.B., Hipólido, D.C., Benedito-Silva, A.A. and Tufik, S. (2004) Sleep deprivation induced by the modified multiple platform technique: quantification of sleep loss and recovery. *Brain Research* 1004(1–2), 45–51. DOI: 10.1016/j.brainres.2004.01.019.
- Mormede, P., Boisseau-Sowinski, L., Chiron, J., Diederich, C., Eddison, J. *et al.* (2018) Bien-être animal: contexte, définition, évaluation [Animal welfare: context, definition and evaluation]. *INRA Production Animale* 31, 145–162.
- Newberry, R.C. (1995) Environmental enrichment: increasing the biological relevance of captive environments. *Applied Animal Behaviour Science* 44(2–4), 229–243. DOI: 10.1016/0168-1591(95)00616-Z.
- Olsson, I.A.S. and Dahlborn, K. (2002) Improving housing conditions for laboratory mice: a review of ‘environmental enrichment’. *Laboratory Animals* 36(3), 243–270. DOI: 10.1258/002367702320162379.
- Reinagel, P. (2018) Training rats using water rewards without water restriction. *Frontiers in Behavioral Neuroscience* 12, 84. DOI: 10.3389/fnbeh.2018.00084.
- Richter, C.P. (1957) On the phenomenon of sudden death in animals and man. *Psychosomatic Medicine* 19(3), 191–198. DOI: 10.1097/00006842-195705000-00004.
- Richter, S.H., Garner, J.P. and Würbel, H. (2009) Environmental standardization: cure or cause of poor reproducibility in animal experiments? *Nature Methods* 6(4), 257–261. DOI: 10.1038/nmeth.1312.
- Russell, W.M.S. and Burch, R.L. (1959) *The Principles of Humane Experimental Technique*. Methuen, London, UK.
- Rutkin, A. (2016) The autistic mouse and other unlikely stories. Online version entitled: lab mice are sending us on a wild goose chase. *New Scientist* 232, 42–43.
- Schulz, D., Henn, F.A., Petri, D. and Huston, J.P. (2016) Rats bred for helplessness exhibit positive reinforcement learning deficits which are not alleviated by an antidepressant dose of the MAO-B inhibitor deprenyl. *Neuroscience* 329, 83–92. DOI: 10.1016/j.neuroscience.2016.04.049.

- Sherwin, C.M., Christiansen, S.B., Duncan, I.J.H., Erhard, H.W., Lay, D.C. *et al.* (2003) Guidelines for the ethical use of animals in applied animal behaviour research. *Applied Animal Behaviour Science* 81, 291–305.
- van de Weerd, H.A. and Day, J.E.L. (2009) A review of environmental enrichment for pigs housed in intensive housing systems. *Applied Animal Behaviour Science* 116(1), 1–20. DOI: 10.1016/j.applanim.2008.08.001.
- Walker, M. and Mason, G. (2018) Using mildly electrified grids to impose costs on resource access: a potential tool for assessing motivation in laboratory mice. *Applied Animal Behaviour Science* 198, 101–108. DOI: 10.1016/j.applanim.2017.09.013.
- Webb, L.E., Veenhoven, R., Harfeld, J.L. and Jensen, M.B. (2019) What is animal happiness? *Annals of the New York Academy of Sciences* 1438(1), 62–76. DOI: 10.1111/nyas.13983.

Future Methodologies and Technological Advances

When I did my PhD back in the early 1990s, the technology used to register animal behaviour was still very much in its infancy, and mostly limited to setting up video cameras above the home pen or test arena. In those days, each camera would be the size of a shoebox, and needed some form of encasing if it was to be used in a humid or dusty environment, such as on a farm. Fast forward to today where, with the advent of the smartphone and fitness trackers, anybody can readily and easily record data on movements, positions and speed. Video cameras in the form of webcams are small, robust and easily mounted almost anywhere, often with large capacity for data storage. Indeed, camera footage is old-school, compared with the automatic tracking devices available on the market today. A lot of the technological advances have been driven by other areas of life, such as professional sports, where no football match is complete without the statistics on the distance run and the number of passes by each player (Fig. 13.1). These data can be collected from loggers sewn into the players' clothes or shoes, but real-time automatic video tracking is also used. Identification of human individuals can now be done by passing your hand once over a sensor that scans your fingers and checks against pre-registered images of your finger prints (MorphoWave Compact; Idemia.com).

When it comes to tracking animals, we are constrained by the size of the animal in terms of what sort of tracking device we can use. For large animals in an open or outdoor environment, the use of Bluetooth and global positioning system (GPS) equipment is possible, because the data collection unit can be attached to the animal's collar or leg without affecting the movements of the animal being monitored. Examples of these are accelerometers, where any change in speed in all three dimensions is logged. The resulting pattern of data can be used to decipher different behavioural elements, such as walking, resting and feeding (e.g. Graf *et al.*, 2015; Zobel *et al.*, 2015). It is interesting to note that although accelerometers can give us fast and automatic access to this type of data, the same information has been obtainable for a long time. An example of this is the

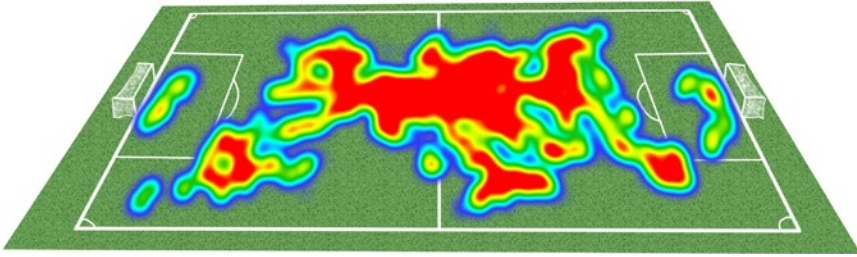


Fig. 13.1. Example of a density map from a tracking system, which can deliver information in real time on the positioning of a player in team sports such as football (image adapted from STATS SportVu).

monitoring of rumination. Attaching an accelerometer to the ear tag of a goat allows us to obtain data on the rumination patterns of these animals, clearly seen as 1 min bouts of chewing the bolus (i.e. the mouthful of regurgitated roughage) interspersed with short pauses as the bolus is swallowed and another brought up (Fig. 13.2a). However, this pattern was already registered automatically more than 60 years ago, when Bell and Lawn fitted goats with a rubber air cushion beneath the lower jaw, which allowed variations in air pressure, caused by the jaw movements, to be recorded. This was done via a flexible rubber tube leading from the cushion over a pulley to four ink pens writing on a moving paper roll, and with an airtight swivel

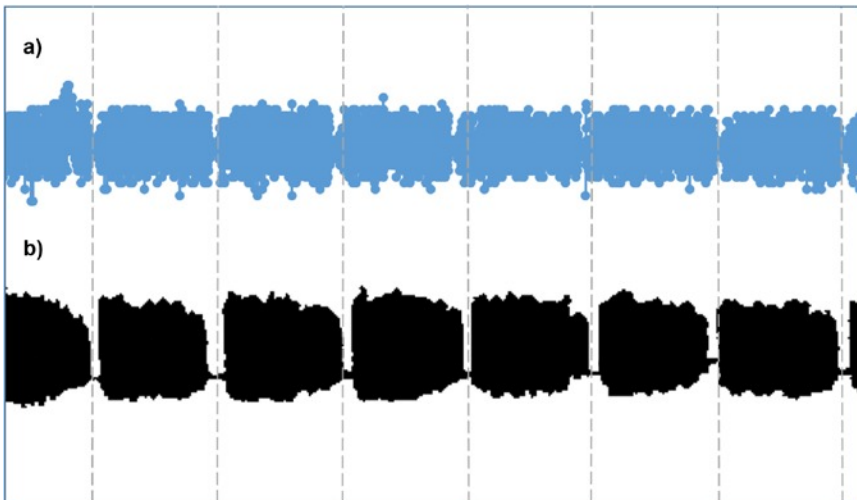


Fig. 13.2. Measurement of rumination in resting goats as a function of time, with vertical dashed lines indicating each minute. (a) Trace from a modern accelerometer fitted to the ear tag of a goat (Nielsen *et al.*, unpublished data from 2017) (b) Trace from jaw movements of a goat traced by ink pens writing on a moving paper roll (data from Bell and Lawn, 1957).

joint preventing the tubes from twisting and allowing free movement of the animal (Bell and Lawn, 1957; Fig. 13.2b). And okay, I may be stretching the meaning of the word ‘automatic’ in this example. Although this method only gave the experimenters reams of paper with the patterns on in contrast to the digitized recordings of the accelerometers, I still found the juxtaposition of the two traces thought-provoking. Indeed, as I write this, the use of accelerometer data still requires some form of calibration to ensure that your recordings truly reflect the behaviour of the animal. Commercial versions of accelerometers exist, e.g. IceTag (www.icerobotics.com), where data interpretation is built into the system, and these devices are available for large farm animals above the size of a goat.

There is no doubt that the progress made by the advancements in digital technology has contributed greatly to our understanding of animal behaviour. We are now able to study movement at a level of detail unheard of just a few years back. By using lasers, for example, to measure the wing movements of a bird trained to fly between two perches (while wearing laser safety goggles!), researchers have challenged the models previously used to calculate lift generated by flapping wings (Gutierrez *et al.*, 2017); information that is pivotal for the development of bioinspired designs used in flying equipment. Another example is the use of artificial intelligence (AI) to control animal-fitted loggers or cameras. Data collection, especially video recording, is often limited by the storage space of the logger or the battery life of the equipment. Some have tried to limit the collection of data to take place only at certain positions (e.g. a certain depth below sea level in seals; Volpov *et al.*, 2015) or when the animal is active (Brown *et al.*, 2012; Nishiumi *et al.*, 2018). Korpela *et al.* (2019) have gone one step further and created AI-assisted loggers that only record when specific types of behaviour occur. We can now log the movement of animals to within a few metres using GPS and satellite relay tags (Cox *et al.*, 2018), although some data quality assurance may still be needed.

Tracking of multiple animals in more confined settings and for extended periods can be achieved using ultra-wideband (UWB) tracking and radio frequency identification device (RFID) technology, allowing more accurate behavioural phenotyping of rodent disease models, and consequently better assessment of potential remedial actions (Howerton *et al.*, 2012). We also have the possibility of more detailed monitoring, including magnetic tracking of eye motion in small, fast-moving animals, such as fish (Plotkin *et al.*, 2008). Computer simulations of animal movements, especially in large groups, have been used to model and predict how certain environmental constraints may affect activity levels and social interactions (Lutnesky and Brown, 2015). Indeed, movement of four-legged creatures has been used to develop legged robots that move more predictably and appropriately for a given terrain (Hwangbo *et al.*, 2019).

Above is mentioned but a few examples of many of the technological advances in behavioural measurements, and the progress currently made

is faster than any book published on the subject can keep up with. For this reason, two other issues spring to mind, which may be more important to mention here than my trying to second-guess the next breakthrough in automatic tracking equipment. The first issue is that we can still achieve a lot, even with simple measures and inexpensive equipment. This is important not least for universities in the developing world, where funding to study animal behaviour may be lacking. Direct observations of an animal in a test arena are usually cheap to carry out, but they do have their limitations in terms of lack of data traceability and difficulties in verifying observer differences. However, with video recording available on any smartphone, and with computers and the Internet more readily accessible, simple types of tracking can now be done almost anywhere.

One example of this is a method used to quantify the way an animal moves in a test situation (Meunier and Nielsen, 2014). Using a permanent marker pen, we drew a black dot on the head and on the back of an animal, in this example a rat pup (Fig. 13.3a). This allowed these dots to be tracked independently using a freely available tracking program (www.kinovea.org). The resulting data, which consist of the x and y coordinates for each black dot in each video frame (at 15 frames/s), could then be used to calculate the direction in which the animal was facing/moving at any given time. This was done by drawing a virtual arrow (a vector) between the black dots, and the direction of this arrow relative to a feature of

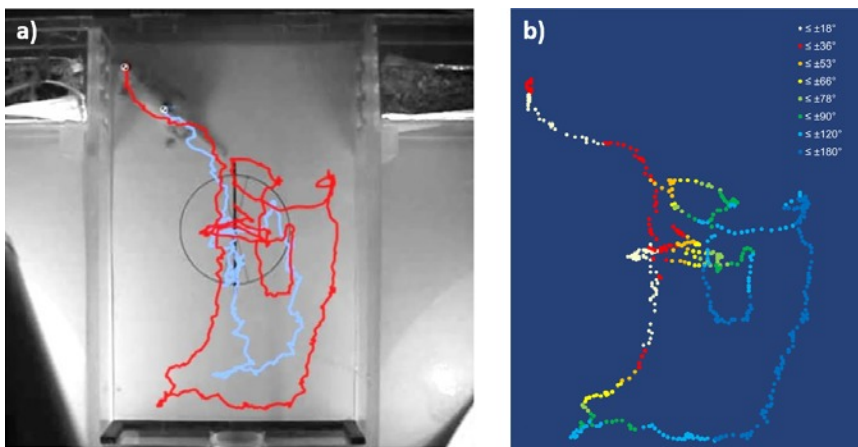


Fig. 13.3. Assessment of search direction of a rat pup in a test arena with two odours. (a) The pup has been marked with black dots on the back and head, and the tracking of these dots is shown in red (head) and blue (back). The direction of movement of the animal at any time follows the vector (arrow) that can be drawn between these two dots. (b) The direction of the pup is shown relative to the odour in the top left-hand corner of the test arena, with the colour of each dot indicating the degree to which the pup is heading for the odour (adapted from Meunier and Nielsen, 2014).

the test arena could then be determined for all time points. [Figure 13.3b](#) shows the position of the head of the pup with a dot for each video frame, and the colour of the dot indicating the angle between the actual direction of movement and the direct line to the goal, in this case the odour of the mother of the rat pup tested. Although this may appear complicated, the calculations are relatively simple, and the equipment used is inexpensive or even free. As is the case with all freeware, it is important that researchers able to contribute financially to the upkeep of these sites should remember to do so, thereby allowing less affluent research groups access to valuable tools.

So, the first issue was to appreciate what can still be achieved by simple means. The second issue I would like to raise is to remember that your data are always only as good as your experimental protocol allows. In other words, you may have the most advanced equipment with precise measurements available at high frequencies, but if your basic experimental plan is inappropriate or even flawed, these data will not contribute to answering your research question. Yes, I know this is common sense, and something all researchers are or should be aware of, but I also think that we all have come across colleagues blinded by the bling of their latest machine, who have proposed or even carried out tests that are statistically weak, and consequently unpublishable. Perhaps this second issue is linked with the first, in that the use of simpler experimental set-ups often makes the interpretation of the data easier, because we are forced to cut the protocol down to its minimum components. Unfortunately, we sometimes try to cram too much into one trial in an attempt to optimize the use of our animals and our expenses, but often with a cost paid in terms of clarity and sometimes even validity of results.

Concluding Remarks

Perhaps I should start this final paragraph by apologising to those of you who expected this chapter to describe all the fantastic futuristic solutions becoming available to get answers from animals in our care. Obviously, measuring animal behaviour in the wild as well as in test situations has come a long way in the past decades, and I can only anticipate that the rapid advances in technology in other fields, such as sport science, will also be of benefit to the animals we study. As I prepare this book manuscript to send to the publisher, interesting new results and articles keep appearing. I will thus not be able to include tests demonstrating the ability of goats to recognize emotion in the calls of conspecifics (Baciadonna *et al.*, 2019), or to dwell on the finding that laying hens that do not use the outdoor range area much appear to be more vigilant than more outgoing hens (Campbell *et al.*, 2019). Other relevant publications have without doubt been overlooked, and I apologize for any major omissions.

I will also take this opportunity to thank you for sticking with me this far. You may only have read this last chapter, but I nevertheless hope some of the messages within this book have piqued your interest, both in terms of animal behaviour and the methods used to ask animals questions. As always, many issues remain unresolved, but the field of applied animal behaviour science is in rapid development. With more focus on animal welfare in society in general, including among farmers, consumers, pharmaceutical companies, wildlife managers and politicians, our abilities to ask animals the right questions are becoming more and more important. This, of course, requires us to obtain the correct answers – even if they sometimes are not the ones we expect or want – and to use the resources available to us appropriately in a world of human expansion and climate change. End of lecture. End of book.

References

- Baciadonna, L., Briefer, E.F., Favaro, L. and McElligott, A.G. (2019) Goats distinguish between positive and negative emotion-linked vocalisations. *Frontiers in Zoology* 16(1), 25. DOI: 10.1186/s12983-019-0323-z.
- Bell, F.R. and Lawn, A.M. (1957) The pattern of rumination behaviour in housed goats. *The British Journal of Animal Behaviour* 5(3), 85–89. DOI: 10.1016/S0950-5601(57)80002-1.
- Brown, D.D., LaPoint, S., Kays, R., Heidrich, W., Kümmeth, F. *et al.* (2012) Accelerometer-informed GPS telemetry: reducing the trade-off between resolution and longevity. *Wildlife Society Bulletin* 36(1), 139–146. DOI: 10.1002/wsb.111.
- Campbell, D.L.M., Taylor, P.S., Hernandez, C.E., Stewart, M., Belson, S. *et al.* (2019) An attention bias test to assess anxiety states in laying hens. *PeerJ* 7(1–2), e7303. DOI: 10.7717/peerj.7303.
- Cox, S.L., Orgeret, F., Gesta, M., Rodde, C., Heizer, I. *et al.* (2018) Processing of acceleration and dive data on-board satellite relay tags to investigate diving and foraging behaviour in free-ranging marine predators. *Methods in Ecology and Evolution* 9(1), 64–77. DOI: 10.1111/2041-210X.12845.
- Graf, P.M., Wilson, R.P., Qasem, L., Hackländer, K. and Rosell, F. (2015) The use of acceleration to code for animal behaviours; a case study in free-ranging Eurasian beavers *Castor fiber*. *PLoS ONE* 10(8), e0136751. DOI: 10.1371/journal.pone.0136751.
- Gutierrez, E., Quinn, D.B., Chin, D.D. and Lentink, D. (2017) Lift calculations based on accepted wake models for animal flight are inconsistent and sensitive to vortex dynamics. *Bioinspiration & Biomimetics* 12(1), 016004. DOI: 10.1088/1748-3190/12/1/016004.
- Howerton, C.L., Garner, J.P. and Mench, J.A. (2012) A system utilizing radio frequency identification (RFID) technology to monitor individual rodent behavior in complex social settings. *Journal of Neuroscience Methods* 209(1), 74–78. DOI: 10.1016/j.jneumeth.2012.06.001.
- Hwangbo, J., Lee, J., Dosovitskiy, A., Bellicoso, D., Tsounis, V. *et al.* (2019) Learning agile and dynamic motor skills for legged robots. *Science Robotics* 4(26), eaau5872. DOI: 10.1126/scirobotics.aau5872.

- Korpela, J.M., Suzuki, H., Matsumoto, S., Mizutani, Y., Samejima, M. *et al.* (2019) Ai on animals: AI-assisted animal-borne logger never misses the moments that biologists want. *BioRxiv*. DOI: 10.1101/630053.
- Lutnesky, M.M.F. and Brown, T.R. (2015) Simulation of movement that potentially maximizes assessment, presence, and defense in territorial animals with varying movement strategies. *Ecological Modelling* 313, 50–58. DOI: 10.1016/j.ecolmodel.2015.06.003.
- Meunier, N. and Nielsen, B.L. (2014) Using scalar products to refine the interpretative value of an orientation choice test. *Chemical Senses* 39(6), 551–560. DOI: 10.1093/chemse/bju024.
- Nishiumi, N., Matsuo, A., Kawabe, R., Payne, N., Huveneers, C. *et al.* (2018) A miniaturized threshold-triggered acceleration data-logger for recording burst movements of aquatic animals. *The Journal of Experimental Biology* 221(6), jeb172346. DOI: 10.1242/jeb.172346.
- Plotkin, A., Paperno, E., Vasserman, G. and Segev, R. (2008) Magnetic tracking of eye motion in small, fast-moving animals. *IEEE Transactions on Magnetics* 44(11), 4492–4495. DOI: 10.1109/TMAG.2008.2002187.
- Volpov, B.L., Hoskins, A.J., Battaile, B.C., Viviant, M., Wheatley, K.E. *et al.* (2015) Identification of prey captures in Australian fur seals (*Arctocephalus pusillus doriferus*) using head-mounted accelerometers: field validation with animal-borne video cameras. *PLoS ONE* 10(6), e0128789. DOI: 10.1371/journal.pone.0128789.
- Zobel, G., Weary, D.M., Leslie, K., Chapinal, N. and von Keyserlingk, M.A.G. (2015) Technical note: validation of data loggers for recording lying behavior in dairy goats. *Journal of Dairy Science* 98(2), 1082–1089. DOI: 10.3168/jds.2014-8635.

Index

- accelerometers 166–167, 168
- acoustic startle response 35–36
- active avoidance test 116
- activity
 - day-old chicks 12–13, 137–138
 - monitoring 12, 24, 166–170
 - in open field test 31, 125–126
- affective state 43, 161
 - behavioural treatments 158–159
- age/ageing 78–82
 - experience 82–85
- aggression (pigs) 36–37, 40–41
- albino rats 129
- animal models of disease 36, 130, 161–162
- animal welfare 90
 - aversive stimuli 80–81, 86–88, 98–100, 116, 156–159
 - definition 159–160
 - different from ethics 160–161
 - laws and guidelines 154–156, 162–163
- anti-depressants 86–88, 162
- anxiety 24, 37, 140
- anxiolytic drugs 86, 157
- arena tests *see* open field test
- artificial intelligence 168
- associative learning
 - classical conditioning 53
 - operant conditioning 57–61, 94
 - see also* punishment; reinforcement
- avoidance tests 52–54, 116–117, 149
- bearded dragons (social learning) 120
- bears
 - clicker training 102–103
 - contrafreeloading 56
- bees (memory) 113–114
- behavioural treatments 158–159
- birds
 - cognitive abilities 110–112
 - quantity estimation 69
 - silhouette experiment 142–145
 - tracking wing movements 168
 - see also* poultry
- breed differences 129–131
- breeding
 - co-selection of other traits 126, 130–131, 133–134
 - selection for/against behavioural traits 124–126, 128–129
- buried food tests 138–139
- cancer detection 73–75
- carp (boldness) 39–40
- cats
 - can recognize their own names 71–72
 - laterality 52
 - virtual fences 100
- cattle
 - colour discrimination 72
 - electric/virtual fences 99, 100
 - Go/No-go test 118–120
 - human approach test 37
 - lameness 9–10
 - odour discrimination 71
 - oestrus 11

- cattle (*continued*)
- recumbency requirement 58–59, 60–61
 - restlessness 16
 - scratching brushes 19, 61
 - training calves where to urinate 100–101
- character tests 29–45
- chickens *see* poultry
- chimpanzees (quantity estimation) 67–69
- choice tests 50–55
- contrafreeloading 56
- choosing a test 19–27, 139–140
- classical conditioning 53
- clicker training 102–103
- closed-economy tests 55
- cognitive ability 107–112, 120–121
- environmental enrichment 88–89
 - memory 24, 112–117
 - see also* learning
- cognitive bias 42–45
- colour discrimination
- chicks 65–66
 - cows 72
- companion animals 90
- see also* cats; dogs
- conditioned place preference/avoidance tests (CPP/CPA) 52–54
- confounding effects 21
- contrafreeloading 56
- coping behaviour 39–40, 130
- cost of testing equipment 23, 169
- coverage of the book 1–5
- cows *see* cattle
- crows (cognitive ability) 110–112
- cylinder test 110–112
- data handling/analysis 3–4, 145–147, 169
- deer (reaction to predator odours) 142
- delayed match-to-sample test 112–114
- demand curves 57
- depression 86–88, 161–162
- desensitization 82–84
- designing a test 19–27, 139–140
- detour test 109–110
- developmental changes in behaviour 81–82
- digital technology 166–168
- discrimination between stimuli
- colour
 - chicks 65–66
 - cows 72
 - habituation–dishabituation test 25, 69–72
 - odour
 - cows 71
 - detection of human disease 73–75
 - rodents 70–71, 73
 - quantity estimation 67–69
 - sound 71–72
- dishabituation 69–70
- dogs
- detection of human cancer by odour 73–75
 - object permanence 108, 108–109
 - quantity estimation 69
 - training
 - punishment 98–99
 - reinforcement 95
- domestic livestock *see* cattle; goats; horses; pigs; poultry; sheep
- domestication 127, 129–130, 130–131
- dominance hierarchy 41–42
- drug testing 86–88, 138–139, 157, 159, 162
- dustbathing 11–12, 81
- eggs
- artificial 138
 - traits co-selected with TI test
 - duration 126
- electric shocks 98–100, 116, 157–158
- elevated plus maze (EPM) tests 37
- emotion 31, 161
- emotional reactivity 34–36
- environment
- enrichment 88–89, 158
 - genotype x environment interactions 127–128
 - housing 7–8
 - social isolation 25–26, 32
 - testing in the home environment 141
- epigenetics 126–127
- equipment for tests 23–24
- hole boards 88, 112
 - mazes 37, 52
 - nose-poke devices 73

- technological advances 166–169
- water maze 79
- error, potential sources 137–139
 - choice tests 50, 52
 - data-related 12–13, 137–138, 145–147
 - discriminatory ability 72–73
 - drug testing in rodents 87–88, 138–139
 - judgement bias test 43–45
 - odour detection 74–75
 - operant conditioning 60–61
 - stress of handling 140–141
- ethics of animal testing 90, 154, 156–163
 - aversive stimuli 80–81, 86–88, 98–100, 116, 156–159
 - drug testing 86–88, 157, 159, 162
 - laws and guidelines 154–156, 162–163
- ethograms 7, 139
- Europe (regulations) 154–155, 162–163
- experience 78, 82–85
- fasting (pre-test) 86, 158
- fearfulness 29
 - heritability 126, 128–129
 - novel object test 32–34
 - open field test 32
 - startle test 34–36
- feed deprivation (pre-test) 54, 86, 158
- feeding behaviour
 - broiler breeders 21–23, 54
 - choice tests in rats 50–51
 - novelty-suppressed feeding 86
 - observed 13–16
 - quantity estimation 67–69
 - trout 131
- female vs male test animals 148–150
- fish
 - boldness 39–40
 - detour test 109–110
 - feeding behaviour 131
 - preference/avoidance 52
- food rewards 94–95, 102, 158
- forced swim test 23, 25, 86–87, 87–88, 148, 162
- foxes (bred for tameness) 129–130
- Gallagher's proximity 80
- Geller Seifter tests 157
- gender 150
- genetically modified animals 161–162
- genetics 124
 - co-selection of other traits 126, 130–131, 133–134
 - epigenetics 126–127
 - genotype x environment interactions 127–128
 - innate vs learnt behaviour 132–133
 - selection for/against behavioural traits 124–126, 128–129
 - strain/breed differences 129–131
- giraffes (object permanence) 108
- Go/No-go tests 117–120
- goats
 - object permanence 108
 - rumination 167–168
 - virtual fences 100
- great tits (size estimation) 69
- grizzly bears
 - clicker training 102–103
 - contrafreeloading 56
- habituation 82–84
 - to handling 140
- habituation–dishabituation test 25, 69–72
- handling of animals 115–116, 140–141
- hens *see* poultry
- heritability of behaviours *see* genetics
- hierarchy tests 41–42
- hole-board discrimination test 88–89, 112
- honey bees (memory) 113–114
- horses
 - habituation 82–84
 - indicating preferences via symbols 20
 - judgement bias test 43
 - negative reinforcement 93–94, 95–96, 157
 - social learning 120
- human approach test 36–37, 115–116
- humans
 - diseases
 - animal models 36, 130, 161–162
 - odour detection 73–75
 - Go/No-go test 117–118
 - object permanence 107
 - training 104
- hunger, in broiler breeders 21–23, 54

- identification of individual animals 8–9
- impulsiveness 117–118
- innate behaviour 7, 132–133, 142
- insects 3
 - memory in bees 113–114
- intelligence 107
 - see also* cognitive ability; learning
- Japanese quail 32, 39, 126
- judgement bias test 42–45
- laterality 52
- Latin squares 145–146
- learning 24, 117–120
 - habituation 82–84
 - innate vs learnt behaviour 132–133
 - in young and old rats 79–81
 - see also* associative learning
- legislation 154–156, 162–163
- lizards (social learning) 120
- locomotor activity
 - day-old chicks 12–13, 137–138
 - monitoring 12, 166–170
 - in open field test 31, 125–126
- Lorenz, Konrad (silhouette experiment) 142–145
- male vs female test animals 148–150
- maternal behaviour
 - pigs 7, 37
 - rats 126–127
- mazes 37, 52
 - Morris water maze 79–81
- memory 24, 112–117
- metamemory 113
- mice
 - anxiety (EPM test) 37
 - electrified grids 157–158
 - forced swim test 148
 - hierarchy 42
 - housing 9, 141
 - odour discrimination 70–71
 - sex differences 149
 - strain differences 125–126, 146
 - suckling a learnt response 132–133
 - tail suspension test 87
 - see also* rodents
- mink
 - choice of environment 55
 - fearfulness/confidence 128–129
 - pulling or pushing 61
- monkeys (reinforcement) 95, 97
- mood 42–43, 161
- Morris water maze 79–81
- Moser's zone 80
- motivational tests 56–61, 157–158
- multiple choice tests 54–55
- negative punishment 98, 104
 - and positive reinforcement 100–101
- negative reinforcement 93–94, 95–98, 157
- nocturnal behaviour 9, 16
- non-test observations 7–16, 139
- nose-poke devices 73
- novel object test 32–34
- novelty-suppressed feeding 86
- object permanence 107–109
- observational studies 7–16, 139
- odour
 - detection of human disease 73–75
 - discrimination
 - in cows 71
 - in dogs 73–75
 - in humans 73
 - in rodents 70–71, 73
 - response to predator odours 132, 142
- oestrus 11
- older animals 78–80
- open field test 30–32
 - re-testing 148
 - social isolation in pigs 25–26
 - variation in mice 125–126, 146
- operant conditioning 57–61, 94
 - see also* reinforcement
- passive avoidance test 116, 149
- passive infrared detectors (PIDs) 12
- Pavlovian conditioning 53
- personality tests 29–45
- pets 90
 - see also* cats; dogs

- pigs
- aggressiveness 36–37, 40–41
 - cognitive ability and housing type 88–89
 - fearfulness 33–34
 - memory 115–116
 - nest-building 7
 - oestrus 11
 - rooting material preference 59–60
 - spray-marking to identify individuals 8–9
 - stress of testing 25–26, 140–141
 - tonic immobility 38, 39
- pilot tests 140
- positive punishment 98–100, 104, 157
- positive reinforcement 93, 94–95, 97–98, 103, 157
- clicker training 102–103
 - in humans 104
 - and negative punishment 100–101
- post-traumatic stress disorder (PTSD) 36
- poultry
- activity of day-old chicks 12–13, 137–138
 - breeding for/against feather pecking 124–125
 - colour discrimination 65–66
 - CPA/CPP tests 54
 - dustbathing 11–12, 81
 - epigenetic changes in domesticated breeds 127
 - fearfulness 126
 - hunger in broiler breeders 21–23, 54
 - light preferences 54–55
 - long-term effects of experience 84–85
 - strain differences 130
 - tonic immobility 38, 39, 126
- predator odours, response to 132, 142
- preference tests 50–61
- principal component analysis 31
- probe tests 80–81
- publication of results 140, 161
- punishment 94, 98–100, 104, 157
- negative, and positive reinforcement 100–101
- quail 32, 39, 126
- quantity estimation 67–69
- rats *see* rodents
- reference memory 112
- reinforcement 93–98, 103, 157
- clicker training 102–103
 - in humans 104
 - and negative punishment 100–101
- reproductive behaviour
- male rats 50, 132
 - oestrus 11
- resident–intruder test 40–41
- re-testing an animal 147–148
- rodents
- acoustic startle response 35–36
 - activity monitoring 169–170
 - age and learning ability 79–81
 - anxiety (EPM test) 37
 - avoidance tests 116, 149
 - buried food test 138–139
 - drug testing 86–88, 138–139, 157, 159, 162
 - electrified grids 157–158
 - feed choice 50–51
 - food/water deprivation 158
 - forced swim test 23, 25, 86–87, 87–88, 148, 162
 - genetically modified 161–162
 - hierarchy 42
 - housing 9, 141
 - innate vs learnt behaviour 132, 132–133
 - maternal care 126–127
 - memory 112–113, 116–117
 - obesity 130
 - odour discrimination 70–71, 73
 - oestrus 11
 - predator odours 132
 - re-testing 148
 - sex differences 149
 - strain differences 125–126, 129–130, 146
 - suckling 132–133
 - tail suspension test 87
 - tickled 89
- sample size 145
- senescence 78–80
- serotonin inhibitors 43
- sex differences 148–150

- sheep
 - genotype x environment interactions 128
 - judgement bias test 43
 - startle test 34–35
- shock collars 98–100
- short-term (working) memory 112–114
- social encounter tests 40–41
- social hierarchy 41–42
- social isolation 25–26, 32
- social learning 120
- sodium lights 9
- sows
 - aggression 36–37
 - nest-building 7
 - oestrus 11
 - see also* pigs
- spatial memory 112
- standardization of tests 30, 162
- startle test 34–36
- statistics 3–4, 145–146
- stress
 - causing epigenetic changes 127
 - experimental induction 159
 - in pigs 25–26, 140–141
 - in rodents 80–81, 86–87
- T-mazes 52
- tail suspension test 87
- test group size/composition 29, 145–146
- thigmotaxis 30–31, 80
- time budgets 8–10
- time-specific activities 11–12
- Tinbergen, Nikolaas (silhouette experiment) 142–145
- tonic immobility (TI) tests 38–40, 126
- Tourette's syndrome 162
- tracking of movement 166–170
- training
 - clickers 102–103
 - humans 104
 - punishment 98–100, 104, 157
 - reinforcement 93–98, 103, 104
- treatment effects, tests of 85–89, 138–139
- trout (feeding behaviour) 131
- USA (regulations) 155, 156
- video recording of tests 30, 146–147, 166
- Vogel conflict tests 157
- volume estimation 67–69
- water deprivation 158
- weight, pre-test 145
- welfare *see* animal welfare
- Whishaw's corridor 80
- wild animals
 - becoming more nocturnal 16
 - clicker training 102–103
 - experience 84
 - predator odours 142
 - training 95, 102–103
- withdrawal response test (human approach test) 36–37, 115–116
- working (short-term) memory 112–114
- Y-mazes 52
- zebrafish (choice) 52

ASKING ANIMALS

AN INTRODUCTION TO ANIMAL BEHAVIOUR TESTING

Birte L. Nielsen

Thought-provoking yet practical, this text provides an introduction to the use of behaviour tests applied to animals. By including illustrative examples from a variety of species, the book inspires the animal scientist to think about what a given behavioural test can be used for and how the results can be interpreted.

The book includes:

- The dos and don'ts of running behaviour tests and interpreting the results.
- Many clear, simple illustrations which make the information readily accessible.
- Down to earth, practical advice, yet a thorough, evidence-based approach.
- Information on behaviour tests for a whole range of species from companion, farm and zoo to laboratory and wild animals.
- Succinct yet comprehensive text, designed to be read cover to cover and stimulate further reading.

This book provides an essential basis for embarking on, and devising, any animal behaviour test. It is valuable to students, established researchers, teachers and practitioners of applied ethology, animal welfare science, and veterinary science.