Does the goal justify the methods? Harm and benefit in neuroscience research using animals

Ana Catarina Vieira de Castro and I Anna S Olsson
Laboratory Animal Science Group, IBMC – Instituto de Biologia Molecular e Celular, Universidade do Porto, Portugal. Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal
in Curr Top Behav Neurosci. 2015, 19: 47-78
http://dx.doi.org/10.1007/7854_2014_319

ABSTRACT

The goal of the present chapter is to open up for discussion some of the major ethical issues involved in animal-based neuroscience research. We begin by approaching the question of the moral acceptability of the use of animals in research at all, exploring the implications of three different ethical theories: contractarianism, utilitarianism and animal rights. In the remainder of the chapter we discuss more specific issues of neuroscience research within what we argue is the mainstream framework for research animal ethics, namely one based on harm-benefit analysis. We explore issues of harms and benefits and how to balance them as well as how to reduce harm and increase benefit within neuroscience research.

1. Introduction

1.1. The ethical dilemma of animal research

Studies on live animals play an important role in neuroscience research. In basic neuroscience research animals are studied to understand the functioning of the nervous system and the mechanisms involved in the diseases that affect it. In applied neuroscience research animals are used to develop and test therapies for such diseases. The ultimate aim of both lines of research is commonly to extrapolate results to the human case. When animals are used as models of human diseases – which constitutes the bulk of animal-based neuroscience research - the object is to induce in them conditions which, at least in some aspects, mimic the conditions that researchers aim to understand in humans and for which they wish to develop appropriate treatment. Disease-oriented research in neuroscience includes the study of both psychiatric and neurological disorders. The former consist of disorders of mood and thought associated with either no apparent signs, or at most only minor physical signs in the motor and sensory systems, and includes diseases such as schizophrenia, depression or anxiety. The latter refer to nervous system disorders that also present somatic signs, and include neurodegenerative diseases, such as Alzheimer's, Parkinson's,

Huntington's disease, Amyotrophic Lateral Sclerosis, or stroke, and pain (Baker et al, 2002). Animal models are used in research into both psychiatric and neurological disorders.

The use of animals for research remains a controversial issue. Most experimental procedures are likely to inflict at least some harm on the animals that are studied. During experimentation, animals may be in relatively limiting conditions, and deprivation of food and water often forms part of behavior testing schedules. Varying degrees of physical or psychological harm can result from the procedures used to induce in animals conditions mimicking the human diseases under study, as well as from the conditions themselves. Distressing or painful interventions may be part of experimental protocols and, not least, most animals are killed at the end of experimental trials. Unlike humans participating as subjects in research, however, animals cannot consent to their own participation. Moreover, since most research is intended to benefit humans and not animals, there is no benefit for the research subjects themselves. While perhaps few researchers would question the desirability of discovering new ways to prevent, alleviate, or cure human diseases, the question remains: Are we, as human beings, morally justified in using animals as tools for research?

The answer to this question, of course, depends on one's ethical framework. In society[1] we find a range of opinions: while some people are outraged by the idea of inflicting suffering in innocent animals, others consider it acceptable to do research on animals, if there are no alternatives available and if the outcome of the research is valuable; at the extreme, some see no problem in animal experimentation at all (e.g., Nuffield Council, 2005). In a pluralist society, there is room for such a variety of opinions to co-exist, but there is also the need for finding a working compromise which can be accepted by a majority, for example in order to draft policy and legislation.

The main driver for regulating the use of animals in experiments is the demand from society to protect these animals. Public concern for the well-being of animals has a long history and the protection given to animals has gradually been formalized, to the extent that animal welfare today is mentioned in several constitutional documents – including those of the European Union. Legislation protecting animals used in research was first introduced in the second half of the 19th century with the enactment of the UK Cruelty to Animals Act (1876), the first legislation to place "conditions on the way in which [experiments on animals] were done, the reasons for which they were done, and the qualifications of persons responsible for performing them" (Bayne et al, 2011). Today, the use of animals in research is commonly regarded as requiring explicit ethical justification. Such justification has become an integral part of the review process for research proposals involving animals and is prescribed in legislation worldwide. Usually, a research proposal is understood to being justifiable only if the benefits of the study can be demonstrated to outweigh the expected harms to the animals used.

Basing ethical justifications for a given procedure on some evaluation of the relative magnitudes of harm and benefits is currently the central requirement in legislation worldwide, as apparent in official guidelines and policy documents for ethical review. For example, the Australian Animal Welfare Act 1999 states that the key principle underlying the application approval process is that "the use of animals in research, testing, and teaching is confined to cases in which there is good reason to believe the benefits (…) are not outweighed by the likely harm to the animals". More recently, the new European Directive 63/2010/EU also makes explicit that "(…) an impartial project evaluation, independent of those involved in the study, should be carried out (…) The project evaluation shall consist in: (…) (d) a harm-benefit analysis of the project, to assess whether the harm to the animals in terms of suffering, pain and distress is justified by the expected outcome taking into account ethical considerations, and may ultimately benefit human beings, animals or the environment".

Before further expanding on the mechanics of this process of weighing harms and benefits of particular research programmes, however, we will first look briefly at the ethics underlying the wider issue of the overall acceptability of any form of animal experimentation.

[1] In this chapter, we use the word 'society' to refer to a wider public in the industrialized countries.

1.2. Ethical theories

Assessment of harms and benefits has become the mainstream, approach in examining the potential justification for individual, animal-based, research protocols. However, this approach presupposes from the outset that animal experimentation is, in principle, acceptable to society, within defined constraints. Hence, assessment of the harms and benefits of a particular piece of research does not address the more fundamental ethical issue of animal experimentation in general. There are various ethical theories regarding human duties to animals, which offer different perspectives on the acceptability of the use of animals in research. Here, we will expand on three of them: contractarianism, utilitarianism and animal rights. These three ethical perspectives, representing a significant proportion of the debate over animal research, reflect three relevant but markedly different views on the issue.

According to the contractarian view (e.g., Carl Cohen, 1997), animals are morally insignificant or lack moral status. Contractarians regard morality as a system of hypothetical contracts that rational individuals tacitly enter into with one another. Animals cannot be direct parties to such contracts, or agreements, because they lack the linguistic and intellectual skills to do so. Hence, animals are not bearers of duties or rights. However, for those who adopt the contractarian view, the way animals are treated is not by definition irrelevant. To the extent that people care about animals, the tacit contracts that constitute morality will contain clauses affording some protection to them. If some people are emotionally attached to certain types of animals and disapprove of their use in experiments, this becomes an ethical concern. For the contractarian, public concern is the overarching issue determining the ethical acceptability of animal experiments. Harm-benefit analysis is only relevant to the extent that it affects public acceptance of research – it is not relevant in itself.

The animal rights view (e.g., Tom Regan, 1989), as the name indicates, defends that some animals (the *experiencing subjects of a life* – usually understood as conscious animals) have rights, at least the basic right always to be respected as an end in themselves and never to be used merely as a means to an end. From this perspective, animal experimentation is considered to be one of the many examples of human use of animals in which this right is disrespected – it implies using animals as instruments for reaching human purposes. Thus, the animal rights view has an abolitionist position as regards animal experimentation. It does not matter that an experiment will cause only minor suffering to the animals or that it will have an extraordinary contribution to humanity: experimentation with sentient animals is in itself unacceptable. Hence, the animal rights view rejects harm-benefit weighing as a measure of the acceptability of animal research.

According to the utilitarian view (e.g., Peter Singer, 1975), morality has one basic rule: always act so as to maximize the well-being of those affected by your actions or, in other words, to create the greatest amount possible of good. In the utilitarian approach to animal ethics, the good to be maximized is defined in terms of pleasure and the absence of suffering. Moral consideration is therefore entitled to every sentient creature. In the utilitarian approach, then, ethical decisions require us to strike the most favorable balance of benefits and harms for all sentient individuals affected by what we do. Utilitarians support the use of animals in research if and only if the harm to animals is outweighed by the benefits of research. In fact, harm-benefit analyses are the core of the utilitarian approach.

The question raised earlier about whether humans are morally justified to use animals for experimental purposes has, thus, no single answer. There are a number of different ethical perspectives, of which we have here advanced three. These three views cover the most common arguments heard in the public debate about the use of animals in experiments. These arguments however focus on the fundamental principle of whether or not it is ethically legitimate to undertake experiments on animals at all. In the rest of the chapter, we will work within a practical reality where such experimentation is legal and accepted by the majority of society, if within certain constraints. Within such a framework we will explore the more applied aspects of animal research ethics: that is, how can animal ethics influence and affect actual animal experiments, their regulation and how they are carried out.

In this more practical application of how ethical considerations may be applied to actual research protocols, issues of harm versus benefit play a central role, as well as considerations of how these two factors may be weighed against each other. It is

important to stress that we are not favoring one particular ethical perspective. Rather, we attempt to explore specific issues of animal research ethics in neurosciences within the utilitarian-based, mainstream, position. Although the harm-benefit weighing has a strong component of utilitarianism, this does not mean that considerations of harm and benefit are only relevant from the utilitarian perspective. Neither contractarians nor animal rights defenders would be against the minimization of the harms caused to the animals used in research and at least contractarians would support the maximization of the benefits to be gained from it.

2. Harms

In this chapter, we follow the standard approach used in ethical evaluation of animal research and consider harms as negative impacts, actually or potentially, caused to the welfare of animals. That includes adverse effects on health, as well as all the adverse subjective experiences animals might undergo, such as pain, fear or anxiety. We also include killing as a harm, based on the understanding that killing rarely is in the animal's own best interest and it prevents it from all potential future positive experiences (see Yeates, 2010).

2.1. What are the harms of neuroscience research?

If we follow the approach outlined in legislation, the ethical acceptability of animal research is determined by balancing its benefits and harms. Harms to animals used in neuroscience research can potentially result from inappropriate transport, housing, handling and care, from the procedures used to inflict the disorders to be studied, from the experimental techniques to administer treatments and monitor parameters, and from badly-conducted euthanasia. Using concrete examples, we discuss specific harms to animals associated with neuroscience research. The cases discussed have been selected to illustrate important issues, with no claim for the list to be exhaustive.

As our first illustration: the $SOD1^{G93A}$ mouse model of Amyotrophic Lateral Schelerosis (ALS) is an example of a genetically-modified neuroscience model. In this model, the onset of this neurological disease is characterized by weakness and tremors of hind limbs. Disease progression then leads to paralysis of hind limbs, accompanied by increased difficulty to eat, drink and swallow, terminating in complete paralysis (Lever et al, 2009; Lever et al, 2010). Mice die of respiratory failure due to paralysis of diaphragm (Solomon et al, 2011). Discomfort is unavoidable at least in the more advanced stages of the disease in this animal model due to the progressive loss of motor capacity which characterizes the disease. In some studies the animals reach an advanced stage in which they have difficulty in reaching food and water provided at the cage top and performing behaviors as simple as chewing. Animals kept beyond this stage will die from the disease. The pain that patients might experience has been a neglected aspect of ALS. In a recent review, Handy et al (2011) raised this concern. Although not generally associated with ALS, pain has been reported to occur in nearly 70% of ALS patients at some point in time. This raises an additional concern regarding the welfare of animals used as models of ALS. They may, like humans, feel pain during disease progression.

Whereas the $SOD1^{G93A}$ mouse phenotype is well known by now, special attention must be given to genetically altered animals whose phenotype is not (yet) thoroughly characterized, as unpredictable changes in physiology or behavior may also occur (Morton and Hau, 2011).

Different welfare concerns can be found in the next illustrative case, rats with neonatal hippocampal lesion (NHL), used as a model of schizophrenia. In the NHL model, the disorder is surgically-induced, in that the hippocampus of these rats is lesioned few days after birth by way of an injection. The animals are then returned to their mothers and weaned normally. This model involves a number of potential stressors for the animal: maternal separation before surgery, the surgery itself, postoperative recovery and presumably postoperative pain and discomfort. At this point, pain and stress appear as the major animal welfare issues. As adults these animals present several behavioral deficits, such as increased response to stress, deficient prepulse inhibition and latent inhibition, impaired social behaviors, and working memory problems (e.g., Lipska, 2004). Although these impairments do not generate physical suffering, they reflect a loss of capacities which may have psychological repercussions for

the animals. For example, it can be assumed that handling and external disturbances are more stressful for animals with hyper-responsiveness to stress or with memory impairment, which may result in a diminished capacity for behavioral habituation. Also, these rats may undergo stress by being housed in groups due to their impaired social behavior.

Another important area of neuroscience research with its own welfare issues is pain research in general, and neuropathic pain studies in particular. Neuropathic pain is defined as a form of chronic pain that results from damage or abnormal function of the central or peripheral nervous system (Abdi et al, 2004; Woolf, 2004). Patients with neuropathic pain frequently report sensory abnormalities including burning sensations, exaggerated responses to noxious stimuli (hyperalgesia), pain sensations resulting from innocuous stimuli (allodynia) and spontaneous pain episodes (dysesthesia) (Gilron et al, 2006). Due to its severity, chronicity and resistance to some classical analgesics (Gilron et al, 2006), it has received much attention in research.

A widely used animal model of neuropathic pain involves the surgical placement of tubing cuffs around the main branch of the sciatic nerve in rats. These animals are expected to undergo substantial pain in the course of the experiments. In this type of animal research, pain is the major welfare concern and the fact that pain is also the characteristic under study makes it more challenging to diminish it than in many other research situations. Besides the direct experience of pain, it has been demonstrated that long-term neuropathic pain can cause anxiety (after 4 weeks) and depression-related behaviors (after 6-8 weeks) (Yalcin et al, 2011). Such harms should be avoided where possible because, unlike pain, they are not a central feature of most studies on neuropathic pain. Although the results presented by Yalcin et al (2011) show that mood disorders and neuropathic pain may be connected, it is not a necessary condition for the study of pain that the animals are left anxious or depressed.

Harms such as those presented in this section affect the well-being of the animals, but they also have important implications for the quality of experimental data. Animals under stress may, for example, have altered physiological parameters which can interfere with the conclusions drawn from the studies. Hence, there are two major reasons for reducing harm to animals in research: to preserve animal welfare and to guarantee the cleanness of experimental data (Russell and Burch, 1959; Smaje et al, 1998).

2.2. How can we assess harms?

In the previous section, we gave some examples of harm inflicted to animals in experimental research in the neurosciences. These harms include effects on animal health but also adverse subjective experiences resulting from the experimental interventions. Health and subjective experiences both form part of what is commonly understood to represent 'animal welfare', and considerable research is directed towards understanding how animal welfare is affected by internal and external factors, as well as towards developing methods for assessing welfare status or changes to welfare. In the context of neuroscience, animal welfare science is particularly relevant when it comes to understanding and measuring subjective experiences, including pain. Here, the two disciplines overlap considerably and there is potential for methods as well as ideas to flow both ways. Two relevant examples are the use of cognitive bias to assess animal welfare and the development of a face expression scale to assess pain in mice.

Based on the idea that the way information is processed by humans is affected by their mood and that thus, those suffering from a mood disorder are likely to make biased judgments (e.g., anxious and depressed people tend to make negative judgments about events and to interpret ambiguous stimuli unfavorably), Mendl and collaborators tested whether laboratory rats also show such cognitive bias. Rats were trained on a discrimination task in which one tone predicted the arrival of food (positive event) and the other the onset of white noise (negative event). When the rats had reliably learned to discriminate the tones, they were presented with ambiguous stimuli, that is, tones that were intermediate between those signaling positive and those signaling negative events. Rats having experienced chronic mild stress (according to a protocol often used in behavioral pharmacology as a model of depression, e.g., Willner, 1997) were more likely to respond to ambiguous cues as if they were predicting a negative event – that is they showed a negative cognitive bias (Harding et al., 2004). This research has received considerable attention in

animal welfare science as a promising method to assess subjective experiences; similar results have been demonstrated for a range of other mammal and also bird species (e.g. Mendl et al, 2009) and the approach may thus have considerable potential to help to quantify harm in animal experiments.

When Mogil and collaborators tested whether facial expression in mice reflects the level of painful stimuli to which they are exposed, their aim was to improve the relevance of animal-based pain research for human benefit. The subjective component is an essential part of the human pain experience, which is not presently available in non-verbal animals, and as the authors argue "[t]his measure of spontaneously emitted pain may provide insight into the subjective pain experience of mice" (Langford et al, 2010). However, it may also be used to develop measures to recognize pain in mice which can underlie decisions to apply analgesia or apply other measures of refinement, though such applications still have to be established.

By increasing our understanding of subjective experiences in animals, measures such as those outlined above may complement clinical measures of animal health when the welfare of animals is to be formally assessed. Such practical assessment is likely to play a larger role in the management of ethical issues in animal research, as formal assessment of actual harm resulting from a given protocol is being given greater emphasis. One example of this is the retrospective assessment of the effects of experimental procedures which is now being implemented at the European level (see 2.3.3).

Currently, however, assessment of probable harm is primarily an exercise of prediction within ethical evaluation for a proposed procedure. Towards this, guidelines and policy documents for the evaluation of animal experiments have suggested lists of criteria to be assessed. These criteria usually include the quality of the facilities, the experience of the personnel caring for and carrying out the procedures on animals, the number of animals, the animal species, the husbandry and housing conditions, the scientific procedures themselves (including the killing method), the duration and the intensity of the pain or distress likely to be inflicted on the animal, the fate of the animals at the end of the experiments and the endpoints to be applied (e.g., APC, 2003; FELASA, 2005; Smith and Boyd, 1991).

Some regulatory systems further ask for classification of the severity of procedures or experiments, that is, the degree of pain or suffering likely to be experienced by animals. In assessing the negative impact on the animal, the duration and frequency of a procedure is considered. Table 1 presents, as examples, the severity classifications adopted by the European Union in Directive 2010/63/EU and by the Canadian Council on Animal Care (CCAC). The new European Directive requires that experiments are attributed a severity classification (non-recovery, mild, moderate or severe) determined by the degree of pain, suffering, distress or lasting harm expected to be experienced by an individual animal during the course of the procedure. Following the European Directive, in assigning the category of severity one shall take into account any intervention or manipulation of an animal within a defined procedure, namely 1) the type of manipulation, handling, 2) the nature of pain, suffering, distress or lasting harm caused by (all elements of) the procedure, and its intensity, the duration, frequency and -multiplicity of techniques employed, 3) the cumulative suffering within a procedure and 4) the prevention from expressing natural behavior including restrictions on the housing, husbandry and care standards. For more than 20 years, evaluation under the Canadian Council on Animal Care (CCAC) has applied a similar scale of "Categories of Invasiveness in Animal Experiments" ranging from experiments on most invertebrates or on live isolates (A) to procedures in vertebrates and some invertebrates which may cause severe pain (E) (CCAC, 1993). *Protocols must be submitted to an appropriate review committee for all studies which involve the use of vertebrates and some invertebrates in Categories B through E. Both the European Directive and the CCAC guidelines give potential* examples of experimental procedures which are considered to be representative of each category. Although neurosciences are poorly represented in these documents, we advanced a few examples in Table 1.

Table 1. Severity classifications of animal experiments from the new European Directive 63/2010/EU (left column) and the Canadian Council on Animal Care (CCAC, 1993, middle column) and examples from neurosciences illustrating the different categories (right column).

EU Directive	CCAC Guidelines	Potential examples in neurosciences
Non-recovery Procedures which are performed entirely under general anesthesia from which the animal shall not recover consciousness		Some studies of neuroconductivity
	A Experiments on most invertebrates or on live isolates	Studies of neurobiology in *C. elegans*
	B Experiments which cause little or no discomfort or stress	Studies of cognitive bias in companion dogs without food or water deprivation
Mild Procedures on animals as a result of which the animals are likely to experience short-term mild pain, suffering or distress, as well as procedures with no significant impairment of the well-being or general condition of the animals	**C** Experiments which cause minor stress or pain of short duration	Mouse model of human repetitive mild traumatic brain injury
Moderate Procedures on animals as a result of which the animals are likely to experience short-term moderate pain, suffering or distress, or long-lasting mild pain, suffering or distress as well as procedures that are likely to cause moderate impairment of the well-being or general condition of the animals	**D** Experiments which cause moderate to severe distress or discomfort	NHL model of schizophrenia
Severe Procedures on animals as a result of which the animals are likely to experience severe pain, suffering or distress, or long- lasting moderate pain, suffering or distress as well as procedures, that are likely to cause severe impairment of the well- being or general condition of the animals		Huntington's disease ALS Learned helplessness
	E Procedures which cause severe pain near, at, or above the pain tolerance threshold of unanesthetized conscious animals	

2.3. How to reduce harms?

According to the mainstream approach on which this chapter is based, for an animal experiment to be ethically acceptable, the expected benefits not only must outweigh potential harms, but the harms caused to the animals must be reduced to a minimum as well, or in other words, animals shall not undergo unnecessary suffering. The Three Rs (replacement, reduction, refinement), proposed by Russell and Burch (1959), are widely recognized principles in the attempt to minimize harms to animals and, hence, to perform ethically acceptable research. Whereas the Replacement and Reduction principles reduce harm by avoiding animal use, the Refinement principle addresses the welfare of individual animals which are actually used in experiments. We will discuss each of the 3Rs in the context of neuroscience research.

2.3.1. Replacement

Replacement is the first of the three Rs, for several reasons*:*

"Replacement enjoys a particular standing among the 3Rs. It was the first of the Rs to be introduced by Russell and Burch (1959), reflecting the intended order in which the Rs were to be considered. Questions about Reduction and Refinement are only relevant if Replacement has first been considered and excluded. The goal of Replacement also has received widespread support, in part because it is the only goal that is fully compatible with the animal rights perspective that animal use solely for human benefit should not be permitted." (Olsson et al, 2011).

The main point of this principle of Replacement is that the use of animals should be replaced by non-animal methods whenever this is possible without compromising the research objective. Replacement methods can be divided into four main types: *in vitro* (e.g., cell lines), *ex vivo* (e.g., tissue cultures), *in silico* methods (e.g., bioinformatics) and research with human volunteers. The idea that studies in human volunteers would be an ethical alternative to the use of animals in research may seem provocative, and it is of course a *sine qua non* condition that such a study meets the ethical standards for research with human subjects. That said, in the neurosciences there seems actually to be real potential for this approach where replacement not only spares animals but also increases the relevance of the research itself. Focusing specifically on pain research, a workshop which brought together neuroscientists with proponents of non-animal research methods came up with a number of suggestions for situations in which studies with human volunteers could replace animal studies. They are all based on the use of low-risk minimally invasive techniques (e. g., functional imaging and microdialysis) in humans, and draw on the fact that it is much easier to evaluate subjective experience - a key aspect of pain – in verbal humans (Langley et al, 2008).

Another example of a replacement strategy in neuroscience research was recently described by Barnett and collaborators (Sorensen et al, 2008; Boomkamp et al, 2012). These authors proposed an *in vitro* method for research on spinal cord injury, a disorder that has depended mainly on animal research. Spinal cord injury is a complex injury, caused by traumatic accidents. Traumatic injury disrupts spinal white matter tracts, resulting in loss of sensory and motor function. This loss of function is generally permanent because the central nervous system has a restricted regenerative capacity (Fawcett and Asher, 1999; Rudge and Silver, 1990). After the initial injury, which results from direct mechanical disruption of spinal cord integrity, glial scars are formed, which inhibit central nervous system repair by creating both physical and biochemical barriers to axonal growth (Boomkamp et al, 2012).

An example of an animal model of spinal cord lesion is a wire knife lesion, generated by inserting the knife into the dorsal column and pulling up a piece of tissue. The method results in a cavity and glial scarring that mimics human spinal cord injury. Disadvantages to rat models of spinal cord injury include the need for large numbers of animals, the severity of the procedure for the animals, the long time frame for results and the high expensive of the experiments (International Animal Research Regulations, Impact on Neuroscience Research, Workshop Summary, Institute of Medicine (US); National Research Council (US), 2012).

In the non-animal model proposed by Barnett and collaborators, embryonic spinal cord cells from rats are layered on top of an astrocyte monolayer derived from embryonic tissue. Growth in culture over time leads to complex axonal/glial interactions resulting in myelinated neurons. This system allows for the study of contact between astrocytes and how they communicate with the axons, which is necessary for understanding the problems in spinal cord injury. The researchers also have induced lesions in the cell culture by cutting with a scalpel to studying axon density and myelination adjacent to the lesion and cell growth into the damaged area (International Animal Research Regulations, Impact on Neuroscience Research, Workshop Summary, Institute of Medicine (US); National Research Council (US), 2012).

Overall, the greater the role of non-animal replacement in research, the fewer animals will be needed in total for research purposes. In this way, replacement is also directly related to the second R, Reduction.

2.3.2. Reduction

The aim with the principle of Reduction is to use the smallest possible number of animals to obtain valid information. Its main ethical purpose is to reduce collective animal harm, understood as the number of animals on which harm is inflicted. One important measure is to use correct and careful statistics, namely by carrying out appropriate power analysis prior to study commencement. Sample sizes can also be decreased by controlling variance associated with different environmental and genetic conditions, as for example, by using uniform housing conditions and inbred animals.

Reduction is probably the most controversial of the 3Rs. There is great political value in bringing down numbers of animals used in experimental procedures as a whole, as the number of animals reported in annual statistics is a very visible and easily understood aspect of research animal ethics. This also holds for replacement - performing fewer experiments is also immediately recognizable in the statistics. However, the problem with reduction is that, as detailed analyses have repeatedly shown, in actual research the number of animals used in an individual experiment is often too small for results to be reliable. This of course has important implications for the validity of the research results. Within a larger review of methods in neuroscience, Button and collaborators (2013) examined the statistical power of animal experiments investigating sex differences in water maze and radial maze performance. The effect (i.e. how large a difference there is between male and female animals) was calculated through a meta-analysis, and the authors then established how many animals a single study would need to detect effects of this magnitude with different levels of statistical power. To achieve 80% power (a common standard), 134 animals would be needed for a water maze experiment and 68 for a radial maze, whereas the average sample sizes were 22 and 24 animals, respectively. The authors commented on the ethical consequences of underpowered studies:

> "There is ongoing debate regarding the appropriate balance to strike between using as few animals as possible in experiments and the need to obtain robust, reliable findings. We argue that it is important to appreciate the waste associated with an underpowered study — even a study that achieves only 80% power still presents a 20% possibility that the animals have been sacrificed without the study detecting the underlying true effect. If the average power in neuroscience animal model studies is between 20–30%, as we observed in our analysis above, the ethical implications are clear.
>
> Low power therefore has an ethical dimension — unreliable research is inefficient and wasteful. This applies to both human and animal research. The principles of the 'three Rs' in animal research (reduce, refine and replace) require appropriate experimental design and statistics — both too many and too few animals present an issue as they reduce the value of research outputs."

Based on this, it does not seem appropriate to apply reduction through uncritically decreasing sample sizes in individual experiments. Additional approaches in experimental design are needed if the aim is to bring down animal numbers. This could include the use of imaging techniques allowing the study of disease progress in the same animals rather than in separate groups for separate time points, or greater use of non-animal approaches before moving to an animal model.

2.3.3. Refinement

Whereas the Replacement and Reduction principles reduce harm by avoiding the use of animals, the Refinement principle addresses the welfare of individual animals which are actually used in experiments. This principle states that all experimental

procedures shall be adjusted to minimize any pain or discomfort they may cause to the animals. Experiments can be refined in several ways, from the use of anesthesia and analgesia, to housing adaptations and the establishment of humane endpoints. Appropriate measures need to be defined for each individual study, taking into account the nature of the harms which need to be mitigated. The scheme for welfare assessment recently proposed by a European working group allows refinement measures to be integrated into the assessment. Table 2 displays examples of refinement measures that can be applied in neuroscience studies whose harms were presented in section 2.2 (and are now summarized in the scheme).

Table 2. Schematic approach for assessing severity proposed by the European Commission Expert Working Group. Adapted from the document at http://ec.europa.eu/environment/chemicals/lab_animals/pdf/examples.pd

Example	**What does this study involve doing to the animals?**	**What will the animals experience? How much suffering might it cause? What might make it worse?**	**How will suffering be reduced to a minimum?**	
		Adverse effects	**Methodology and interventions**	**Endpoints**
Genetically-modified SOD1^{G93A} mouse model of ALS		Discomfort associated with motor capacity loss and difficulties to eat, drink and swallow	Housing adaptations (e.g., placing mashed food at a low level and adjusting of bedding to facilitate movement)	
		Pain	Use of painkillers	
		Animals may reach complete paralysis		Euthanasia of the animals as soon as possible to avoid unnecessary suffering
NHL rat model of schizophrenia	Maternal separation	Anxiety associated with maternal separation	Reduce duration of separation up to a minimum	
	Surgery for hippocampus lesion	Pain and discomfort associated with surgery	Appropriate anesthesia and analgesia	
		Potential stress resulting from handling and external disturbances	Handling and external disturbances avoided up to a minimum	
		Potential stress resulting from group housing	Rats housed in groups should be monitored for anxiety behaviors related to social contact and maybe housed individually	
Rat model of neuropahic pain	Placement, through surgery, of tubing cuffs around the main branch of the sciatic nerve	Pain	Adequate anaesthesia and analgesia during and immediately after surgery	
		Possible anxiety and depression-related behaviors		Early endpoints to avoid the development of anxiety and depression-related behaviors

The principle of the Three Rs is already present in much legislation. For example, the new European Directive states that "To ensure that the way in which animals are bred, cared for and used in procedures within the Union is in line with that of the other international and national standards applicable outside the Union, the principles of replacement, reduction and refinement should be considered systematically when implementing this Directive". Although the 3Rs principle was not explicitly referred to in previous European legislation, researchers were asked to use animals only when necessary, to use as few animals as possible and to use procedures having as little impact as possible.

Unfortunately, systematic reviews of the implementation of refinement measures in biomedical research indicate that the present situation is far from ideal. For example, between 2000 and 2002 pain relief was administered in only around 20 % of studies subjecting rodents to potentially painful procedures (Richardson and Flecknell, 2005). In 2009, humane endpoints were only reported in about 20% of studies of mice models of the neurodegenerative disorder Huntington's Disease, with no significant increase in the reporting of this refinement measure during the preceding ten-year period (Franco et al, 2012). There is thus considerable potential for improvement in the application of refinement.

2.3.4. Is species choice a way to reduce harm?

In this final section regarding harm, we will address an idea that is recurrent in the discussion of ethical evaluation and regulation of research: that research will be more or less harmful depending on the animal species chosen (see also chapter 3 of this book). There are sometimes obvious physical justifications, having to do with the size of the animal in relation to the minimum amount of tissue needed for analysis or the minimum size of lesion determined by human dexterity and instruments used – in such cases the smaller the animal the larger the proportional impact will be.

But there is also the widespread idea that animals of different species vary in their capacity for subjective experience. This idea is put forward in the European Directive which requires that if several methods are available one shall choose those that "involve animals with the lowest capacity to experience pain, suffering, distress or lasting harm" (Directive 2010/63/EU, Article 13). This seems to indicate that animals can be different on their capacity to suffer – but no guidelines are given for how to assess this capacity. Smith and Boyd (1991) proposed a systematic method of assessment consisting of a checklist of neuroanatomical/ physiological and behavioral criteria to determine whether a nonhuman animal has the capacity for pain, stress, and anxiety. On the neuroanatomical side the criteria include 1) the possession of receptors sensitive to noxious stimuli (nociceptors), 2) the possession of higher brain centers (especially a structure analogous to the human cerebral cortex), 3) the possession of nociceptors connected to these higher brain structures and 4) the possession of opioid-type receptors. On the behavioral side the criteria include 5) responses to painful stimuli modified by analgesics, 6) avoidance or escape responses to painful stimuli, 7) responses to noxious stimuli that persist and 8) the capacity to associate neutral with noxious stimuli. However, looking at how taxonomically distinct animals used in research fare in this assessment it becomes clear that 1) if we complement the information available in the original 1991 analysis with contemporary knowledge about fish, at least all vertebrate animals meet the criteria for pain and 2) information about non-vertebrate sentience is too limited to allow species to be identified as less sentient with reasonably reliability. That is, it is highly unclear what animal researchers are to choose to ensure "lowest capacity to experience". Colin Allen (2004) proposed to use learning abilities as indicators of capacity to suffer. This would include operant learning - which appears to require a brain (Grau, 2002), and certain kinds of classical conditioning (e.g., trace conditioning). However, this also does not help much to draw the distinction between species. There are no significant differences in the learning abilities between mammal species, and most likely all vertebrates and even some invertebrates would still fall within the same category in terms of their "capacity to experience"[2].

On the other side of the sentience coin, we find the concern that some species might have a *higher* capacity to experience. In a position paper, a European Science Foundation working group argued that non human primates (NHPs) have a greater potential

[2] See Chapter 2 of this book for an extensive discussion on animal consciousness.

for suffering since they are : “distinguished by the very advanced nature of their social, cognitive, sensory and motor functions” (ESF, 2009). A reasonable interpretation of this is that NHPs will be more harmed by research than other laboratory animals. But in which way?

We have analyzed this question in some detail elsewhere (Olsson and Sandøe, 2010). In summary, in terms of capacity for sentience, it is unclear how most NHPs are different from other mammals which also share the capacity for experiencing pain and distress. Capacity for self-awareness may affect potential for suffering, but reasonable evidence to attribute this capacity only exists for great apes. The biological difference with clearest welfare relevance between NHPs and other mammal species used in research seems to be that primate species are not fully domesticated, making it more challenging to meet their needs in captive housing. On the other hand, there are also aspects in which primates may be better off in research than, for example, rodents: primates are usually trained to collaborate rather than restrained, their greater similarity to human beings facilitates the recognition of signs of poor welfare and higher concern for their welfare might encourage scientists to be more careful in how primates are treated.

Very recently, working on an analogy with pediatric research ethics, Fenton (2014) advocated that the cognitive capacities of chimpanzees may allow them to dissent from participating in research. It is not consensual whether chimps fulfill the conditions set, including for example whether they are capable of planning the future - which is far from being consensual (e.g., Shettleworth, 2010; Suddendorf et al., 2009). Furthermore it remains to be seen how to allow such dissent to be expressed in practice in a meaningful way. However, if these hurdles were overcome, this may be an interesting approach to develop a research ethics for non-human primates which not only respects but actually relies on their cognitive capacities.

In summary, with present knowledge there is little support for establishing differences which can motivate species choice to be a useful measure to reduce animal harm. Instead, the differences that society and the research community tend to make between less and more ethically problematic species are best understood in the light of the socio-zoological scale. This scale rates animals in terms of how greatly they are valued by humans, and places companion animal species and non-human primates at the top and rodents, fish and invertebrates quite further down (Arluke and Sanders, 1996). That the socio-zoological scale is based on what humans think about animals rather than on the characteristics of the animals themselves does not mean that it is ethically irrelevant. But, in our opinion, the difference between using a rhesus macaque or a fish in a given experiment is better described as more or less harmful to public sensitivity than more or less harmful to the animal.

3. Benefits

3.1. What are the benefits of neuroscience research?

We now turn to the other side of the harm-benefit equation, the benefits. Overall, research in neuroscience aims to deliver benefits for scientific knowledge and for human health and welfare, that is, benefits for humans. On the one hand, basic research is conducted with the aim of understanding the functioning of the nervous system and the mechanisms underlying the diseases that affect it. On the other hand, applied research is carried out to develop treatments for such disorders.

3.2. How can we quantify benefits?

3.2.1. Assessment of potential benefits

It is very difficult or almost impossible to predict accurately whether a research project will improve our understanding of important mechanisms or lead to the development of therapeutics. Science has a considerable element of unpredictability; even when armed with well-defined hypotheses and carefully executed experiments, it is impossible to guarantee that a research project delivers its intended benefits, in particular if these are defined on the level of the practical impact the study will have in the scientific field or in society. Especially with basic research it is difficult to anticipate the direction of the findings (i.e., whether they will support the researchers’ hypotheses) and the long-term impact of such results for human health and the society. Nevertheless, assessing benefits is fundamental if we are to weigh them against harms in order to justify animal

experimentation. Also, using animals for research with no clear or intended benefit would be unethical virtually for every ethical position.

Official documents provide some, although little, guidance as to how to evaluate benefits. An expert working group set up by the Federation of European Laboratory Animal Science Associations (FELASA, 2005) described and explored a set of principles for how to conduct ethical reviews of laboratory animal use and proposed an outline scheme for the assessment of benefits and harms in scientific projects involving animals. On the benefit side, the questions to be answered included:

- How will the results add to existing knowledge? What practical applications, if any, are envisaged at this stage?
- What is the potential value of these insights and/or applications?
- Are the objectives of the project original, timely and realistic?
- How does the present proposal relate to what was done before? What progress was made in previous studies, and what scientific or other benefits have resulted?
- What is the relevance of this project to other studies in this field of research and what might be the implications for other areas of research, if any?

Similar questions are proposed by other policy documents and reports. Both The Canadian Council on Animal Care guidelines (CCAC, 1997) and the recent Working document on Project Evaluation and Retrospective Assessment, for example, claim for clear statements of the scientific objectives and potential value of the study in terms of originality and importance of the new information, as well as the need for the experimental project.

Many of these issues are challenging to evaluate, to say the least. There are practical challenges having to do with the difficulty of predicting outcomes, but also ethical challenges in terms of judging whether a certain scientific objective is more valuable than another. For which purpose animals are used plays a role in determining acceptability of research in society (see Lund et al, 2012), but it is unclear to what extent this is also reflected in practical decision-making, and official documents guiding ethical review focus more on the assessment of the likelihood that the proposed benefits will be achieved.

3.2.2. Assessment of likelihood that potential benefits will be achieved

At least as long as benefit is understood in terms of knowledge gains, these questions are of a more technical nature and, thus, are easier to evaluate and can be assessed more objectively. They include evaluating 1) the appropriateness of the animal model and the scientific approach, 2) the validity of the experimental design, 3) the staff competence, 4) the appropriateness and quality of facilities and 5) the communication of results (e.g., APC, 2003; FELASA, 2005; Smith and Boyd, 1991). This kind of evaluation will tell whether a proposed study will be able to provide reliable answers to the questions it poses, without making any judgment on the relevance of these questions.

The choice of an appropriate animal model is crucial for a research project to succeed. However, critical analysis of what characterizes a good animal model is curiously rare in the scientific literature. Most review papers on animal models limit themselves to an overview of the existing models in a field and discussion of the results obtained in studies using them. In neuroscience, it is widely recognized that existing animal models are insufficient (e.g., Micale et al, 2013). The animal models of psychiatric diseases have been especially criticized (e.g., Schapiro, 1998; Rollin and Rollin, 2014), mainly due to the impossibility for any animal model to accurately mimic all aspects of mental illnesses, some of which are unique to humans (American Psychiatric Association, 2000). These criticisms have led to greater attention to questions of model validity in neurosciences than in, for example, infectious disease research. In a 2012 NIH workshop aimed to increase the utility and translation of animal models in neuroscience research, experts highlighted the need for improved animal models and better matching between animal model research and the human disease phenotype. Several other joint efforts in the scientific community indicate that there is a heightened awareness of the present shortcomings of animal models in this field. It remains to be seen how well this translates

into better practice. An important consequence of using animal models of unreliable validity is that scientific progress may be retarded and, as a consequence, animals are used unnecessarily.[3]

Regarding the second issue, the experimental design of animal experiments, systematic reviews have recurrently shown significant shortcomings in how studies are designed and carried out in neurosciences. Most of this knowledge comes from extensive systematic reviews and meta-analyses of experimental animal research in acute ischemic attack (stroke). In this field, a wide number of neuroprotective drugs which had proven effective in animal models, later failed to work in clinical trials on humans (van der Worp et al, 2005). In many of these experiments, the efficacy of the treatment was probably overestimated as a result of design bias. Often animals were not randomly allocated to treatments, and researchers who were not blinded when they administered the treatment or assessed the outcome, may unconsciously have influenced the measurements (van der Worp et al, 2005; Crossley et al, 2008). Additionally, there were obvious significant methodological differences between pre-clinical and clinical trials, in that the animals used were generally young and healthy before the experimentally induced stroke, while human patients were often elderly and hypertense (Macleod and Sandercock, 2005). The same sort of shortcomings was also identified in studies on intracerebral hemorrhage (MacLellan et al, 2012).

The likelihood of the proposed benefits being delivered also depends on the competence of the people involved in the project, namely on the experience the researchers have in conducting research in the field or in using the proposed animal model, on their general scientific capacity, and on their level of training in laboratory animal science. The fact that the experiments are conducted in facilities with the necessary and appropriate conditions and equipments is also essential to guarantee that the experiments produce reliable results.

Finally, if the intended benefits of research are to be achieved in practice, the results of the experiments must be made public; hence communication is central. Publication in peer reviewed journals is a central feature of modern academic research and, as is well known, the performance of today's researchers is measured largely on the basis of the number of publications they have in influential journals. However, it is generally difficult to get negative results (no effect of treatment) published. As a direct consequence of this, publications are likely to reflect only part of the research that has been carried out in a field – the research in which differences were found between treatment groups. This has wide-ranging ethical consequences. Importantly, it affects the number of animals used in research. Also, poor concepts, hypotheses, and models may survive, notwithstanding a vast amount of contradictory data, all of this merely because these data are not made available to the scientific community. Publication bias - that positive results are more likely to be published, has indeed been pointed out as a major problem in neuroscience research (Sena et al, 2010).

3.3. How can we improve benefits?

In order to overcome such shortcomings and improve the benefits delivered by animal research, some issues are particularly pertinent.

First of all, more rigorous preclinical study methodology is clearly necessary. A relevant handicap of animal-based neuroscience research has been the difficulty in translating apparent therapeutic successes from in vivo preclinical studies into human clinical trials. Given this remarkably low success rate, some attempts have been made to bring researchers together to discuss how to do preclinical research in a given field in the best way. An example is the European ALS/MND group, which organized two workshops, one in 2006 and one in 2009 (Ludolph et al, 2007; 2010). Another example is the 2011 meeting of the National Institute of Neurological Disorders and Stroke (NINDS), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), the International Rett Syndrome Foundation (IRSF), the Rett Syndrome Research Trust (RSRT), the National Institutes of Health (NIH), the US Food and Drug Administration (FDA), the pharmaceutical industry and private

[3] In this context, it is interesting to note that NIH recently announced it will only fund clinical trials in psychiatry which address underlying biological mechanisms (Reardon, 2014). Although it is too early to tell whether this will affect how and which animal models are used in this field, this measure may contribute further to raising the validity requirements in this field of research.

foundations in the USA (Katz et al, 2012). These meetings aimed to identify the reasons for the translational failure in ALS and Rett's syndrome research, respectively, and to establish guidelines to standardize and improve preclinical research in these fields.

The major barriers to translational success identified in these workshops were a lack of rigorous standards and transparency in reporting preclinical studies, similar to what was reported by der Worp (2005) in ischemia stroke research and by MacLellan et al (2012) in intracerebral hemorrhage research. The subsequent published guidelines from these groups outline some of the principles and standards of good study design and report when conducting preclinical trials of candidate therapeutics – e.g., allocation concealment, blinded assessment of outcome, random allocation of subjects to experimental groups and other methods designed to minimize bias and Type 1 ('false positive') errors. By raising standards and awareness, these initiatives strive to increase the reliability, reproducibility and predictive value of preclinical research, and ultimately improve the likelihood of success on clinical translation.

Research on spinal cord injury has also encountered a low translational rate. Results of many treatment strategies showed some beneficial effects in animals on the anatomical/ histological and/or the functional level, but subsequent pilot clinical trials failed or were abandoned. However, in this case, the problem does not appear to result from experimental design shortcomings. A major reason for such translational failure, which has been less discussed, may derive from the high control of variability in the animal experiments (Fili and Schwab, 2012). Animal studies are designed to minimize all variables within an experiment (age, sex, and strain of animals; extent and location of the lesion; application of medication, etc), which is in contrast to the clinical situation, where large heterogeneities exist at all levels and thus may hamper the detection of potential treatment effects. This implies that beneficial effects of a given treatment in animal models have to be robust, functionally meaningful, and reproducible in independent laboratories. Minor functional improvements as a consequence of a novel therapy or effects at only the anatomical/histological level are unlikely to materialize as detectable functional benefits in clinical trials, where the patient population is more heterogeneous and the assessment of functional outcome is even more difficult (Fili and Schwab, 2012). In summary, improving the standards for considering therapeutic treatments as ready for clinical trials would be likely to make the use of animals more profitable and increase the translational relevance of animal models.

Secondly, we believe that more rigorous and extensive evaluations of animal models are necessary. As mentioned above, there are fields where it is widely recognized that existing animal models are insufficient. As stated by Micale et al (2013) regarding neuropsychiatry, "The field desperately needs better animal models of depression and SCZ [schizophrenia] because of the partial efficacy of present pharmacological treatment. Without improved models of human disease, we cannot know whether particular molecular and cellular findings in animals are relevant to the clinical situations". We propose that, for improved ethical animal experimentation, animal research in these situations should be avoided until improved models are developed. It is not clear that in these cases the benefits of research justify the harms caused to the animals.

Lastly, publication of negative findings from well-conceived and performed studies should be encouraged. It can help investigators to evaluate and ultimately abandon the development of invalid and irrelevant animal models, which may be hampering the progress of neuroscience. Importantly, this will result in efforts for the development of good animal models and their validation, which will improve the likelihood that the benefits of research are delivered.

4. Weighing harms and benefits

Harm-benefit assessment is widely regarded as the standard approach for assessing the ethical justification of experiments within the framework of regulations. It is advanced in several guidelines and recommendations for ethical review as the way to determine the acceptability of animal experiments (e.g., APC, 2003; CCAC, 1997; FELASA 2005) and it is also explicitly required in legislation in many countries (e. g., EU through Directive 2010/63 and Australia through Australian Animal Welfare Act 1999).

4.1. Committees

The decision to authorize a research project with animals is usually based on some form of assessment of the relative harms and benefits, with that evaluation, in most countries carried out by a review committee. Such review may take place at an institutional, regional or national level, with different countries using different organizations. Animal Ethics Committees (AECs) - or Institutional Animal Care and Use Committees (IACUCs) as they are called in the USA and Canada, thus have the major responsibility for the performance of these harm-benefit assessments. These committees may have an advisory (e.g. UK) or regulatory role (e.g. Canada). Their function is to ensure that animal experiments are only performed when they are ethically justified, i.e., that the impact on animal welfare is minimized and that the proposed benefits of the research are likely to be achieved.

Typically, review committees are composed of a number of people representing different competences, including scientists, animal technicians/caretakers, veterinarians, non-animal researchers, animal welfare and/or animal ethics experts and lay people. Such composition is, in many countries, effectively mandated by law, including the USA, Australia and European countries such as Sweden, Germany, Denmark and Switzerland. The majority of members belong to the scientific community and in the case of institutional committees they also typically belong to the institution. Non-institutional members and lay people constitute a minor percentage of the committees' composition. The largest representation seems to be in Sweden, where half of the committee members are non-scientists (laypersons or representatives of animal protection NGOs). The Australian Code currently states that at least one third of committee members should be laypersons or representatives of animal welfare groups. In the USA, according to the legislation,"... at least one member... shall not be affiliated with [the research] facility other than as a member of the Committee." Moreover, the unaffiliated member may not be a close relative of a person affiliated with the facility. The idea of such composition is, by having members that are external to the institution, to avoid potential conflicts of interests in project evaluations and, by having lay people, to provide representation for general community interests and to ensure that institutional committees go beyond a purely scientific analysis in making decisions on laboratory animal care and use (Dresser, 1999).

There has been some research into how this committee review system works. Dresser (1989) surveyed the responses of 32 IACUCs when asked to review four hypothetical protocols. Results suggested that that committees showed little consensus in their approach to assessing the justification for animal use, although there was broad agreement on the need to refine particular procedures. Forsmann (1993) analyzed the written reports from the regional committees in Sweden between 1979 and 1989 and found that they acted primarily as technical committees, with refinement of procedures being the predominant focus of de-liberations. Hagelin et al (2002) found that between 1989 and 2000 the practice had not changed in Sweden: the majority of the modifications continued to relate to the goals of refinement. However, the authors concluded that the AECs were being effective in improving animal welfare. Houde et al (2003) observed three Canadian IACUCs over a 1-year period and the data revealed that most comments were of a technical nature, with 16% related to what had been defined as the "explicit ethical categories" of the 3Rs. In interviews with members of Swedish AECs, Ideland (2009) found that they focused on methodological improvements instead of on weighing research aims against animal suffering. More recently, Schuppli (2011) conducted an ethnographic study involving participant observation and in-depth interviews with members of university AECs in Canada and also found that the major focus of protocol review was reducing harms to animals, with some members stating that assessing ethical justification was the role of scientific peer review. In summary, these studies showed that the focus of committees was mainly reducing harms to animals, with less importance given to the justification of research and evaluation of potential benefits, despite this being stressed in policy as a goal of AECs.

Another focus of research on ethics committees has been the degree of agreement in the decision-making process both between and within committees. In her 1989 study in which 32 IACUCs evaluated four animal research protocols, Dresser found low levels of inter-committee agreement. More recently, Plous and Herzog (2001) evaluated both inter and intra-committee agreement. The authors randomly selected IACUCs from US universities and colleges. Next, each IACUC was asked to submit its three most

recently reviewed protocols involving animal behavior. Each protocol was randomly assigned to be reviewed a second time by a participant of another IACUC. Once they had received the reviews from individual committee members, the IACUCs were asked to meet as a group and render a final evaluation for each of the three protocols. The authors found a low level of agreement both between and within committees. Although strong conclusions and generalizations cannot be drawn only from these two qualitative studies, these results are important indicators that the review process performed by ethics committees may be working with low levels of agreement.

4.2. The challenge of harm-benefit weighing

The studies referred to above indicate two major problems in the process of ethical review. On the one hand, it appears that committees are neglecting the assessment of benefits and their weighing against harms and that they are focusing mainly on reducing harms. On the other hand, the reliability of the ethical review process appears questionable, as studies pointed to a poor agreement in protocol evaluation both by the members of the same committee and by different committees. Despite this general low agreement, some aspects appear to gather consensus, namely the evaluation of animal suffering (Plous & Herzog, 2001) and the evaluation and improvement of technical and methodological issues (Ideland, 2009). Hence, there is some suggestion that problems in the reliability of the ethical review process appear to result mainly from the lack of agreement on benefit evaluation and harm-benefit weighing.

One explanation for the quasi-absence and the lack of agreement on benefit assessment and harm-benefit weighing may be the fact that guidelines for ethical review are very general, especially as regards how to evaluate benefit, or how to weigh it against harms (e.g., Directive 63/2010/EU, IACUC Guidebook, CCAC Guidelines). Hence, they are not very helpful to these respects. Stafleu (1994) further suggested that ethics committee members have probably little notion of how they are to balance the significance of research against the interests of the animals. In support of this view, more recently, Schuppli (2011) found a large variation in how harm-benefit analysis was applied by ethics committee members. It seems probable that this is due to a general absence of concrete guidance in the guidelines as to how to evaluate benefit and how to weigh it against harm. Of the two elements in the harm-benefit assessment, the question of harm (or more precisely, how to reduce it) has been central to research and teaching in laboratory animal science worldwide over the last couple of decades and has thus had a longer history of research and greater contextual framework underpinning its assessment. In fact, harms may be easier to identify and define than the benefits (Voipio, 2004). What we currently know about the mechanisms of pain in vertebrates and of their physiological and behavioral needs allows us to make relatively objective predictions of the impact of research on them. Guidelines and examples are commonly provided to help bring some uniformity to harm assessment (Orlans, 1996). Classification systems for ranking the degree of animal pain and distress have been elaborated for ethical reviews (e.g., Directive 63/2010/EU; CCAC, 1997). The fact that these are more immediate and explicitly visible may also make it easier in any evaluation to assess harms, rather than benefits which are more general and further away in time.

In the meantime, a few ethical schemes or scoring systems have been proposed to help in the ethical review process (e.g., Bateson, 1986; Delpire et al, 1999; Smith & Boyd, 1991; Porter 1992; Stafleu et al, 1999). These systems are meant to support the process of assessing and weighing the various considerations that determine the ethical acceptability of an animal experiment and to help people to do so in a systematic way. In an attempt to make them as objective as possible, some of these schemes are based on mathematical calculations (e.g., Stafleu et al, 1999; Porter, 1992). These have a series of categories to which scores are attributed. These scores are then incorporated in a computational rule, whose result, when compared to a cut-off value, determines whether an experiment is or is not ethically acceptable. In Stafleu's scheme, for example, scores are calculated for the human interests, the relevance of the animal experiment and the harms caused to the animals. To calculate a score for each one of these, several points are considered. For example, for calculating a score for the human interest, assessors are advised to consider whether the benefits are for human health, for scientific knowledge or for economy, with each one deserving a different weight. For calculating a score for the relevance of the animal experiment, issues such as the

methodological quality, the quality of the research group and the necessity of the experiment are scored. For the harm score, the duration and intensity of discomfort, the number of animals involved and their psychological complexity are considered. In the end, the score for human interests is multiplied by the score for the relevance of the animal experiment and this value is compared with the harm score for animals. This will determine the acceptability of the experiment. If the former is higher than the latter, the experiment is acceptable, otherwise it is unacceptable.

However, not everyone supports the use of schemes based on detailed algorithms. The argument is that they try to imply accurate measures where there can be none - harms and benefits are not quantifiable in this way. This idea was widely defended in 2009, at a workshop organized by the Cooperation Group for Laboratory Animal Sciences within the Finnish Ministry of Education. This workshop gathered participants representing the scientific community, animal welfare organizations and regulators from Nordic and Baltic countries and The Netherlands. A classification of harms and benefits into three degrees (low, medium and high; e.g., Bateson, 1986) was preferred over the scoring systems (Voipio, 2004). In Bateson's model, the probability of benefits, the quality of research and the animal suffering are attributed a classification of low, medium or high. If these three dimensions are then considered as the three dimensions of a cube, there will be a zone where the probability of benefits and the quality of research will be high and the animal suffering will be low – if a research proposal falls into this zone it is ethically acceptable, and there will be a zone where the probability of benefits and the quality of research will be low and the animal suffering will be high – if a research proposal falls into this zone it is ethically unacceptable and should not be approved. A revised version of the Bateson model is proposed in the most recent guidelines prepared by the European Expert Working Group for Project Evaluation for how to evaluate projects under Directive 2010/63/EU.

Finally, the fact that AECs are not assessing benefits may have to do with some confusion over the relation between AEC review and scientific peer-review by granting agencies – some committee members believe that ethical justification is decided by scientific peer-review (e.g., Graham, 2002; Schuppli, 2011).

Overall, there seem to be a discrepancy between the intended content of an ethics review and what is currently being evaluated in actual practice by animal ethics/ animal care and use committees. Committees spend much effort on evaluating issues related to animal harm, and on considerations of how to reduce said harm (by way of applying the 3Rs). Benefit, on the other hand, is not always addressed and when it is, there is little agreement between committees and between individual members. We will end this section by reflecting on what can be done to change this discrepancy.

There are several possible ways to improve how benefit is evaluated and weighed against harm. A first possibility for improvement would be the development of more detailed guidelines for ethical review. Specifically, more detailed classification systems for benefits and harms might help. A good example is the report recently produced by the European Commission Expert Working Group. The fact that Plous and Herzog (2001) found a relatively high intra-committee agreement when the committee members used a pain scale to rate animals' pain or stress, suggests that when committee members are given detailed classification criteria they may achieve a relatively high degree of agreement. Elsewhere it was already advanced that the use of detailed classification systems is essential for careful ethical analysis (Orlans, 1997). A second way ethical review effectiveness could be improved would be by clarifying the roles of AECs and scientific peer-review. As referred above, AECs may not be assessing benefits and weighing them against harms because they believe it is not their role. Finally, giving the committee members enhanced training in ethics and/or philosophy could also help, as suggested by Houde et al (2009) and Schuppli and Fraser (2005).

However, another possibility would be to change the aims of ethics review. Based on the variety of problems resulting from the evaluation of benefit and from weighing benefit against harm, we wonder whether benefit assessment should not be openly (and expressly) discarded and whether ethical evaluation should not focus on applying the three Rs and thus upon improving animal welfare only. For all intents, this largely seems to be what is currently happening anyway. An interesting comparison can be made with farm animals: While broad societal concern exists around the intensive rearing of animals for human consumption, regulation quite strictly focuses on the improvement of animal welfare or, in other words, the reduction of harms caused to the

animals, and the harms for the animal are not weighed against the benefits. Is there a significant reason for demanding a different approach in laboratory animals?

Another possibility would be to restrict full harm-benefit analyses to experiments likely to cause) severe harm to the animals, though this implies that in all other situations the benefit would be considered to outweigh the harm without explicit reflection. Still, since it is such a difficult exercise, we could reserve such explicit reflection for those situations where animal suffering appears so great that a strong justification is needed for the research to be ethically acceptable.

5. Summary

- Studies on live animals play a crucial role in neuroscience research, but their use as research subjects is a controversial issue.
- Different ethical theories have different positions as regards the morality of using animals in research. Contractarianism, utilitarianism and animal rights are three ethical perspectives that represent a great proportion of the debate.
- Harm-benefit weighing, a utilitarian-based perspective, is currently the mainstream approach for determining the acceptability of animal experiments in practice and it is prescribed in legislation worldwide: If the benefits outweigh the harms, the research is considered justified.
- Harms caused to animals in neuroscience research include effects on animal health but also adverse subjective experiences. These harms can range, among others, from the loss of capacities to behavioral deficits, to pain and anxiety.
- Harm assessment is one part of the ethics evaluation of animal experiments. Guidelines and policy documents have advanced lists of criteria to be assessed. Usually, these criteria include the quality of the facilities, the experience of the personnel, the number of animals, the animal species, the husbandry and housing conditions, the scientific procedures themselves (including the killing method), the fate of the animals at the end of the experiments and the endpoints to be applied.
- The harms inflicted to research animals have implications not only for the well-being of the animals, but also for the quality of experimental data. There are thus two major reasons for reducing harm to animals in research: to preserve animal welfare and to guarantee the cleanness of experimental data.
- The Three Rs (replacement – replacing animals with non-animal alternatives, reduction – reducing the number of animals used, refinement – adjusting experimental procedures to minimize any pain or distress) are widely recognized principles for minimizing harms to animals.
- A recurrent idea when considering the harms to animals is that research will be more or less harmful depending on the animal species chosen, which is based on the idea that animals of different species vary in their capacity for subjective experience. We argue that this claim is of limited validity in distinguishing between different vertebrate species, and that the moral difference between using different animals is better described as an issue of public sensitivity than as an issue of animal harm.
- Overall, the benefits of neuroscience research encompass increased scientific knowledge and the improvement of human health and welfare.
- Official documents provide some, although little, guidance as to how to evaluate benefits. Usually it includes statements of the scientific objectives and potential value of the study in terms of originality and importance of the new information, timeliness and the need for the experimental project. However, it is very difficult or almost impossible to assess what will be the benefits of a given research project
- Besides harm and benefit assessment, ethical reviews also involve the assessment of the likelihood that the potential benefits will be achieved. This involves evaluating the appropriateness of the animal model and the scientific approach, the validity of the experimental design, the staff competence, the appropriateness and quality of facilities and the communication of results.
- Benefits can be improved in neuroscience research through more rigorous preclinical study methodology, more rigorous and extensive evaluations of animal models and publication of negative results.
- There seems to be a discrepancy between what the ethics review is said to consist in and what Animal Ethics Committees (AECs) actually evaluate. Committees spend much effort on evaluating issues having to do with animal harm and how to reduce

it (3Rs), but benefit is not always addressed and when it is, there is little agreement between committees and between individual members.

- Explanations for this discrepancy may include the fact that guidelines for ethical review are very general, especially as regards benefit assessment and harm-benefit weighing, and the fact that there is some confusion over who bears responsibility for benefit assessment - AECs or scientific peer-review.

- Improvement of ethical review may be achieved by the development of more detailed guidelines, by clarifying the roles of AECs and scientific peer-review and giving the committee members enhanced training in ethics and/or philosophy. Another possibility would be to change the aims of ethics review. Benefit assessment could be discarded and ethical evaluation could focus only on applying the three Rs and improving animal welfare, or AECs could be required to evaluate benefits and perform harm-benefit analysis only when the proposed experiments appear to be highly severe for the animals.

Acknowledgments

We wish to thank the editors of this book for the challenge to write this chapter, the opportunity to publish it and for constructive comments on previous versions of the text. This work was funded by FEDER funds through the Operational Competitiveness Programme – COMPETE and by National Funds through FCT – Fundação para a Ciência e a Tecnologia under the projects FCOMP-01-0124-FEDER-037277 (PEst-C/SAU/LA0002/2013) , FCOMP-01-0124-FEDER-029527 (PTDC/MHC-ETI/4890/2012) and Projeto "NORTE-07-0124-FEDER-000001 - Neurodegenerative disorders" cofinanciado pelo Programa Operacional Regional do Norte (ON.2 – O Novo Norte), ao abrigo do Quadro de Referência Estratégico Nacional (QREN), através do Fundo Europeu de Desenvolvimento Regional (FEDER) e por fundos nacionais através da FCT - Fundação para a Ciência e Tecnologia.

REFERENCES

1. Abdi S, Haruo A, Bloomstone J (2004) Electroconvulsive therapy for neuropathic pain: a case report and literature review. Pain Physician 7: 261–263.
2. Allen, C. (2004) Animal pain. Noűs 38: 617-643.
3. APA (2000) American Psychiatric Association: diagnostic and statistical manual of mental disorders, 4th ed. Washington DC.
4. APC (2003) Review of Cost-Benefit Assessment in the Use of Animals in Research. Report of the Cost-benefit Working Group of the Animal Procedures Committee. London: Home Office, Communication Directorate.
5. Arluke A, Sanders CR (1996) Regarding animals. Temple University Press, Philadelphia.
6. Australian Animal Welfare Act. 1999. (accessed online at http://www.legislation.govt.nz/act/public/1999/0142/latest/DLM49664.html)
7. Baker M, Kale R, Menken M (2002) The wall between neurology and psychiatry. BMJ 324:1468-1469.
8. Bateson P (1986) When to experiment on animals. New Scientist 109: 30-32.
9. Bayne K, Howard BR, Kurosawa TM, Nájera MEA (2011) An overview of global legislation, regulation and policies. In: Hau J and Schapiro SJ (eds) Handbook of Laboratory Animal Science: Essential Principles and Practices, vol 1, 3rd edn. Boca Raton, FL, pp 39-64.
10. Boomkamp SD, Riehle MO, Wood J et al (2012) The Development of a Rat In Vitro Model of Spinal Cord Injury Demonstrating the Additive Effects of Rho and ROCK Inhibitors on Neurite Outgrowth and Myelination. GLIA 60: 441–456.
11. Button KS, Ioannidis JPA, Mokrysz C et al (2013) Power failure: why small sample size undermines the reliability of neuroscience. Nature Reviews Neuroscience 14, 365-376.
12. CCAC (1993) Guide to the Care and Use of Experimental Animals. Ottawa, Canada: Canadian Council on Animal Care.
13. CCAC (1997) CCAC Guide to Protocol Review. Ottawa, Canada: Canadian Council on Animal Care.
14. Cohen C (1997) Do animals have rights? Ethics and Behavior, Harvard University.
15. Crossley NA, Sena E, Goehler J et al (2008) Empirical evidence of bias in the design of experimental stroke studies—a metaepidemiologic approach. Stroke 39 (3): 929–934.
16. Delpire VC, Mepham TB, Balls M (1999) A proposal for a new ethical scheme addressing the use of laboratory animals for biomedical purposes, Alternatives to Laboratory Animals 27: 869-881.
17. Dresser R (1989) Developing standards in animal research review. Journal of the American Veterinary Medical Association 194: 1184-1191.
18. Dresser R (1999) Community Representatives and Nonscientists on the IACUC: What Difference Should It Make? ILAR Journal 40 (1): 29-33.
19. ESF-EMRC Position on the Proposal for a Directive on the Protection of Animals used for Scientific Purposes — 2nd Edition March 2009 (accessed February 2014 at http://www.esf.org/fileadmin/Public_documents/Publications/AnimalsProtection_2ndEd.pdf).
20. European Union. 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 No. L 276/33.
21. Fawcett JW, Asher RA (1999) The glial scar and central nervous system repair. Brain Res Bull 49:377–391.
22. FELASA (2005) Principles and practice in ethical review of animal experiments across Europe: A report prepared by the FELASA working group on ethical evaluation of animal experiments.
23. Fenton A (2014) Can a chimp say “no”? Reenvisioning chimpanzee dissent in harmful research. Cambridge Quarterly of Healthcare Ethics 23: 130-139.
24. Fili L, Schwab ME (2012) The Rocky Road to Translation in Spinal Cord Repair. Ann Neurol 72:491–501.

25. Forsman B (1993) Research ethics in practice: The Animal Ethics Committees in Sweden 1979-1989. In: The Royal Society of Arts and Sciences in Gothenburg (eds.), Studies in Research Ethics No. 4. Goteborg, Sweden: Centre for Research Ethics.

26. Franco NH, Olsson IAS (2012) How sick must your mouse be? An analysis of Huntington's disease research using animal models. ATLA – Alternatives to Laboratory Animals 40:271-83.

27. Galvin SL, Herzog HA (1992) The ethical judgement of animal research. Ethics and Behavior 2: 263-286.

28. Gilron I, Watson CP, Cahill CM, Moulin DE (2006) Neuropathic pain: a practical guide for the clinician. CMAJ 175:265–75.

29. Graham K (2002) A study of three IACUCs and their views of scientific merit and alternatives, Journal of Applied Animal Welfare Science 5: 75-81.

30. Grau, J (2002) Learning and Memory without a Brain. In: Bekoff M, Allen C and Burghardt GM (eds) The Cognitive Animal. Cambridge, MA: MIT Press, pp 325-334.

31. Hagelin J, Hau J, Carlsson HE (2003) The refining influence of ethics committees on animal experimentation in Sweden, Laboratory Animals 37: 10-18.

32. Handy CR, Krudy C, Boulis N, Federici T (2011) Pain in Amyotrophic Lateral Sclerosis: A Neglected Aspect of Disease. Neurology Research International.

33. Harding EJ, Paul ES, Mendl M (2004) Animal behaviour: cognitive bias and affective state. Nature 427:312.

34. Houde L, Dumas C, Leroux T (2009) Ethics: Views from IACUC members. ATLA 37: 291-296.

35. Ideland M (2009) Different views on ethics: how animal ethics is situated in a committee culture, Journal of Medical Ethics 35: 258-261.

36. International Animal Research Regulations, Impact on Neuroscience Research, Workshop Summary, Institute of Medicine (US); National Research Council (US), Washington (DC): National Academies Press (US); 2012.

37. Katz DM, Berger-Sweeney JE, Eubanks JH et al (2012) Preclinical research in Rett syndrome: setting the foundation for translational success. Disease Models & Mechanisms 5: 733-745.

38. Langford DJ, Bailey AL, Chanda ML et al (2010). Coding of facial expressions of pain in the laboratory mouse. Nature Methods 7: 447 – 449.

39. Langley CK, Aziz Q, Bountra C et al (2008) Volunteer studies in pain research--opportunities and challenges to replace animal experiments: the report and recommendations of a Focus on Alternatives workshop. Neuroimage 42(2):467-73.

40. Lever TE, Gorsek A, Cox KT et al (2009) An animal model of oral dysphagia in amyotrophic lateral sclerosis. Dysphagia 24: 180-195.

41. Lever TE, Simon E, Cox KT et al (2010) A mouse model of pharyngeal dysphagia in amyotrophic lateral sclerosis. Dysphagia 25: 112-126.

42. Lipska B (2004) Using animal models to test a neurodevelopmental hypothesis of schizophrenia. J Psychiatry Neurosci 29(4):282-286.

43. Ludolph AC, C Bendotti, E Blaugrund et al (2007) Guidelines for the preclinical in vivo evaluation of pharmacological active drugs for als/mnd: Report on the 142nd enmc international workshop. Amyotroph Lateral Scler 8: 217-223.

44. Ludolph AC, C Bendotti, E Blaugrund et al (2010) Guidelines for preclinical animal research in als/mnd: A consensus meeting. Amyotrophic Lateral Sclerosis 11: 38-45.

45. Lund, TB, Mørkbak_MR, Lassen J et al (2012) Painful dilemmas: A study of the way the public's assessment of animal research balances costs to animals against human benefits. Public Understanding of Science 0:1-17.

46. MacLellan CL, Paquette R, Colbourne F (2012) A critical appraisal of experimental intracerebral hemorrhage research. Journal of Cerebral Blood Flow & Metabolism 32: 612–627.

47. Macleod M, Sandercock P (2005) Systematic reviews improve clinical research design – can they help improve animal experimental work? RDS News Winter: 8–10.

48. Mendl M, Burman OHP, Parker RMA, Paul ES (2009) Cognitive bias as an indicator of animal emotion and welfare: emerging evidence and underlying mechanisms. Applied Animal Behavior Science 118:161-181.

49. Micale V, Kucerova J, Sulcova A (2013) Leading compounds for the validation of animal models of psychopathology. Cell Tissue Res. 354:309–330.

50. Morton DB, Hau J (2011) Welfare assessment and humane endpoints. In: Hau J and Schapiro SJ (eds) Handbook of Laboratory Animal Science: Essential Principles and Practices, vol 1, 3rd edn. Boca Raton, FL, pp 535-572.

51. Nuffield Council on Bioethics (2005) The ethics of research involving animals. London.

52. OLAW (2002) Institutional Animal Care and Use Committee Guidebook. USA.

53. Olsson IAS, Franco NH, Weary DM, Sandøe P (2011) The 3Rs principle: Mind the ethical gap! Altex 29, Special Issue. Proceedings of WC8. (http://www.altex.ch/en/index.html?id=90)

54. Olsson IAS, Sandøe P (2010) "What's wrong with my monkey?" An ethical perspective on germline transgenesis in marmosets. Transgenic Research 19: 181-186

55. Orlans, FB (1996) Invasiveness scales for animal pain and distress. Lab Animal, 25(6): 23-25.

56. Orlans, FB (1997) Ethical decision making about animal experiments, Ethics and Behavior 7:163-171.

57. Plous S, Herzog, H (2001) Reliability of protocol reviews for animal research, Science 293: 608-609.

58. Porter DG (1992) Ethical scores for animal experiments, Nature 356: 101-102.

59. Prentice ED, Crouse DA, Mann MD (1992) Scientific merit review: the role of the IACUC, Institute for Laboratory Animal Research News 34: 15-19.

60. Reardon S (2014) NIH rethinks psychiatric trials. Nature 507: 288.

61. Regan T (1983) The case for animal rights. University of California Press.

62. Richardson CA, Flecknell PA (2005) Anaesthesia and post-operative analgesia following experimental surgery in laboratory rodents: are we making progress? Altern Lab Anim 33(2):119-27.

63. Rollin MDH, Rollin BE (2014) Crazy like a fox: validity and ethics of animal models of human psychiatric disease. Cambridge Quarterly of Healthcare Ethics 23: 140-151.

64. Rudge JS, Silver J (1990) Inhibition of neurite outgrowth on astroglial scars in vitro. J Neurosci 10:594–3603.

65. Russell W, Burch R (1959) The principles of humane experimental technique. London: Methuen.

66. Schapiro KJ (1998) Animal models of human psychology: Critique of science, ethics and policy. Seattle, WA: Hogrefe & Huber.

67. Schuppli CA (2011) Decisions about the use of animals in research: ethical reflection by Animal Ethics Committee Members. Anthrozoos: A Multidisciplinary Journal of The Interactions of People & Animals, 24(4): 409-425.

68. Schuppli CA, Fraser D (2005). Factors influencing the effectiveness of research ethics committees. J Med Ethics 33(5): 294–301.

69. Sena ES, van der Worp HB, Bath PMW et al(2010) Publication Bias in Reports of Animal Stroke Studies Leads to Major Overstatement of Efficacy. Plos Biology 8(3):1-8.

70. Shettleworth SJ (2010) Clever animals and killjoy explanations in comparative psychology. Trends in Cognitive Sciences 14(11): 477-481.

71. Singer P (1975) Animal Liberation. New York: Avon Books.

72. Smaje LH, JA Smith, R Ewbank et al (1998) Advancing refinement of laboratory animal use. Laboratory Animals 32: 137-142.

73. Smith, J.A. and Boyd, K.M. 1991. Lives in Balance. The Ethics of Using Animals in Biomedical Research. Oxford, UK: Oxford University Press.

74. Solomon JA, Tarnopolsky MA, Hamadeh MJ (2011) One universal common endpoint in mouse models of amyotrophic lateral sclerosis. PLoS One 6: e20582.

75. Sorensen A, Moffat K, Thomson C, Barnett SC (2008) Astrocytes, But Not Olfactory Ensheathing Cells or Schwann Cells, Promote Myelination of CNS Axons In Vitro. GLIA 56:750–763.

76. Stafleu FR, Tramper R, Vorstenbosch J, Joles JA (1999) The ethical acceptability of animal experiments: a proposal for a system to support decision-making, Laboratory Animals 33: 295-303.

77. Stafleu FR (1994) The ethical acceptability of animal experiments as judged by researchers (PhD Thesis). Utrecht University.

78. Suddendorf T, Corballis MC, Collier-Baker E (2009) How great is great ape foresight? Animal Cognition 12(5): 751-754.

79. van der Staay FJ, SS Arndt, RE Nordquist (2009) Evaluation of animal models of neurobehavioral disorders. Behav Brain Funct 5: 11.

80. Van der Worp HB, de Haan P, Morrema E et al. (2005) Methodological quality of animal studies on neuroprotection in focal cerebral ischaemia. J Neurol 252:1108–1114.

81. Voipio H-M, Kaliste E, Hirsjärvi P et al (2004) Nordic-European Workshop on Ethical Evaluation of Animal Experiments. Scand. J. Lab. Anim. Sci. 31,251–267.

82. Woolf CJ (2004) Dissecting out mechanisms responsible for peripheral neuropathic pain: implications for diagnosis and therapy. Life Sci 74:2605–10.

83. Willner P (1997) Validity, reliability and utility of the chronic mild stress model of depression: a 10-year review and evaluation. Psychopharmacology 134: 319-329.

84. Yalcin I, Bohren Y, Waltisperger E et al (2011) A Time-Dependent History of Mood Disorders in a Murine Model of Neuropathic Pain. Biol Psychiatry 70: 946–953.

85. Yeates JW (2010) Death is a welfare issue. J Agric Environ Ethics 23:229–241.

86. National Competent Authorities for the implementation of Directive 2010/63/EU on the protection of animals used for scientific purposes: Working document on Project Evaluation and Retrospective Assessment (accessed January 2014 at http://ec.europa.eu/environment/chemicals/lab_animals/pdf/examples.pdf)