The Humane Society Institute for Science and Policy

2019

Refinement on the Way Towards Replacement: Are We Doing What We Can?

Kathrin Herrmann
Johns Hopkins Bloomberg School of Public Health

Refinement on the Way Towards Replacement: Are We Doing What We Can?

Kathrin Herrmann
Freie Universität Berlin, Department of Veterinary Medicine, Institute of
Pharmacology and Toxicology, Berlin, Germany; current address: Johns
Hopkins University, Bloomberg School of Public Health,
Baltimore, MD, United States
kherrma1@jhu.edu

1 Introduction

> [R]efinement is never enough, and we should always seek further for
> reduction and if possible replacement.
>
> RUSSELL and BURCH, 1959, Chapter 4

Russell and Burch introduced the principles of replacement, reduction, and refinement of animal experimentation in 1959 in their groundbreaking book, *The Principles of Humane Experimental Technique,* to eradicate inhumanity towards non-human animals (hereinafter referred to as animals). They utilized the term *inhumanity* to indicate negative mental states experienced by animals used in research and the procedures that cause such mental states. Their goal was to avoid the use of animals wherever possible and to improve significantly the treatment of the animals still deemed indispensable, while improving the quality of scientific and medical research and testing (Russell and Burch, 1959). Since the 1990s, the 3Rs have slowly gained more acceptance within the animal research community. They have been recognized by organizations such as the Council of Europe (1986) and the World Organisation for Animal Health (2018), and they have been implemented in law in several countries, for example in Germany and in the UK (Herrmann, Köpernik and Biedermann, 2009; Zurlo, Rudacille and Goldberg, 1996).

Today, the principles are not only embedded in legislation in the European Union (EU) but around the world (Bayne et al., 2015). In the EU, Directive 2010/63/EU on the protection of animals used for scientific purposes came into effect in 2013, thereby requiring all EU Member States to implement the

3Rs fully. The EU Directive is more far-reaching compared to other legislation since it promotes a strong shift away from animal experimentation, with its goal being "full replacement of procedures on live animals for scientific and educational purposes as soon as it is scientifically possible" (European Parliament, 2010, Recital 10). Furthermore, the EU Directive mandates that replacement should be the first priority, followed by reduction and then refinement to be implemented if animal use is deemed absolutely unavoidable (European Parliament, 2010, Recital 11). Russell and Burch (1959, Chapter 7) proposed the following hierarchy: "Suppose, for a particular purpose, we cannot use replacing techniques. Suppose it is agreed that we shall be using every device of theory and practice to reduce to a minimum the number of animals we have to employ. It is at this point that refinement starts, and its object is simply to reduce to an absolute minimum the amount of distress imposed on those animals that are still used."

As a result of the incorporation of the 3Rs into legislation, which has mainly been driven by ever-increasing societal concerns (cf. Clemence and Leaman, 2016; European Citizens' Initiative, 2016; Jones, 2017; Pew Research Center, 2015, 2018), it would seem reasonable to expect changes within the research industry, particularly replacement of animals with non-animal models. However, the cumulative effect of any such replacements has not prevented the overall number of animals used from steadily increasing since the 2000s (European Commission, 2013; Taylor et al., 2008; Taylor and Rego, 2016). When looking at the 3Rs and their impact, it seems that refinement, the R of *ultima ratio*, is receiving the most attention by the laboratory animal science community (AALAS, n.d.; FELASA., 2016), especially in basic and applied research where the majority of animals are utilized (in the EU, 65% of animals; cf. Daneshian et al., 2015). A survey conducted with participants of laboratory animal science training courses in four European countries found that refinement was seen as more feasible and more pressing than replacement and reduction of animal use (Franco, Sandøe and Olsson, 2018).

Due to this focus, the chapter starts by exploring the application of several refinement methods in practice, commencing with current housing and husbandry standards and a discussion about the benefits of a "culture of care", followed by assessing important experimental refinements. To further assess the quality of animal-based research, it reviews necessary refinements in planning, conduct, and reporting practices of animal studies. The chapter then moves on to look at feasible ways to reduce and replace animal use by, first discussing tools to appraise animal studies whose application could lead to a significant reduction of animal experiments and thus numbers of animals used. It subsequently reflects on what the scientific community has

been doing to move towards replacement of animals in research, testing, and education. Finally, the chapter concludes with recommendations for steps to be taken to work towards using non-animal, human-relevant approaches to biomedical research and testing aimed to protecting human health.

2 Refinement of Animal Housing and Husbandry

> Husbandry is a factor for contingent inhumanity in all types of experiment.
> RUSSELL and BURCH, 1959, Chapter 4

Animals used in research, testing, and education spend their lives in a captive environment that is very different from their natural environment. Refined housing gives animals the opportunity to cope with some of the stressors imposed by life in the laboratory (Mason, 2006). Improving their living conditions by trying to meet some of the animals' basic behavioral needs is called *environmental refinement* or *environmental enrichment* (EE). Krech, Rosenzweig and Bennett (1960) were the first to report biochemical changes in the brains of rats kept in a complex housing environment and augmented with daily exposure to novel items in an open field. They coined the term EE when describing this paradigm (Benefiel, Dong and Greenough, 2005). Environmental Enrichment is defined as "[a]ny modification in the environment of captive animals that seeks to enhance the physical and psychological well-being of the animals by providing stimuli which meet the animals' species-specific needs" (Baumans and van Loo, 2013). It includes complex social and inanimate object stimulation (Rosenzweig, 1966). Its positive behavioral effects were first described in rats by Hebb in 1947, who kept them as companion animals in his home. He observed that the rats living in a more complex, stimulating environment learned better and more quickly (Hebb, 1947). In addition to enhancing cognition, EE also promotes neuronal activation, signaling and plasticity in a number of brain regions (Nithianantharajah and Hannan, 2006). In the beginning, research on EE was conducted primarily to assess changes in behavior and brain development. With the increased concern for animal welfare and the establishment of animal welfare science as a specific discipline, has EE been applied to improve the animals' daily lives.

Aside from being driven by animal welfare and health concern, many EE-related research projects have also assessed the influence of poor housing conditions on research data. Garner (2005), van Praag, Kempermann and Gage (2000), and Würbel (2001, 2007), among others, demonstrated that life in barren cages leads to abnormal brain development and to physiological and behavioral malfunction. Standard non- to little-enriched cages can cause a variety

of abnormal behaviors, such as stereotypies (abnormal repetitive behaviour patterns) (see e.g., Würbel and Stauffacher, 1994, 1996; Würbel, Stauffacher and Holst, 1996) and inactivity while awake, observed for example in rhesus monkeys (Hennessy et al., 2014) and mice. Inactivity appears to be an alternative to stereotypic behavior and indicates a depression-like state (Fureix et al., 2016).

Nonetheless, for a period of time, a number of laboratory animal scientists strongly believed that standardizing the animals' environment—by housing animals in barren cages—was essential to control environmental variables (e.g., Bayne, 2005; Eskola et al., 1999; Gärtner, 1999; Tsai et al., 2002, 2003, 2006). The assumption was that standardization was crucial to minimize both variation in the data and the risk of obtaining conflicting results in replicate studies. Many laboratory animal scientists were concerned that implementing EE would add undesirable variation to their responses to experimental treatments (e.g., Bayne, 2005; Eskola et al., 1999; Gärtner, 1999; Tsai et al., 2002, 2003, 2006). However, eight mouse strains kept under such uniform, standardized conditions, and tested on highly standardized behavioral tests in different laboratories, showed significant laboratory dependent variations (Crabbe, Wahlsten and Dudek, 1999). Since then, studies by Augustsson et al. (2003), van de Weerd et al. (2002), Wolfer et al. (2004), and Würbel (2007) have demonstrated that housing conditions can be enriched without increasing variability in experimental results. Additional experiments using mice confirmed earlier research findings that basic environmental enrichments (shelters and nesting material) can be used without compromising the research data (André et al., 2018). Furthermore, this study showed that data from mice who had access to shelters and nesting material is comparable to previous data collected under barren housing conditions, consistent with earlier findings (see Augustsson et al., 2003). The authors concluded that the influence of enrichment on research outcomes was trivial, and that nesting material and shelters could be used without negative impact on study outcomes or loss of comparability to previous data obtained from animals living in impoverished cages. (André et al., 2018).

In the future, rather than using more animals in new experiments on this topic, a systematic review (SR) could be undertaken to provide an overview of the accessible evidence and new knowledge without further animal use. It would also point out knowledge gaps and assess the quality and validity of the conducted animal studies (for more on SRs of animal experimentation, see e.g., Systematic Review Center for Laboratory Animal Experimentation, SYRCLE, n.d. a).

This so-called standardization fallacy (Würbel, 2000), the belief that homogenization of study populations (using the same strain, age, sex, weight, housing conditions, etc.) is an essential part of good experimental design, appears to be one driver for the irreproducibility of results and for the lack of external validity (Bailoo, Reichlin and Würbel, 2014). External validity is the

extent to which experimental results can be used as a basis for generalizations to other human and non-human animal populations in other environmental conditions (van der Worp et al., 2010). This is why authors, including Richter, Garner and Würbel (2009), Richter et al. (2010), Würbel (2000), and Würbel and Garner (2007), promote systematic environmental heterogenization, which is a "controlled and systematic variation of the properties of any given animal (or animal population) and its environment within a single experiment" (Richter, 2017, p. 344). Voelk et al. (2018) compared 440 single- and multi-laboratory preclincial animal studies that had used the same overall number of animals. They compared effect size estimates and found that the studies conducted in one laboratory only, in most cases did not predict effect size correctly, whereas multi-laboratory studies generated more consistent and accurate results. Within-study standardization was identified as a major cause of poor reproducibility. Thus, Voelk et al. (2018) advocate for multi-laboratory design with no increase of overall number of animals being necessary to enhance reproducibility and, potentially, external validity.

EE combined with systematic heterogenization contributes to improved quality of animal experiments (Richter, Garner and Würbel, 2009; Richter et al., 2010; Würbel, 2000; Würbel and Garner, 2007), whereas failure to provide animals with living conditions that meet their species-specific needs jeopardizes both their welfare and experimental validity (e.g., Bailey, 2018; Balcombe, 2010; Bayne and Würbel, 2014; Garner, 2005; Messmer et al., 2014; Olsson et al., 2003; Poole, 1997; Sherwin, 2004; Würbel, 2001, 2007; Würbel and Garner, 2007).

2.1   *Examples of Environmental Refinement*

An example of an extensively researched refinement method is providing mice with various types and sufficient amounts of nesting material to build nests, creating a microclimate needed for breeding and for preventing cold stress (Gaskill et al., 2009, 2012; Gaskill and Garner, 2014; Hess et al., 2008). The thermoneutral zone of mice lies between 26°C and 34°C (Gordon, 1993); and standard temperatures in animal vivariums range between 20°C and 24°C. During their inactive phase, mice prefer temperatures of 30°C–32°C (Gordon, 2012). A proper nest is, therefore, essential for reducing cold stress, which not only compromises animal well-being but also scientific data (Gaskill et al., 2009; Karp, 2012; Messmer et al., 2014). Gaskill et al. (2013) additionally demonstrate its negative effect on breeding performance. Nest building is a species-specific behavior of mice, the absence of which can be used as an indicator of illness (Gaskill and Pritchett-Corning, 2016). Another example involves gerbils, who have a high motivation to dig, since they naturally build and live in burrows. In standard laboratory conditions, where there is not enough substrate to dig tunnels, gerbils show stereotypic digging behavior (Wiedenmayer, 1997). One

solution, based on research conducted by Waiblinger and König (2004), is a nesting box with an attached tunnel. The artificial burrow system seems to help reduce stereotypic digging behavior.

Jirkhof (2015) found that housing conditions that meet the needs of mice help them recover better and faster from experimental procedures. The influence of environment on diseases, such as cancer, has also been demonstrated; for example, by Cao et al. (2010). In colon cancer and melanoma research, mice living in an enriched environment showed reduced tumor growth and increased remission compared to those living in a non-enriched environment (Cao et al., 2010). Rabbits who received special positive attention from their care givers showed a markedly increased resistance to the development of atherosclerosis compared to rabbits who received no extra attention (Nerem, Levensque and Cornhill, 1980).

2.2   *Discussion on Environmental Refinement*

It has been established that animals in a monotonous environment frequently display abnormal behaviors, such as stereotypies (Garner, 2005; Garner and Mason, 2002; Gross et al., 2012; Howerton, Garner and Mench, 2008; Würbel and Stauffacher, 1994, 1996; Würbel, Stauffacher and Holst, 1996). Furthermore, research has demonstrated the importance of environmental refinement, not only for animal welfare and for decreasing the negative health effects of life in captivity, but for its benefits for research outcomes in terms of their reliability, replicability, and validity (e.g., Abou-Ismail and Mahboub, 2011; Garner, 2005; Weed and Raber, 2005).

Due, at least in part, to enforcement of animal protection laws, housing conditions for laboratory animals have improved over the past decade. In the EU, the Commission Recommendation of 18 June 2007 on guidelines for the accommodation and care of animals used for experimental and other scientific purposes (Commission of the European Communities, 2007)—which was later largely adopted by Directive 2010/63/EU (European Parliament, 2010, Annex III)—helped to enhance the captive environment of laboratory animals. However, exceptions to these minimum requirements may be demanded by researchers for certain experiments. Examples include housing social species, such as rats, pigs, or non-human primates, individually and away from their social groups; or not providing rodents with sufficient nesting material and shelters, to allow easier and quicker monitoring. Yet, in most cases, a solution that considers the animals' well-being and does not further compromise their welfare could probably be found.

Moreover, it should be noted that the term most frequently used when talking about an improved living environment, *environmental enrichment*, can be misleading, since it suggests that the standard cage conditions should be

considered normal or species-typical. However, captive conditions have little in common with the natural habitat of every single species used in research. For example, Lahvis (2017) points out that the floor area in a standard mouse cage is 280,000-times smaller than the animal's natural home range. For rhesus macaques, he calculated it is 7 million-fold smaller. Along with the difference in the size of the animals' habitats, the stimulation provided in laboratories is also different from what animals encounter in their natural environments. Burghardt (1996) argues that it would be more accurate to use the term *controlled deprivation*, since all captive environments deprive animals of some natural stimuli. He points out that these restrictions have various, and oftentimes unpredictable, consequences for the welfare of captive animals (Burghardt, 1996, 1999). In fact, a study by Gross et al. (2012) showed that around 12% of mice who lived in enriched cages which contained nesting material, a shelter and a climbing structure, still revealed stereotypic behavior. Moreover, evidence indicates that when stereotypies are not observed, a potential reason could be that they are only displayed when nobody is watching, e.g., in the nocturnal phase (Wells, 2017); or, since highly stereotypic animals seem to cope better than their identically-treated conspecifics, non-stereotypic animals present an even more abnormal, depression-like state as an alternative to stereotypic behavior (Mason, 2006). It has been shown that sustained, uncontrolled stress can, at least in some mouse strains, foster *learned helplessness* (Cabib, 2006).

2.3   *Challenges in the Implementation of Refined Housing*

The enforcement of animal protection laws has contributed to somewhat improved housing conditions for laboratory animals over the past decade. However, despite the mounting evidence of welfare and scientific problems associated with standardized housing, the implementation of animal husbandry knowledge in laboratories has in the author's experience been a major and elusive challenge.

It is increasingly recognized that experimental animals experience serious and repeated stress and distress, caused by life in the laboratory. Besides being a welfare concern, there are multiple factors that adversely affect the animal's biological systems and thus the data collected from these animals (Bailey, 2018). Examples for stressors and thus potential influences on data, besides the confinement itself, include ultrasonic noises (Baldwin, Primeau and Johnson, 2006; Turner et al., 2005), bedding material and cage cleaning (Burn et al., 2006), handling, blood collection, and orogastric gavage (Balcombe, Barnard and Sandusky, 2004), and the experimenters (Chesler et al., 2002) and their sex (Baldwin, Primeau and Johnson, 2006; Sorge et al., 2014).

Numerous studies have shown that animals living in captive environments are generally abnormal and unhealthy, as such environments change their

behavior as well as immune, nervous, and endocrine functionality. Examples include their altered response to infection (Gurfein et al., 2014), altered immune response (Beura et al., 2016; Messmer et al., 2014), increased rates of obesity, Type II diabetes, high blood pressure, and premature death (Martin et al., 2010), altered brain development (Bennett et al., 1964; Kempermann, Kuhn and Gage, 1997; Lewis et al., 2006; Rosenzweig and Bennett, 1969; Rosenzweig et al., 1962), decreased strength and endurance (During et al., 2015), altered sleep, activity patterns, and blood pressure (Martire et al., 2012), altered growth rates (Serrat, King and Lovejoy, 2008), altered organ development, metabolic, growth, and reproduction rates and behavior (Gordon, 2012), and enhanced tumor growth (Cao et al., 2010; Li et al., 2015). As such, untreated control animals do not represent healthy individuals, since they are metabolically abnormal (Martin et al., 2010). To date, there are only a few studies comparing wild versus confined animals, but they all show immense biological differences in physiology, such as structure variation of the visual cortex among caged and free-roaming Norway rats (Campi et al., 2011), lower levels of cholesterol in wild versus captive animals (Schmidt et al., 2006), and immune system dissimilarities (Beura et al., 2016).

We must acknowledge that even if laboratory animal housing is enriched, it cannot be enriched to an extent that it has no negative effect on the animal's welfare (e.g., Burghardt, 1996; Gross et al., 2012). Well-being can only be achieved if the animal experiences positive welfare states, which require a responsive environment the animal can engage with. Studies show that animals prefer complex environments and are motivated to work for them (Anselme, Robinson and Berridge, 2013; Sherwin et al., 2004). Current minimum legal requirements for animal housing in the European Union, laid out in Directive 2010/63/EU, are still insufficient in meeting all needs of all animals; although they are held to be the most progressive in the world. As shown, problems of confinement are manifold. Animals' lives in captivity are monotonous and, therefore, lead to boredom (Burn, 2017; Meagher and Mason, 2012), learned helplessness and depression (Cabib, 2006; Špinka and Wemelsfelder, 2011), and abnormal behaviors. The effects pose serious welfare concerns and raise concerns about the validity and translatability of data obtained from these unhealthy individuals.

2.4   *Potential Improvements*

In assuming an ethical responsibility to improve the lives of captive animals (Gruen, 2014), the goal of husbandry refinement should be not only to reduce stressors but to promote well-being. It is apparent that current housing conditions do not achieve that. One step towards improving animal housing is to provide cages that allow for more natural behaviors. Makowska and Weary (2016a) investigated the frequency of burrowing, climbing, and standing

upright of rats held in pairs in standard (behaviorally restrictive) laboratory cages in comparison with rats in cages allowing these behaviors (larger cages with lower floors, filled with moist soil, holding five rats per cage) over a period of 13 months. Although climbing bouts decreased with age, standing upright and especially burrowing were still frequent behaviors in older rats. Stretching is a corrective response to stiffness caused by immobility or positional stress (Bertolucci, 2011). Makowska and Weary (2016a) found that standard-housed rats performed 9 times more lateral stretches than rats housed in the semi-naturalistic environment. The authors proposed that standard-housed rats were stretching frequently in an attempt to alleviate stiffness from low mobility associated with standard housing. Improved welfare of the rats housed in the semi-naturalistic cages was observed in an anticipatory behavior test that assessed differences in reward sensitivity performed when the rats were 19 and 21 months old (Makowska and Weary, 2016b).

From the animals' perspective, an even better approach would be the radical solution for housing refinement proposed by Lahvis (2017). Lahvis suggests that research animals should live in the wild or at least roam freely in a large, captive environment under naturalistic conditions. He is confident that with available technologies (e.g., cameras, transponders, magnetometers, pressure sensors, global positioning systems), this novel approach could be accomplished for many experiments. Lahvis (2017) advises that biomedical researchers should work together with behavioral ecologists to develop sufficiently complex environments in order to ensure that test subjects produce scientific data not influenced by husbandry.

3 A "Culture of Care" for Animals as Refinement

The term *culture of care* has frequently been referred to by members of the laboratory animal science community to demonstrate "a commitment to improving animal welfare, scientific quality, care of the staff and transparency for the stakeholders." (Norecopa, 2016a). For instance, a working document on the development of a common education and training framework to fulfill Directive 2010/63/EU requirements mentions the culture of care numerous times (National Competent Authorities for the implementation of Directive 2010/63/EU on the protection of animals used for scientific purposes, 2014). Individuals responsible for the welfare of animals should establish and maintain high standards to champion a culture of care among both husbandry and scientific staff (European Commission, 2014). Entire sessions at conferences have been dedicated to this topic, including sessions at the

European Society for Alternatives to Animal Testing (EUSAAT) Congress in 2015 (EUSAAT, 2015) and at the 10th European World Congress on Alternatives and Animal Use in the Life Sciences in Seattle in 2017 (von Aulock, 2017).

Reinhardt (2003, p. 123) identifies compassion for laboratory animals as a refinement: "Kindness and concern for animals in the laboratory often have been stigmatized as subjective, emotional qualities that can undermine the 'objectivity' of biomedical and psychological research." However, since there is evidence that the human-animal bond helps animals to cope with stressful situations in the laboratory (Wolfle, 1987), compassion for laboratory animals should not be dismissed as emotional and subjective but as a sound methodological base for scientifically valid animal-based research (see Mahoney, 1992; Reinhardt, 2003). Compassion implies an acute awareness of an animal's state of emotional, behavioral, and physical well-being and the urge to provide them with the conditions essential for optimal well-being (Reinhardt, 2003). According to Herzog, "there is every reason to believe that individuals who care about their wards on a personal level actually treat the animals better." (2002, p. 30). Morton highlights that, ideally, the staff assessing pain in animals should have an empathetic attitude toward them (Morton, 2000). Such a mindset can also be seen as a protection mechanism to control unrelated, potentially data-influencing, variables (Reinhardt, 2003). Brown (2014) states, "Although there are laws and regulations that govern working with research animals, institutions involved in research, testing, and teaching using laboratory animals should strive to go beyond what is legally required and work to establish a 'culture of care' to ensure animals are treated with compassion and respect." Brown highlights that this culture of care for animals not only benefits animals but the quality of science as well.

3.1   *From Theory to Practice*

How far a culture of care is being implemented on an institutional level is unknown. Personal experiences of this author—as an inspector of animal research institutions in Germany between 2007 and 2016 (Herrmann, 2013; Herrmann and Ratsch, 2010; Maurin, 2012)—revealed differences regarding the level of care for animals within the same institutions, with individual care givers acting more or less compassionately towards their animals. An *institutional* culture of care agenda could not be identified.

The European Commission (EC) (2014) recommends the implementation of such a culture, and other countries have taken steps, in this direction. For example, New Zealand's National Animal Ethics Advisory Committee guide is called *A Culture of Care*: A *Guide for People Working with Animals in Research, Testing, and Teaching* (National Animal Ethics Advisory Committee, 2002). Several pharmaceutical companies, such as Sanofi-Aventis

and Merck (Klein and Bayne, 2007), and commercial breeding companies (Brown, 2014) are reported to have established culture of care programs; however, no external review or assessment of these programs has been published.

3.2 *Towards a Culture of Care and Compassion for Animals*

There is potential for a positive impact of a culture of care on animal use and welfare. But how can we implement such a culture? Schuppli et al. (2017) used a new educational approach to test if exposure to socialized rats, who were trained to fulfill several tasks, fostered compassion among animal experimenters. Six rats were trained using positive reinforcement techniques to, for example, jump onto a scale, or to lift objects. Participants observed these rats and engaged with handling them. After the class, researchers (17) discussed their feelings and reactions. Main findings included that all participants were impressed by the rats' abilities and the close relationship with their trainers. They assumed that this positive animal-human interaction decreased stress in the rats. However, various views existed in regard to potential effects on data. The experimenters expressed unease about emotional difficulties in "sacrificing" their experimental animals after having bonded with them (Schuppli et al., 2017). This highlights one of the major obstacles: When animal researchers develop compassion for their research subjects, they face moral difficulties (see Birke, Arluke and Michael, 2007; Gluck, 2016) and moral harms (see Chapter 13 in this Volume, Johnson and Smajdor, 2019) just as animal caretakers and technicians do. However, this could be an important starting point in moving towards a culture of compassion for all animals which could contribute to their replacement efforts.

4 Refinement of Experimental Procedures

There are several essential refinement methods to reduce the pain, distress, anxiety, and suffering inflicted during the course of experimenting on the animals. Handling and restraint techniques are a source of potential distress and anxiety (Balcombe, Barnard and Sandusky, 2004; Hurst and West, 2010; Meijer et al., 2006); and these techniques have been investigated in experimental studies on stress (Johnson, Sharp and Miller, 2000). To avoid negative effects on behavior, tail handling of mice should be replaced by using tunnels or cupping mice in the open hand (Gouveia and Hurst, 2013). These and other non-aversive handling practices should be implemented industry-wide, since they have been shown to reduce anxiety (Hurst and West, 2010) and optimize the performance of mice in behavioral tests (Gouveia and Hurst, 2017). A recent study by Clarkson et al. (2018) concluded that particular handling methods can

not only cause anxiety, but they can also alter the hedonic value of reward. Tail-handled mice demonstrated a decreased responsiveness to reward and, potentially, a more depressive-like state compared to tunnel handled conspecifics (Clarkson et al., 2018).

For surgical procedures, basic experimental refinements include: proper acclimatization of the animals to the room where anesthesia will be induced (Flecknell, 2018a); optimal anesthesia, peri- and postoperative analgesia; and adequate postoperative monitoring and care, including pain management (Flecknell, 2016; Herrmann and Flecknell, 2018a). The application of humane endpoints also prevents needless suffering. A humane endpoint (or "less-inhumane endpoint," see Balls, 1999, p. 1) represents "[t]he earliest indicator in an animal experiment of (potential) pain and/or distress that, within its scientific context and moral acceptability, can be used to avoid or limit adverse effects by taking actions, such as humane killing, terminating the study, or alleviating the pain and distress." (Hendriksen, Morton and Cussler, 2011, p. 344).

The way an animal is killed is another subject for refinement. Animal care policies in many countries stipulate that death must be painless, and fear and anxiety should be minimized (e.g., Charbonneau et al., 2010; European Parliament, 2010). Less inhumane killing comprises the use of the least distressing and least painful methods that cause rapid loss of consciousness and subsequent death (see e.g., Leary et al., 2013).

The application of our steadily increasing knowledge on experimental refinements should benefit the over 115 million animals who are used annually in research, testing, and education around the world (Knight, 2008; Taylor et al., 2008). However, this benefit cannot be achieved unless the knowledge is translated into practice. In cases where research workers plan to use, for example, less than optimal anesthesia or analgesia protocols, or do not provide other standard veterinary practices, they need to scientifically justify this and demonstrate that the anticipated benefits of the experiments still outweigh the harms inflicted upon the animals (Herrmann and Flecknell, 2018b). Due to the multitude of available means, solutions can be found, in most cases, that help prevent needless animal suffering (Herrmann and Flecknell, 2018a).

4.1   *The Use of Experimental Refinements in Practice*
Several structured and systematic literature reviews have given some insight on certain experimental refinements, notably, killing methods (Pound and Nicol, 2018; Uhlig et al., 2015) and the use of anesthetics and analgesics (Bertrand, Sandersen and Flecknell, 2018; Carbone and Austin, 2016; Coulter, Flecknell and Richardson 2009; Coulter et al., 2011; Pound and Nicol, 2018; Richardson, and Flecknell, 2005; Stokes, Flecknell and Richardson, 2009; Uhlig et al., 2015). For example, animal research involving surgical procedures carried out on

diverse species and published in peer-reviewed journals has been analyzed with regard to analgesic and anesthetic administration (Coulter et al., 2011; Coulter, Flecknell and Richardson, 2009; Richardson and Flecknell, 2005). Stokes, Flecknell and Richardson (2009) focused on studies conducted in two time periods (2000–2001 and 2005–2006), assessing trends in the administration of analgesics and anesthetics to laboratory mice and rats undergoing surgical procedures. The study showed a trend of improvement in terms of safer anesthetic regimens used in the later period examined; however, the findings of this study and an earlier review assessing analgesic use in rodents (Richardson and Flecknell, 2005) show that there was still significant scope for refinement, especially with respect to perioperative care.

A systematic review of anesthesia, analgesia and euthanasia methods used in anesthesiology, respiratory and critical care research in top-10 impact factor ranked journals journals pointed to insufficient reporting of experimental studies with small laboratory mammals. Despite the poor reporting, the review found shortcomings in the application of refinement (Uhlig et al., 2015). Another recent attempt to assess trends in pain management, this time in papers published before 2011 and from 2014 to 2015, further confirmed that reporting (and probably the use) of experimental refinement methods is still poor (Carbone and Austin, 2016). The review demonstrated that scientific publications still cannot be relied upon to present a detailed description of analgesia and anesthesia protocols, not to mention other experimental refinements.

Another approach employed by the author of this chapter, with Flecknell (2018 a, b, c), was to retrospectively review proposals for authorization of basic and applied animal research studies to learn which experimental refinements were proposed. Over 500 applications submitted to the German competent authorities in 2010 were reviewed. German law stipulates that all possible refinements that are planned in an animal study are described in detail in its proposal. The review's goal was to evaluate the intended application of and, thus, the awareness about possible refinements. Among other results, postoperative analgesia was not proposed for 30% of surgeries; and, in the majority of cases, its scientific necessity was not further discussed. Following 10% of procedures, animals were to be given pain relieving medication only if the investigators decided that it was necessary; however, structured assessments to detect pain were absent (Herrmann and Flecknell, 2018a).

4.2   *Discussion on Refinement of Experimental Procedures*

Structured and systematic literature reviews and the work of this author found strong indications for flaws in the administration of experimental refinement. Refinement methods need to be fully employed in order to minimize stressors that can lead to distress, such as suffering from postoperative pain, or living

in a barren cage. The biological consequences of stress and distress compromise rigor, reliability, and relevance of data collected from these animals (see Bailey, 2018 for a review on how stress of laboratory life and experimentation can adversely affect research data). Animal researchers are responsible for the animals they use (National Health and Medical Research Council, 2013). Thus, they and their animal care staff should know enough about animal behavior to properly assess the health and well-being of their test subjects. In the European Union, they are legally required to be skilled, educated, and equipped to detect and relieve suffering accordingly (European Parliament, 2010, Article 24).

There are several challenging areas of refined care and use that should be addressed. For example, there is a need for automated, remote, 24/7 cage-side monitoring to identify abnormal behavior, which is especially important when assessing the welfare of genetically modified animals, as well as for prey species who tend to mask their medical condition or psychological state. Additionally, there is a need for further development and implementation of valid pain-assessment techniques to determine the efficacy of treatment in the individual animal due to individual variations in pain response. While there is necessity for further research into certain areas of experimental refinement, it is essential that we apply the knowledge we already have, so that immediate improvements in animal welfare can be achieved.

5   Refinement of Experimental Design, Conduct, and Reporting

There have been quality problems throughout medical and biomedical research (Begley and Ellis, 2012; Harris, 2017; Pound and Bracken, 2014; Prinz, Schlange and Asadullah, 2011). "The scandal of poor medical research" with human subjects was discussed in a British Medical Journal (BMJ) editorial in 1994 (Altman, 1994). A biostatistician took a prominent stance against the unethical misuse of statistics (Altman, 1980). In a follow up 20 years later, another BMJ editorial called, "Medical research—still a scandal," concluded that matters have become worse (Smith, 2014). It is apparent that the quality of *in vivo* research with animal and human subject demands urgent improvement. Weaknesses in design, conduct, and analysis of biomedical and public health research studies yield misleading results and, thus, waste resources (Ioannidis et al., 2014). Since legally-required animal data forms the basis of decisions to move forward to human clinical trials, flawed animal research is additionally problematic.

Aside from evidence that many animal experiments that are performed never get published (Scherer et al., 2018), a large part of what gets published is

incorrect (e.g., Harris, 2017; Ioannidis, 2005; Freedman, Cockburn and Simcoe, 2015). Ioannidis (2005) argues that it is highly probable that most published findings are indeed false. He drew his conclusion after conducting simulation studies and SRs. He calculated that, at best, only one in three publications took basic precautions to minimize bias (Ioannidis, 2005). Freedman, Cockburn and Simcoe (2015) estimated that more than 50% of all preclinical studies in the United States are unreliable, and that the financial damage of these irre-producible preclinical studies is US$28 billion per year. Their analysis revealed that about 20% of the studies had an untrustworthy experimental design, one quarter used media that contained contaminated cells and antibodies, and in 18% of studies the data analysis was poor. All of these issues have contributed to the so-called *reproducibility crisis* in animal research (e.g., Aarts et al., 2015; Baker, 2016; Begley and Ioannidis, 2015; Bracken, 2009; Collins and Tabak, 2014; Freedman, Cockbury and Simcoe, 2015; Ioannidis, 2005; Perel et al., 2007; Pound et al., 2004; Pound and Bracken, 2014; Reichlin, Vogt and Würbel, 2016; Scannel and Bosley, 2016; Würbel, 2016). A review of the literature by Bailoo, Reichlin and Würbel (2014) strongly suggests that experimental design and conduct of laboratory animal research are in need of improvement. A study by Vogt et al. (2016) revealed that animal researchers working in Switzerland do not apply basic principles of study design to avoid bias and do not properly report their study outcomes. They also found that neither the Swiss regulatory authority nor the international journals and their peer reviewers had adequate knowledge to recognize these flaws.

In an attempt to improve the quality of research reports, several checklists and guidelines have been put in place, such as Consolidated Standards of Reporting Trials (CONSORT) and Standards for Reporting of Diagnostic Accuracy (STARD) for human clinical trials, Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) for SRs and meta-analyses, and Gold Standard Publication Checklist (GSPC), Animal Research: Reporting of *In Vivo* Experiments (ARRIVE) guidelines for animal research (Glasziou et al., 2014; Hooijmans, Leenaars and Ritskes-Hoitinga, 2010) and HARRP, the harmonized animal research reporting principles, which are a recent attempt by ICLAS (International Council for Laboratory Animal Science) in harmonizing animal research reporting to further improvements in the scientific rigor of animal experiments (Osborne et al., 2018). The ARRIVE guidelines (Kilkenny et al., 2010) are the most widely known for reporting of animal-based experiments. These guidelines have recently been complemented by the Planning Research and Experimental Procedures on Animals: Recommendations for Excellence (PREPARE) guidelines, which help to ensure quality when preparing animal studies (Smith et al., 2017).

The ARRIVE guidelines were adopted by more than 1000 scientific journals, and more than 20 funding agencies were expected to endorse them in 2010 (Baker et al., 2014; Enserink, 2017). Two years later, Baker et al. (2014) assessed the degree to which they had been endorsed by reviewing journals, such as Nature and PloS, and found that there was little improvement. The knowledge about and the use of reporting guidelines, such as the ARRIVE guidelines, is still not widespread, as a study by Reichlin, Vogt and Würbel (2016) has shown. Reichlin et al. asked animal experimenters in Switzerland to complete a questionnaire regarding their use of measures against risk of bias. Only 16% responded. The ARRIVE guidelines were known by less than half (43.7%). Furthermore, Carbone and Austin (2016) found no increase in reporting of analgesic use in the articles published in journals that had agreed to endorse the ARRIVE guidelines. Results of a recent randomized controlled trial of close to 1,700 scientists, who submitted papers to the scientific journal *PLoS One*, suggest that scientists are either ignoring the guidelines or are still unaware of their existence. Another finding was that, even when an ARRIVE checklist was completed, the correlating papers were actually not more compliant, which may indicate that researchers do not know what is expected of them and why providing this information is crucial, emphasizing the importance of proper training (Enserink, 2017; Hair et al., 2018).

5.1   *Sources of Bias in Animal-based Research*

There is a large number of potential sources of bias in animal research. Not surprisingly, most published animal studies have some risk of bias (Macleod et al., 2015). Safeguards to avoid bias in study design, conduct, and analysis include randomization of treatment groups to eliminate systematic differences between them, blinding of investigator to treatment and to handling of data, and reporting on sample size estimation (Macleod, 2011). Analytical errors may account for close to a quarter of the irreproducible studies (Freedman, Cockburn and Simcoe, 2015); thus, knowledge about statistical methods is essential. Other suggested items for reporting include: a clear description of the hypotheses tested or primary and secondary objectives of the study, housing and husbandry, including welfare-related assessments and interventions, adverse events, and interpretation of results, taking into account the hypotheses/study objectives (Kilkenny et al., 2010). Part of the reproducibility and translatability crisis is considered to be due to poor experimental design and conduct of animal experiments (Bailoo, Reichlin and Würbel, 2014; Ioannidis, 2005; Macleod, 2011; van der Worp, 2010; Würbel, 2016), including the influences of inappropriate animal housing (Lahvis, 2017) and handling (Gouveia and Hurst, 2017), insufficient pain relief (Carbone and Austin, 2016; Herrmann and Flecknell, 2018a),

as well as the absence of other refinements, such as careful monitoring, early humane endpoints, and less inhumane killing methods to reduce pain, suffering, and distress (Herrmann and Flecknell, 2018b and 2018c). The other, and perhaps larger, part is due to insurmountable species differences (Pound and Bracken, 2014; Pound and Ritskes-Hoitinga, 2018), which Russell and Burch already discussed 60 years ago (1959, Chapter 5).

Another source of bias is selective reporting when publishing results of animal experiments (Briel et al., 2013; Ioannidis, 2012; Landis et al., 2012; Lees et al., 2012; Macleod et al., 2004; Pound and Bracken, 2014; Sena et al., 2010; Tsilidis et al., 2013; Würbel, 2016). One problem relates to *negative* findings— studies for which the original hypotheses were not proven. Some of these are not published at all, which has long been recognized as a source of publication bias. The second problem relates to studies that are reported incompletely. For example, only the parts that demonstrate that the treatment is effective are reported, with whole experimental groups excluded from reporting. This is selective outcome and analysis reporting bias (Ioannidis, 2012). These partially or unreported studies may be repeated by others and thus represent an unnecessary waste of animal lives. Incomplete reporting of published findings makes it impossible to replicate studies (Begley and Ellis, 2012). Because negative findings are often not published (Scherer et al., 2018), the value of published findings is over-estimated, which, in part, could explain some of the difficulties in translating promising preclinical results into effective therapies for human disease (Bath et al., 2009; Mergenthaler and Meisel, 2012; Sena et al., 2010).

Yet another pitfall is researchers' freedom of flexibility in data collection, analysis, and reporting, which dramatically increases false-positive rates in the literature and, therefore, contributes to misleading animal research data and overestimation of its significance. Regardless of the nominal endorsement of a maximum false-positive rate of 5% ($p \leq .05$), standards for disclosing details of data collected and analyzed make false positive results very likely (Simmons, Nelson and Simonsohn, 2011). The authors describe this as *p-hacking*. Oftentimes, an experimenter is more likely to find evidence that an effect exists falsely than to find evidence that it does not correctly. This occurs because of the investigators' degree of freedom with regard to the amount of data collected and analyzed, the exclusion of certain observations made, the comparison or combination of conditions, the variables considered, and so forth. It is uncommon for researchers to make these decisions before undertaking experiments. Their exploratory behavior is explained as ambiguity in how best to make these decisions and the desire to find statistically significant results.

Confirmatory bias is another potential pitfall, since people tend to interpret ambiguous information in such a way that it supports a justifiable

conclusion that matches their own aspirations (e.g., Dawson, Gilovich and Regan, 2002). *HARKing* (i.e., Hypothesizing After the Results are Known) (Kerr, 1998) is another common and problematic practice in science. Statistical tests to differentiate true effects from random noise are designed for confirmatory research, not exploratory research. Thus, when researchers change their *a priori* hypotheses after obtaining their results, this leads to false conclusions.

An additional area that urgently needs refinement is transparency and data sharing to avoid publication bias and needless repetition of studies. Openness is a cornerstone of science and could help in reducing the reproducibility problem science is facing (Errington et al., 2014; Harris, 2017; McNutt, 2014). It is essential to discover and correct errors. The Food and Drug Administration Modernization Act of 1997 (FDAMA) (Food and Drug Administration, 2018) requires scientists to register their hypotheses and endpoints in advance, if they plan to run a clinical trial on potential new pharmaceutical drugs (ClinicalTrials.gov). This new law went into effect in 2000. It also requires pharmaceutical companies to publish their results, thus, avoiding publication bias. Despite the insufficient enforcement of the law, as many scientists still do not report the results of their studies (Harris, 2017), the indispensability of such provisions is demonstrated by the findings of a study conducted by Kaplan and Irvin (2015). They assessed whether the FDAMA had any effect on study outcomes. Before the law was in place, 57% of drugs or supplements showed benefits; after the law was in place, 8% of the studies published confirmed their preregistered hypotheses (Kaplan and Irvin, 2015). Such a prospective registration process is currently exceptional for animal-based studies, but it is unquestionably required in order to enhance transparency, reduce selective reporting bias, and prevent duplication. The Center for Open Science, a nonprofit where researchers can register their hypotheses *a priori* (https://cos.io) and Preclinical Trials, a platform for registration at the outset of all types of animal studies (www.preclinicaltrials.eu), will hopefully improve the current situation. An additional measure to improve transparency, and potentially reproducibility, is data sharing, which is a requirement for publication by some major journals but many researchers still refuse to share. By sharing data, errors can be discovered (e.g., Salzberg et al., 2001). This is especially important in animal research, since it helps reduce the number of animals used and sheds light on the real value of animal derived data.

5.2   *Necessary Steps*
The improved quality of human clinical trials was achieved by strategies to minimize bias, *a priori* power analysis and further biostatistics, clear definition of the primary and secondary endpoints, data monitoring and auditing, internationalization and inclusion of multiple centers, external steering committees and safety monitoring, rigid publication standards, trial registries, and

more (Dirnagl and Fisher, 2012). The lessons learned from the improvement of human clinical trial quality should be adopted by preclinical (Dirnagl and Fisher, 2012) and all other biomedical research fields (Hartung, 2013), where relevant and with appropriate changes, since flawed research is unscientific and unethical. The ethical issues with research involving animals become extra critical as needless animal suffering must be avoided, and as preclinical animal data generally forms the basis for decisions whether to proceed to human clinical trials. Thus, in order to adhere to the 3Rs, the following efforts are crucial:

– Education and ongoing training of researchers in experimental design, statistical methods, and model selection (Justice and Dhillon, 2016).
– Close assistance in study design by institutional animal welfare bodies and by biostatisticians.
– As a possible solution for the problem of false positives making their way into the literature, some researchers suggest the *p*-value threshold should be reduced to 0.005 (Chawla, 2017). Others say researchers should select and justify *p*-value thresholds for their experiments, before collecting any data. These levels should be based on factors such as the potential impact of a discovery. These thresholds could then be evaluated via *registered reports*, a type of scientific article in which methods and proposed analyses are peer-reviewed *before* any experiments are conducted (Chawla, 2017).
– Transparency must be improved as it is crucial to document all anticipated or exploratory steps in the study. Prospective registration of all animal studies with their hypotheses and endpoints is essential to prevent selective-reporting biases (Ioannidis, 2012) and avoid study duplications (Preclinical Trials, n.d.).
– Disclosure and openness are critical elements of science for self-correction, and they can help avoid poor practices, such as HARKing.
– The use of preparation and reporting guidelines, such as the PREPARE guidelines (Smith et al., 2017) combined with the ARRIVE guidelines (Kilkenny et al., 2010), should be a mandatory, legally required part of funding applications, project license applications, as well as publications. Education on how to fill out the checklists and present the required information in the publication, as well as a focus on enforcement of compliance to both by journals, is critical (Eisen, Ganley and MacCallum, 2014; Enserink, 2017; Hair et al., 2018; Herrmann and Flecknell, 2018a).
– Raw data, analyses, and protocols must be made available to allow other researchers to verify results. This can easily be achieved by using data repositories (e.g., https://datadryad.org or https://figshare.com).
– Reporting of all study outcomes to avoid traditional reporting bias and selective outcome and analysis reporting bias should be mandatory (Ioannidis, 2012).

- Retrospective assessments of animal studies (see EC Expert Working Group for Project Evaluation and Retrospective Assessment, 2013, pp. 28–32) should be performed comprehensively and by independent experts; and all results should be published to enhance transparency, minimize publication bias, identify animal models lacking external validity, and, thus, improve future research.
- Mandatory data sharing so that other scientists can build on the work and discover errors faster (cf. the error in the Human Genome Project discovered by Salzberg et al., 2001). Data sharing should be compulsory, especially when research is publicly funded.

It is equally important that funding and regulatory bodies, animal ethics committees, animal welfare bodies, journal editors, and peer reviewers have a detailed knowledge of these topics in order to recognize flawed research studies. This requires effective and thorough education and training of funders, animal ethics and welfare committees, and regulatory body members on how to assess animal research proposals (Vogt et al., 2016). Furthermore, in order to review these applications in-depth, enough time and manpower are a prerequisite.

6   Refinement: Are We Doing What We Can?

As presented in this chapter, knowledge about and implementation of refinement of husbandry, experimental procedures and design, conduct, and reporting appears to still be patchy. Since adoption of refinement strategies has been inconsistent, it would seem that rather than use additional animals to carry out more refinement research, we should focus on the comprehensive application of existing refinements in animal laboratories as well as on reducing and replacing animals.

6.1   *But What about the Refinement of Animal Models?*
Animal models ought to describe a biological phenomenon that the model species has in common with the target species. Significance and validity, in terms of the translatability of results produced in an animal model to the human condition, "depend on the selection of a suitable animal model," writes Hau (2008, p. 4), which is why comprehensive knowledge about comparative anatomy and physiology is essential. A majority of animal models developed with the expectation to study the origin, disposition, and treatment of human disorders and is created through experimental induction, genetic modification, or breeding of disease-causing mutations (Hau, 2008, p. 4). These presumed predictive models are used to find treatments or to assess the toxicity of drugs and other chemicals (Hau, 2008). Hence, they cause conditions

associated with pain and distress up to severe, long-lasting suffering for these animals.

Some laboratory animal scientists focus on the refinement of animal models in an attempt to reduce the suffering caused. Examples for refinement recommendations of animal models include those described for mice and rats who are utilized as models of ischemic stroke (Percie du Sert et al., 2017), for rheumatoid arthritis (Hawkins et al., 2015), in experimental autoimmune encephalomyelitis (EAE) (Wolfensohn et al., 2013a), as models and in procedures involving seizures, convulsions, and epilepsy (Lidster et al., 2016; Wolfensohn et al., 2013b), and as models of sepsis and septic shock (Lilley et al., 2015). If the gathering of such recommendations does not involve additional harmful animal experiments, and in case these guidelines are then applied in practice, they could lead to an improvement of the individual animal's life.

However, due to failure of numerous models to predict human outcomes (e.g., Joffe et al., 2016; Mak, Evaniew and Ghert, 2014; Pharmaceutical Research and Manufacturers of America, 2015, 2016), and due to limited funding, it seems crucial to first assess carefully which research methods and models to use. In the case of sepsis models, for example, there have been multiple publications highlighting the differences in human and mouse immunology (e.g., Mestas and Hughes, 2004; Rittirsch, Hoesel and Ward, 2007; Seok et al., 2013; Shay et al., 2013; Payne and Crooks, 2007). After over 20 years of unsuccessful research in this field, a number of scientists finally investigated why, out of the approximately 150 new compounds that were developed for the treatment of sepsis using mice, not one had beneficial effects for humans. They identified around 5,000 genes that are activated or deactivated by inflammation in humans who suffered from sepsis, trauma, or burns. They went on to look for the same genes in one commonly-used strain of mice and realized that there was no correlation (Seok et al., 2013). As a consequence of the dissimilarity of mouse and human immune systems, the entire field of sepsis research in mice has been called into question, regarding its predictive value for humans. Paradoxically, funding for this kind of animal research, which is also known for causing severe levels of animal suffering, is still ongoing (Leist and Hartung, 2013). At the same time, human-based sepsis research has led to clinical trials of effective therapies (van der Poll, 2012).

7 Reduction and Replacement: Are We Doing What We Can?

Most animal research is being justified as indispensable to furthering human healthcare. However, despite measures being taken to improve the quality of animal-based research, the translational success rate from animal studies to

humans is low: Less than 12% of drugs entering clinical trials result in an approved medicine (Pharmaceutical Research and Manufacturers of America, 2015, 2016); and between 51% and 89% of preclinical studies are not reproducible (Freedman, Cockburn and Simcoe, 2015; Harthorne and Schachner, 2012). There is an ongoing debate among scientists as to why animal models fail to be predictive: Is this mainly due to poor scientific rigor and reporting, to species differences, or to the fact that today we mainly deal with complex, oftentimes, chronic ailments of which many are not well understood and, thus, impossible to model in other animals?

As a consequence of the failure to translate findings to humans, new criteria for mouse models have been described (Justice and Dhillon, 2016). Hoping to enhance animal models of stroke, Dirnagl and Fisher (2012) call for international, multicenter, preclinical Phase III-type studies of promising new ischemic stroke therapies in animals before moving to clinical trial. As Phase III studies would be based on prior studies and would use various strains and species (Dirnagl and Fisher, 2012), as well as older animals with various comorbidities (e.g., diabetes mellitus, obesity, and hypertension) (Mergenthaler and Meisel, 2012), the severity of these experiments and the numbers of animals involved would markedly rise. Of close to 100 interventions that improved the outcome in animal stroke models, which were tested in clinical trials, only one intervention improved the outcome in human patients (O'Collins et al., 2006). Despite decades of research, most translational stroke trials that aim to extrapolate basic research findings into clinical treatments, particularly in the area of neuroprotection, have failed (Mergenthaler and Meisel, 2012). The authors admit that, to date, there is no ideal animal model for stroke, and that more complex models are needed to improve translational success in experimental stroke research (Mergenthaler and Meisel, 2012). Thus, at the time of writing, Mergenthaler and his colleague Stachelscheid are developing human stem cell-derived 2D and 3D models for stroke (vfa, 2017). Building on the latest *in vitro* research to model human brain development and disease, they plan to employ a recently established protocol for generating 3D brain tissue, so-called cerebral organoids, from human pluripotent stem cells that can be applied to study a number of human brain diseases (Lancaster and Knoblich, 2014). Renner et al. (2017) further examined the development and potential differentiation of cerebral organoids, which hold great potential to advance human-relevant stroke research.

7.1   *Potential for Reduction by Critical Appraisal of Animal Studies*

Several unsuccessful animal models have been discussed, such as for Alzheimer disease (Cavanaugh, Pippin and Barnard, 2014; Pippin, Cavanaugh and

Pistollato, 2019, Chapter 20 in this Volume; Pistollato et al., 2016), for stroke (Shuaib et al., 2007; van der Worp et al., 2010), for tuberculosis (Fonseca et al., 2017); for asthma (Mullane and Williams, 2014), for HIV/AIDS, for neurological, menopausal human therapy, and for cancer research as well as drug development (Pippin, 2012). Since only disease models with high predictive validity are likely to yield positive results and treatments for humans, it is critical to assess the reliability, reproducibility, and validity of the animal model first. With the overall low quality and predictive validity of the majority of research studies, it has become evident that animal-based studies require rigorous evaluation (Pound et al., 2004). A solid methodological approach would be to systematically review and to perform meta-analyses of animal studies, as SRs are seen by experts in the field of evidence-based medicine as the highest level of medical evidence (Hooijmans, Leenaars and Ritskes-Hoitinga, 2010).

7.1.1 Systematic Reviews (SRs)

A systematic review (SR) is a literature review that focuses on a specific question with the aim to identify and assess all relevant studies in order to generate new, high-quality evidence. Thus, it enables evidenced-based decision making (Norecopa, 2017). A SR may contain a meta-analysis. In a meta-analysis, the results of a number of independent studies are statistically combined to calculate the average effect of studies addressing the same question, which may lead to more reliable conclusions and may help to minimize needless duplication of animal studies (Hooijmans et al., 2014a). SRs conform with the implementation of the 3Rs concept (Ritskes-Hoitinga, 2016), as their application leads to a more evidence-based choice of animal models (e.g., de Vries et al., 2012; Sloff et al., 2014; Zeeff et al., 2016). They help decrease unnecessary animal studies, the evidence they produce should further responsible animal use, and they increase scientific quality (van Luijk, 2016), as they are an excellent tool to assess study quality by evaluating the internal, external, and construct validity of the models. Internal validity is the degree to which the design, conduct, and analysis of the experiment remove potential bias, so that the interpretation of a causal relationship between an experimental treatment and variation in an outcome measure is secured (Bailoo, Reichlin and Würbel, 2014). The extent to which animal data gives a basis for generalization to other animal and human populations, including other environmental circumstances, represents the external validity; and construct or predictive validity shows how good the model is, the rate to which the sampling properties are representative for the entities they ought to represent (Bailoo, Reichlin and Würbel, 2014; Würbel, 2017). An example for a SR on internal validity is the study of Macleod et al. (2008) and for construct validity, the work of Sena et al. (2010), both focusing on reasons

for translational failure of experimental stroke. SRs are a significant tool to identify quality issues with primary animal studies. For example, a recent SR on the welfare implications of toe clipping and ear notching revealed that the underlying animal experiments were too flawed to draw conclusions (Wever et al., 2017). SRs are excellent to assess the risk of bias in animal studies and thus to evaluate the reliability of the available evidence (van Luijk et al., 2014). Perel et al. (2007) systematically reviewed the success of treatments in animals and in humans, with head injury, hemorrhage, thrombosis due to acute ischemic stroke, acute ischemic stroke, and osteoporosis as well as preventive medication in neonatal respiratory distress syndrome, with their applications in humans with these impairments. Their conclusion was that the incongruity between animal and human studies may be due to bias or to the failure of the animal models to mimic clinical disease (Perel et al., 2007).

SRs of animal studies are still much less common than in the clinical setting, where they are frequently used to make evidence-based decisions on healthcare; but awareness of the benefits of the utility of SRs of animal research has been increasing (Hooijmans et al., 2014; van Luijk et al., 2014). The Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies (CAMARADES) group, at the University of Edinburgh in the UK, and the Systematic Review Center for Laboratory Animal Experimentation (SYRCLE), at Radboud University Medical Center in the Netherlands, provide a supporting framework for groups who are or want to get involved in the SR and meta-analysis of data from experimental animal studies and offer advice and training (CAMARADES, 2014; SYRCLE, n.d. b). SYRCLE has published a step-by-step guide on how to identify all relevant animal studies (Norecopa, 2017), as well as a tool similar to the Cochrane tool for assessing risk of bias in randomized clinical trials (Higgins et al., 2011), to assess the risk of bias in animal-based studies (Hooijmans et al., 2014). It is important to receive proper training first, as one needs to be aware of the pitfalls and limitations of these tools, and how they can be misused and/or misleading (Gurevitch et al., 2018). Various types of reporting biases, together with the limited methodological quality of some studies on which meta-analyses and SRs are based, can impede their conduct and interpretation (e.g., Benatar, 2007). When publication bias against negative animal studies exists, it will lead to an overestimate of the value of animal studies. It is likely that if unpublished studies were to be included, then SRs would show more studies with no effect in animals (Akhtar, Pippin and Sandusky, 2009). Checklists and tools have been proposed to help improve SRs and meta-analyses (Hooijmans et al., 2014; Moher et al., 2009).

The use of SRs should be standard practice within animal-based research, in the same way it has become a vital part of clinical research (Hooijmans, Leenaars and Ritskes-Hoitinga, 2010; Hooijmans and Ritskes-Hoitinga, 2013;

Pound and Bracken, 2014; Pound et al., 2004; Sandercock and Roberts, 2002). SRs should be conducted prior to a new animal study to assess the validity of the proposed animal model and to avoid needless animal use (Ritskes-Hoitinga and Wever, 2018). For example, in refinement research, SRs are an efficient way to gather new high-quality data without having to experiment on additional animals. As shown in this chapter, the implementation of new knowledge about refinements to improve animal welfare has proven very difficult. A prominent example is the use of carbon dioxide to kill animals. Extensive research conducted on this welfare topic has produced overwhelming evidence against its use, but these findings still have not led to the abolishment of this common practice. At the time of writing, Turner et al. are conducting a SR on the use of carbon dioxide as a killing method for mice and rats. Their protocol (SYRCLE, n.d. c), as well as the protocols of others, are published on the SYRCLE website and, since 2018, protocols of SRs relevant to human health can be registered at the international prospective register of SRs, called PROSPERO (https://www.crd.york.ac.uk/prospero/).

7.1.2 Other Retrospective Assessments

Conducting retrospective assessments (RAs) is a useful way to identify disease models and research methods that may be of limited value. Since 2013, RAs are mandatory for certain animal studies in the European Union (European Parliament, 2010, Article 39). Members of the animal research inspectorates have been required to assess the outcomes of animal studies that were classified as severe and/or use non-human primates. The animal researcher has to submit the necessary documents so that the competent authority can evaluate whether the study objectives were met, the actual harm inflicted, and whether the severity of procedures coincided with the prospective assessments, and the number of animals used. In addition, the competent authorities must appraise any component that can advance the implementation of the 3Rs (European Parliament, 2010, Article 39).

These RAs could be extremely effective in facilitating a critical review of the use of animals in scientific procedures, if there are sufficient and qualified personnel to conduct them, as the EC's aim with these RAs is to identify 3Rs improvements and enhance transparency to the public (EC Expert Working Group for Project Evaluation and Retrospective Assessment, 2013). Publication of RA results of all studies, including those that produced negative results and may not be published elsewhere, would likely be of significant value. It would increase the knowledge base in a range of disciplines, reduce risks of duplication of studies, and inform the design of future research (EC Expert Working Group for Project Evaluation and Retrospective Assessment, 2013). However, only about one sixth of all EU Member States agreed to make the RA results publicly

available. This is not enough to meet the EC's goals. To achieve maximum benefit, access to study results should be given not only to regulatory authorities but to independent experts, in order for them to perform critical reviews of these data. And all RA results have to be made publicly available. It is possible to do so and still protect intellectual property by redacting and anonymizing certain parts of the documentation.

7.1.3   Necessary Steps

As outlined earlier, the scientific and ethical justification for animal models of human diseases depends on their providing an opportunity to investigate disease biology and to determine potentially beneficial therapies for humans (Benatar, 2007). Thus, only after an animal model has proven to have satisfactory predictive value for humans, should it be refined as much as possible to reduce pain, suffering, distress, and lasting harm. If proven of no value, it should be abandoned. Such models should no longer receive regulatory approval nor funding, nor should they be accepted by scientific journals. SRs and meta-analyses of animal models as well as RAs of all animal experiments performed by independent experts would benefit animals and human patients, as they help to identify flawed studies and to eliminate misleading, invalid models, and experimental designs. Such a rigid quality control of animal-based research would most certainly lead to a significant reduction of animal use and, thus, to an increased effort to find more animal-free, robust, human biology-based models.

7.2   *Is the Biomedical Research Industry Shifting away from Animal Use?*

> The problem is that it hasn't worked, and it's time we stopped dancing around the problem. […] We need to refocus and adapt new methodologies for use in humans to understand disease biology in humans.
>
> ZERHOUNI, former head of the US National Institutes of Health, quoted in McManus, 2013

There is growing recognition that instead of focusing efforts on trying to refine animal experiments, a primary focus on human-relevant data is needed (Collins, 2011; Giri and Bader, 2015; Langley et al., 2015, 2017; Zerhouni, 2014), as a significant challenge that medical research is facing today is the understanding and possible treatment of chronic, complex diseases of which many are not well understood and, thus, cannot be modeled in other animals (Tsukamoto, 2016). Tsukamoto asks in a Drug Discovery Today editorial: "How can we replicate human diseases that develop later in life and/or result from a prolonged unhealthy lifestyle, far beyond the lifespan of rodent animals? What makes us expect that the outcomes from carefully controlled animal

experiments can be duplicated in patients with substantial heterogeneity across various aspects (age, gender, genetics, lifestyle, disease stage, etc.)?" Transgenic mice commonly used as disease models, oftentimes contain multiple copies of presumed disease-causing transgenes, and it is dubious "whether phenotypes seen in mice as a result of this 'genetic exaggeration' have any relevance to the corresponding human diseases" (Tsukamoto, 2016). Zerhouni (2014) calls for a new approach that redirects the drug-development paradigm that commences with the patient to explore the genetic foundation of molecular changes inherent to human pathophysiology.

As Russell and Burch remarked in 1959, "refinement is never enough, and we should always seek further for reduction and if possible replacement" (Chapter 4). Since 1959, we have gathered immense knowledge about animals and their consciousness, which has led to the public acknowledgment by a group of prominent neuroscientists that other animals are conscious too: The Cambridge Declaration on Consciousness (Low, 2012). Since 1959, the technology revolution has also immensely changed the field of life sciences and, hence, provides us with the tools to move away from using animals (Langley et al., 2015, 2017).

Current legislation, reflecting societal concerns, as well as the scientific failures of animal research should function to drive research, testing, and education away from using live animals. Some areas of education and training are already using animal-free teaching approaches, for ethical reasons and educational advances (see e.g., Bones et al., 2019, Chapter 23; Pawlowski et al., 2019, Chapter 22 in this Volume). In the area of chemical-toxicity testing, some progress has already been made in finding advanced non-animal methods, initiated, for example, through the pioneering *Toxicology in the 21st Century* (*Tox21*), a US federal initiative (National Research Council, 2007; National Toxicity Program, 2004; Rovida et al., 2015; Zurlo, 2012). However, the general tendency in toxicology is to introduce new methods without eradicating all the old (animal-based) ones (Rovida et al., 2015). Still, the acceptance of animal-free alternatives by regulators without additional animal-based tests, in the pharmaceutical and food-toxicity testing fields, should be possible when proven scientifically qualified for the specific context of use. However, awareness and acceptance of scientically-valid, non-animal methods is still low among regulators as well as research workers (Ramirez et al., 2015).

The high failure rate of drugs in the clinical phase (Begley and Ellis, 2012; Food and Drug Administration, 2004; Hutchinson and Kirk, 2011; Kola and Landis, 2004; Olson et al., 2000) indicates not only poor scientific quality and cognitive bias but also that animals are not good models for humans (e.g., Greek and Kramer, 2019, Chapter 17 in this Volume; Kramer and Greek, 2018; Knight, 2019, Chapter 14 in this Volume; Leist and Hartung, 2013); and the same

applies to food-safety testing in animals (Rovida et al., 2015). Already back in 2000, an eye-opening report (Olson et al., 2000) was published about the results of a multinational pharmaceutical company survey, which served to better understand the concordance of the toxicity of pharmaceuticals in humans compared with other animals. The weakness of animal studies to predict the human toxicity of drugs became apparent, as results revealed a human toxicity concordance rate of 71% when tested in multiple rodent as well as non-rodent species. When they compared humans with rodent species only, there was a 43% correlation; humans compared with non-rodent species showed a 63% match. Drug toxicity studies in animals are long-lasting and, hence, may cause severe suffering; and they are frequently not predictive for effects in humans (Hartung, 2009).

Cumulative knowledge is essential for scientific progress. Thus, there is increasing awareness of the importance of data sharing and collaboration to shift the paradigm away from using unsound animal models for drug toxicity testing. The human toxome project, a systematic mapping of the entirety of toxicity pathways, is ongoing in the area of chemical risk assessment. Rovida et al. (2015) suggested that this project should be extended to include the assessment of efficacy and adverse effects of drugs and food ingredients. Continued reliance on animal models appears implausible to enhance the current poor rate of clinical approval of new treatments. This is why Humane Society International initiated the Biomedical Research for the 21st Century (BioMed21) Collaboration. The BioMed21 Collaboration is working internationally with health experts, regulatory and research agencies, funding bodies, and others to develop innovative research roadmaps that focus on understanding human disease pathophysiology. The goal is to further this human-focused approach to studying, preventing, and treating disease (BioMed 21 Collaboration, n.d.). A central recommendation of the BioMed21 2015 workshop was to use the Adverse Outcome Pathway (AOP) concept in biomedical research. AOP, an important concept in toxicology, describes a logical sequence of causally-linked biological events that lead from the first action of a compound to an eventual adverse effect on human health (Langley et al., 2017). Furthermore, it was recommended that technological advances should be combined in human-specific tools and models. The importance of funding these new approaches was highlighted as well as the need for faster validation and acceptance by the scientific community, funding bodies, and scientific journals, who mostly still postulate the use of animals (Langley et al., 2015, 2017).

BioMed21 is a rare example for a non-animal-based approach in the area of applied research, which—together with the field of basic research—uses the majority of animals. Overall, there is little evidence that these fields are reducing the use of animals, as the 3Rs posit we must. Quite the contrary: Animal use

has been increasing in the new century (Taylor and Rego, 2016), mainly due to an increasing generation and use of genetically altered animals (Bailey, 2019, Chapter 19 in this Volume; Carbone, 2004; Ormandy, Schuppli and Weary, 2009; Ram, 2019, Chapter 15 in this Volume), which has, in recent years, been fueled by excitement over new technologies, such as CRISPR, an easier genetic modification technique that will most probably lead to a further steep increase in animal numbers and species modified (Bailey, 2019). These new technologies, however, have not kept their promise of improving translation between animal models and human health, as they have failed to increase the efficacy and the safety of drugs (Hunter, 2011). For a detailed discussuion on the scientific and ethical issues of the genetic modification of animals, see Chapter 19 in this Volume (Bailey, 2019).

7.2.1   Funding

Progress in the development of replacement methods seems to be *limited most by the availability of funds*. Some governments and non-governmental organizations around the world are providing scarce funding, especially when compared to funds available for biomedical and life research as a whole. It is unclear how much of the annual worldwide funds—an estimated US$100 billion for biomedical research alone (Chalmers and Glasziou, 2009) and up to US$240 billion for all the life sciences (Røttingen et al., 2013)—are currently used for research centered around the use of animals, as it is not differentiated in the statistics (e.g., in Germany, BMBF, 2017). Daneshian (2016) estimated that in 2015, funds for projects with animals in Germany, including animal research facilities, were about €1920 million; funds for replacement methods ranged around €6.45 million. These financial means, mainly derived from German taxes, are distributed in opposition to Germany's declared political goal of working towards replacement of animal use at the national level (BMEL, 2015) as well as the EU level (European Parliament, 2010, Recital 10).

In preclinical human model development, the *Tissue Chips for Disease Modeling and Efficacy Testing* initiative, funded by US National Center for Advancing Translational Sciences (NCATS) of the National Institutes for Health (NIH), is a rare example. Its goal is to explore human microphysiological systems as potential facilitators of drug development in numerous disease areas. Its budget is approximately US$15 million, annually, for 13 two-year projects (NCATS, 2017); while NIH, being the biggest funder and research organization in the world, has annual funds of about US$39 billion for medical research alone (NIH, 2019). The EU framework program for research and innovation, Horizon 2020 (European Commission, n.d.), has, at the time of writing, supported 16 research projects devoted to alternative methods to animal testing, with a total of €90 million (European Parliament, 2017). The main research activities

are targeted towards developing complex *in silico* and *in vitro* human-based systems for better and more cost-effective safety and efficacy testing of chemicals, nanoparticles, vaccines, and drugs (European Parliament, 2017).

Between 1981 and 2015, the German Federal Ministry for Education and Research (Bundesministerium für Bildung und Forschung, BMBF) gave €160 million in funding for over 500 3Rs research projects. Aside from not exclusively funding replacement projects, the funds dedicated to the 3Rs were sparse; for example, in the 6-year period between 2010 and 2015, less than €20 million were available (BMBF, 2016). The UK National Centre of the Replacement, Refinement and Reduction of Animals in Research (NC3Rs) is the largest funder of such research in the United Kingdom (Burden et al., 2015). Between 2004 and 2014, NC3Rs awarded 200 grants worth approximately US$54 million (Burden et al., 2015); the annual overall budget of NC3Rs is approximately €11.2 million (NC3Rs, n.d. d). In contrast, the German national 3Rs center, Zentrum zum Schutz der Versuchstiere (Bf3R), has an annual budget of €1.5 million to run all of its operations (Bundesinstitut für Risikobewertung, 2016) and provides approximately €350,000 to external replacement and refinement research groups per year (Bundesinstitut für Risikobewertung, 2018). Replacement research has to compete with refinement research for these limited funds (BMBF, 2016; NC3Rs, n.d. a).

A donor that exclusively provides money for the first R is the cosmetic company Lush, which in 2012 established the Lush Prize in collaboration with the UK not-for-profit group, Ethical Consumer Research Association (Redmond, 2019, Chapter 27 in this Volume). Lush provides £250,000 in funding each year for the main prize categories, with additional funds for regional awards in Asia and the Americas (Lush Prize, n.d.). An example for a charity providing some funding is People for the Ethical Treatment of Animals (PETA) International Science Consortium (PISC), which, in June 2017, awarded funding to develop four *in vitro* exposure systems to researchers from institutions in the United Kingdom, United States, and Belgium that are leaders in the development of non-animal methods to test the toxicity of airborne substances (PETA International Science Consortium, n.d. a). PISC assists with funding where promising *in vitro* or *in silico* techniques require further development or validation in order to gain regulatory acceptance. PISC focuses on toxicology and until, 2017, it has contributed about €2.9 million towards improving and implementing non-animal research methods (PETA International Science Consortium, n.d. b). The Alternatives Research & Development Foundation (ARDF) funds and promotes the development and validation of non-animal methods in biomedical research, product testing, and education and has provided US$3.25 million in funds since 1993 (Alternatives Research & Development Foundation, 2018).

The US National Anti-Vivisection Society (NAVS) provides some grants through the International Foundation for Ethical Research (IFER) for early career scientists to develop humane, human-relevant alternatives that replace animal use (NAVS, 2018). Overall, there are a few local and international initiatives and prizes but most focus on animal testing, while non-animal approaches in basic and applied research lag behind. Moreover, to ensure the field of animal-free, human-based research methods and approaches is continually and substantially growing, increased, stable governmental funding must be provided.

7.2.2   Education and Training

Another obstacle in shifting the current research paradigm is the limited availability of educational and training courses on animal-free methods and approaches in all areas of biomedical science, but especially in basic and applied research, since current available guidance documents and databases as well as courses almost exclusively focus on testing alternatives. There are some efforts being made to improve experimental design, conduct, and reporting; for example, online resources are available at some of the national 3Rs centers, such as at Norecopa, Norway's National Consensus Platform for the advancement of the 3Rs (Norecopa, 2016b) and the UK NC3Rs (NC3Rs, n.d. b, c), since quality issues of biomedical research has become apparent.

By EU law, the researcher must be well informed about state-of-the-art developments in the field of investigation, and animals must only be used if all possible alternatives are considered to be inadequate (EC Joint Research Centre, 2013). The EC Joint Research Centre's EU Reference Laboratory for Alternatives to Animal Testing – European Centre for the Validation of Alternative Methods (EURL ECVAM) Search Guide (EC Joint Research Centre, 2013) and Data Base Service on Alternative Methods to animal experimentation (DB-ALM) (EC Joint Research Centre, 2017) ought to assist with the search for alternatives to animal use. However, even for experts in the respective field, it is a lengthy and difficult task, as existing search systems do not support the necessary search strategies.

Altertox Academy, formerly CAAT Academy, offers hands-on training, but primarily for toxicologists, in human-relevant alternative methods and technologies (Altertox Academy, 2018). Education and training courses, mandatory for all animal researchers in the EU, include one animal-free methods module (e.g., FELASA B courses), but of a 40 hour FELASA B course, about one hour is dedicated to replacements, and generally only alternatives used in toxicology testing are covered (e.g., Berliner Kompaktkurse, 2017, p. 23). In 2016, the University of California (UC) San Diego offered a course that introduces participants to the available non-animal research methods, their efficacy, and how to

identify and implement them. It covered more areas than just regulatory toxicology (UC San Diego, 2018). However, detailed courses with extensive modules for all areas of the biomedical sciences currently do not exist.

7.2.3   Search Engine for Alternative Methods

What is urgently needed—aside from specific education and training courses—is an unbiased, freely available search engine that is able to find correlations regarding scientific purpose between animal experiments and alternative methods and, at the same time, 3Rs-relevant deviances in the methodologies (*in vitro* versus *in vivo*). Scientists from the Leibniz Institute for Social Sciences (GESIS) and the German Federal Institute for Risk Assessment (Bundesinstitut für Risikobewertung, BfR) have laid the foundation for such a search engine, using machine learning. The project is called SMAFIRA, which stands for "smart feature-based interactive ranking algorithm." The goal of SMAFIRA is to develop automated but mechanistically transparent search procedures that focus on such deviations and, thus, to provide an improved automatic support to search for non-animal methods (fisaonline, n.d.). This search engine will drastically reduce the number of documents scientists have to go through (GESIS, n.d.). A first version of the SMAFIRA search engine is anticipated to be available in the second half of 2019 (Daniel Butzke, BfR, personal communication, January 2019).

8   Ways to Work Towards Replacement

Directive 2010/63/EU, a progressive animal protection legislation in the field, sums up some important steps that have to be taken to work towards a paradigm shift, when it states (emphasis added): "The availability of alternative methods is highly dependent on the progress of the research into the development of alternatives. [...] the Commission and the Member States should contribute through research and by other means to the **development and validation of alternative approaches**." (Recital 46). Article 47 declares: "The **Commission** and the **Member States shall contribute to the development and validation of alternative approaches** which could provide the same or higher levels of information as those obtained in procedures using animals [...], and they **shall take such other steps as they consider appropriate to encourage research in this field**. [...] Member States should, at national level, ensure the **promotion of alternative approaches and the dissemination of information** [...]".

8.1   *Political Engagement*

The needed political engagement that Directive 2010/63/EU demands from its Member States to move towards an animal-free world of scientific experimentation was made a priority by the Dutch government in 2016. The Netherlands National Committee for the protection of animals used for scientific purposes (NCad) has developed a vision and plan of action for moving away from laboratory animal use. The Dutch goal is to phase out the utilization of animals in a number of fields by 2025, namely in regulatory testing of chemicals, food ingredients, pesticides and (veterinary) medicines, and biological products, such as vaccines (NCad, n.d.). The Committee also plans to steadily reduce animal involvement in regulatory preclinical research and basic research: "If we are to make the transition to non-animal research methods, we must make a paradigm shift away from existing mindsets and practices" (NCad, n.d., p. 3), a task which seems to be impossible without political involvement. The Dutch strategy holds the potential to act as a driver for other countries to follow this path.

8.2   *Legislative Change*

There is a need for regulators who are brave to move legislative change forward. The reason for the continued use of animals for regulatory testing is legislative, as existing policies require that new drug candidates are tested on animals before they can be assessed in human clinical trials, regardless of the fact that these animal tests are often unreliable in assessing safety and efficacy in humans (Greek and Kramer, 2019, Chapter 17 in this Voume). These regulations need to be amended according to scientific knowledge, and serious efforts need to be made to accelerate the development of advanced, humane, and human-relevant models (Archibald, Coleman and Drake, 2019, Chapter 18 in this Volume).

8.3   *Redeployment of Funds*

Absolutely essential for the paradigm change towards advanced, animal-free science and better healthcare is the redirection of funding. The limited funding for replacement research, oftentimes, has to compete with refinement research (e.g., BMBF, 2016; NC3Rs, n.d. a). These scarce funds should be used to further human biology-based approaches. Also, regarding taxpayers' money, the national governments, arguably, have the responsibility to use the funds in the name of a society that has repeatedly voiced that more needs to be done to replace animals in science. Moreover, our society is ethically evolving, with evidence of dwindling acceptance for animal suffering in the name of science. And it is being increasingly acknowledged that the continued reliance on animal models is unlikely to improve significantly the currently poor rate of

clinical approval of new treatments. Thus, animal-based research also contributes to resources being wasted (Harris, 2017; Ioannidis et al., 2014; Keen, 2019, Chapter 10 in this Volume).

Aside from redeploying funds, partially to preclinical human-relevant disease research (Langley et al., 2017) and to clinical rather than basic research (Pound and Bracken, 2014), a large part of funding should be dedicated to disease prevention efforts. To combat the increasing prevalence of dementia, for example, human-focused, non-animal models and methods, such as computational methods, advanced brain imaging techniques, and epidemiological studies should be given funding preference (Pistollato et al., 2016). Another extremely important area of disease prevention is basic public healthcare (Marks, 2012) as well as nutrition and lifestyle education. In addition, funds should also be used for pollution control, as pollution is currently found to be the largest environmental cause of disease and premature death around the world (Landrigan et al., 2017). The World Health Organization (WHO) estimated that around 3 million people die prematurely every year due to air pollution alone (Watts et al., 2017). In 2015, diseases caused by pollution were responsible for about 16% (9 Mio.) of all human deaths worldwide, which is three times more than deaths from tuberculosis, malaria, and AIDS combined and 15 times more than all wars and other means of violence together (Landrigan et al., 2017).

8.4   *Education and Training*

Education as well as re- and ongoing training about how to conduct *state-of-the-art science* and report it properly, as well as education on *research ethics* and *bioethics* are crucial. They enable students and scientists to gain a solid grounding in science based on non-animal models, while sincerely embracing the hierarchy of the 3Rs. Such learning objectives should be made available and should be mandatory for everyone planning to work or working in biomedical science. Education and retraining are the most important means to move away from the current thought culture and practice of animal use towards a new, humane research paradigm.

8.5   *Scientific Collaboration*

As Russell and Burch observed in 1959, "As we shall see, replacement is widely used in some fields, while in others it is very far from being exploited to the full, if at all. Moreover, such developments have been largely empirical, and largely independent of each other" (Chapter 5). At the moment, 3Rs experts are divided into replacement experts, on the one hand, and refinement experts, on the other. Animal welfare bodies and national committees in the EU (Directive 2010/63/EU, Recital 48), for example, are supposed to advise

scientists about the application of the 3Rs but seem to have little to no knowledge about available replacements and novel animal-free approaches to scientific questions (van Luijk et al., 2012; van Luijk et al., 2013). To achieve the ultimate goal in shifting the focus from refinement of animal use to replacement of animal use the animal research community needs to engage with replacement experts. National 3Rs centers should be equipped with a majority of experts in replacement methods, and a close collaboration between replacement experts and animal researchers appears crucial in moving towards animal replacement. To accelerate the development of new human biology-based approaches, a multidisciplinary approach is essential for bringing together the newest technologies and experts from various disciplines (Langely et al., 2017; Noor, 2019, Chapter 25 in this Volume).

9 Final Remarks

Looking into the future of animal-based science, Carbone (2004) wrote that morality and politics will continue to be the drivers for replacement research. Since the introduction of the principles, it has been widely held that animal researchers have an ethical responsibility to minimize any pain, distress, fear, suffering, and harm caused to animals when keeping them confined and utilizing them for invasive experiments without their consent. To apply the knowledge gained through animal welfare and refinement research is good veterinary and scientific practice, but it is not a substitute for reduction and replacement of animal experimentation. Indeed, Balls warned "that refinement can be used as a convenient way of showing commitment to the 3Rs, while ensuring that animal experimentation is seen as respectable and can be allowed to continue, while the fundamental ethical questions raised by it are avoided" (2010, p. 21). Thus, we have to be on guard that refinement is not used as a whitewashing tool, but its full application, which is an ethical imperative, must be guaranteed during the transition to human-relevant, animal-free methodologies.

Aside from extensive flaws in the way the majority of animals are housed and treated, and the poor conduct and reporting of many animal studies, the general lack of transparency around the use of animals in research as well as the low rate of critical appraisal of animal experiments are apparent. These failings have led to incorrect data and an overestimation of their significance (Cohen, 2018). Unnecessary harm inflicted upon these animals and, in the case of medical research, the harms done to patients who suffer from adverse reactions to drugs that were tested safe in animals or who are urgently waiting for treatments are serious issues that need to be addressed. A commitment to adhere to the 3Rs and to good scientific practice as well as to address societal

concerns about the use of animals in science would require a strong shift away from animals towards the use of human-relevant approaches.

References

Aarts, A.A., J.E. Anderson, C.J. Anderson, P.R. Attridge, A. Attwood and A. Fedor (2015). Estimating the Reproducibility of Psychological Science. *Science*, 349(6251), pp. 1–8.

Abou-Ismail, U.A. and H.D. Mahboub (2011). The Effects of Enriching Laboratory Cages Using Various Physical Structures on Multiple Measures of Welfare in Singly-Housed Rats. *Laboratory Animals*, 45(3), pp. 145–153.

Akhtar, A.Z., J.J. Pippin and C.B. Sandusky (2009). Animal Studies in Spinal Cord Injury: A Systematic Review of Methylprednisolone. *Alternatives to Laboratory Animals*, 37(1), pp. 43–62.

AALAS (n.d.). AALAS National Meeting Abstract Archive. [online] Available at: https://www.aalas.org/national-meeting/abstract-archive [Accessed 13 August 2018].

Altertox Academy (2018). *Hands-on lab trainings and events in alternative methods in toxicology.* [online] Available at: http://academy.altertox.be [Accessed 12 August 2018].

Altman, D.G. (1994). The Scandal of Poor Medical Research. *British Medical Journal*, 308(6924), pp. 283–284.

Altman, D.G. (1980). Statistics and Ethics in Medical Research. Misuse of Statistics is Unethical. *British Medical Journal*, 281(6249), pp. 1182–1184.

Alternatives Research & Development Foundation (2018). *ARDF Annual Open Grant Program.* [online] Available at: http://www.ardf-online.org/ardf-grants.html [Accessed 12 August 2018].

André, V., C. Gau, A. Scheideler, J.A. Aguilar-Pimentel, O.V. Amarie, L. Becker, L. Garrett, W. Hans, S.M. Hölter, D. Janik and K. Moreth (2018). Laboratory Mouse Housing Conditions Can Be Improved Using Common Environmental Enrichment Without Compromising Data. *PLoS Biology*, 16(4), p. e2005019. [online] Available at: http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2005019 [Accessed 18 April 2018].

Anselme, P., M.J. Robinson and K.C. Berridge (2013). Reward Uncertainty Enhances Incentive Salience Attribution as Sign-tracking. *Behavioural Brain Research*, 238, pp. 53–61.

Archibald, K., R. Coleman and T. Drake (2019). Replacing animal tests to improve safety for humans. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill, pp. 417–442.

Augustsson, H., H.A. van de Weerd, C.L. Kruitwagen, V. Baumans (2003). Effect on Enrichment on Variation and Results in Light/Dark Test. *Laboratory Animals*, 37(4), pp. 328–340.

Bailey, J. (2018). Does the stress of laboratory life and experimentation on animals adversely affect research data? A critical review. *Alternatives to Laboratory Animals*, 46(5), pp. 291–305.

Bailey, J. (2019). Genetic modification of animals: scientific and ethical issues. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill pp. 443–479.

Bailoo, J.D., T.S. Reichlin and H. Würbel (2014). Refinement of Experimental Design and Conduct in Laboratory Animal Research. *ILAR Journal*, 55(3), pp. 383–391.

Baker, M. (2016). 1,500 Scientists Lift the Lid on Reproducibility. *Nature News*, 533(7604), p. 452.

Baker, D., K. Lidster, A. Sottomayor and S. Amor (2014). Two Years Later: Journals Are Not Yet Enforcing the ARRIVE Guidelines on Reporting Standards for Preclinical Animal Studies. *PLoS Biology*, 12(1), p. e1001756. [online] Available at: http://jour nals.plos.org/plosbiology/%20article?id=10.1371/journal.pbio.1001756 [Accessed 12 August 2018].

Balcombe, J. (2010). Laboratory rodent welfare: thinking outside the cage. *Journal of Applied Animal Welfare Science*, 13(1), pp. 77-88.

Balcombe, J.P., N.D. Barnard and C. Sandusky (2004). Laboratory Routines Cause Animal Stress. *Journal of the American Association for Laboratory Animal Science*, 43(6), pp. 42–51.

Baldwin, A.L., R.L. Primeau and W.E. Johnson (2006). Effect of Noise on the Morphology of the Intestinal Mucosa in Laboratory Rats. *Journal of the American Association for Laboratory Animal Science*, 45, pp. 74–82.

Balls, M. (1999). The biomedical sciences and the need for less inhumane animal procedures. In: C.F.M. Hendriksen and D.B. Morton, eds., *Humane Endpoints in Animal Experiments for Biomedical Research*. London: Royal Society of Medicine Press, pp. 1–4.

Balls, M. (2010). The Principles of Humane Experimental Technique: Timeless Insights and Unheeded Warnings. *Alternatives to Animal Experimentation*, 27(2), pp. 144–148. [online] Available at: http://www.altex.ch/resources/altex_2010_2_144_148_Balls .pdf [Accessed 12 August 2018].

Bath, P.M.W., L.J. Gray, A.J.G. Bath, A. Buchan, T. Miyata and A.R. Green (2009). Effects of NXY-059 in Experimental Stroke: An Individual Animal Meta-analysis. *British Journal of Pharmacology*,157(7), pp. 1157–1171.

Baumans, V. and P. van Loo (2013). How to Improve Housing Conditions of Laboratory Animals: The Possibilities of Environmental Refinement. *The Veterinary Journal*, 195(1), pp. 24–32.

Bayne, K. (2005). Potential for Unintended Consequences of Environmental Enrichment for Laboratory Animals and Research Results. *ILAR Journal*, 46(2), pp. 129–139.

Bayne, K. and H. Würbel (2014). The Impact of Environmental Enrichment on The Outcome Variability and Scientific Validity of Laboratory Animal Studies. *Revue*

Scientifique et Technique de l'OIE, 33(1), pp. 273–280. [online] Available at: https://www.oie.int/doc/ged/D13675.PDF [Accessed 12 August 2018].

Bayne, K., G.S. Ramachandra, E.A. Rivera and J. Wang (2015). The Evolution of Animal Welfare and the 3Rs in Brazil, China, and India. *Journal of the American Association for Laboratory Animal Science*, 54(2), pp. 181–191.

Begley, C.G. and L.M. Ellis (2012). Drug Development: Raise Standards for Preclinical Cancer Research. *Nature*, 483(7391), pp. 531–533.

Begley, C.G. and J.P.A. Ioannidis (2015). Reproducibility in Science. *Circulation Research*, 116(1), pp. 116–126.

Benatar, M. (2007). Lost in Translation: Treatment Trials in the SOD1 Mouse and in Human ALS. *Neurobiology of Disease*, 26(1), pp. 1–13.

Benefiel, A.C., W.K. Dong and W.T. Greenough (2005). Mandatory "Enriched" Housing of Laboratory Animals: The Need for Evidence-Based Evaluation. *ILAR Journal*, 46(2), pp. 95–105.

Bennett, E.L., M.C. Diamond, D. Krech and M.R. Rosenzweig (1964). Chemical and Anatomical Plasticity of Brain. *Science*, 146(3644), pp. 610–619.

Berliner Kompaktkurse (2017). *Laboratory Animal Science—Basic Course Main Focus: Mice/Rats*, p. 23. [online] Available at: https://www.berliner-kompaktkurse.de/assets/files/downloads/ph_kk_2017-01_web.pdf [Accessed 1 July 2017].

Bertolucci, L.F. (2011). Pandiculation: nature's way of maintaining the functional integrity of the myofascial system?. *Journal of Bodywork and Movement Therapies,* 15(3), pp. 268–280.

Bertrand, H.G.M.J., C. Sandersen and P.A. Flecknell (2018). Reported analgesic and anaesthetic administration to non-human primates undergoing experimental surgical procedure: 2010–2015. *Journal of Medical Primatology,* 47, pp. 217–225.

Beura, L., S. Hamilton, K. Bi, J. Schenkel, O. Odumade, K. Casey, E. Thompson, K. Fraser, P. Rosato, A. Filali-Mouhim, R. Sekaly, M. Jenkins, V. Vezys, W. Haining, S. Jameson and D. Masopust (2016). Normalizing the Environment Recapitulates Adult Human Immune Traits in Laboratory Mice. *Nature*, 532(7600), pp. 512–516.

BioMed21 Collaboration (n.d.). *About us*. [online] Available at: http://biomed21.org/about-us/ [Accessed on 23 May 2018].

BMBF (2016). *Alternativen zum Tierversuch. Informationen zur Förderung.* [online] Availabale at: https://www.bmbf.de/pub/Alternativmethoden_zum_Tierversuch.pdf [Accessed 1 August 2017].

Birke, L., Arluke, A. and Michael, M. (2007). *The Sacrifice: How Scientific Experiments Transform Animals and People.* Purdue University Press.

BMEL (2015). *"Mein langfristiges Ziel ist es, Tierversuche komplett zu ersetzen."* Bundesminister Schmidt eröffnet das Deutsche Zentrum zum Schutz von Versuchstieren. [online] Berlin, 25. September 2015. Available at: https://www.bmel.de/DE/Tier/Tierschutz/_texte/TierschutzTierforschung.html;jsessionid=DB4703F553CD731C7EA41D8F642C806E.2_cid367?docId=6671414 [Accessed 12 August 2018].

Bones, V.C., R.C.M. Garcia, G.G. Alves, R.L. Paixão, A.A. Rocha, K.V. Capilé and R. Bachinski (2019). Humane education: the tool for scientific revolution in Brazil. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill.

Bracken, M.B. (2009). Why animal studies are often poor predictors of human reactions to exposure. *Journal of the Royal Society of Medicine*, 102(3), pp. 120–122.

Briel, M., K.F. Müller, J.J. Meerpohl, von E. Elm, B. Lang, E. Motschall, V. Gloy, F. Lamontagne, G. Schwarzer and D. Bassler (2013). Publication bias in animal research: A systematic review protocol. *Systematic Reviews*, 2, p. 23. [online] Available at: https://systematicreviewsjournal.biomedcentral.com/track/pdf/10.1186/2046-4053-2-23 [Accessed 12 August 2018].

Brown, M.J. (2014). Creating a culture of care. *NC3Rs News & Blog Online*. [online] Available at: https://www.nc3rs.org.uk/news/creating-culture-care [Accessed 12 August 2018].

Bundesinstitut für Risikobewertung (2016). *Fragen und Antworten zum Deutschen Zentrum zum Schutz von Versuchstieren (Bf3R).* [online] Available at: http://www.bfr.bund.de/cm/343/fragen-und-antworten-zum-deutschen-zentrum-zum-schutz-von-versuchstieren-bf3r.pdf [Accessed 12 August 2018].

Bundesinstitut für Risikobewertung (2018). *Bf3R Research Funding in the area of 3R – Replacement, Reduction and Refinement.* [online] Available at: https://www.bfr.bund.de/en/bf3r_research_funding_in_the_area_of_3r___replacement__reduction_and_refinement-62825.html [Accessed 12 August 2018].

Burden, N., K. Chapman, F. Sewell and V. Robinson (2015). Pioneering Better Science through the 3Rs: An Introduction to the National Centre for the Replacement, Refinement, and Reduction of Animals in Research (NC3Rs). *Journal of the American Association for Laboratory Animal Science*, 54(2), pp. 198–208.

Burghardt, G.M. (1996). Environmental enrichment or controlled deprivation? In: G.M. Burghardt, J.T. Bielitzki, J.R. Boyce and D.O. Schaefer, eds., *The Well-being of Animals in Zoo and Aquarium Sponsored Research*. Greenbelt, MD: Scientists Center for Animal Welfare, pp. 91–101.

Burghardt, G.M. (1999). Deprivation and Enrichment in Laboratory Animal Environments, *Journal of Applied Animal Welfare Science*, 2(4), pp. 263–266.

Burn, C.C. (2017). Bestial Boredom: A Biological Perspective on Animal Boredom and Suggestions for Its Scientific Investigation. *Animal Behaviour*, 130, pp. 141–151.

Burn, C.C., A. Peters, M.J. Day and G.J. Mason (2006). Long-Term Effects of Cage-Cleaning Frequency And Bedding Type on Laboratory Rat Health, Welfare, and Handleability: A Cross-Laboratory Study. *Laboratory Animals*, 40(4), pp. 353–370.

Cabib, S. (2006). The Neurobiology of stereotypy II: The role of stress. In: G. Mason and J. Rushen, eds., *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare*, 2nd ed. Oxfordshire, UK: CAB International.

CAMARADES (2014). *Bringing Evidence to Translational Medicine*. [online] Available at: http://www.dcn.ed.ac.uk/camarades/default.htm [Accessed 12 August 2018].

Campi, K.L., C.E. Collins, W.D. Todd, J. Kaas and L. Krubitzer (2011). Comparison of Area 17 Cellular Composition in Laboratory and Wild-Caught Rats Including Diurnal and Nocturnal Species. *Brain, Behavior and Evolution*, 77, pp. 116–130.

Cao, L., X. Liu, E. Lin, C. Wang, E. Choi, V. Riban, B. Lin and M. During (2010). Environmental and Genetic Activation of a Brain-Adipocyte BDNF/Leptin Axis Causes Cancer Remission and Inhibition. *Cell*, 142(1), pp. 52–64.

Carbone, L. (2004). What Animals Want: Expertise and Advocacy in Laboratory Animal Welfare Policy. New York: Oxford University Press.

Carbone, L. and J. Austin (2016). Pain and Laboratory Animals: Publication Practices for Better Data Reproducibility and Better Animal Welfare. *PLoS One*, 11(5). [online] Avaiable at: http://journals.plos.org/plosone/article/related?id=10.1371/journal.pone.0155001 [Accessed 12 August].

Cavanaugh, S.E., Pippin, J.J. and Barnard, N.D. (2014). Animal Models of Alzheimer Disease: Historical Pitfalls and a Path Forward. Alternatives to Animal Experimentation, 31(3), pp. 279–302.

Center for Open Science (n.d.). *Our Mission is to Increase Openness, Integrity, and Reproducibility of Research*. [online] Available at: https://cos.io/about/mission/ [Accessed 12 August 2018].

Chalmers, I. and P. Glasziou (2009). Avoidable Waste in the Production and Reporting of Research Evidence. *Obstetrics & Gynecology*, 114(6), pp. 1341–1345.

Charbonneau, R., L. Niel, E. Olfert, M. von Keyserlingk and C. Griffin (2010). CCAC Guidelines on: Euthanasia of Animals Used in Science. *Canadian Council on Animal Care.* Ottawa ON, Canada, pp. 1–32. [online] Available at: https://www.ccac.ca/Documents/Standards/Guidelines/Euthanasia.pdf [Accessed 12 August 2018].

Chawla, S.D. (2017). 'One-size-fits-all' threshold for *P* values under fire. Scientists hit back at a proposal to make it tougher to call findings statistically significant. *Nature News*. [online] Available at: https://www.nature.com/news/one-size-fits-all-threshold-for-p-values-under-fire-1.22625?WT.mc_id=TWT_NatureNews [Accessed 12 August 2018].

Chesler, E.J., S.G. Wilson, W.R. Lariviere, S.L. Rodriguez-Zas and J.S. Mogil (2002). Influences of Laboratory Environment on Behavior. *Nature Neuroscience,* 5, pp. 1101–1102.

Clarkson, J.M., D.M. Dwyer, P.A. Flecknell, M.C. Leach and C. Rowe (2018). Handling method alters the hedonic value of reward in laboratory mice. *Scientific Reports*, 8(1), p. 2448. [online] Available at: https://www.nature.com/articles/s41598-018-20716-3 [Accessed 12 August 2018].

Clemence, M. and J. Leaman (2016). *Public Attitudes to Animal Research in 2016*. A report by Ipsos MORI for the Department for Business, Energy & Industrial Strategy, Ipsos MORI Social Research Instititute. [online] Available at: https://www

.ipsos.com/sites/default/files/publication/1970-01/sri-public-attitudes-to-animal
-research-2016.pdf [Accessed 12 August 2018].

Cohen, D. (2018). Oxford TB Vaccine Study Calls into Question Selective Use of Animal
Data. *British Medical Journal*, 360, p. j5845.

Collins, F.S. (2011). Reengineering Translational Science: The Time Is Right. *Science
Translational Medicine*, 3(90), p. 90cm17. [online] Available at: https://www.ncbi
.nlm.nih.gov/pmc/articles/PMC5101940/ [Accessed 12 August 2018].

Collins, F.S. and L.A. Tabak (2014). NIH Plans To Enhance Reproducibility. *Nature,* 505,
pp. 612–613.

Commission of the European Communities (2007). Recommendation of 18 June 2007 on
Guidelines for the accommodation and Care of Animals Used for Experimental and
Other Scientific Purposes (2007/526/EC). *Official Journal of the European Communi-
ties*, L197, pp. 1–89. [online] Available at: https://eur-lex.europa.eu/legal-content/
EN/TXT/PDF/?uri=CELEX:32007H0526&from=EN [Accessed 15 August 2017].

Coulter, C., P. Flecknell and C. Richardson (2009). Reported Analgesic Administration
to Rabbits, Pigs, Sheep, Dogs and Non-Human Primates Undergoing Experimental
Surgical Procedures. *Laboratory Animals*, 43(3), pp. 232–238.

Coulter, C., P. Flecknell, M. Leach and C. Richardson (2011). Reported Analgesic Ad-
ministration To Rabbits Undergoing Experimental Surgical Procedures. *BMC Veteri-
nary Research*, 7(1), p. 12. [online] Available at: https://doi.org/10.1186/1746-6148-7-12
[Accessed 12 August 2018].

Council of Europe (1986). *Convention for the Protection of Vertebrate Animals Used for
Experimental and other Scientific Purposes.* European Treaty Series No. 123. [online]
Available at: https://rm.coe.int/168007a67b [Accessed 12 August 2018].

Crabbe, J.C., D. Wahlsten and B.C. Dudek (1999). Genetics of Mouse Behavior: Interac-
tions with Laboratory Environment. *Science*, 284(5420), pp. 1670–1672.

Daneshian, M. (2016). Vergleiche der Förderungen von Alternativen in Deutschland –
Europa – USA. In: *Forschung ohne Tierversuche – verliert Deutschland den Anschluss?
Zum Weltaktionstag gegen Tierversuche.* Talk, Berlin. [online] Program available at:
https://www.urania.de/forschung-ohne-tierversuche-verliert-deutschland-den
-anschluss [Accessed 6 August 2017].

Daneshian, M., F. Busquet, T. Hartung and M. Leist (2015). Animal Use for Science in
Europe. *Alternatives to Animal Experimentation,* 32, pp. 261–274.

Dawson, E., T. Gilovich and D.T. Regan (2002). Motivated Reasoning and Performance
on the was on Selection Task. *Personality and Social Psychology Bulletin*, 28(10),
pp. 1379–1387.

de Vries, R.B., P. Buma, M. Leenaars, M. Ritskes-Hoitinga and B. Gordijn (2012). Re-
ducing the Number of Laboratory Animals Used in Tissue Engineering Research by
Restricting the Variety of Animal Models. Articular Cartilage Tissue Engineering as
a Case Study. *Tissue Engineering Part B: Reviews*, 18(6), pp. 427–435.

Dirnagl, U. and M. Fisher (2012). International, Multicenter Randomized Preclinical Trials in Translational Stroke Research: It's Time To Act. *Journal of Cerebral Blood Flow & Metabolism*, 32, pp. 933–935.

During, M.J., X. Liu, W. Huang, D. Magee, A. Slater, T. McMurphy, C. Wang and L. Cao (2015). Adipose VEGF Links the White-To-Brown Fat Switch with Environmental, Genetic, and Pharmacological Stimuli in Male Mice. *Endocrinology,* 156, pp. 2059–2073.

EC Expert Working Group for Project Evaluation and Retrospective Assessment (2013). 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. [online] Available at: http://ec.europa.eu/environment/chemicals/lab_animals/pdf/Endorsed_PE-RA.pdf [Accessed 12 August 2017].

EC Joint Research Centre (2013). *The EURL ECVAM Search Guide. Good Search Practice on Animal Alternatives.* EUR 24391 EN, Re-edition. [online] Available at: https://ec.europa.eu/jrc/en/publication/books/eurl-ecvam-search-guide-good-search-practice-animal-alternatives [Accessed 12 August 2017].

EC Joint Research Centre (2017). *EURL ECVAM Data Base Service on Alternative Methods to Animal Experimentation* (DB-ALM). [online] Availble at: https://ecvam-dbalm.jrc.ec.europa.eu/methods-and-protocols [Accessed 12 August 2018].

Eisen, J.A., E. Ganley and C.J. MacCallum (2014). Open Science and Reporting Animal Studies: Who's Accountable? *PLoS Biology,* 12(1), p. e1001757. [online] Available at: http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001757 [Accessed 12 August 2018].

Enserink, M. (2017). Sloppy reporting on animal studies proves hard to change. *Science*, 357(6358), pp. 1337–1338.

Errington, T.M., E. Iorns, W. Gunn, F.E. Tan, J. Lomax and B.A. Nosek (2014). An Open Investigation of the Reproducibility of Cancer Biology Research. *Elife*, 3, p. e04333.

Eskola, S., M. Lauhikari, H.M. Voipio, M. Laitinen and T. Nevalainen (1999). Environmental Enrichment May Alter the Number of Rats Needed To Achieve Statistical Significance. *Scandinavian Journal of Laboratory Animal Science,* 26, pp. 134–144.

European Citizen's Initiative (2016). *Stop Vivisection.* [online] Available at: http://www.stopvivisection.eu/de [Accessed 12 August 2018].

European Commission (2013). Report from the Commission to the Council and the European Parliament. Seventh Report on the Statistics on the Number of Animals used for Experimental and other Scientific Purposes in the Member States of the European Union. *COM (2013) 859 final.* [online] Available at: https://eur-lex.europa.eu/resource.html?uri=cellar:e99d2a56-32fc-4f60-ad69-61ead7e377e8.0001.03/DOC_1&format=PDF [Accessed 12 August 2018].

European Commission (2014). National Competent Authorities for the implementation of Directive 2010/63/EU on the Protection of Animals Used for Scientific Purposes.

A working document on the development of a common education and training framework to fulfil the requirements under the Directive. [online] Available at: http://ec.europa.eu/environment/chemicals/lab_animals/pdf/Endorsed_E-T.pdf [Accessed 15 July 2018].

European Commission (n.d.). *Horizon2020. The EU Framework Programme for Research and Innovation.* [online] Availble at: https://ec.europa.eu/programmes/horizon2020/en/find-your-area [Accessed 12 August 2018].

European Parliament (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 Communities*, L276, pp. 33–79. [online] Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32010L0063 [Accessed 23 July 2017].

European Parliament (2017). *Parliamentary questions: Question Reference E-004265/2017.* Answer given by Mr Moedas on behalf of the Commission. [online] Available at: http://www.europarl.europa.eu/doceo/document/E-8-2017-004265-ASW_EN.html [Accessed 12 August 2018].

EUSAAT (2015). *19th European Congress on Alternatives to Animal Testing, 16th Annual Congress of EUSAAT*. Alternatives to Animal Experimentation Proceedings, 4(2). [online] Available at: http://www.eusaat-congress.eu/images/2015/Abstractbook_EUSAAT_2015_Linz_2015.pdf [Accessed 12 August 2018].

FELASA (2016). *Detailed program FELASA 2016.* [online] Available at: http://felasa2016.eu/wp-content/uploads/2016/06/DETAILED_PROGRAMME_FELASA_2016_V3.6.pdf [Accessed 12 August 2018].

fisaonline (n.d.). SMAFIRA – Smart Feature-based interactive ranking (SMAFIRA) Project. [online]. Available at: https://fisaonline.de/en/find-institutions/federal-research-institutions/?tx_fisaresearch_federalresearchinstitutions%5Bp_id%5D=4377&tx_fisaresearch_federalresearchinstitutions%5Baction%5D=projectDetails&tx_fisaresearch_federalresearchinstitutions%5Bcontroller%5D=Projects&cHash=1223e8b07c976b94f16b67e2f0e4af1b#more [Accessed 23 Janaury 2018].

Flecknell, P.A. (2016). *Laboratory Animal Anaesthesia.* 4th ed., London: Academic Press.

Flecknell, P.A. (2018). Rodent Analgesia: Assessment and Therapeutics. *The Veterinary Journal,* 232, pp. 70–77.

Fonseca, K., P. Rodrigues, I. Olsson and M. Saraiva (2017). Experimental Study Of Tuberculosis: From Animal Models To Complex Cell Systems and Organoids. *PLoS Pathogens*, 13(8) [online] Available at: https://doi.org/10.1371/journal.ppat.1006421 [Accessed 12 August 2018].

Food and Drug Administration (2018). *Food and Drug Administration Modernization Act (FDAMA)* [online] Available at: https://www.fda.gov/RegulatoryInformation/LawsEnforcedbyFDA/SignificantAmendmentstotheFDCAct/FDAMA/default.htm [Accessed 12 August 2018].

Food and Drug Administration (2004). *Innovation or Stagnation? Challenge and Opportunity on the Critical Path to New Medical Products.* US Department of Health and Human Services, Food and Drug Administration, pp. 1–32. [online] Available at: http://www.who.int/intellectualproperty/documents/en/FDAproposals.pdf [Accessed 12 August 2018].

Franco, N.H., Sandøe, P. and Olsson, I.A.S. (2018). Researchers' attitudes to the 3Rs—An upturned hierarchy?. *PloS One*, 13(8), p. e0200895. [online] Available at: https://doi.org/10.1371/journal.pone.0200895 [Accessed 17 August 2018].

Freedman, L.P., I.M. Cockburn, T.S. Simcoe (2015). The Economics of Reproducibility in Preclinical Research. *PLoS Biology,* 13(6), p. e1002165. [online] Available at: https://doi.org/10.1371/journal.pbio.1002165 [Accessed 15 July 2017].

Fureix, C., M. Walker, L. Harper, K. Reynolds, A. Saldivia-Woo and G. Mason (2016). Stereotypic Behaviour in Standard Non-enriched Cages Is an Alternative To Depression-like Responses in C57BL/6 Mice. *Behavioural Brain Research*, 305, pp. 186–190.

Garner, J.P. (2005). Stereotypies and Other Abnormal Repetitive Behaviors: Potential Impact on Validity, Reliability, and Replicability of Scientific Outcomes. *ILAR Journal*, 46(2), pp. 106–117.

Garner, J.P. and G.J. Mason (2002). Evidence for a Relationship Between Cage Stereotypies and Behavioural Disinhibition in Laboratory Rodents. *Behavioural Brain Research*, 136(1), pp. 83–92.

Gärtner, K. (1999). Cage Enrichment Occasionally Increases Deviation of Quantitative Traits. In: *International Joint Meeting 12th ICLAS General Assembly and Conference and 7th FELASA Symposium.* Madrid: SECAL, pp. 207–210.

Gaskill, B. and J. Garner (2014). Letter-To-The-Editor on "Not so hot: Optimal Housing Temperatures for Mice To Mimic the Thermal Environment of Humans". *Molecular Metabolism*, 3(4), pp. 335–336. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4060287/ [Accessed 12 August 2017].

Gaskill, B., C. Gordon, E. Pajor, J. Lucas, J. Davis and J. Garner (2012). Heat or Insulation: Behavioral Titration of Mouse Preference for Warmth or Access To a Nest. *PLoS One*, 7(3), p.e32799. [online] Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0032799 [Accessed 12 August 2017].

Gaskill, B. and K. Pritchett-Corning (2016). Nest Building as an Indicator of Illness in Laboratory Mice. *Applied Animal Behaviour Science*, 180, pp. 140–146.

Gaskill, B., S. Rohr, E. Pajor, J. Lucas and J. Garner (2009). Some Like It Hot: Mouse Temperature Preferences in Laboratory Housing. *Applied Animal Behaviour Science*, 116(2–4), pp. 279–285.

Gaskill, B., C. Winnicker, J. Garner and K. Pritchett-Corning (2013). The Naked Truth: Breeding Performance in Nude Mice with and without Nesting Material. *Applied Animal Behaviour Science*, 143(2–4), pp. 110–116.

GESIS (n.d.). SMAFIRA Computer-aided modeling of signatures and features of relevant scientific content. Available at: https://www.gesis.org/en/research/external-funding-projects/archive/smafira/ [Accessed 17 August 2017].

Giri, S. and A. Bader (2015). A Low-Cost, High-quality New Drug Discovery Process Using Patient-derived Induced Pluripotent Stem Cells. *Drug Discovery Today*, 20(1), pp. 37–49.

Glasziou, P., D.G. Altman, P. Bossuyt, I. Boutron, M. Clarke, S. Julious, S. Michie, D. Moher and E. Wager (2014). Reducing Waste from Incomplete or Unusable Reports of Biomedical Research. *The Lancet*, 383(9913), pp. 267–276.

Gluck, J.P. (2016). Voracious Science & Vulnerable Animals. A Primate Scientist's Ethical Journey. The University of Chicago Press: Chicago and London.

Gordon, C.J. (1993). *Temperature regulation in laboratory rodents*. New York: Cambridge University Press.

Gordon, C.J. (2012). Thermal Physiology of Laboratory Mice: Defining Thermoneutrality. *Journal of Thermal Biology*, 37(8), pp. 654–685.

Gouveia, K. and J.L. Hurst (2013). Reducing Mouse Anxiety During Handling: Effect of Experience with Handling Tunnels. *PLoS One*, 8(6), p. e66401.

Gouveia, K. and J.L. Hurst (2017). Optimising reliability of mouse performance in behavioural testing: the major role of non-aversive handling. *Scientific Reports*, 7, p. 44999, [online] Available at: https://www.nature.com/articles/srep44999 [Accessed 12 August 2018].

Greek, R. and L.A. Kramer (2019). The scientific problems with using non-human animals to predict human response to drugs and disease. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill.

Gross, A., S. Richter, A. Engel and H. Würbel (2012). Cage-induced Stereotypies, Perseveration, and the Effects of Environmental Enrichment in Laboratory Mice. *Behavioural Brain Research*, 234(1), pp. 61–68.

Gruen, L. (2014). *The Ethics of Captivity*. New York: Oxford University Press.

Gurevitch, J., J. Koricheva, S. Nakagawa and G. Stewart (2018). Meta-analysis and the Science of Research Synthesis. *Nature*, 555(7695), pp. 175–182.

Gurfein, B.T., O. Davidenko, M. Premenko-Lanier, J.M. Milush, M. Acree, M.F. Dallman, C. Touma, R. Palme, V.A. York, G. Fromentin, N. Darcel, D.F. Nixon and F.M. Hecht (2014).Environmental Enrichment Alters Splenic Immune Cell Composition and enhances Secondary Influenza Vaccine Responses in Mice. *Molecular Medicine,* 20(1), pp. 179–190.

Hair, K., Macleod, M. R., Sena, E. S. and IICARus Collaboration (2018). A randomised controlled trial of an Intervention to Improve Compliance with the ARRIVE guidelines (IICARus). bioRxiv, p. 370874.

Hartshorne, J. and A. Schachner (2012). Tracking replicability as a method of post-publication open evaluation. *Frontiers in Computational Neuroscience*, 6, p. 8.

[online] Available at: https://www.frontiersin.org/articles/10.3389/fncom.2012.00008/full [Accessed 12 August 2018].

Hartung, T. (2009). Toxicology for the twenty-first century. *Nature*, 460(7252), pp. 208–212.

Hartung, T. (2013). Food for Thought Look Back in Anger—What Clinical Studies Tell Us About Preclinical Work. *Alternatives to Animal Experimentation*, 30(3), pp. 275–291. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3790571/ [Accessed 12 August 2018].

Harris, R. (2017). *Rigor mortis. How sloppy science creates worthless cures, crushes hope, and wastes billions*. New York: Basic.

Hau, J. (2008). Animal models for human disease: An overview. In: P.M. Conn, ed., *Sourcebook of Models for Biomedical Research*. Totowa: Humana Press Inc., pp. 3–9.

Hawkins, P., R. Armstrong, T. Boden, P. Garside, K. Knight, E. Lilley, M. Seed, M. Wilkinson and R. Williams (2015). Applying Refinement to the Use of Mice and Rats in Rheumatoid Arthritis Research. *Inflammopharmacology*, 23(4), pp. 131–150.

Hebb, D.O. (1947). The Effects of Early Experience on Problem Solving at Maturity. *American Psychologist*, 2, pp. 306–307.

Hendriksen, C., D. Morton and K. Cussler (2011). Use of humane end points to minimise suffering. In: B. Howard, T. Nevalainen and G. Perretta, eds., *The Cost Manual of Laboratory Animal Care and Use*. Florida: CRC Press, pp. 333–353.

Hennessy, M.B., McCowan, B., Jiang, J. and Capitanio, J.P. (2014). Depressive-like behavioral response of adult male rhesus monkeys during routine animal husbandry procedure. *Frontiers in Behavioral Neuroscience*, 8, p. 309. [online] Available at: https://www.frontiersin.org/articles/10.3389/fnbeh.2014.00309/full [Accessed 17 August 2018].

Herrmann, K. (2013). Directive 2010/63/EU—A Chance for More Humane Education?. In: *12th FELASA 2013 SECAL Congress "Animal Research: Better Science with Fewer Animals"*. Barcelona: SECAL, Abstracts of scientific papers 12th FELASA SECAL Congress, Journal of the American Association for Laboratory Animal Science, 52(3), p. 389. [online] Available at: https://www.researchgate.net/publication/282816209_jaalas_FELASA_2013 [Accessed 12 August 2018].

Herrmann, K. and P.A. Flecknell (2018a). Retrospective review of anesthetic and analgesic regimens used in animal research proposals. *Alternatives to Animal Experimentation,* 36(1), pp. 65–80. [online] Available at: https://www.altex.org/index.php/altex/article/view/780/1234 [Accessed 14 September 2018].

Herrmann, K. and Flecknell, P.A. (2018b): Application of humane endpoints and humane killing methods in animal research proposals – a retrospective review. *Alternatives to Laboratory Animals*, 46(6), pp. 1–17.

Herrmann, K. and Flecknell, P.A. (2018c): Severity classification of surgical procedures and application of health monitoring strategies in animal research proposals – a retrospective review. *Alternatives to Laboratory Animals*, 46(5), pp. 273–289.

Herrmann, K., K. Köpernik and M. Biedermann (2009). Ein Leitfaden für die Teilprüfung der "Unerlässlichkeit" im Hinblick auf "Refinement". In: D. Borchers and J. Luy, eds., *Der Ethisch Vertretbare Tierversuch, Kriterien und Grenzen*, Padaborn: mentis, pp. 219–234.

Herrmann, K. and H. Ratsch (2010): Bessere Haltungsbedingungen für Labortiere: Überwachung von Versuchstierhaltungen – die Frage des Enrichments. *Deutsches Tierärzteblatt*, 4, pp. 492–499.

Herzog, H. (2002). Ethical Aspects of Relationships Between Humans and Research Animals. *ILAR Journal*, 43(1), pp. 27–32.

Hess, S.E., S. Rohr, B.D. Dufour, B.N. Gaskill, E.A. Pajor and J.P. Garner (2008). Home Improvement: C57BL/6J Mice Given More Naturalistic Nesting Materials Build Better Nests. *Journal of the American Association for Laboratory Animal Science*, 47(6), pp. 25–31.

Higgins, J.P., D.G. Altman, P.C. Gøtzsche, P. Jüni, D. Moher, A.D. Oxman, J. Savović, K.F. Schulz, L. Weeks and J.A. Sterne (2011). The Cochrane Collaboration's Tool for Assessing Risk of Bias in Randomised Trials. *British Medical Journal*, 343, p. d5928.

Hooijmans, C.R., J. IntHout, M. Ritskes-Hoitinga and M.M. Rovers (2014a). Meta-analyses of Animal Studies: An Introduction of a Valuable Instrument To Further Improve Healthcare. *ILAR journal*, 55(3), pp. 418–426. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4276598/ [Accessed 15 July 2017].

Hooijmans, C.R., M. Leenaars and M. Ritskes-Hoitinga (2010). A Gold Standard Publication Checklist To Improve The Quality Of Animal Studies, To Fully Integrate The Three Rs, And To Make Systematic Reviews More Feasible. *Alternatives to Laboratory Animals*, 38(2), pp. 167–182. [online] Available at: http://repository.ubn.ru.nl/bitstream/handle/2066/89153/89153.pdf [Accessed 15 August 2017].

Hooijmans, C.R. and M. Ritskes-Hoitinga (2013). Progress in Using Systematic Reviews of Animal Studies To Improve Translational Research. *PLoS Medicine*, *10*(7), p. e1001482. [online] Available at: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001482 [Accessed on 14 November 2017].

Hooijmans, C.R., M.M. Rovers, R.B. de Vries, M. Leenaars, M. Ritskes-Hoitinga and M.W. Langendam (2014b). SYRCLE's Risk of Bias Tool for Animal Studies. *BMC Medical Research Methodology*, 14(1), p. 43. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4230647/ [Accessed on 14 November 2017].

Howerton, C.L., J.P. Garner and J.A. Mench (2008). Effects of a Running Wheel-igloo Enrichment on Aggression, Hierarchy Linearity, and Stereotypy in Group-housed Male CD-1 (ICR) Mice. *Applied Animal Behaviour Science*, 115(1), pp. 90–103.

Hunter, A.J. (2011). Have Animal Models of Disease Helped or Hindered the Drug Discovery Process. *Annals of the New York Academy of Sciences*, 1245, pp. 1–2.

Hurst, J.L. and R.S. West (2010). Taming Anxiety in Laboratory Mice. *Nature Methods*, 7, pp. 825–826. [online] Available at: https://www.nature.com/nmeth/journal/v7/n10/full/nmeth.1500.html [Accessed on 14 January 2017].

Hutchinson, L. and R. Kirk (2011). High Drug Attrition Rates—Where Are We Going Wrong?. *Nature Reviews Clinical Oncology*, 8(4), pp. 189–190.

Ioannidis, J.P.A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8), p. e124. [online] Available at: https://doi.org/10.1371/journal.pmed.0020124 [Accessed 15 July 2017].

Ioannidis, J.P.A. (2012). Why Science Is Not Necessarily Self-Correcting. *Perspectives on Psychological Science*, 7(6), pp. 645–654. [online] Available at: http://journals.sagepub.com/doi/full/10.1177/1745691612464056 [Accessed 15 July 2017].

Ioannidis, J.P.A., S. Greenland, M.A. Hlatky, M.J. Khoury, M.R. Macleod, D. Moher, K.F. Schulz and R. Tibshirani (2014). Increasing Value and Reducing Waste in Research Design, Conduct, and Analysis. *The Lancet*, 383(9912), pp. 166–175.

Jirkof, P. (2015). Effects of Experimental Housing Conditions on Recovery of Laboratory Mice. *Lab Animal*, 44(2), pp. 65–70.

Joffe, A.R., M. Bara, N. Anton and N. Nobis (2016). Expectations for the Methodology and Translation Of Animal Research: A Survey of the General Public, Medical Students, and Animal Researchers in North America. *Alternatives to Laboratory Animals*, 44(4), pp. 361–381.

Johnson, E.A., D.S. Sharp and D.B. Miller (2000). Restraint as a Stressor in Mice: Against the Dopaminergic Neurotoxicity of d-MDMA, Low Body Weight Mitigates Restraint-induced Hypothermia and Consequent Neuroprotection. *Brain Research*, 875(1), pp. 107–118.

Johnson, J. and A. Smajdor (2019). Human wrongs in animal research—A focus on moral injury and reification. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill.

Jones, J.M. (2017). Americans Hold Record Liberal Views on Most Moral Issues. *Gallup Poll Social Series*. [online] Available at: http://www.gallup.com/poll/210542/americans-hold-record-liberal-views-moral-issues.aspx?g_source=2017+poll+animals&g_medium=search&g_campaign=tiles [Accessed 10 June 2017].

Justice, M.J. and P. Dhillon (2016). Using the Mouse to Model Human Disease: Increasing Validity and Reproducibility. *Disease Model Mechanisms*, 9(2), pp. 101–103.

Kaplan, R.M. and V.L. Irvin (2015). Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased Over Time. *PLoS One*, 10(8), p. e0132382. [online] Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0132382 [Accessed 10 June 2017].

Karp, C.L. (2012). Unstressing Intemperate Models: How Cold Stress Undermines Mouse Modeling. *Journal of Experimental Medicine*, 209(6), pp. 1069–1074.

Keen, J. (2019). Wasted money in US biomedical and agricultural animal research. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill.

Kempermann, G., H.G. Kuhn and F.H. Gage (1997). More Hippocampal Neurons in Adult Mice Living in an Enriched Environment. *Nature*, 386(6624), p. 493.

Kerr, N.L. (1998). HARKing: Hypothesizing After the Results Are Known. *Personality and Social Psychology Review*, 2(3), pp. 196–217.

Kilkenny, C., W.J. Browne, I.C. Cuthill, M. Emerson and D.G. Altman (2010). Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. *PLoS Biology*, 8(6), p. e1000412. [online] Available at: http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1000412 [Accessed 14 May 2017].

Klein, H. and K. Bayne (2007). Establishing a Culture of Care, Conscience, and Responsibility: Addressing the Improvement of Scientific Discovery and Animal Welfare Through Science-based Performance Standards. *ILAR Journal*, 48(1), pp. 3–11.

Knight, A. (2008). 127 Million Non-Human Vertebrates Used Worldwide for Scientific Purposes in 2005. *Alternatives to Laboratory Animals*, 36(5), pp. 494–496.

Knight, A. (2019). Critically evaluating animal research. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill pp. 321–340.

Kola, I. and J. Landis (2004). Can the Pharmaceutical Industry Reduce Attrition Rates? *Nature Reviews Drug Discovery,* 3(8), p. 711–715.

Kramer, L.A. and R. Greek (2018). Human Stakeholders and the Use of Animals in Drug Development. *Business and Society Review,* 123(1), pp. 3–58.

Krech, D., M.R. Rosenzweig and E.L. Bennett (1960). Effects of Environmental Complexity and Training on Brain Chemistry. *Journal of Comparative and Physiological Psychology,* 53, pp. 509–514.

Lahvis, G.P. (2017). Unbridle Biomedical Research From the Laboratory Cage. *Elife*, 6. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5503508/ [Accessed 4 July 2017].

Lancaster, M.A. and J.A. Knoblich (2014). Generation of Cerebral Organoids From Human Pluripotent Stem Cells. *Nature Protocols*, 9(10), pp. 2329–2340.

Landis, S.C., S.G. Amara, K. Asadullah, C.P. Austin, R. Blumenstein, E.W. Bradley, R.G. Crystal, R.B. Darnell, R.J. Ferrante, H. Fillit and R. Finkelstein (2012). A Call for Transparent Reporting To Optimize The Predictive Value of Preclinical Research. *Nature*, 490(7419), pp. 187–191.

Landrigan, P.J., R. Fuller, N.J. Acosta, O. Adeyi, R. Arnold, A.B. Baldé, R. Bertollini, S. Bose-O, J.I. Boufford, P.N. Breysse and T. Chiles (2017). The Lancet Commission on Pollution and Health. *The Lancet*, pp. 462–512.

Langley, G., I. Adcock, F. Busquet, K. Crofton, E. Csernok, C. Giese, T. Heinonen, K. Herrmann, M. Hofmann-Apitius, B. Landesmann, L. Marshall, E. McIvor, A. Muotri, F. Noor, K. Schutte, T. Seidle, A. van de Stolpe, H. Van Esch, C. Willett and G. Woszczek (2017). Towards a 21st Century Roadmap for Biomedical Research and Drug Discovery: Consensus Report and Recommendations. *Drug Discovery Today*, 22(2), pp. 327–339.

Langley, G., C.P. Austin, A.K. Balapure, L.S. Birnbaum, J.R. Bucher, J. Fentem, S.C. Fitzpatrick, J.R. Fowle III, R.J. Kavlock, H. Kitano and B.A. Lidbury (2015). Lessons from

Toxicology: Developing a 21st Century Paradigm for Medical Research. *Environmental Health Perspectives*, 123(11), p. A268–A272.

Leary, S., W. Underwood, R. Anthony, S. Cartner, D. Corey, T. Grandin, C. Greenacre, S. Gwaltney-Brant, M.A. McCrakin, R. Meyer and D. Miller (2013). *AVMA Guidelines for the Euthanasia of Animals,* 2013 ed. [online] Available at: https://www.avma.org/KB/Policies/Documents/euthanasia.pdf [Accessed 4 July 2017].

Lees, J.S., E.S. Sena, K.J. Egan, A. Antonic, S.A. Koblar, D.W. Howells and M.R. Macleod (2012). Stem Cell-based Therapy for Experimental Stroke: A Systematic Review and Meta-analysis. *International Journal of Stroke*, 7(7), pp. 582–588.

Leist, M. and T. Hartung (2013). Inflammatory Findings On Species Extrapolations: Humans Are Definitely No 70 kg Mice. *Archives of Toxicology*, 87(4), pp. 563–567.

Lewis, M.H., M.F. Presti, J.B. Lewis and L.A. Turner (2006). The neurobiology of stereotypy I. Environmental complexity. In: G.J. Mason and J. Rushen, eds., *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare,* 2nd ed. Oxfordshire: CAB International, pp. 190–226.

Li, G., Y. Gan, Y. Fan, Y. Wu, H. Lin, Y. Song, X. Cai, X. Yu, W. Pan, M. Yao, J. Gu and H. Tu (2015). Enriched Environment Inhibits Mouse Pancreatic Cancer Growth and Down-regulates the Expression of Mitochondria-related Genes in Cancer Cells. *Scientific Reports*, 5, p. 7856. [online] Avaiable at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4297951/ [Accessed 1 July 2017].

Lilley, E., R. Armstrong, N. Clark, P. Gray, P. Hawkins, K. Mason, N. López-Salesansky, A. Stark, S. Jackson, C. Thiemermann and M. Nandi (2015). Refinement of Animal Models of Sepsis and Septic Shock. *Shock*, 43(4), pp. 304–316.

Lidster, K., J.G. Jefferys, I. Blümcke, V. Crunelli, P.A. Flecknell, B.G. Frenguelli, W.P. Gray, R. Kaminski, A. Pitkänen, I. Ragan and M. Shah (2016). Opportunities for improving animal welfare in rodent models of epilepsy and seizures. *Journal of Neuroscience Methods*, 260, pp. 2–25.

Low, P. (2012). *Cambridge Declaration on Consciousness.* [online] Avaiable at: http://fcmconference.org/img/CambridgeDeclarationOnConsciousness.pdf [Accessed 1 July 2017].

Lush Prize (n.d.). *About the Lush Prize.* [online] Available at: http://lushprize.org/awards/ [Accessed 11 July 2018].

Macleod, M.R. (2011). Why Animal Research Needs To Improve: Many of The Studies That Use Animals To Model Human Diseases Are Too Small and Too Prone To Bias To Be Trusted. *Nature*, 477(7366), pp. 511–512.

Macleod, M.R., T. O'Collins, D.W. Howells and G.A. Donnan (2004). Pooling of Animal Experimental Data Reveals Influence Of Study Design and Publication Bias. *Stroke*, 35(5), pp. 1203–1208.

Macleod, M.R., A. Lawson McLean, A. Kyriakopoulou, S. Serghiou, A. de Wilde, N. Sherratt, T. Hirst, R. Hemblade, Z. Bahor, C. Nunes-Fonseca, A. Potluru, A. Thomson,

J. Baginskitae, K. Egan, H. Vesterinen, G. Currie, L. Churilov, D. Howells and E. Sena (2015). Risk of Bias in Reports of In Vivo Research: A Focus for Improvement. *PLoS Biology*, 13(10), p. e1002273. [online] Available at: https://core.ac.uk/download/pdf/45473665.pdf [Accessed 11 July 2017].

Macleod, M.R., H.B. van der Worp, E.S. Sena, D.W. Howells, U. Dirnagl and G.A. Donnan (2008). Evidence for the Efficacy of NXY-059 in Experimental Focal Cerebral Ischaemia Is Confounded by Study Quality. *Stroke*, 39(10), pp. 2824–2829.

Mahoney, C.J. (1992). Some Thoughts on Psychological Enrichment. *Lab Animal,* 21(5), pp. 27–37.

Mak, I.W., N. Evaniew and M. Ghert (2014). Lost in Translation: Animal Models and Clinical Trials in Cancer Treatment. *American Journal of Translational Research*, 6(2), pp. 114–118.

Makowska, I.J. and D.M. Weary (2016a). Differences in Anticipatory Behaviour Between Rats (Rattus Norvegicus) Housed In Standard Versus Semi-naturalistic Laboratory Environments. *PLoS One*, 11(1), p. e0147595. [online] Available at: https://doi.org/10.1371/journal.pone.0147595 [Accessed 8 July 2017].

Makowska, I.J. and D.M. Weary (2016b). The Importance of Burrowing, Climbing, and Standing Upright for Laboratory Rats. *Royal Society Open Science*, 3(6), p. 160136. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4929907/ [Accessed 8 July 2017].

Marks, J. (2012). Accept No Substitutes: The Ethics of Alternatives. *Hastings Center Report*, 42(s1).

Martin, B., S. Ji, S. Maudsley and M.P. Mattson (2010). "Control" Laboratory Rodents Are Metabolically Morbid: Why It Matters. *Proceedings of the National Academy of Sciences*, 107(14), pp. 6127–6133.

Martire, V.L., A. Silvani, S. Bastianini, C. Berteotti and G. Zoccoli (2012). Effects of Ambient Temperature on Sleep and Cardiovascular Regulation in Mice: The Role of Hypocretin/Orexin Neurons. *PLoS One*, 7(10), p. e47032. [online] Available at: https://doi.org/10.1371/journal.pone.0047032 [Accessed on 23 February 2017].

Mason, G.J. (2006). Stereotypic Behaviour in Captive Animals: Fundamentals and Implications for Welfare and Beyond. In: G. Mason and J. Rushen, eds., *Stereotypic Animal Behaviour: Fundamentals and Applications to Welfare*, 2nd ed. Oxfordshire: CAB International, pp. 325–356.

Maurin, J. (2012). Blutige Bisswunden, hungernde Mäuse. *TAZ*. [online] Available at: http://www.taz.de/!5095484/ [Accessed on 23 February 2017].

McManus, R. (2013). Ex-Director Zerhouni Surveys Value of NIH Research. *NIH Record*, LXV(13). [online] Available at: https://nihrecord.nih.gov/newsletters/2013/06_21_2013/story1.htm [Accessed on 23 February 2017].

McNutt, M., (2014). Journals Unite for Reproducibility. *Science*, 346(6210), p. 679.

Meagher, R.K. and G.J. Mason (2012). Environmental Enrichment Reduces Signs of Boredom in Caged Mink. *PLoS One*, 7(11), p. e49180. [online] Available at: https://doi.org/10.1371/journal.pone.0049180 [Accessed on 23 February 2017].

Meijer, M.K., B.M. Spruijt, L.F.M. Van Zutphen and V. Baumans (2006). Effect of Restraint and Injection Methods on Heart Rate and Body Temperature in Mice. *Laboratory animals*, 40(4), pp. 382–391.

Mergenthaler, P. and A. Meisel (2012). Do Stroke Models Model Stroke?. *Disease Models & Mechanisms*, 5(6), pp. 718–725.

Messmer, M.N., K.M. Kokolus, J.W.L. Eng, S.I. Abrams and E.A. Repasky (2014). Mild Cold-stress Depresses Immune Responses: Implications for Cancer Models Involving Laboratory Mice. *Bioessays*, 36(9), pp. 884–891.

Mestas, J. and C. Hughes (2004). Of Mice and Not Men: Differences Between Mouse and Human Immunology. *The Journal of Immunology*, 172(5), pp. 2731–2738.

Moher, D., A. Liberati, J. Tetzlaff, D.G. Altman, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), p. e1000097. [online] Available at: https://doi.org/10.1371/journal.pmed.1000097 [Accessed 12 January 2017].

Morton, D. (2000). A Systematic Approach for Establishing Humane Endpoints. *ILAR Journal*, 41(2), pp. 80–86.

Mullane, K. and M. Williams (2014). Animal Models of Asthma: Reprise or Reboot?. *Biochemical Pharmacology*, 87(1), pp. 131–139.

National Animal Ethics Advisory Committee (2002). *A Culture of Care: A Guide for People Working With Animals in Research, Testing and Teaching.* Wellington, New Zealand. [online] Available at: https://anzccart.org.nz/app/uploads/2017/03/culture-of-care.pdf [Accessed 10 June 2017].

National Research Council (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy.* Washington, D.C.: National Academies Press.

National Toxicology Program (2004). *A National Toxicology Program to the 21st Century. A Roadmap for the Future.* [online] Available at: https://ntp.niehs.nih.gov/ntp/about_ntp/ntpvision/ntproadmap_508.pdf [Accessed 11 January 2017].

NCad (n.d.). *NCad opinion Transition to non-animal research.* [online] Available at: https://www.ncadierproevenbeleid.nl/documenten/rapport/2016/12/15/ncad-opinion-transition-to-non-animal-research [Accessed 11 July 2017].

NCATS (National Center for Advancing Translational Sciences) (2017). *Tissue Chips for Disease Modeling and Efficacy Testing.* [online] Available at: https://ncats.nih.gov/tissuechip/projects/modeling [Accessed 14 July 2017].

NAVS (2018). *Fund Smarter Science.* [online] Available at: https://www.navs.org/what-we-do/fund-smarter-science/#.W1UAPi3Myu5 [Accessed 10 July 2018].

NC3Rs (n.d. a). *Funding Scheme Priority Areas.* [online] Available at: https://www.nc3rs.org.uk/funding-scheme-priority-areas#historic [Accessed 16 July 2018].

NC3Rs (n.d. b). *Experimental Design*. [online] Available at: https://nc3rs.org.uk/exper imental-design [Accessed 14 July 2017].

NC3Rs (n.d. c). *Research Hubs*. [online] Available at*:* https://nc3rs.org.uk/resource -hubs [Accessed 14 July 2017].

NC3Rs (n.d. d). *Our funding*. [online] Available at*:* https://www.nc3rs.org.uk/about-us/ funders [Accessed 16 July 2018].

Nerem, R.M., M.J. Levensque and J.F. Cornhill (1980). Social Environment as a Factor of Diet Induced Atherosclerosis. *Science,* 208, pp. 1475–1476.

NIH (2019). *Budget.* [online] Available at: https://www.nih.gov/about-nih/what-we -do/budget [Accessed 14 February 2019].

Nithianantharajah, J. and A. Hannan (2006). Enriched Environments, Experience- dependent Plasticity and Disorders of the Nervous System. *Nature Reviews Neuro- science*, 7(9), pp. 697–709.

NHMRC (National Health and Medical Research Council) (2013). *Australian code for the care and use of animals for scientific purposes.* 8th ed. [online] Available at: https://www.nhmrc.gov.au/guidelines-publications/ea28 [Accessed 11 July 2017].

Noor, F. (2019). The changing paradigm in preclinical toxicology: *in vitro* and *in silico* methods in liver toxicity evaluations. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill.

Norecopa (2016a). *Culture of care.* [online] Available at: https://norecopa.no/more- resources/culture-of-care [Accessed 11 August 2018].

Norecopa (2016b). *Design and reporting of animal experiments*. [online] Available at: https://norecopa.no/more-resources/experimental-design-and-reporting [Ac- cessed 12 August 2018].

Norecopa (2017). *Systematic Review Centre for Laboratory Animal Experimentation* (*SYRCLE*). [online] Available at: https://norecopa.no/3r-guide/systematic-review -centre-for-laboratory-animal-experimentation-syrcle [Accessed 12 August 2018].

O'Collins, V.E., M.R. Macleod, G.A. Donnan, L.L. Horky, B.H. van der Worp and D.W. Howells (2006). 1,026 Experimental Treatments in Acute Stroke. *Annals of Neu- rology*, 59(3), pp. 467–477.

Olson, H., G. Betton, D. Robinson, K. Thomas, A. Monro, G. Kolaja, P. Lilly, J. Sanders, G. Sipes, W. Bracken, M. Dorato, K. Van Deun, P. Smith, B. Berger and A. Heller (2000). Concordance of the Toxicity of Pharmaceuticals in Humans and in Animals. *Regulatory Toxicology and Pharmacology*, 32(1), pp. 56–67. [online] Available at: https://www.gwern.net/docs/statistics/meta-analysis/2000-olson.pdf [Accessed 12 August 2017].

Olsson, I.A.S., C.M. Nevison, E.G. Patterson-Kane, C.M. Sherwin, H.A. van de Weerd and H. Würbel (2003). Understanding Behaviour: The Relevance of Ethological Approaches in Laboratory Animal Science. *Applied Animal Behaviour Science*, 81(3), pp. 245–264.

Ormandy, E.H., C.A. Schuppli, and D.M. Weary (2009). Worldwide Trends in the Use of Animals in Research: The Contribution of Genetically-modified Animal Models. *Alternatives to Laboratory Animals*, 37, pp. 63–68.

Osborne, N., Avey, M.T., Anestidou, L., Ritskes-Hoitinga, M. and Griffin, G. (2018). Improving animal research reporting standards: HARRP, the first step of a unified approach by ICLAS to improve animal research reporting standards worldwide. EMBO Reports , 19(5), p. e46069. [online] Available at: http://embor.embopress.org/content/19/5/e46069 [Accessed 19 April 2018].

Pawlowski, J., D. Feinstein, M.L. Crandall and S. Gala (2019). Modernizing biomedical training: replacing live animal laboratories with human simulation. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill, pp. 551–66.

Payne, K.J. and G.M. Crooks (2007). Immune-cell Lineage Commitment: Translation From Mice to Humans. *Immunity*, 26(6), pp. 674–677.

Percie du Sert, N., A. Alfieri, S.M. Allan, H.V. Carswell, G.A. Deuchar, T.D. Farr, P. Flecknell, L. Gallagher, C.L. Gibson, M.J. Haley and M.R. Macleod (2017). The IMPROVE Guidelines (Ischaemia Models: Procedural Refinements Of in Vivo Experiments). *Journal of Cerebral Blood Flow & Metabolism*, p. 0271678X17709185. [online] Available at: http://journals.sagepub.com/doi/pdf/10.1177/0271678X17709185 [Accessed 18 August 2017].

Perel, P., I. Roberts, E. Sena, P. Wheble, C. Briscoe, P. Sandercock, M. Macleod, L.E. Mignini, P. Jayaram and K.S. Khan (2007). Comparison of Treatment Effects Between Animal Experiments and Clinical Trials: Systematic Review. *British Medical Journal*, 334, pp. 197–206. [online] Available at: https://www.bmj.com/content/334/7586/197 [Accessed 12 July 2017].

PETA International Science Consortium (PISC) (n.d. a). *Awards of VITROCELL In Vitro Inhalation Exposure Systems.* [online] Available at: https://www.piscltd.org.uk/vitrocell-prize/ [Accessed 10 August 2017].

PETA International Science Consortium (PISC) (n.d. b). *Funding to Advance the Development and Use of Non-Animal Methods.* [online] Available at: https://www.piscltd.org.uk/funding/ [Accessed 10 August 2018].

Pew Research Center (2015). *Public and Scientists' Views on Science and Society: Use of Animals in Scientific Research.* [online] Available at: http://www.pewinternet.org/2015/01/29/public-and-scientists-views-on-science-and-society/pi_2015-01-29_science-and-society-03-05/ [Accessed 10 July 2017].

Pew Research Center (2018). Americans are divided over the use of animals in scientific research. [online] Available at: http://www.pewresearch.org/fact-tank/2018/08/16/americans-are-divided-over-the-use-of-animals-in-scientific-research/. [Accessed 10 July 2017].

Pharmaceutical Research and Manufacturers of America (2015). *Biopharmaceutical research industry profile.* Washington, DC: PhRMA. [online] Available at: http://

phrma-docs.phrma.org/sites/default/files/pdf/2015_phrma_profile.pdf [Accessed 10 July 2017].

Pharmaceutical Research and Manufacturers of America (2016). Biopharmaceutical research industry profile. Washington, DC: PhRMA. [online] Available at: http:// phrma-docs.phrma.org/sites/default/files/pdf/biopharmaceutical-industry-profile .pdf [Accessed 10 December 2017].

Pippin, J.J. (2012). Animal Research in Medical Sciences: Seeking a Convergence of Science, Medicine, and Animal Law. *South Texas Law Review,* 54, pp. 469–511.

Pippin, J.J., S.E. Cavanaugh and F. Pistollato (2019). Animal Research for Alzheimer Disease: Failures of Science and Ethics. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill.

Pistollato, F., E.L. Ohayon, A. Lam, G.R. Langley, T.J. Novak, D. Pamies, G. Perry, E. Trushina, R.S.B. Williams, A.E. Roher, T. Hartung, S. Harnard, N. Barnard, M.C. Morris, M. Lai, R. Merkley and P.C. Chandrasekera (2016). Alzheimer Disease research in the 21st century: past and current failures, new perspectives and funding priorities. *Oncotarget*, 7(26), pp. 38999–39016. [online] Available at: https://www.ncbi.nlm .nih.gov/pmc/articles/PMC5129909/ [Accessed 10 July 2017].

Poole, T. (1997). Happy Animals Make Good Science. *Laboratory Animals*, 31(2), pp. 116–124.

Pound, P. and M.B. Bracken (2014). Is Animal Research Sufficiently Evidence Based To Be a Cornerstone of Biomedical Research?. *British Medical Journal,* 348, p. g3387.

Pound, P., S. Ebrahim, P. Sandercock, M.B. Bracken and I. Roberts (2004). Where Is the Evidence That Animal Research Benefits Humans?. *British Medical Journal*, 328(7438), pp. 514–517.

Pound, P. and C.J. Nicol (2018). Retrospective Harm-benefit Analysis of Preclinical Animal Research for Six Treatment Interventions. *PLoS One,* 13(3), p. e0193758. [online] Available at: http://journals.plos.org/plosone/article?id=10.1371/journal .pone.0193758 [Accessed 28 March 2018].

Pound, P. and Ritskes-Hoitinga, M. (2018). Is it possible to overcome issues of external validity in preclinical animal research? Why most animal models are bound to fail. *Journal of Translational Medicine*, 16(1), p. 304. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6223056/ [Accessed 11 November 2018].

Preclinical Trials (n.d.). *International Register for Preclincial Protocols.* [online] Available at: https://www.preclinicaltrials.eu [Accessed 10 December 2017].

Prinz, F., T. Schlange and K. Asadullah (2011). Believe It or Not: How Much Can We Rely on Published Data on Potential Drug Targets?. *Nature Reviews Drug Discovery*, 10(9), pp. 712–712.

Ram, R. (2019). Extrapolation of animal research data to humans: an analysis of the evidence. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill.

Ramirez, T., S. Beken, M. Chlebus, G. Ellis, C. Griesinger, S. De Jonghe, I. Manou, A. Mehling, K. Reisinger, L. Rossi, J. van der Laan, R. Weissenhorn and U. Sauer (2015). Knowledge sharing to facilitate regulatory decision-making in regard to alternatives to animal testing: Report of an EPAA workshop. *Regulatory Toxicology and Pharmacology*, 73(1), pp. 210–226.

Redmond, C. (2019). When is an alternative not an alternative? Supporting progress for absolute replacement of animals in science. In: K. Herrmann and K. Jayne, eds., *Animal Experimentation: Working Towards a Paradigm Change*, Vol. 22. Leiden: Brill.

Renner, M., M.A. Lancaster, S. Bian, H. Choi, T. Ku, A. Peer, K. Chung and J.A. Knoblich (2017). Self-organized Developmental Patterning and Differentiation in Cerebral Organoids. *The EMBO Journal*, p. e201694700. [online] Available at: http://emboj .embopress.org/content/36/10/1316.long [Accessed 10 May 2017].

Reichlin, T.S., L. Vogt and H. Würbel (2016). The Researchers' View of Scientific Rigor—Survey on the Conduct and Reporting of In Vivo Research. *PLoS One*, 11(12), p. e0165999. [online] Available at: https://doi.org/10.1371/journal.pone.0165999 [Accessed 10 January 2017].

Reinhardt, V. (2003). Compassion for Animals in the Laboratory: Impairment or Refinement of Research Methodology?. *Journal of Applied Animal Welfare Science*, 6(2), pp. 123–130.

Richardson, C.A. and P.A. Flecknell (2005). Anaesthesia and Postoperative Analgesia Following Experimental Surgery in Laboratory Rodents: Are We Making Progress?. *Alternatives to Laboratory Animals*, 33, pp. 119–127.

Richter, S.H. (2017). Systematic Heterogenization for Better Reproducibility in Animal Experimentation. *Lab Animal*, 46(9), pp. 343–349.

Richter, S.H., J.P. Garner and H. Würbel (2009). Environmental Standardization: Cure or Cause of Poor Reproducibility in Animal Experiments?. *Nature Methods*, 6(4), pp. 257–261.

Richter, S.H., J.P. Garner, C. Auer, J. Kunert and H. Würbel (2010). Systematic Variation Improves Reproducibility of Animal Experiments. *Nature Methods*, 7(3), pp. 167–168.

Ritskes-Hoitinga, M. (2016). Systematic Reviews of Animal Studies Equal the Implementation of the 3Rs. In: *20th European Congress on Alternatives to Animal Testing, 17th Annual Congress of EUSAAT*. Linz. [online] Availabe at: https://norecopa.no/ media/7597/systematic-reviews.pdf [Accessed 15 December 2017].

Ritskes-Hoitinga, M. and K. Wever (2018). Improving the Conduct, Reporting, and Appraisal of Animal Research. Editorial. *British Medical Journal*, 360, p. j4935.

Rittirsch, D., L. Hoesel and P. Ward (2007). The Disconnect Between Animal Models of Sepsis and Human Sepsis. *Journal of Leukocyte Biology*, 81(1), pp. 137–143.

Rosenzweig, M.R. (1966). Environmental Complexity, Cerebral Change, and Behavior. *American Psychologist,* 21, pp. 321–332.

Rosenzweig, M.R. and E.L. Bennett (1969). Effects of Differential Environments on Brain Weights and Enzyme Activities in Gerbils, Rats, and Mice. *Developmental Psychobiology*, 2(2), pp. 87–95.

Rosenzweig, M.R., D. Krech, E.L. Bennett and J.F. Zolman (1962). Variation in Environmental Complexity and Brain Measures. *Journal of Comparative and Physiological Psychology*, 55(6), p. 1092.

Røttingen, J.A., S. Regmi, M. Eide, A.J. Young, R.F. Viergever, C. Årdal, J. Guzman, D. Edwards, S.A. Matlin and R.F. Terry (2013). "Mapping of available health research and development data: what's there, what's missing, and what role is there for a global observatory?" *The Lancet*, 382 (9900), pp. 1286–1307.

Rovida, C., S. Asakura, M. Daneshian, H. Hofman-Huether, M. Leist, L. Meunier, D. Reif, A Rossi, M. Schmutz, J. Valentin, J. Zurlo and T. Hartung (2015). Toxicity Testing in the 21st Century Beyond Environmental Chemicals. *Alternatives to Animal Experimentation*, 32(3), pp. 171–181.

Russell, W.M.S. and R.L. Burch (1959). *The Principles of Humane Experimental Technique*. Potters Bar, Hertfordshire, England: Universities Federation for Animal Welfare. [online] Available at: http://altweb.jhsph.edu/pubs/books/humane_exp/het-toc [Accessed 12 May 2017].

Salzberg, S.L., O. White, J. Peterson and J.A. Eisen (2001). Microbial Genes in the Human Genome: Lateral Transfer or Gene Loss?. *Science*, 292 (5523), pp. 1903–1906.

Sandercock, P. and I. Roberts (2002b). Systematic Reviews of Animal Experiments. *The Lancet*, 360, p. 586.

Scannell, J.W. and Bosley, J. (2016). When Quality Beats Quantity: Decision Theory, Drug Discovery, and the Reproducibility Crisis. *PLoS One*, 11(2), p.e0147215. [online] Available at: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0147215 [Accessed 14 September 2017].

Scherer, R.W., Meerpohl, J.J., Pfeifer, N., Schmucker, C., Schwarzer, G. and von Elm, E. (2018). Full publication of results initially presented in abstracts. Cochrane Database of Systematic Reviews 2018, Issue 11. Art. No.: MR000005. DOI: 10.1002/14651858. MR000005.pub4.

Schmidt, D.A., M.R. Ellersieck, M.R. Cranfield and W.B. Karesh (2006). Cholesterol Values in Free-ranging Gorillas (Gorilla Gorilla Gorilla and Gorilla Beringei) and Bornean Orangutans (Pongo Pygmaeus). *Journal of Zoo and Wildlife Medicine*, 37(3), pp. 292–300.

Schuppli, C., A. Walterhouse, V. Chew, N. Hammound, L. Kolody, B. Tan, J. Makowska, S. Mcnamara, J. Sato-Reinhold, V. Wong and D. Weary (2017). A Better Life for Research Animals by Fostering a Culture of Compassion Amongst Researchers. In: *10th World Congress on Alternatives and Animal Use in the Life Sciences (WC10)*. Seattle: Alternatives to Animal Experimentation Proceedings, 6(1), p. 13. [online] Available at: http://www.altex.ch/resources/WC10_entire_issue1.pdf [Accessed 12 September 2017].

Sena, E.S., H.B. van der Worp, P.M. Bath, D.W. Howells and M.R. Macleod (2010). Publication Bias in Reports of Animal Stroke Studies Leads To Major Overstatement of Efficacy. *PLoS Biology*, 8(3), p. e1000344. [online] Available at: http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1000344 [Accessed 12 August 2017].

Seok, J., H.S. Warren, A.G. Cuenca, M.N. Mindrinos, H.V. Baker, W. Xu, D.R. Richards, G.P. McDonald-Smith, H. Gao, L. Hennessy and C.C. Finnerty (2013). Genomic Responses in Mouse Models Poorly Mimic Human Inflammatory Diseases. *Proceedings of the National Academy of Sciences*, 110(9), pp. 3507–3512.

Serrat, M.A., D. King and C.O. Lovejoy (2008). Temperature Regulates Limb Length in Homeotherms By Directly Modulating Cartilage Growth. *Proceedings of the National Academy of Sciences*, 105 (49), pp. 19348–19353.

Shay, T., V. Jojic, O. Zuk, K. Rothamel, D. Puyraimond-Zemmour, T. Feng, E. Wakamatsu, C. Benoist, D. Koller and A. Regev (2013). Conservation and Divergence in the Transcriptional Programs of the Human and Mouse Immune Systems. *Proceedings of the National Academy of Sciences*, 110(8), pp. 2946–2951.

Sherwin, C.M. (2004). The influences of Standard Laboratory Cages on Rodents and the Validity of Research Data. *Animal Welfare*, 13(1), pp. 9–15.

Sherwin, C.M., E. Haug, N. Terkelsen and M. Vadgama (2004). Studies on the Motivation for Burrowing by Laboratory Mice. *Applied Animal Behaviour Science*, 88(3), pp. 343–358.

Shuaib, A., K. Lees, P. Lyden, J. Grotta, A. Davalos, S. Davis, H. Diener, T. Ashwood, W. Wasiewski and U. Emeribe (2007). NXY-059 for the Treatment of Acute Ischemic Stroke. *New England Journal of Medicine*, 357(6), pp. 562–571.

Simmons, J.P., L.D. Nelson and U. Simonsohn (2011). False-positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, 22 (11), pp. 1359–1366.

Sloff, M., R. de Vries, P. Geutjes, J. IntHout, M. Ritskes-Hoitinga, E. Oosterwijk and W. Feitz (2014). Tissue Engineering in Animal Models for Urinary Diversion: A Systematic Review. *PLoS one*, 9(6), p. e98734. [online] Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0098734 [Accessed 30 July 2017].

Smith, A.J., R.E. Clutton, E. Lilley, K.E.A. Hansen and T. Brattelid (2017). PREPARE: Guidelines for Planning Animal Research and Testing. *Laboratory Animals*, pp. 135–141. [online] Available at: http://journals.sagepub.com/doi/pdf/10.1177/0023677217724823 [Accessed 5 August 2017].

Smith, R. (2014). Medical Research—Still a Scandal. [Blog] *British Medical Journal Blog*. Available at: https://blogs.bmj.com/bmj/2014/01/31/richard-smith-medical-research-still-a-scandal/ [Accessed 30 August 2017].

Sorge, R., L. Martin, K. Isbester, S. Sotocinal, S. Rosen, A. Tuttle, J. Wieskopf, E. Acland, A. Dokova, B. Kadoura, P. Leger, J. Mapplebeck, M. McPhail, A. Delaney, G. Wigerblad, A. Schumann, T. Quinn, J. Frasnelli, C. Svensson, W. Sternberg and J. Mogil (2014). Olfactory Exposure to Males, Including Men, Causes Stress and Related Analgesia in Rodents. *Nature Methods*, 11(6), pp. 629–632.

Špinka, M. and F. Wemelsfelder (2011). Environmental challenge and animal agency. In: M.C. Appleby, J.A. Mench, I.A.S. Olsson and B.O. Hughes, eds., *Animal Welfare,* 2nd ed. Wallingford: CABI International, pp. 27–43.

Stokes, E., P. Flecknell and C. Richardson (2009). Reported Analgesic and Anaesthetic Administration To Rodents Undergoing Experimental Surgical Procedures. *Laboratory Animals*, 43(2), pp. 149–154.

Systematic Review Center for Laboratory Animal Experimentation (SYRCLE) (n.d. a). *Systematic reviews of animal studies*. [video] Available at: https://www.radboudumc .nl/en/research/radboud-technology-centers/animal-research-facility/systematic -review-center-for-laboratory-animal-experimentation [Accessed 12 July 2017].

SYRCLE (n.d. b). *Tools and guidelines*. [online] Available at: https://www.radboudumc .nl/en/research/radboud-technology-centers/animal-research-facility/systematic -review-center-for-laboratory-animal-experimentation [Accessed 12 July 2017].

SYRCLE (n.d. c). Protocols. Turner et al. (2017). The use of carbon dioxide as a method for euthanasia of laboratory mice and rats – a systematic review. [online] Available at: https://issuu.com/radboudumc/docs/the_use_of_carbon_dioxide_as_a_meth?e=28 355229/48248733 [Accessed 12 January 2018].

Taylor, K. and L. Rego (2016). EU statistics on Animal Experiments for 2014. *Alternatives to Animal Experimentation Proceedings*, 33(4), pp. 465–468.

Taylor, K., N. Gordon, G. Langley and W. Higgins (2008). Estimates for Worldwide Laboratory Animal Use in 2005. *Alternatives to Laboratory Animals*, 36(3), pp. 327–342.

Tsai, P.P., U. Pachowsky, H.D. Stelzer and H. Hackbarth (2002). Impact of Environmental Enrichment in Mice. 1: Effect of Housing Conditions on Body Weight, Organ Weights and Haematology in Different Strains. *Laboratory Animals*, 36(4), pp. 411–419.

Tsai, P.P., H.D. Stelzer, H.J. Hedrich and H. Hackbarth (2003). Are the Effects of Different Enrichment Designs on the Physiology and Behaviour of DBA/2 Mice Consistent?. *Laboratory Animals*, 37(4), pp. 314–327.

Tsai, P.P., H.D. Stelzer, A. Schraepler and H. Hackbarth (2006). Importance and Effects of Enrichment on Physiology, Behaviour and Breeding Performance in Mice. *Alternatives to Animal Experimentation*, 23, pp. 96–98.

Tsilidis, K., O. Panagiotou, E. Sena, E. Aretouli, E. Evangelou, D. Howells, R. Salman, M. Macleod and J.P.A. Ioannidis (2013). Evaluation of Excess Significance Bias in Animal Studies of Neurological Diseases. *PLoS Biology*, 11(7), p. e1001609. [online] Available at: http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001609 [Accessed 12 January 2018].

Tsukamoto, T. (2016) Animal Disease Models for Drug Screening: The Elephant in the Room? *Drug Discovery Today*, 21, pp. 529–530.

Turner, J.G., J.L. Parrish, L.F. Hughes, L.A. Toth and D.M. Caspary (2005). Hearing in Laboratory Animals: Strain Differences and Nonauditory Effects of Noise. *Comparative Medicine*, 55(1), pp. 12–23.

UC San Diego (2018). *Biomedical Research using Non Animal Models*. [online] Available at: https://extension.ucsd.edu/courses-and-programs/biomedical-research-using -non-animal-models [Accessed 12 July 2018].

Uhlig, C., H. Krause, T. Koch, M.G. de Abreu and P.M. Spieth (2015). Anesthesia and Monitoring in Small Laboratory Mammals Used in Anesthesiology, Respiratory and Critical Care Research: A Systematic Review on the Current Reporting in top-10 Impact Factor Ranked Journals. *PLoS One,* 10(8), p. e0134205. [online] Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0134205 [Accessed 18 July 2017].

van de Weerd, H.A., Aarsen, E.L., Mulder, A., Kruitwagen, C.L., Hendriksen, C.F. and Baumans, V. (2002). Effects of environmental enrichment for mice: variation in experimental results. *Journal of Applied Animal Welfare Science*, 5(2), pp. 87–109.

van der Poll, T. (2012). Experimental Human Sepsis Models. *Drug Discovery Today: Disease Models*, 9(1), pp. e3–e9.

van der Worp, H.B., D.W. Howells, E.S. Sena, M.J. Porritt, S. Rewell, V. O and M.R. Macleod (2010). Can Animal Models of Disease Reliably Inform Human Studies? *PLoS Medicine*, 7, pp. e1000245. [online] Available at: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000245 [Accessed 15 August 2017].

van Luijk, J. (2016). Improving Animal Research Using a Science Driven Approach: Sytematic Reviews of Animal Studies. In: *20th European Congress on Alternatives to Animal Testing, 17th Annual Congress of EUSAAT*. Linz. [online] Availabe at: https://norecopa.no/media/7599/systematic-reviews-judith-van-luijk.pdf [Accessed 15 December 2017].

van Luijk, J., B. Bakker, M.M. Rovers, M. Ritskes-Hoitinga, R.B. de Vries and M. Leenaars (2014). Systematic Reviews of Animal Studies: Missing Link in Translational Research?. *PLoS One*, 9(3), p. e89981. [online] Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0089981 [Accessed 15 August 2017].

van Luijk, J., Y. Cuijpers, L. van der Vaart, T.C. de Roo, M. Leenaars and M. Ritskes-Hoitinga (2013). Assessing the Application of the 3Rs: A Survey Among Animal Welfare Officers in the Netherlands. *Laboratory Animals*, 47(3), pp. 210–219.

van Luijk, J., M. Leenaars, A.M. van Dongen, L. van der Vaart and M. Ritskes-Hoitinga (2012). Outcomes of a Dutch Workshop on Improvements for the 3Rs in Daily Practice. *Alternatives to Animal Experimentation*, 29(4), p. 440–443. [online] Available at: https://www.altex.org/index.php/altex/article/view/439/447 [Accessed 15 August 2017].

van Praag, H., G. Kempermann and F.H. Gage (2000). Neural Consequences of Environmental Enrichment. *Nature Reviews Neuroscience*, 1(3), pp. 191–198.

vfa (Die forschenden Pharmaunternehmen) (2017). *Tierschutz. Preise des Landes Berlin für tierfreie Forschung zu Grippeviren und Schlaganfall verliehen.* [online] Available at: https://www.vfa.de/de/arzneimittel-forschung/forschungsstandort-deutschland/preise-des-landes-berlin-fuer-tierfreie-forschung-zu-grippeviren-und-schlaganfall-verliehen.html [Accessed 12 October 2017].

Voelkl, B., Vogt, L., Sena, E.S. and Würbel, H. (2018). Reproducibility of preclinical animal research improves with heterogeneity of study samples. *PLoS Biology,* 16(2), p. e2003693.

Vogt, L., T.S. Reichlin, C. Nathues and H. Würbel (2016). Authorization of Animal Experiments Is Based on Confidence Rather Than Evidence of Scientific Rigor. *PLoS Biology, 14 (12),* p. e2000598. [online] Available at: http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2000598 [Accessed 12 July 2017].

von Aulock, S. (2017). The 3Rs in Action, Session VIII-3: Establishing a Culture of Care Through Assessment, Transparency, and Communication. In: *10th World Congress on Alternatives and Animal Use in the Life Sciences* (*WC10*). Seattle: Alternatives to Animal Experimentation Proceedings, 6(1), p. 213. [online] Available at: http://www.altex.ch/resources/WC10_entire_issue1.pdf [Accessed 22 August 2017].

Waiblinger, E. and B. König (2004). Refinement of Gerbil Housing and Husbandry in the Laboratory. *Alternatives to Laboratory Animals*, 32 (1A), pp. 163–169.

Watts, N., M. Amann, S. Ayeb-Karlsson, K. Belesova, T. Bouley, M. Boykoff, P. Byass, W. Cai, D. Campbell-Lendrum, J. Chambers and P.M. Cox (2017). The Lancet Countdown on Health and Climate Change: From 25 Years of Inaction to a Global Transformation for Public Health. *The Lancet*, 391 (10120), pp. 581–630.

Weed, J. and J. Raber (2005). Balancing Animal Research With Animal Well-being: Establishment of Goals and Harmonization of Approaches. *ILAR Journal*, 46(2), pp. 118–128.

Wells, S. (2017). Pain Assessment and New Innovations. In: *10th World Congress on Alternatives and Animal Use in the Life Sciences* (*WC10*). Seattle: Alternatives to Animal Experimentation Proceedings, 6(1), p. 222. [online]Available at: http://www.altex.ch/resources/WC10_entire_issue1.pdf [Accessed 22 August 2017].

Wever, K.E., F.J. Geessink, M.A. Brouwer, A. Tillema and M. Ritskes-Hoitinga (2017). A systematic review of discomfort due to toe or ear clipping in laboratory rodents. *Laboratory Animals*, 51(6), pp. 583–600.

Wiedenmayer, C. (1997). Causation of the Ontogenetic Development of Stereotypic Digging in Gerbils. *Animal Behaviour*, 53(3), pp. 461–470.

Wolfensohn, S., P. Hawkins, E. Lilley, D. Anthony, C. Chambers, S. Lane, M. Lawton, H. Voipio and G. Woodhall (2013a). Reducing Suffering in Experimental Autoimmune Encephalomyelitis (EAE). *Journal of Pharmacological and Toxicological Methods*, 67(3), pp. 169–176.

Wolfensohn, S., P. Hawkins, E. Lilley, D. Anthony, C. Chambers, S. Lane, M. Lawton, S. Robinson, H. Voipio and G. Woodhall (2013b). Reducing Suffering in Animal Models and Procedures Involving Seizures, Convulsions and Epilepsy. *Journal of Pharmacological and Toxicological Methods*, 67(1), pp. 9–15.

Wolfer, D., O. Litvin, S. Morf, R. Nitsch, H. Lipp and H. Würbel (2004). Laboratory Animal Welfare: Cage Enrichment and Mouse Behaviour. *Nature*, 432(7019), pp. 821–822.

Wolfle, T.L. (1987). Control of Stress Using Non-drug Approaches. *Journal of the American Veterinary Medical Association,* 191, pp. 1219–1221.

World Organisation for Animal Health (OIE) (2018). *Terrestrial Animal Health Code, Ch. 7.8.* Paris: World Organisation for Animal Health, first adopted in 2010; most recent

update adopted in 2013. [online] Available at: http://www.oie.int/index.php?id
=169&L=0&htmfile=chapitre_aw_research_education.htm [Accessed 12 May 2017].

Würbel, H. (2000). Behaviour and the Standardization Fallacy. *Nature Genetics,* 26,
p. 263.

Würbel, H. (2001). Ideal homes? Housing Effects on Rodent Brain and Behaviour.
Trends in Neuroscience. 24, pp. 207–211.

Würbel, H. (2007). Environmental Enrichment Does Not Disrupt Standardisation of
Animal Experiments. *Alternatives to Animal Experimentation*, 24, pp. 70–73.

Würbel, H. (2016). Tierversuch und Irrtum. In: M. Fehlmann, M. Michel and R. Nieder-
hauser, eds. *Tierisch! Das Tier und die Wissenschaft: Ein Streifzug durch die Disziplin-
en*, Reihe Zürcher Hochschulforum, 55, pp. 97–106.

Würbel, H. (2017). More Than 3Rs: The Importance of Scientific Validity for Harm-benefit
Analysis of Animal Research. *Nature Lab Animal*, 46(4), pp. 164–166. [online] Avail-
able at: https://www.nature.com/articles/laban.1220.pdf [Accessed 12 June 2017].

Würbel, H. and J.P. Garner (2007). Refinement of Rodent Research Through Environ-
mental Enrichment and Systematic Randomization. *National Centre for the Replace-
ment, Refinement, and Reduction of Animals in Research (NC3Rs).* [online] Available at:
https://www.nc3rs.org.uk/sites/default/files/documents/Refinementenvironmen
talenrichmentandsystematicrandomization.pdf [Accessed 12 August 2017].

Würbel, H. and M. Stauffacher (1994). Standardhaltung für Labormäuse—Probleme
und Lösungsansätze. *Tierlaboratorium*, 17, pp. 109–118.

Würbel, H. and M. Stauffacher (1996). Prevention of Stereotypy in Laboratory
Mice: Effects on Stress Physiology and Behaviour. *Physiology & Behavior*, 59(6),
pp. 1163–1170.

Würbel, H., M. Stauffacher and D. Holst (1996). Stereotypies in Laboratory Mice—
Quantitative and Qualitative Description of the Ontogeny of 'Wire-gnawing'and
'Jumping'in Zur: ICR and Zur: ICR nu. *Ethology*, 102(3), pp. 371–385.

Zeeff, S.B., C. Kunne, G. Bouma, R.B. de Vries and A.A. Te Velde (2016). Actual usage and
Quality of Experimental Colitis Models in Preclinical Efficacy Testing: A Scoping
Review. *Inflammatory Bowel Diseases*, 22(6), pp. 1296–1305.

Zerhouni, E.A. (2014). Turning the Titanic. *Science Translational Medicine*, 6(221),
p. 221ed2.

Zurlo, J. (2012). No Animals Harmed: Toward a Paradigm Shift in Toxicity Testing. *Hast-
ings Center Report*, 42(s1). [online] Available at: http://animalresearch.thehastings
center.org/report/no-animals-harmed-toward-a-paradigm-shift-in-toxicity-testing/
[Accessed 12 June 2017].

Zurlo, J., D. Rudacille and A.M. Goldberg (1996). The Three R's: The Way Forward. *Envi-
ronmental Health Perspectives*, 104(8), pp. 878–880.