Animal Experimentation in Research: The 3Rs Principle and the Validity of Scientific Research

Guidelines of the Permanent Senate Commission on Animal Protection and Experimentation of the DFG for the Design and Description of Animal Experimental Research Projects

DFG

Version: July 2019

Contact:

Dr. Sonja Ihle
Life Sciences 1 Division
Phone: 0228/885-2362
E-Mail: sonja.ihle@dfg.de

Table of Contents

1 Summary

The right to freedom of research as enshrined in the Basic Law for the Federal Republic of Germany and the constitutional objective of animal welfare are two important values with high relevance to animal experimentation in research.

The use of animals for scientific purposes can be ethically justified only if an animal experiment is “indispensable” and the expected scientific outcome justifies the harm imposed on the animals in the course of the experiment.

A crucial ethical guideline in animal experimentation is the 3Rs principle (Replace, Reduce, Refine)[1]: Animal experiments may be performed only if no other suitable methods are available to investigate the research question and if the number of animals and the harms imposed on them are limited to the unavoidable minimum. The consistent application of the 3Rs principle is very much in the interest of science itself, since impaired welfare of experimental animals can also compromise the validity of research findings. The effort to maximise the scientific validity and replicability of research findings while observing animal welfare must always be the foundation of study design in animal experimentation.

In practice, tensions may arise between measures taken to advance animal welfare in research and the requirements for ensuring scientific validity. Because of the interdependence of these two aspects, policies for implementing the 3Rs principle should not be considered in isolation, but should rather be integrated in the study design and included in descriptions of research projects.

With these guidelines, the DFG Senate Commission on Animal Protection and Experimentation aims to contribute to the debate on quality in biomedical research and to help define the specific requirements for conducting animal experiments. In addition, this publication supports researchers in the design and adequate description of research projects involving animal experimentation.

---

1 Russell, W. M. S., Burch, R. L. (1959). The principles of humane experimental technique. London: Methuen.

2 The Relationship between Animal Welfare in Research and Scientific Quality

Scientific research is based on the aspiration to achieve results of the highest possible quality and validity. In discussions surrounding the "replication crisis" in research,[2,3] the DFG in its statements regarding the replicability of research findings has emphasised the need to ensure the quality of research as a fundamental standard in science. It has therefore initiated a discourse on subject-specific reflection concerning quality and replicability as well as on various aspects of quality assurance.[4,5]

Animal experimentation must ensure the scientific validity of its findings and at the same time the appropriate and responsible care and use of experimental animals. High animal-welfare standards are a prerequisite for high-quality research, and conversely, animal experiments can only be justified if the quality of the research and thus the validity of its results are guaranteed.

Due to the fundamental imperative to protect animals, researchers bear a special moral responsibility when using animals in experiments. The 3Rs principle provides guiding criteria for animal welfare policies in research by calling for the refinement, reduction and replacement of animal experiments. Measures to implement the 3Rs principle and the objective to maximise scientific validity, which is inherent to all scientific research, should not be considered

---

2 Ioannidis, J. (2005). Why most published research findings are false. PLoS Medicine, 2(8): 696–701.

3 McLeod, M. R. et al. (2014). Increasing value, reducing waste. The Lancet, 383(9912): 101–104.

4 Deutsche Forschungsgemeinschaft (2017). Replicability of Research Results: A Statement by the German Research Foundation.

5 Deutsche Forschungsgemeinschaft, Permanent Senate Commission on Key Questions in Clinical Research, Quality in Clinical Research Working Group (2018). Replizierbarkeit von Ergebnissen in der Medizin und Biomedizin: Stellungnahme der Arbeitsgruppe „Qualität in der Klinischen Forschung" der DFG-Senatskommission für Grundsatzfragen in der Klinischen Forschung.

separately. Furthermore, the expected harm imposed on the animals and the intended knowledge gain from the research project provide the basis for a harm – benefit analysis to determine whether the use of animals for research purposes can be justified.[6]

It is thus in the very interest of science to strive for high scientific standards in animal experimentation and to incorporate animal welfare policies in the project design.

To support this endeavour, the focus of these guidelines is deliberately limited to those aspects of the 3Rs principle that are relevant when designing, planning and implementing projects that feature animal experimentation. Specific areas of conflict are illustrated by examples. Overarching issues of quality assurance in animal experimentation in general are addressed in the statement by the "Quality in Clinical Research Working Group", which is endorsed by the Senate Commission on Animal Protection and Experimentation.

6 Deutsche Forschungsgemeinschaft, Permanent Senate Commission on Animal Protection and Experimentation (2016). Animal Experimentation in Research.

3 More than 3Rs: The 3Rs Principle and Scientific Validity

3.1 Legitimising Animal Experimentation by Means of Harm – Benefit Analysis

The use of animals in research is governed by the principle of proportionality. This rule-of-law principle is applied in cases of conflict between different fundamental rights, legal interests or legal principles under the Basic Law (here: general personal rights, integrity of life and limb, freedom of research, animal welfare). Accordingly, experiments on animals are legitimate if the expected benefit of the research outcomes outweighs the harms inflicted on the animals. The tool for evaluating the proportionality of an animal experiment is the harm–benefit analysis. It entails determining whether an experimental project with a legitimate purpose is (a) suitable, (b) necessary and (c) appropriate in order to achieve the intended gain in knowledge.[7]

Suitability and necessity are prerequisites for performing a harm – benefit analysis; they cover the scientific justification of the experiment (suitability) and potential alternatives to the use of animals and to harmful methods (necessity). Only if an experimental project appears to be scientifically sensible and suitable and if no alternatives to animal experiments are available can a harm – benefit analysis be made. Such an analysis examines whether the experiments are appropriate, that is whether the expected benefit of the project justifies the harm imposed on the experimental animals. Only if this is the case can the experimental project be authorised.

---

7 Wienbracke, M. (2013). Der Verhältnismäßigkeitsgrundsatz. Zeitschrift für das Juristische Studium (ZJS). No. 2: 148 – 155.

3.2 The 3Rs Principle as a Normative Criterion for Animal Welfare in Research

According to the German Animal Welfare Act (TierSchG), every animal experiment must be submitted for authorisation to the competent authority before it can be carried out (§ 8 TierSchG). The authorisation procedure includes an explicit review of whether the 3Rs principle (Replace, Reduce, Refine) in its various aspects will be implemented in the best possible way. Applicants must demonstrate that the desired gain in knowledge cannot be achieved without the use of sentient animals (Replace), with fewer animals (Reduce) or with less harmful methods (Refine). The possibility of using methods without experimental animals (alternative methods) or of using animals considered to be less sentient (relative replacement) must be taken into consideration. The minimisation of the number of animals requires careful experimental design with sample size calculation (e.g. power analysis). Numerous methods are available to minimise the harm inflicted on animals. For example, improving housing conditions through environmental enrichment, gradual familiarisation of animals with the experimental conditions, use of non-invasive techniques, and optimised anaesthesia and analgesia methods can shift the harm–benefit ratio in favour of an animal experiment.

3.3 The 3Rs Principle in the Context of Validity and Replicability

Careful implementation of the 3Rs is a necessary but not sufficient prerequisite for the ethical justification of animal experiments. Regardless of the severity of procedures imposed on the animals, the study design must meet stringent scientific quality requirements in terms of objectivity, validity, and replicability. Fundamental to every researcher's responsibility is the assurance of scientific quality with the aim of maximising scientific validity. Furthermore, this aim must always be a key criterion in the scientific review of research projects.

A positive evaluation of the scientific validity (see Section 3.4) and the application of the 3Rs principle are prerequisites for determining the appropriateness

of an experimental project in the final harm–benefit analysis. In keeping with the principle of proportionality, the criterion of scientific validity may be given precedence over the implementation of the 3Rs in view of animal welfare, particularly in consideration of the fundamental right of freedom of research, which can only be restricted to the extent that it conflicts with other fundamental rights.[8] The 3Rs principle can only fulfil its purpose if there is a suitable experimental design, and the Reduce principle must not be misunderstood in such a way that the sample size of a study is lowered at the expense of scientific validity (see Section 4.2). Whether an animal experiment is appropriate when taking into account its scientific quality and necessity, and can thus be authorised must be determined in the final harm – benefit analysis.

3.4 Aspects of Scientific Validity in Animal Experimentation

In contrast to the 3Rs principle as a central component in the evaluation of the ethical justifiability of animal experimentation, no similarly pithy concept exists for assessing the suitability of animal experiments. In general, however, an examination of the scientific validity of experimental findings can provide good guidance. The explanatory value of experimental findings largely depends on three different aspects of scientific validity, namely the quality of the animal model or experimental model, the quality of the postulated cause–effect relation, and the degree of generalisability. Accordingly, Würbel (2017) categorises these into construct validity, internal validity, and external validity:[9]

Construct validity refers to the validity of the animal model or experimental model and of the methods used to investigate the research question of an experimental project. It describes the accuracy with which an experimental setup

8 Wienbracke, M. (2013). Loc. cit.

9 Würbel H. (2017). More than 3Rs: The importance of scientific validity for harm-benefit analysis of animal research. Lab Animal, 46(4): 164–166.

measures what it claims to measure.[10] The assessment of construct validity should therefore focus on empirical evidence of the biological congruence between properties of the test population and those of the target population, as well as on the significance of the primary outcome variable of the experiment for the function to be measured.[11,12]

Internal validity refers to the quality of the postulated cause–effect relation. Internal validity is present if a change in the dependent variable (a treatment effect) can be demonstrably attributed to variation in the experimental treatment. An assessment of internal validity must therefore take into account not only fundamental aspects of experimental design and the use of appropriate control groups, but also all applicable criteria of good research practice in order to avoid risks of bias (e.g. randomisation, blinding, sample calculation, definition of primary and secondary target variables, definition of inclusion and exclusion criteria, definition of the statistical analysis plan, etc.).

External validity refers to the degree to which experimental results can be generalised beyond the specific conditions of a given experiment. In empirical research, the generalisability of findings often takes priority. The replicability of results alone depends on a minimum of external validity, since experimental conditions inevitably differ between replicate experiments (both within and between experimental laboratories). The assessment of external validity should therefore take into account aspects of the experimental design that allow conclusions to be drawn about replicability and generalisability (e.g. to other animal models, to other experimental conditions). This includes, among others, the

---

10 Koob, G. F., Heinrichs, S. C., Britton K. (1998). Animal models of anxiety disorders. The American Psychiatric Press Textbook of Psychopharmacology. (2. Schatzberg AF, Nemeroff CB, editor). Washington DC – London: American Psychiatric Press: 133–144.

11 Willner, P. (1984). The validity of animal models of depression. Psychopharmacology, 83(1): 1–16.

12 Belzung, C., Lemoine, M. (2011). Criteria of validity for animal models of psychiatric disorders: focus on anxiety disorders and depression. Biology of Mood & Anxiety Disorders, 1(1): 9.

consideration of both sexes, the division of an experiment into several independent replicates, the systematic variation of one or more independent variables (e.g. different mouse lines, housing conditions, experimental setups or experimenters), and the conduct of multi-laboratory studies.

These three aspects of the validity of findings from animal experimentation cover essential criteria for assessing the scientific validity and replicability of animal experiments. In analogy to the 3Rs principle, which enables evaluation of the necessity of animal experiments, the application of these three aspects of validity can facilitate the evaluation of the suitability of animal experiments to achieve the intended benefit.

4 Areas of Tension between Scientific Validity and the 3Rs

Researchers have a great stake in implementing animal welfare policies in their projects, and not just for ethical considerations. Conditions that compromise the welfare of laboratory animals (pain, suffering, harm) can impair the validity of research findings. Not only for reasons of animal welfare but also with regard to the quality of research, minimising the harms imposed on animals is therefore a cardinal concern of researchers. The 3Rs principle and scientific validity go hand in hand in most cases.

In some areas of experimental design, however, there is profound tension between the implementation of the 3Rs principle and the maximisation of scientific validity. This can also affect project evaluation in authorisation procedures. The application of binding criteria of scientific validity (such as construct validity, internal and external validity) helps to determine the impact of 3Rs policies on key aspects of scientific validity.

Areas of tension exist with regard to all 3Rs; when it comes to the use of certain animal species, the determination of the number of animals (sample size) and the application of certain experimental methods. In order to help researchers to plan, describe and justify their projects, some of these areas of tension will be illustrated and possible solutions suggested below.

4.1 Selection of Animal Species

Replace does not only refer to the replacement of animal experiments by alternative methods (in silico, in vitro, etc.), but also includes relative replacement, i.e. switching to “lower” animals (invertebrates) or “lower” vertebrates or mammals (e.g. mice instead of primates, fish instead of mammals). The background of the concept of relative replacement is the widespread view that “higher” vertebrate species (such as primates, but also dogs, cats) suffer

more from the effects of experimentation than “lower” species.[13] This view is also reflected in EU Directive 2010/63/EU on the protection of animals used for scientific purposes, which grants special protection to primates and certain other animal species. The postulated hierarchy of suffering of “higher” versus “lower” species under given experimental conditions is questionable from a scientific point of view and is also controversial from an ethical point of view.[14,15] The Animal Welfare Act is based on the assumption that all vertebrates (and some invertebrates) are sentient and does not rule this out for other species.[16] There exist as yet no binding criteria to establish an ethical hierarchy within vertebrates based of varying degrees of sentience. Relative replacement can also be quite critical from a scientific point of view if the selection of the animal model impacts scientific validity. In some cases, this may diminish the usefulness of the research (knowledge gain) and thereby also its ethical justification.

- *In addition to an ethical and legal review of the postulated ethical hierarchy within the animal species protected by the Animal Welfare Act, there is also a need for greater problem awareness. When planning experiments, researchers should carefully consider the choice of their animal models and justify them in the harm–benefit analysis on the basis of scientific validity criteria and the species-specific disposition to suffer under the given experimental conditions.*

13 For example, the Animal Welfare Act (§ 7 (1) (1)) stipulates that animal experiments must be limited to the indispensable minimum, taking into account the experimental animals’ species-specific ability to suffer from the effects of the experiment.

14 Rippe, K. P. (2003). Tierethik. Bioethik. Ed. M. Düwell & K. Steigleder. Frankfurt a. M.

15 Ach, J. S., Borchers, D. (eds.) (2018). Handbuch Tierethik: Grundlagen – Kontexte – Perspektiven. Stuttgart: J. B. Metzler.

16 Animal Welfare Act § 8 (a) (4).

4.2 Determination of the Number of Animals

Reduce refers mainly to the number of animals used in an experiment (total sample size). From a scientific point of view, larger samples would often be desirable, but from an animal welfare point of view, smaller samples are preferred. A formal sample size calculation (e.g. by power analysis) is in many cases an appropriate way to determine the minimum number of animals.[17] Since putative progress in animal welfare in research is measured and communicated to the public primarily on the basis of annual statistics on animals used in experimentation, there is also considerable political pressure to reduce the number of animals. Due to the abovementioned views assuming an ethical hierarchy among vertebrates, this applies in particular to experiments involving "higher" mammals.

Excessive minimisation of the sample size based on the 3Rs principle (Reduce) becomes problematic if it comes at the price of subcritical scientific validity. This is particularly the case when important comparison or control groups are omitted or samples are used that are too small. It should be borne in mind that the ethical "costs" of samples that are too small can be higher than those of samples that are too large, because in the worst case, animal experiments below a certain sample size may become worthless due to a lack of statistical power.

Researchers must therefore be able to justify the number of animals used in their experiments on the basis of scientific criteria (such as the internal and external validity of the research findings). There are numerous tools for experimental design (e.g. the Experimental Design Assistant, https://eda.nc3rs.org.uk/) and for calculating and justifying the sample size (e.g. G*Power, http://www.gpower.hhu.de/). In many experiments, it is also possible to use adaptive experimental design (e.g. group-sequential design, interim analysis with adjustment of sample size, adaptive randomisation, etc.) in order to minimise the number of animals without diminishing scientific va-

17 Festing, M. F. W. (2018). On determining sample size in experiments involving laboratory animals. Laboratory Animals, 52(4): 341–350.

lidity. For more complex experimental approaches, however, it is advisable to seek professional support from biostatisticians with expertise in animal experimental design.

- *Determining the number of experimental animals requires a careful balance between striving to reduce the number of animals and ensuring scientific validity. These considerations should be explained in the project description.*

4.3 Standardisation

Animal experiments are usually carried out under controlled laboratory conditions. Controlled experimental conditions are often desirable in basic research in order to exclude any confounding variables. Animals and housing conditions are usually rigorously standardised by housing animals that are genetically largely identical (e.g. inbred lines) under identical conditions (in terms of cage, feed, management). Standardisation has also been recommended as a means of minimising the number of animals (Reduce), since treatment effects can be statistically demonstrated with smaller sample sizes when there is less variation in the experimental results.

Although standardisation can improve the precision of experimental findings (less variation), excessive standardisation affects the external validity and thus the generalisability of the results.[18] Therefore, rigorous standardisation of experimental conditions can at best minimise the number of animals used per experiment, but in turn may necessitate a greater number of independent experiments – and thus ultimately more animals – in order to demonstrate reproducibility under varying conditions and thus generalisability.[19]

---

18 Richter, S. H. et al. (2009). Environmental standardization: cure or cause of poor reproducibility in animal experiments? Nature Methods, 6: 257–261.

19 Voelkl, B., Vogt, L., Sena, E. and Würbel, H. (2018). Reproducibility of preclinical animal research improves with heterogeneity of study samples. PLoS Biology, 16(2), e2003692.

Researchers should therefore design experiments in such a way that they can also draw conclusions about external validity (i.e. generalisability across both sexes as well as genetic and environmental variation). A suitable method for this is controlled heterogenisation of the experimental population by means of multi-factorial experimental designs, whereby both sexes, several genotypes, and/or several environmental or experimental conditions can be taken into account without requiring more animals.[20] Particularly robust results can be achieved with multi-laboratory studies.[21] Although not always feasible, this appears to be a useful addition to research methodology, especially in large research networks.

- *When planning experiments, researchers should take into account external validity, i.e. replicability and generalisability, and justify the study design accordingly.*

4.4 Pilot Studies and Exploratory Experiments

Many experimental projects are part of large-scale research programmes. Exploratory studies are often carried out as a first step to generate hypotheses, followed by pilot studies to optimise the experimental design with a view towards confirmatory studies to test specific hypotheses. For exploratory and pilot studies, as well as in some areas of basic research, it is not possible to perform accurate sample size calculations.[22] However, the studies not only serve to generate promising hypotheses and optimise experimental designs, but also help to minimise the number of animals used in subsequent large-scale research programmes. Otherwise, there is a risk of testing

---

20 Shaw, R., Festing, M. F. W., Peers, I. and Furlong, L. (2002). Use of Factorial Designs to Optimize Animal Experiments and Reduce Animal Use. ILAR Journal, 43(4): 223–232.

21 Voelkl, B., Vogt, L., Sena, E. and Würbel, H. (2018). Loc cit.

22 Mayer, B., Muche, R. (2013). Formal sample size calculation and its limited validity in animal studies of medical basic research. Tierärztliche Praxis, Ausgabe K: Kleintiere – Heimtiere, 41(6): 367–374.

unproductive hypotheses or testing hypotheses with immature experimental designs and inappropriate methods, which can significantly impair scientific validity.

- *Exploration and pilot studies should be designed such that, with a view towards follow-up studies or large-scale research programmes, they enable better focus and thus minimisation of the total number of animals required.*

4.5 Replicate Experiments

According to the Animal Welfare Act, animal experiments are indispensable only if they expand the current state of scientific knowledge and if the gain in knowledge outweighs the harm imposed on the animals. This must also be taken into account when justifying replicate experiments in order to assess the replicability of experimental findings.

However, replicability is a cornerstone of scientific evidence. Without verification of research results, there is a high risk of unproductive follow-up experiments based on promising but non-replicable findings. On the one hand, this underscores the importance of experimental approaches that allow an assessment of the external validity and replicability of experimental findings (e.g. multi-factorial designs, multi-laboratory studies; see Section 4.3). On the other hand, researchers must have the possibility to verify the replicability of previous findings independently.

- *Researchers should consider and justify replicate experiments (e.g. to assess the replicability of experimental findings). Under certain circumstances, necessary replicate experiments may be incorporated as positive or negative controls into the design of new experiments.*

5 Appendix

Reporting Guidelines

ARRIVE Guideline (NC3Rs)
Kilkenny, C. et al. (2010): *Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research.* PLoS Biology, 8(6), e1000412.
https://www.nc3rs.org.uk/arrive-guidelines

HARRP Guidelines (ICLAS)
Osborne, N. et al. (2018): *Improving animal reporting standards.* EMBO Reports, 19(5), e46069.
http://iclas.org/

PREPARE Guidelines (NORECOPA)
Smith, A. et al. (2017): *PREPARE: Guidelines for planning animal research and testing.* Laboratory Animals, 52(2).
https://norecopa.no/prepare

Overview of additional reporting guidelines:
https://www.nlm.nih.gov/services/research_report_guide.html

Databases for preregistration of animal experimental studies

Animal Study Registry (Federal Institute for Risk Assessment)
https://www.animalstudyregistry.org/

PreClinicalTrials.EU
https://www.preclinicaltrials.eu/

Further information and links

What are the 3Rs? (NC3Rs)
https://www.nc3rs.org.uk/the-3rs

Experimental Design Assistant (NC3Rs)
https://eda.nc3rs.org.uk/

G*Power (Heinrich-Heine University Düsseldorf)
http://www.gpower.hhu.de/

6 Members of the Permanent Senate Commission on Animal Protection and Experimentation of the DFG

Professor Brigitte Vollmar, Rostock (chair)

Professor Peter Dabrock, Erlangen

Professor Bernd Fleischmann, Bonn

Professor Christiane Herden, Gießen

Professor Nils Hoppe, Hannover

Professor Lars Lewejohann, Berlin

Professor Heidrun Potschka, Munich

Professor Stefan Schlatt, Münster

Professor René H. Tolba, Aachen

Professor Stefan Treue, Göttingen

Professor Hanno Würbel, Bern

Commission secretariat

Dr. Cornelia Exner, Marburg

Dr. Valeska Stephan, Rostock

DFG Head Office

Dr. Sonja Ihle, Bonn

Deutsche Forschungsgemeinschaft

German Research Foundation

Kennedyallee 40 · 53175 Bonn

Postal address: 53170 Bonn

Tel: +49 228 885-1

Fax: +49 228 885-2777

postmaster@dfg.de

www.dfg.de/en