

PEER REVIEW HISTORY

BMJ Open Science publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://openscience.bmj.com/pages/wp-content/uploads/sites/62/2018/04/BMJ-Open-Science-Reviewer-Score-Sheet.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Behavioral effects of methylphenidate in an animal model of attention-deficit/hyperactivity disorder, the spontaneously hypertensive rats: a systematic review and meta-analysis protocol
AUTHORS	

VERSION 1 - REVIEW

REVIEWER 1	Hanno Würbel
REVIEW RETURNED	31-01-18

GENERAL COMMENTS	<p>Recommendation * Major Revision</p> <p>In a position paper, Strech and Dirnagl propose extending the 3Rs (replace, reduce, refine) concept for humane experimental techniques by another set of 3Rs (robustness, registration, reporting) to promote the value of animal research. This is certainly a laudable proposal given that research on animals can only be humane if it generates value for society. No matter how few animals are used and how little harm is imposed on them, if the research is not meaningful or the results lack validity, the animals have been wasted for no good reason. For the exact same reason, I have recently proposed complementing the 3Rs by 3Vs (construct validity, internal validity, external validity) in project evaluation of animal research (Würbel 2017, cited in the manuscript). Strech and Dirnagl aim for a bigger picture, including other dimensions of research integrity beyond rigour and robustness, such as pre-registration and proper reporting.</p> <p>As much as I support every single one of their recommendations, I have several issues with the conceptualization and branding of their proposal.</p> <p>- Going for the big picture, the authors write, “research is only valuable if it is innovative, useful, robust, and accessible”, suggesting that these four aspects describe the value of research. However, they then note that</p>
-------------------------	---

“innovation” and “usefulness” are hard to gauge and therefore restrict their concept of value to “robustness” and “accessibility”. Therefore, from the outset it seems that 2 further Rs are missing, because the authors failed to find a practicable operationalisation for them (or an adequate R-term).

- They use the term “robustness” to cover all aspects of scientific validity. I am not convinced that robustness is the right term here. Perhaps, “rigour” would be more accurate? For example, the NIH has recently defined “scientific rigour” as “the strict application of the scientific method to ensure robust and unbiased experimental design, methodology, analysis, interpretation and reporting of results. This includes full transparency in reporting experimental details so that others may reproduce and extend the findings”. Here, scientific rigour includes all aspects of scientific validity (internal and external validity), as well as reporting and (methods) reproducibility – which would mean that the 6th R “reporting” proposed here would be obsolete as it is already included in the definition of “rigour”. The authors argue that “robustness” is included in Articles 21 and 22 of the Declaration of Helsinki; however, while these two articles contain general descriptions of scientific requirements and research protocols, they do not use the term “robustness”.

- Whereas “robustness” (as used by the authors) and accurate “reporting” are necessary aspects of scientific rigour, it is at least questionable whether this is true for “preregistration”. For example, although it makes sense that both exploratory studies and pilot studies need formal approval by the authorities, I am not convinced that it makes sense to ask for preregistration of such studies. This is relevant insofar as adherence to each of the 3Rs is mandatory by law; thus, extending the 3Rs concept by Rs that are not (and should not be) legally mandatory appears problematic. (See also the issue of “diluting the 3Rs principles” below.)

- The authors write, “regulators, ethics boards, scientists, and funders should add robustness, registration, and reporting to their criteria when planning, licensing, or funding animal experiments.” However, reporting cannot be assessed by funders or regulators but only by journal editors. Since the 3Rs concept is a major pillar in project evaluation and a formal

part of the harm-benefit analysis (see below), the inclusion of a criterion that is not applicable to project evaluation is not ideal.

- There have been quite some other proposals already for adding further Rs to the 3Rs concept, none of which has had a lasting impact. It is rather unlikely that the present proposal will succeed in establishing a strong brand. For example, the German Max-Planck Society has added a 4th R for “responsibility”, others have proposed “reproducibility” or “rehabilitation” as a 4th R. There have also been proposals for adding 2Rs, e.g. “robustness and reproducibility” or “rigor and reproducibility”

(<https://onlinelibrary.wiley.com/doi/full/10.15252/embr.201744428>), and – similar to these authors – others have previously used another set of 3Rs (“rigor, reproducibility and robustness”) to describe research quality (https://www.nature.com/polopoly_fs/1.21707!/menu/main/topColumns/topLeftColumn/pdf/543613a.pdf). One author has referred to this use of R-terms (Replication, Reproducibility, Rigor, Robustness) as a “rhumba of R’s”. Therefore, I am not convinced that yet another Rs concept is a fruitful way forward.

- I think there are good reasons for leaving the 3Rs intact and using a separate concept for scientific value. First, the 3Rs concept is not just a strong concept that has been adopted almost universally as a guiding principle for humane animal experimentation; it is also a very strong brand. Adding other Rs will likely dilute the concept, and its message, especially if the resulting (4Rs, 5Rs, or 6Rs) concept is not as coherent as the original 3 Rs concept. Second, the 3Rs play a critical role in the harm-benefit analysis (HBA) of animal research, which is at the heart of project evaluation and authorisation. The HBA is based on two preconditions that need to be met: the proposed use of animals needs to be suitable for achieving the proposed objectives (scientific rigour), and it needs to be unavoidable (no further potential for replacement, reduction, and refinement). If these preconditions are met, an ethical evaluation (the actual HBA) can be made as to whether the expected benefits (scientific and societal value) outweighs the expected harms. From the conception of the HBA it should become clear that the 3Rs act on the harm side (with the aim to minimize the harms imposed on animals), while scientific rigour acts on the benefit

side (with the aim to maximize the benefit of the research). Adding more Rs to the 3Rs concept will likely blur this important distinction.

REVIEWER 1	Tom Matheson
REVIEW RETURNED	31-01-18

GENERAL COMMENTS	<p>Recommendation * Minor Revision</p> <p>This position paper makes a well-reasoned argument that funding bodies and publishers should consider not only Replacement, Reduction and Refinement in animal experimentation, but also Robustness (of experimental design and statistics), Registration (of experiments prior to work commencing) and more thorough Reporting of outcomes. These are entirely reasonable points to consider and I am very supportive of highlighting them to provoke further detailed discussions about how to implement them effectively. My points below are relatively minor and aimed at tightening up the arguments put forward.</p> <p>Abstract.</p> <p>The phrase (line 4) “lacks scientific value” is rather vague. Do you mean ‘has no scientific value’, or ‘has less than the maximum possible scientific value’ or something in between? The term (line 7) “safeguard” might better be ‘improve’ or ‘strengthen’. It may be judicious to make it clear whether you are referring to research on any animal (including insects, molluscs etc), or just those vertebrates and cephalopods regulated by e.g. the Home Office in the UK.</p> <p>Introduction and main text.</p> <p>I am not completely convinced by your statement (lines 5-7) that the 3Rs do not consider (“represent”) “value for science and society”. Although the 3Rs acronym might not explicitly address this, I would argue that the current implementation in grant funding committees</p>
-------------------------	--

does seek to ensure (for example) that studies provide sufficient power, and that there are suitable pathways to tangible benefits ('impact'). Perhaps reconsider how you phrase this section and the strong inference expressed in lines 12-13 that scientific and societal value are not currently "considered". I think the claim (line 7) that "research is only valuable if it is innovative..." is moot. Replication for example might be important, but not innovative. I think it would be useful to define "registration" (Abstract and Introduction line 11) early on in the article. There is a typo in "biorXiv": it should be 'BioRxiv'. Where you cite ref 16, it would be useful to describe the '3Vs' that Würbel proposes. The final section on National Centres (NC3Rs etc) partially addresses some of my points above about current implementations, but it is too late here, and left hanging. You need to make a clearer point. Note that the correct acronym is 'NC3Rs'.

Summary

The Summary is largely repetitive of the Abstract, and adds little. I would have liked a more concrete proposal about the next steps.

REVIEWER 1	S Clare Stanford
REVIEW RETURNED	31-01-18

GENERAL COMMENTS

Recommendation *
Major Revision

The authors are right to point out that the existing 3Rs (Replacement, Refinement, Reduction) all focus on animal welfare. It is also true that research that cannot be trusted (e.g., through flawed design or bias) has no value. In fact, I would go further and say it can be harmful. To help address these problems, this article proposes that there should be three more Rs: Robustness, Registration and Reporting. None of these suggestions is new, as the list of citations confirms. The novelty of this article lies in proposing the three additional Rs, as a mantra. I think the authors are on to something and that it could catch on. However,

although I am sure everyone would agree with adding these Rs, as an aspiration, I have several comments on this article and some suggestions for revisions.

1. The Introduction includes the comment that research is ethical only if it is of value to science and society. That is true, but these extra Rs do not address some key elements that contribute to the value of the research. In fact, the text draws attention to the challenging question of how we evaluate the 'value' of research (to science and society) and whether its value meets ethical standards. For instance, the UK has already banned research on cosmetics and tobacco, but debates are developing about the ethics of some other research objectives, especially those directed at resolving problems that are deemed by many people to arise from 'choice' of life-style. The authors rightly state that 'value' is an essential element of Robustness, so perhaps they should either offer suggestions on how to evaluate 'value' to society or define their concept of value more precisely?

2. Introduction (para 1, lines 7 – 8). I agree that the usefulness of a research study might not be apparent for many years, but I disagree that the same is the case for 'innovation'. Novelty is a key requirement for a recommendation for grant funding or publication.

3. The authors should be more specific about their recommendation for Registration and their feasibility / obstacles. For instance, should the registration be openly accessible, or merely held in a secure archive for scrutiny, after the work has been done? If the latter, who would be granted access? Would this be an element of the peer review of the completed work? How would registration affect ownership of intellectual property? How would the information to be registered differ from that already required for a protocol carried out under Good Laboratory Practice (or are we heading for a requirement for all research to be GLP compliant?) To avoid Registration being a flimsy aspiration, I think the authors should express specific views on these points. Perhaps a comment on why many major research-funders are insisting on Open Access for publications but not Registration (yet), would be appropriate too?

4. My main criticism of this article is that each of the original 3Rs (Replacement, Refinement and Reduction) is a single, simple concept and can be used in a check-list for compliance. The same is true for Registration and Reporting, but Robustness is an odd-ball in the series. The article (e.g., paragraph 4). makes it clear that Robustness is a multifactorial concept, which rests on a valid rationale, construct, experimental design, data analysis and so on. Following on from this, all the other Rs can be appraised prospectively and objectively. That is not the case for Robustness because it draws on so many different criteria, some of which might not have been identified. This problem is reflected in the text of this article, which describes different aspects of Robustness in different ways in different paragraphs. So, while totally supporting Robustness as an aspiration, I am not convinced that Robustness would be useful as a 6th R, in practice.

5. I agree with the authors that Wurbel (2017) [ref 16 in the text] does emphasize the importance of robustness, in the context of harm / benefit analysis, and proposes that consideration of the 3Vs (construct, internal and external validity) would help. But, in the end, unlike the other Rs, Robustness is assured only when an experiment has been independently and successfully replicated: i.e., the findings are evidently reproducible. Although this could compromise Reduction, there has already been some discussion on whether independent replication should be a factor to consider when assessing the quality of research. With that in mind, it is worth considering whether Reproducibility or Replicability should be the 6th R instead?

6. Para 5. This is a bit misleading. The assured security of protocols designed using the EDA has been given utmost attention. It was not developed with the intention of serving as an official archive for Registration, as intended in this article, although there is a facility for authorized 'sharing' (e.g. between collaborators). The statement that the EDA allows pre-registration needs to be clarified (bearing in mind the authors' views on points in (5), above).

7. I totally agree with the statements in the Summary, that "a multi-stakeholder discourse is needed to conceptualise the specifics of

	<p>Robustness". In my experience, this has been happening for several years already and has led to important changes in the way research is carried out and reported. However, some factors are controversial and show no sign of being resolved (e.g., whether to abandon the $P < 0.05$ threshold and how to deal with the risk of false positives). Such debates merely strengthen my view that the parameters for robust science are not yet defined adequately and so Robustness is not yet ready to be awarded the status of an independent 'R'.</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Response on reviewers' comments for Manuscript ID bmjos-2018-000048 - "3Rs missing: Animal research without scientific value is unethical"

Dear Editors,

Dear Reviewers,

we like to thank you for the thoughtful and helpful reviews. Please see our comments to the points raised below:

Reviewer(s)' Comments to Author:

Hanno Würbel

Recommendation *

Major Revision

Comments *

In a position paper, Strech and Dirnagl propose extending the 3Rs (replace, reduce, refine) concept for humane experimental techniques by another set of 3Rs (robustness, registration, reporting) to promote the value of animal research. This is certainly a laudable proposal given that research on animals can only be humane

if it generates value for society. No matter how few animals are used and how little harm is imposed on them, if the research is not meaningful or the results lack validity, the animals have been wasted for no good reason. For the exact same reason, I have recently proposed complementing the 3Rs by 3Vs (construct validity, internal validity, external validity) in project evaluation of animal research (Würbel 2017, cited in the manuscript). Strech and Dirnagl aim for a bigger picture, including other dimensions of research integrity beyond rigour and robustness, such as pre-registration and proper reporting.

As much as I support every single one of their recommendations, I have several issues with the conceptualization and branding of their proposal.

- Going for the big picture, the authors write, "research is only valuable if it is innovative, useful, robust, and accessible", suggesting that these four aspects describe the value of research. However, they then note that "innovation" and "usefulness" are hard to gauge and therefore restrict their concept of value to "robustness" and "accessibility". Therefore, from the outset it seems that 2 further Rs are missing, because the authors failed to find a practicable operationalisation for them (or an adequate R-term).

Authors' response: We thank the reviewer for this very good commentary. We agree that the term "innovative" is misleading here. We removed this term in the revised version. See also the comments from the other two reviewers. The term useful should point into the direction of high quality reporting according to, for example, the ARRIVE guidelines. But as it might confuse the reader we removed this term as well.

- They use the term "robustness" to cover all aspects of scientific validity. I am not convinced that robustness is the right term here. Perhaps, "rigour" would be more accurate? For example, the NIH has recently defined "scientific rigour" as "the strict application of the scientific method to ensure robust and unbiased experimental design, methodology, analysis, interpretation and reporting of results. This includes full transparency in reporting experimental details so that others may reproduce and extend the findings". Here, scientific rigour includes all aspects of scientific validity (internal and external validity), as well as reporting and (methods) reproducibility – which would mean that the 6th R "reporting" proposed here would be obsolete as it is already included in the definition of "rigour".

Authors' response: We agree with the reviewer that the term rigour is currently used in an overarching manner. For the purpose of guiding principles, however, it is preferable to have more specific and distinct recommendations. We therefore believe that differentiating robustness principles from reporting principles is preferable for concrete guidance. See also our new text at the end of the manuscript where we defend our additional 3R against the "rhumba of R" criticism:

New text: "Rigour as an R principle is too broad, at least in its current use. Rigour is often used interchangeably with "scientific value" as it comprises robustness, non-selective reporting and could also comprise registration."

The authors argue that “robustness” is included in Articles 21 and 22 of the Declaration of Helsinki; however, while these two articles contain general descriptions of scientific requirements and research protocols, they do not use the term “robustness”.

Authors’ response: We agree that the Articles 21 and 22 are not sufficiently representative for claims on robustness. In the revised version we refer to the widely established ethical framework for clinical research from Emanuel et al. that includes “scientific validity” as one of seven ethical principles making clinical research ethical. It is probably the most cited paper in research ethics (> 2000 citations in Google Scholar). In the explanation of the framework the authors further describe scientific validity as “Use of accepted scientific principles and methods, including statistical techniques, to produce reliable and valid data”.

- Whereas “robustness” (as used by the authors) and accurate “reporting” are necessary aspects of scientific rigour, it is at least questionable whether this is true for “preregistration”. For example, although it makes sense that both exploratory studies and pilot studies need formal approval by the authorities, I am not convinced that it makes sense to ask for preregistration of such studies. This is relevant insofar as adherence to each of the 3Rs is mandatory by law; thus, extending the 3Rs concept by Rs that are not (and should not be) legally mandatory appears problematic. (See also the issue of “diluting the 3Rs principles” below.)

Authors’ response: We agree that the guiding principle of registration can become more or less important. Future developments for animal study registries might focus on confirmatory animal studies for several reasons. But in principle, registration would be relevant for exploratory studies as well. We again refer to the development of ethical principles in clinical research: In 2013, registration became a general principle in the Declaration of Helsinki. Registration, however, became a legal requirement only for subgroups of clinical research. Ethics review of animal studies should assess whether registration is recommendable. Adherence to the registration principle can also mean sufficient justification for why registration is not possible or necessary. By the way, the same logic applies to the replacement principle. Adherence of the replacement principle does not imply that every animal study is in fact replaced.

- The authors write, “regulators, ethics boards, scientists, and funders should add robustness, registration, and reporting to their criteria when planning, licensing, or funding animal experiments.” However, reporting cannot be assessed by funders or regulators but only by journal editors. Since the 3Rs concept is a major pillar in project evaluation and a formal part of the harm-benefit analysis (see below), the inclusion of a criterion that is not applicable to project evaluation is not ideal.

Authors' response: We agree that even for clinical research it is still not standard practice of ethics review to assess the applicants' reporting performance retrospectively or to demand reporting as part of the approval decision. The Declaration of Helsinki, however, made reporting a core principle for the work of ethics committees. In Article 23 it says: "After the end of the study, the researchers must submit a final report to the committee containing a summary of the study's findings and conclusions.". In Article 36, the Declaration highlights the general importance of reporting and of adherence to reporting guidelines: "Researchers, authors, sponsors, editors and publishers all have ethical obligations with regard to the publication and dissemination of the results of research. Researchers have a duty to make publicly available the results of their research on human subjects and are accountable for the completeness and accuracy of their reports. All parties should adhere to accepted guidelines for ethical reporting.". Furthermore, several international funders just recently joined the WHO joint statement and signed the following: "Within 12 months of becoming a signatory of this statement, we each pledge to develop and implement a policy with mandated timeframes for prospective registration and public disclosure of the results of clinical trials that we fund, co-fund, sponsor or support. We each agree to monitor registration and endorse the development of systems to monitor results reporting on an ongoing basis. We agree to share challenges and progress in the monitoring of these policies. We agree that transparency is important and therefore the outputs from the monitoring process will be publicly available".

These examples illustrate how "reporting" could indeed become an assessment criterion within review procedures.

- There have been quite some other proposals already for adding further Rs to the 3Rs concept, none of which has had a lasting impact. It is rather unlikely that the present proposal will succeed in establishing a strong brand. For example, the German Max-Planck Society has added a 4th R for "responsibility", others have proposed "reproducibility" or "rehabilitation" as a 4th R. There have also been proposals for adding 2Rs, e.g. "robustness and reproducibility" or "rigor and reproducibility" (<https://onlinelibrary.wiley.com/doi/full/10.15252/embr.201744428>), and – similar to these authors – others have previously used another set of 3Rs ("rigor, reproducibility and robustness") to describe research quality (https://www.nature.com/polopoly_fs/1.21707!/menu/main/topColumns/topLeftColumn/pdf/543613a.pdf). One author has referred to this use of R-terms (Replication, Reproducibility, Rigor, Robustness) as a "rhumba of R's". Therefore, I am not convinced that yet another Rs concept is a fruitful way forward.

Authors' response: When preparing our paper we also reviewed the other existing proposals for adding Rs to the 3R framework. We were astonished how inconsistent or meaningless these proposals were. This in fact convinced us that we should come up with a consistent and meaningful proposal. In the revised manuscript, we included a section on this type of counter-argument:

NEW TEXT: "Rhumba of Rs"?

“In the previous sections we already commented on potential counter-arguments against the introduction of a complementary set of 3R principles. These counter-arguments addressed the relevance or implementability of registration, robustness, or reporting in a direct way. Another type of counter-argument is more indirect: Does it make sense at all to add new R principles? In discussion with our peers we perceived this counter-argument in at least two ways: First, other papers already and unsuccessfully proposed new Rs such as responsibility, reproducibility, rigour, or relevance. These contributions did not impact on animal research but rather heat up a “rhumba of Rs”. We think that former proposal of new Rs were unsuccessful because they were either circular, too broad, or did not provide direct guidance. Responsibility as an R principle is clearly circular, as it cannot specify how to act responsible. Reproducibility as an R principle does not provide direct guidance. It is a desired characteristic of animal research that strongly depends on robustness and non-selective reporting. Rigour as an R principle is too broad, at least in its current use. Rigour is often used interchangeably with “scientific value” as it comprises robustness, non-selective reporting and could also comprise registration. “

One could of course argue that proposals for expanding the ethical framework of animal research should take the opportunity to get rid of the “R” terminology. We partly agree with this argument and would welcome a revised guidance that overcomes the R terminology and only focuses on the consistent and comprehensive description of guiding principles for animal research. The distinction of guiding principles for animal welfare on the one hand and guiding principles for scientific value, however, would still build a core element of such an ethical framework. Furthermore, registration, robustness, and reporting describe the relevant guiding principles for scientific value in a comprehensive and consistent way. They should show up in such a revised ethical framework even if this framework got rid of the R terminology. To raise awareness across the different stakeholder groups in animal research that “scientific value” is completely ignored in the 3R framework we believe that using the R terminology for another while makes sense.

- I think there are good reasons for leaving the 3Rs intact and using a separate concept for scientific value. First, the 3Rs concept is not just a strong concept that has been adopted almost universally as a guiding principle for humane animal experimentation; it is also a very strong brand. Adding other Rs will likely dilute the concept, and its message, especially if the resulting (4Rs, 5Rs, or 6Rs) concept is not as coherent as the original 3 Rs concept.

Authors’ response: At this point we clearly disagree with the reviewer. The main message of our paper is that the 3R concept is indeed not strong and not coherent at all. In contrast, the 3R concept probably contributed to the fact that animal research currently lacks value and often result into waste. We further argued why our 6R framework is consistent (in contrast to former inconsistent 4R or 5R

proposals) and stronger than the current 3R concept. We added the following text in the manuscript:

NEW TEXT: The second counter-argument against any modification of the 3R framework is based on the assumption that the current 3R framework is a strong concept especially because it is established all over the world. Adding new Rs bear the risk to dilute this strong concept and we might ultimately end up with weaker protection of animal welfare. However, we do not find it plausible to believe that a consistent set of three new guiding principles that all center around the complementary basic principle of “scientific value” will dilute the very distinct basic principle of animal welfare. In contrast, we assume that the incomprehensiveness of the current 3R approach contributed to the fact that animal research often lacks value.

Second, the 3Rs play a critical role in the harm-benefit analysis (HBA) of animal research, which is at the heart of project evaluation and authorisation. The HBA is based on two preconditions that need to be met: the proposed use of animals needs to be suitable for achieving the proposed objectives (scientific rigour), and it needs to be unavoidable (no further potential for replacement, reduction, and refinement). If these preconditions are met, an ethical evaluation (the actual HBA) can be made as to whether the expected benefits (scientific and societal value) outweighs the expected harms. From the conception of the HBA it should become clear that the 3Rs act on the harm side (with the aim to minimize the harms imposed on animals), while scientific rigour acts on the benefit side (with the aim to maximize the benefit of the research). Adding more Rs to the 3Rs concept will likely blur this important distinction.

Authors’ response: The HBA example nicely illustrates the strengths of our 6R framework. As the reviewer indicates the 3R for animal welfare help to operationalize the harm side and the 3R for scientific value operationalize the benefit side. It seems that the reviewer use scientific value and rigour interchangeable as basic principle.

Tom Matheson

Recommendation *

Minor Revision

Comments *

This position paper makes a well-reasoned argument that funding bodies and publishers should consider not only Replacement, Reduction and Refinement in animal experimentation, but also Robustness (of experimental design and statistics), Registration (of experiments prior to work commencing) and more thorough Reporting of outcomes. These are entirely reasonable points to consider and I am very supportive of highlighting them to provoke further detailed discussions about how to implement them effectively. My points below are relatively minor and aimed at tightening up the arguments put forward.

Abstract.

The phrase (line 4) "lacks scientific value" is rather vague. Do you mean 'has no scientific value', or 'has less than the maximum possible scientific value' or something in between? The term (line 7) "safeguard" might better be 'improve' or 'strengthen'. It may be judicious to make it clear whether you are referring to research on any animal (including insects, molluscs etc), or just those vertebrates and cephalopods regulated by e.g. the Home Office in the UK.

Authors' response: Thank you very much for the helpful comments. We revised the abstract to acknowledge your two first comments. We do refer to the third point in the discussion part.

Introduction and main text.

I am not completely convinced by your statement (lines 5-7) that the 3Rs do not consider ("represent") "value for science and society". Although the 3Rs acronym might not explicitly address this, I would argue that the current implementation in grant funding committees does seek to ensure (for example) that studies provide sufficient power, and that there are suitable pathways to tangible benefits ('impact'). Perhaps reconsider how you phrase this section and the strong inference expressed in lines 12-13 that scientific and societal value are not currently "considered".

Authors' response: We revised the text to make clear that "value" might be considered in current ethics review of animal research. But we still find our argument valid that the current 3Rs do not address or do not represent scientific value directly. The 3Rs simply do not provide any guidance on how to increase scientific value. It is great to hear that current ethics review procedures do seek to ensure power or randomization but the current 3Rs did not tell them to do so. It seems as they already follow the guiding principle for robustness. We also mention this aspect in the revised manuscript.

I think the claim (line 7) that “research is only valuable if it is innovative...” is moot. Replication for example might be important, but not innovative.

Authors’ response: We agree, and deleted the term

I think it would be useful to define “registration” (Abstract and Introduction line 11) early on in the article.

Authors’ response: We added some explanation in the introduction but prefer to explain registration in more detail after the introduction.

There is a typo in “biorXiv”: it should be ‘BioRxiv’.

Authors’ response: We corrected the type accordingly.

Where you cite ref 16, it would be useful to describe the ‘3Vs’ that Würbel proposes.

Authors’ response: We briefly explained the 3Vs.

The final section on National Centres (NC3Rs etc) partially addresses some of my points above about current implementations, but it is too late here, and left hanging. You need to make a clearer point. Note that the correct acronym is ‘NC3Rs’.

Authors’ response: Thank you very much for this comment. We revised this part to make the point clearer and we corrected the typo.

Summary

The Summary is largely repetitive of the Abstract, and adds little. I would have liked a more concrete proposal about the next steps.

Authors’ response: We added content to make the proposals more concrete.

NEW TEXT: The Basel Declaration should consider making the normative framework for animal research more comprehensive and coherent. National centres for the 3R should consider revising their branding and explicitly addressing the ethical rationale underlying their recent policies for registration, robustness and reporting. To this end, a multi-stakeholder discourse is needed to i) conceptualize the specifics of robustness, ii) develop platforms to register animal studies, and iii) clarify funding and approval requirements related to results reporting.

S Clare Stanford

Recommendation *

Major Revision

Comments *

The authors are right to point out that the existing 3Rs (Replacement, Refinement, Reduction) all focus on animal welfare. It is also true that research that cannot be trusted (e.g., through flawed design or bias) has no value. In fact, I would go further and say it can be harmful. To help address these problems, this article proposes that there should be three more Rs: Robustness, Registration and Reporting. None of these suggestions is new, as the list of citations confirms. The novelty of this article lies in proposing the three additional Rs, as a mantra. I think the authors are on to something and that it could catch on. However, although I am sure everyone would agree with adding these Rs, as an aspiration, I have several comments on this article and some suggestions for revisions.

1. The Introduction includes the comment that research is ethical only if it is of value to science and society. That is true, but these extra Rs do not address some key elements that contribute to the value of the research. In fact, the text draws attention to the challenging question of how we evaluate the 'value' of research (to science and society) and whether its value meets ethical standards. For instance, the UK has already banned research on cosmetics and tobacco, but debates are developing about the ethics of some other research objectives, especially those directed at resolving problems that are deemed by many people to arise from 'choice' of life-style. The authors rightly state that 'value' is an essential element of Robustness, so perhaps they should either offer suggestions on how to evaluate 'value' to society or define their concept of value more precisely?

Authors' response: We think that the reviewer point to an important issue but that this issue should be discussed in another broader context. The broader question is what research questions for animal research are legitimate at all. This is a fundamental question that should not be evaluated case by case. The 3R framework and our proposal for a 6R framework becomes relevant only for research projects that address legitimate research questions. It might be worthwhile again to refer to frameworks for clinical research. Whether clinical research with cognitively impaired people or basic research with embryonic stem cells is ethical or not cannot be captured with those principles that guide the review of individual clinical studies (informed consent, risk benefit assessment). It needs a more fundamental discourse and general, project-overarching decision.

2. Introduction (para 1, lines 7 – 8). I agree that the usefulness of a research study might not be apparent for many years, but I disagree that the same is the case for 'innovation'. Novelty is a key requirement for a recommendation for grant funding or publication.

Authors' response: We agree that the term "novelty" is more appropriate in this context. With respect to comments from other reviewers we substantially revised this section.

NEW TEXT: Individual research projects are only valuable if they enable a knowledge gain, apply robust study designs, and report their results in a non-selective manner. Whether a research project will contribute to innovation in health care is hard to gauge for several reasons. One reason is that scientific breakthroughs may take years to manifest.

3. The authors should be more specific about their recommendation for Registration and their feasibility / obstacles. For instance, should the registration be openly accessible, or merely held in a secure archive for scrutiny, after the work has been done? If the latter, who would be granted access? Would this be an element of the peer review of the completed work? How would registration affect ownership of intellectual property? How would the information to be registered differ from that already required for a protocol carried out under Good Laboratory Practice (or are we heading for a requirement for all research to be GLP compliant?) To avoid Registration being a flimsy aspiration, I think the authors should express specific views on these points. Perhaps a comment on why many major research-funders are insisting on Open Access for publications but not Registration (yet), would be appropriate too?

Authors' response: We extended our description of animal study registries, referenced papers addressing the spectrum of attitudes towards such registries, and further highlighted recent developments from NC3Rs and Bf3R that provide registration opportunities.

4. My main criticism of this article is that each of the original 3Rs (Replacement, Refinement and Reduction) is a single, simple concept and can be used in a checklist for compliance. The same is true for Registration and Reporting, but Robustness is an odd-ball in the series. The article (e.g., paragraph 4). makes it clear that Robustness is a multifactorial concept, which rests on a valid rationale, construct, experimental design, data analysis and so on. Following on from this, all the other Rs can be appraised prospectively and objectively. That is not the case for Robustness because it draws on so many different criteria, some of which might not have been identified. This problem is reflected in the text of this article, which describes different aspects of Robustness in different ways in different paragraphs. So, while totally supporting Robustness as an aspiration, I am not convinced that Robustness would be useful as a 6th R, in practice.

Authors' response: We further described the Robustness principle to avoid misunderstanding. We introduced robustness as a principle that can be assessed prospectively via criteria for internal, external, and construct validity. See the papers from Würbel and Kimmelman. Based on a systematic review of methodological guidance papers for preclinical research Jonathan Kimmelman and colleagues distinguished 14 criteria for validity that can be assessed for individual animal studies. Hanno Würbel is one of the reviewers for this paper and argued that while robustness and registration can be evaluated or required prospectively, this is less obvious for reporting. See our comments above in this regard.

5. I agree with the authors that Wurbel (2017) [ref 16 in the text] does emphasize the importance of robustness, in the context of harm / benefit analysis, and proposes that consideration of the 3Vs (construct, internal and external validity) would help. But, in the end, unlike the other Rs, Robustness is assured only when an experiment has been independently and successfully replicated: i.e., the findings are evidently reproducible. Although this could compromise Reduction, there has already been some discussion on whether independent replication should be a factor to consider when assessing the quality of research. With that in mind, it is worth considering whether Reproducibility or Replicability should be the 6th R instead?

Authors' response: We included a section on the question why our R terminologies are preferable over other Rs that have been suggested such as reproducibility, rigour or responsibility.

NEW TEXT: "Rhumba of Rs"?

In the previous sections we already commented on potential counter-arguments against the introduction of a complementary set of 3R principles. These counter-arguments addressed the relevance or implementability of registration, robustness, or reporting in a direct way. Another type of counter-argument is more indirect: Does it make sense at all to add new R principles? In discussion with our

peers we perceived this counter-argument in at least two ways: First, other papers already and unsuccessfully proposed new Rs such as responsibility, reproducibility, rigour, or relevance. These contributions did not impact on animal research but rather heat up a “rhumba of Rs”. We think that former proposal of new Rs were unsuccessful because they were either circular, too broad, or did not provide direct guidance. Responsibility as an R principle is clearly circular, as it cannot specify how to act responsible. Reproducibility as an R principle does not provide direct guidance. It is a desired characteristic of animal research that strongly depends on robustness and non-selective reporting. Rigour as an R principle is too broad, at least in its current use. Rigour is often used interchangeably with “scientific value” as it comprises robustness, non-selective reporting and could also comprise registration.

6. Para 5. This is a bit misleading. The assured security of protocols designed using the EDA has been given utmost attention. It was not developed with the intention of serving as an official archive for Registration, as intended in this article, although there is a facility for authorized ‘sharing’ (e.g. between collaborators). The statement that the EDA allows pre-registration needs to be clarified (bearing in mind the authors’ views on points in (5), above).

Authors’ response: We specified our comment on the EDA.

NEW TEXT: The new NC3Rs Experimental Design Assistant (EDA) not only supports the development of robust study protocols but also allows to time-stamp the resulting protocols. With the option to make such time-stamped protocols publicly available, the EDA facilitates pre-registration of protocols on a voluntary basis

7. I totally agree with the statements in the Summary, that “a multi-stakeholder discourse is needed to conceptualise the specifics of Robustness”. In my experience, this has been happening for several years already and has led to important changes in the way research is carried out and reported. However, some factors are controversial and show no sign of being resolved (e.g., whether to abandon the $P < 0.05$ threshold and how to deal with the risk of false positives). Such debates merely strengthen my view that the parameters for robust science are not yet defined adequately and so Robustness is not yet ready to be awarded the status of an independent ‘R’.

Authors’ response: As described in our response on comment 4 we believe that robustness can already be judged for several important criteria. We agree that the concept of robustness should and most probably will be further developed. But this is the same with replacement, refinement and reduction. The important question is whether an established or newly proposed guiding principle can positively impact on the ethical practice of animal research. Implementing the robustness principle in the review of individual animal studies and in the overall governance will have such a positive impact.

VERSION 2 - REVIEW

REVIEWER 1	Hanno Würbel
REVIEW RETURNED	31-01-18

GENERAL COMMENTS	<p>Recommendation * Accept</p> <p>The authors have adequately addressed all of my previous comments, and I am happy with the changes they made in the manuscript. There is one last point, which I did not mention in my previous review, but which I find important to consider. In their response letter, the authors write that "the 3R concept is indeed not strong and not coherent at all. In contrast, the 3R concept probably contributed to the fact that animal research currently lacks value and often result into waste."</p> <p>Furthermore, they argue that awareness needs to be raised "across the different stakeholder groups in animal research that "scientific value" is completely ignored in the 3R framework". Here, I think they confuse the 3Rs principles as published by Russell and Burch with common, but inaccurate, interpretations of the 3Rs. Russell and Burch were advocates of robust research. They considered rigorous and robust research as first principles, and explicitly took the stance that the 3Rs can only be applied to rigorous and robust research.</p> <p>Moreover, they discuss many cases of how the 3Rs actually help improve the quality of the research. It is true that there might seem to be conflicts between the 3Rs concept and good science. For example, it is often argued, that the second R 'reduce' promotes underpowered study designs. However, this could only happen if the 3Rs were taken as first principles, regardless of scientific validity. Nothing could be further away from the 3Rs principles as proposed by Russell and Burch.</p>
-------------------------	--

REVIEWER 2	Tom Matheson
REVIEW RETURNED	31-01-18

GENERAL COMMENTS	<p>Recommendation * Minor Revision</p> <p>The authors have addressed my concerns. I think the paper provides a useful focus for further discussion, regardless of whether one agrees with the position put forward or not. Some of the language in the revisions could perhaps be edited for English style. For example, 'incomprehensiveness' in the last sentence would be better phrased as 'relatively narrow focus of'.</p>
-------------------------	---

REVIEWER 3	S Clare Stanford
REVIEW RETURNED	31-01-18

GENERAL COMMENTS	<p>Recommendation * Accept</p> <p>In this revised version of their article, the authors have reorganized much of the material and strengthened the focus of their main points. They have also added some helpful text to clarify their lines of thought. I do not dispute their view that each of the additional Rs has merit and processes are already in train to ensure their continual development and promotion. Nevertheless, I still hold the view that mandatory addition of three more Rs is premature. This is because the specifications and practicalities for the new Rs (Preregistration, Reporting compliance and, especially, prospective evaluation of Robustness) are currently too vague. I acknowledge that this is a personal view and so I recommend publication on the basis that readers should be allowed to make up their own mind on whether or not adoption of 6Rs, as a checklist, would be helpful. Further points, dealing with the authors' responses to my comments on the original version of this manuscript, are: General 1. Re: point [1] Here, I asked for clarification of the authors' concept of 'value' and how that might be assessed. Instead, their response redirects the question to whether or</p>
-------------------------	---

not a program of research tackles a 'legitimate' research question. That is a different point, which remains unanswered by the authors, other than to draw parallels with clinical studies. 2. In my point [3], I suggested that the article should be more specific about their recommendation for mandatory 'Registration', taking account of the lack of any agreed details on what precisely should be registered and the many foreseeable problems arising from such a process. The revised text deals with this to some extent in that it provides more information on existing platforms for registration, but it does not address the concerns I raised in my original comment. The NC3Rs EDA does indeed offer an excellent opportunity to register details of the scientific objectives, experimental protocol and data analysis and it is good to have its potential as a platform for registration highlighted in this version of the article. However, there remain points that are not covered by the EDA (e.g. procedural quality assurance and details of assays for deriving the outcome measures) and yet these are fundamental for GLP (and regulatory approval of clinical applications in some jurisdictions). There is a clear need for agreement on what aspects of any study should be preregistered and how compliance would be monitored, by whom and when, and the implications for IP. In the revised article, the authors acknowledge these problems, but offer no recommendations. 3. Re: my points [4] and [5,] I suggested that it is difficult to assess Robustness prospectively, because it is a multifactorial concept and some of the key factors might not be known when the study is designed. In their response, the authors suggest that Robustness is synonymous with validity. I agree that it is possible to confirm robustness for certain aspects of any study (blinding, randomisation and so on). Indeed, appraisal of the research objectives and the study design already underpin the harm / benefit analysis, which is an essential component of the process of assessing a program research for licensed approval in the UK. Also, many Journals (including the BMJ) now pay far more attention to study design when assessing articles for publication. However, that does not resolve my concern that 'Robustness' can be confirmed only retrospectively – particularly in respect of assessing the external validity of animal

	<p>studies, which is the point at which many fail. In their response to my point [7], the authors state that Robustness “can already be judged for several important criteria”. That is true but, in my view, ‘several criteria’ is not adequate to qualify for an ‘R’. For that reason, I believe that Robustness is a concept that is far more difficult to define and confirm, prospectively, than Replacement, Refinement and Reduction. 4. The Section on the Rhumba of the 3Rs serves well as a rebuttal of Wurbel’s comment on the original manuscript, but merely reinforces the controversy on whether there should be any additional Rs and the lack of consensus on what they should be. 5. Overall, I still believe that addition of three new Rs would not be as helpful as implied in this article, not least because essential practicalities have not been addressed. Who would audit the extra Rs? Also, at what stage of the research process would they be regarded as mandatory: funding application (how would that work for Registration and Reporting?) / ethics and welfare assessment (ditto?) / regulatory application / publication peer review)? I agree with Hanno Wurbel’s comments that, in addition to the cost of rebranding the 3Rs, the resource implications make it likely that most of the bodies involved in these various steps of the process would not want to be responsible for an audit of 6Rs compliance and current legislation does not require them to do so.</p>
--	--

VERSION 2 – AUTHOR RESPONSE

Responses to Reviewer Comments

Tom Matheson

Recommendation *

Minor Revision

Comments *

Please leave your comments for the authors below

The authors have addressed my concerns. I think the paper provides a useful focus for further discussion, regardless of whether one agrees with the position put forward or not. Some of the language in the revisions could perhaps be edited for English style. For example, 'incomprehensiveness' in the last sentence would be better phrased as 'relatively narrow focus of'.

Response: We are happy to have addressed the reviewer's concerns and also tried to further improve the English style including the mentioned example.

Hanno Würbel

Recommendation *

Accept

Comments *

Please leave your comments for the authors below

The authors have adequately addressed all of my previous comments, and I am happy with the changes they made in the manuscript. There is one last point, which I did not mention in my previous review, but which I find important to consider. In their response letter, the authors write that "the 3R concept is indeed not strong and not coherent at all. In contrast, the 3R concept probably contributed to the fact that animal research currently lacks value and often result into waste." Furthermore, they argue that awareness needs to be raised "across the different stakeholder groups in animal research that "scientific value" is completely ignored in the 3R framework". Here, I think they confuse the 3Rs principles as published by Russell and Burch with common, but inaccurate, interpretations of the 3Rs. Russell and Burch were advocates of robust research. They considered rigorous and robust research as first principles, and explicitly took the stance that the 3Rs can only be applied to rigorous and robust research. Moreover, they discuss many cases of how the 3Rs actually help improve the quality of the research. It is true that there might seem to be conflicts between the 3Rs concept and good science. For example, it is often argued, that the second R 'reduce' promotes underpowered study designs. However,

this could only happen if the 3Rs were taken as first principles, regardless of scientific validity. Nothing could be further away from the 3Rs principles as proposed by Russell and Burch.

Response: We thank Hanno Würbel for his comment that Russell and Burch were advocates of robust research. We agree and added a remark that they explicitly mention A) the need of statistics to determine sample sizes that are as small as possible and B) randomization as a measure to reduce the threat of variance. We did not find a mention on “blinding” and we did not find any comment on the problems stemming from selective reporting and from the lack of transparency regarding past and ongoing studies.

S. Clare Stanford

Recommendation *

Accept

Comments *

Please leave your comments for the authors below

In this revised version of their article, the authors have reorganized much of the material and strengthened the focus of their main points. They have also added some helpful text to clarify their lines of thought. I do not dispute their view that each of the additional Rs has merit and processes are already in train to ensure their continual development and promotion. Nevertheless, I still hold the view that mandatory addition of three more Rs is premature. This is because the specifications and practicalities for the new Rs (Preregistration, Reporting compliance and, especially, prospective evaluation of Robustness) and are currently too vague. I acknowledge that this is a personal view and so I recommend publication on the basis that readers should be allowed to make up their own mind on whether or not adoption of 6Rs, as a checklist, would be helpful. Further points, dealing with the authors’ responses to my comments on the original version of this manuscript, are:

General 1. Re: point [1] Here, I asked for clarification of the authors’ concept of ‘value’ and how that might be assessed. Instead, their response redirects the question to whether or not a program of research tackles a ‘legitimate’ research question. That is a different point, which remains unanswered by the authors, other than to draw parallels with clinical studies.

Response: We did not want to “redirect” the value-question to a question about legitimacy. We used this term synonymously. Our analogy from clinical research should only demonstrate that there are questions about the value/legitimacy of animal research that need to be discussed independently of individual animal studies. Our paper only addresses those areas of animal studies that society accepts as being of “sufficient value” in general. Neither the standard 3Rs nor the additional 3Rs help with clarifying whether animal research on “life-style” issues is legitimate/of sufficient value or not.

2. In my point [3], I suggested that the article should be more specific about their recommendation for mandatory ‘Registration’, taking account of the lack of any agreed details on what precisely should be registered and the many foreseeable problems arising from such a process. The revised text deals with this to some extent in that it provides more information on existing platforms for registration, but it does not address the concerns I raised in my original comment. The NC3Rs EDA does indeed offer an excellent opportunity to register details of the scientific objectives, experimental protocol and data analysis and it is good to have its potential as a platform for registration highlighted in this version of the article. However, there remain points that are not covered by the EDA (e.g. procedural quality assurance and details of assays for deriving the outcome measures) and yet these are fundamental for GLP (and regulatory approval of clinical applications in some jurisdictions). There is a clear need for agreement on what aspects of any study should be preregistered and how compliance would be monitored, by whom and when, and the implications for IP. In the revised article, the authors acknowledge these problems, but offer no recommendations.

Response: We fully agree that detailing mandatory information for registration is important. We mention this in the revised summary section. But trying to provide these information would go beyond the scope of this paper. We believe that for didactic reasons our paper should not even try to provide a satisfying set of recommendations on how to translate the 6R framework into practice. We are very keen to follow-up our ideas and to develop more specific recommendations together with the mentioned stakeholder groups. But we believe that this paper should stop by outlining the core idea of a second, complementary basic principle for ethical animal research (“scientific value”) and the additional idea of how to operationalize this basic principle via three guiding principles.

3. Re: my points [4] and [5,] I suggested that it is difficult to assess Robustness prospectively, because it is a multifactorial concept and some of the key factors might not be known when the study is designed. In their response, the authors suggest that Robustness is synonymous with validity. I agree that it is possible to confirm robustness for certain aspects of any study (blinding, randomisation and so on). Indeed, appraisal of the research objectives and the study design already underpin the harm / benefit analysis, which is an essential component of the process of assessing a program research for licensed approval in the UK. Also, many Journals (including the BMJ) now pay far more attention to study design when assessing articles for publication. However, that does not resolve my concern that ‘Robustness’

can be confirmed only retrospectively – particularly in respect of assessing the external validity of animal studies, which is the point at which many fail. In their response to my point [7], the authors state that Robustness “can already be judged for several important criteria”. That is true but, in my view, ‘several criteria’ is not adequate to qualify for an ‘R’. For that reason, I believe that Robustness is a concept that is far more difficult to define and confirm, prospectively, than Replacement, Refinement and Reduction.

Response: We would argue that external validity could be judged prospectively. We cite Hanno Würbel who outlined this in the paper “More than 3Rs: the importance of scientific validity for harm-benefit analysis of animal research”: “[...]assessment of external validity should be based on evidence for experimental design features that enhance, or facilitate inference about, the reproducibility and generalizability of the expected results¹. This includes splitting experiments into multiple independent replicates (batches)²³, introducing systematic variation (heterogenization) of relevant variables (for example, species/strains of animals, housing conditions, tests, etc.)^{15,24,25}, or implementing multi-center study designs²⁶.”

4. The Section on the Rhumba of the 3Rs serves well as a rebuttal of Würbel’s comment on the original manuscript, but merely reinforces the controversy on whether there should be any additional Rs and the lack of consensus on what they should be. 5. Overall, I still believe that addition of three new Rs would not be as helpful as implied in this article, not least because essential practicalities have not been addressed. Who would audit the extra Rs? Also, at what stage of the research process would they be regarded as mandatory: funding application (how would that work for Registration and Reporting?) / ethics and welfare assessment (ditto?) / regulatory application / publication peer review)? I agree with Hanno Würbel’s comments that, in addition to the cost of rebranding the 3Rs, the resource implications make it likely that most of the bodies involved in these various steps of the process would not want to be responsible for an audit of 6Rs compliance and current legislation does not require them to do so.

Response: See also our comments above on the scope of the paper. We extended and revised the summary section to further emphasize the need of further multi-stakeholder discussion and to further specify the content of these discussions including the issue of legislation.

BMJ Open Science

BMJ Open Science is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open Science is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://openscience.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmj@bmj.com

3Rs missing: Animal research without scientific value is unethical

Daniel Strech 1,2*, Ulrich Dirnagl 1-3

*corresponding: daniel.strech@charite.de

1 QUEST Center for Transforming Biomedical Research, Berlin Institute of Health (BIH), Berlin, Germany

2 Charité - Universitätsmedizin Berlin, Berlin, Germany

3 NeuroCure Clinical Research Center, Charité - Universitätsmedizin Berlin, Berlin, Germany

Abstract

The current, widely established 3R framework for the ethical use of animals in research consists of three guiding principles, i.e., **R**eplacement, **R**eduction, and **R**efinement, all aiming to safeguard the overarching ethical principle of animal welfare. However, animal welfare alone does not suffice to make animal research ethical if the research lacks scientific value. Animal research lacks scientific value if it is not sufficiently robust or if the results are only selectively reported. Against this background, we argue that three guiding principles are missing, i.e., **R**obustness, **R**egistration, and **R**eporting, all of which aim to safeguard the value of animal research. To establish a new 6R framework, we need a multi-stakeholder discourse to conceptualize the specifics of robustness, to develop platforms to register animal studies and to incentivize the non-selective reporting of results.

Intro

Framed by Russel and Burch more than 60 years ago, the 3Rs (**R**eplacement, **R**eduction and **R**efinement) have become the guiding principles for the ethical use of animals in research ¹. Although universally accepted, there is an ongoing discourse on their improvement, uptake and implementation ². Here, we argue that with their current focus on animal welfare, the 3Rs lack an important ethical dimension. Research on animals is only ethical if it generates value for science and society, a dimension that is not represented by the 3Rs. Research is only valuable if it is innovative, useful, robust, and accessible. Innovation and usefulness may not be immediately apparent and are hard to gauge, as scientific breakthroughs may take years to manifest. Robustness, on the other hand, results from research with high scientific rigour. Research is accessible if others know about it (registration) and if its results are faithfully reported. We posit that while the current 3Rs are important for upholding animal welfare, the dimension of scientific and societal value needs to be considered when planning or conducting animal research. We therefore propose the

addition of three additional Rs, i.e., **Robustness**, **Registration**, and **Reporting**, to the guiding principles for the ethical use of animals in research (see figure 1). Adopting 6Rs will change the way animal experiments are appraised and improve the welfare of animals.

→ Figure 1

Why do we need to complement the 3R framework now?

Over the past five years, several empirical studies and expert analyses have demonstrated that four challenges endanger the value of animal research. First, animal research often lacks measures to reduce validity threats, such as biases or a lack of statistical power^{3,4}. Second, animal research faces a substantial publication bias, i.e., null and negative results often end up in the file drawer^{5,6}. Third, animal studies are often not reproducible^{7,8} and, fourth, do not play the role they are claimed to play in justifying early human research^{9,10}. Each of these threats reduces the value of the research results, potentially leading to the unnecessary use and suffering of experimental animals.

Why robustness, registration, and reporting?

Our core argument is that the current 3R principles for animal research, despite their importance, are biased towards animal welfare but lack principles safeguarding the value of animal research. Furthermore, each of the additional 3R principles is important in itself and not replaceable by the other two. Animal studies, for example, can be robust but reported in a biased or otherwise inappropriate way. Alternatively, they can be appropriately reported but not robust. Both scenarios compromise the study's value. In times where approximately 50% of animal studies are not reported, only the pre-registration of animal study protocols allows the identification of biased, delayed, or un-reported results. Finally, ethics frameworks for human research already address all three value principles for the same moral reasons. The Declaration of Helsinki, for example, includes robustness (articles 21 and 22), registration (article 35), and reporting (article 36) as obligatory principles¹¹.

How do we implement the new 3Rs in current practices?

The reporting principle is relatively easy to implement. Preprint servers (such as biorXiv) or Open Access journals with post-publication review allow the immediate and accessible reporting of all types of research results, including null and negative results. Adherence to reporting guidelines, such as ARRIVE¹², aims to improve the evaluation and utilization of study results. Dedicated tools for implementing the registration principle in animal research equivalent to registries for human studies (such as ClinicalTrials.gov) are currently under development. There is an urgent need for platforms that allow swift protocol registration with an embargoing option for a certain time period¹³. The robustness principle is more difficult to implement: How can we gauge scientific rigor? More specifically: When is sample

size calculation or blinded outcome assessment necessary? How can the external and construct validity of individual studies be improved? Recent expert proposals to better distinguish between exploratory and confirmatory study designs in animal research have provided preliminary answers^{14 15}. Recently, initial guidance on how to implement a more systematic assessment of animal study robustness in standard review procedures was published¹⁶. Such distinctions require complex judgements. Ethics review boards for animal studies, however, already require complex judgements regarding the welfare principles, and in many jurisdictions, already consider a study's robustness.

National centres for the 3Rs (NC3R, Bf3R) already promote the new three value principles in several ways. The revised NC3R guidelines for primate research, for example, explicitly require robustness and reporting¹⁷. The new NC3R Experimental Design Assistant (EDA) not only supports the development of robust designs but also allows pre-registration¹⁸.

Summary

Animal research is ethical only when it is of scientific and social value. The past years have demonstrated that the value of animal research and thus its capacity to improve human health are threatened by a lack of robustness and biased or un-reported results. Three ethical principles (**R**obustness, **R**egistration, and **R**eporting) help to safeguard the value of animal research. The current, widely established ethical framework for animal research (3Rs = **R**eplacement, **R**eduction, and **R**efinement) misses the value dimension by solely focusing on the equally important animal welfare dimension. We recommend complementing the current 3R framework (for animal welfare) with the second set of 3Rs (for research value). Regulators, ethics boards, scientists, and funders should add robustness, registration, and reporting to their criteria when planning, licensing, or funding animal experiments. To this end, a multi-stakeholder discourse is needed to conceptualize the specifics of robustness and rigour and to develop platforms to register animal studies following the example of the clinical study community.

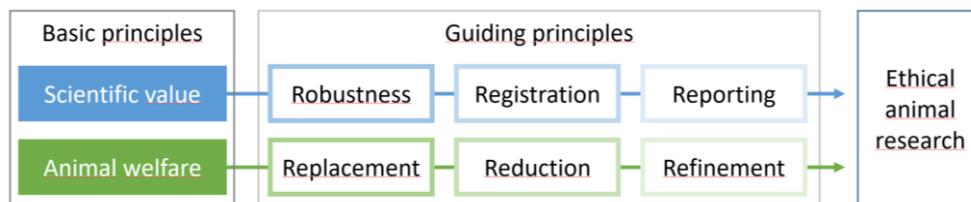
References

1. Russell WMS, Burch RL. *The Principles of Humane Experimental Technique*. London: Methuen 1959.
2. RSPCS, LASA. *Guiding principles on good practice for Animal Welfare and Ethical Review Bodies*. A report by the RSPCA Research Animals Department and LASA Education, Training and Ethics Section; http://www.lasa.co.uk/PDF/AWERB_Guiding_Principles_2015_final.pdf, 2015.
3. Vogt L, Reichlin TS, Nathues C, et al. Authorization of Animal Experiments Is Based on Confidence Rather than Evidence of Scientific Rigor. *PLoS biology* 2016;14(12):e2000598. doi: 10.1371/journal.pbio.2000598
4. Macleod MR, Lawson McLean A, Kyriakopoulou A, et al. Risk of Bias in Reports of In Vivo Research: A Focus for Improvement. *PLoS biology* 2015;13(10):e1002273. doi: 10.1371/journal.pbio.1002273
5. Sena ES, van der Worp HB, Bath PM, et al. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. *PLoS biology* 2010;8(3):e1000344. doi: 10.1371/journal.pbio.1000344
6. Mueller KF, Briel M, Strehl D, et al. Dissemination bias in systematic reviews of animal research: a systematic review. *PLoS One* 2014;9(12):e116016. doi: 10.1371/journal.pone.0116016

7. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. *Nature* 2012;483(7391):531-3. doi: 10.1038/483531a
8. Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 2011;10(9):712. doi: 10.1038/nrd3439-c1
9. Wieschowski S, Chin WWL, Federico C, et al. Preclinical efficacy studies in investigator brochures: Do they enable risk-benefit assessment? *PLoS biology* 2018;16(4):e2004879. doi: 10.1371/journal.pbio.2004879
10. Kimmelman J, Federico C. Consider drug efficacy before first-in-human trials. *Nature* 2017;542(7639):25-27. doi: 10.1038/542025a
11. World Medical Association. Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects, Fortaleza, 2013.
12. Kilkenny C, Browne WJ, Cuthill IC, et al. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS biology* 2010;8(6):e1000412. doi: 10.1371/journal.pbio.1000412
13. Wieschowski S, Silva DS, Strech D. Animal Study Registries: Results from a Stakeholder Analysis on Potential Strengths, Weaknesses, Facilitators, and Barriers. *PLoS biology* 2016;14(11):e2000391. doi: 10.1371/journal.pbio.2000391
14. Kimmelman J, Mogil JS, Dirnagl U. Distinguishing between exploratory and confirmatory preclinical research will improve translation. *PLoS biology* 2014;12(5):e1001863. doi: 10.1371/journal.pbio.1001863
15. Mogil JS, Macleod MR. No publication without confirmation. *Nature* 2017;542(7642):409-11. doi: 10.1038/542409a
16. Wurbel H. More than 3Rs: the importance of scientific validity for harm-benefit analysis of animal research. *Lab Anim (NY)* 2017;46(4):164-66. doi: 10.1038/labani.1220 [published Online First: 2017/03/23]
17. NC3R. Non-human primate accommodation, care and use; <https://www.nc3rs.org.uk/sites/default/files/documents/Guidelines/NC3Rs%20guidelines%20-%20Non-human%20primate%20accomodation%20care%20and%20use.pdf>. 2017
18. du Sert NP, Bamsey I, Bate ST, et al. The Experimental Design Assistant. *Nat Methods* 2017;14(11):1024-25. doi: 10.1038/nmeth.4462 [published Online First: 2017/09/30]

Figure

Figure 1: Two basic principles for animal research ethics translate into six practice-guiding principles (6R)



3Rs missing: Animal research without scientific value is unethical

Daniel Strech 1,2*, Ulrich Dirnagl 1-3

*corresponding: daniel.strech@charite.de

1 QUEST Center for Transforming Biomedical Research, Berlin Institute of Health (BIH), Berlin, Germany

2 Charité - Universitätsmedizin Berlin, Berlin, Germany

3 NeuroCure Clinical Research Center, Charité - Universitätsmedizin Berlin, Berlin, Germany

Abstract

The current, widely established 3R framework for the ethical use of animals in research consists of three guiding principles, i.e., **R**eplacement, **R**eduction, and **R**efinement, all aiming to safeguard the overarching ethical principle of animal welfare. However, animal welfare alone does not suffice to make animal research ethical if the research has not sufficient scientific value. The scientific value of animal studies strongly decreases if it is not sufficiently robust or if the results are selectively reported. Against this background, we argue that three guiding principles are missing, i.e., **R**obustness, **R**egistration, and **R**eporting, all of which aim to increase the value of animal research. To establish a new 6R framework, we need a multi-stakeholder discourse to conceptualize the specifics of robustness, to develop platforms for the registration of animal studies and to incentivize the non-selective reporting of results.

Intro

Framed by William Russel and Rex Burch more than 60 years ago, the 3Rs (**R**eplacement, **R**eduction and **R**efinement) have become the guiding principles for the ethical use of animals in research ¹. Although universally accepted, there is an ongoing discourse on their improvement, uptake and implementation ². Here, we argue that with their current focus on animal welfare, the 3Rs lack an important ethical dimension. Research on animals is only ethical if it generates value for science and society, a dimension that is not represented by the current 3Rs.

Individual research projects are only valuable if they enable a knowledge gain, apply robust study designs, and report their results in a non-selective manner. Whether a research project will ultimately contribute to innovation in health care is hard to gauge for several reasons. One reason is that scientific breakthroughs may take years to manifest. Robustness,

on the other hand, can be judged on the research project level. If we want to better understand what research questions are still insufficiently addressed we need individual projects to be accessible via animal study registries open to the public. Furthermore, only if protocols are prospectively registered we are able to identify selective reporting of study results.

We posit that while the current 3Rs are important for upholding animal welfare, the dimension of scientific value needs to be considered when planning, reviewing, and conducting animal research. We therefore propose the addition of three additional Rs, i.e., **Robustness, Registration, and Reporting**, to the guiding principles for the ethical use of animals in research (see figure 1).

Figure 1

Why do we need to complement the 3R framework now?

Over the past five years, several empirical studies and expert analyses have demonstrated that three challenges endanger the value of animal research. First, animal research often lacks measures to reduce validity threats such as biases or a lack of statistical power³⁴. Second, animal research faces a substantial publication bias, i.e., null and negative results often end up in the file drawer⁵⁶. Third, publication of results often lack important information that are needed for a critical appraisal (e.g. information on study design or attrition of animals)⁷⁸. These challenges negatively affect the reproducibility of animal studies⁹¹⁰ and the relevance of animal studies in justifying early human research¹¹¹². In summary, these threats reduce the value of the research results, potentially leading to inefficient allocation of public funds, to ill-advised clinical research, and to the unnecessary use and suffering of experimental animals,

Why robustness, registration, and reporting?

Our core argument is that the current 3R principles for animal research, despite their importance, are biased towards animal welfare but lack principles that aim to increase the value of animal research. Furthermore, each of the additional 3R principles (robustness, registration, reporting) is important in itself and not replaceable by the other two. Animal studies, for example, can be robust but reported in a biased or otherwise inappropriate way. Alternatively, they can be appropriately reported but not robust. Both scenarios compromise the study's value. In times where approximately 50% of animal studies are not reported¹³, only the pre-registration of animal study protocols allows the identification of biased, delayed, or un-reported results. Finally, ethics frameworks for human research already address all three value principles for the same moral reasons. The Declaration of Helsinki, for example, includes registration (article 35) and reporting (article 36) as obligatory principles

¹⁴. The widely acknowledged framework for clinical research “What makes clinical research ethical” from Ezekiel Emanuel et al. highlights robustness (scientific validity) as one of the basic ethical principles ¹⁵.

How do we implement the new 3Rs in current practices?

The reporting principle is relatively easy to implement. Beside standard peer-review journals new publication formats allow accessible reporting of all types of research results, including null and negative results, such as preprint servers (e.g. bioRxiv), Open Access journals (e.g. BMJ Open Science, PLoS One), or journals with post-publication review (e.g. f1000research). Adherence to reporting guidelines, such as ARRIVE ⁷, further aims to improve the evaluation and utilization of study results. Several leading research funders such as the Wellcome Trust, the Horizon 2020 program, or the Bill and Melinda Gates Foundation just recently signed the WHO Joint Statement and thus indicated to make reporting requirements a part of funding decisions for clinical trials ¹⁶. Similarly, ethics review of individual animal studies could implement a requirement for timely and non-selective results reporting and evaluate compliance.

Dedicated tools for implementing the registration principle in animal research equivalent to registries for human studies (such as ClinicalTrials.gov) have already been launched (www.preclinicaltrials.eu) or are currently under development (see below). These platforms allow swift protocol registration with an embargoing option for a certain time period. The registration principle will increase the value of research but how will it affect the efficiency of animal research? In a recent study experts from all relevant stakeholder groups in animal research expressed their attitudes on potential strengths and weaknesses of animal study registries ¹⁷. Some highlighted their concerns that animal study registration might aggravate administrative burdens and the theft of ideas. Others emphasized the opposite viewpoint that improved transparency via such registries might ultimately make animal research more efficient.

The robustness principle is more difficult to implement: How can we gauge robustness of individual animal studies? More specifically: When is sample size calculation or blinded outcome assessment necessary? How can the external and construct validity of individual studies be improved? Recent expert proposals to better distinguish between exploratory and confirmatory study designs in animal research have provided preliminary answers ^{18 19}. Initial guidance on how to implement a more systematic assessment of animal study robustness in standard review procedures was recently published by Hanno Würbel ²⁰. Würbel distinguishes three dimensions of validity (internal, external and construct validity, 3Vs) and recommends assessing each dimension within the harm-benefit analysis (HBA) for individual animal studies. With this proposal he is in line with recent guidance from Jonathan Kimmelman on how to assess the validity of animal studies within approval procedures for phase I/II clinical trials ²¹. Assessing robustness of individual studies requires complex judgements. Ethics review boards for animal studies, however, already require complex

judgements regarding the welfare principles, and in many jurisdictions, already consider a study's robustness.

In line with our recommendation to add guiding principles for scientific value to the ethical framework for animal research are recent activities from national centres for the 3Rs such as the UK NC3Rs or the German Bf3R. Both already promote the new 3R principles for scientific value in several ways. The revised NC3Rs guidelines for primate research, for example, explicitly require robustness and reporting²². The new NC3Rs Experimental Design Assistant (EDA) not only supports the development of robust study protocols but also allows to time-stamp the resulting protocols. With the option to make such time-stamped protocols publicly available, the EDA facilitates pre-registration of protocols on a voluntary basis²³. At a 3R symposium in Berlin in November 2018, the director of the German Center for the Protection of Animals in Research (Bf3R), Gilbert Schönfelder, announced the launch of an Animal Study Registry in early 2019. We very much welcome these recent developments but want to highlight that they do not follow logically from any of the three animal welfare principles. They make sense only when considering scientific value as a complementary set of ethical principles.

“Rhumba of Rs”?

In the previous sections we already commented on potential counter-arguments against the introduction of a complementary set of 3R principles. These counter-arguments addressed the relevance or implementability of registration, robustness, or reporting in a direct way. Another type of counter-argument is more indirect: Does it make sense at all to add new R principles? At least two arguments were raised: First, other papers already and unsuccessfully proposed new Rs such as responsibility, reproducibility, or rigour. These contributions did not impact on animal research but rather heat up a “rhumba of Rs”²⁴. We think that former proposals of new Rs were unsuccessful because they were either circular, too broad, or did not provide direct guidance. Responsibility as an R principle is clearly circular, as it cannot specify how to act responsibly. Reproducibility as an R principle does not provide direct guidance. It is a desired characteristic of animal research that strongly depends on robustness and non-selective reporting. Rigour as an R principle is too broad, at least in its current use. Rigour is often used interchangeably with “scientific value” as it comprises robustness, non-selective reporting and could also comprise registration.

The second counter-argument against any modification of the 3R framework is based on the assumption that the current 3R framework is a strong concept especially because it is established all over the world. Adding new Rs bears the risk to dilute this widely accepted concept, ultimately leading to a weaker protection of animal welfare. However, we do not find it plausible to believe that a consistent set of three new guiding principles that all center around the complementary basic principle of “scientific value” will dilute the very distinct

basic principle of animal welfare. In contrast, we posit that the incomprehensiveness of the current 3R approach contributed to the fact that animal research often lacks value.

Summary

Animal research is ethical only when it is of scientific and social value. The past years have demonstrated that the value of animal research and thus its capacity to improve human health are threatened by a lack of robustness and biased or un-reported results. Three ethical principles (**R**obustness, **R**egistration, and **R**eporting) help to safeguard the value of animal research. The current, widely established ethical framework for animal research (3Rs = **R**eplacement, **R**eduction, and **R**efinement) misses this value dimension by solely focusing on the equally important animal welfare dimension. We recommend complementing the current 3R framework (for animal welfare) with the second set of 3Rs (for research value). Regulators, ethics boards, scientists, and funders should add robustness, registration, and reporting to their criteria when planning, licensing, or funding animal experiments. Guidances such as the Basel Declaration should consider making the normative framework for animal research more comprehensive and coherent²⁵. National centres for the 3R should consider revising their branding and explicitly addressing the ethical rationale underlying their recent policies for registration, robustness and reporting. To this end, a multi-stakeholder discourse is needed to i) conceptualize the specifics of robustness, ii) develop platforms to register animal studies, and iii) clarify funding and approval requirements related to results reporting.

References

1. Russell WMS, Burch RL. *The Principles of Humane Experimental Technique*. London: Methuen 1959.
2. RSPCS, LASA. *Guiding principles on good practice for Animal Welfare and Ethical Review Bodies*. A report by the RSPCA Research Animals Department and LASA Education, Training and Ethics Section; http://www.lasa.co.uk/PDF/AWERB_Guiding_Principles_2015_final.pdf, 2015.
3. Vogt L, Reichlin TS, Nathues C, et al. Authorization of Animal Experiments Is Based on Confidence Rather than Evidence of Scientific Rigor. *PLoS biology* 2016;14(12):e2000598. doi: 10.1371/journal.pbio.2000598
4. Macleod MR, Lawson McLean A, Kyriakopoulou A, et al. Risk of Bias in Reports of In Vivo Research: A Focus for Improvement. *PLoS biology* 2015;13(10):e1002273. doi: 10.1371/journal.pbio.1002273
5. Sena ES, van der Worp HB, Bath PM, et al. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. *PLoS biology* 2010;8(3):e1000344. doi: 10.1371/journal.pbio.1000344
6. Mueller KF, Briel M, Strech D, et al. Dissemination bias in systematic reviews of animal research: a systematic review. *PLoS One* 2014;9(12):e116016. doi: 10.1371/journal.pone.0116016
7. Kilkeny C, Browne WJ, Cuthill IC, et al. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS biology* 2010;8(6):e1000412. doi: 10.1371/journal.pbio.1000412

8. Holman C, Piper SK, Grittner U, et al. Where Have All the Rodents Gone? The Effects of Attrition in Experimental Research on Cancer and Stroke. *PLoS biology* 2016;14(1):e1002331. doi: 10.1371/journal.pbio.1002331
9. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. *Nature* 2012;483(7391):531-3. doi: 10.1038/483531a
10. Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 2011;10(9):712. doi: 10.1038/nrd3439-c1
11. Wieschowski S, Chin WWL, Federico C, et al. Preclinical efficacy studies in investigator brochures: Do they enable risk-benefit assessment? *PLoS biology* 2018;16(4):e2004879. doi: 10.1371/journal.pbio.2004879
12. Kimmelman J, Federico C. Consider drug efficacy before first-in-human trials. *Nature* 2017;542(7639):25-27. doi: 10.1038/542025a
13. Kilkeny C, Parsons N, Kadyszewski E, et al. Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLoS One* 2009;4(11):e7824. doi: 10.1371/journal.pone.0007824
14. World Medical Association. Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects, Fortaleza, 2013.
15. Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA* 2000;283(20):2701-11.
16. WHO. Joint statement on public disclosure of results from clinical trials; www.who.int/ictrp/results/jointstatement/en/. 2017
17. Wieschowski S, Silva DS, Strech D. Animal Study Registries: Results from a Stakeholder Analysis on Potential Strengths, Weaknesses, Facilitators, and Barriers. *PLoS biology* 2016;14(11):e2000391. doi: 10.1371/journal.pbio.2000391
18. Kimmelman J, Mogil JS, Dirnagl U. Distinguishing between exploratory and confirmatory preclinical research will improve translation. *PLoS biology* 2014;12(5):e1001863. doi: 10.1371/journal.pbio.1001863
19. Mogil JS, Macleod MR. No publication without confirmation. *Nature* 2017;542(7642):409-11. doi: 10.1038/542409a
20. Wurbel H. More than 3Rs: the importance of scientific validity for harm-benefit analysis of animal research. *Lab Anim (NY)* 2017;46(4):164-66. doi: 10.1038/labam.1220 [published Online First: 2017/03/23]
21. Kimmelman J, Henderson V. Assessing risk/benefit for trials using preclinical evidence: a proposal. *J Med Ethics* 2016;42(1):50-3. doi: 10.1136/medethics-2015-102882
22. NC3R. Non-human primate accommodation, care and use; <https://www.nc3rs.org.uk/sites/default/files/documents/Guidelines/NC3Rs%20guidelines%20-%20Non-human%20primate%20accomodation%20care%20and%20use.pdf>. 2017
23. du Sert NP, Bamsey I, Bate ST, et al. The Experimental Design Assistant. *Nat Methods* 2017;14(11):1024-25. doi: 10.1038/nmeth.4462 [published Online First: 2017/09/30]
24. Steward O. A Rhumba of "R's": Replication, Reproducibility, Rigor, Robustness: What Does a Failure to Replicate Mean? *eNeuro* 2016;3(4) doi: 10.1523/ENEURO.0072-16.2016 [published Online First: 2016/07/16]
25. Basel Declaration Society. Basel Declaration; available at <https://www.basel-declaration.org/basel-declaration/> 2010 [

Figure

Figure 1: Two basic principles for animal research ethics translate into six practice-guiding principles (6R)

