Politics and the Life Sciences (2026), 45: 1, 1–11
doi:10.1017/pls.2025.10008

PERSPECTIVE ESSAY

Why (and how) total numbers should matter for animal experimentation policy

Nico Dario Müller

Department of Arts, Media, and Philosophy, University of Basel, Basel, Switzerland
Email: nicodario.mueller@unibas.ch

Abstract

In many countries, overall animal experimentation is not significantly decreasing or becoming less severe. Does this show that these countries' programs to promote alternatives and the "three Rs" of "replace, reduce, refine" are failing? Scholars and activists sometimes take this for granted, but representatives of "three Rs" programs have disagreed. This article makes two contributions to the debate: one conceptual and one normative. First, it draws attention to the distinction between evaluating impact (whether a program makes a difference) and evaluating sufficiency (whether a program makes enough of a difference to achieve its goals). Total numbers are typically unhelpful in assessing impact, but depending on goals, they can be relevant in assessing sufficiency. Second, this article argues that an overall decrease in harm to animals in experimentation is a sensible policy goal. This article concludes with suggestions for how to go beyond the "three Rs" to effect overall change.

Keywords: animal experimentation; 3Rs; three Rs; evaluation; program; phase-out; transition strategy

Introduction

In the early twenty-first century, total animal experimentation numbers[1] worldwide increased dramatically, from an estimated 115 million animals used in 2005 to around 192 million in 2015 (Taylor & Alvarez, 2019, p. 210). While China notably increased its animal use during the period, many European countries also saw increases or stagnation (Taylor & Alvarez, 2019, p. 207). In Switzerland, for example, the overall number of animals used for scientific purposes has remained roughly stable since 1996—around 600,000 animals per year—with a roughly stable distribution across different severity degrees (FSVO, 2024). In the European Union (plus Norway after 2018), the number of animals used for regulatory purposes and in routine production has decreased steadily since 2015, but the number of animals used in basic, applied, and translational research has remained stagnant around 6–7 million per year (European Commission, 2025). Across all purposes in the European Union, the percentage of animals used for procedures of different severities has been shifting slightly, with fewer animals being used for "mild," more for "moderate," and a roughly stable number used for "severe" procedures (ibid.).

Many countries today have programs or institutions that support the "three Rs" of "replace, reduce, refine." Should it be concerning, then, if overall harm to animals is not significantly decreasing, especially in research? Does it provide a reason to reconsider existing programs? One might take this for granted,

[1]This article uses "total animal experimentation numbers" or "total numbers" as a shorthand for the numbers associated with overall harm done to animals for research and testing purposes in a given jurisdiction—most importantly the number of animal experiments conducted that involved the infliction of harm, the number of animals used in such experiments, and the proportion of different degrees of severity among overall experiments—which are usually reported in official animal experimentation statistics.

© The Author(s), 2025. Published by Cambridge University Press on behalf of The Association for Politics and the Life Sciences. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press

but in fact, the question is controversial. Some scholars and activists have argued that rising or stagnating total animal experimentation numbers are a sign of failure on the part of "three Rs" programs, but representatives of the programs themselves have disagreed, arguing that total numbers obscure too much to be helpful in program evaluation (see the "Debates over 3Rs program evaluation" section for examples).

To move the debate forward, the present article contributes two points: one conceptual and one normative. First, it draws a distinction between two purposes for which total numbers might be relevant —they might matter for an assessment of *program impact* or for an assessment of *program sufficiency* relative to its goals (section "How total numbers might matter: program impact versus sufficiency"). While total numbers are at best indirectly relevant for an assessment of impact, they are directly relevant for assessing sufficiency, provided that program goals have implications in terms of total numbers. However, program goals are often vague. To see whether programs are sufficient to achieve their goals, policymakers first need to set clear goals.

Second, this article turns to the normative question of whether making a difference at the level of total animal experimentation numbers *should* be a policy goal (section "Why total numbers should matter"). Of course, that depends on normative background assumptions—but which ones? One might expect that only agents who value animals particularly highly would have reason to call for an overall decrease in harm to animals in experimentation. However, the article will argue that this policy goal is sensible from a widely agreeable "middle ground" standpoint which acknowledges that animals have a claim to protection but considers their claims less important than those of humans. The goal of this section will not be to list every possible critical concern about animal experimentation, which might also include concerns about scientific quality (Akhtar, 2015), financial cost (Bottini & Hartung, 2009; Keen, 2019; Meigs et al., 2018), and the psychological toll the practice can take on personnel (Baysal et al., 2024; Johnson & Smajdor, 2019; King & Zohny, 2022; LaFollette et al., 2020). Rather, the aim is to provide one robust argument for a policy goal of decreasing overall harm to animals in experimentation —even in cases where it avoids those other critical concerns. The article ends with some suggestions for what policies beyond the "three Rs" would have to look like in order to be able to decrease overall harm by overcoming reliance on animal experimentation (section "Designing policies that fare better"), before concluding (section "Conclusion").

In sum, this article argues that total animal experimentation numbers should not be dismissed wholesale when evaluating success and failure in animal experimentation policy, though program goals should be more clearly defined. From a widely agreeable standpoint, when the numbers indicate stagnating or even rising overall harm inflicted on animals in experiments, it is indeed time to reconsider "three Rs" programs.

Debates over 3Rs program evaluation

Today, many governments and institutions undertake efforts to promote the "three Rs" of "replace, reduce, refine" (Bayne et al., 2015; Neuhaus et al., 2022a, 2022b; see generally Russell & Burch, 1959). Scholars have argued that it is a sign of failure on the part of these programs when total animal experimentation numbers are stagnating or rising (Bailey, 2024; Blattner, 2019; Herrmann, 2019). Animal advocates have endorsed the same claim (Bertrand, 2024; Marshall et al., 2022; PETA, 2024), while others have stated sarcastically that the 3Rs "succeeded in preserving the status quo" (Stop Vivisection, 2016). It is sometimes assumed that the "reduce" principle calls for a decrease in overall animal experimentation (as pointed out by Olsson et al., 2012, p. 333). The traditional and still dominant interpretation, however, is that it calls for reducing the sample size within studies (Lauwereyns, 2018, p. 17; Russell & Burch, 1959, Ch. 4), which is compatible with an overall stagnation or increase if more individual studies are conducted. And while the "replace" principle calls for substituting animals with other models or approaches, new animal models can be innovated in the meantime so that replacement progress is compatible with an overall increase of animals used (Müller, 2023). Thus, the implications of

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press

the "three Rs" for animal experimentation at the level of total numbers are not as straightforward as one might think.

The view that rising animal experimentation numbers indicate a failure of "three Rs" programs has been disputed by representatives of those programs. For instance, a 2012 report by the British NC3Rs on the institution's evaluation framework contains a special annex titled "The Home Office statistics are not a gauge of progress in the 3Rs" (NC3Rs, 2012, p. 49). The annex argues that "the number of animals reported in the annual statistics is influenced by a range of scientific and strategic factors independent of the 3Rs" (NC3Rs, 2012, p. 51), such as investment decisions by major funding bodies and technological breakthroughs like gene modification techniques that open up new possibilities to engineer and use animal models (ibid.). Furthermore, the report continues, "reductions in animal use for some studies have been achieved but this may not be apparent if there is an overall increase in the number of such studies performed" (NC3Rs, 2012, p. 51), and "developments that avoid animal use are not easily counted" (NC3Rs, 2012, p. 52).

Along similar lines, scholars associated with a Swiss program for the "three Rs" have more recently argued:

> To assess the effectiveness of the 3Rs principle as a policy instrument for advancing humane animal experimentation, appropriate parameters for measuring the effects of the 3Rs are still needed. […] First, the metrics track the numbers of animals used rather than those replaced, and second, most replacement approaches are not recognized as such. (Grimm et al., 2023, p. 6)

> There is a persistent perception that there is a "missing effect" of the implementation of the 3Rs. However, […] proper outcome measures for replacement, reduction and refinement alike are largely undeveloped, and therefore one is entitled to question whether the effect really is missing, or whether instead we are simply unable to properly measure it on a wider, global scale. (Grimm et al., 2023, p. 10)

The central argument advanced by NC3Rs (2012) and Grimm et al. (2023) is straightforward: Overall numbers are irrelevant for "three Rs" evaluation because they are influenced by factors beyond the "three Rs" and because they only track the animal experiments conducted, not the ones avoided. However, as the next section will argue, the force of this argument depends on what we are evaluating: a program's ability to make a difference, or its ability to achieve a given policy goal.

How total numbers might matter: program impact versus sufficiency

Despite stark differences in tone, critics and representatives of "three Rs" programs might not truly disagree over the role of total numbers but might rather be talking past each other. There is a difference between the question of whether a program has an *impact*—that is, whether it affects any changes in the world—and the question of whether a program is *sufficient*—that is, whether it achieves its goals. A program can have great effects while failing to have the specific effects it is meant to have. In such situations, it is easy to have pseudo-disagreements about whether the program is "working."

By way of analogy, consider the following vignette:

> *Goalkeeper:* An association football coach is furious because their team has been on a losing streak, conceding an increasing number of goals with each match. The coach tells the goalkeeper she needs to train harder, but the goalkeeper protests: The number of goals *conceded* is misleading, as it ignores other important variables (e.g., the number of shots on target) and it does not track how many goals were *prevented*. Say that the goalkeeper is indeed among the best in the league when it comes to the number of shots on target blocked, and has even been improving. It is just that the

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press

number of shots on target has been rising even more rapidly, masking the goalkeeper's improvement in the overall result.

In this scenario, both the coach and the goalkeeper have a point. The coach has reasons to be unhappy with the number of goals conceded because this number is directly relevant to whether the team wins or loses. The goalkeeper is right to point out that they are doing their best from their position. The problem is that the two do not mean the same by "success" and "failure." The coach is evaluating the *sufficiency* of current efforts: whether they are enough to adequately address the problem (the threat of losing games). The goalkeeper, by contrast, is evaluating *impact*: whether current efforts are making a difference. The coach and the goalkeeper do not truly disagree. Both can agree that factors beyond the goalkeeper's influence, such as maybe the team's defense strategy, require revision to win games.

The case of "three Rs" programs is similar to an extent. When program representatives object to the use of total numbers to assess their own success or failure, they are talking about impact. The NC3Rs program evaluation report does so unambiguously, articulating a framework of "inputs–outputs/outcomes–interim impacts–mature impacts" (2012, pp. 6–7), where impacts are essentially the difference made by the initial inputs. Grimm et al. (2023) focus on "effects" and "effectiveness," both of which, in context, again denote impact. However, policymakers and governments may be more interested in whether a program can adequately address a given problem and reach given goals—not just in whether they are making *some* difference.

However, there is also an important disanalogy between the goalkeeper vignette and animal experimentation: A football team definitely wants to concede fewer goals, but whether a given political body wants to conduct less (or less severe) overall animal experimentation is often unclear.

The European Union Directive 2010/63 does mention "full replacement" as a "final goal" in its Article 10, but this has never been concretized into more specific reduction goals (as a European Parliament resolution highlighted; EU Parl, 2021). Most individual countries likewise do not design animal experimentation policies to achieve outcome-oriented goals, let alone specific reduction goals. They rather lay down licensing procedures (Olsson et al., 2024) and more recently provide resources for "three Rs" programs in a continuous, open-ended fashion. An exception is Germany, whose government in 2024 started a participatory process to develop a "reduction strategy" for overall animal experimentation (BMEL, 2024), whose fate at the time of writing is, however, uncertain, as the governing coalition has changed (Tierschutzbund, 2025). There exist other, partial exceptions. In 2023, the European Commission announced a reduction plan for regulatory animal tests, though it refused to set specific reduction goals for animal experimentation overall (EC, 2023). From 2019 to 2024, the United States Environmental Protection Agency worked on a phase-out plan for the animal tests it requires and funds, though it subsequently abandoned the plan (D. Grimm, 2024). Two other federal agencies—the National Institute of Health (NIH) and the Food and Drug Administration (FDA)—in 2025 announced commitments to reduce their reliance on animal testing (FDA, 2025; NIH, 2025). Perhaps the most noteworthy exception for present purposes is the British NC3Rs, which has stated that it aims to prevent a rise in total animal experimentation in the United Kingdom from 2015 to 2025 (NC3Rs, 2014).

In these exceptional cases, where governments or institutions have committed to a reduction or the prevention of a rise in total animal experimentation numbers, a rise in the numbers unambiguously indicates that efforts have been *insufficient*. This does not contradict the NC3Rs' claim that total numbers are unhelpful in assessing *impact*. The problem is that in most countries, no specific outcome-oriented goals exist for animal experimentation policy. And in the absence of clear enough goals, program sufficiency cannot be meaningfully evaluated.

Vagueness about program goals opens up a space for political maneuvering. Those who, for fiscal or other reasons, oppose the creation of additional programs can point out that existing programs are doing good work and are making an impact—it would certainly be worse without them! This tacitly equates impact with sufficiency, setting no goal for the program other than doing whatever it already does. Conversely, those wishing to revise or increase efforts can stress that the total numbers are not changing in the right direction, tacitly suggesting that achieving a particular overall change is part of the program's

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press

goals, which may be unrealistic, given how the program is set up. Due to the inherent limitations of the "three Rs" framework—it says nothing about the allocation of research funds, for example (Müller, 2024a; NC3Rs, 2012, p. 51)—programs to promote the "three Rs" might be ill-equipped to achieve significant change at the level of total animal experimentation numbers, just as a goalkeeper alone cannot make up for bad defense. In this way, goal vagueness sets "three Rs" programs up for uncomfortable and unproductive debates about their supposed success or failure whenever animal experimentation statistics are released.

As a first result of this discussion, we can conclude that "three Rs" program representatives are right to point out that total numbers are dubious indicators of program impact. But what often interests policymakers, and *should* interest them, is the sufficiency of programs to achieve policy goals. The problem is that in animal experimentation policy, the goals are often unclear. To achieve a more productive debate, policymakers should make up their minds about what they want "three Rs" programs to achieve and articulate clear goals.

While setting policy goals is of course a political matter, it can be informed by philosophical arguments. Policymakers may want to stick to views that are compatible with widely shared values and normative assumptions that already underpin existing policy. From this standpoint, as the next section will argue, the policy goal of decreasing overall harm to animals is sensible.

Why total numbers should matter

For present purposes, assume (a) that most people view animals as having moral status and a nontrivial claim to their own welfare, but (b) that most people do not value animals so highly as to consider most animal experimentation to be all-things-considered wrong. In other words, most people think that animals are worth protecting for their own sake, but not in the same way and to the same extent as humans. They occupy some middle ground between anthropocentrism (the view that only humans have moral claims and animals matter only indirectly) and antispeciesism (the view that all interests should be considered equally irrespective of species) or an animal rights view (the view that animals have moral rights that trump mere human interests).[2] Animal experiments are morally acceptable from this standpoint as long as they yield enough benefit for humans or, less frequently, other animals.

Many governments have good reason to think along the lines of this middle-ground view not just because their individual citizens endorse it, but also because they are already tacitly committed to it. An increasing number of legal systems give animals some special status as sentient beings and protect them for their own sake (Kotzmann, 2023). Many legal systems regulate animal experiments through licensing procedures based on harm–benefit analysis (Bayne et al., 2024; Brønstad et al., 2016; Laber et al., 2016). This reflects a commitment to both the good of animal welfare and the good of free research as well as health and safety. Assuming that most people and their governments thus occupy a middle ground between anthropocentrism and antispeciesism or animal rights, do they have good reasons to adopt the policy goal of an overall decrease in harm to animals in experimentation?

This might not be obvious at first sight. Total numbers clearly matter when the problem consists in all-things-considered wrongs or evils. However, according to the middle ground position, not all animal experiments are all-things-considered wrongs. Whether they are depends on how much benefit is generated and how much harm is inflicted. If there is any indication that many animal experiments create more harm than benefit (e.g., because of poor scientific quality or because species differences

[2]These assumptions may, of course, be false in some contexts, and for those contexts, the argument of this section does not hold without reservations. For instance, most people may lack clearly defined moral views about animals. They might also be internally conflicted (see Dhont & Hodson, 2019; McGlacken, 2022; McGlacken & Hobson-West, 2022) or have a value-action gap (Vigors, 2018). Though quantitative data about people's attitudes to animals are unfortunately scarce, the mere fact that policies in line with the middle ground position have not sparked major outrage suggests that they are broadly in line with majority views. If anything, evidence suggests that most Europeans apparently want more animal protection, not less (European Union, 2023).

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press

undermine translatability, Akhtar, 2015), a sensible policy goal would be to decrease the number of just these net-harmful experiments specifically, while the number of net-beneficial experiments may be allowed to stagnate or even increase. However, given that existing licensing procedures often already involve a form of harm–benefit analysis, at most, this could imply a policy of double-checking whether the procedures work as intended (e.g., by conducting retrospective harm–benefit analysis, as was done by Pound & Nicol, 2018). Though this may be worthwhile, it does not yet justify a policy goal of decreasing overall harm to animals in experimentation more generally. If the problem is bad science, the goal should be to reduce bad science—not animal-harming science.

However, not just all-things-considered wrongs should be avoided, but so should *pro tanto* wrongs. To use a simple example: Saving a young person rather than an old person from a burning building (when one cannot save both) might be the right decision because the young person's claim to rescue is greater due to their future prospects. Still, the agent has reasons to regret their failure to save the old person even if they made all the morally right decisions (Williams, 1965). That greater moral claims outweighed the senior's claim to rescue does not mean there is no problem—to the contrary, having to choose is a moral problem in itself. And it certainly does not mean that we need not be concerned if such cases are statistically on the rise. As Brigid Brophy pointed out over 50 years ago, if you constantly have to choose whom or what to save from burning buildings, you need better fire safety (Brophy, 1972, p. 137). The same is true for any nontrivial *pro tanto* wrong.

One could object that the state is not an individual moral agent and does not feel regret over value trade-offs. The law has ways to navigate conflicts between the goods it protects, such as the aforementioned harm–benefit analysis for animal experiments (Bayne et al., 2024; Brønstad et al., 2016; Laber et al., 2016; Olsson et al., 2024). As far as the law is concerned, sacrificing one good for the sake of another is not a problem if the proper procedures are followed. However, the state's procedures should not be arbitrary. They should aim to protect all affected goods as fully as possible. And it is always better for the protection of two goods to avoid their conflict in the first place than merely to sacrifice one for the other when the conflict arises. Thus, although a state may not *regret* having to weigh animal welfare against scientific advances, it should still try to reduce the overall number of occasions on which it needs to choose.

Therefore, adopting the policy goal of decreasing total harm to animals in science and testing is sensible—it has good reasons supporting it—as long as we assume that harming animals for the sake of experimentation-derived benefits is regrettable to a nontrivial extent, even if we take it to be justified all-things-considered.

However, being supported by good reasons is not enough for a policy goal to be worth adopting. The goal must also not have stronger reasons *against* it. The main worry one might have about the policy goal of decreasing harm to animals in experimentation is that it threatens other social goods, such as the progress of science or health and safety. However, the above argument does not suggest that the policy goal should be to decrease harm to animals in experimentation *no matter what.* Rather, it should be to reduce the incidence of conflicts between animal welfare and science or health and safety, and only in this way to change the total animal experimentation numbers. In other words, the goal should be to decrease overall harm to animals in experimentation without diminishing the social goods that are enabled by science and health and safety. This requires that either the functions of animal-harming experiments are fulfilled by other approaches, or that any minus in benefits from such experiments is outweighed by greater benefits from other approaches, even if the benefits are different in kind.

This policy goal—decreasing conflicts between animal welfare and the goods enabled by science and regulatory testing—is not the same as *increasing* the number of animals replaced or animal experiments avoided. The former goal, but not the latter, would be achieved without effort in a world where major research trends just happened to go in the direction of fields that are naturally animal-free. In this world, no researchers would have actively replaced animals, but the goods of animal welfare and science would not be at odds as often. So the two goals differ in what they substantively demand. They also differ in measurability. Total numbers can feasibly be collected, as they are already officially reported in many jurisdictions. By contrast, counting the number of animals replaced or animal experiments avoided

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press

would require a great deal of controversial speculation about what *would* have happened absent intervention. Fortunately, this speculation is not necessary to assess progress toward the policy goal suggested by the argument in this section.

In sum, the acknowledgment that animal welfare matters, even if it does not matter in the same way or to the same extent as human welfare, implies that a policy goal of reducing harm to animals at the level of total numbers is sensible. We should aim to decrease the number of occasions on which we have to choose which of two goods to sacrifice. This argument does not show that most existing "three Rs" programs are failing *on their own terms.* For that, the goal of decreasing overall harm to animals in experimentation would have to be officially adopted. However, many governments and institutions would have good reasons to adopt that goal, and once they do, stagnant or rising total animal experimentation numbers will show many current "three Rs" programs to be insufficient, impactful though they may be.

Designing policies that fare better

So far, this article has argued that total numbers can be relevant for program sufficiency depending on policy goals and that a policy goal of decreasing overall harm to animals without jeopardizing science or health and safety is sensible. If this goal were adopted by jurisdictions in which total numbers are stagnating, then the respective "three Rs" programs would be clearly insufficient. To all this, one might respond that "three Rs" programs are still the best one can do to work in the direction of that goal. However, there exist alternative approaches that could be taken in addition to "three Rs" programs.

Notably in the Netherlands, efforts to move away from animal experimentation have increasingly been framed in terms of a technology transition rather than the "three Rs" (Denktank, 2015; NCad, 2016).[3] A similar shift is visible in the previously mentioned steps toward reduction or phase-out strategies in Germany, the European Union, and the United States (see section "How total numbers might matter: program impact versus sufficiency"). Academics (Baumgartl-Simons & Hohensee, 2019; Herrmann, 2019; Hutchinson et al., 2022; Marshall et al., 2022; Müller, 2024b) and NGOs (Cruelty Free Europe, 2022; PETA, 2024) have made various suggestions for how to flesh out such strategies. Given that the policy goal is not just to decrease animal experimentation, but to decrease harm to animals without jeopardizing science or health and safety, a strategic approach that monitors intended and unintended consequences is needed (see Müller, 2024b). A reduction or phase-out strategy should furthermore be flexible enough to allow for course corrections when milestones are missed (see Müller, 2024b).

A challenge that reduction or phase-out efforts face is to tackle the environmental factors pointed out by the NC3Rs that affect animal experimentation, but lie beyond the influence of "three Rs" programs, such as funding allocation, technological developments, and regulation. The "three Rs" approach does not address these factors because it focuses on the individual researcher as the main locus of control. No amount of "R" principles that make prescriptions only for researchers will overcome this limitation. However, one could draw on frameworks developed in other policy areas where a transformation in socio-technical systems is sought (notably, climate and biodiversity policy), such as transformative governance (Visseren-Hamakers et al., 2021), transition management (Loorbach, 2010), or mission-oriented innovation policy (Lindner et al., 2021; Mazzucato, 2018). All of these approaches heavily feature stakeholder involvement and combine the top-down with the bottom-up. The range of possible

[3]In fact, some stakeholders have questioned the role of total numbers in evaluating transition progress, echoing the arguments of "three Rs" representatives (NCad, 2024, p. 16). Other stakeholders, however, argued that "we should not abandon the goal of reducing the number of reported animal experiments, as we ultimately aim to see a reduction in that area" (ibid.). Compared with a "three Rs" framing, a "transition" framing more strongly suggests a policy goal that implies a change in total animal experimentation numbers. However, any vagueness about policy goals still opens up the room for political maneuvering discussed in the "How total numbers might matter: program impact versus sufficiency" section. Therefore, even within a "transition" framing, clear policy goals are indispensable.

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press

measures to be taken in the course of a transformative approach to animal experimentation policy is by no means restricted to prohibitive or other regulatory measures. A "planning without banning" approach (Müller, 2024c) suggests that a decrease in harm to animals in research should be achieved through non-prohibitive, stimulating interventions in the environment of scientific work comprised of funding opportunities, infrastructure and equipment, research networks, and educational programs, among other factors. While developing viable reduction or phase-out strategies is a challenging task, governments could feasibly create working groups or competence centers to accomplish it.

Additional efforts beyond the current "three Rs" programs would require additional resources. They would also threaten the interests of some researchers, administrators, and industries that benefit from current arrangements. Very likely, the debate will thus feature "norm antipreneurs" (Bloomfield, 2015), who often defend the status quo by arguing that there is no problem to solve, and therefore that current programs do not produce "outcomes that vary *radically* and *damagingly* from expectations" (Bloomfield, 2015, p. 323). A vague statement of goals, thus of the government's and the public's expectations, makes this argument exceedingly easy to make. Activists and policymakers, if they want to decrease conflicts between animal welfare and science or health and safety in the long run, would thus be well-advised to insist on clear goal-setting for "three Rs" programs.

Conclusion

This article has argued that representatives of "three Rs" programs are right to an extent when they dismiss total numbers as indicators of program success because they focus on program impact. However, policymakers should be more interested in whether programs are sufficient to attain policy goals, for which total numbers can very well be relevant. The problem is that clear outcome-oriented goals are often missing in animal experimentation policy. Thus, if policymakers want to be able to tell whether these programs are succeeding, they should make it clear what change they expect them to effect. This article has furthermore argued that there are good reasons to adopt an overall decrease in harm to animals in experimentation as a policy goal. Almost anyone can agree that it is a pity if we have to decide whether to protect animals or attain the benefits of science and regulatory testing—the goal should be to not have to choose as often. If policymakers follow this advice and adopt an overall decrease of harm to animals in experimentation as a policy goal, then stagnating or rising total numbers are indeed markers of program failure, and existing programs for the "three Rs" need to be reconsidered or complemented with new approaches. These would have to tackle the environmental factors that affect overall animal experimentation numbers, such as funding, regulation, and the focus of innovation activities.

Acknowledgments. The author thanks Angela Martin, Christian Rodriguez-Perez, and two anonymous reviewers for this journal for their helpful feedback on earlier versions of this article.

Financial support. This work was supported by the Swiss National Science Foundation's NRP 79 (Grant No. 447944_214850). The author makes it transparent that he is involved in one of the programs discussed in this article, the NRP 79, as a principal investigator of a research project, but he is not involved in, or personally affected by, the program's evaluation.

References

Akhtar, A. (2015). The flaws and human harms of animal experimentation. *Cambridge Quarterly of Healthcare Ethics*, **24**, 407–419. https://doi.org/10.1017/S0963180115000079

Bailey, J. (2024). It's time to review the three Rs, to make them more fit for purpose in the 21st century. *Alternatives to Laboratory Animals*, **52**(3), 155–165. https://doi.org/10.1177/02611929241241187

Baumgartl-Simons, C., & Hohensee, C. (2019). How can the final goal of completely replacing animal procedures successfully be achieved? In K. Herrmann, & K. Jayne (Eds.), *Animal experimentation: Working towards a paradigm change* (pp. 88–123). Brill.

Bayne, K., Guillen, J., France, M. P., & Morris, T. H. (2024). Legislation and oversight of the conduct of research using animals: A global overview. In H. Golledge, & C. Richardson (Eds.), *The UFAW handbook on the care and management of*

laboratory and other research animals (9th ed., pp. 101–121). John Wiley & Sons Ltd. https://doi.org/10.1002/9781119555278.ch8

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

Baysal, Y., Goy, N., Hartnack, S., & Canu, I. G. (2024). Moral distress measurement in animal care workers: A systematic review. *BMJ Open*, **14**, Article e082235. https://doi.org/10.1136/bmjopen-2023-082235

Bertrand, K. (2024). *Exposing the gaps: How the 3Rs fail to protect animals in research*. Science Advancement & Outreach, A Division of PETA. https://www.scienceadvancement.org/reflections/exposing-the-gaps-how-the-3rs-fail-to-protect-animals-in-research/

Blattner, C. (2019). Rethinking the 3Rs: From whitewashing to rights. In K. Herrmann, & K. Jayne (Eds.), *Animal experimentation: Working towards a paradigm change* (pp. 168–193). Brill.

Bloomfield, A. (2015). Norm antipreneurs and theorising resistance to normative change. *Review of International Studies*, **42**(2), 310–333.

BMEL. (2024). *Tierversuche durch Alternativmethoden nachhaltig reduzieren: BMEL startet Beteiligungsprozess für Reduktionsstrategie mit Experten-Treffen* (Press Release No. 89). Bundesministerium für Ernährung und Landwirtschaft. https://www.bmel.de/SharedDocs/Pressemitteilungen/DE/2024/089-tierversuche.html

Bottini, A. A., & Hartung, T. (2009). Food for thought… on the economics of animal testing. *ALTEX*, **26**(1), 3–16. https://doi.org/10.14573/altex.2009.1.3

Brønstad, A., Newcomer, C. E., Decelle, T., Everitt, J. I., Guillen, J., & Laber, K. (2016). Current concepts of harm–benefit analysis of animal experiments—report from the AALAS-FELASA Working Group on Harm–Benefit Analysis—Part 1. *Laboratory Animals*, **50**(1S), 1–20. https://doi.org/10.1177/0023677216642398

Brophy, B. (1972). In pursuit of a fantasy. In S. Godlovitch, R. Godlovitch, & J. Harris (Eds.), *Animals, men and morals: An enquiry into the maltreatment of non-humans* (pp. 125–145). Taplinger Publishing Company.

Cruelty Free Europe. (2022). *Reducing and replacing animal experiments: Europe needs an action plan*. https://crueltyfreeeurope.org/sites/default/files/Cruelty%20Free%20Europe%20Reducing%20and%20replacing%20animal%20experiments%20Europe%20needs%20an%20action%20plan%20%281%29.pdf

Denktank. (2015). *In transitie! Nederland internationaal toonaangevend in proefdiervrije innovaties*. Denktank Aanvullende Financiering alternatieven voor dierproeven. https://www.transitieproefdiervrijeinnovatie.nl/binaries/proefdiervrije-innovatie/documenten/rapporten/15/10/15/in-transitie/Advies_In_transitie_Nederland_internationaal_toonaangevend_in_proefdiervrije_innovaties.pdf

Dhont, K., & Hodson, G. (Eds.) (2019). *Why we love and exploit animals: Bridging insights from academia and advocacy*. Routledge. https://doi.org/10.4324/9781351181440

EC. (2023). *Communication from the commission on the European Citizens' Initiative (ECI) "Save cruelty-free cosmetics—Commit to a Europe without animal testing"* [No. C(2023) 5041 final]. European Commission.

EU Parl. (2021). *Plans and actions to accelerate a transition to innovation without the use of animals in research, regulatory testing and education* (p. P9_TA(2021)0387) [The transition to innovation without the use of animals in research, regulatory testing and education (2021/2784(RSP))]. European Parliament.

European Commission. (2025). *ALURES—animal use reporting—EU System EU statistics database on the use of animals for scientific purposes under Directive 2010/63/EU*. https://webgate.ec.europa.eu/envdataportal/content/alures/section1_number-of-animals.html

European Union. (2023). *Attitudes of Europeans towards animal welfare* [Special Eurobarometer 533—Report]. https://europa.eu/eurobarometer/api/deliverable/download/file?deliverableId=88297

FDA. (2025). *Roadmap to reducing animal testing in preclinical safety studies*. Food and Drug Administration. https://www.fda.gov/media/186092/download?attachment

FSVO. (2024). *Bericht Tierversuchsstatistik 2023*. Federal Food Safety and Veterinary Office. https://www.blv.admin.ch/dam/blv/de/dokumente/tiere/tierversuche/bericht-tierversuchsstatistik-2023.pdf.download.pdf/DE_Bericht_Tierversuchsstatistik_2023.pdf

Grimm, D. (2024). EPA scraps plan to end mammal testing by 2035. *Science*. https://doi.org/10.1126/science.zbmpq89

Grimm, H., Biller-Andorno, N., Buch, T., Dahlhoff, M., Davies, G., Cederroth, C. R., Maissen, O., Lukas, W., Passini, E., Törnqvist, E., Olsson, I. A. S., & Sandström, J. (2023). Advancing the 3Rs: Innovation, implementation, ethics and society. *Frontiers in Veterinary Science*, **10**, Article 1185706. https://doi.org/10.3389/fvets.2023.1185706

Herrmann, K. (2019). Refinement on the way towards replacement: Are we doing what we can? In K. Herrmann, & K. Jayne (Eds.), *Animal experimentation: Working towards a paradigm change* (pp. 3–64). Brill.

Hutchinson, I., Owen, C., & Bailey, J. (2022). Modernizing medical research to benefit people and animals. *Animals*, **12**, Article 1173.

Johnson, J., & Smajdor, A. (2019). Human wrongs in animal research: A focus on moral injury and reification. In K. Herrmann, & K. Jayne (Eds.), *Animal experimentation: Working towards a paradigm change* (pp. 305–318). Brill.

Keen, J. (2019). Wasted money in United States biomedical and agricultural animal research. In K. Herrmann, & K. Jayne (Eds.), *Animal experimentation: Working towards a paradigm change* (pp. 244–272). Brill.

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press

King, M., & Zohny, H. (2022). Animal researchers shoulder a psychological burden that animal ethics committees ought to address. *Journal of Medical Ethics*, **48**(5), 299–303. https://doi.org/10.1136/medethics-2020-106945

Kotzmann, J. (2023). Sentience and intrinsic worth as a pluralist foundation for fundamental animal rights. *Oxford Journal of Legal Studies*, **43**(2), 405–428. https://doi.org/10.1093/ojls/gqad003

Laber, K., Newcomer, C. E., Decelle, T., Everitt, J. I., Guillen, J., & Brønstad, A. (2016). Recommendations for addressing harm–benefit analysis and implementation in ethical evaluation—report from the AALAS-FELASA Working Group on Harm–Benefit Analysis—Part 2. *Laboratory Animals*, **50**(1S), 21–42. https://doi.org/10.1177/0023677216642397

LaFollette, M. R., Riley, M. C., Cloutier, S., Brady, C. M., O'Haire, M. E., & Gaskill, B. N. (2020). Laboratory animal welfare meets human welfare: A cross-sectional study of professional quality of life, including compassion fatigue in laboratory animal personnel. *Frontiers in Veterinary Science*, **7**, Article 114. https://doi.org/10.3389/fvets.2020.00114

Lauwereyns, J. (2018). *Rethinking the three R's in animal research*. Springer International Publishing.

Lindner, R., Edler, J., Hufnagl, M., Kimpeler, S., Kroll, H., Roth, F., Wittmann, F., & Yorulmaz, M. (2021). *Mission-oriented innovation policy: From ambition to successful implementation* [Policy Brief No. 02/2021]. Fraunhofer ISI. https://doi.org/10.24406/publica-fhg-416799

Loorbach, D. (2010). Transition management for sustainable development: A prescriptive, complexity-based governance framework. *Governance*, **23**(1), 161–183.

Marshall, L. J., Constantino, H., & Seidle, T. (2022). Phase-in to phase-out-targeted, inclusive strategies are needed to enable full replacement of animal use in the European Union. *Animals*, **12**, Article 863.

Mazzucato, M. (2018). *Mission-oriented research & innovation in the European Union: A problem-solving approach to fuel innovation-led growth*. European Commission, Directorate-General for Research and Innovation. https://www.sbtse.gr/images/Mission-oriented_ResearchInnov.pdf

McGlacken, R. (2022). Constrained, contingent, and conflicted: Complicating acceptance of animal research through an analysis of writing from the UK Mass Observation Project. In D. Bruce, & A. Bruce (Eds.), *Transforming food systems: Ethics, innovation and responsibility* (pp. 245–250). Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-939-8_37

McGlacken, R., & Hobson-West, P. (2022). Critiquing imaginaries of "the public" in UK dialogue around animal research: Insights from the Mass Observation Project. *Studies in History and Philosophy of Science*, **91**, 280–287. https://doi.org/10.1016/j.shpsa.2021.12.009

Meigs, L., Smirnova, L., Rovida, C., Leist, M., & Hartung, T. (2018). Animal testing and its alternatives—the most important omics is economics. *ALTEX*, **35**(3), 275–305. https://doi.org/10.14573/altex.1807041

Müller, N. D. (2023). The 3Rs alone will not reduce total animal experimentation numbers: A fundamental misunderstanding in need of correction. *Journal of Applied Animal Ethics Research*, **5**(2), 269–284. https://doi.org/10.1163/25889567-bja10042

Müller, N. D. (2024a). Beyond anthropocentrism: The moral and strategic philosophy behind Russell and Burch's 3Rs in animal experimentation. *Science and Engineering Ethics*, **30**(44), 1–15. https://doi.org/10.1007/s11948-024-00504-1

Müller, N. D. (2024b). Phase-out planning for animal experimentation: A definition, an argument, and seven action points. *ALTEX*, **41**(2), 260–272. https://doi.org/10.14573/altex.2312041

Müller, N. D. (2024c). Planning without banning: Animal research and the argument from avoidable harms. *Ethical Theory and Moral Practice*, **28**, 111–124. https://doi.org/10.1007/s10677-024-10455-y

NC3Rs. (2012). *Evaluating progress in the 3Rs: The NC3Rs framework*. National Centre for the Replacement, Refinement & Reduction of Animals in Research. https://www.nc3rs.org.uk/our-evaluation-framework

NC3Rs. (2014). *Our vision: 2015–2025*. National Centre for the Replacement, Refinement & Reduction of Animals in Research. https://nc3rs.org.uk/sites/default/files/2021-09/NC3Rs%20Our%20Vision%202015-2025.pdf

NCad. (2016). *Transition to non-animal research: On opportunities for the phasing out of animal procedures and the stimulation of innovation without laboratory animals*. Netherlands National Committee for the protection of animals used for scientific purposes.

NCad. (2024). *Evaluation of the NCad policy advice "transition to animal-free research."* Netherlands National Committee for the protection of animals used for scientific purposes. https://english.ncadierproevenbeleid.nl/binaries/ncad-english/documenten/reports/24/10/03/evaluation-transition-advice/Evaluation+of+the+NCad+Policy+Advice+%E2%80%98Transition+to+non-animal+research%E2%80%99.pdf

Neuhaus, W., Reininger-Gutmann, B., Rinner, B., Plasenzotti, R., Wilflingseder, D., De Kock, J., Vanhaecke, T., Rogiers, V., Jírová, D., Kejlová, K., Knudsen, L. E., Nielsen, R. N., Kleuser, B., Kral, V., Thöne-Reineke, C., Hartung, T., Pallocca, G., Leist, M., Hippenstiel, S., … Spielmann, H. (2022a). The rise of three Rs centres and platforms in Europe. *Alternatives to Laboratory Animals*, **50**(2), 90–120. https://doi.org/10.1177/02611929221099165

Neuhaus, W., Reininger-Gutmann, B., Rinner, B., Plasenzotti, R., Wilflingseder, D., De Kock, J., Vanhaecke, T., Rogiers, V., Jírová, D., Kejlová, K., Knudsen, L. E., Nielsen, R. N., Kleuser, B., Kral, V., Thöne-Reineke, C., Hartung, T., Pallocca, G., Rovida, C., Leist, M., & Spielmann, H. (2022b). The current status and work of three Rs centres and platforms in Europe. *Alternatives to Laboratory Animals*, **50**(6), 381–413. https://doi.org/10.1177/02611929221140909

NIH. (2025). *NIH to prioritize human-based research technologies* [Press release]. National Institutes of Health. https://www.nih.gov/news-events/news-releases/nih-prioritize-human-based-research-technologies

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press

Olsson, I. A. S., Franco, N. H., Weary, D. M., & Sandøe, P. (2012). The 3Rs principle: Mind the ethical gap! In *ALTEX Proceedings: Proceedings of the 8th World Congress on Alternatives and Animal Use in the Life Sciences, Montreal 2011* (pp. 333–336). Johns Hopkins University Press.

Olsson, I. A. S., Sandøe, P., Hawkins, P., & Jennings, M. (2024). Ethics review of animal research. In H. Golledge, & C. Richardson (Eds.), *The UFAW handbook on the care and management of laboratory and other research animals* (9th ed., pp. 281–296). John Wiley & Sons Ltd. https://doi.org/10.1002/9781119555278.ch18

PETA. (2024). *Research modernization NOW*. PETA.

Pound, P., & Nicol, C. J. (2018). Retrospective harm–benefit analysis of pre-clinical animal research for six treatment interventions. *PLOS ONE*, **13**(3), Article e0193758.

Russell, W. M. S., & Burch, R. L. (1959). *The principles of humane experimental technique*. Methuen.

Stop Vivisection. (2016). *Stop Vivisection counter-conference: A response to the Commission's lack of accountability*. http://www.stopvivisection.eu/sites/default/files/press_release_3_12_16_.pdf

Taylor, K., & Alvarez, L. R. (2019). An estimate of the number of animals used for scientific purposes worldwide in 2015. *Alternatives to Laboratory Animals*, **47**(5–6), 196–213. https://doi.org/10.1177/0261192919899853

Tierschutzbund. (2025). *Reduktionsstrategie zu Tierversuchen nicht veröffentlicht* [Press release]. Deutscher Tierschutzbund. https://www.tierschutzbund.de/ueber-uns/aktuelles/presse/meldung/reduktionsstrategie-zu-tierversuchen-nicht-veroeffentlicht/

Vigors, B. (2018). Reducing the consumer attitude–behaviour gap in animal welfare: The potential role of "nudges." *Animals*, **8**(12), Article 232. https://doi.org/10.3390/ani8120232

Visseren-Hamakers, I. J., Razzaque, J., McElwee, P., Turnhout, E., Kelemen, E., Rusch, G. M., Fernández-Llamazares, A., Chan, I., Lim, M., Islar, M., Gautam, A. P., Williams, M., Mungatana, E., Saiful Karim, M., Muadian, R., Gerber, L. R., Lui, G., Liu, J., Spangenberg, J. H., & Zaleski, D. (2021). Transformative governance of biodiversity: Insights for sustainable development. *Current Opinion in Environmental Sustainability*, **53**, 20–28.

Williams, B. (1965). Ethical consistency. *Proceedings of the Aristotelian Society Supplement*, **39**, 103–124.

Cite this article: Müller, N. D. (2026). Why (and how) total numbers should matter for animal experimentation policy. *Politics and the Life Sciences*, **45**, 1–11. https://doi.org/10.1017/pls.2025.10008

https://doi.org/10.1017/pls.2025.10008 Published online by Cambridge University Press