Can Knowledge Itself Justify Harmful Research?

JEFF SEBO and DAVID DEGRAZIA

In our paper, we argue for three necessary conditions for morally permissible animal research: (1) an assertion (or expectation) of sufficient net benefit, (2) a worthwhile-life condition, and (3) a no-unnecessary-harm/qualified-basic-needs condition.[1] We argue that these conditions are necessary, without taking a position on whether they are jointly sufficient. In their excellent commentary on our paper, Matthias Eggel, Carolyn Neuhaus, and Herwig Grimm (hereafter, the authors) argue for a friendly amendment to one of our three conditions.[2] In particular, they argue for replacing the first condition—expectation of sufficient net benefit (ESNB)—with an expectation of knowledge production (EKP).[3] In this reply, we will explain why we are open to this proposed amendment, but not yet convinced.

As the authors explain, ESNB is satisfied when, and only when, the study in question can be expected to produce sufficient net benefit. In contrast, EKP is satisfied when, and only when, the study in question can be expected to produce knowledge worth having (about which more below). Of course, if a study produces knowledge worth having, then it might produce further benefits as well. But if we accept EKP, then we should be considering whether a study can be expected to produce the former benefit, not the latter.

It is noteworthy that the authors tacitly assume that, on our view, knowledge is not a benefit. However, we do not make this claim in our paper, and we are open to the possibility that this claim is false, either because knowledge is intrinsically good or because knowledge satisfies our desire for knowledge. For example, suppose a study provided substantial evidence that amphibians have a divided consciousness, rather than a single unified consciousness.[4] This would be extremely interesting and might be relevant to other questions about animal minds and the evolution of consciousness. Arguably, this knowledge should count as a benefit whether or not it leads to further benefits. If we were to adopt this position, then there would be greater compatibility between ESBN and EKP than the authors realize. Nevertheless, in what follows, we will assume—for the sake of a worthwhile argument—that, on ESNB, knowledge is not, in itself, a benefit (or at least is not, in itself, enough of a benefit to justify the intentional or foreseeable imposition of substantial harm on animals).

Before we consider the relative merits of ESNB and EKP against the background of this assumption, we want to emphasize how much we all agree about. We all agree that animal research is morally justified only if it meets a sufficient expected benefit condition together with other conditions. Moreover, given the connection between epistemic and practical benefits in this context (studies that have the former will also tend to have the latter), our different interpretations of this benefit condition will have similar implications in many cases. Insofar as we disagree, our disagreement is relatively minor.

Acknowledgment and disclaimer: Work on this paper was supported, in part, by intramural funds from the National Institutes of Health Clinical Center. The views expressed here are the authors' own and do not reflect the position or policy of NIH or any other part of the federal government.

Cambridge Quarterly of Healthcare Ethics (2020), **29**, 302–307.
© Cambridge University Press 2020.
doi:10.1017/S0963180119001087

https://doi.org/10.1017/S0963180119001087 Published online by Cambridge University Press

Still, the differences between ESNB and EKP are theoretically important and, at least in some cases, practically important as well, so they merit careful consideration in spite of our general agreement.

In particular, we will focus on two related arguments the authors make for replacing ESNB with EKP.

The first argument the authors make for replacing ESNB with EKP is that ESNB is, in a certain respect, *underinclusive.* That is, it will preclude approval of studies that should in fact be approved. For example, suppose researchers propose to conduct a study that will seriously harm and ultimately kill 1000 amphibians of different frog, newt, salamander, and toad species. They expect that this study will contribute to knowledge worth having about the nature of amphibian consciousness, but they have no expectation that this knowledge will contribute to any further benefits, such as health or welfare benefits for human or nonhuman animals. Instead, and at most, they see this knowledge as having the *potential* to contribute to further benefits, without yet knowing what these benefits might be, or how likely this knowledge might be to contribute to them (we will return to this complication below). In this case, ESNB would preclude approval of the study (since the expectation of knowledge worth having would not, in itself, be enough to justify seriously harming and killing all these animals on the interpretation of ESNB that we are assuming for the sake of argument). Yet, the authors suggest, precluding such a study would be implausibly strict. The expectation of knowledge worth having should, in itself, be enough to justify seriously harmful studies (assuming other relevant conditions are met) whether or not we can expect further benefits or can expect that the benefits will outweigh the costs.

The second argument the authors make for replacing ESNB with EKP is that ESNB is, in a certain respect, *overinclusive.* That is, it will permit approval of studies that should not in fact be approved. For example, suppose researchers propose to conduct a study that will seriously harm and ultimately kill 1000 amphibians. They do not expect that this study will produce knowledge worth having at all. However, they do expect that this study will lead to further studies (for example by securing continued employment and other professional benefits for the research team), and that some of these further studies will produce knowledge worth having as well as other benefits. In this case, ESNB might permit approval (since the expectation of these benefits might be enough to justify seriously harming and killing the animals, depending on the details). Yet, the authors suggest, this is implausible. The aim of research is to produce knowledge worth having (and to do so by means of research itself, not by means of the institutional side effects of research). So, the only kind of benefit that should count in favor of a study, in the context of institutional review, is the expectation that this research will produce knowledge worth having (by means of the research itself). If a study has other benefits as well, then this is great. But it should not be part of what justifies the study.[5]

We think that both of these arguments are reasonable (especially the second), but we are not fully persuaded by either of them yet, for reasons we will now explain.

Consider first the claim that ESNB is underinclusive. The authors suggest that ESNB is underinclusive because it precludes approval of studies that can be expected to produce knowledge worth having if this knowledge cannot, in turn, be expected to produce further benefits. Before we can assess this

https://doi.org/10.1017/S0963180119001087 Published online by Cambridge University Press

suggestion, we need to clarify our terms. In particular, what exactly makes knowledge worth having in the relevant sense? The authors do not answer this question, other than to say that knowledge worth having "advances an important interest."[6] Yet, no matter how we unpack this idea, this answer will raise a problem for their argument that EKP is more plausible than ESNB. To see why, consider three possible interpretations of knowledge worth having.

First, suppose we say that knowledge can be worth having in the relevant sense whether or not it has the potential to produce further benefit at all. On this interpretation, we believe, EKP would not be a plausible alternative to ESNB. For example, suppose that researchers are considering whether to conduct a study that would seriously harm and ultimately kill 1000 amphibians. Suppose further, that they know two facts about this study: first, that it can be expected to produce knowledge about the evolution of consciousness, and, second, that this knowledge cannot possibly produce any further benefit at all. (We realize that this latter assumption is unrealistic, and we will return to that point in a moment. Still, we can imagine at least some cases in which it might be warranted. For example, it might be warranted in a case where the researchers are working in a totalitarian state, and they know that the state will suppress any and all research about the evolution of consciousness.) In this case, EKP would permit approval of this study in spite of the fact that the researchers reasonably believe that it will not produce any benefits at all other than knowledge. Yet this is implausible. In order for a significantly harmful, nonconsensual, nontherapeutic study to be morally permissible, it must at least have the *potential* to produce benefits other than knowledge.

Second, then, suppose we say that knowledge is worth having in the relevant sense if and only if it at least has the potential to produce further benefit (whether or not it can be *expected* to do so). On this interpretation, EKP is more plausible than before, since it would no longer permit approval of seriously harmful studies that have no chance at all of producing benefits other than knowledge. However, this interpretation is not much different from the previous interpretation in practice, since we rarely, if ever, know in practice that a seriously harmful study can be expected to produce knowledge that, in turn, cannot possibly produce any further benefit at all. Instead, and at best, our epistemic state tends to be more like the one described initially: We expect that this study *will* be harmful, that it *might* produce knowledge, and that it *could* produce further benefit, without being sure what this further benefit might be and/or how likely it is to result. This is part of what makes science so interesting. But it is also why, if mere potential of further benefit can be enough to make knowledge worth having in the relevant sense, then EKP will permit approval of virtually any study that can be expected to produce knowledge, no matter how harmful this study is and no matter how unlikely or minor the further benefits are (a possible exception being the kind of unrealistic case described in the previous paragraph). Once again, this is implausible. We need a standard that can meaningfully constrain seriously harmful, nonconsensual, nontherapeutic research, not a standard that would simply rubber stamp all such research as long as there is at least *some* possibility that it will produce at least *some* benefit.

Third, then, suppose we say that knowledge is worth having if and only if it can be expected to produce sufficient further benefit (leaving the idea of expectation of sufficient further benefit imprecise for now; we will return to

https://doi.org/10.1017/S0963180119001087 Published online by Cambridge University Press

what it might mean below). On this interpretation, EKP is once again more plausible than before, since it can now serve as a meaningful constraint on seriously harmful, nonconsensual, nontherapeutic research. In particular, whereas on the previous interpretations, EKP would permit approval of virtually any seriously harmful study that can be expected to produce knowledge, on this interpretation it would permit approval of a seriously harmful study that can be expected to produce knowledge only if this knowledge can, in turn, be expected to produce sufficient further benefit. However, note two things. First, on this interpretation, EKP is more like ESNB, since it would require us to calculate the expected benefits of a study beyond knowledge production (which, the authors claim, we might not always be able to do). Second, however, on this interpretation, EKP, unlike ESNB, still fails to compare expected benefits with expected costs and harms. It would therefore permit approval of studies that will predictably, say, harm 100 animals for each animal that they predictably benefit. Yet this, too, is implausible. If the expected costs and harms of a study are orders of magnitude greater than the expected benefits, then the mere fact of expected benefits—even substantial expected benefits—should not be enough to permit approval of the study.

The upshot of this discussion is that we face a familiar tradeoff in our thinking about the ethics of animal research. The less net beneficial we require a study to be on expectation, the more we risk 'false positives,' i.e., net harmful studies that pass institutional review. Whereas, the more net beneficial we require a study to be on expectation, the more we risk 'false negatives,' i.e., net beneficial studies that fail institutional review. Any principle of sufficient benefit will face this issue, whether we formulate it in terms of knowledge worth having or in terms of sufficient net benefit. We will not say exactly how we should resolve this issue here. However, we can make several observations about how we should think about it.

First, the expected consequences of individual studies are difficult, but not impossible, to assess. We often have at least some information about whether there is, say, no chance, a low chance, a medium chance, or a high chance that a study will produce no benefit, weak benefit, moderate benefit, or strong benefit. We also have a variety of ethical tools that we can use in cases involving risk and uncertainty, such as an expected value principle (useful in cases where we have reliable information about probabilities and utilities, about either individual studies or sets of similar studies) and a precautionary principle (useful in cases where we do not).

Second, which epistemic standard is appropriate is plausibly a contextual matter. For example, if we know that a study will cause a lot of harm, then it is plausible that the probability or level of benefit should be relatively high. In contrast, if we know that a study will not cause a lot of harm, then it is plausible that the probability or level of benefit can be lower. Either way, we should at least consider the probability and level of benefit, the probability and level of harm, and their interaction.

Third, insofar as the appropriate epistemic standard is difficult to meet, we need to accept this outcome rather than evasively shift the goalposts. In particular, if we want to tell whether a seriously harmful study is justified, then we at least have to try to compare the expected benefits and harms, even if doing so is hard and/or leads to a negative verdict. Otherwise we will not be conducting actual ethical review but instead simply extending the appearance of ethical legitimacy to a practice that may be morally wrong.

https://doi.org/10.1017/S0963180119001087 Published online by Cambridge University Press

Finally, it is useful to remember that, when we face this kind of tradeoff in the case of human subjects research, we judge that the bar for seriously harmful, nonconsensual, nontherapeutic research on vulnerable populations is extremely, perhaps unsurpassably, high. Granted, we might think that nonhuman subjects research is relevantly different. Either way, the bar in this case should at least be higher than: "Is there *any chance at all* that this seriously harmful study will produce knowledge that produces *any benefit at all*?" And once we grant this point, we are already well on our way to a standard that resembles (or surpasses) ESNB, since we have no choice but to consider the messy, complex realities about how research interacts with the world.

Consider now, more briefly, the claim that ESNB is overinclusive. We agree with the authors that the aim of research is knowledge production, and that the expectation of epistemic benefit (and, in our view, of resulting practical benefit) should be a central part of what justifies seriously harmful research, if anything does. For example, if a researcher proposed conducting a study that would harm and kill 1000 amphibians, not on the grounds that she expected to learn anything at all but rather on the grounds that she expected to draw a paycheck, then we could all agree that the Institutional Animal Care and Use Committee (IACUC) should reject this study. Does this verdict mean that our ethical assessment of research should not take into account its expected positive side effects at all? This is less clear.

Suppose that a researcher proposes to conduct a study that would (a) harm and kill 1000 amphibians as a means to produce knowledge *and* (b) dump thousands of gallons of toxic waste into a local river as a side effect. In this case, we can all agree that the IACUC should take both of these impacts into account when deciding whether or not to approve the study (to say nothing of other costs and harms, including opportunity costs). The mere fact that the harm to animals is intended and the harm to the environment (and other kinds of harm) are merely foreseen is not enough to make it the case that the harm to animals is relevant to ethical review and the harm to the environment (and other kinds of harm) are not.

Granted, since this example involves *negative* side effects, it is not a factor for EKP, which focuses on positive effects (though, as we have claimed, we think that this is a mistake). Still, this case is a reminder that expected side effects do sometimes matter in ethical review. What we need, then, is an account of which side effects matter, how much they matter, and why.

Our current view is that all expected positive and negative impacts, intended as well as merely foreseen, should matter equally. Does that mean that IACUCs should be approving seriously harmful, nonconsensual, nontherapeutic research when the only expected benefit is, say, continued employment for researchers? Of course not. But the primary reason is not that these expected positive side effects are irrelevant, but rather that they are not nearly strong enough to justify the expected costs and harms (intended as well as foreseen) of such research.

However, we think that these questions about which effects should matter and how much they should matter are difficult and important. So, we are open to changing our minds about this issue. All that we want to emphasize for now is that there are pros and cons either way. Insofar as we consider side effects, we risk focusing too much on questions unrelated to the aim of research (which, as the authors claim, researchers tend to be worse equipped to answer) and not enough on questions related to the aim of research (which, as the authors claim,

https://doi.org/10.1017/S0963180119001087 Published online by Cambridge University Press

researchers tend to be better equipped to answer). Insofar as we do not consider side effects, we risk failing to ask questions that we need to ask in order to know whether a particular study is justified.

When we put all this together, the upshot is this. We do not think that ESNB should be replaced with a weaker requirement, that is, with a requirement that allows seriously harmful, nonconsensual, nontherapeutic studies to proceed *whether or not* they can be expected to produce knowledge that, in turn, can be expected to produce sufficient further benefit. We are open to the possibility that it should be replaced with a stronger requirement, that is, a requirement that allows a such studies to proceed only if they can be expected to produce a certain kind of benefit (such as knowledge that, in turn, can be expected to produce sufficient further benefit). But we will need to think about whether or not this revision makes sense, all things considered.

With all that said, we appreciate the authors' perspective. We share their view that researchers tend to exaggerate the societal benefits of animal research in order to justify it, and that these exaggerated claims are often difficult to evaluate in advance.[7] But in our view, the solution to this problem is not to lower the bar for passing institutional review. If anything, this solution would make it easier, not harder, for morally impermissible studies to pass institutional review. Instead, the solution is to ensure that oversight committees are empowered to assess whether a study can be expected to produce sufficient net benefit, among other issues. We need a full accounting of the probability and level of all expected costs and benefits of proposed studies, relative to alternatives, and a willingness to approve studies only if they actually pass the test (as we all agree is necessary in the case of human subjects research). This, in turn, will require systematic reform within the scientific research community, not mere revision of ESNB (especially if this revision is plausible only to the extent that it resembles ESNB).

In any case, we want to thank the authors for engaging with us on this issue. Again, no matter what the best expression of the expected benefit principle is, we can all agree that animal research should be held to a higher ethical standard than it typically is. This is what matters most. However, the details matter too, and we look forward to further discussion of exactly what details would be optimal.

Notes

1. DeGrazia D, Sebo J. Necessary conditions for morally responsible animal research. *Cambridge Quarterly of Healthcare Ethics* 2015;24 (4):420–30.
2. Eggel M, Neuhaus C, Grimm H. Re-evaluating benefits in the moral justification of animal research: A comment on 'Necessary conditions for morally responsible animal research.' *Cambridge Quarterly of Healthcare Ethics* 2020;29(1):131–43.
3. Because what is at issue in this exchange of views is the *prospective* justification of particular animal trials, rather than a general justification for animal trials, we will understand the first condition as an *expectation* of sufficient net benefit (ESNB) rather than as an assertion of sufficient net benefit (ASNB).
4. We draw this idea from a study that found frogs to have two independent—that is, non-integrated—visual systems (Ingle D. Two visual systems in the frog. *Science* 1973;181:1053–5) and a discussion in Godfrey-Smith P. *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. New York, NY: Farrar, Straus, and Giroux; 2016:89–90.
5. See note 2, Eggel et al. 2020, at 134.
6. See note 2, Eggel et al. 2020, at 137. The authors also stress that, in order to justify harmful research, the knowledge produced should be generalizable. In what follows, we will assume that the knowledge produced in all the examples we discuss is generalizable.
7. See note 2, Eggel et al. 2020, at 136.

https://doi.org/10.1017/S0963180119001087 Published online by Cambridge University Press