Special Section: Moving Forward in Animal Research Ethics

The Upper Limits of Pain and Suffering in Animal Research

A Moral Assessment of the European Union's Legislative Framework

TOM L. BEAUCHAMP and DAVID B. MORTON

Abstract: The control of risk and harm in human research often calls for the establishment of upper limits of risk of pain, suffering, and distress that investigators must not exceed. Such upper limits are uncommon in animal research, in which limits of acceptability are usually left to the discretion of individual investigators, institutions, national inspectors, or ethics review committees. We here assess the merits of the European Directive 2010/63/EU on the Protection of Animals Used for Scientific Purposes and its accompanying instruments, such as guides and examples. These documents present a body of legislation governing animal research in the European Union. We argue that the directive supplies a promising approach, but one in need of revision. We interpret the directive's general conception of upper limits and show its promise for the establishment of high-quality policies. We provide a moral rationale for such policies, address the problem of justified exceptions to established upper limits, and show when causing harm is and is not wrongful. We conclude that if the standards we propose for improving the directive are not realized in the review of research protocols, loose and prejudicial risk-benefit assessments may continue to be deemed sufficient to justify morally questionable research. However, a revised EU directive and accompanying instruments could have a substantial influence on the ethics of animal research worldwide, especially in the development of morally sound legal frameworks.

Keywords: animal research; animal suffering; animal welfare; beneficence; deontic constraints; nonmaleficence; upper limits of suffering

Introduction

Research animals often have their welfare compromised in a variety of ways due to scientific objectives. The quality of housing and husbandry conditions is limited in almost all animal research. Although the harms caused to animals in research are sometimes relatively minor, they are often nontrivial and enduring. These harms may not be reducible beyond a certain point when they are controlled by regulations or required by the research design. However, as many scientists are aware, causing pain, distress, and suffering to animals also can skew—and even invalidate—the scientific data, because of the physiological and psychological side effects of the harms inflicted.[1,2]

The control of risk and inflicted harm in animal research contrasts sharply with the control of risk and harm in human research, in which laws, regulations, and guidelines governing nontherapeutic research with human subjects who are incompetent to consent have been developed during the last 35 years in virtually

We are indebted to David DeGrazia, David Lamb, James Yeates, John Gluck, Hope Ferdowsian, Lise Bitsch, Mette Ebbesen, and Penny Hawkins for critical assessments of this article.

Cambridge Quarterly of Healthcare Ethics (2015), **24**, 431–447.

doi:10.1017/S0963180115000092

all countries. These rules often require, or at least suggest, a *threshold* (a place or point on a continuum of harming that investigators must not cross over in research) or *upper limit* (a limit to the amount of harm that investigators can cause). Despite subtle differences of meaning, for our purposes we take these metaphors of threshold and upper limit to be synonymous. These limits in human research are commonly expressed using terms such as "minimal risk" and "minor increment beyond minimal risk," which are moral categories expressing acceptable and unacceptable levels of risk.[3,4,5,6]

Comparable thresholds or upper limits of pain, suffering, and distress are not recognized in most countries for nonhuman animal subjects, who have no opportunity to decline involvement in research and normally experience some level—and, in some research, a high level—of pain or suffering. The problem of unacceptable upper limits is commonly left to the discretion of individual investigators, institutions, or ethics review committees. Little legislation or scholarly literature carefully describes upper limits, and legislation uniformly lacks an articulated moral justification of any limits imposed.

This article examines, from a moral perspective, the merits and defects of perhaps the most influential and explicit existing legislation on upper limits of animal suffering: Directive 2010/63/EU on the Protection of Animals Used for Scientific Purposes, established by the European Parliament and the Council of the European Union.[7] This EU directive (hereafter "the directive," a term we use to include its associated instruments) does not claim to present a body of moral rules. Rather, it presents a body of legal rules and a legal framework governing animal research in the member states of the European Union. However, as the directive specifically recognizes in some of its provisions, its account is based on a moral perspective; and the rules underlying the legal framework are clearly moral rules. We argue that the directive is a promising approach but is also in need of some development in its moral foundations regarding upper limits to animal suffering. At the same time, we recognize that the directive forms a legally binding piece of legislation on animal research that may be the best of its type in the world.

The European Union comprises 28 member states. Its regulations, directives, and decisions are legally binding and provide a legal framework for interstate trade to avoid unfair practices and create a "level playing field." Directives set out results to be achieved, and the mechanisms to achieve these are left to the member states through transposition into the national legislation, and each member state is responsible for its implementation.[8] The European Commission monitors to see that directives are adequately transposed and that their implementation is effective. Directives cannot contain moral provisions or guidelines as such, only legally binding provisions. However, the European Commission can adopt recommendations to complement legislation.[9] Our comments are related primarily to the moral positions found in or presupposed in the directive. We duly appreciate that a directive cannot explicitly address some of the moral problems we mention because directives are based on a legal framework for compliance and, if necessary, prosecution.

We find promise in what we will call the directive's general conception of an upper limit of harm, meaning a threshold that distinguishes acceptable from unacceptable amounts of pain, suffering, distress, and the like (including fear, anxiety, and frustration). The directive's general conception lacks support by explicit argument, but it has the starting point of a position that could be so defended. The directive

also lacks a clear, principled position on whether exceptions to an upper limit are justified in some circumstances, but once more it provides a start on such an account. We argue that amendments can and should be made to upgrade this general conception.

In the first section we interpret the directive's concise presentation of its general conception. Interpretation is necessary because the directive inadequately explains its commitments to upper limits and does too little to connect these limits to specific procedures used in animal research (even though it covers all forms of suffering from birth until death). In the second and third sections we show the directive's promise for the delineation of upper limits and its current shortcomings. In the fourth section we constructively criticize the directive's handling of the problem of justified exceptions to supposedly firm upper limits. The fifth section establishes the moral rationale for a policy of firm upper limits, and the sixth section adds a brief account of the conditions under which causing harm does and does not amount to wrongdoing.

The Directive's General Conception of Upper Limits and Moral Prohibitions

We interpret the directive's requirement that upper limits of pain, suffering, and distress not be exceeded as an embryonic system for identifying levels of acceptable and unacceptable harm. An early passage in the preamble aims at limiting levels of harm by raising the standards to protect research animals as new knowledge is gained:

> New scientific knowledge is [now] available in respect of factors influencing animal welfare as well as the capacity of animals to sense and express pain, suffering, distress and lasting harm. It is therefore *necessary to improve the welfare of animals* used in scientific procedures by *raising the minimum standards* for their protection in line with the latest scientific developments. (Preamble 6)

Preamble 6 does not delineate an upper limit, but it does state that investigators, reviewers, government officials, and the like are obligated to adjust the standards they use in determining unacceptable levels of pain, distress, and suffering as new knowledge becomes available.

The meaning of the term "scientific procedure" is critical for practical applications under the directive. "Procedure" receives a specific definition: "'Procedure' means any use, invasive or non-invasive, of an animal for experimental or other scientific purposes, with known or unknown outcome, or educational purposes, which may cause the animal a level of pain, suffering, distress or lasting harm equivalent to, or higher than, that caused by the introduction of a needle in accordance with good veterinary practice" (Article 3[1]).

The following passage in the directive is its sole provision specifically about upper limits of pain, suffering, and distress and is perhaps its most innovative moral requirement: "*From an ethical standpoint*, there should be *an upper limit of pain, suffering and distress above which animals should not be subjected* in scientific procedures. To that end, the performance of procedures that result in severe pain, suffering or distress, which is likely to be long-lasting and cannot be ameliorated, *should be prohibited*" (Preamble 23).

Immediately preceding this moral prohibition in Preamble 23 is a provision delineating a severity classification system: "To enhance transparency, facilitate the project authorization, and provide tools for monitoring compliance, *a severity classification of procedures should be introduced* on the basis of estimated levels of pain, suffering, distress and lasting harm that [are] inflicted on the animals" (Preamble 22).

In our interpretation, the directive's "ethical standpoint" functions as both a moral and a legal position on an upper limit and is directly connected to this "severity classification," although the directive does not specifically so state. The link between these two preamble statements does not emerge until the last section of the directive, an annex that provides the "classification of procedures" that are proclaimed morally mandatory in Preamble 22. Annex VIII, entitled "Severity Classification of Procedures," states that procedures in animal research should be classified as either "non-recovery," "mild," "moderate," or "severe" on a case-by-case basis. The annex also contains "assignment criteria" and attempts to state how to determine the severity of a procedure.

The severity classification is not, as it stands, merely a graded set of what the directive calls the severe category (see Annex VIII, Section I). The classification's full set of categories includes "non-recovery, mild, moderate, and severe procedures." Strictly speaking, the annex presents not a graded continuum of severity but a continuum running from mild to severe. The important matter is that the three categories of mild, moderate, and severe make Annex VIII the only point in the directive where a threshold or upper-limit line is drawn (implicitly, never explicitly). It is apparently drawn where moderate procedures end and severe procedures begin. Procedures up to and including moderate severity are presumed to be morally tolerable, though of course they are in need of justification; severe procedures that are prolonged or cannot be ameliorated are presumably prohibited. However, this interpretation of the general conception is still somewhat in doubt. Research that falls in the severe class may turn out to be justified if the procedures are not long lasting. As Preamble 23 puts it, "severe pain, suffering or distress, which is likely to be long-lasting and cannot be ameliorated, *should be prohibited*." Apparently only procedures resulting in both severe and long-lasting harms are prohibited. If so, severe harms that are not long lasting are not prohibited. (See, further, the fourth section of this article, "The Problem of Justified Exceptions to an Upper Severity Limit.")

Later in the document, Article 15 offers a modest normative extension of the statements about upper limits, severity, and what is morally prohibited. It requires the following:

> Classification of severity of procedures
>
> 1. Member States shall ensure that all procedures are classified as "non-recovery," "mild," "moderate," or "severe" on a case-by-case basis using the assignment criteria set out in Annex VIII.
> 2. Subject to the use of the safeguard clause in Article 55(3) [see subsequently in the section entitled "The Problem of Justified Exceptions to Established Upper Limits"], Member States shall ensure that a procedure is not performed if it involves severe pain, suffering or distress that is likely to be long-lasting and cannot be ameliorated. (Article 15[1–2])

Article 15 helps clarify the general conception—in particular, (1) the idea of exceeding upper limits by using severe procedures that are morally prohibited and (2) classification of the procedures that are severe. Note the recommendation that member states "shall ensure" that a procedure is "not performed": this language is best interpreted to mean that states "are obligated to ensure" that a procedure is "prohibited." The statement that a procedure is not to be performed "if it involves severe pain, suffering or distress that is likely to be long-lasting and cannot be ameliorated" is best interpreted as prohibiting such procedures when they are long lasting and incapable of being made better by ameliorating conditions. Accordingly, only enduring and unmitigable procedures in the severity category are prohibited. This way of understanding the directive brings increased clarity and coherence to its overall account, which lacks an argued or principled defense of upper limits and gives no reasons why the severe category, and it alone, identifies prohibited territory.

With the interpretation of the directive behind us, we turn now to the strengths and weaknesses of this general conception.

The Promise of the Directive for Identifying Threshold Limits

We start with some key provisions in the directive that we will hereafter assume, without argument, to be either morally correct or at least morally promising. First, we assume the acceptability of the directive's general conception of a policy that articulates which procedures cross over a threshold demarcated by the severity of procedures. We also accept its apparent view that if moderate pain, suffering, and distress are likely to be long lasting and cannot be reduced or otherwise ameliorated, they can and should be upgraded on the continuum of harm—for example, moving from moderate to severe—or, if it is prolonged, from mild to moderate—harm. Second, we assume, though here we reach beyond any explicit wording in the directive, that this document entails that legal, professional, or other policies governing animal research are deficient if they do not put into practice an upper limit of pain and suffering that is at or above the upper limit proposed in the directive. Finally, we do *not* assume that the moral claims in the directive apply only to member states, because we regard these moral standards as generalizable for animal research wherever it occurs. That is, these rules present a moral minimum for a policy of upper limits that should be observed in all animal research, irrespective of its sponsor, location, origin, or purpose—whether or not the authors of the directive conceived the general conception as universally applicable.

Critical Shortcomings in the Directive as a Moral Conception

Two problems confront the directive's account of upper limits: (1) The directive fails to specify the notion of an overall upper limit that would let an investigator or a reviewer know where in practice the threshold line should be drawn. (2) Given that exceptions to upper-limit requirements are permitted and the level of pain or suffering can be exceeded if a piece of research is important enough for human health, how can such an exception to upper-limit constraints be justified? (See especially the phrase "exceptional and scientifically justifiable reasons" in Article 55(3) [safeguard clause], quoted subsequently.) We address the first problem in this section, and the second in the following section on justified exceptions.

For thresholds to be meaningful requirements that take seriously the idea of a strict upper limit, investigators and review committees must be given guidelines that include a way of *grading*, not merely *listing*, research procedures as to their noxious, aversive, and painful properties. These guidelines would be based on prior experience of the actual impact of the procedures on the animals and would stipulate where on the graded continuum of harm an upper limit of permissible harm can be approved by genuinely independent review committees in the case of each research protocol, procedure, or series of procedures that assess cumulative severity.

In regard to committee review, the directive's listed expertise for assessments of protocols is notably limited and ideally should include other skills. Article 26 states that "the animal-welfare body shall include at least the person or persons responsible for the welfare and care of the animals and, in the case of a user, a scientific member. The animal-welfare body shall also receive input from the designated veterinarian or the expert referred to in Article 25." It is noteworthy that the designated veterinarian is not a mandated member of the animal welfare board.

An article on project evaluation states the following:

> Project evaluation
>
> The competent authority carrying out the project evaluation *shall consider* [*but it is not mandated that there be*] expertise in particular in the following areas: (a) the areas of scientific use for which animals will be used including replacement, reduction and refinement in the respective areas; (b) experimental design, including statistics where appropriate; (c) veterinary practice in laboratory animal science or wildlife veterinary practice where appropriate; (d) animal husbandry and care, in relation to the species that are intended to be used. (Article 38[3])

Unlike guidelines from some other countries (e.g., New Zealand and Australia), a member of the animal protection movement is not mentioned, nor is a member trained in ethics, nor is a citizen to help reflect the public's interest.

In regard to independence, Preamble 39 refers to the need for "an impartial project evaluation independent of those involved in the study," and Article 59, with reference to the "competent authorities," states that there must be in place "the expertise and infrastructure required to carry out the tasks; and [freedom from] any conflict of interests as regards the performance of the tasks." Clause 4 in Article 38 states, in addition, that "the project evaluation process shall be transparent … [and] shall be performed in an impartial manner and may integrate the opinion of independent parties." No mention is made of parties *external* to the institution as consultants.

The directive allows member states to upgrade the level of its protections. The states are free to retain stricter national measures if they were in force before the directive was adopted (on November 9, 2010). A member state also may proceed to expand the scope of prohibited activities. The implementation of new provisions must be stricter, not weaker:

> There is a demand in certain Member States to maintain more extensive animal-welfare rules than those agreed upon at the level of the Union. In the interests of the animals, and provided it does not affect the functioning of

> the internal market, *it is appropriate to allow the Member States certain flexibility to maintain national rules aimed at more extensive protection of animals* in so far as they are compatible with the Treaties of the European Union (TFEU)….
>
> Member States may … maintain provisions in force on 9 November 2010, aimed at ensuring more extensive protection of animals falling within the scope of this Directive than those contained in this Directive. (Preamble 7 and Article 2)

Some member states have acted under this provision. Denmark provides an informative example of stricter measures adopted to strengthen provisions about acceptable upper limits. The Danish legislation on experimental animals—the Dyreforsøgsloven—differs from the directive by upgrading the idea of a higher level beyond the directive's three levels of mild, moderate, and severe.[10] In effect, the legislation adds a fourth level that is not contained in the general conception of the EU directive. Crossing the threshold of this highest level is entirely prohibited in Denmark, and it does not matter whether the pain or suffering is long lasting. The best translation in English of the central term used for the fourth level in the Dyreforsøgsloven is probably "intense," meaning strong, severe, and extreme pain, suffering, or discomfort (J. L. Ottesen, personal communication to D.M., 2014). The legislation states that "animals must not experience intense pain, other intense suffering, or intense anxiety and must be euthanized in cases in which such conditions are likely to exist when the anesthesia wears off or palliative treatment stops working."[11] Intense pain and the like are disallowed because they exceed the threshold. If at the end of the experiment an animal is likely to remain in a state of moderate pain, suffering, or distress after the anesthetic has worn off, or if an animal is likely to suffer lasting harm (including severe impairment of function) at the intense level, euthanasia is required. In principle, no time limit is set in this provision. Causing experiences of intensity, as here described, is prohibited even if it lasts for only a few seconds. Whether this provision will be interpreted as independent of time has not yet been tested in the actual practice of research. These categories will be difficult to interpret in practice, especially because the concept of intense pain and suffering is underanalyzed in the legislation and may not turn out in practice to mean more than the level of severe pain and suffering mentioned in the directive. The Danish law also does not provide specific experiments or something comparable to the annex in the directive.

One of the authors of the present article (D.M.) carried out an informal survey of veterinarians and national legislation in various countries regarding (1) whether there is an upper limit of pain, suffering, and distress, and, if so, (2) how it is described in guidebooks or legislation. Members of the European Union follow the EU directive, but the survey showed that some form of upper limit is recognized in other countries, using a specific terminology to characterize a prohibited type of pain, suffering, or distress. The following terms are used in regard to pain, suffering, or distress: "intense"; "long lasting and unrelieved"; "very severe, long lasting," with lethality as an endpoint; "unrelievable and severe"; and "severe"; another policy described the upper limit as a "major departure from the animal's usual state of health and well-being." The survey also found that Australia, Canada, New Zealand, and the United States have no upper limit of any description.

Most of the terms used to characterize upper limits are not described in a quantitative manner that could be used in practice to terminate a study or to determine if a study should proceed in the first place based on its potential for severe pain, distress, or suffering. There are no detailed conceptual analyses of key terms and no moral argument in support of the choice of words or the selected threshold. It is left unstated how to interpret or even to find the differences between categories such as "severe," "very severe," and "intense." Another undefined term used in some countries is "intolerable," which is vague, possibly tautologous in context, and in need of interpretation for animal research. With regard to duration, phrases such as "long-lasting" and "enduring" could be analyzed in terms of specific durations of time, but they are not so treated; and their meaning from the perspective of the experience of research animals is not considered. We expect that, in the near future, severity levels, duration, and upper limits will be more carefully delineated, because the work now available is primitive by comparison to what it should be, from both a moral and a scientific point of view.

Virtually all existing legislation or policy, including the directive, lacks a careful and detailed analysis of what we regard as the two essential dimensions that must be articulated in assessment of severity of harm—namely, the *intensity* and the *duration* of the harm. These two dimensions can be conjoined with the total number of animals experiencing this degree of suffering. This category could presumably articulate what is permissible in the way of total suffering, using a utilitarian calculation, but it does not articulate deontic constraints on utilitarian reasoning or other ways in which an upper limit might be fixed. Moreover, the total suffering of animals cannot be calculated by a precise objective measurement, as it involves complex, subjective experiences. We are not, of course, suggesting that a utilitarian calculus is a decisive consideration about whether a procedure is morally acceptable or unacceptable.

When a series of procedures is carried out on an animal, the cumulative suffering is difficult to compute, because the extent and nature of a prior harm may affect an animal's perception of a subsequent harm. In an account of upper limits, these essential components should receive a detailed conceptual analysis and should be connected to an upper limit that can be measured and avoided in practice. In treating intensity, the degree, depth, or level of the pain, suffering, or distress should be characterized in a manner that allows for measurement and provides upper limits connected to these measurements. Duration is expressible in terms of the time that the pain, suffering, or distress persists, for example, seconds, minutes, hours, days, weeks, months, or years.

As a final observation in this section, we suggest that, however difficult or controversial, the measurement or assessment of actual severity (or suffering) should apply to each and every animal rather than being merely predictive based on retrospective suffering that has occurred in an experimental group previously studied. It should be based on outcome and not only on the technical procedure that has been carried out and assumed to be the same for each animal (as exemplified in the directive, Annex VIII). Some technical procedures in the mild or moderate categories, if carried out poorly, may result in higher levels of suffering. In the severe category, such poor execution may cause a level of suffering greater than "long lasting," which is not defined or explained in the directive. How suffering is measured is key, and this assessment is often based on predefined clinical signs that can be measured (e.g., percentage of body weight loss, strength of escape

strategies, and intensity of vocalization) or assessed (e.g., abdominal pain or colicky behavior), or on the fact that certain techniques may be assumed to be painful based on human experiences or experiences and assessments in other animals of the same species or in a closely related species.[12]

The Problem of Justified Exceptions to an Upper Severity Limit

Whether and, if so, how to allow for exceptions to upper severity limits is a critical—but also deficient—part of the directive guidelines. Despite its statements about upper ranges of harm that are prohibited, a prominent provision in the directive is that the upper limit of pain or suffering can justifiably be exceeded if a given piece of research turns out to be important enough for human health or some similarly important undertaking. The directive's upper limits are therefore not rigidly fixed ceilings that trump research investigations. It appears that nothing is absolutely or categorically prohibited in animal research in the EU directive generally (perhaps by contrast to Denmark, where such a prohibition does at least seem to hold). The directive draws on the same utilitarian reasoning that justifies all animal research in terms of human benefit, but its reasoning is not deontic constraint by fundamental obligations to animals. If this interpretation is correct, the concept of upper limits in the directive may seem to exert such a weak constraint that any research protocol involving animals can be justified if its benefits are at a sufficiently high level; and no account is provided of what constitutes a sufficiently high level. How, then, can the conduct of research that exceeds the upper limit genuinely be prohibited?

How Unyielding Are Firm Upper Limits?

The directive permits justified exceptions under a "safeguard clause" that could just as well be called a "justified exception clause":

> Classification of Severity of Procedures
>
> *Subject to the use of the safeguard clause* in Article 55(3), Member States shall *ensure that a procedure is not performed if it involves severe pain, suffering or distress* that is likely to be long-lasting and cannot be ameliorated….
>
> Safeguard Clause
>
> Where, for exceptional and scientifically justifiable reasons, *a Member State deems it necessary to allow the use of a procedure involving severe pain, suffering or distress that is likely to be long-lasting and cannot be ameliorated*, as referred to in Article 15(2), it may adopt *a provisional measure to allow such procedure*. Member States *may* decide not to allow the use of non-human primates in such procedures. (Articles 15[2] and 55[3])

Even the causing of long-lasting and severe pain, suffering, or distress can be justified under this safeguard clause, though the exception must be "provisional." Articles 15 and 55, in effect, assert that there is no absolutely binding upper limit.

The directive is weak on the meaning of a "provisional measure" and on which conditions in particular justify exceptions to upper limits. However, it is fairly

clear about two related matters. First, the exception may be sought if (and, presumably, only) "it is essential for the preservation of the species or in relation to an unexpected outbreak of a life-threatening or debilitating clinical condition in human beings." Under this provision, a provisional measure may use any species, including great apes (see Article 55[2], quoted previously). Second, exceptions must be justified in individual cases based on a balancing of the risks and possible benefits of the research—where the risks are to the animal subjects, and the benefits are likely to be exclusively for human populations. Various passages in the directive dealing with harm-benefit assessments seem to reduce to the now widely accepted position that although suffering should be as limited as possible, under exceptional conditions the limit can be justifiably overridden under the condition of a sufficient anticipated benefit.

The directive's clearest general statement of this idea about justification is found in an article that does not specifically discuss exceptions to upper limits: "The project evaluation shall consist in … (d) a harm-benefit analysis of the project, to assess whether the harm to the animals in terms of suffering, pain and distress is justified by the expected outcome taking into account ethical considerations, and may ultimately benefit human beings, animals or the environment" (Article 38, Project Evaluation 2).

We interpret the directive to hold that exceptions to nonabsolute upper limits of pain and suffering are justified if critically important benefits are deemed probable after a thorough scientific and moral investigation by an impartial project evaluation. A member state therefore can make an exception and can override the upper-limit threshold of severe pain, suffering, or distress even when the harm is likely to be long lasting and cannot be ameliorated. The directive also does not appear to rule out the possibility that vastly important research might be justified even if anticipated benefits are to some extent improbable—for example, somewhat less than fifty percent likely to succeed. On this interpretation, the conception of overall expected benefit should be understood in terms of both the quantity of possible benefit and the probability of its realization.

Subsequent clauses state that such exceptions must be referred back to the commission, which must then reach a decision. Final decisionmaking must be referred to a committee of experts from each member state, which by a qualified majority[13] will decide to accept or reject the exception. The commission then authorizes the exception for a defined period or revokes it (Article 55[4]). Without commission authorization, the research may not proceed, no matter how scientifically justifiable the applicant believes it to be in light of expected benefits to human populations. Requiring this form of oversight is critical from a moral point of view, but a provision should be added to the directive to the effect that the membership of the committee that will authorize or forbid the research must be constituted so that the group as a whole has a thorough understanding of both the science and the ethics involved, as well as an understanding of public concerns about animal use and public priorities.

When acting under an authorized exception, it remains morally essential in the directive for investigators to cause only the minimum amount of pain, suffering, or distress needed to achieve the scientific objective of a promising benefit. Investigators must seek to limit the level of pain, suffering, distress, and lasting harm to the lowest possible level, using the fewest number of animals possible.

This goal conforms to two of the three Rs (replacement, reduction, and refinement) put forward by Russell and Burch in 1959,[14] which underpin much existing legislation on animal research worldwide.[15]

An Evaluation of the Directive's Provisions on Justified Exceptions

The directive's solution to the problem of justified exceptions to upper limits may seem a disappointing failure to make good on its promise of prohibited activities. It might even be concluded that the directive has no upper limits because they can, in principle, always be exceeded. However, the directive cannot be so easily faulted. That there are both valid principles of obligation and justifiable exceptions to those principles is a standard approach in many areas of ethical deliberation. This view must be given some place in a policy of upper limits in animal research, just as it has a place in human research.

For example, few, if any, moral rules are more important when using human research subjects than the obligation to obtain informed consent. But various situations in medicine and public health justify overriding this requirement in emergencies, especially in extreme public health emergencies.[16] Likewise, severe restrictions on a person's movements by use of quarantine can be justified in order to prevent the spread of an infectious disease such as Ebola. A prima facie (or, alternatively, pro tanto) obligation in law and ethics is not always an actual obligation, and it must be determined in light of the full set of obligations and their weight in a circumstance. Both limitations of liberty and risk of harm may be justified when there is a serious threat to health that can only be controlled by such measures.

It would be preferable if the directive contained a carefully crafted body of articles that specify what counts as a justification for overriding an upper limit, but there is no such statement. The directive does provide an interesting example of allowable exceptions to general rules, but the example is directed at the exceptional use of species normally forbidden to be used in research and is not directed at exceptions to prohibited upper limits for a type of procedure:

> The use of great apes, as the closest species to human beings with the most advanced social and behavioural skills, should be permitted only for the purposes of research aimed at the preservation of those species and where action in relation to a life-threatening, debilitating condition endangering human beings is warranted, and no other species or alternative method would suffice in order to achieve the aims of the procedure. The Member State claiming such a need should provide information necessary for the Commission to take a decision. (Preamble 18)

Although not about upper limits, this statement expresses how what is ordinarily impermissible becomes permissible. The circumstance "where action in relation to a life-threatening, debilitating condition endangering human beings is warranted, and no other species or alternative method would suffice in order to achieve the aims of the procedure" is the central condition that justifies exceptions.

A similar conclusion was reached in December 2011 by a U.S. Institute of Medicine committee that was largely composed of research scientists. Its report, which became U.S. federal policy, held that use of chimpanzees in biomedical research

was no longer justified in research sponsored by the National Institutes of Health (NIH), unless three conditions could be met:

1. The knowledge gained must be necessary to advance the public's health.
2. There must be no other research model by which the knowledge could be obtained, and the research cannot be ethically performed on human subjects.
3. The animals used in the proposed research must be maintained either in ethologically appropriate physical and social environments or in natural habitats.[17,18]

The primary point of this report is to state why use of chimpanzees in research should *not* be allowed, but these three conditions form a framework for exceptions to the otherwise prohibited use of chimpanzees in NIH-funded research (though the general moral conclusions in this report reach beyond particular sources of funding). These exceptive conditions—which have not as yet been invoked and are not expected to be invoked—provide a good starting framework of the sort that we hypothesize will be needed to fill out the account of justified breaches of the upper limits demanded in the EU directive (although some of the points arguably may be covered in the directive, using different terminology).

Another example of a "prohibition" is in the use of neuromuscular blocking agents without an anesthetic or analgesic cover—a proscription widely found in international legislation on animal research (see Article 14[3] of the EU directive). The reason for outlawing this excessive harm is moral rather than scientific. When blocking agents are used, an animal is paralyzed but still has sensory awareness. If the agents are used, guidelines state that animals must be monitored to ensure that an adequate plane of anesthesia is maintained to eliminate sensory awareness. Nonuse of this adequate plane is a violation of an exceptionally firm threshold that must not be exceeded except in truly unanticipated emergency circumstances.

General principles and rules typically admit of at least some exceptions (e.g., killing and possibly torture), even if a few rules such as "do not enslave" and "do not rape" are exceptionless. Almost all general principles can be justifiably overridden in some circumstances by other moral norms with which they come into contingent conflict. Principles, obligations, and rights are not unconditional merely because they are universally valid. However, all obligations must be acted on unless they conflict on a particular occasion with another obligation that is of overriding importance. One's actual obligation is then determined by an examination of the respective weights of the competing obligations.

Nonetheless, demanding standards of justification should be in place to show that an exception is warranted. It is particularly important to state (1) that justified exceptions are for truly extreme circumstances not expected to arise in the ordinary course of research and (2) that final review of the proposed exceptions must be done by a truly independent review team of impartial judges (though we recognize the considerable difficulties that lie in the way of stating the conceptual conditions of impartiality and in locating truly impartial persons). The directive's requirement that "an impartial project evaluation independent of those involved in the study should be carried out as part of the authorization process" states a necessary condition of a credible process of authorizing exceptions to a threshold limit, but it fails to include another necessary condition: reviews should be

conducted only by a group of people with a range of expertise in the relevant science, ethics, and public policy.

We also suggest that all exceptions should be disclosed to the general public and that these disclosures should be accompanied by the moral argument used to justify the exception in each specific case. The justification presented needs to incorporate a realistic evaluation of benefits. Here profoundly difficult problems arise about the knowledge base required to accurately predict the probability of obtaining desired benefits in detail, even if the overall objectives may be clear.

The Moral Rationale for a Policy of Firm Upper Severity Limits

We now step back from the directive and ask why, from a moral point of view, strict upper-limit conditions such as the general conception proposed in the EU directive are needed. To answer this question, we use the language of two general moral principles that are widely invoked in both human research and animal research: nonmaleficence and beneficence.

The *principle of nonmaleficence* requires that we refrain from actions that cause harms to others. It is, or at least should be, noncontroversial that many interventions that cause harms in the form of pain, suffering, distress, or enduring impairment of key functions in humans will have a similar effect on other sentient animals, given their similarity to humans in the capacity for such experiences. For research animals, as for humans, pain is pain, suffering is suffering, and distress is distress, wherever they occur—in animal laboratories no less than human healthcare centers. As levels of these harms increase, they could reach the level of brutal, inhumane, and merciless actions. The more investigations approach these levels, the more a policy of firm upper limits is needed. The history of research on animals and humans down to the present day has shown how easily obligations to abstain from harming can disappear from view, causing both human and animal research to be viewed, with good reason, as sometimes at the level of the brutal and inhumane.

The *principle of beneficence* has also played a major role in the development of protections against harm in both research involving human subjects and research involving animal subjects. Research institutions and investigators are obligated to take actions to reduce and limit the pain and suffering they cause and to provide for the welfare of the subjects under their care. It is a beneficent action to provide ample space for animals to move freely and, if they are social animals, to interact with a group that provides species-typical relationships. To say that this action is beneficent is not to say that the animals are better off than they would be if they lived in a different situation, nor is it to suggest that benefits for the animals have been maximized. It is only to say that the action is to their benefit. Many recommendations and policies in animal research ethics incorporate appeals to some form of beneficence, even if only implicitly. For example, when laboratories are criticized for having inadequate housing and for failing to meet basic needs, the goal of the criticisms is to assert that those in charge of the laboratories are morally obligated to supply better conditions and are morally deficient if they fail to do so.

Principles of nonmaleficence and beneficence can both be invoked in defense of this position. Investigators arguably are obligated to not cause harm to animals by depriving them of their basic needs, and they also have continuing obligations to supply conditions supportive of good welfare, such as adequate space for

movement, exercise, play, and desirable natural behaviors such as foraging, appropriately protective housing, access to compatible conspecifics (for social animals), hydration and nutritious food, veterinary care, and the like.

The term "beneficence" is sometimes conceived as a normative but optional ideal that is more like charity, humanity, or decency, rather than being seen as a matter of genuine moral obligation. However, obligations to set upper limits or thresholds of the sort considered in this article are not optional, charitable ideals or self-assumed principles of stewardship. They are basic, nonoptional obligations owed directly to animals. The now century-old history of government and institutional struggles with animal research and welfare guidelines shows that general considerations of nonmaleficence and beneficence underlie the moral obligations that all research investigators have to animals under their care, not merely the obligations that they have to institutions or to sources of funding for the research (and to the public, which directly or indirectly funds most animal research and testing).

Minimization of pain, suffering, and distress is among the main objectives of the institution of morality and is arguably its single most important objective, though we do not here defend this proposition. Our concern is that minimizing animal use and suffering in research should be among the main objectives of research institutions. The general principles of nonmaleficence and beneficence are not confined to a particular species, such as human beings, great apes, or companion animals—though much in moral treatises and in law presumes that certain species, especially the human species first and great apes second, are owed preferential treatment. General moral principles, in their abstract form, place no restriction on the range of individuals affected, and therefore they do not exclude particular classes or species such as birds, rats, and mice; nor do they apply only to species with a certain genetic makeup. As a substantive matter, species should make a difference only if there is a morally relevant difference that is species based, and one would have to show by evidence and argument in any given case that such a morally relevant difference does exist. It is open to argument that moral principles of beneficence and nonmaleficence apply only to humans, but the claim must be argued, not merely asserted, as has traditionally been the case; and we believe that the most penetrating work in animal research ethics undermines rather than supports this claim.

The moral integrity, capacity for sympathy and empathy, and sense of accountability of research scientists should also be considered as part of the moral landscape. The moral character of persons who are able to inflict intense, long-lasting pain or suffering on an animal in research calls for moral evaluation and, of course, is subject to critical scrutiny. All investigators who conduct research on animals must take responsibility for their actions, as must institutions that sponsor or conduct such research.

Firm upper limits acknowledge that some information must be foregone in both animal research and human research if the only way to obtain it is by the infliction of excruciating pain, suffering, and distress. In many cases information has been obtained in unethical ways, though they perhaps were considered ethical at the time. Total pancreatectomy leading to diabetes is a good example. Pancreatectomized animals not only later died of diabetes but were unable to digest food due to the lack of digestive enzymes. In other cases, experiments on learned helplessness in dogs and primates were used to develop drugs against depression. These models

caused severe distress in the animals—from inescapable electric shock, treadmills, swimming fatigue, and the like. These experiments were banned in the U.K., but, subsequently, drugs were developed using these models and were imported to treat human patients. The problem with this research was its severity in terms of the intensity of the stress and also its duration and sometimes repeated exposure, in a manner similar to the use of neuromuscular blocking agents with repeated exposure. This research beyond the upper limit should have been foregone but was not.

Harm and Wronging in Research with Animals

As the previous section on justified exceptions suggests, the term "harm," as used in our discussions of causing pain and suffering, does not entail a wrongful injuring or invasion of interests; nor does it refer to an intentional harm or a moral evil. "Harm" refers to a thwarting, defeating, or setting back of a physical or mental interest of an individual, whether or not it was inflicted deliberately.[19] Causing or failing to alleviate pain, suffering, distress, or enduring impairment of function in an individual is to cause or fail to prevent a setback to critical interests.

One basic moral question about research involving animals asks when the causing of such harm is justified. Where no adequate moral justification exists for an inflicted harm, the individual animals affected have been morally wronged. However, if causing a harm is justified, then the action may be the right action to perform despite the harming that occurs. A complication is that, even if a harm-causing act is the right act to perform in the circumstances, the actor may still be causing harm in a manner that understandably elicits regret or remorse because the agent has not been able to discharge a firm obligation to the subject(s) harmed. One may appropriately feel remorse even while believing that a justified action was performed under the circumstances.

One way of thinking about problems of upper limits is that if it is not justified to exceed fixed upper levels of pain, suffering, and distress with nonconsenting human subjects in nontherapeutic research, and if animal subjects are relevantly similar to human subjects in the relevant respects, then exceeding the same levels of pain, suffering, and distress would likewise not be justified in the use of animals in research—or at least a justification would be required to show why what is unjustified with human subjects is justified with animal subjects. When human interests and animal interests are relevantly similar and their welfare is contingent on not being constrained, coerced, deprived of basic needs, and placed in pain or terror, it is difficult to see what, if anything, would justify treating the interests of animals as dissimilar to human interests.

Animals have been so treated on many occasions because societies allow the animals' interests to count for less than human interests, especially if there is reason to believe that a significant improvement in some area of human health requires animal suffering during testing. This assessment of so-called scientific necessity is frequently mistaken, as we noted earlier in discussing chimpanzee research. Moreover, many supposed benefits that have regularly been taken to justify research with animals have been directed not to either human or animal health but to basic research and to discovering more about the biology of body functions. Such research could turn out to be valuable for human and animal health, but it often has had only a tenuous connection to health benefits at the time it is proposed for approval.

Conclusion

If the ways of handling problems of upper limits we have proposed are not taken to heart in legislation and the review of research protocols, loose and prejudicial risk-benefit assessments may continue to be deemed sufficient to justify morally questionable research. One of our aims in assessing the EU directive and its accompanying instruments has been to highlight some areas that need further development from a moral point of view. In particular, more precise definitions of "intensity" and "severity" are needed, and clearly articulated upper limits should be established on the allowed intensity of pain, suffering, and distress. Finally, it should be ensured that specifications and enforcement provisions are in place for a full and genuinely independent and impartial review. A revised directive, so amended, should prove worthy of acceptance in animal research worldwide.

Notes

1. Morton DB. A systematic approach for establishing humane endpoints. *ILAR Journal* 2000;41: 80–86.
2. Akhtar A. The flaws and human harms of animal experimentation. *Cambridge Quarterly of Healthcare Ethics* 2015;24:407–419.
3. United States Code of Federal Regulations, at 45 CFR 46 Subpart D (§46.404 and §46.406), contains an influential regulation governing nontherapeutic research involving children that limits risk to a "minor increase over minimal risk."
4. National Human Research Protections Advisory Committee (NHRPAC). Clarifying a specific portion of 45 CFR 46 subpart D that governs children's research; available at http://ctep.cancer.gov/investigatorResources/childhood_cancer/docs/nhrpac16.pdf (last accessed 7 Oct 2014).
5. Rid A. Setting risk thresholds in biomedical research: Lessons from the debate about minimal risk. *Monash Bioethics Review* 2014;32:63–85.
6. Binik A. Why the debate over minimal risk needs to be reconsidered. *Journal of Medicine and Philosophy* 2014;39:387–405.
7. Directive 2010/63/EU on the Protection of Animals Used for Scientific Purposes. *Official Journal of the European Union* L 276/33; adopted 2010 Sept 22; available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32010L0063 (last accessed 26 Dec 2014). Quotations from this directive, as presented herein, often have italics added for emphasis.
8. European Commission. Applying EU law. *The European Commission at Work*; available at http://ec.europa.eu/atwork/applying-eu-law/index_en.htm; see also http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012E/TXT (both last accessed 19 Feb 2015).
9. National Competent Authorities for the Implementation of Directive 2010/63/EU on the Protection of Animals Used for Scientific Purposes. *Working Document on Project Evaluation and Retrospective Assessment*. Brussels; 2013 Sept 18–19; available at http://ec.europa.eu/environment/chemicals/lab_animals/pdf/Endorsed_PE-RA.pdf (last accessed 19 Feb 2015).
10. Lov om Dyreforsøg [Act on Animal Testing, Animal Studies]; available at https://www.retsinformation.dk/Forms/R0710.aspx?id=162938 (last accessed 11 Jan 2015). See §7 stk 4 (Article 7, point 4).
11. Available at http://forsoegsdyrenes-vaern.dk/forsog-pa-dyr/lovgivning/bekendtgorelse-af-lov-om-dyreforsog/ (last accessed 28 Jul 2015).
12. National Competent Authorities for the Implementation of Directive 2010/63/EU on the Protection of Animals Used for Scientific Purposes. *Working Document on a Severity Assessment Framework*. Brussels; 2012 July 11–12; available at http://ec.europa.eu/environment/chemicals/lab_animals/pdf/Endorsed_Severity_Assessment.pdf (last accessed 28 Jul 2015); accompanying examples available at http://ec.europa.eu/environment/chemicals/lab_animals/pdf/examples.pdf (last accessed 28 Jul 2015).
13. Europa. Qualified majority. *Summaries of EU Legislation*; available at http://europa.eu/legislation_summaries/glossary/qualified_majority_en.htm (last accessed 19 Feb 2015).
14. Russell WMS, Burch RL. *The Principles of Humane Experimental Technique*. London: Methuen, 1959; reprinted: Potters Bar, UK: Universities Federation for Animal Welfare and Dover; 1992.

15. See, for example, Australian Government, National Health and Medical Research Council. *Australian Code for the Care and Use of Animals for Scientific Purposes*. 8th ed. Canberra: National Health and Medical Research Council; 2013. This Australian code sets no upper limit, but it consistently demands that researchers minimize pain and distress.
16. Meisel A. The "exceptions" to the informed consent doctrine: Striking a balance between competing values in medical decisionmaking. *Wisconsin Law Review* 1979;1979:413–88.
17. Committee on the Use of Chimpanzees in Biomedical and Behavioral Research, Institute of Medicine. *Chimpanzees in Biomedical and Behavioral Research: Assessing the Necessity*. Washington, DC: National Academies Press; 2011; available at http://iom.edu/~/media/Files/Report%20Files/2011/Chimpanzees/chimpanzeereportbrief.pdf (last accessed 25 Dec 2014).
18. Council of Councils, National Institutes of Health. *Council of Councils Working Group on the Use of Chimpanzees in NIH-Supported Research: Report*; 2013 Jan 22; available at dpcpsi.nih.gov/council/pdf/FNL_Report_WG_Chimpanzees.pdf (last accessed 28 Oct 2014).
19. Feinberg J. *The Moral Limits of the Criminal Law, Vol. 3: Harm to Self*. New York: Oxford University Press; 1986.