Ethics in experiments on live cattle: a pragmatic approach

Isabelle Veissier, Véronique Deiss, Mette Herskin, Emer Kennedy, Kenny Rutherford

▶ To cite this version:

HAL Id: hal-03137788

https://hal.inrae.fr/hal-03137788

Submitted on 10 Feb 2021

Methods in cattle physiology and behaviour research –
Recommendations from the SmartCow consortium

ZB MED Publication Portal
for Life Sciences

PUBLISSO

www.publisso.de

Ethics in experiments on live cattle: a pragmatic approach

Isabelle Veissier[1]
Véronique Deiss[1]
Mette Herskin[2]
Emer Kennedy[3]
Kenny Rutherford[4]

[1] Université Clermont Auvergne, INRAE, Saint-Genès-Champanelle, France

[2] Department of Animal Science, Aarhus University, Tjele, Denmark

[3] Teagasc, Animal and Grassland Research and Innovation Centre, Fermoy, Ireland

[4] Animal Behaviour and Welfare, SRUC, Roslin, United Kingdom

Introduction

There are ongoing philosophical, moral, and societal debates about experiments on live animals. Along with a consequentialist approach, it may be considered that an action – here an experiment on live animals – is morally acceptable if the knowledge it aims to produce can result in an overall benefit. In other words, it is acceptable if the constraints imposed on the animals involved are outweighed by the larger benefits expected for others (humans or animals). Alternatively, it can be argued that, because animals are sentient beings, they have a right not to be used by others. There may also be points of view in between, considering that animal experimentation is acceptable on the condition that no alternative ways to gain the relevant knowledge are possible and animal suffering is limited as much as possible. What is considered as an important knowledge that justifies experimentation may differ as it refers to individual values. It is not the intention of this chapter to hold a moral debate about animal experimentation (for reviews, see for instance [1], [2]). Rather, we take a pragmatic approach that we hope will guide experimenters in their decision to undertake an experiment (or not), after having made explicit all arguments for or against that experiment.

A – Reducing animal suffering: The 3Rs approach

In 1959, the zoologist William Russell and the microbiologist Rex Burch published *The Principles of Humane Experimental Technique* [3]. These principles are known as the 3Rs: Replacement, Reduction, and Refinement. The 3Rs are now used worldwide, including in legislation (e.g. the EU Directive 2010/63/UE, Article 1).

1. Replacement

Russell and Burch insisted on the need to strive to replace sentient species by non-sentient species or by non-living models, e.g. cells, tissues or computer models. For instance, ruminal digestion processes have often been studied by fixing a cannula on the animals' rumen in order to monitor how feeds are digested by the rumen microbiota. *In vitro* fermentation techniques have been developed where small amounts of feeds are incubated with ruminal fluid (Example 1 detailed below). Although obtaining ruminal fluid still requires that few animals are cannulated this approach means that many plants or compounds can be tested *in vitro* [4]. Another possibility to replace experiments is to collect data from farms. Such an approach is largely used by epidemiologists, e.g. to identify risk factors for a disease or the kinetics of its transmission. A similar approach has been used to study factors affecting the welfare of dairy cows [5]. Nowadays, many farms are equipped with sensors so that data can be collected without any handling of the animals. This offers large datasets that can be processed for research purposes, as done by Veissier et al. [6], that identified patterns of cows' activities in relation to their health status. Finally, meta-analyses of published results can also be used to clarify mechanisms [7].

Example 1: Replacing *in vivo* studies on ruminal digestion by *in vitro* fermentation (personal communication from V. Niderkorn, INRAE)

Often, studies on ruminal degradation of feeds are performed using *in situ* methods, with nylon bags containing feeds incubated in the rumen. The substrates are diluted in the rumen and do not interfere with fermentation. Such a technique requires the animals to be fixed with a canula in the rumen, which is a very invasive method. In addition, measuring ruminal fermentation parameters and methane emissions is done using *in vivo* experiments with cannulated animals and heavy equipment, respectively. We now increasingly use *in vitro* fermentation, mimicking digestion within the rumen. Plant substrates are incubated in anaerobic conditions at 39°C in culture bottles containing buffered rumen juice from sheep or cows (Figure 1). After 24 h of fermentation, parameters such as pH, total gas and methane production, dry matter disappearance, concentrations and profiles of total volatile fatty acids, and ammonia in the fermentation medium, are measured (Figure 2). This method requires only three animal donors of rumen fluid which are cannulated, and can provide information on the rumen digestion of plant substrates or on effects of specific bioactive compounds or additives. *In vitro* fermentation can thus be used in a screening approach, by which many plants, compounds, or additives can be tested, allowing selection of the most promising ones according to the objectives of the study. Only these may be tested *in situ* or *in vivo* (that is provided to the animals in the food) for further validation.
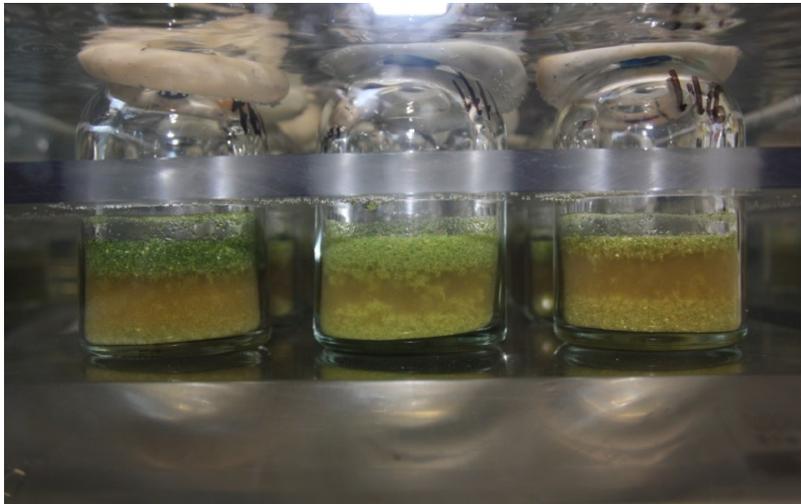

Figure 1: Batch fermenters containing forage substrate and buffered rumen fluid (photo by V. Niderkorn)
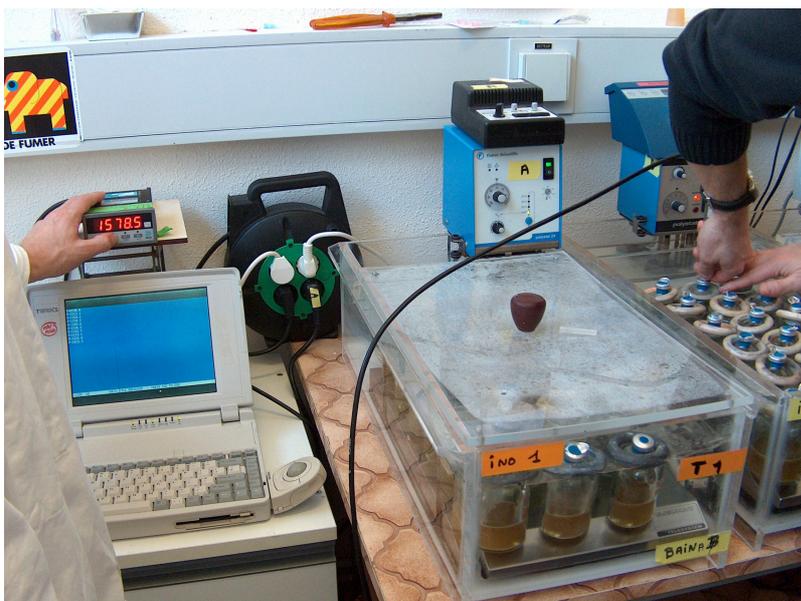

Figure 2: Measuring gas production using a pressure transducer (photo by V. Niderkorn)

2. Reduction

According to Russell and Burch, in the absence of full replacement, researchers must limit the use of sentient animals to experiments considered essential. What is considered essential remains under question (see discussion in the next section). One can at least limit the number of animals in an experiment to the minimum needed to ensure meaningful results. This number can be estimated by taking into account the variability between individuals in a population and the expected effect of a treatment, as well as the risks one is ready to take when rejecting or accepting a hypothesis (alpha and beta risks), in a so-called *power analysis* [8]. The inter-individual variability can be estimated *a priori* from pilot studies or from the literature on the condition that similar measurements have already been reported. Alternatively, one can check *a posteriori* that the number of animals were sufficient to highlight effects of treatments applied in an experiment, given the observed variability and the magnitude of differences; if that number is not sufficient, then the experiment can be extended on more animals. Other authors call for a *New Statistics* approach whereby the null hypothesis significance test is replaced by estimates of effect sizes and confidence intervals [9].

Special attention shall be drawn to experimental designs and statistical models to ensure that the analyses will be powerful and is able to highlight significant effects. The power of a statistical analysis can be increased by reducing the error term. Therefore, limiting errors due to measurements (by using reliable measurement methods) is essential. In addition, the variability between animals can be reduced by using animals of same age, weight, production, and genetic background. However, the effort to reduce animal numbers may limit translation of findings to relevance under real-life conditions, if – for instance – only one sex or breed are represented in a study, this may lessen how generalizable the results are to a wider population. The inherent trade-off here is often unavoidable but experimenters should at least be aware that it exists when making experimental design decisions. To overcome some diversity among animals, it is important to take full account of all factors that can affect results, e.g. by adding covariates such as the genetic potential. In case of a study using repeated measurements, choosing appropriate matrix correlation also contributes to increase the chance to highlight significant differences between treatments.

3. Refinement

Refinement refers to the use of methods that alleviate or minimize potential pain, suffering or distress, enhance animal welfare, and seek to optimise model validity. This requires defining good practices to perform specific measurements whereby the constraints on animals are minimised. For instance, digestibility measurements are often performed by measuring closely the intake of animals and their excretions in faeces, which requires the animals to be kept in small individual crates where faeces can be collected apart from urine. 'Four-star' digestibility crates were designed to offer as little discomfort as possible to the animals while allowing precise collection of samples ([10], Example 2) and studies are being carried out to reduce the duration of measurements in such stalls. Whenever possible, a more severe procedure should be replaced by an alternative method with less negative impact on the animal, e.g. digestibility measurements in individual crates can be replaced by the use of proxies in blood and faeces ([11], [12], Example 3). In addition, refinements to the overall life experience of the animal can often be made whilst they are not on study; this includes ensuring that living conditions allow good welfare, e.g. physical comfort, possibility to express behaviour (incl. enough space), appropriate social groups, absence of stress (incl. a high standard of human handling, perhaps alongside appropriate training or habituation to experimental procedures and measurements), and good health.

There may be antagonisms among the 3Rs, e.g., reducing the number of animals results in more suffering for the remaining animals. In that case, it is usually recommended to prioritise refinement against reduction [13].

Apart from aligning research projects to moral values, large advantages of following the 3Rs principles have been obtained, from limiting the cost of research by using alternatives to animals or reducing their number to obtaining more reliable results by reducing the constraints imposed on animals. Indeed, stress responses – whether due to poor handling, social isolation or another cause – interfere with the metabolism and immune functions and this can distort results of experiments [14]. If a painful procedure is applied, the pain should be minimised by using painkillers – unless pain is the focus of the study. Some scientists argue that painkillers interfere with the results of experiments. Actually, pain also can affect the results of experiments. Peterson et al. [15] reviewed the effects of analgesics and of pain in common

animal models used in therapeutic studies and provided guidance on the decision whether or not to use painkillers. Overall, painkillers must be used when a painful procedure is used, unless the experimenter can demonstrate that withholding of painkillers is necessary for the study.

The best advantages were obtained when there is an institutional strategy to commit to the 3Rs. Indeed, analysing projects from pharmaceutical industries, Tornqvist et al. [16] showed that applying the 3Rs can result in a large decrease in the number of animals used, thanks to improvement of the design of experiments, refining methods of investigation, and better coordination of research projects. The authors concluded that: "*in silico*-, *in vitro*- and *in vivo*-methods all hold the potential for applying the reduction R and should be consequently coordinated at a strategic level".

Example 2: 'Four-stars' stalls to measure digestibility in cattle (personal communication from H. Tournadre, INRAE)

For 'gold standard' measurement of digestibility, cattle are maintained in stalls restraining their movements to separately collect faeces and urine. We designed a device to improve the housing of cattle for digestibility measurements so that the constraints for the animal are reduced while the sampling of faeces and urine are easier and more precise (Figure 3).
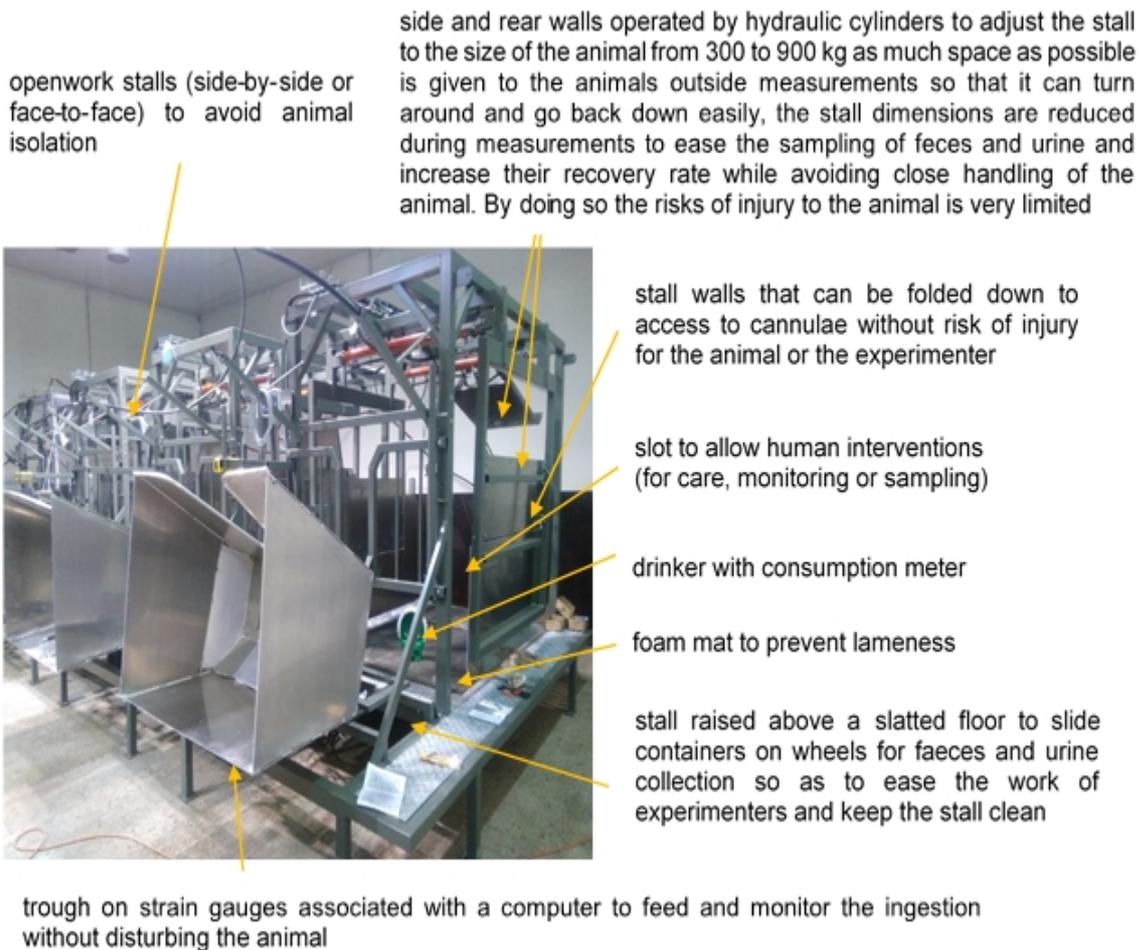

openwork stalls (side-by-side or face-to-face) to avoid animal isolation

side and rear walls operated by hydraulic cylinders to adjust the stall to the size of the animal from 300 to 900 kg as much space as possible is given to the animals outside measurements so that it can turn around and go back down easily, the stall dimensions are reduced during measurements to ease the sampling of feces and urine and increase their recovery rate while avoiding close handling of the animal. By doing so the risks of injury to the animal is very limited

stall walls that can be folded down to access to cannulae without risk of injury for the animal or the experimenter

slot to allow human interventions (for care, monitoring or sampling)

drinker with consumption meter

foam mat to prevent lameness

stall raised above a slatted floor to slide containers on wheels for faeces and urine collection so as to ease the work of experimenters and keep the stall clean

trough on strain gauges associated with a computer to feed and monitor the ingestion without disturbing the animal

Figure 3: Four-star stalls to measure digestibility. The access is facilitated by an inclined plane (photo by H. Tournadre).

Example 3: Proxies of total tract digestibility and urinary N from faeces and blood samples (personal communication from G. Cantalapiedra-Hijar et D. Andueza)

Total tract digestibility and nitrogen (N) excretion are often measured by separately collecting faeces and urine from cattle placed in so-called metabolic cages for at least 10–14 d. We propose alternative methods to predict total tract digestibility and urinary N excretion that require only restraining animals for a few minutes to collect faeces, blood, or milk. The formula to calculate total digestibility or N excretion will be available soon (articles in prep.).

We propose to assess total tract digestibility by faecal visible/near infrared spectroscopy (VIS/NIRS). Faeces are obtained by hand rectal collection while animals are temporally immobilized by head holding devices. The spectra obtained from faeces allow diets to be discriminated according to their total tract *organic matter digestibility* (OMD, Figure 4). It may also reflect inter-individual variations in digestibility from animals fed the same diet.
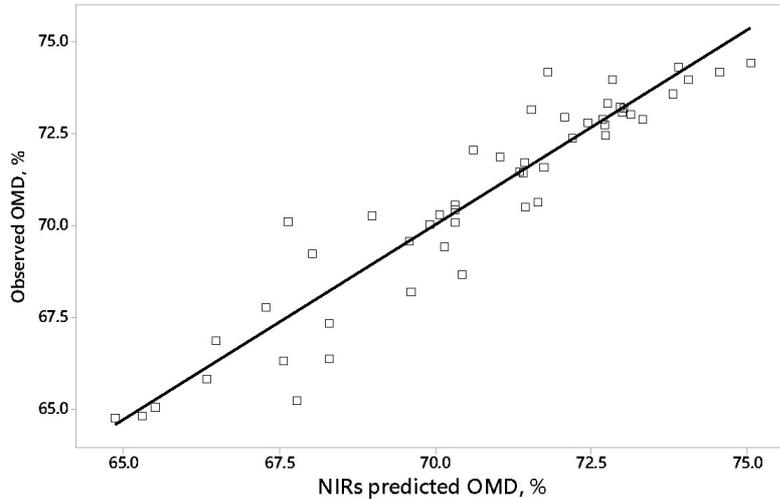

Figure 4: Relation between organic matter digestibility predicted by NIRS from faeces and measured organic matter digestibility by the 'gold standard' method using metabolic cages. Each symbol represents an experimental diet within a particular study. r^2=0.89 (P<0.001).

We propose to assess urinary N excretion by two biomarkers: the natural ^{15}N enrichment of animal proteins (in plasma or milk) over the consumed diet and the urea concentration (in plasma or milk) (Figure 5). Both biomarkers vary with N partitioning between anabolism (animal proteins in body or milk) and catabolism (urinary N). Measuring these biomarkers require blood sampling, e.g. by venepuncture of the coccygeal vein, and milk sampling, taken during conventional milking.
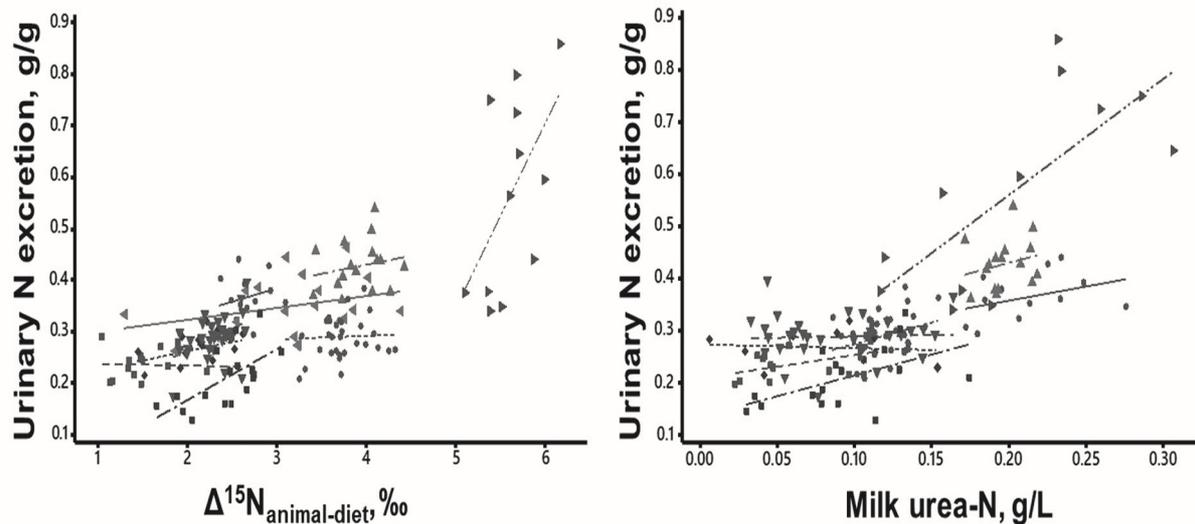

Figure 5: Relation between urinary nitrogen excretion measured by 'gold standard' methods using metabolic cages and either natural ^{15}N enrichment of animal proteins over the diet ($\Delta^{15}N_{animal-diet}$) or milk urea nitrogen concentration.

B – How to assess and minimise harms imposed on animals

The 3Rs approach prompts scientists to better consider the constraints they apply on experimental animals. Several methods have been proposed to assess these constraints together with ways to alleviate them.

1. Assessment of constraints imposed to animals

According to the Nuffield Council on Bioethics [1], assessing the harms imposed on animals includes the evaluation of clinical signs, the study of animals' choices, the familiarity with ethological and ecological data, and the consideration of physiological and neurological features.

Clinical signs of poor welfare can be detected from the behaviour of animals and physiological indices of stress. Such signs are widely described in the literature. To cite the most common ones:

- Acute responses can be seen in attempts to escape from a situation, distress vocalisations, tachycardia and increased blood cortisol as observed when animals not used to handling are put in a holding cage [17]. More subtle signs can be seen in ear postures, with animals turning their ears back when exposed to a frightening stimulus [18].
- When animals are submitted to prolonged stressful situations, their behavioural reactivity can be altered leading to hyperactivity or on the contrary to hyporeactivity (sometimes referred to as 'apathy') and the functioning of the corticotropic axis is modified [19]. More subtle signs can be a reduction in food or water consumption, decrease in body weight, isolation from peers, or hypo activity.

In addition, diseases lead to clinical signs such as dull hair or hyperthermia or more specific signs such as nasal discharge, diarrhoea, dyspnoea,… which often do not require veterinary skills to be noticed [20]. Indices of pain have been extensively described in laboratory animals. Although less documented in farm animals, pain indicators have been proposed, e.g. for cattle [21], [22]:

- Isolation: a cow experiencing pain is often isolated from the rest of the herd and her activity is less synchronized with that of the other cows.
- Low head posture: an active cow has generally the head at or above the line of the spinal column whereas a cow not active, not sleeping nor ruminating, and with the head below the line of the spinal column can be in pain.
- Facial expression of pain: strained look, furrows above nostrils, ears back or in a 'lamb' position (low ears with the opening facing downwards).
- Antalgic 'guarding' postures to reduce stimulation on the painful area: a cow with a painful limb may stand unsteadily, sometimes the body leaning against a wall, or back arched, or with weight shifting on hind legs to release the weight on a painful forelimb.
- Behaviours directed towards the painful area (up to selfmutilation in extreme cases)
- Vocalisations: Pain can induce more vocalisations with a high frequency.
- Reduced feed intake, decrease in time spent ruminating, this affecting milk yield in dairy cows and body condition
- Failure to perform normal behaviours, such as selfgrooming or play, or changes in time spent lying (either increase or decrease)
- Clinical signs: erect hair, shivering of muscles, shivering or panting

Most of these indicators are not specific. For instance, fever can induce isolation-seeking and a stressed cow can have her ears backward and flattened. It is necessary to combine indicators before concluding that the animal is or is not in pain. In case of doubt, additional observations should be performed such as looking at the threshold for withdrawal movements in response to a noxious stimulation (an animal in pain often reacts at a lower threshold, which is called hyperalgesia) or checking if the animal displays such a withdrawal movement to a stimulation that would normally not induce withdrawal (allodynia). A scoring system was proposed by which a given clinical sign is assigned a specific weight and the final score is expressed on a 0–16 scale [21]. The authors consider that a cow is probably in pain if the score is 5 or more. However, different conditions will lead to different responses, and such scales should only be used for guidance.

The study of animal choices provides knowledge on animals' preferences vs. avoidances. One can allow animals choices between several options or ask them to work to obtain a putative preferred option [23]. For instance, using such methods, it was shown that cattle avoid people shouting or making rapid

erratic movements, that they search the company of familiar animals, and prefer to lie down on a soft and dry bedding, etc. [24], [25], [26]. This knowledge helps to design animal living conditions that offer what animals favour and to avoid producing stress during experimental procedures.

The study of the ethological repertoire and the ecology of a species provides knowledge on the behavioural repertoire of a species, which is the range of behaviours the animals from this species are likely to express in their natural environment. There is evidence that animals are motivated to express at least some of the behavioural patterns from their repertoire, even when not in their natural habitat: performing this behaviour is rewarding and preventing it is frustrating, possibly leading to abnormal behaviour. For instance, calves are motivated to suck a teat and to eat solid foods. By contrast, for a long time, veal calves were fed only liquids that they had to drink from buckets, which resulted in non-nutritive activities (cross sucking, nibbling, tong-playing) [27], [28]. Similarly tethering prevents cows from walking although they are motivated to do so [29]. The conditions in which experimental animals are kept should allow them to express activities they are motivated for, at least during some parts of the day.

Last, **knowing the physiological and neurological features of a species** can help to assess their degree of consciousness or the extent to which they can feel pain. For a long time only vertebrates and cephalopods were thought to be able to experience pain. However, this assumption was challenged by studies on invertebrates such as prawns [30]. Scientists now question the ethics of experimentation on insects [31]. Across species, animals should be given the benefit of the doubt regarding their ability to suffer. The same applies when considering life stages within a species, particularly in the highly debated area of foetal and neonatal sentience [32].

All these aspects of animal welfare should be assessed *a priori* – that is during the design of an experiment – from the knowledge available in the literature, from pilot studies or from experienced experts. Then during an experiment, the state of the animals should be monitored closely in order to detect any signs of reduced welfare.

2. Monitoring animal welfare and setting of end-points

Animals should be monitored closely during experiments. As a minimum, there should be a daily observation taking into account clinical and behavioural signs as well as feed intake and milk production, if applicable. The animals should be weighed or their body condition scored regularly. In addition, for each experiment that is likely to impair the welfare of animals, it is necessary to adjust the monitoring to the disorders that may occur: indices of stress if the procedure is likely to be stressful, monitoring of specific clinical signs or blood parameters in case a disease is induced.

Such a monitoring will enable the experimenter to take corrective actions, e.g. administration of pain killers or reduction of stress. In addition, it is essential to document the condition of the animals during studies so that the results are reproducible. Thresholds should be decided beforehand to decide that corrective actions are needed. Such thresholds are called 'Endpoints':

"A humane endpoint can be defined as the point at which an experimental animal's pain and/or distress can be terminated, minimized, or reduced by actions such as killing the animal humanely, terminating a painful procedure, or providing treatment to relieve pain and/or distress." [33]

The purpose of the application of endpoints is to be able to predict severe pain, distress, suffering, or impending death, and to act **before** the animal experiences these effects. Typically an experiment should be terminated in case of [33]:

- significant loss of body weight,
- loss of appetite and change in water consumption,
- important changes in physical appearance (e.g. posture, coat appearance),
- severe clinical signs, e.g. high body core temperature, increased heart rate and respiration rate, hematological indictors of a disease,
- marked changes in behaviour, e.g. selfmutilation, abnormal vocalizations, changes in responsiveness to external stimuli.

For each experiment where endpoints are required, the experimenter, in consultation with a veterinarian, must develop a checklist specific to the research program. All signs need to be measurable or described precisely, and a clear limit has to be set, e.g. beyond which weight loss or which score on a pain scale the experiment must be ended. Then a plan must be defined to specify the actions to be undertaken when an endpoint is reached.

C – Balancing issues between constraints on animals and expected benefit of an experiment

Following the 3Rs approach and the application of appropriate end-points should result in the minimisation of harm for animals used in experiments. Yet, one may still contend that such remaining harm is not justified. This poses the question of how to balance issues between animal harm and potential benefits from the research results. In his seminal paper, Bateson [34] proposed to weigh up the degree of animal suffering against the quality of the research. Bateson did not, however, clearly define how research quality should be judged; although one may understand that he means the importance of the expected results. Bateson [34] then added that the certainty of medical benefits needs to be taken into account. It is not clear if this dimension is independent from the quality of research or is part of the quality of research, focused on benefits to humans. Anyway, the scene was set that issues needs to be balanced between constraints for animals and benefits that could be taken from the knowledge gained from an experiment.

How to consider the potential benefits of an experiment? The Australian National Health and Medical Research Council (1997) [35] proposed that any experimentation on animals should be appraised according to five potential objectives: understanding humans or animals; maintaining or improving the health and well-being of humans and animals; improving breeding techniques, ecology, and education. In some cases, the links between the expected results and one or several of these objectives can be easily estimated beforehand. However, there are many examples where a piece of knowledge was produced from so-called blue-sky research and lead to very important application that had not been foreseen. Grimm et al. [36] argued that "whether practical benefits are realized is (a) impossible to predict and (b) exceeds the scope and responsibility of researchers". Forcing researchers to address the potential impacts of their research often leads them to promise too much from their research, which can in turn impair scientific credibility. Grimm et al. [36] thus proposed to take into account the contributions that the research will make to a specific scientific area or to a research program. They concluded that the value of the expected knowledge from research can be understood in terms of (a) the scientific value of knowledge – i.e. its contribution to a given research field – and (b) the societal value of knowledge – i.e. potential benefits for a given society. Moreover, deciding if a research field is important is a political decision.

The rigour and the reproducibility of experiments are also part of their quality [37], [38]. Research results do not always make their way up to publication. From our own experience in reviewing manuscripts, the poor design of experiments which prevent from drawing firm conclusions is actually the major cause of rejection. An unpublished piece of research using animals is definitively a waste of animals since no scientific knowledge will be produced. Poor reporting of the exact methods used is another cause of rejection. Editors of scientific journals report of a 'reproducibility crisis', in which a result cannot be consolidated by similar findings obtained by other researchers. The poor reproducibility can come from poor statistical analyses, incorrect data interpretation, technical issues, inappropriate experimental design (e.g. absence of control group), reporting only a subset of data that fits to the hypothesis, or even results falsification (summarised in [38]). This problem is amplified by the fact that negative results are often not published although they can be of significant value. According to Pritt and Hammer [38], poor animal welfare can also be a cause of poor reproducibility. Reporting precisely how an experiment was carried out and what were the statistical models should improve the reproducibility of results. To fill in this gap, the UK National Centre for the Replacement Refinement & Reduction of Animals in Research published the ARRIVE guidelines (Animal Research Reporting of *in vivo* experiments, available at https://arriveguidelines.org/). Several publishers now ask authors of manuscripts to follow the ARRIVE guidelines. Aske and Vaugh [37] argued that it would be more efficient that funding bodies required that applicants provide a clear description of their intended work. This however can sometimes be difficult: the first study of a project may be easy to describe precisely but experimenters may learn from this study and refine the following studies accordingly. We suggest that the analysis of how the study is described could take place during the ethical evaluation of a protocol.

In conclusion, **three main issues** must be considered to evaluate the acceptability of a study: **the constraints imposed on animals**, **the expected benefits of the research** (for the society or a discipline), and **the likelihood of obtaining research** (quality of the design).

D – Who should decide and how?

Balancing issues between pros and cons of an experiment is based on values: the values attributed to animals and their suffering, the values attributed to a research piece. Scientists alone cannot decide on these values [34]. Ethical committees for experiments on animals have thus been set up in various countries and organisations to help to take decisions. They generally include veterinarians, scientists, animal technicians and other people such as ethicists, animal protectionists, or laypeople.

Several metrics have been proposed to balance issues regarding experiments on animals and help ethical committees in their evaluation of protocols. In the context of human medical research, Liguori et al. [39] propose a metric for assessing expected benefits and animal harm, with 3-level scales.

Expected human benefits:

- *Small* is for an increase in basic biological or medical knowledge.
- *Medium* is for new or better treatment for a nonlifethreatening disease.
- *Large* is for new or better treatment for a lifethreatening disease.

(Note that in this context of human medical research, only benefits for human health were considered. A similar reasoning could be applied to define other benefits for humans or benefits to animals whereby for instance benefits are qualified high if there is large impacts on animal production, health and welfare (which in turn would benefit to humans).

Expected animal harm:

- *Mild* is for shortterm mild pain or distress with no significant impairment of the well-being or general condition.
- *Moderate* is for shortterm moderate pain, suffering, or distress, or long-lasting mild pain, suffering, or distress, likely to cause moderate impairment of the well-being or general condition.
- *Severe* is for severe pain, suffering, or distress, or longlasting moderate pain, suffering, or distress, likely to cause severe impairment of the well-being or general condition.

Then the probability that the results can be translated from animals to humans is estimated. The authors propose an explicit weighing scheme to combine the three scores, with a pre-defined threshold above which an experiment shall not be carried out. Such a metric approach allows harmonising the evaluation of protocols. However, it appears very technocratic and it largely limits ethical discussion among committee members. In addition, one can question if a ratio between harms and benefits is adequate to take ethical decisions: if an experiment is supposed to produce intense pain, is it acceptable even if the potential benefits are very high? If the research is of poor quality, can an experiment be carried out because it does not impose any constraints on animal? This raises the question of acceptable minima, which cannot be addressed in a cost/benefit analysis.

By contrast, a discourse approach is based on discussions between committee members to assess and weigh the various issues before taking a decision. Such an approach is criticised on the ground that there is no pre-defined method and thus the process lacks transparency, consistency and fairness [40]. Grimm et al. [40] recommended mixing the metric and the discourse approaches. A metric approach can guide the discussions among committee members and such discussions can contribute to establish metrics. Summarising the points raised in the present paper, we propose a list of issues to be addressed during an ethical evaluation of a protocol (Table 1). We do not establish scores and let ethical committee members discuss the decision to be taken based on the results of the assessment of each issue.

E – Conclusions and further discussion

There is strong consensus in the scientific community and beyond that the 3Rs are not optional. They need to be addressed in all animal experiments. As mentioned above, this requirement is explicit in EU legislation (Directive 2010/63/UE, Article 1). Taking into consideration ethical issues in animal experimentation shall not be seen as a limitation to experiments but rather as a help to better perform such experiments:

- Reducing the imposition of suffering on animals can help to reduce the variability between individuals and to ensure soundness of results.
- Taking full account of the 3Rs approach helps design adequate studies. The reduction and refinement should be part of the study design.
- A clear description of what is planned to do helps assessing protocols and increases the robustness and reproducibility of results.

Researchers should not consider ethical requirements as arguments to be filled in an application to an ethical committee. Rather they should include an ethical reasoning when designing their research, as part of their quality system, whereby research is made as efficient as possible in terms of reducing animal use by improving measurement methods – making them more precise –, taking most information from experiments by sharing protocols with other researchers and allowing re-use of samples and data, and making clear account of what has been done. We encourage leaders of research programs and institutions to put in place an 'animal experimentation policy' to encourage following the 3Rs principles and sharing of information between researchers on their methods, their procedures, their experiments and data.

An ethical attitude in research using animals will become even more critical in the context of Responsible Research and Innovation (RRI) introduced by Owen et al. [41] and put forward in the Horizon 2020 research program of the EU. The RRI concept is based on:

- **Anticipation** of potential impacts of research
- **Reflection** on the motivations for a given work and the sets of values related to this work
- **Deliberation** with a broad range of stakeholders (sometimes called Inclusion)
- **Responsiveness**, that is the ability for adaptive learning

As argued by [42], RRI can help meeting ethical principles in experiments on animals. RRI explicitly recognises the need to assess the potential benefits of a research, it forces clarity regarding values related to assessing harms and benefits, it calls for discussion with a broad range of experts and non-experts to take decisions about research, as in ethical committees, and calls for addressing adequately values and expectations to refine research projects.

Table 1. List of issues to be addressed when evaluating a protocol. This list should help to check that all necessary measures are taken to reduce animal constraints and that the remaining constraints are counterbalanced by the expected benefits from the experiment and the likelihood of obtaining reliable results. Addressing these issues should lead to a discussion among ethical committees to decide to accept or reject a proposed experiment. It can also be used to request scientists to refine their proposal.

Issue	Outcome
A – Is there an institution/personal strategy to commit with the 3Rs approach (including checking that no other studies are carried at the same time with a similar approach, reuse of collected data, meta-analysis of published work,…)	YES/NO
B – How the 3Rs were taken into account in the current work: • Have alternatives to the use of animals been sought and adequately assessed? • Is the number of animals at minimum (incl. reliability of how this was calculated) • Have the constraints on animals been minimised?	YES/NO
C – What are the (remaining) harms being imposed on animals, taking into account the sensitivity of the species, the number of animals and the procedures applied (including conditions before experiment)?	low – moderate – high
D – What is the importance of the expected results: • What knowledge will be produced that benefits a given research program or a discipline? • Can that knowledge benefit society?	detail the argumentation very likely – maybe – unlikely
E – Are the expected results likely to be obtained: • Is the protocol described in sufficient precise terms? • Do the experimental design sound and the statistical models ensure to obtain exploitable results?[1]	If not, ask for more information. YES/NO (NO may lead to rejection or to requesting changes.)

[1]The experimental design should be clearly described in terms of number and type of animals per treatment incl. controls. Potential confounding factors should be adequately addressed in the experimental design and/or the statistical models. The statistical models planned shall be able to highlight potential effects with no doubt.

References

1. Nuffield Council on Bioethics. The ethics of research involving animals. Plymouth, UK: Nuffield Council on Bioethics; 2005.
2. Liou S. The ethics of animal experimentation [Internet]. Stanford; 2010 [cited 2020 Dec 23]. Availabe from: https://hopes.stanford.edu/animal-research/2010
3. Russell WMS, Burch RL. The principles of humane experimental technique [Internet]. Baltimore: Johns Hopkins University; 1992 [cited 2020 Dec 23]. Available from: https://caat.jhsph.edu/principles/the-principles-of-humane-experimental-technique
4. Niderkorn V, Barbier E, Macheboeuf D, Torrent A, Mueller-Harvey I, Hoste H. In vitro rumen fermentation of diets with different types of condensed tannins derived from sainfoin (Onobrychis viciifolia Scop.) pellets and hazelnut (Corylus avellana L.) pericarps. Anim Feed Sci Tech. 2020;259. DOI: 10.1016/j.anifeedsci.2019.114357
5. de Boyer des Roches A, Veissier I, Coignard M, Bareille N, Guatteo R, Capdeville J, et al. The major welfare problems of dairy cows in French commercial farms: an epidemiological approach. Anim Welfare. 2014;23(4):467-78. DOI: 10.7120/09627286.23.4.467

6. Veissier I, Mialon M-M, Sloth KH. Short communication: Early modification of the circadian organization of cow activity in relation to disease or estrus. J Dairy Sci. 2017;100(5):3969-74. DOI: 10.3168/jds.2016-11853

7. Sauvant D, Letourneau-Montminy MP, Schmidely P, Boval M, Loncke C, Daniel JB. Review: Use and misuse of meta-analysis in Animal Science. Animal. 2020;14(S2):s207-s22. DOI: 10.1017/S1751731120001688

8. Cohen J. Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates; 1988.

9. Kruschke JK, Liddell TM. The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. Psychon B Rev. 2018;25(1):178-206. DOI: 10.3758/s13423-016-1221-4

10. Tournadre H. "Four-star" stalls to measure digestibility in cattle [Une stalle « 4 étoiles » pour les mesures de digestibilité chez les bovins]. Innover avec l'Inra. 2019;111. Available from: http://prod.inra.fr/layout/set/newsletter/Kiosque/Newsletters/La-lettre-aux-entreprises/Campagne/NL-innover-111-2019

11. Nasrollahi SM, Nozière P, Dewhurst RJ, Chantelauze C, Cheng L, Froidmont E, et al. Natural 15N abundances in plasma and urea-N concentration in milk as biomarkers of urinary N excretion in dairy cows: a meta-analysis. 6th EAAP International Symposium on Energy and Protein Metabolism and Nutrition 9-12 September 2019; Belo Horizonte, Brazil 2019.

12. Andueza D, Picard F, Pourrat J, De la Torre A, Devant M, Reynolds CK, et al. Faecal-NIRS for predicting animal-to-animal variation in feed organic matter digestibility in cattle. EAAP annual conference, 1-4 December 2020; webinar2020.

13. Swiss Academy of Medical Sciences (SAMS), Swiss Academy of Sciences (SCNAT). Ethical principles and guidelines for experiments on animals [Internet]. Basel, Bern: SAMS/SCNAT; 2005. Available from: https://private.aaalac.org/intlRefs/IntRegs/Siwtzerland/Ethical%20Principles%20and%20Guidelines 2005. 6 p. (last visit 2020-12-23)

14. Poole T. Happy animals make good science. Lab Anim. 1997;31(2):116-24. DOI: 10.1258/002367797780600198

15. Peterson NC, Nunamaker EA, Turner PV. To treat or not to treat: The effects of pain on experimental parameters. Comp Med. 2017;67(6):469-82.

16. Törnqvist E, Annas A, Granath B, Jalkesten E, Cotgreave I, Öberg M. Strategic Focus on 3R Principles Reveals Major Reductions in the Use of Animals in Pharmaceutical Toxicity Testing. PLOS ONE. 2014;9(7):e101638. DOI: 10.1371/journal.pone.0101638

17. Veissier I, Le Neindre P, Trillat G. Adaptability of calves during weaning. Biologie du Comportement (Biology of Behaviour). 1989;14:66-87.

18. Boissy A, Aubert A, Désiré L, Greiveldinger L, Delval E, Veissier I. Cognitive sciences to relate ear postures to emotions in sheep. Anim Welfare. 2011;18:47-56.

19. Mormède P, Andanson S, Auperin B, Beerda B, Guemene D, Malmkvist J, et al. Exploration of the hypothalamic-pituitary-adrenal function as a tool to evaluate animal welfare. Physiol Behav. 2007;92(3):317-39. DOI: 10.1016/j.physbeh.2006.12.003

20. Welfare Quality. Welfare Quality® assessment protocol for cattle (fattening cattle, dairy cows, veal calves). Lelystad, The Netherlands: Welfare Quality® Consortium; 2009.

21. De Boyer Des Roches A, Faure M, Lussert, A., Herry V, Rainard P, Durand D, et al. Behavioral and patho-physiological response as possible signs of pain in dairy cows during Escherichia con mastitis: A pilot study. J Dairy Sci. 2017;100(10):8385-97. DOI: 10.3168/jds.2017-12796

22. Gleerup KB, Andersen PH, Munksgaard L, Forkman B. Pain evaluation in dairy cattle. Appl Anim Behav Sci. 2015;171:25-32. DOI: 10.1016/j.applanim.2015.08.023

23. Dawkins MS. Battery hens name their price: consumer demand theory and the measurement of ethological "needs". Anim Behav. 1983;31:1195-205.

24. Munksgaard L, De Passillé AM, Rushen J, Thodberg K, Jensen MB. The ability of dairy cows to distinguish between people. Proceedings of the 29th International Congress of the International Society for Applied Ethology, 3-5 August 1995, at Exeter, UK. 1995:19-20.

25. Fregonesi JA, Veira DM, von Keyserlingk MAG, Weary DM. Effects of bedding quality on lying behavior of dairy cows. J Dairy Sci. 2007;90(12):5468-72. DOI: 10.3168/jds.2007-0494

26. Raussi S, Niskanen S, Siivonen J, Hanninen L, Hepola H, Jauhiainen L, et al. The formation of preferential relationships at early age in cattle. Behav Proc. 2010;84(3):726-31. DOI:

10.1016/j.beproc.2010.05.005

27. Veissier I, Ramirez de la Fe AR, Pradel P. Non-nutritive oral activities and stress responses of veal calves in relation to feeding and housing conditions. Appl Anim Behav Sci. 1998;57(1-2):35-49.

28. Veissier I, Chazal P, Pradel P, Le Neindre P. Providing social contacts and objects for nibbling moderates reactivity and oral behaviors in veal calves. J Anim Sci. 1997;75(1-2):356-65.

29. Veissier I, Andanson S, Dubroeucq H, Pomiès D. The motivation of cows to walk as thwarted by tethering. J Anim Sci. 2008;86(10):2723-9. DOI: 10.2527/jas.2008-1020

30. Barras C. Invertebrates can feel pain, suggests study on prawns. New Sci. 2007;196(2629):14. DOI: 10.1016/S0262-4079(07)62820-6

31. Freelance CB. To Regulate or Not to Regulate? The Future of Animal Ethics in Experimental Research with Insects. Sci Eng Ethics. 2019;25(5):1339-55. DOI: 10.1007/s11948-018-0066-9

32. Mellor DJ, Diesch TJ. Onset of sentience: The potential for suffering in fetal and newborn farm animals. Appl Anim Behav Sci. 2006;100(1-2):48-57. DOI: 10.1016/j.applanim.2006.04.012

33. CCAC (Canadian Council on Animal Care). Guidelines on Choosing an Appropriate Endpoint in Experiments Using Animals for Research, Teaching and Testing. Ottawa, ON: CCAC; 1998.

34. Bateson P. When to experiment on animals. Conflicts of interest between experimenters and their critics might be resolved by weighing up the degree of suffering againts the value of the research. New Sci. 1986:30-2.

35. National Health and Medical Research Council. Australian Code of Practice for the Care and Use of Animals for Scientific Purposes. 6th ed. Canberra: Australian Government Pub. Service; 1997.

36. Grimm H, Eggel M, Deplazes-Zemp A, Biller-Andorno N. The road to hell Is paved with good intentions: why Harm-Benefit analysis and its emphasis on practical benefit jeopardizes the credibility of research. Animals. 2017;7(9). DOI: 10.3390/ani7090070

37. Aske KC, Waugh CA. Expanding the 3R principles. Embo Rep. 2017;18(9):1490-2. DOI: 10.15252/embr.201744428

38. Pritt SL, Hammer RE. The Interplay of ethics, animal elfare, and IACUC oversight on the reproducibility of animal studies. Comparative Med. 2017;67(2):101-5. Available from: https://europepmc.org/backend/ptpmcrender.fcgi?accid=PMC5402729&blobtype=pdf

39. Liguori GR, Jeronimus BF, de Aquinas Liguori TT, Moreira LFP, Harmsen MC. Ethical Issues in the use of animal models for tissue engineering: Reflections on legal aspects, moral theory, three Rs Strategies, and Harm-Benefit analysis. Tissue Eng Part C-Me. 2017;23(12):850-62. DOI: 10.1089/ten.tec.2017.0189

40. Grimm H, Olsson IAS, Sandoe P. Harm-benefit analysis – what is the added value? A review of alternative strategies for weighing harms and benefits as part of the assessment of animal research. Lab Anim. 2019;53(1):17-27. DOI: 10.1177/0023677218783004

41. Owen R, Stilgoe J, Macnaghten P, Gorman M, Fisher E, Guston D, et al. A Framework for Responsible Innovation. 2013. p. 27–50. DOI: 10.1002/9781118551424.ch2

42. McLeod C, Hartley S. Responsibility and Laboratory Animal Research Governance. Sci Technol Hum Val. 2018;43(4):723-41. DOI: 10.1177/0162243917727866

Corresponding authors: Kenny Rutherford, SRUC, Animal Behaviour and Welfare Roslin, United Kingdom, E-mail: kenny.rutherford@sruc.ac.uk