The Principles of Humane Experimental Technique: Timeless Insights and Unheeded Warnings

Michael Balls
FRAME, Nottingham, UK

Summary
In The Principles of Humane Experimental Technique*, Russell and Burch said that "the central problem is that of determining what is and what is not humane, and how humanity can be promoted without prejudice to scientific and medical aims". They then explained how the Three Rs can be used to diminish or remove direct inhumanity ("the infliction of distress as an unavoidable consequence of the procedure employed") and contingent inhumanity ("the infliction of distress as an incidental and inadvertent by-product of the use of a procedure"). They concluded that "Replacement is always a satisfactory answer, but Reduction and Refinement should, whenever possible, be used in combination".*
Many of the commonsense insights in The Principles *are no less relevant today than they were in 1959. However, their warnings about the limited value of models and, in particular, the danger of succumbing to the high-fidelity fallacy (whereby it is assumed that the best models for humans are always placental mammals, because they are more like humans than other animals), appear to have largely gone unheeded. Of particular importance is their discussion on toxicity testing, which they saw as one use of laboratory animals "which is an urgent humanitarian problem, for it regularly involves considerable and sometimes acute distress". How, then, can it be that mammalian models are still routinely used in attempts to detect chemical carcinogens and reproductive toxins, despite the fact that the relevance to humans of the data they provide has not been, and perhaps could never be, satisfactorily established? Nevertheless, there are signs that some significant changes in attitude are taking place, particularly in the USA, which could be more in line with the main thrust of The Principles, the belief that good science and human technique inextricably go hand-in-hand.*

Keywords: animal experimentation, reduction, refinement, replacement, Russell & Burch, Three Rs, toxicity testing

1 Introduction

On 21 August 2007, in Tokyo, Japan, during the Opening Ceremony of the *6th World Congress on Alternatives and Animal Use in the Life Sciences*, we paid tribute to the rich and varied life of W. M. S. Russell, who, with R. L. Burch, gave us the Three Rs concept, in their book, *The Principles of Humane Experimental Technique* (Russell and Burch, 1959; Balls, 2008). Now, two years later, in Rome, Italy, toward the end of the *7th Congress,* we are celebrating the 50th anniversary of the publication of *The Principles*.

I firmly believe that *The Principles* contains timeless insights into how we should think about the use of laboratory animals for research and testing, which are as relevant today as they were in 1959, and which can guide us as we seek to achieve genuine progress, whilst maintaining the highest standards in terms of both scientific methodology and animal welfare. The book also contains warnings about how fundamental mistakes can be made, which compromise the value of the science and threaten the welfare of the animals.

My concern is that, although a large number of people say they are committed to supporting the Three Rs concept of *Reduction*, *Refinement* and *Replacement*, as put forward by Russell and Burch, most of them are unaware of the detailed implications of these insights and warnings, because they have not read the book itself. The result is that I am disappointed that the great benefits afforded by a careful consideration and dedicated application of *The Principles* have not been achieved. I therefore hope that this Congress will mark a new beginning – a much-needed, renaissance of the Three Rs.

As one of the initiatives to celebrate its own 40th anniversary, FRAME has made an abridged version of *The Principles* available, with the cooperation and support of Cleo Paskal, W. M. S. Russell's Literary Executor. The principal aim of *The Three Rs and the Humanity Criterion* (Balls, 2009) is "to seek to retain the remarkable concepts and flavour of the original, whilst clarifying some of the English language employed, as well as reducing some of the lengthy discussions based on uses of animals in the 1950s which are no longer practised".

Plenary lecture held at the 7th Worldcongress in Rome 2009

What I plan to do here, is to list some of the insights, then say why the failure to heed some of the warnings severely limits the impact of the Three Rs, and, as a result, compromises the high standards of scientific practice and animal welfare which Russell and Burch sought. Quotations from *The Principles* are shown in italics in the sections 2 to 4.

2 The insights

The main principle on which Russell and Burch based their analysis is that *it is widely recognised that the humanest possible treatment of experimental animals, far from being an obstacle, is actually a prerequisite for successful animal experiments*. They considered the central problem to be that of *determining what is and what is not humane, and how humanity can be promoted without prejudice to scientific and medical aims*.

They began with *the concept of inhumanity and its relation to those of pain and distress,* then turned to *the positive aspect – the analysis of methods of diminishing inhumanity in experimentation*.

They said *that we must first distinguish direct and contingent inhumanity. By the former, we mean the infliction of distress as an unavoidable consequence of the procedure employed. By the latter, we mean the infliction of distress as an incidental and inadvertent by-product of the use of the procedure, which is not necessary for its success*.

Their thesis was that *inhumanity can be, and is being, diminished or removed under the three broad headings of Replacement, Reduction, and Refinement – the Three Rs of humane technique:*

- *Reduction means reduction in the numbers of animals used to obtain information of a given amount and precision*.
- *Refinement means any decrease in the incidence or severity of inhumane procedures*.
- *Replacement means the substitution for conscious living higher animals of insentient material. It is always a satisfactory answer, but reduction and refinement should, wherever possible, be used in combination*.

3 The warnings

The Principles also contains a number of important warnings, but it will only be possible to discuss two of them here.

First, the tendency to misunderstand the nature of models, and especially, the use of animals as models for man: *A perfect model of the human organism would obviously be indistinguishable by any test from its original. Any other* in vivo *model must depart in some degree from the original. There are two factors governing the way in which the model differs from the original. Fidelity means overall proportionate difference, and discrimination means the extent to which the model reproduces one particular property of the original*.

The point is that, however great the overall similarity between the original and a model may be, if there are significant differences in the specific properties being studied, the model will not be useful. Also, however, great the differences between the original and a model may be, if there are sufficient similarities in the specific properties being studied, the model may be a useful one. Clearly, high fidelity/high discrimination models are most useful, but, where this is not possible, a low fidelity/high discrimination model is preferable to a high fidelity/low discrimination one.

Russell and Burch go on to say that *Progress in replacement has been restricted by certain plausible, but untenable assumptions* about models, which have led to *the high-fidelity fallacy*. The major premise is that *the highest possible fidelity is always desired in medical research and testing*, and that, *for man*, a *member of another placental mammal species would be a model of higher fidelity than a bird or a microbe*. This assumption can have disastrous consequences in terms of the data produced, and can also lead to unnecessary, and therefore unacceptable, animal suffering.

Most of the macaques used in the UK are involved in toxicity testing for the pharmaceutical industry. When asked about the 16% increase in 2008, a senior scientist from Global R&D of a leading pharmaceutical company said that this was driven by a move toward more biological medicines (Gill, 2009): "These treatments need to be tested in a human-like model, and old world primates are closer relatives of humans than new world primates." But what about the TGN 412 scandal, where the "human-like model" did not reveal the acute adverse effects of this humanised product, which later occurred in human volunteers? Macaques should not be used merely because of their overall similarity to humans, but only when it has been established, in advance, that they are appropriate models for use in a particular study.

Russell and Burch were concerned about toxicity testing on more-general grounds, since they considered it to be *an urgent humanitarian problem, for it regularly involves considerable and sometimes acute distress, and to be an activity where the high-fidelity fallacy may be more prevalent and influential at the legal level, rather than at the laboratory level*. They clearly foresaw the problem of persuading regulators to accept the use of scientifically-advanced, replacement alternative methods instead of the animal tests with which they are more comfortable.

4 Progress of the Three Rs

4.1 Reduction

At the 5th World Congress, in Berlin in 2005, I said that "the progressive *reduction* in the numbers of animal experiments which had been foreseen when the new legislation was passed in the 1980s seems to have come to an end, especially as more and more mice are sacrificed on the altar of genetic exploitation. Also, far from working together toward the zero option of the use of non-human primates, there is pressure to build more and more primate research centres (Balls, 2006)." Sadly, the situation has worsened since 2005, rather than improved.

In Britain, the number of scientific procedures on living animals in 2008 was higher than the number in 1987 (Anon, 2009), the first year after the *Animals (Scientific Procedures) Act 1986*

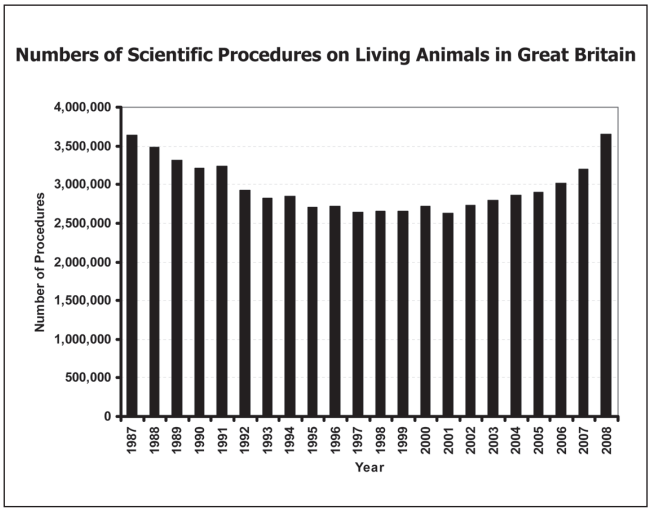

Fig. 1: Numbers of scientific procedures on living animals in Great Britain, 1987 to 2008*.

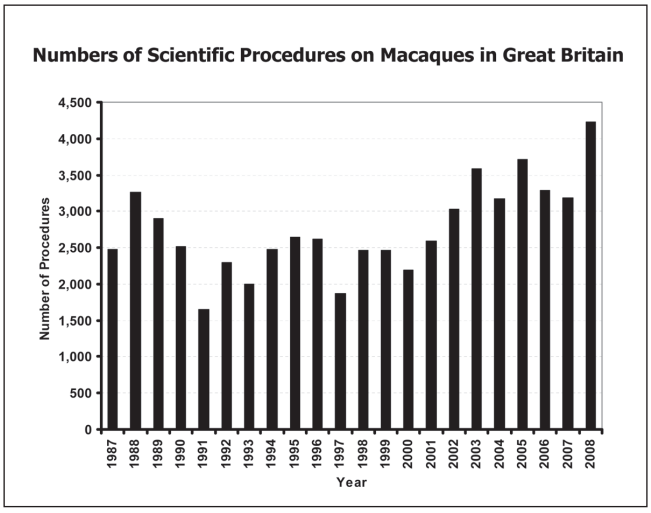

Fig. 2: Numbers of scientific procedures on macaques in Great Britain, 1987 to 2008*.

came into force (Fig. 1), and the number of procedures applied to old world primates (macaques) in 2008 (4230) was much higher than in 1987 (2470; Fig. 2).

The situation in Britain is no doubt mirrored in other countries, and the production and use of genetically modified mice is widely used to excuse the overall situation. The chilling prospect is that there is much worse to come, given a recent report in *Nature* (Abbot, 2009): "European investment could see knockout rats catching up with mutant mice in medical research. The European Commission has approved the world's first major systems-biology programme to study the rat. Known as EURATRANS – for European large-scale functional genomics in the rat for translational research – the multimillion-€ project includes collaborators in the United States and Japan." Will this be followed by a move to set up programmes for producing large numbers of transgenic non-human primates?

One of the main scientific points of emphasis in *The Principles*, the need for high quality experimental design and statistical analysis, has been largely ignored. Indeed, there is much evidence to support the contention that scientists, regulators, universities, industries, governments and grant-giving bodies are content to tolerate bad science. There are some praiseworthy efforts to redress this situation, such as the training schools run by the FRAME Reduction Steering Committee, in collaboration with the University of Manchester, and with the support of the European Commission's COST programme (Howard et al., 2009).

Virtually no time has been specifically devoted to *Reduction* at this 7th World Congress, so it is impossible to avoid the conclusion that it is the forgotten R, even though Russell and Burch saw it as *of great importance, and of all the modes of progress, it is the one most obviously, immediately, and universally advantageous in terms of efficiency*.

4.2 Refinement

There has been considerable progress concerning the husbandry and use of laboratory animals, not least because of greater recognition of the importance of laboratory animal technicians and laboratory animal veterinarians.

That is to be welcomed, but there is also a danger that refinement can be used as a convenient way of showing commitment to the Three Rs, whilst ensuring that animal experimentation is seen as respectable and can be allowed to continue, while the fundamental ethical questions raised by it are avoided.

This is not my area of expertise, so I will not dwell on it further. However, I do wonder whether the activities linked to ethical review processes and institutional animal care and use committees, however positive they may be in terms of refinement, have any significant effects in relation to reduction and replacement. We should remember that Russell and Burch said that, *in general, refinement is never enough, and we should always seek further for reduction and if possible replacement*.

4.3 Replacement

The position adopted by academia and research-funding bodies has long been that new techniques emerge during the natural development of a science, so deliberately seeking replacement alternatives for the vast array of procedures applied to laboratory animals in the basic sciences is not necessary. However, it could be argued that the legislation which regulates animal experimentation imposes legal and ethical obligations on all concerned, which should not be so easily avoided. From the scientific point of view, the high-fidelity fallacy deserves far greater recognition and resultant action, especially in the case of animal models of human disease, where insufficient about the disease

* Figures 1 and 2 were kindly provided by Michelle Hudson, and are based on the annual statistics of animal procedures regulated under the terms of the *Animals (Scientific Procedures) Act 1986*, published each year by The Stationery Office, London.

is known for sound judgements to be made about the relevance, or otherwise, of the model.

Nevertheless, it is toxicity testing where, as Russell and Burch recognised, the greatest concerns arise. For example: Why is it believed that the rodent bioassay can tell us what chemicals are likely to be carcinogenic in humans, when dosing is based on the maximum tolerated dose, and the mouse is a poor model for the rat and vice versa? Why is it believed that the current regulatory reproductive toxicity tests can tell us what chemicals are likely to be reprotoxic in humans, when so many false positives and false negatives occur that it is impossible to judge whether the test procedures can even identity reprotoxins in the animal models themselves? Why are animal data still widely regarded as the "gold standard" to be matched by non-animal tests, when the reliability and relevance of the tests concerned cannot be established, even for the animals? The Draize rabbit eye irritancy test data are so variable that the test cannot reliably be used to identify potential eye irritants in rabbits, so why is it believed that the data have any relevance to humans?

5 Moving backward in Europe

Despite much positive talk by politicians and senior officials about the importance of the Three Rs and their commitment to them, it could be argued that, in actual fact, Europe is going backwards. The great promises of the 1990s have not been delivered. Three examples will suffice, although this is not the place to discuss them in detail.

1. The REACH system: totally unworkable, proposed by ill-informed ambitious civil servants, taken up by ill-informed ambitious politicians, and then by ill-informed ambitious governments. It was clear from the early drafts of the Commission's White Paper that nobody had any coherent or defensible idea of the numbers of chemicals that would need to be registered, the number of additional animal tests that would be required, or how human health and the environment would be afforded greater protection. There was no mention of non-animal tests or their validation. Later on the potential value of replacement alternatives was grudgingly accepted, and, as it became clear that the numbers and costs of various aspects would be much, much higher than had been expected, they came to be seen as a way of saving face and reducing embarrassment. We now have an expensive agency in Helsinki, which is producing thousand upon thousand of "guidance" documents. What is the value of accumulating so-called "missing" data, if its value and genuine usefulness have not been established? There is a likelihood that, since the science is being driven by the politics, the validation process itself will be corrupted. Instead of waiting until they have been independently shown to be reliable and relevant for their stated purposes, replacement alternative tests may come to be accepted because they are "suitable" (i.e. politically convenient). But aren't plausibility and suitability, based on the high-fidelity fallacy, among the reasons why we have so many useless animal tests? What will be the consequences, if "suitable" tests are found, in time, to have been "unsuitable" after all, and who will accept the responsibility for their failure?

2. The 7th Amendment to the Cosmetics Directive: a ruse of no value, seemingly designed to convince politicians and a gullible public that something is being done. The situation with regard to cosmetic ingredients is no less unsatisfactory. Many of the chemicals used in cosmetics are also used for other purposes, and the REACH system will apply to them. If the testing of cosmetic ingredients in compliance with the Cosmetics Directive comes to be banned, which companies will admit to doing any animal testing for that purpose? Won't they say that the testing was done for compliance with the REACH system, and won't some of them try to stick to "not tested in animals for cosmetics purposes" labelling, while conveniently and dishonestly omitting the last three words? In addition, the definition of a "cosmetic" used in Europe is increasingly unsatisfactory, as cosmetic products are produced which actively alter the biological properties of the components of the skin.

3. Draft proposals for a Directive to replace Directive 86/609/EEC: one of the worst pieces of draft legislation ever published, which even foresees circumstances in which Member States could permit experiments on Great Apes. The Commission's proposals were produced after years of discussion with all kinds of stakeholders, but what emerged was not a draft directive at all. Rather than having its intentions spelled out clearly, in a way which could be implemented as a law, the result was a curate's egg mishmash of ideas which were either ill-conceived or in need of further discussion and development. As a result, hundreds of amendments have been put forward, and the result is a threat to both good science and sound animal welfare. Despite this totally unsatisfactory situation, there is great political pressure to get something in the statute book.

I mention these three points, because I fear that they illustrate the fact that Europe is going down a slippery slope as far as the Three Rs and a sensible balance between science and animal welfare are concerned. In the long run, fine words and catchy slogans count for nothing – it is sustainable actions of high quality that matter. As Jesus said, "Ye shall know them by their fruits" (Matthew's Gospel, 1611 translation). The problem is, who has the power and the desire to intervene and see that the downward trend is reversed?

6 Moving forward in the USA?

Meanwhile in the USA, a number of very promising developments are taking place, and in particular, the follow up to the publication by the US National Academy of Sciences of *Toxicity Testing in the 21st Century – A Vision and a Strategy* (National Research Council, 2007) and by the US Food and Drug Administration of the *Critical Path Initiative* (FDA,

2004). What these two documents have in common is the recognition that animal models can no longer be relied on in drug development and in toxicity testing in general, and that more effort should be put into the development, evaluation, acceptance and use of what we would describe as replacement alternative methods and strategies, particularly when they are of direct relevance to humans.

To be fair, I must recognise that promising developments are also taking place in Europe in relation to speed and safety in drug discovery, as was shown, for example, at a symposium held in London in 2008 (Gard and Clotworthy, 2009).

7 Concluding remarks

The published proceedings of the 7th World Congress will reveal a wealth of activity, mainly focused, perhaps, on the possibility of replacing animal procedures by more-modern methods, based on the remarkable progress being made in cellular and molecular biology. In a way, then, the Three Rs concept is to the fore as we celebrate the 50th anniversary of the publication of *The Principles*.

Nevertheless, as I have pointed out, there are grave causes for concern, especially as the number of animal procedures conducted each year continues to increase, and legislative changes, especially in Europe, threaten to perpetuate and expand that increase even more.

I hope that many of the grandiose statements made in apparent support of good and ethical science based on the Three Rs will lead to identifiable and excellent outcomes, which will demonstrate a genuine renaissance in line with Russell and Burch's outstanding concept.

In particular: significant *reduction* in animal use should be achieved, without further delay, through better experimental design and statistical analysis; *refinement*, however welcome, should not be seen as an end in itself; and much greater resources should be invested in the dedicated search for *replacement* alternatives. Meanwhile, Russell and Burch's warning about the high fidelity fallacy should be taken much more seriously and acted upon.

The way in which Russell and Burch put it cannot be repeated too often: *If we are to use a criterion for choosing experiments, that of humanity is the best we could possibly invent. The greatest scientific experiments have always been the most humane and attractive, conveying that sense of beauty and elegance which is the essence of science at its most successful.*

So, let us all take this opportunity to renew our commitment to live up to this ideal, with total sincerity, then go home, and get on with the job.

References

Abbot, A. (2009). Return of the rat. *Naturenews 460*, 788. Published online on 11 August 2009.

Anon (2009). *Statistics of Scientific Procedures on Living Animals*, Great Britain 2008. HC 800. London: The Stationery Office.

Balls, M. (2009). *The Three Rs and the Humanity Criterion*. Nottingham: FRAME.

Balls, M. (2008). Professor William Russell (1925-2006): Doyen of the Three Rs. *AATEX 14, Spec. Issue*, 1-7.

Balls, M. (2006). The Three Rs: Looking back … and forward. *ALTEX 23, Spec. Issue*, 29-32.

FDA (2004). *Challenge and Opportunity on the Critical Path to New Medicinal Products*. Available at: http://www.fda.gov/oc/initiatives/critical path/whitepaper.pdf. (Accessed 15.04.05). Silver Spring, MD: US Department of Health and Human Services, Food and DRUG Administration.

Gard, T. and Clotworthy, M., eds (2009). Proceedings of the Speed and Safety in Drug Discovery Symposium. *ATLA 37, Suppl. 1*, 1-65.

Gill, V. (2009). Upward jump in lab animal tests. *BBC Online*, 21 July 2009. Available at: http://news.bbc.co.uk/1/hi/sci/tech/8161253.stm (Accessed 21.07.09).

Howard, B., Hudson, M. and Preziosi, R. (2009). More is less: reducing animal use by raising awareness of the principles of efficient study design and analysis. *ATLA 37*, 33-38.

National Research Council (2007). *Toxicity testing in the 21st century: A vision and a strategy*. Washington, DC: National Academy Press.

Russell, W. M. S. and Burch, R. L. (1959). *The Principles of Humane Experimental Technique*. London: Methuen.

St Matthew (1611). Jesus teaches with authority. Chapter 7, verse 16. New Testament, Authorised King James Version. London: Church of England.

Correspondence to

Prof. Michael Balls
FRAME
Russell & Burch House
96–98 North Sherwood Street
Nottingham NG1 4EE, UK
e-mail: michael.balls@btopenworld.com