ANIMAL TESTING, 3R MODELS AND REGULATORY ACCEPTANCE

Technology Transition in a Risk-averse Context

Marie-Jeanne Schiffelers

Assessment Committee
Dr. K-H. Buchheit, European Directorate for the Quality of Medicines, Strasbourg
Prof. dr. W.D.J. Kremer, Faculty of Veterinary Medicine, Utrecht University
Prof. dr. A.J. Meijer, Utrecht University School of Governance
Prof. dr. A.H. Piersma, Institute for Risk Assessment Sciences, Utrecht University
Prof. dr. B.A.M. van der Zeijst, Leiden University Medical Center

ISBN: 978-90-393-6567-0

Lay-out: Persoonlijk Proefschrift
Cover: John Wright of Derby. “An Experiment on a Bird in an Air Pump.” (1768) National Gallery, London
Printing: Ipskamp Drukkers

ANIMAL TESTING, 3R MODELS AND REGULATORY ACCEPTANCE

Technology Transition in a Risk-averse Context

Dierproeven, 3V-modellen en de acceptatie in het regulatoire domein
Technologie transitie in een risicomijdende context

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen

op vrijdag 17 juni 2016 des middags te 2.30 uur

door

Marie-Jeanne Winanda Adolph Schiffelers

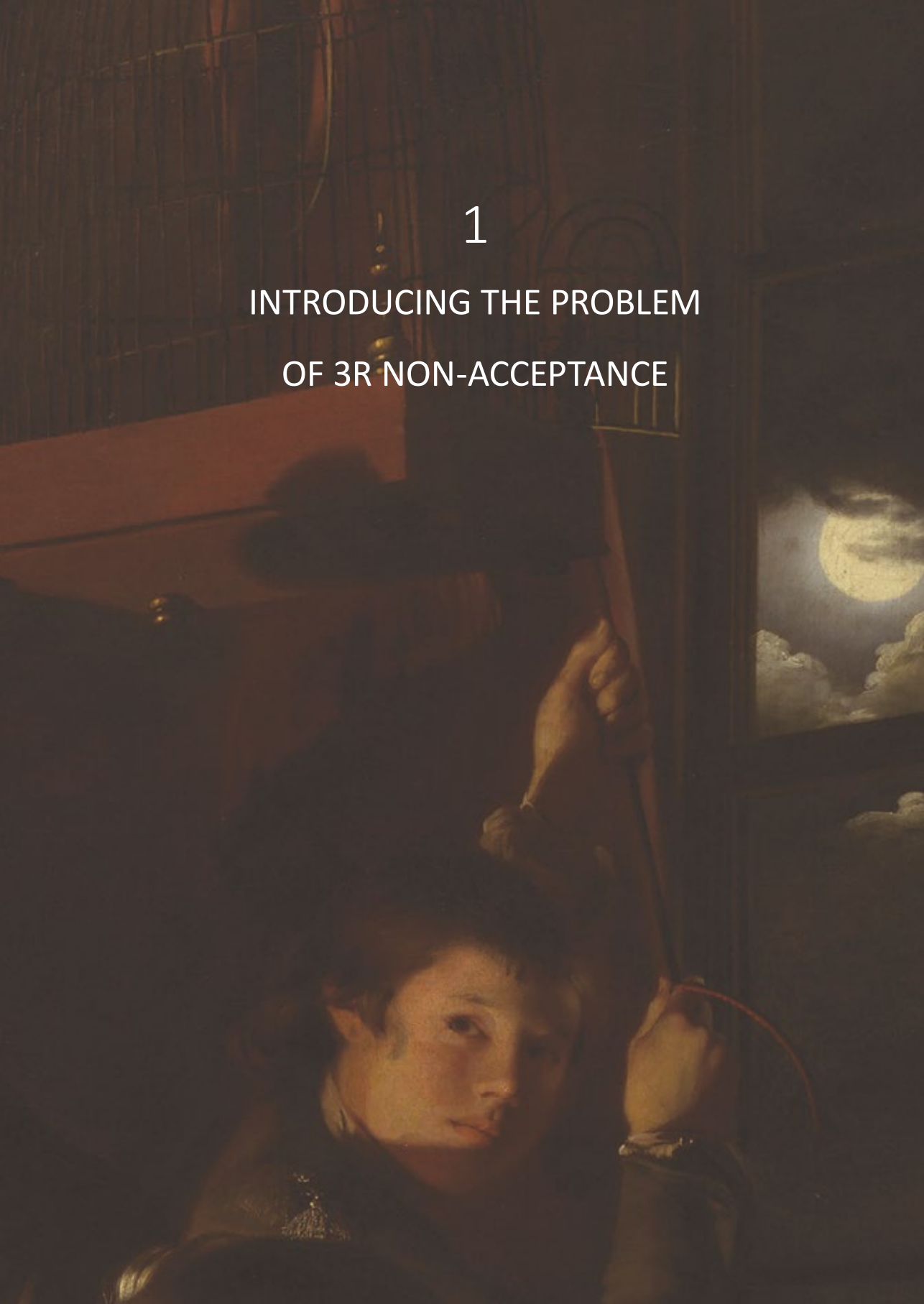
geboren op 2 oktober 1970 te Hoensbroek

Promotoren: Prof. dr. B. J. Blaauboer
Prof. dr. C. F. M. Hendriksen
Prof. dr. W.E. Bakker

This thesis was financially supported by the Doerenkamp-Zbinden Foundation & TNO

The printing of this thesis was financially supported by the Willy van Heumenfonds

TABLE OF CONTENTS



1

INTRODUCING THE PROBLEM OF 3R NON-ACCEPTANCE

*"It's better not to change ten times
than to make nine changes for the better and
one for the worse."*

Civil servant European Commission
as quoted in research Schiffelers et al., 2005

1.1 The central research problem

In Europe, approximately 11,5 million laboratory animals are used annually for a variety of purposes such as education and training, basic research and safety assessment, and efficacy testing of substances and products (EC, 2013). Animal testing or *in vivo* testing raises concerns in terms of its scientific value, moral issues and the costs connected to the use of laboratory animals (see section 2.2 for an elaboration on this). In response to these concerns Russell and Burch introduced in 1959 the 3Rs principle to 'replace, reduce and refine' animal models as much as possible. In their influential book "*The Principles of Humane Experimental Technique*" they presented the 3Rs framework to make progress both in terms of animal welfare as well as scientifically. *"They advocated using scientific ingenuity to replace, reduce and refine the use of animals wherever feasible without compromising scientific rigor"* (Stephens and Mak, 2014, p.2).[1] These 3R approaches have the potential of combining better science and advanced relevance of the test, with fewer animals and less animal suffering and faster and cheaper test results. Ever since, an increasing number of 3R models have become available.[2]

Many of these 3R models are already extensively used for product Research & Development (R&D) and product testing. However, their acceptance and use in the area of safety assessment and efficacy testing of products and substances such as pharmaceuticals and chemicals is repeatedly referred to as highly challenging (e.g. Richmond, 2002; Garthoff, 2005; Hendriksen, 2006; Krul et al., 2006, Blaauboer and Andersen, 2007, Cooper and Jennings, 2008, Bottini et al., 2008, Lilienblum et al., 2008, Hartung and Daston, 2009, Kooijman 2013).

1 The term 'alternatives' was introduced by animal protection organizations in the 1960s and 1970s and is thereby more politically charged as the term 3Rs (Stephens and Mak, 2014). Furthermore the term 'alternatives' predominantly refer to the replacement option. For this reason the term 3Rs is used as the default in the context of this thesis.

2 **Replacement methods** are those methods which no longer entail the use of live animals. *In vivo* testing is replaced by full *in vitro* (cell culture) models, analytical methods, *in silico*/computational models, ethical humane studies or a combination of these options. Out of the 3Rs replacement is the preferred option. However, full replacement options are scarse especially when it comes to more complex endpoints such as carcinogenicity, reproductive toxicity, immunogenicity and mutagenicity. Most replacement methods that have been developed so far are cell cultures measuring a few parameters. There is now a development going on towards more complex systems: e.g. organotypic cell cultures which model certain organs. The future perspective is the development of multi-organ models in which the metabolisms can be mimicked e.g. the human on a chip.
Reduction methods are methods that use fewer animals compared to the conventional animal model due to a well-designed protocol and analysis of the findings. Reduction methods enable researchers to obtain comparable levels of information from fewer animals, or to obtain more information from the same number of animals.
Refinement methods ease or minimize potential pain, suffering, or distress, and enhance animal welfare for the animals used. These include anesthetic and analgesic regimes for pain relief, housing and care measurements to meet the animals' natural needs, and humane endpoints (Stephens and Mak, 2014).

Regulatory acceptance and use of 3R models refers to the formal and actual acceptance and the use of these alternative approaches by regulatory authorities (i.e. licensing organizations[3], standard setting bodies[4] and product assessors[5]) and manufacturers for licensing, safety assessment and quality control purposes of regulated substances. The process of regulatory acceptance and use is further explained in section 2.4. Regulatory testing is laid down in product requirements (see section 2. 6) and habitually involves the use of laboratory animals. About 25% of the laboratory animals in Europe are used for regulatory purposes (see section 2.3). It entails procedures which frequently cause serious pain and distress to the animals involved. In addition, regulatory testing is often repetitive in nature -e.g. every new vaccine batch needs to undergo quality control testing to prove the safety and potency of the batch-. As a result, the slow acceptance of 3R models in this regulatory domain is often referred to as a serious problem by those committed to the 3Rs.

Slow or non-acceptance of 3R models is also objectionable in terms of the European legal obligation to use alternative approaches such as 3R models where available. Article 13 of Directive 2010/63/EU on the protection of animals used for scientific purposes states that: *''Member States shall ensure that a procedure is not carried out if another method or testing strategy for obtaining the result sought, not entailing the use of a live animal, is recognized under the legislation of the Union.*" (EU, 2010). Nonetheless the acceptance and use of 3R models in the regulatory domain is repeatedly observed as a difficult process (e.g. Schiffelers et al., 2007, Guy et al., 2008 a and b, Vandebriel and Opperhuizen, 2011, Van den Berg, 2011, Long and Griffin, 2012, Kooijman 2013).

1.2 Research questions

This thesis offers a deeper understanding of the process of regulatory acceptance and use in order to grasp this challenging process and examine possibilities to enhance 3R acceptance in the regulatory domain. For this purpose, a systematic scientific analysis is adopted of factors –i.e. drivers and barriers– influencing the central process. This analysis will help those committed to the 3Rs to comprehend and deal with the complexity of the issue and find ways to facilitate regulatory acceptance and use where possible. To analyze the central issue of this thesis and prescribe ways to address it, the following open-ended questions are formulated:

3 Such as the European Medicines Agency (EMA) and the European Chemicals Agency (ECHA)

4 Such as the European Pharmacopeia (Ph.Eur) for pharmaceuticals and vaccines and the OECD for chemicals.

5 Such as Official Medicines Control Laboratories (OMCLs) for pharmaceuticals and vaccines and ECHA/ National Competent authorities -many of them are ministries or agencies in the environmental sector- for chemicals.

Q1. How can regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes be defined? [6]

Q2. Which theoretical perspectives are needed to comprehend the process of regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes, and to find suitable ways of enhancing the process?

Q3a. Which factors influence the regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes?

Q3b. How do these factors influence the regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes?

Q4. How can these factors be influenced in order to optimize the process of regulatory acceptance and use of 3R models for risk assessment and efficacy testing?

This research focuses on the European setting. The European requirement to use 3R models where available and suitable (EU, 2010) has no equivalent in other parts of the world. This means that one can expect that there is a favorable climate within Europe regarding 3R models and that Europe in this sense can function as a frontrunner in terms of the acceptance and use of 3R model. Where relevant, a broader view is taken into account, for example the US level in Chapter 6 and OECD level in Chapter 7.

Furthermore, this research concentrates on the pharmaceuticals/biologicals and chemicals sectors. Regulatory animal testing is conducted for a range of products and substances varying from pharmaceuticals and biologicals, to food and feed products and from pesticides to industrial chemicals. The choice to focus on the sectors of chemicals and pharmaceuticals –including vaccines– is instigated by the fact that most laboratory animals in the regulatory domain are used for pharmaceutical/biological products and for 'other' toxicological evaluations (see section 2.3). Furthermore, these sectors include many of the animal models causing serious ethical and scientific concerns in terms of the repetitiveness of tests (e.g. vaccine batch control, see Chapter 6), the levels of pain and distress (e.g. induced by challenge tests which are still used for established vaccines, see Chapter 6), variability of test results (e.g. regarding the NIH test for rabies vaccine batch control, see Chapter 6) and the numbers of animals used (e.g. reproductive toxicity testing which accounts for a large percentage of the animals used in the chemical sector, see Chapter 7).

1.3 Regulatory acceptance and use of 3R models: the nature of the problem

In order to examine the central research questions, defining the nature of the problem is essential. A previous study of Schiffelers et al. 2007 (see Appendix I) revealed that regulatory acceptance and use of validated 3R models for risk assessment and efficacy testing purposes is a highly complex problem. The issue crosses geographical, institutional and sectorial borders and involves many different stakeholders – both public and private –

6 Efficacy testing in this context also refers to potency testing even though in practice there is a difference between the two. Real efficacy tests are conducted in the target animal, whereas potency testing uses a surrogate model.

often with diverging perspectives. Furthermore, it is characterized by a multilevel playing field and a highly risk-averse context. In this sense the issue of regulatory acceptance and use of 3R models has many features of a wicked problem (Rittel and Webber, 1973). Wicked problems are societal problems that display many interdependencies and multi-causality and they are often characterized by internally conflicting objectives and are often instable. Legislation, scientific evidence and perspectives are changing at the same time that policy makers are trying to address the problem. Wicked problems are socially complex, which means that there are many stakeholders involved with often as many different opinions and of which no one bears the full responsibility for the problem. In addition, there are no clear solutions which in turn might lead to unforeseen consequences (APS, 2007). As a result, they are difficult to define and solve and often overpower normal problem solving and project management approaches (APS, 2007).

The problem of regulatory acceptance and use of 3R models might not have a broad societal impact compared to well-known wicked problems like climate change, healthcare problems or problems of social injustice, but the issue reveals many of the features described above. It is difficult to clearly define the issue of regulatory acceptance and use of 3R models. To start with, the problem definition is observed to differ among different stakeholders and is therefore hard to objectify. In addition, there are no clear facts and figures with regard to regulatory non-acceptance. Several sources indicate that regulatory testing accounts for approximately a quarter of animal tests conducted in Europe, however there are no solid numbers available to define the magnitude of the problem of regulatory non-acceptance (see section 2.3.1). The barriers and drivers which are observed to influence the process of regulatory acceptance and use are also subjected to different perceptions which are the product of diverging situations and stakeholders involved. Therefore, it is important to obtain a clear and in-depth understanding of specific cases of regulatory acceptance and use. To fully grasp the complexity of the problem, an integrative approach is needed, which is found in the theory of technology transitions (see Chapter 4).

1.4 The relevance of innovation in the domain of regulatory animal testing

What is clear however, is that there is growing criticism linked to the limitation of the conventional animal model.[7] Nonetheless, these models manage to maintain a firm position within the regulatory domain. For example, in regulatory testing some animal models have been used in approximately the same manner for 40 to 70 years without rigorous evaluations. This is also the case for many of the animal models used for quality control purposes of vaccines (see Chapter 6) and in the chemicals domain where: *"…the review literature on the limitations of basic toxicological tools is astonishingly scarce"* while *"… in vitro tests undergo the most extensive evaluation of any model in the life sciences"* (Hartung and Daston, 2009, p.233).

7 For example regarding the limited predictive value of the animal models (e.g. Hendriksen and Van der Gun, 1995 on animal models and alternatives in quality control of vaccines, Ekwall et al., 1998 for acute toxicity, York and Steiling, 1998 for eye irritation, Bruckner et al., 2003 for rabies vaccine potency testing, Basketter et al., 2004 for skin irritation and Bremer et al., 2005 for reproductive toxicity).

Science is swiftly progressing, offering new testing strategies which prove much more robust when compared to the often variable *in vivo* models. Many authors (e.g. Richmond, 2002; Bruckner et al., 2003, Garthoff, 2005; Hendriksen, 2006; Blaauboer and Andersen, 2007, Cooper and Jennings, 2008, Bottini et al., 2008, Lilienblum et al., 2008, Hartung and Daston, 2009, Vandebriel and Opperhuizen, 2011, Romberg et al., 2012, Long and Griffin, 2012, Kooijman et al., 2013) stipulate the need to revise this strong dependence on animal studies and to tackle the slow acceptance of innovative 3R models. This is important for a variety of reasons:

- The strong and ever increasing understanding of the mechanistic foundations of pathophysiological processes are not sufficiently echoed in current *in vivo* testing (Hartung and Daston,2009, Hendriksen, 2009).
- New technologies confront hazard and risk assessment with new challenges which cannot be effectively tackled with the existing methods. (Hartung and Daston,2009) "*The introduction of new toxicity testing methodologies is intended to improve current risk assessment processes, making them more relevant by incorporating mechanistic information, accounting for differences in dose between in vitro and in vivo conditions, and focusing on human relevance."* (Blaauboer and Andersen, 2007, p.386.).
- Present approaches are often too conservative because of the large uncertainties that are connected to animal testing (Cooper and Jennings, 2008, Hartung and Daston,2009). Plus, most of the current testing strategies are not validated for the purposes of predicting human health risks (Blaauboer and Andersen, 2007) and some are highly variable in the test results they generate (e.g. Bruckner et al., 2003, Cooper and Jennings,2008).
- There is a need for assay systems that enable a higher throughput of large numbers of chemical substances that must be tested in the context of REACH (Hartung and Daston, 2009) or that better reflect progress being made in current production processes (De Mattia et al., 2011).
- Animal welfare is very often compromised by *in vivo* testing and the general public increasingly refutes the use of animal studies for purposes which do not absolutely necessitate them (Hartung and Daston, 2009) (see for example the European Civil Initiative).[8]

In other words, 3R models are in many instances observed to be valuable in terms of scientific progress, as well as in terms of animal welfare (e.g. to reduce animal numbers and levels of pain inflicted to the animals used). Nonetheless, it is very difficult for these innovations to enter the existing regulatory domain; i.e. to become part of the regulatory requirements, get accepted and used for safety and efficacy testing purposes of man-made substances such as chemicals and pharmaceuticals. For these reasons the process of regulatory acceptance and use of 3R models is an important issue to look into.

8 "Stop Vivisection" is the third European Citizens' Initiative submitted to the European Commission on 3 March 2015. It was signed by 1.17 million citizens. The Initiative asks the Commission to abrogate Directive 2010/63/EU on the protection of animals used for scientific purposes and put forward a new proposal aimed at phasing out the practice of animal experimentation, making compulsory the use - in biomedical and toxicological research - of data directly relevant for the human species. http://ec.europa.eu/citizens-initiative/public/initiatives/finalised/details/2012/000007 :consulted May 2015

1.5 Outline thesis

In this thesis the following steps are taken to answer the research questions (Q1, Q2, Q3a and Q3b: see section1.2) and the prescriptive question (Q4) (Figure 1 offers a schematic outline of the research design).

Chapter 2 defines regulatory acceptance and use of 3R models and describes the context in which it has to be placed (Q1). To this end, it contains a brief history of animal testing for regulatory purposes and the rise of 3R models in this specific area. It offers facts and figures regarding regulatory testing and the use of animal and 3R models for regulatory purposes and defines the process of regulatory acceptance and use of 3R models. In addition the regulatory framework in which regulatory acceptance and use has to be placed and the central stakeholders which play a role in this process are defined.

Chapter 3 presents the research approach and the research methods that are applied to investigate the influence of drivers and barriers (independent variables) on the regulatory process of acceptance and use (dependent variable). Moreover, the Chapter describes the philosophy underlying this research, which entails the ambition to contribute to understanding the existing tardiness and possible ways to address it.

Chapter 4 introduces the theoretical framework that is applied to explore and analyse the core process of regulatory acceptance and use of 3R models and the factors that influence this process and thereby targets Q2. Wicked problems are characterized by the fact that there are multiple causal factors that influence the central process. Addressing the complexity and uncertainties of such a problem requires covering the big picture as well as the interrelationships between the causal factors influencing the central process (APS, 2007). Therefore, an explorative integrative perspective is adopted to create a realistic picture of the examined issue. Different theoretical lines were investigated, such as innovation-, implementation- and risk regulation literature. Based on this, it was concluded that technology transition literature offers the integrative perspective needed. This means that the central query of this thesis i.e. how to understand and augment the slow transition from the conventional way of product risk assessment to a more innovative way of testing is first and foremost outlined as a technology transition issue.

Based on the theoretical insights as described in Chapter 4 and earlier empirical findings (Schiffelers et al., 2007, see Appendix I) the technology transition perspective is combined with a risk regulation perspective in Chapter 5 to create a tailor-made perspective for the analyses of empirical findings collected in the context of this thesis. Drivers and barriers which are observed to play a role in the transition towards the 3Rs are placed in a multilevel perspective. In this perspective three influencing levels are distinguished in technology transitions: the micro or niche level at which innovations are developed and tested; the meso level or the sociotechnical regime which encompass existing rules and regulations, testing infrastructure and connected knowledge and expertise; and the macro level or the sociotechnical landscape which represents broader societal developments e.g. political, economic, geographical and demographical developments in which

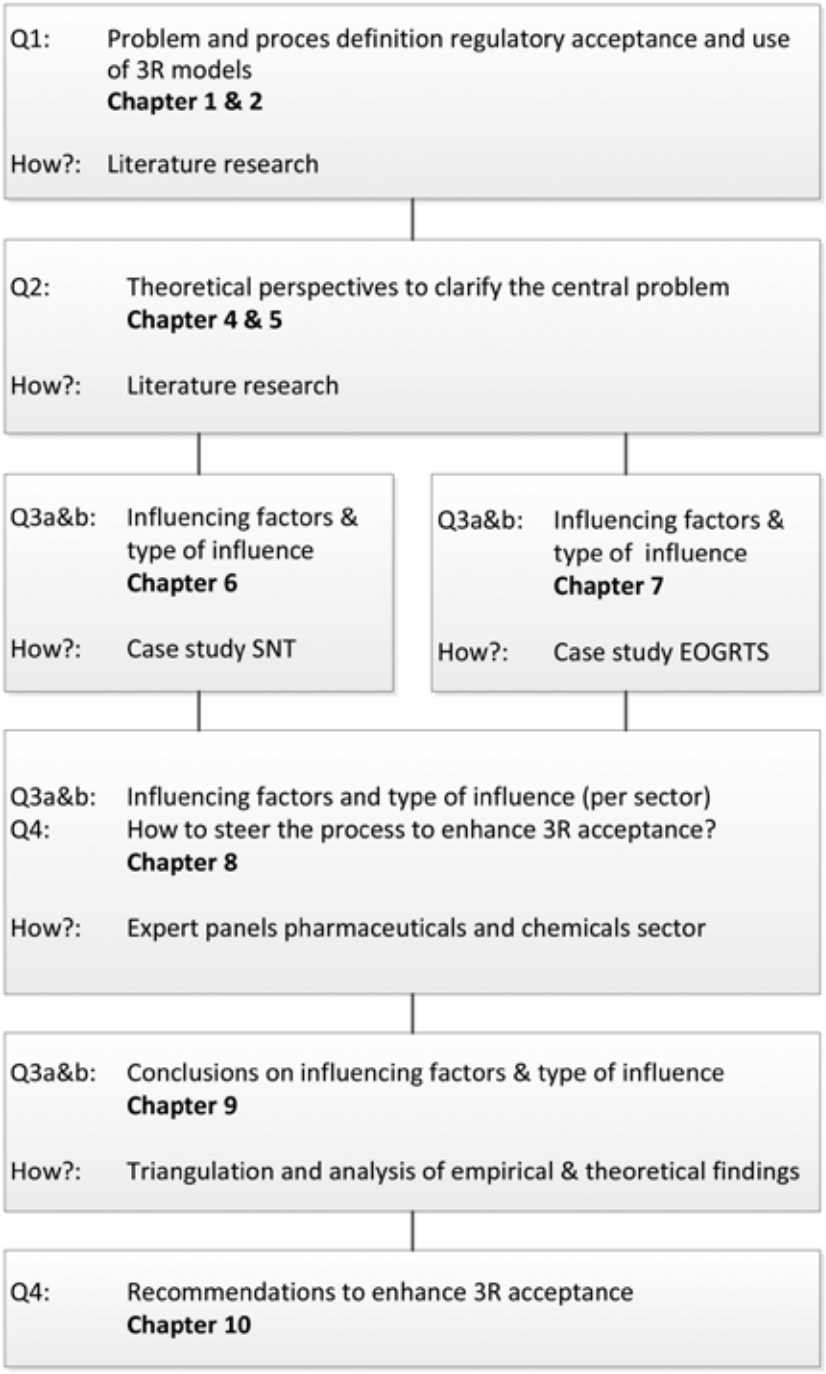


Figure 1. Schematic outline of research design

technology transitions are inserted. Through this analytical model Chapter 5 offers the remaining parts of the puzzle to answer Q2.

Chapter 6 and Chapter 7 each describe a case study to provide an in depth examination of the process of regulatory acceptance and use and the main factors that influence this process. Both case studies focused on describing and inspecting the deeper dynamics of regulatory acceptance and use and thereby target Q3a and Q3b. The veterinary rabies vaccine case study on the acceptance and use of the Serum Neutralization Test (SNT) in Europe to replace the NIH challenge test for potency testing purposes is presented in Chapter 6. This case study was anticipated by a survey on drivers and barriers on 3R acceptance for rabies vaccine potency testing purposes. The results of this survey (see Appendix II) served as input for the case study. Chapter 7 presents the case study on reproductive toxicity testing of chemicals. It describes the process of acceptance and use of the Extended-One Reproductive Toxicity Test (EOGRTS) in the context of the OECD and the EU.

Chapter 8 covers the results of two expert panels that were conducted to check the validity of the earlier found influencing factors (Schiffelers et al., 2007 and 2012) in the specific context of the product sectors of pharmaceuticals and chemicals. In addition, options to enhance this process were discussed for the three central stakeholder groups (regulators, manufacturers and academia). Chapter 8 thereby targets Q3a, Q3b and Q4.

The empirical chapters 6, 7 and 8 offer an overview of the main influences and dynamics we came across during this research. They thereby reflect the context of discovery which was the focal point of this research.

Subsequently, an inductive approach is adopted in Chapter 9 to reflect on the empirical findings through the use of the technology transition and risk regulation perspective as presented in the Chapters 4 and Chapter 5. Through this context of justification conclusions could be drawn with regard to the main influencing factors and there mutual interactions. This chapter thereby offers the answers to Q3a and Q3b.

Lastly, Chapter 10 offers an overview of optimization options (Q4) with the aim to overcome the existing barriers 3R models face in the regulatory domain and has a focus which is *"oriented towards action"* as Flyvbjerg calls it (Flyvbjerg, 2006: see philosophy underlying this research in Chapter 3). In addition this chapter discusses the scope and the limitations of this research and identifies the aspects that require further research and future discussion.

Table 1 summarizes the chapters of this thesis, the research questions, the research methods and the scientific articles that have been published in peer reviewed journals in the context of this PhD project. The chapters 5 to 8 consist of previously published articles. These chapters start by offering by a short readers guide to clarify the contribution of this chapter to the thesis.

Table 1. Summary of chapters, research steps & methods and publications

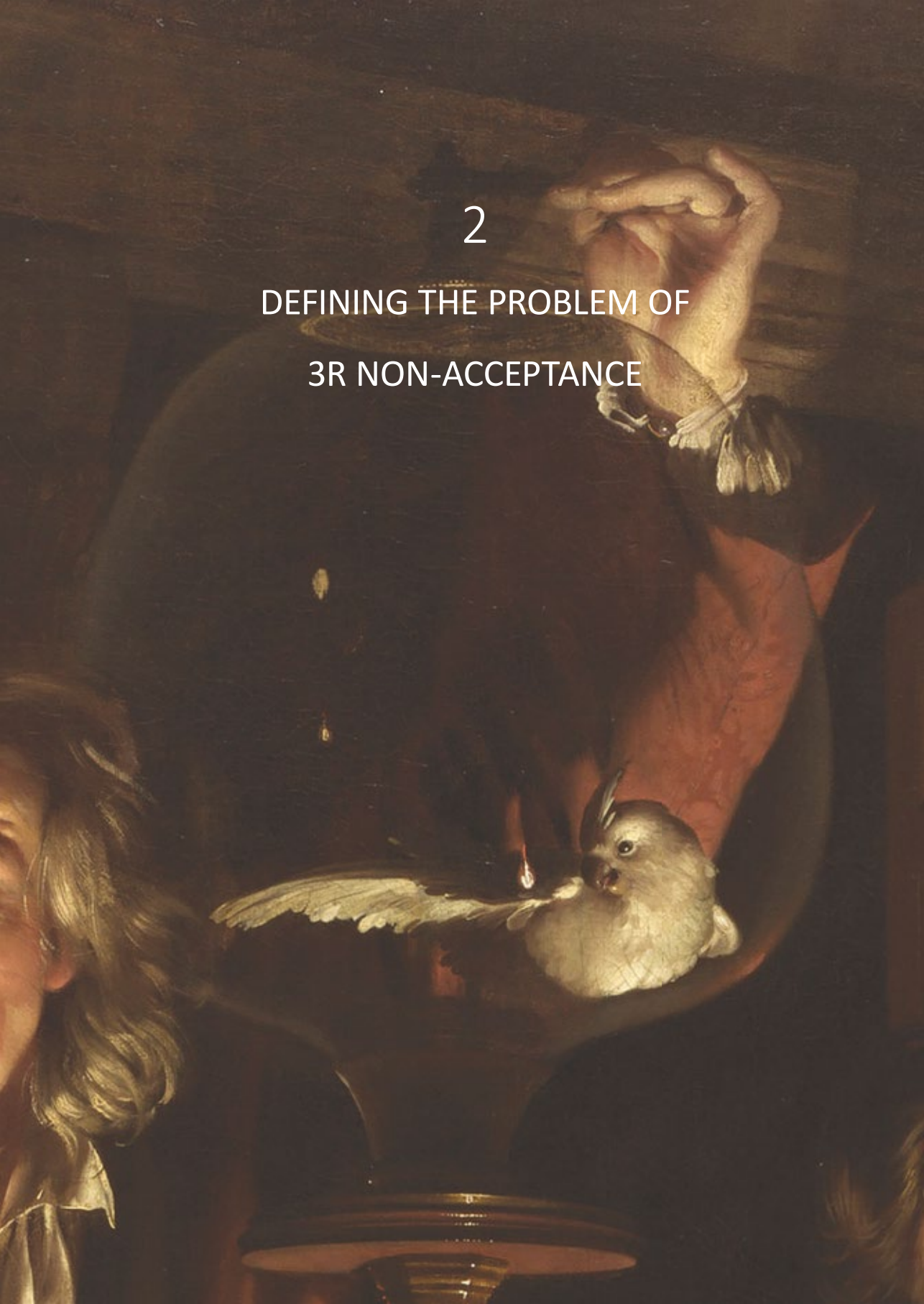
Chapter	Goal	Research method	Publication
Ch. 1	Introduction to research problem	Literature research	-
Ch. 2	Definition of research problem	Literature research	-
Ch. 3	Explication of research approach	Literature research	
Ch. 4	Investigation theories to analyze research problem	Literature research	-
Ch. 5	Presentation analytical frame & risk regulation perspective. Presentation 3R acceptance model	Literature research and previous empirical work (see Appendix I)	Schiffelers et al., 2012 *Altex*
Ch. 6	Case study 1 to examine drivers and barriers and acceptance process in the pharma viz. vaccines context	Case study veterinary rabies vaccine potency testing (SNT) (see also Appendix II)	Schiffelers et al., 2015a *Altex*
Ch. 7	Case study 2 to examine drivers and barriers and acceptance process in the chemicals context	Case study reproductive toxicity testing (EOGRTS)	Schiffelers et al., 2015b *Regulatory Toxicology and Pharmacology*
Ch. 8	Specification of drivers and barriers for pharma & chemicals sector and investigation optimizing options	Expert panels	Schiffelers et al., 2014b *Regulatory Toxicology and Pharmacology*
Ch. 9	Conclusions	Triangulation of empirical findings and theoretical reflection on these findings	
Ch. 10	Optimizing options	Based on empirical findings and theoretical reflections	
App. I	Initial investigation anticipating this thesis	Inventory of influencing factors based on literature research and interviews	Schiffelers et al., 2007 *Altex*
App. II	Investigation in preparation of SNT case	Survey on drivers and barriers to replace the NIH test in rabies vaccine potency testing	Schiffelers et al., 2014a *Biologicals*

1.6 Intended contributions to the scientific and social debate

Although many perceptions exist with regard to the process of regulatory acceptance and use of 3R models and why it is such a demanding process and numerous publications make notice of factors influencing the process of acceptance and use, there is little research available using a systematic approach in comprehending this process and its underlying dynamics. The current scientific debate on the use of 3R models in the field of safety and efficacy strongly focuses on the technical possibilities and limitations of the tests (e.g. validation of the test and interpretation/extrapolation of the test results). In other words the focus lies with influencing factors at the micro level of technology transitions (see Chapter 4). Far less attention is given to the developments at the meso-(socio technical regime) and the macro level (sociotechnical landscape) of technological transitions. And even less attention is being paid to the interaction between these levels. This thesis aims at filling that gap by offering such a systematic analysis and bringing insights from technology transition literature into the domain of 3R acceptance in the regulatory domain. This approach brings insight into individual factors – i.e. drivers and barriers – which are perceived to influence the process of technology transition towards 3R models and the way in which they are seen to influence this process. Moreover, the interaction between these factors is examined.

The initial study conducted by Schiffelers et al. in 2007 was a first step in adopting such a systemic approach. This study summarized general categories of influencing factors. However, influencing variables may differ from case to case and no case-specific information was obtained at that stage. Through this thesis case-specific data were collected as well as sector- specific information. The other way round, the characteristics of the acceptance of 3R models in the regulatory domain, in which risk minimization is observed to play a dominant role, offers the possibility to examine the specificities and the effects of the level of risk aversion on technology acceptance.

Furthermore, understanding the adoption of technology transition from a social science perspective is often ignored or overlooked as being too complex (NRC, 2004). This thesis holds the ambition to align the seemingly contradictive scientific and societal debates regarding technology transitions in general, and the use of alternative test approaches to replace, reduce and refine animal models in the regulatory domain in particular. By drawing the scenery of existing and often competing social constructions in this field (e.g. the dichotomy between the technical and ethical perspective or the innovative and precautionary perspective) and by examining the connected underlying motivations, a new way of looking, analyzing and dealing with the existing problem is offered which may help in solving the remaining controversies.



2

DEFINING THE PROBLEM OF 3R NON-ACCEPTANCE

"The first step toward change is awareness.
The second step is acceptance."

Nathaniel Branden
Canadian–American psychotherapist and writer
1930-2014

2.1 Introduction

In this chapter, the central issue of this thesis i.e. the regulatory acceptance and use of 3R models is specified. To this end several steps are taken. First of all, a short description is offered of the history of regulatory testing and the upcoming of the 3Rs principle. Subsequently, facts and figures about regulatory animal testing and the use of 3R models for regulatory purposes are given and the substages of the process of regulatory acceptance and use of 3R models are defined. To conclude, the central stakeholders and the regulatory frame in which regulatory testing has to be placed are described. This information is gathered through literature research and the analysis of policy documents. These steps offer a definite answer to Q1: *How can regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes be defined?*

2.2 Brief history of animal experimentation for regulatory purposes

The history of regulatory animal testing and the rise of the 3Rs principle are briefly described to apprehend the developments concerning 3R acceptance and use in the regulatory domain. The history of animal experimentation is characterized by Franco in his historical perspective on animal experiments in biomedical research (Franco, 2013). Although there are already notices of animal experiments in Ancient Greece, the use of live animals for experimental purposes took a giant leap in the 19th century when Charles Darwin introduced the principle of human beings being biologically comparable to animals in his seminal work *'On the Origin of Species'* (1859). This principle was adopted by Claude Bernard in 1865 in his book *'Introduction à l'étude de la medicine experimentale* which can be regarded as the starting point of animal testing for risk and efficacy assessment purposes. Bernard established animal experimentation as part of the standard scientific method to evaluate the effects of substances on human beings by concluding that experiments on animals are entirely conclusive for the toxicology of man.[9] Ever since, the number of animals used for experimental purposes rose sharply and medical science made huge progress. The 20th century showed world-changing developments in medical knowledge, such as the discovery of many therapeutics (e.g. hormones, antibiotics, new and safer vaccines, insulin, hemodialysis, chemo and radiotherapy for cancer etc.) and new diagnostic methods and surgical techniques, which dramatically improved the life expectancy in many countries around the world. According to Balls, many if not most of the benefits of these developments were based upon animal experimentation (Balls, 2009).

With the increasing number of products developed in the mid-20th century (e.g. pharmaceuticals, biologicals, agrochemicals, industrial chemicals and consumer products) and the occurrence of several incidents (such as the Cutter incident in 1955 and the Thalidomide incident in the late 1950s), the need for evaluating products in terms of risk and efficacy also increased. This resulted in a broad range of rules and regulations (see also section 2.3.1 and 2.5) on both sides of the Atlantic to regulate the risks connected to these products.

9 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3123518/ :consulted May 2015.

The initial tests to assess these risks were largely based on animal models and as these animal based tests became part of the regulatory product requirements, *in vivo* testing became institutionalized in the area of product assessment and efficacy testing.

In the 19th and 20th century a dichotomy became apparent between scientifically meaningful and medically relevant animal research on the one hand, and an increasing opposition to vivisection throughout Europe based on ethical arguments on the other (Franco, 2013). As a response to this discussion, William Russell and Rex Burch presented the 3Rs framework for the replacement, reduction and refinement of animal experiments to make progress both in terms of animal welfare as scientifically (Russell and Burch, 1959). Throughout the decades that followed the publication of Russell and Burch, many initiatives were taken to develop testing models that use less or no animals or are less stressful to the animals (see for example Blaauboer, 2015, Stephens and Mak, 2014, Hoonakker et al., 2011, Charton, 2008, for overviews of available 3R models). Under the influence of the Cosmetics Directive the development of 3R models for acute local toxicity endpoints took a leap forward and ever since, a broad range of replacement alternatives has become available for skin penetration, skin corrosion, skin irritation, and phototoxicity. However, replacing animal use for chronic endpoints is much more challenging due to the complexity of these endpoints.

In the last decades, the initial controversy between scientific and ethical arguments gradually shifted to a debate on diverging scientific arguments. Animal models face extrapolation problems due to the profound differences in anatomy, physiology, and genetics between the laboratory animal and the target species (e.g. Martić-Kehl, Schibli & August, 2012, Langley, 2009, Matthews, 2008, Knight, 2007, Hackam and Redelmeier, 2006, Bailey, 2005, Gerde, 2005, Pound et al., 2004, Piersma et al., 2014). Klein et al. for example indicated already in 1981 that the "*analgesic aspirin most probably would not have been marketed today because of its teratogenicity in rodents"* (Klein et al., 1981: as cited in Piersma et al., 2014, p.876). And "*nickel, the most important skin sensitizer in humans, is negative in the Local Lymph Node Assay, because mice lack the receptor to trigger the immune response."* (Schmidt et al. 2010: as cited in Piersma et al., 2014, p.876). As a result, there is increasing doubt about the scientific value of animal models for human beings (see Hendriksen and Van der Gun, 1995, Pound and Bracken, 2014, Van Meer, 2013). This leads to an increased attention for alternative approaches to animal use in the life sciences (Franco, 2013). Nonetheless, these approaches face difficulties in penetrating the area of risk and efficacy assessment.

2.3 Defining the magnitude of the central problem

Regulatory acceptance and use of 3R models is often referred to as a persistent and challenging problem by those committed to the 3Rs. However, the scope of the problem remains largely undefined. It proves to be very difficult to obtain figures to create a clear picture of the problem of regulatory non-acceptance and use of 3R models. This is, amongst other things, the result of diverging and blurry process definitions (e.g. where

does the development and production of a product end and the product assessment start) and of incomplete or even non-existing statistics when it comes to the use of 3R models for regulatory purposes. Consequently, it remains largely unclear which percentage of the available and validated 3R models is actually accepted and used for regulatory purposes, at which level and for which exact purposes. This means that we depend upon incomplete or incomparable information sources to draw the picture of regulatory non-acceptance of 3R models. In an attempt to deal with this lack of consistent information we consult and combine quantitative and qualitative sources of information on regulatory animal testing and 3R non-acceptance to give an overall idea of the magnitude of the problem.

2.3.1 Regulatory animal testing

Many of the toxicity and efficacy tests[10] described in these regulatory requirements were developed in the first half of the 20th century. Later on they became the routine procedures for product safety or efficacy testing for chemical or pharmaceutical products (Stephens and Mak, 2014) and as such part of the requirements to regulate product safety/efficacy. Regulatory animal testing thereby became firmly embedded in the regulatory requirements to which products such as chemicals and pharmaceuticals are subjected.

Regulatory animal testing can to a certain extent be quantified using the seventh statistical report on the number of animals used for experimental and other scientific purposes in the member states of the European Union as a starting point (EC, 2013).

The total number of animals used within the European Union for experimental and other scientific purposes in 2011[11] was just below 11,5 million (EC, 2013). Animal experiments within Europe are mainly conducted for basic- and applied research (e.g. biomedical research, efficacy testing of drugs and vaccines and toxicology tests) (see Figure 2). Regulatory testing is mainly covered by the categories *'Production and quality control of products for human medicine and dentistry/veterinary medicine' (respectively 10,97% and 2,94%)* as well as the category *'Toxicological and other safety evaluation' 8,75%)* [12] (see Figure 2).[13] [14]

10 Such as the LD50 test for acute systemic toxicity, the Draize test for eye irritancy and the NIH mouse potency challenge assay for efficacy testing of -rabies- vaccines.

11 With one Member State reporting for 2010.

12 Most toxicity testing is undertaken in the context of legal and regulatory requirements governing the use of particular types of chemicals in different parts of the world. (Nuffield Council on Bioethics, 2005).

13 This pie diagram reveals the main purposes of experiments in percentages. It reflects the situation in the 28 European member states in the year 2011 (with one MS reporting over the year 2010) (EC, 2013).

14 Almost 50% of the animals used in the category toxicological and other safety evaluation is done for the endpoints acute and sub-acute toxicity. Nearly 15% of the animals was used for the endpoints carcinogenicity, mutagenicity and toxicity to reproduction, while 22% was done to cover other toxicological and safety evaluation (EC, 2013). In particular, complex toxicological endpoints, such as carcinogenicity and reproductive toxicity, entail significant numbers of animals per test. Overall, products intended for medicine, dentistry and veterinary medicine require the highest proportion of animals for different types of tests i.e. approximately 39%.

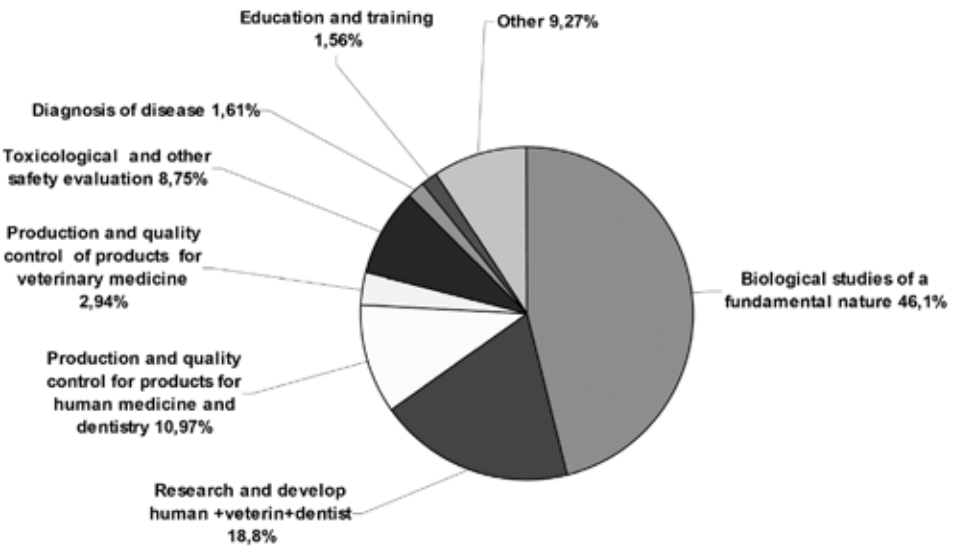


Figure 2. Purposes of experiments (Source: EC, 2013)

Together, the categories *'Production and quality control of products for human medicine and dentistry/veterinary medicine'* and *'Toxicological and other safety evaluation'* account for around a quarter (i.e. 22,66%) of the animal experiments conducted in the EU in 2011. Although not all of the tests within these categories will be conducted for regulatory purposes, it is fairly safe to state that most of the tests within these categories are done to meet regulatory requirements (EC, 2013).[15]

The Dutch statistics on the use of laboratory animals of 2013 show comparable percentages. In that year 22,25% of all animal procedures in the Netherlands were conducted to conform to legislative or regulatory requirements, for a single country (including countries outside Europe such as the US), the EU, the Council of Europe (CoE) or to meet any combination of these levels (NVWA, 2014).

Thus, according to these sources regulatory animal testing amounts to about a quarter of the animal studies in Europe[16] (Hartung and Daston, 2009, EC, 2013). This represents a total of about 2.6 million laboratory animals that are used for regulatory purposes in Europe each year. When extrapolated to a global figure, based on Taylor et al.'s estimate of 115.3 million animals used worldwide, approximately 26 million animals are used in regulatory testing (Taylor et al., 2008) and this may well be a conservative estimation (Knight, 2008).

15 Of the laboratory animals used in the EU, mice are the most common used species accounting for 61% of the total laboratory animal use, followed by rats at 14%. (EC, 2013). For regulatory testing various species are used. Again, rodents are the most commonly used species, however also larger animals including rabbits, dogs and primates are used for regulatory purposes (Nuffield Council on Bioethics, 2005)

16 No solid comparable data are available for other regions

Wagner et al. state that the number of animal used for experimental purposes *"is likely to rise in the coming years, given the ever increasing number of new products being developed and the significant increase in concern over the last decades regarding both hazards to human health and environmental pollution"* (Wagner et al., 2012, p.303).

2.3.2 3R non-acceptance in the regulatory domain

The core issue of this thesis however is not regulatory animal testing but the slow acceptance of 3R models in the regulatory domain. Regulatory acceptance of 3Rs methods is observed to include long lag periods of up to 11 years (Wagner et al., 2012) or longer. Apart from such indicators it proves very difficult to quantify problem 3R non-acceptance in the regulatory domain. The Alttox website offers an overview of 3R models for endpoints that have gained a certain status in the regulatory domain.[17] The website reveals 89 methods that were formally accepted for regulatory purposes at either the OECD level, the European or national level (see Appendix III). For vaccines safety and efficacy testing, such overviews have been provided by Castle, 1996 and Hoonakker et al., 2011.
Adler et al. published an analysis of the status of alternatives connected to the ban on animal testing for cosmetic ingredients. They point out that *"another decade will be needed to bring the science of alternatives up to the level would allow for full regulatory implementation."* (Adler et al., 2011: as cited in Piersma et al., 2014 p. 877).

However, these sources do not offer information on the numbers of 3R models that have up to date failed in gaining a regulatory status. Furthermore, it must be noted that these overviews only give information on the formal status of a 3R model. They do not offer any information on the actual regulatory acceptance and use by industry (ARA and UI: see section 2.3.5.) of a 3R model in daily practice.

In the absence of quantitative figures this section designates several qualitative sources of information that are available regarding the challenging implementation process of 3Rs in the regulatory domain.

- The problem of slow acceptance is acknowledged by experts in the field (e.g. Richmond, 2002, Scheel and Brekelmans, 2007, Bottini et al., 2008, Leist et al., 2008, Hartung and Daston, 2009, Storer, 2010, Stokes et al. 2012).
- Wagner et al. (2012) state that *"The provisions of the Directive 2010/63/EU that require alternative methods to be used instead of animal tests wherever available are not fully implemented in data requirements of relevant EU legislation, which has been the subject of serious criticism (Schiffelers et al., 2007). Our study found this criticism to be legitimate."* (Wagner, 2012, p.331). This legitimization is based upon the analysis of data requirements of EU legislation dealing with chemicals, biocidal products, plant protection products, and Novel Foods. Through this analysis Wagner et al. identified numerous endpoints in these data requirements that still require testing on animals for risk assessment, even though accepted alternatives are available.
- For chemicals, the European Chemicals Agency states in their latest report on the use of alternatives in testing for the REACH regulation, that companies are not fully implementing the use of – available – animal alternative methods (ECHA, 2014).

17 http://alttox.org/mapp/table-of-validated-and-accepted-alternative-methods/ :consulted May 2015

- Scholtz et al. even state that *"until now alternative approaches have only rarely been used in regulatory settings"* (Scholtz et al., 2013, p.507). This situation is less extreme for pharmaceutials and vaccines. For many of the animal tests as described in the European Pharmacopoeia (Ph. Eur.) monographs there are already alternative options available (see Hoonakker et al., 2011). And in an increasing number of instances these are accepted by the Ph. Eur. for example for routine-based quality control of vaccines batches.

To deal with the problem of regulatory acceptance and use there are even specific fora that have been established with the purpose to stimulate the use of 3R models in the regulatory domain. Examples are:

- The European Partnership for Alternative Approaches to Animal Testing (EPAA), a voluntary collaboration between the European Commission, European trade associations, and companies from seven industry sectors with the overall aim to replace, reduce and refine animal use in regulatory testing.
- EURL ECVAM's[18] Network for Preliminary Assessment of Regulatory Relevance (PARERE) which aims at advancing the process of regulatory acceptance of alternative methods.

These aspects show that regulatory acceptance and use of 3R models is an area of concern. To be able to examine the process of acceptance and use we will now define what is precisely meant by regulatory acceptance and use of 3R models in the context of this thesis.

2.4 Defining the process of regulatory acceptance and use

Regulatory acceptance and use of 3R models is a process which, for the sake of clarity, needs to be divided into three substages:

1. The formal incorporation of a 3R model into regulatory requirements (FI);
2. The actual regulatory acceptance of a 3R model by regulatory authorities (ARA) and;
3. The use of the 3R model for regulatory purposes by end users such as industry (UI).

Regulatory authorities[19] and industry are defined as the central stakeholders in the process of replacing, reducing or refining conventional animal models by alternative test methods (Schiffelers et al., 2005).

18 The Joint Research Centre (JRC) hosted European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) is actively involved in the development and promotion of alternative approaches to animal testing with the aim to replace, reduce or refine the use of laboratory animals (the 'Three Rs') in the safety assessment of chemicals and the quality control of biologicals (e.g. vaccines).

19 Appendix IV offers information on the regulatory authorities one has to think about in the context of this research.

2.4.1 Formal Incorporation (FI)

The first step of regulatory acceptance refers to the formal adoption or formal incorporation (FI) of a (validated) test method by a regulatory agency/authority.[20] This substage covers the formal uptake of a 3R model into regulatory requirements. Depending on the sector, FI can be accomplished at a national, a European (e.g. the monographs of the European Pharmacopoeia or the Test Method regulations of REACH) and/or a global level (e.g. the OECD test Guidelines). Generally speaking the most favourable situation is FI at the highest possible level. If an alternative for example is accepted at the OECD level, the connected lower levels are informed about the test and are stimulated to incorporate the test at their own level. However, this level of regulatory acceptance is a challenging exercise. For example, to amend an OECD testing protocol, all 30 member states must agree to the alterations. Furthermore, FI at this level does not automatically lead to FI at the lower levels.

2.4.2 Actual Regulatory Acceptance (ARA)[21]

The process of regulatory acceptance and use of a 3R model is not completed after its FI. To begin with, it requires the actual regulatory acceptance and use of the 3R model by regulatory authorities for regulatory purposes (ARA). *"Developing and validating alternative methods for regulatory purposes can only make sense if at the end of the process regulatory authorities accept it for registration or authorisation of a product or compound."* (EPAA, 2007). ARA means that a testing model is accepted by a specific regulatory authority to demonstrate product safety or efficacy. Habitually the FI of a 3R model is needed before ARA will occur. However, there are examples in which a model is not described in regulatory requirement but is nonetheless accepted by certain regulatory authorities. The other way round also occurs. This means that a 3R model is formally accepted as an alternative for the conventional animal models but is not (yet) accepted by certain regulatory authorities. Full ARA means that a 3R model is accepted by all regulatory authorities. In practice, full ARA is very challenging and ARA by one regulatory level is no guarantee for ARA at other regulatory levels.

2.4.3 Use by Industry (UI)

Use by Industry (UI) is strongly connected to FI and ARA. The broader the Formal Incorporation (FI) and Actual Regulatory Acceptance (ARA) of a 3R model the more likely its UI will be. FI or ARA of a 3R model by one of the lower level authorities is only of use to a manufacturer if he wants to market his product in that particular market. Manufacturers in the fields of pharmaceuticals, biologicals, chemicals etc. however, almost always operate in

20 http://alttox.org/ttrc/validation-ra/ :consulted May 2015

21 ARA together with the step of Use by Industry (UI) is often referred to as the implementation phase. In the field of policy science however, implementation would cover the whole process from the initial intention to work towards alternatives to the actual uptake. Therefore the term implementation might cause confusion. For this reason we have specified the term implementation by the substages of actual acceptance and use of a 3R model for regulatory purposes by regulatory authorities (ARA) and the use of a 3R models to meet regulatory requirements by industry (UI).

a global market. This means that they will have to adhere to the different requirements simultaneously, often resulting in a situation in which manufacturers stick to the strictest test regimes that they will encounter. And even though *in vitro* and other alternative tests play an important role in corporate decision-making on product formulation and product safety and efficacy, they are not necessarily considered definitive in the regulatory context. Corporations often follow up on their alternative testing with the historical animal-based methods to assure that there product will be accepted.[22]

Manufacturers and regulatory authorities at a national or supranational level both have the potential to play a pivotal role in terms of the acceptance of 3R models for regulatory purposes since they are respectively the producers and the assessors of products and substances. There is a substantial amount of legislation offering discretionary space to assessors and manufacturers of products/substances in choosing the most suitable test model. These requirements often only define the endpoints a product has to be tested on. Both stakeholder groups therefore have the possibility of using the available discretionary space offered in product requirements to use 3R methods. The following section offers an extensive description of the regulatory framework in which 3R acceptance in the regulatory domain has to be placed. This is important for understanding the legal possibilities and constraints that 3R models face.

2.5 Regulatory framework

Regulatory acceptance and use of 3R models within Europe is in fact the outcome of a deliberation of these central stakeholders between two types of legislation i.e. horizontal and vertical legislation (see Figure 3).

2.5.1 Horizontal legislation

The first type of legislation is 'horizontal legislation' pertaining to animal experimentation and multilateral agreements which includes pieces of legislation that aim at regulating the use of animals for scientific purposes, such as Directive 2010/63/EU on the protection of animals used for scientific purposes (EU, 2010). This Directive is the updated version of Directive 86/609/EEC which was largely based on European Treaty (ETS) 123 of the Council of Europe (CoE). However, whereas the primary goal of ETS 123 was the moral obligation of the CoE member states to protect laboratory animals, the primary goal of directive 86/609/EEC was a level playing field for the different stakeholders involved in the field of animal experimentation (De Leeuw, 2004). More recently, animal welfare became enshrined as a value of the Union in Article 13 of the Treaty on the Functioning of the European Union (TFEU). This value was also adopted in Directive 2010/63/EU (EU, 2010) which states in Article 12 that: *"Animals have an intrinsic value which must be respected. There are also the ethical concerns of the general public as regards the use of animals in procedures. Therefore, animals should always be treated as sentient creatures and their use in procedures should be restricted to areas which may ultimately benefit human or*

22 http://alttox.org/validation-and-acceptance-status-of-alternatives-2/ :consulted May 2015

animal health, or the environment. The use of animals for scientific or educational purposes should therefore only be considered where a non-animal alternative is unavailable. Use of animals for scientific procedures in other areas under the competence of the Union should be prohibited."

Through this revision, the 3Rs became the primary principle of the European legislative documents regulating animal use in science. There is no global equivalent to the European legislation which is this explicitely formulated. Nonetheless, there is also commitment to the implementation of the 3R-principles at other regulatory levels such as the OECD[23] and the European Pharmacopoeia (see section 2.5.2).

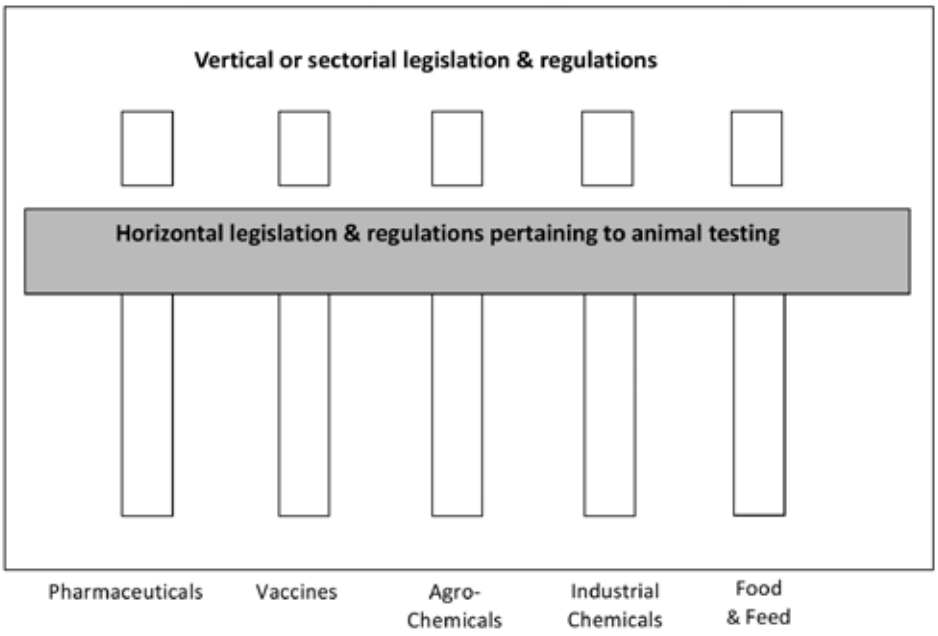


Figure 3. Horizontal and vertical legislation (Source: Schiffelers et al., 2005)

23 The OECD is also important in terms of another piece of horizontal legislation relevant in this context; i.e. Mutual Acceptance of Data (MAD) (OECD, 2005). MAD states that *"data generated in the testing of chemicals in an OECD member country in accordance with OECD Test Guidelines and OECD Principles of Good Laboratory Practice (GLP), shall be accepted in other member countries. Adhering non-member countries, of which there are an increasing number, should also accept this data. The data should be accepted for purposes of assessment, and other uses relating to the protection of man and the environment."* According to the OECD website, the OECD Council Decisions regarding MAD save thousands of animals every year by avoiding double testing.

2.5.2 Vertical legislation

The second type of legislation which is pivotal in terms of the process of regulatory acceptance and use of 3R models is the 'vertical' or 'sectorial legislation'. This regulates the activities of particular products, for example in terms of their licensing and marketing, to protect human beings, animals and the environment from potential adverse effects. In Europe alone there is an estimated amount of 800 regulations and requirements that prescribe animal experimentation for safety assessment purposes (De Leeuw, 2004). Several European examples in the field of pharmaceuticals and industrial chemicals are:

- REACH (the EU Chemicals Regulation (EC) No. 1907/2006) dealing with the Registration, Evaluation, Authorization and Restriction of Chemical substances in the EU.
- The Test Methods Regulation (TMR), Regulation (EC) No. 440/2008, which lays down the legally binding EU standard test methods to determine the hazardous properties of chemicals.
- Directive 2001/83/EC on the Community code relating to medicinal products for human use.
- Directive 2001/82/EC on the Community code relating to veterinary medicinal products.
- European Pharmacopoeia defining the requirements for the qualitative and quantitative composition of medicines (including vaccines), the tests to be carried out on medicines and on substances and materials used in their production.

Manufacturers operating at a global level will have to face a broad set of additional requirements for that specific product in other parts of the world. Different products have to meet different standards. This means that sectorial legislation for different products often cover diverse objectives and unalike mechanisms or procedures to achieve these objectives. The following objectives can be defined: licensing and market authorization of products (e.g. pharmaceuticals/biologicals, industrial and agrochemicals), batch control in terms of safety and/or efficacy of new batches of a certain product (e.g. biologicals and shellfish) and industry responsibility and market surveillance schemes (e.g. cosmetics). Often a mix of these objectives is aimed at. Whereas the Cosmetics Directive primarily requires consumer safety, the European chemicals legislation REACH aims at a combination of the protection of human health and environmental protection while the pharmaceuticals legislation aims at quality, safety and efficacy of products in order to protect public and animal health.

Besides, the status of the requirements differs. In some sectors test methods are bound by regulation, in others they are subject to guidance (EPAA, 2007) offering unalike levels of discretion to deviate from the tests described.[24] The Cosmetics Directive for example provides a fairly flexible testing framework whereas the crop protection directive establishes a strict framework for active substances and plant protection products in which endpoints and test methods are stipulated, regardless of tonnage bands. Testing requirements under REACH mainly depend on the volume of chemical substances to be registered

24 Most regulations do not exclude the application of other methods, such as in-house tests. This means that deviation from described test is possible if needed for specific scientific reasons or to meet certain product characteristics (EPAA, 2006).

for marketing.[25] Below a more detailed description is given of the European framework regulating the two sectors this thesis focuses on (i.e. pharmaceuticals including vaccines and – industrial – chemicals).

2.5.2.1 Regulation of pharmaceuticals including vaccines

Pharmaceuticals and vaccines in the European Union are regulated through a stepwise approach. To start with, these products have to obtain a marketing authorization before being sold. Marketing authorization can only be granted after evaluating the risks and benefits of the product. This evaluation is based on the dossier provided by the manufacturer presenting the data collected during the product development and clinical trials.[26] If the product is intended to be registered for the entire European market, the 'Centralized Procedure' is followed which falls under the responsibility of the European Medicines Agency (EMA). The application is evaluated by the EMA. However, the final decision is taken by the European Commission that issues a marketing authorization valid throughout the EU, Iceland, Liechtenstein and Norway.[27] An alternative route is formed by the national procedures which are reserved for products that are licensed in a single country.

In the case of vaccines, once a license for marketing is obtained, each new batch of vaccines is subjected to quality control to assess its safety and efficacy before being released for use. In addition, all vaccines and pharmaceuticals are monitored after release onto the market for potentially adverse events (pharmacovigilance).

The European Pharmacopoeia (Ph. Eur.) provides the scientific and technical standards to guarantee a minimum quality of pharmaceutical products including vaccines. The Ph. Eur. intends to harmonize the safety and quality control of products. It does not contend with authorization matters. Groups of experts from countries that have signed the Convention of the Ph. Eur. formulate the requirements for pharmaceuticals and immunobiologicals (e.g. vaccines). The tests which are recommended to use are laid down in the monographs of the Ph. Eur. These include tests for registration purposes (for pharmaceuticals including vaccines) and batch quality control (for vaccines) and very often consist of animal models.

Of the animals used in Europe for the production and quality control of products for human medicine and dentistry and for veterinary medicine[28] 47% was used to satisfy requirements from the EU, the Council of Europe, national legislation and legislation outside of the EU simultaneously. Testing conducted to satisfy EU legislation including the European Pharmacopoeia covered 35,9% (see Figure 4).

25 Under REACH testing can in some instances even be waived based on exposure considerations.

26 http://www.vaccineseurope.eu/about-vaccines/eu-regulatory-framework-for-vaccines/ :consulted May 2015

27 http://www.vaccineseurope.eu/about-vaccines/eu-regulatory-framework-for-vaccines/ :consulted May 2015

28 These catergories account for 13,9 % of the total number of animals used for experimental purposes in Europe

When validated alternative methods become available, the monographs should be revised to incorporate the alternative method. Furthermore, the Ph. Eur. leaves a certain amount of discretionary space to the regulatory authorities and manufacturers to choose the method they see as most suitable. However, the FI of a 3R model into the monographs and the available discretionary space do not automatically lead to the ARA by national authorities and/or UI by manufacturers as will be elucidated in Chapter 6 of this thesis.

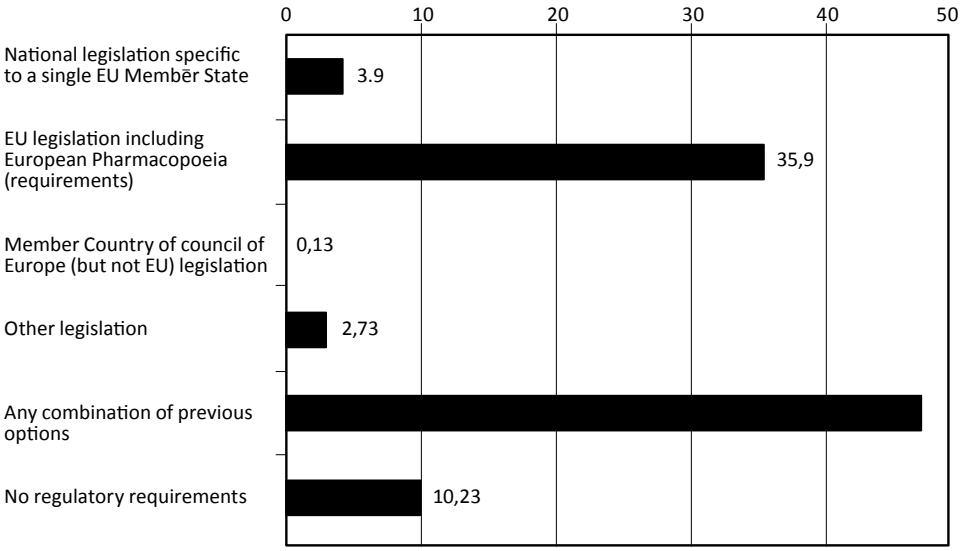


Figure 4. Percentages of animals used for regulatory requirements for production & quality control of products and devices for human medicine, dentistry and for veterinary medicine. (Source: EC, 2013)

2.5.2.2 Regulation of chemicals

The European Union (EU) has modernized its European chemicals legislation through REACH, an integrated system for the registration, evaluation, authorization and restriction of chemicals. Its objective is to improve the protection of human health and the environment, even as maintaining competitiveness and strengthening the spirit of innovation in Europe's chemicals industry. At the same time the European Chemicals Agency (ECHA) was set up, to deal with the day-to-day management of REACH requirements and to ensure the consistency of the decision-making at Community level.[29] The Agency also manages the registration process and plays a key role in the evaluation process and the authorization and restriction procedures.

REACH requires firms, that manufacture and import chemicals, to evaluate the risks resulting from the use of those chemicals and to take the necessary steps to manage any identified risk(s). The burden of proof that these chemicals are safe lies with the industry.

29 http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3Al21282 :consulted May 2015

Registration is the key element of the REACH system. Manufacturers are obliged to register their chemicals in a central database if manufactured or imported in quantities of more than one ton or more per year. Substances may only be produced or placed on the European market once registered.[30] Registration requires that manufacturers and importers provide information on the properties and use of chemicals and the safety measures to be taken when using them. The type of data required relates to the production volume of and the risks connected to the substance.[31] [32]

The European Chemicals Agency (ECHA) manages the database, receives registration dossiers and develops technical guides with the aim to assist manufacturers, importers and the competent authorities in implementing the REACH requirements. Evaluation enables ECHA to check whether the industry fulfills its obligations and whether they avoid tests on vertebrate animals when unnecessary. Two types of evaluation are provided for: dossier evaluation and substance evaluation. Dossier evaluation is compulsory for any application which needs to carry out tests specified in Annexes IX and X to the Regulation (i.e. the most stringent tests, mostly involving the use of vertebrate animals). The aim is to minimize the need for such experiments and to check the conformity of a registration.[33]

Next to ECHA, each of the EU member states has its own authority with the competence and resources to carry out the tasks assigned to them under REACH. Member state experts/representatives are members of ECHA's Management Board, the Agency's Committees, the Forum and several networks.

In addition to REACH, it is important to refer to the OECD Guidelines for the testing of chemicals. These are a collection of the most relevant internationally agreed testing methods used by governments, industry and independent laboratories to assess the safety of chemical products. They are primarily used in regulatory safety testing and subsequent chemical notification and registration.[34] The OECD guidelines provide a strong guidance in the chemical, pesticides and cosmetics field as to which model needs to be used. However, they are non-binding and regulatory authorities such as ECHA have discretionary space to choose the test protocol they perceive as most suitable and are allowed to conduct or ask for additional tests, if they feel this is necessary to guarantee product safety within their region. This implies that the FI of a 3R model into the OECD guidelines does not lead to an automated ARA at the lower levels (European, national) (see Chapter 7 of thesis for an example of this).

30 http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3Al21282 :consulted May 2015

31 http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3Al21282 :consulted May 2015

32 e.g. extensive toxicity tests are required for substances of very high concern (SVHC's) such as carcinogenic, mutagenic and reprotoxic (CMR) substances and persistent, bioaccumulative and toxic (PBT) substances and for substances manufactured or imported in quantities of more than 1000 tonnes.

33 http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3Al21282 :consulted May 2015

34 http://www.oecd.org/chemicalsafety/testing/oecdguidelinesforthetestingofchemicalsandrelated documents.htm :consulted May 2015

Of the animals used in Europe for toxicological or other safety evaluation,[35] 56% was done to meet a combination of regulatory requirements. Testing required under EU legislation represented 21,27% of the tests conducted in this area (see Figure 5).[36]

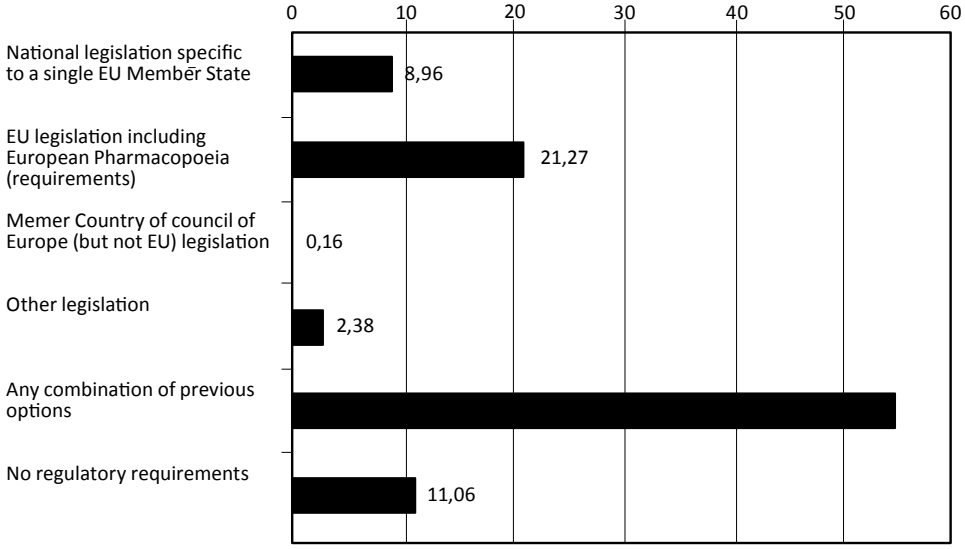


Figure 5. Percentages of animals used for regulatory requirements for toxicological and other safety evaluation. (Source EC, 2013)

2.5.3 Relation between horizontal and vertical legislation

Vertical legislation has to take horizontal legislation into account. This means that Directive 2010/63/EU, has to be taken into consideration by vertical pieces of legislation. Several pieces of legislation have effectuated this by referring to the spirit of the Directive. The European Pharmacopoeia Commission for example, in the view of the Directive 2010/63/EU, started to evaluate the texts of the Pharmacopoeia that recommend alternatives to animal tests, in order to make this information available to users and thereby encouraging the use of 3R models.[37] A comparable statement is made by the European Medicines Agency (EMA) in their draft guideline on regulatory acceptance of 3R testing approaches. *"Directive 2010/63/EU on the protection of animals used for scientific purposes, which is fully applicable to regulatory testing of human and veterinary medicinal products, unambiguously fosters the application of the principle of the 3Rs (Replacement, Reduction and Refinement) when considering choice of methods to be used."* [38]

35 8,75% of the total number of animals used for experimental purposes in the EU

36 Testing reported under 'no regulatory requirements' for example refers to in-house methods to verify the safety and efficacy of veterinary biologicals and medicinal products carried according to company's or known international standards (EC, 2013)

37 http://www.edqm.eu/en/European-Pharmacopoeia-news-43.html :consulted May 2015

38 http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/10/WC500174977.pdf: consulted May 2015

Furthermore, REACH requires that animal testing is used *"only when there are no other scientifically reliable ways of assessing the potential effects on humans or the environment."* [39] As a result testing on vertebrate animals under REACH is only allowed as a last resort and registrants of industrial chemicals are obliged to use alternatives to animals whenever possible.[40] However, untill now the Cosmetics Directive is the only Community regulatory framework with the aim of successively phasing out animal testing. It has established a prohibition to test finished cosmetic products and cosmetic ingredients on animals (testing ban), and a prohibition to market in the European Community, finished cosmetic products and ingredients included in cosmetic products which were tested on animals (marketing ban).[41]

2.6 Conclusion

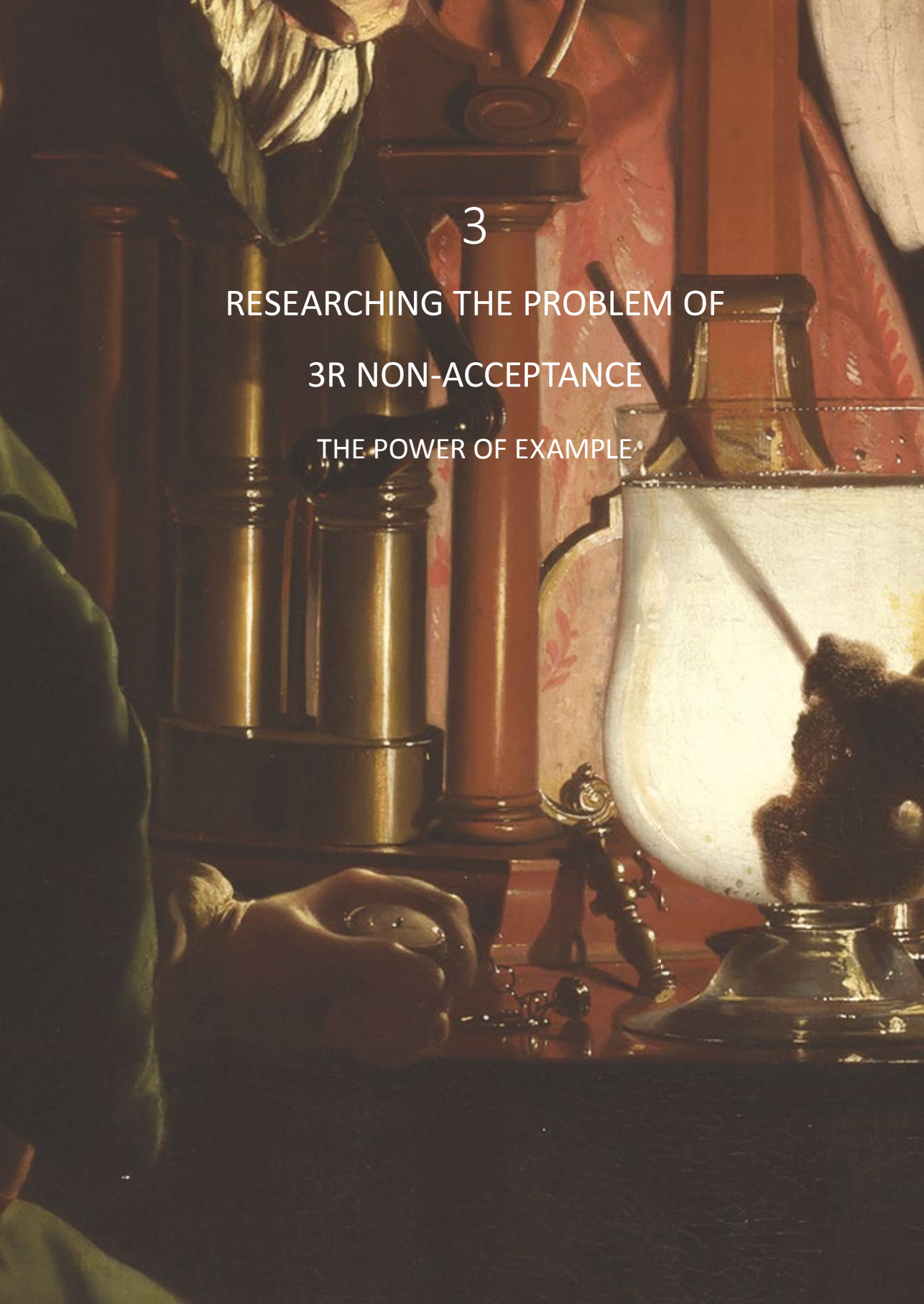
In this chapter several aspects have been describes to answer Q1: *How can regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes be defined?* To start with, regulatory testing has been put in a historical context to comprehend the developments so far. The book of Russell and Burch (1959) was a starting point in terms of the 3Rs principle. From that stage on, the attention for alternative approaches increased for moral, scientific and economic reasons and many 3R models have been developed ever since. However, their regulatory acceptance and use still proves challenging. Even though the problem of non-acceptance is difficult to quantify, there are many indicators stipulating the need of targeting this problem (see section 2.3). Regulatory acceptance and use is a process which consists of three substages: the Formal Incorporation of 3R models into regulatory product requirements (FI), the Actual Regulatory Acceptance of these models by regulatory authorities (ARA) and the Use of the models by industry to meet regulatory product requirements (UI). The central actors within this process are regulatory authorities and manufacturers. They are responsible for the safety and efficacy evaluation of the substances prior to their release to the European market. This evaluation is regulated by a broad set of regulatory requirements indicating the tests that need to be conducted.

However, most requirements offer discretionary space to these central actors to choose the most suitable test options and to introduce alternative ways of testing. In addition Directive 2010/63/EU stipulates the need to use 3R models where available. As such, both regulatory authorities and manufacturers have possibilities to enhance the challenging process of 3R acceptance and use in the domain of risk assessment and efficacy testing.

39 http://alttox.org/mapp/toxicity-testing-overview/ :consulted May 2015

40 European Commission's Joint Research Centre's European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) provided the European Chemicals Agency -ECHA- in September 2014 with a review of non-animal approaches for the assessment of several human health and ecotoxicological endpoints. Although this report, entitled *Alternative Methods for Regulatory Toxicology -A State-of-the-Art Review*, is not official EU policy, it provides information on "alternative" methods that ECHA can take into consideration in implementing regulatory processes for REACH, Biocidal Products, and Classification, Labelling, and Packaging (CLP).

41 http://ec.europa.eu/consumers/sectors/cosmetics/files/doc/antest/(2)_executive_summary_en.pdf :consulted May 2015



3

RESEARCHING THE PROBLEM OF 3R NON-ACCEPTANCE

THE POWER OF EXAMPLE

"…true expertise is based on intimate experience with thousands of individual cases and on the ability to discriminate between situations, with all their nuances of difference, without distilling them into formulas or standard cases."

Bent Flyvbjerg (2006, p.23)
Danish economic geographer

3.1 Introduction

This chapter describes the research approach of this thesis. Previous research (Schiffelers et al. 2007: see Appendix I) identified a variety of influencing factors in the process of non-acceptance of 3R models for regulatory purposes. However, the conclusion drawn was that every instance of non-acceptance is characterized by its own specificities in terms of the type of product, the region, the regulatory framework, the existing techniques and the proposed innovations. Reality is shaped by the interaction of stakeholders with these specific artifacts. In other words, the process of regulatory acceptance and use is viewed upon as a process in which individual stakeholders, organizations, institutions and networks give meaning to the techniques they work with as well as to the new developments that arise. These meanings are largely shaped by their context (see section 3.2. for an elaboration on this conceptualization). This also counts for the drivers and barriers that stakeholders experience when it comes to regulatory acceptance and use of 3R models. These are also dependent on the specific context in which regulatory acceptance of a specific 3R model has to be accomplished. To do justice to these specificities the main research approach used in this thesis is that of case studies (see section 3.4). In section 3.3 an overview is given of the research design of this thesis.

3.2 Ontology [42]

The assumptions about the nature of reality and the sources of knowledge (ontology), which are at the basis of this research, including the connected methodological choices (epistemology), are a reflection of Bent Flyvbjerg's philosophy described in his book *Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again* (Flyvbjerg, 2001). Flyvbjerg discusses how social sciences are often judged in terms of natural sciences and are expected to adhere to typical epistemic qualities, i.e., producing an epistemic theory that is predictive and explanatory. However, in the social sciences there is no single reality and social phenomena are very much shaped by their context. Early induction would hold the danger of narrowing down the perspective too early and thereby loosing valuable insights and the broader picture, a risk that is beautifully illustrated by the Indian tale of the six blind men and the elephant (see Appendix V). Therefore using a comprehensive perspective is very important. This is even more important in the context of this thesis in which a social sciences perspective is adopted to examine the decisions made by a multidisciplinary field and in which a broad range of values, realities and criteria are observed to play a role.

Due to the strong context dependence, social sciences have not and will not succeed in producing general, predictive, context-independent theories. The value of social sciences lies in its aspect of *phronesis,* which was already introduced by Aristotle and is *"often translated as 'prudence' or 'practical common sense' concerns values and goes beyond episteme (analytical, scientific knowledge) and techne (technical knowledge or know-how)…."* (Flyvbjerg, 2012, p.26). *Phronesis* balances instrumental rationality by

42 The ontology reveals which views people have with respect to the reality they live in. It is the philosophical study of the nature of being, becoming, existence, or reality, as well as the basic categories of being and their relations.

value rationality which means that it ensures that scientific and technical development does not take place without ethical checks and balances. For this, a reflexive approach to reality is adopted that involves contextual structure, value weight considerations, and balance in decision-making (Kuschel, 2012). This balancing of the instrumental rationality by value rationality was also embraced by Weber in his ideas regarding the establishment of society. In Weber's division between instrumental and value rationality (*Wertrationalität*) *"value rationality produces acceptance not by the merit of a logical internal structure, but by the non-rational elements present in the concepts* (Kuschel, 2012, p.6). Weber embraced value rationality to counterbalance the pretention of his timeframe, in which society could be constructed purely based on objective facts and figures. Value rationality was needed according to Weber to prevent the society from becoming too technocratic.

A scientist who conducts research based on value rationality sees it as his task to analyze the values that stakeholders associate with a certain theme to gain insight in dominant perspectives and their consequences in a specific domain (Flyvbjerg, 2012). In the field examined by this thesis there is a clear tendency to look at the process of non-acceptance in a very technocratic way (e.g. more data and efforts to optimize 3R testing strategies are needed to solve the problems). Even though such technological efforts are needed, a primarily technocratic approach is insufficient when it comes to tackling the existing barriers and looking for potential drivers. This research therefore aims at broadening this technocratic perspective by exploring the connected stakeholder perceptions.

Social research issues are best understood through narrative inquiries that develop descriptions and interpretations of the phenomenon from the perspective of involved stakeholder groups. This is a manifestation of social constructionism, which is one of the key concepts of sociology.[43] Social constructionism investigates *"the process whereby people continuously create, through their actions and interactions, a shared reality that is experienced as objectively factual and subjectively meaningful"* (Wallace and Wolf 1999, p. 277). These meanings jointly form the social reality (Boeije, 2005) and understanding this social reality is crucial when trying to comprehend the developments regarding a certain issue. The barriers and drivers influencing regulatory acceptance and use of 3 R models are also socially constructed and thereby rooted in underlying beliefs and understandings. For this reason, an interpretative research approach is adopted in the context of this thesis. Interpretative research aims at comprehending this social reality and the underlying dynamics through which it is formed (Tijmstra & Boeije, 2011).

Furthermore, the goal of social research with a *phronetic* approach is to come to analyses and interpretations regarding the values and interests in society with the goal of social change. It concerns *"the analysis of values as a point of departure for managed action." a*nd well managed action requires contextualism (Flyvbjerg, 2012). Therefore, *phronesis* strongly values the 'power of example'. That's an additional reason why this research focuses on empiricism through examining the experiences and perceptions of involved stakeholders, as will be described in the following paragraph.

43 Social constructionism looks at the ways social phenomena are created, institutionalized, known, and made into tradition by humans.

3.3 Research design

This thesis makes use of a research design that allows for nuances to exist, which enables analyzing the multi-causality and interdependencies between the forces at hand. In Figure 1 (Chapter 1) a schematic representation of this research design is given. It begins with the description and definition of the central problem (Chapters 1 and 2) and the introduction of the analytical frame which supports the categorization of drivers and barriers influencing this central problem (Chapters 4 and 5). The thesis then zooms in on two concrete examples (case studies) to examine the specificities of 3R acceptance in the area of vaccines (pharmaceuticals) (Chapter 6) and in the area of (industrial) chemicals (Chapter 7). Within these case studies 'reality' (i.e. the underlying social constructions) regarding 3R non-acceptance was reconstructed through a combination of research methods (see section 3.4.2). Subsequently, the thesis zooms out again to put the findings into the broader perspective of the two product sectors of pharmaceuticals -including vaccines - and the chemicals (Chapter 8). For this purpose two expert panels were organized (see section 3.5). Through this step the broader usability of the findings was tested and optimizing options were collected. To conclude, triangulation of the research methods took place in the light of the earlier found theoretical insights. This has led to a final overview of influencing factors and the way they interact (Chapter 9) and a series of possibilities to enhance the current process (Chapter 10).

3.4 Case study approach

To be able to collect detailed information about a certain phenomenon while taking its destined context into account (Swanborn, 2003), a case study approach is used. Through these case studies specific examples of 3R (non) acceptance are reconstructed. It provides a circumstantial way of looking and offers the possibility to thoroughly examine the process dynamics of a certain case. In addition, case studies enable the co-existence of different perspectives. *"A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident."* (Yin, 2003, p.13).

Blatter and Haverland define case study research as "*a non-experimental research approach that differs from large N-studies in the following four characteristics:*

1. *A small number of cases;*
2. *A large number of empirical observations per case;*
3. *A huge diversity of empirical observations for each case; and*
4. *An intensive reflection on the relationship between concrete empirical observations and abstract theoretical concepts" (Blatter and Haverland, 2012, p. 19).*

In-depth case study research is necessary to understand a complex social issue. *"The closeness of the case study to real-life situations and its multiple wealth of details are important ….for the development of a nuanced view of reality."* (Flyvbjerg, 2006, p.6). *"From both an understanding-oriented and an action-oriented perspective, it is often more important to*

clarify the deeper causes behind a given problem and its consequences than to describe the symptoms of the problem and how frequently they occur" (Flyvbjerg, 2001, p.78). For this reason, this research predominantly uses constructivistic qualitative research methods with case study research as the central research approach.

3.4.1 Causal process tracing

Regulatory acceptance and use is a process which is influenced by a broad variety of drivers and barriers (Schiffelers et al., 2007). To trace the causal processes between the dependent variable (regulatory acceptance and use of 3R models) and various independent variables (drivers and barriers), causal process-tracing (CPT) is used (George and Bennett, 2005; Blatter and Haverland, 2012). CPT aims at revealing how a certain situation (Y) is caused. This leads to research questions which are mostly framed as a *How* questions with the goal to appoint by what steps X led to Y. *"Thus, causal mechanisms provide more detailed and in a sense more fundamental explanations than general laws do. The difference between a law and a mechanism is that between a static correlation ("if X, then Y") and a "process" ("X leads to Y through steps A, B and C")* (George and Bennett, 2005, p.141). CPT enables tracking down information about specific events and steps within a certain process through the analysis of available documents and interviewing the central actors within this process (Tansey, 2007). The aim is to observe how events unfold over time and to discover the story behind it (Blatter and Blume, 2008). CPT has the following basic characteristics (Blatter and Haverland, 2012, p. 81):

1. It applies configurational thinking, which entails the assumption that the plurality of causal factors work together to create an outcome.
2. It takes the term 'process' seriously by adopting methodological and theoretical concepts that embrace the fact that causality plays out in time and space. This requires:
    a. The determination of the temporal order by which the causal process unfolds (comprehensive storylines);
    b. Empirical observations that provide a certain level of certainty regarding the Pathway leading from cause to effect (smoking guns) and;
    c. Empirical information that allows us to specify the underlying action-formation mechanism[44] that link causes and effects (confessions).

Causal mechanisms are highly dependent on the social context in which they are situated. This means that CPT does not strive for conclusive generalization but for 'possibilistic' generalization, which means that the outcomes lead to knowledge about the causal configurations (combinations of social mechanisms) that enable a specific outcome (Blatter and Haverland, 2012). Case studies offer this possibility to examine the operation of causal mechanisms in individual cases in detail (George and Bennet, 2005).

44 The action-formation mechanism refers to general assumptions about the behaviour of individuals which can be found in different theories (Blatter and Haverland, 2012) such as the Game Theory based on the basic assumption that individuals act rationally.

The steps of causal process tracing must be supported by theoretical models of causal mechanisms (George and Bennet, 2005). This model of causal mechanisms for the elucidation of the process of regulatory acceptance and use of 3R models is offered by the multilevel perspective of technology transitions as will be further clarified in Chapters 4 and 5 of this thesis. The theoretical framework of technology transitions (TT) enables analyzing transition processes which evolve in time and are the result of the interaction of drivers and barriers at the micro-, meso- and macro level (Kemp, 1994, Schot et al., 1994, Rip and Kemp, 1998, Geels, 2002). Explanations for TT cannot be found in forthright relations between variables. They have to be examined for conditions that may predict certain comparable outcomes in similar situations.

3.4.2 Case selection

Phronesis raises the value of the 'power of example'. For this, it is important to carefully choose examples through strategic sampling. A way of strategic sampling is looking for a *critical case* having strategic importance in relation to the general problem (Flyvbjerg, 2006). To find critical cases Flyvbjerg advises to look for either 'most likely' or 'least likely' cases, i.e. cases which are likely to either clearly confirm or irrefutably falsify propositions. In the context of this thesis the choice has been made to look for two specific examples of acceptance of 3R models which are viewed upon as critical cases by stakeholders in the field. Critical in that respect that they are frequently referred to as well known examples of the challenging process 3R models often face in the field of testing for regulatory purposes and as such 'most likely cases'. For this the case studies in this thesis had to meet the following criteria:

- The existing regulatory test is an animal model which is under discussion;
- There is a model available to reduce, replace or refine the existing animal model (3R model);
- And this 3R model is in the process of becoming regulatorily accepted/used.

The SNT case (Serum Neutralization Test for potency testing purposes of inactivated veterinary rabies vaccines) and the EOGRTS case (Extended One-Generation Reproductive Toxicity Study for reproductive toxicity testing purposes of chemical substances) meet all three criteria.
More specifically, both case studies entail a series of features that reflect the relevance of the transition towards available 3R models.
The SNT case is a relevant exemplar for the following reasons:

- The conventional animal model - the NIH challenge test - is highly disputed both scientifically and with regard to animal welfare, but is nonetheless still valued as the standard protocol. It involves high animal numbers and high levels of pain and distress.
- Rabies vaccines belong to the so called 'established vaccines' which can vary in composition. Therefore every batch has to be subjected to safety and potency tests. This makes established vaccines responsible for the bulk of the animal use in vaccine batch release testing.
- Validated 3R models are already available for many years.
- There are many regulatory levels involved e.g the European Directorate for the Quality of Medicines (EDQM) and the European member states Official Medicines

Control Labs (OMCL's), but also the US Food and Drug Administration (FDA) and the US Department of Agriculture (USDA), which offered the opportunity to examine diverse perspectives of different authorities.

The EOGRTS case is a relevant exemplar because:

- The conventional animal model - the two-generation test - is disputed both scientifically and with regard to animal welfare. It involves high animal numbers. The standard protocol is estimated to use nearly 40% of the laboratory animals under REACH (Janer et al., 2007a) and is thereby one of the major users of rodents in safety test programs.
- A validated 3R model - the EOGRTS - is available that can act as a reduction model to the two-generation study leading to a reduction of 40% in terms of animal use per test while at the same time being more informative.
- The EOGRTS has been formally adopted in OECD guidelines in July 2011 (OECD, 2011a) which is considered to be a big success in terms of the formal acceptance of this model for regulatory purposes. However, the study faced additional problems at the stage of actual acceptance and use by regulatory authorities and industries for regulatory purposes.
- There are many regulatory levels involved e.g. the Organisation for Economic Cooperation and Development (OECD)-, the US Environmental Protection Agency (EPA), the European Chemicals Agency (ECHA), the European Commission (EC) and European member states, which offered the opportunity to examine diverse perspectives of different authorities.

Additionally, it is important that the cases are 'accessible' to identify the needed information in order to make convincing causal claims. The fact that the SNT and the EOGRTS were, at the time of investigation, already quite far in the process of becoming accepted/used for regulatory purposes, offered the possibility to depict the full process and the different causes (independent variables) influencing the process of regulatory acceptance and use (dependent variable)

3.4.3 Operationalization

The transition from theory to empirical research is called 'operationalization', for which three steps are important (Van Thiel, 2014). Defining the central concepts of the research is the first step, which is needed to delineate what exactly will be studied. The second step consists of defining the different ways in which the central concepts can be expressed in the real world, also referred to as the 'variables'. The third and final step is to decide, for each variable, which influence it asserts to the original construct. This paragraph targets the further description of the central concepts and the initial definition of the variables, and their potential influence on regulatory acceptance and use of 3R models. In following a phronetic approach, the following three value-rational questions were core:

1. *Where are we going?*
2. *Is this development desirable?*
3. *What, if anything, should we do about it?*

In the context of the case studies, these questions are operationalized as follows:

1. *How did the process of regulatory acceptance and use of the SNT/ EOGRTS model unfold over time?*
2. *Which drivers and barriers at the micro-, meso- and macro level of the multilevel perspective of TT have influenced this process at the stages of FI, ARA and UI?*
3. *Is the result of this process desirable in terms of the European ambition to use 3R models wherever possible (see Chapter 1)?*
4. *Which options are available to steer the process in the direction of the European ambition?*

For operationalization purposes the following concepts require further specification: Drivers and barriers refer to the factors influencing the central process of regulatory acceptance and use of 3R models for regulatory purposes. Drivers are those factors that are observed to stimulate this process. Barriers are the factors which are observed to withhold this process. Drivers and barriers in the context of this thesis refer to the actual influences as perceived by the stakeholders involved and not the hypothetical ones. Hypothetical drivers are defined as options to enhance the process. They have the potential to stimulate the process but are not (yet) observed to actually do so.

Drivers and barriers that were retrieved in anticipation of the interviews through earlier research (see Appendix I), through document analysis and orientating interviews in the EOGRTs case and through document analysis, meetings and a survey in the SNT case (see section 3.4.4) were classified after the macro- meso- and micro level of the multilevel perspective on TT. These drivers and barriers served as so-called 'sensitizing concepts' (Bowen, 2006). Sensitizing concepts are interpretive devices used as a starting point for a qualitative study. They draw attention to important features of social interaction and provide guidelines for research in specific settings and provide starting points for building analysis (Charmaz, 2003: as cited in Bowen, 2006) (see Appendix VI for an overview of the sensitizing concepts). However, these concepts were not used to actively steer the interviews. During the interviews the researcher stayed as close as possible to the concepts as proposed by the respondents in order to avoid narrowing down the perspective at an early stage.

The micro-, meso- and macro level of the multilevel perspective of TT are defined in Chapter 4 of this thesis. For the operationalization it is important to mention that factors are labelled as:

- micro elements when connected to the niche in which innovations are developed and put to the test, (e.g. -dis-advantages of the new technique, the competences of the innovators and the strengths and weaknesses of the innovation process);
- meso elements when being part of the existing sociotechnical regime (e.g. the standard operating procedures, rules and regulations regulating the product assessments and the connected test methods);
- and macro elements when related to the surrounding sociotechnical landscape (e.g. cultural, economic, demographical and geographical developments).

It must be noted that it is sometimes difficult to draw a clear line between these levels. This has led to the re-categorization of some of these elements in the course of this research.

Furthermore, the process of regulatory acceptance and use of 3R models is divided into three substages FI, ARA and UI which are defined in Chapter 2. However, as with the levels micro-, meso- and macro, the phases cannot always be sharply distinguished. The substages for example, cannot always be placed in a chronological order and the substages may well take place in parallel e.g.: FI may still be in the making while ARA and/or UI are already partly occurring.

3.4.4 Data collection and data analysis

In the inductive CPT approach, the separation between data generation and data analysis is difficult to distinguish (Blatter and Haverland, 2012). "*Inductive analysis means that the patterns, themes, and categories of analysis come from the data; they emerge out of the data rather than being imposed on them prior to data collection and analysis"* (Patton, 1980, p. 306). In this iterative (cyclic) process, data collection and data analysis are alternated until a saturation point is reached. Through this process the sensitizing concepts are specified during the course of the research based on the empirical findings collected.

The case study's strength is its ability to deal with a variety of evidence (Yin, 2008). Different methods, i.e. literature research, document analysis, (semi)structured interviews and the attendance of meetings - in the SNT case -, were combined in order to get to a comprehensive representation of the situation examined (Yin, 2003). Research methods are elucidated below in order to clarify their specific contribution to the case study. The precise aspects connected to the case, such as information regarding the respondents are described in methodological sections of Chapter 6 and Chapter 7.

Document analysis: The collection of findings started with document analyses to provide an overview of the regulatory framework, the stakeholders involved, existing testing practices and variables - drivers and barriers - influencing the regulatory acceptance and use of respectively the SNT and the EOGRTS. The examined sources consisted of scientific publications, meeting reports, (institutional) websites and press releases and correspondence between stakeholders.

Interviews: A series of about twenty in-depth interviews was conducted per case study to collect the perspectives of stakeholders on the process of acceptance and use of respectively the SNT and the EOGRTS. The interviews were semi-structured, asking open-ended questions designed to reconstruct the process and identify the drivers and barriers per subsequent substage, i.e. FI, ARA and UI (see 3.4.3 for an operationalization of these central research questions).

The selection of respondents was done through a combination of criterion and snowball sampling (Patton, 2001). Through criterion sampling a small group of respondents was chosen using the selection criterion of being a (scientific, legal and/or political) experts with experience in/or knowledge of/involvement in the specific case study. This first

sampling was done through document analyses searching for involved stakeholders in the field. Involvement means having been able to closely follow or take part in – parts of – the process of acceptance and use of that specific 3R model. Next, the population was broadened though snowball sampling asking each respondent for other suitable candidates. Suitability was defined as direct or indirect involvement in one or more of the substages of FI, ARA or UI. This might have led to a certain level of bias in the sample, since the people directly or indirectly involved might be more positive about the 3R model under discussion and the strategy that was followed. To diminish the risk of such a bias, each respondent was explicitly asked for potential candidates with different perspectives on the case, both in terms of stakeholder groups as in terms of opinion regarding the 3R model. In the SNT case study, the researcher was able to observe the discussions that took place during international workshops and select respondents that represented different points of view. In the EOGRTS case study the researcher did not have this opportunity and relied on the information retrieved from document analyses and other respondents. The respondents came from stakeholder groups among European legislators, European and national regulatory authorities, industry/contract research organizations, academia, animal welfare organizations and consultants in this specific area.

The interviews began with the introduction of the respondent, his or her involvement in the process examined in the context of this case study, the position of his/her organization regarding the 3R model, a brief process reconstruction and a more specific chronology of the involvement of the respondent in this process. Next, a series of questions was asked about barriers and drivers per substage (FI, ARA and UI). This started with an open question of which barriers/drivers have influenced the process of FI/ARA/UI of this 3R model. Depending on the involvement of the respondent one or more of the substages were examined. Subsequently, the sensitizing concepts were checked. Lastly, interviewees were asked to give their opinion on optimizing options to enhance the process of regulatory acceptance and use of 3R models. The questions were the same for every respondent, but the focus differed depending on the respondent's involvement in the process.

For most of the interviews, an audio recording and a transcription was made. In those cases where interviews were not recorded the interviews were transcribed and made available to the respondent for validation of the findings. Next, the transcripts where analyzed to make an inventory of drivers and barriers per substage and of the optimization possibilities. To analyze the empirical findings the field notes of the meetings and the transcripts of the interviews were coded. To begin with, the findings were categorized after stakeholder group and in sets of overarching themes (e.g. process reconstruction using the three substages FI, ARA and UI, drivers, barriers and recommendations). TT literature is used to reflect upon the case study findings, also known as 'pattern matching' (Yin, 2008). The events were placed in a chronological order and the drivers and barriers were classified into the levels of the 3R acceptance model (see Chapter 5) - i.e. the niche or micro level, the sociotechnical regime or meso level and the sociotechnical landscape or macro level -.

Meetings: Between 2010 and 2012 six international meetings on 3R models for - rabies - vaccine testing were attended (see Chapter 6 and Appendix II). The official reports of these meetings were examined for factors that potentially drive or withhold regulatory acceptance and use of 3R models for – rabies – vaccine potency testing purposes. The meetings also offered the researcher the possibility to be an *'observer as participant'*. "*In this role, the researcher or observer has only minimal involvement in the social setting being studied. There is some connection to the setting but the observer is not naturally and normally part of the social setting"* (Gold, 1958).[45] This participatory observation was used as a tool to collect data about the stakeholders and their interactions. Field notes were taken to register the observations, conversations and features of the context in which this took place. No such meetings have been attended for the EOGRTS case study since most of these meetings had already taken place at the stage that the case study was conducted. Furthermore, the fora at which the decisions took place in the context of the EOGRTS case study (OECD, ECHA) were not open to external observers. Instead the meeting reports were examined where available and accessible.

Survey: In addition, the rabies case study which is described in Chapter 6 was anticipated by a survey on drivers and barriers on 3R acceptance for – rabies - vaccine potency testing purposes. The results of this survey, which were published in 2014 in the journal Biologicals (Schiffelers et al., 2014: see Appendix II), offered input on the drivers and barriers in this field. (For a detailed description of the methodology used for this surveys see Appendix II). Due to time restrictions and difficulties in finding sufficient respondents, no such survey has been conducted in anticipation of the EOGRTS case study. To deal with this lacuna available documents were examined and several orientating interviews (N=3) where organized to get an idea of the potential drivers and barriers and of the specific context of this case study.

3.5 Expert panels

In spring 2012 two expert panels were organized to make an inventory of the existing perceptions on drivers and barriers influencing 3R acceptance in the sectors of pharmaceuticals and chemicals. The goal was to check the validity of the earlier retrieved empirical and theoretical factors. Moreover, a series of optimization options was discussed. Both the pharmaceuticals and the chemicals panel included a total of 20 experts in the fields of safety assessment, regulatory testing and 3R models. The participants derived from the following stakeholder groups:

- Regulatory authorities, legislators & policy makers
- Industry
- Academia & research organizations

Both panels aimed at clarifying the process of regulatory acceptance and use of 3R models and at the examination of possibilities to enhance this process. These goals were targeted through the following three subsequent steps: firstly, an inventory of barriers and drivers

45 http://www.qualres.org/HomeGold-3648.html

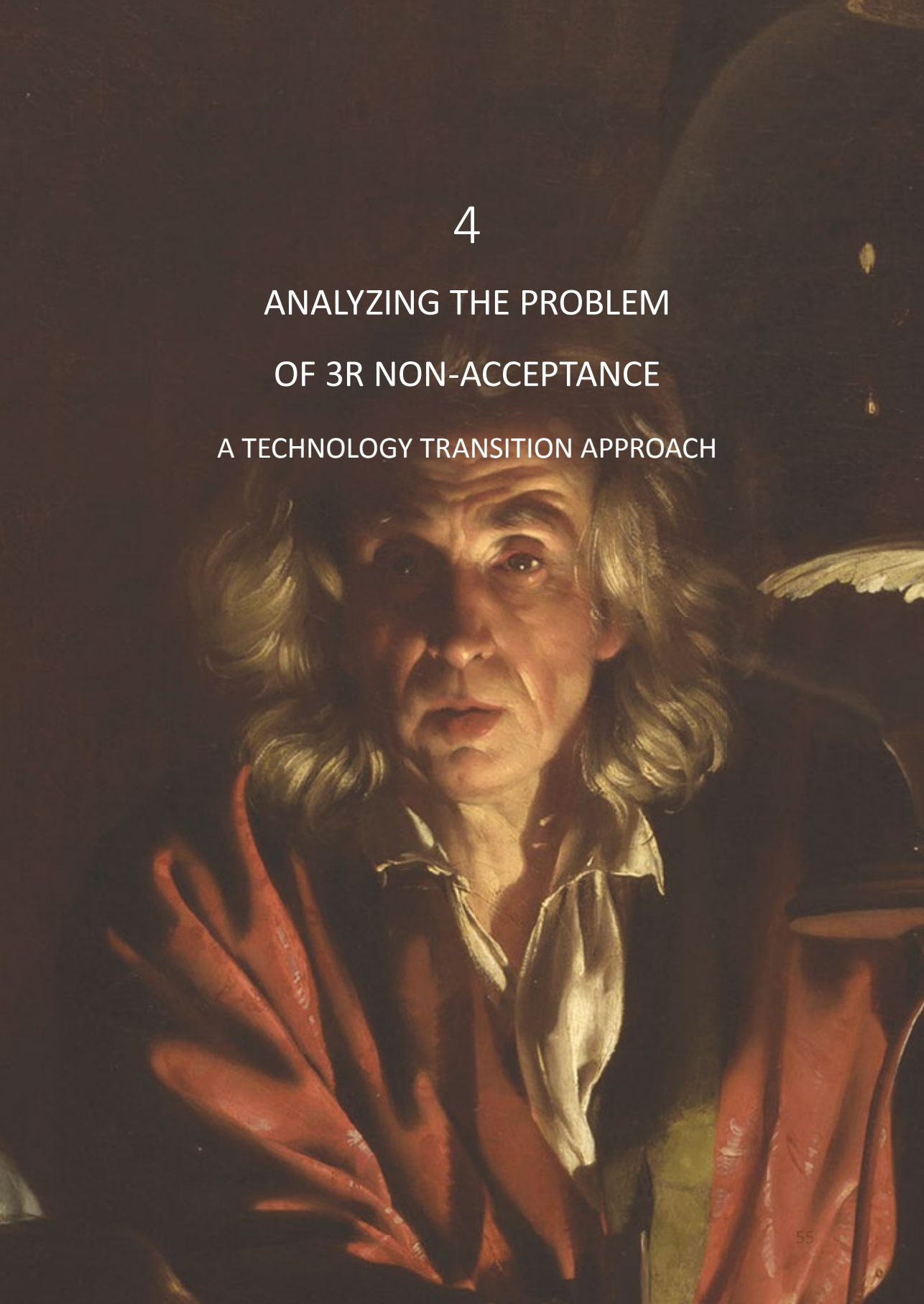
was made; secondly, a prioritization was made of the factors in terms of their influence on regulatory acceptance and use; thirdly, actions were identified that can be pursued by the stakeholder groups in order to optimize the process of regulatory acceptance. For a full description of the methodology used for the expert panels, see Chapter 8.

3.6 Reliability and validity of the collected data

The term 'reliability' is a concept used for testing or evaluating quality in quantitative research. However, it is often also used in reference to qualitative research. Patton (2001) for example states that validity and reliability are two factors that any qualitative researcher should be concerned about while designing a study, analyzing results and judging the quality of the study (Golafshani, 2003, p. 601).

"Constructivism values multiple realities that people have in their minds.." (Golafshani, 2003 p. 604) and as a result many case studies have been based on stakeholders' perceptions. To minimize the risk of arbitrary subjectivism several control mechanisms have been adopted. To start with, discrepancies have been reduced by organizing a series of interviews with respondents from different stakeholder groups and with different perspectives and opinions regarding the central topic. Subsequently, the interview reports were analyzed and compared. During the interviews the researcher repeatedly summarized what was said by the respondent and the interview-, and expert panel reports were checked by the respondents /experts.

Furthermore, different sources of data have been used and the findings of these sources were compared. This control mechanism, also referred to as triangulation, is a well-known strategy for improving the validity and reliability of research or evaluation of findings. Mathison (1988) elaborates on this by saying: *"Triangulation has risen an important methodological issue in naturalistic and qualitative approaches to evaluation [in order to] control bias and establishing valid propositions because traditional scientific techniques are incompatible with this alternate epistemology."*(Mathison, 1988: as cited in Golafshani, 2003, p.603). Therefore, to acquire valid and reliable multiple and diverse realities, multiple methods of searching and gathering data were used. Through the adoption of multiple methods, such as, document analysis, observations, interviews, a survey, interview recordings and expert groups more valid, reliable and diverse constructions of realities were uncovered. To further improve the analysis and understanding of the findings, the interpretation of the data was checked through three additional steps. Firstly, in every empirical step taken, several co-researchers/supervisors (co-authors of the manuscripts) were involved in the gathering and analysis of the data. Secondly, the final concept of the manuscript was pre-read by an expert in the field of the specific case study before it was submitted to a peer-reviewed journal. And thirdly the findings were checked through the process of the peer review.



4

ANALYZING THE PROBLEM OF 3R NON-ACCEPTANCE

A TECHNOLOGY TRANSITION APPROACH

"Technology transition is difficult, in part, because we have underestimated just how much effort is required for such transfer to occur effectively" [46]

Everett Rogers
Communication Scholar and Sociologist
1931-2004

46 E.M. Rogers (2003) Diffusion of innovation, p.152.

4.1 Introduction

This thesis aims at creating a better understanding of the challenging process of 3R non-acceptance and at providing the involved stakeholders with valuable insights to overcome the existing delays 3R models face in the regulatory domain. To accomplish this goal the following chapter targets research question Q2: *Which theoretical perspectives are needed to comprehend the process of regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes and find suitable ways of enhancing the process?* This means that theory in the context of this research does not serve falsification purposes but is utilized to get a better grip on the complexity of the central research issue and to find ways to overcome the existing barriers.

The initial research prior to this thesis (Schiffelers et al. 2007: see Appendix I) was valued for the useful overview it offered on factors influencing regulatory acceptance and use of 3R models. However, from this research it was also concluded that regulatory acceptance and use is highly contextual and complex. The issue knows many different perceptions and controversies, e.g. on the pros and cons of the existing and the alternative models, on the main influencing factors and on the way forward. To understand and deal with this wicked problem (see Chapter 1) a broad view is needed which enables the combination of perspectives of involved organizations and stakeholders. *"By their nature, wicked issues are imperfectly understood, and so initial planning boundaries that are drawn too narrowly may lead to a neglect of what is important in handling the wicked issues." …"What is needed is thinking capable of grasping the big picture, including the interrelationships between the full range of causal factors and policy objectives." (*APS. 2007, p. 11).

From a first examination of innovation-, implementation- and risk regulation literature, it became clear that many of the mechanisms that were observed in the initial research (Appendix I) are covered by the literature on Technology Transitions (TT) / System Innovations (SI). In Chapter 5 the specificities of a risk-averse context are taken into account and combined with the multilevel perspective on technology transitions. Risk aversion was observed to be a dominant contextual feature in terms of regulatory acceptance and use of 3R models in the research prior to this thesis (see Appendix I). This specification is important to comprehend the impact on innovations of the striving for risk minimization.

Combining theoretical concepts from different disciplines to create the necessary integrative /multidisciplinary approach is an illustration of 'appreciative theory' (Nelson and Winter, 1982). Appreciative theory *"aims to capture the basics of what actually is going on."* (Nelson, 2006). *"In its role of providing a framework for appreciation, a theory is a tool of inquiry, and in skilful applied research that tool is used flexibly, bent to fit the problem, and complemented by any other tools that happen to be available and that appear to be useful."* (Nelson and Winter, 1982, p. 46).

This chapter focuses on TT and SI literature which provides the multidisciplinary and integrative approach that is needed to cover the contextuality and complexity of 3R acceptance and use for regulatory purposes. This literature is examined to:

- Define what is meant by technology transition and system innovation and clarify why it is a suitable approach for the analysis of the process of regulatory acceptance and use of 3R models (section 4.2);
- Explore the theoretical factors barriers and drivers in terms of stimulating or obstructing such system innovation/technology transition processes (section 4.3);
- Identify intervention strategies to enhance technology innovation/technology acceptance processes (section 4.4).

In the following chapters the link between these theoretical notions and the empirical findings will be made.

4.2 The Technology Transitions (TT) and System Innovations (SI) perspectives

This section elucidates the TT and SI perspectives. The TT perspective (4.2.1) offers a multilevel systemic overview of factors influencing the transitions process and is thereby the broadest view. The SI perspective (4.2.2) provides insights in institutional activities and interactions that can initiate innovation. In addition the Technology Innovation System (TIS) (4.2.3) is described which combines the TT and SI perspective and provides information on encompassing niches through conjoint activities.

4.2.1 The TT perspective: an integrative multilevel approach

The theoretical approach that has been primarily used to analyse the phenomenon of the acceptance and use of 3R models, is the perspective of Technology Transitions. *"Technological Transitions (TT) are defined as major technological transformations in the way societal functions such as transportation, communication, housing, feeding, are fulfilled."* (Geels, 2002, p.1257). This theoretical perspective entails a holistic approach on the conversion from one technological system to another. "*TT do not only involve technological changes, but also changes in elements such as user practices, regulation, industrial networks, infrastructure, and symbolic meaning"* (Geels, 2002, p.1257). The acceptance and use of 3R models in the regulatory domain can be viewed upon as a TT, since this group of technologies has the potential to contribute to far reaching changes in the way products and substances are produced and assessed. This would also lead to major changes in the whole system and infrastructure (e.g. required knowledge, training and education, testing facilities) that is in place to enable regulatory testing.[47]

The perspective of TT stems from the sociology of technology. In this perspective, technology in itself has no power. TT literature therefore uses a multilevel framework to scrutinize the transformation from one technological system to another. This framework, which has been described amongst others by Kemp (1994), Schot et al. (1994), Rip and

47 It must be noticed that although the term 3Rs refers to a diverse group of technologies, they have a clear mutual goal, i.e. offering alternatives to the existing test regimes which are still largely based on animal models.

Kemp (1998), Geels and Kemp (2000), covers the different aspects that a new technology has to cope/interact with. It consists of the following three levels:

1. The micro level or niches
   Niches function as incubation rooms in which radical innovations emerge and mature. *"Niches are crucial for TT, since they provide the seeds for change"* (Geels, 2002, p.1261). Niches offer protection from the severe selection process of the regime and provide facilities such as locations for learning processes and possibilities to build social networks supporting innovations. Within the niches there is a dominant design and actors improvise to work out the best design and find out what users want (Kemp, 2010).
2. The meso level or sociotechnical regime
   "*The key concept of the (TT) framework is the sociotechnical regime, a coherent, highly interrelated and stable structure at the meso-level characterized by established products and technologies, stocks of knowledge, user practices, expectations, norms, regulations, etc.* (Markard and Truffer, 2008, p.603). The sociotechnical (ST) regime, which represents this patchwork of rules, knowledge, standard procedures and existing infrastructure, is built upon Nelson and Winter's (1982) concept of 'technological regimes' Within these regimes organizations and actors have developed firm organizational and cognitive routines which result in routine-based behaviour and technological trajectories. As a result, ST regimes are characterized by stability and function as selection and retention mechanisms (Geels, 2002). "*From the evolutionary perspective, a regime represents the selection environment for technological development in a certain field or sector, thus exerting a significant barrier for radical innovations to diffuse."* (Markard and Truffer, 2008, p.603). However, this is stability of a dynamic kind in which innovation still occurs but in a more incremental way.
3. The macro level or sociotechnical landscape
   The sociotechnical landscape is the external structure or context in which innovations are embedded. It consists of a set of slow changing broad societal factors such as economy, demography, broad political coalitions, cultural and normative values and environmental aspects. These societal factors influence innovation or transition processes. The metaphor 'landscape' is chosen because of the association with the relative 'hardness' and the material context of society. "*The context of landscape is even harder to change than that of regimes. Landscapes do change, but more slowly than regimes."* (Geels, 2002, p.1260).

Figure 6 depicts the relation between the three levels which is characterized as a nested hierarchy: *"The nested character of these levels means that regimes are embedded within landscapes and niches within regimes"* (Geels, 2002, p.1261). *". . . technological niches and sociotechnical regimes are similar kinds of structures, although different in size and stability. Both niches and regimes have the character of organizational fields (community of interacting groups). For regimes, these communities are large and stable, while for niches they are small and unstable. Both niche and regime communities share certain rules that coordinate action. For regimes, these rules are stable and well-articulated; for niches, they are unstable and 'in the making'."* (Geels and Schot, 2007. p.7).

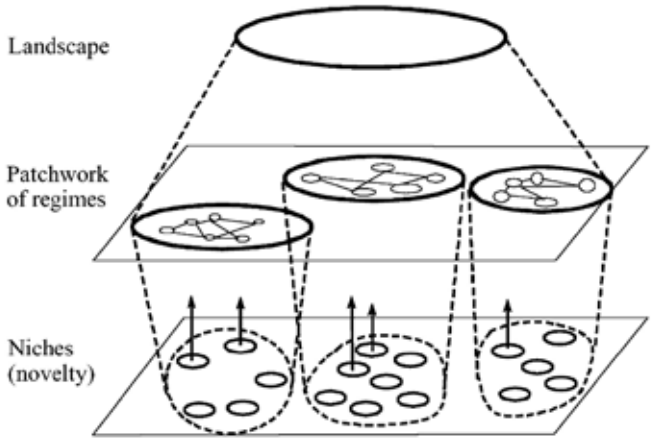


Figure 6. Multiple levels as a nested hierarchy (Source: Geels, 2002).

In the context of this thesis these levels would for example contain the following elements:

- The niches are formed by those stakeholders/organizations/platforms which have started to work on a 3R model. Within these niches the 3R models are developed, tested in small scale studies and later on in interlaboratory studies with the goal to validate the test for broader use.
- The regimes consist of the existing regulatory requirements, the standard testing procedures, the connected testing infrastructure, the existing knowledge and experience which is based on the standard operating procedures etc. in which animal models are often still the leading paradigm.
- The landscape is the broader context in which regulatory testing has to be inserted. In the context of this thesis the landscape is formed by the broad developments in society, with a focus on Europe in terms of cultural values related to animal testing and product safety. However, also broader economic, political and geographical developments are observed to influence the central research issue at this macro level. A further specification of the influences at the micro-, meso- and macro level in the area of regulatory acceptance and use of 3R models is given in Chapter 5.

4.2.1.1 Alignment of the three levels

Rip and Kemp distinguish three phases for Technology Transitions to occur (Rip and Kemp, 1996; Kemp 2010) (see Chapter 5 Figure 9 for a depiction of these phases). In the first phase, radical innovations emerge in niches, often outside or on the fringe of the existing regime. The supporting networks are small and unstable and the innovation is no match yet for sociotechnical (ST) regime. This is the phase in which a 3R model is for example used by a manufacturer for in-house purposes such as R&D, however its broader use for example for regulatory purposes is still far from achieved. In the second phase, the

innovation is used in small market niches, developing a technical trajectory of its own. A dominant design is slowly becoming apparent. Nonetheless, the technology is of no major threat yet to the regime. In the area of the 3Rs this for example would be the case for validated methods which have been tested extensively and in an interlaboratory setting. The test has proven its usability in a broader setting, however its FI and ARA is not yet accomplished. The third phase is that of a *"wider breakthrough"* and of competition with the existing ST regime which leads to new types of structuring (Kemp, 2010). In this phase the 3R models have reached the level of FI and maybe also already ARA and/or UI. The 3R models which are described in the case studies can be placed in this stage. Both 3R models start to compete with the existing ST regime. However, they are still at a stage in which they exist in parallel to the conventional testing methods. Both models were not yet able to fully overthrow the existing regime.

Rip and Kemp have taken the innovation as the starting point. Geels labels the usage of this starting point as a bias which needs to be balanced by a more profound focus on the regime and landscape level resulting in Figure 7 below (Geels, 2002). Geels' figure distinguishes seven dimensions in the ST regime: technology, user practices and application domains (markets), symbolic meaning of technology, infrastructure, industry structure, policy and techno-scientific knowledge. "*The regular ongoing incremental processes are represented with relatively long arrows. Although the different dimensions are linked and co-evolve, they also have internal dynamics. This may result in 'tensions', represented in Fig. 7 with shorter diverging arrows, indicating uncertainty and differences of opinion. Tensions may lead to periods in which linkages are weakening* (Geels, 2002, p.1262.) These tensions between the seven dimensions create drivers or barriers for new technologies which can stimulate or obstruct further progress. Whether or not progress is made depends on the outcome of the interplay between the drivers and barriers.

In the area of regulatory testing several tensions can be observed. To start with, there are many new technologies which have become available to test the safety and efficacy testing of products. At the same time, the symbolic meaning of animal models is shifting. Animal models are increasingly being questioned for a combination of ethical and scientific reasons (see Chapter 2). Furthermore, European policy (Directive 2010/63/EU on the protection of animals use for scientific purposes) stimulates the use of 3R models and an increasing number of 3R models are already being used by industry for R&D and production purposes. As mentioned, it then depends on the balancing of the drivers and barriers whether a 3R model will be able to enter the regulatory domain and compete with the existing ST regime. The way in which this balancing act takes place in practice is described in the case studies in Chapters 6 and 7.

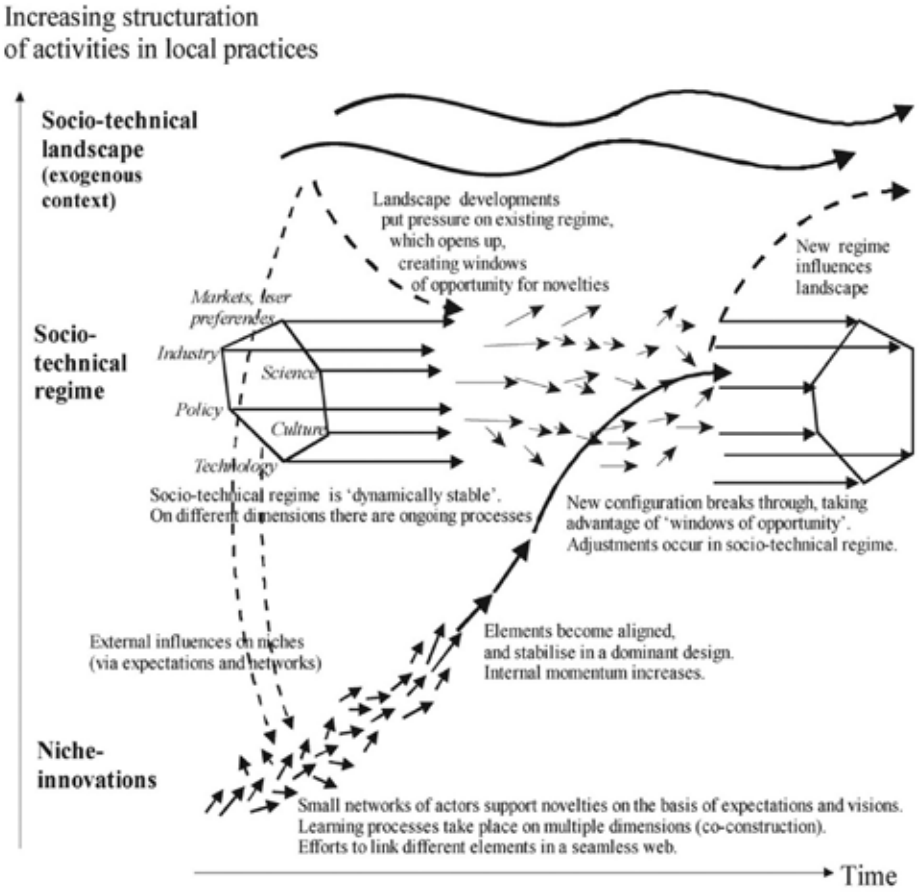


Figure 7. Multi-level perspective on transitions
(Source:Geels and Schot, 2007: adapted from Geels, 2002, p. 1263).

The different levels of the multilevel framework are very useful in understanding the complex dynamics of sociotechnical change. Due to the fact that the elements in a ST regime are inter-connected, radically new technologies often are incompatible with the existing regime and as a result have difficulties in breaking through. However, changes at one or more of the three levels can lead to alterations in the status quo and thereby create options for new technologies to break out of their niche. *"Sociotechnical change is described as a process of shifting assemblies of associations and substitutions, a reweaving of elements. Changes in one element in the network can trigger changes in other elements."* (Geels, 2002, p.1259). Technological transition in this perspective is the result of the alignment between a heterogeneous set of elements at these three different levels (Geels, 2002) which is referred to by Rip and Kemp (1998) as *"configurations that work"*. "*Regimes may face landscape pressure from social groups objecting to certain features (pollution,*

capacity problems and risks) and may be challenged by niche developments consisting of alternative technologies and product systems." (Kemp, 2010, p.293). *"It is the alignment of developments (successful processes within the niche reinforced by changes at regime level and at the level of the sociotechnical landscape) which determine if a regime shift will occur"* (Kemp et al., 2001, p.277).

4.2.2 The SI perspective: insights in institutional activities and interactions

Next to the broad TT perspective the SI perspective is often utilized by innovation scholars to analyze the institutional activities and interactions connected to system innovation processes. Freeman (1987) defines an Innovation System as: *"The network of institutions in the public and private sectors whose activities and interactions initiate, import, modify, and diffuse new technologies."* The overall role of the innovation system is described as the generation, diffusion and use of an innovation (Edquist, 2005). The central idea behind this approach is that factors connected to technological change are not only related to individual firms or research institutes, but certainly also to the broad societal structure in which firms and knowledge institutes, are rooted (Lundvall, 1988). Innovation and diffusion of technology is determined by individual actors and by the innovation systems in which they are embedded. It is both an individual and a collective act (Jacobsson and Bergek, 2004, Hekkert et al., 2007). And even though individual actors are rooted in an institutional context, they can also change or adapt existing institutions or create new ones (Markard and Truffer, 2008). The character of an innovation system for example can be changed over time by the entry of new actors, but also by other developments such as new laws, and events such as crises, change of political culture and demographical developments. Therefore, to understand technological change, one needs insight into innovation system dynamics (Hekkert et al., 2007).

4.2.2.1 Connecting sub-functions within the innovation system

Whether or not system developments will lead to technological change depends on how the sub-functions of the innovation system are connected. Various sub-functions are defined, such as entrepreneurial activities, the generation and diffusion of knowledge through networks, the guidance of search processes, the creation of markets, the mobilization of resources and the creation of legitimacy (Bergek et al., 2005; Hekkert et al., 2007) (see Markard and Truffer, 2008). The accomplishment of one of these sub-functions is likely to affect the fulfillment of other functions. Connections between these sub-functions may bring about processes of 'cumulative causation' (Jacobsson and Bergek, 2004) or virtuous cycles ('motors') or vicious cycles (e.g. Hekkert et al., 2007: see Markard and Truffer, 2008). In other words, the quality of the overall system function - i.e. its ability to generate, diffuse and use an innovation - depends on the quality and the interaction of the sub-functions (Markard and Truffer, 2008). The SI approach describes different fuction combinations depending on the technology and the context- that may serve as triggers. In the light of this thesis it is important to look for possible answers to the questions: which connections are needed and what can be done to stimulate such connections? In section 4.3 on drivers and barriers, we will elaborate on the functions which, in earlier empirical work on innovation system dynamics, were found to be likely starting points for such connections.

4.2.3 An integrated systemic approach: Technological Innovation System

The TT multilevel framework and the SI approach have different approaches in analyzing innovations, each with their own strengths and weaknesses (Markard and Truffer, 2008). Where the multilevel framework of the TT approach is very useful for a comprehensive analysis of a technology transition through the inclusion of all three levels and their interplay, the innovation system approach proves strong in analyzing the roles and strategies of the different actors and their interaction and offers concrete options to stimulate progress. This means that combining the strengths of both the TT and the IS approach is desirable.

Markard and Truffer explain that combining both approaches is not only desirable since the two approaches have complementary strengths which makes an integration of both approaches valuable, but also feasible because both approaches share many theoretical and conceptual notions. They introduce the concept of technological innovation systems (TIS) which combines both approaches: *"A technological innovation system is a set of networks of actors and institutions that jointly interact in a specific technological field and contribute to the generation, diffusion and utilization of variants of a new technology and/or a new product."* (Markard and Truffer, 2008).

They define a TIS in a way that is compatible with the multilevel concept by combining the existing three levels with a fourth conceptual element. The four levels then become:

1. The niches or application contexts, in which radical innovations emerge and mature;
2. The technological innovation system, which might encompass niches and is characterized by emergent institutions and conjointly produced resources;
3. The sociotechnical regimes that represent the dominant production structure, which challenges the TIS;
4. The landscape with parameters that influence regimes and innovations without being In turn influenced (Markard and Truffer, 2008, p.611-612).

A TIS can interact with one or more sociotechnical regimes (see Figure 8). It depends on the institutional networks and the interaction (mutual set of actors) of a technological innovation system (TIS) with a regime, what the level of opposition will be to transform it in the way proposed by the niche. The different niches/innovations may help each other in weakening dominant regimes. This is the case in situations of niche accumulation but may even be the case when the niches offer competing innovations. Even though they are competing they all have the same overall goal: offer alternatives to the existing regime (Markard and Truffer, 2008).

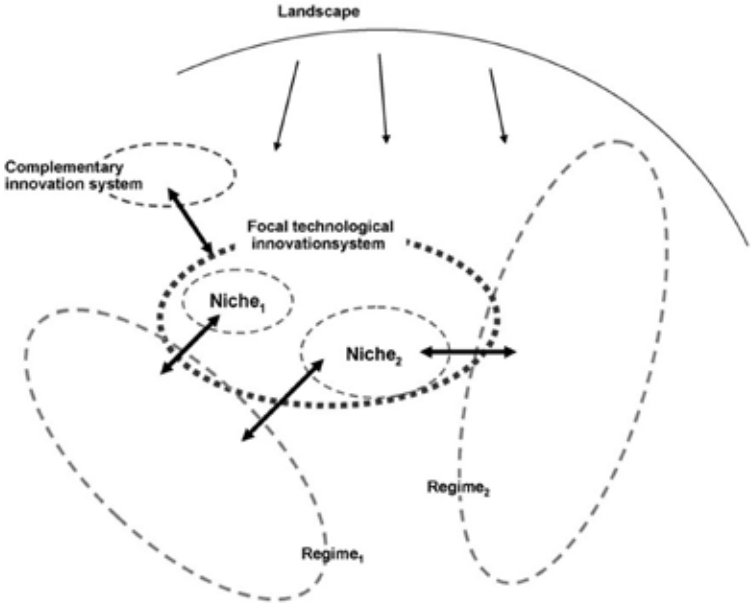


Figure 8. Technological innovation system and interactions with the conceptual elements of the multi-level framework. (Source: Markard and Truffer, 2008)

4.3 Barriers and drivers found in TT and SI literature

The challenges 3R models face in becoming accepted and used are typical for many new technologies. Asplund and Sandin (1999) and Cozijnsen et al. (2000) for example observed that out of every five projects initiated only one was actually realized (see Van der Panne et al., 2003). Even though this number refers to single product innovations initiated by firms, it is an indication of the challenging innovation process. For system innovation the challenge is likely to be even bigger due to the fact that they are subjected to far more variables and interdependencies.

4.3.1 Barriers

In this section, barriers that potentially work against the diffusion of alternative technologies are discussed. For this, the categorisation as distinguished by Kemp, Schot and Hoogma is adopted (Kemp et al., 1998).[48]

1. Technological factors

   This category of barriers refers to the misfit of alternative technologies with the

48 A large amount of publications focus on innovations in firms and consequently on the factors that specifically refer to the innovation process in individual firms (see Van der Panne et al., 2003 for a literature review on success and failure of innovation within firms). This thesis however focuses on the factors that are identified in TT and SI literature to influence system innovations/technology transitions/regime shifts.

existing regime and the requirement to fit it in the prevailing system. The new technology is often insufficiently developed in terms of user needs and end users require further optimization. Many of the arguments brought forward by stakeholders in this research are technological barriers referring to the limitations/drawbacks of the 3R models and the uncertainties on how to interpret the test results obtained through this model.

2. Governmental policy and regulatory framework
   Government policy may be a barrier in the sense that the signals it sends out about new technologies are often abigous. Some policy goals will stimulate a certain new technology while others are observed to refute with the ambition to stimulate the alternative. This is aggravated by the fact that different policy levels often pursue different policy goals and as such become contradictory (Hoogma, 2000).
   Stimulation of innovation through policy requires a coherent, consistent and congruent policy regime. However, achieving consistency, coherence and congruence is no easy task. Literature on political science and public policy discloses the limitations of 'rational/comprehensive' planning of policy (Lindblom, 1959) (see Meadowcroft, 2009 and Kern and Howlett, 2009). Reality reveals a chaotic and conditional character of policy making (Cohen et al., 1972). Government actions to alter socio technical systems will be seriously challenged and it has to be notified that *"…the movement from one dynamic equilibrium to another typically stretches over several generations, around 25 to 50 years*" (Meadowcroft, 2009, p.324).
   It is even seriously questioned whether the conscious shaping of such societal changes is even possible (Pressman and Wildavsky 1973; Majone and Wildavsky 1978). As long as there is no unambiguous policy to stimulate the alternative technology, manufacturers remain uncertain about the market developments and will be reluctant to invest in the risky alternatives. This is underlined by Wallace (1995) who states that unpredictable and inconsistent policies frustrate innovation by creating uncertainty for prospective innovators (Wallace 1995: as cited in Taylor, Rubin & Hounshell, 2005). This is in line with the observations in the field of regulatory testing. Here a conflict can be observed between horizontal and vertical legislation (see Chapter 2) leading to uncertainty in terms of the final decision with regards to the acceptance of the 3R model (see Wagner et al., 2012). Furthermore, existing regulations frequently conflict with the new technology. For example, in the of case environmentally benign vehicles as described by Kemp et al. 1998: strict safety requirements in Japan drive up the price of on-board gas tanks and refuelling stations which withholds cleaner gas vehicles from entering the market. *"Adaptations of legislation are often cumbersome, partly because some of the actors may oppose them"* (Kemp et al., 1998, p.178).
3. Cultural and psychological factors
   Manufacturers and technology users have an idea about what a certain technology is and what it should be able to do. This image is shaped by the dominant technology to which all kinds of values are attributed. The alternative technology often does not match with that image and the unfamiliarity with the new technology and the comparison with the dominant technology then lead to skepticism about the alternative (Hoogma, 2000).
   Mismatches are the effect of images that are incongruent with the dominant image

and of opposing interests of the regime and the niche. *"While regimes generate incremental innovations that strengthen the regime, niches create and protect radical innovations, which may lead to destabilization and far-reaching changes in established regimes." (*Markard and Truffer, 2008). The uncertainty which is connected to new technologies often results in resistance to innovation. In connection to the acceptance and use of 3R models, stakeholders often bring up the 'Not Invented Here' syndrome, which refers to the resistance which is said to exist against alternatives methods amongst those who were not involved in developing and/or validating the test protocol. Schumpeter (1942) states that there are always winners and losers of new technologies and the ones that are expected to lose by the introduction of the new technology are likely to oppose them.

4. Demand factors
   The usability of new technologies still needs to be demonstrated. This may well mean that they do not meet the demands of the consumers/end users. Introducing the technology would thus require a change of preferences of the end user or a change of the technology to meet user demands. Insecurities and aversions of the end users towards the new technology lead to cautiousness from the introducers of the new technology. *"The manufacturers think that consumer demands cannot be changed, and therefore they often refer to them as the most important barriers"*. The result is that manufacturers anticipate on what the potential reaction of the end-users will be and as a result prefer to "*avoid risks by building on current consumer* (or end-user) *preferences"* (Kemp et al., 1998, p.179).
5. Production factors
   By choosing the alternative technology, manufacturers face the risk that parts of their core technical production process and R&D activities, as well as organizational modes of control and marketing strategies/competences could become superfluous. Manufacturers are alleged to only take that risk if they then can use/produce the alternative technology for the wide market. However, new technologies are often poorly developed and expensive (Kemp et al., 1998) and thereby unfit for broad use/commercialization. In the context of this research, manufacturers often refer to the fact that they make products for the global market. If 3R models are accepted by regulatory authorities, this often only applies for a certain region. This means that the manufacturers still have to conduct conventional testing for other regions.
6. Factors related to infrastructure and maintenance
   Connected to the production factors, is the problem of the so called *"sunk investment in the existing infrastructure."* (Kemp et al., 1998, p. 180). The introduction of an alternative technology often requires an adaptation of the existing infrastructure and the available expertise to be able to control and maintain the new technology. In addition, it often requires complementary inventions e.g. new test infrastructures, new skills and new ways of interpretation of the results (Fagerberg, 2006). Existing knowledge may become obsolete or superfluous. As a result "t*he groups in charge of the current infrastructure form a strong lobby for their own interest"* (Kemp et al., 1998, p. 180). When it comes to 3R acceptance, the fact that manufacturers often refer to the problem about conventional testing that is still required in other regions, is an additional argument to maintain the connected knowledge and infrastructure.

7. Factors related to -societal and environmental- effects of new technologies
   The alternative technology may solve problems but is at the same time likely to bring along many uncertainties (e.g. nanotechnology and genetic engineering) and may cause unknown effects. According to Olson et al. (2014) uncertainty and unanticipated consequences are both a barrier and driver and are often mentioned as critical factors. In this research, the aspects of uncertainty and unanticipated consequences are observed to strongly influence the diffusion of alternative testing methods since it has to become accepted in a risk-averse context (see Chapter 5).

All in all, the established regimes or technological pathways have a serious amount of excluding power (Nelson and Winter, 1977, Dosi, 1982, Geels, Kemp et al., 1998). It's unclear what exactly causes this exclusion, however it is suggested that there are two key factors that lead to this excluding effect i.e. the consensus of engineering beliefs and the shared knowledge about the key parameters which are connected to the existing technology (Nelson and Winter, 1977) and beliefs about what is requested by the market/ end user (Kemp, Schot and Hoogma, 1998). Dosi speaks in this context of *'ex ante selection'* along the line of existing technological trajectories (Dosi, 1982). Due to the on-going globalisation, organizations are more and more interdependent of each other when it comes to the acceptance of new innovations and change of techniques. Specific norms, values, convictions, rules and procedures are often shared by organizations in a certain network or product sector. This 'collective memory' facilitates cooperation and mutual learning but also complicates the questioning of the old routines. In other words networks and sectors are increasingly institutionalised at an international level and within the strength of the institution lies the problem to change (Vermeulen, 2011, p.16-17).

According to Kemp, *"One of the key reasons why technological progress often proceeds along certain trajectories is that the prevailing technology and its design has already benefited from all kinds of evolutionary improvements in terms of costs and performance characteristics, from a better understanding at the user side, and from the adaptation of the socio-economic environment in terms of accumulated knowledge, capital outlays, infrastructure, available skills, production routines, social norms, regulations and lifestyles"* (Kemp, 1994, p.5-6). In other words, it has shown its value, whereas the new technology still has to demonstrate what it's worth.

4.3.2 Drivers

While regimes exhibit a high degree of stability and coherence, they are also dynamic and challenged by alternatives. This means that the path dependency and stability of regimes is relative. *"…regimes are continually subject to competitive selection pressures exerted by other regimes and by new sociotechnical configurations in niches. Often these pressures are weak and incoherent, but at other times they become stronger."* (Berkhout et al, 2003 p.21).

The literature on technology transition and system innovations identifies a variety of critical success factors or drivers which are observed to stimulate new technologies in escaping out of their niches and becoming part of the regulatory regime. Harder and

Benke (2006) stipulate that a combination of critical success factors is needed for technology transfer to occur. This section describes a series of critical success factors found in innovation literature and which are applicable in the context of this thesis.

1. Champions, entrepreneurial leaders and collaborations

   Champions are seen as critical participants in the successful outcome of the technology transfer (Harder and Benke, 2006). Osborne and Brown (2005: as cited in Meijer, 2013, p.5) identify the role of 'champion', 'supporter' and 'advocate' that are all about creating support and ensuring resources for innovation. They are also described as internal entrepreneurs (Tushman and Nadler, 1986), process promoters, salesman of the new idea (Chakrabarti and Hauschildt, 1989) and agents of change (Walter et al., 2011).

   Champions - "*individuals who actively pursue new product ideas to apply such ideas to product innovation and bring them to market"* - are pivotal to innovation processes (Walter et al., 2011 p. 586). *"Champions take up new ideas and aggressively fight bureaucratic, as well as social and political barriers to turn such ideas into successful innovations"* (Schon, 1963: as cited in Walter et al., 2011, p.586). Without them new ideas are likely to be disregarded. To be successful, champions have to be personally committed to the idea (Chakrabarti and Hauschildt, 1989).

   They display four behaviours i.e. pursuing innovative ideas, building networks, persisting under adversity, and taking responsibility, i.e. 'ownership' of the idea.

   Furthermore, innovation literature focuses on individuals holding executive positions in the organization; generally referred to as entrepreneurial leaders. *"A strong executive leadership is needed to drive innovations in the public sector through rhetorical leadership and coalition-building"*. (Doig and Hargrove, 1987: as cited in Meijer, 2013, p.5). *"Entrepreneurial leaders provide resources and guidance, and they lead by example."* (Harder and Benke, 2006, p.22). Meijer adds to the idea of champions and entrepreneurial leaders by distinguishing five roles for realizing innovations:

   - *Creators* are the 'intellectual leaders', the generators of the new idea and new ways of thinking. They manage to break through perceptual barriers.
   - *Innovation entrepreneurs* connect the idea to an existing problem.
   - *Test managers* manage to realize a successful test of the idea.
   - *Innovation packagers* embed the innovation into organizational structures and routines.
   - *Innovation diffusers* coordinate the large-scale roll-out of innovations and create incentives for other organizations to adopt the innovation.

   Each role fulfils its own task in the process from the actual innovation to its diffusion. *"This perspective helps to transform our understanding of the role of individuals in public innovation from an individualized one (great individuals) to a distributed one (great collaborations)"* (Meijer, 2013, p.6). Meijer concludes that hero innovators do not exist but distributed heroism does. Harder and Benke (2006), in the context of transportation technology transfer, refer to the selection of the various participants to form a partnership as an important factor to facilitate technology transfer. *"Building a network of partners is the key element in technology transfer"* (Rouach, 2003: as cited in Mamat and Roslan, 2012, p.169). Partnerships are needed to attract the right participants and are important to leverage resources (Harder and Benke, 2006).

2. Technology transfer strategy
   To be able to succeed in the technology transfer, the transfer partners must have a strategy/implementation plan (Harder and Benke, 2006) with clear shared goals and an overall planning to guide the transition process. This plan should also describe the specific tasks and connected roles for each of the partners.
3. Pilot projects and the search for promising connections
   It is imperative to start with the new technology in smaller contexts of specialized application where it can gradually show its value (Harder and Benke, 2006*). "The importance of learning by using is…widely recognized as important."* (Rip and Kemp, 1998, p.348). Pilot projects offer the opportunity to obtain practical experience with the innovation in a contained manner. Innovators can subsequently build upon the experiences retrieved through these pilots. Furthermore, technologies that are successful in creating a regime shift tend to succeed in involving systems of related techniques*. "Interrelatedness can be actively sought…. resulting in new technologies (with new rules and procedures) based on merging of previously distinct ones"* (Rip and Kemp, 1998, p.350). This is an example of niche accumulation which was already referred to in section 4.2.3. For niche accumulation purposes it is important to keep an open mind to new developments in other domains.
4. Early involvement of users and effective communication
   Implementation success is often related to the early involvement of end users in the process of the research and development of an innovation (Harder and Benke, 2006). Kemp et al. (1998) point to the need for close interaction between the technological development and the user environment. User-producer interface is seen as an important item in explaining innovation successes (Rip and Kemp, 1998). This aspect is also referred to by other scholars. Markard and Truffer (2008) and Geels (2002) for example, state that the level of institutional overlap between the niche, the TIS and the regime can be seen as an indicator for successful TT. In governing the potential success of the niche, the connection between stakeholders in the regime and stakeholders in the niche is therefore very important. A critical success factor for technology transfer is therefore the effective communication between the technologists who understand the technology, and the end users who will have to work with it. Both stakeholder groups have their social views (e.g. engineering ideas, management beliefs and the perception of users) with regard to the new technology. These social views are usually highly subjective and dynamic and may work in favor or against the new technology (Kemp et al., 1998). It is therefore very important that the interface between the parties involved in technology transfer is managed well (Boulter and Bendell 2002) to discuss and align the social views.

The critical success factors as described above enable active steering of technology transitions. These success factors can be combined in a strategic manner to stimulate innovations to compete with the existing regime. Embracing these critical success factors is especially promising in situations where the function of a sociotechnical regime is already under discussion.

De Haan and Rotmans (2011) describe three main mechanisms in which the function of sociotechnical regimes is compromised and the system is forced to respond and adapt, eventually leading to the destabilisation and change of a sociotechnical regime.

- If the system/regime is challenged by its environment, on which it depends to maintain an appropriate influx and outflux of resources like energy, financial capital, goods, labor or information, it is said to suffer from tensions. De Haan and Rotmans distinguish structural tension, which refers to problems with the physical, infrastructural, economical, formal and legal aspects and cultural tension, relating to problems concerning cognitive, discursive, normative and ideological aspects. Examples of signs of tensions are the depletion of resources, a hostile political climate, environmental/ethical awareness and the public opinion.
- If the system is comprised by its own composition, two conditions can be distinguished:
    - Firstly, the structures and cultures within a system have to match. In case of mismatching, for example if the system becomes rigid, one speaks of stress. In stress situations the regime proves inadequate or internally inconsistent in coping with the dominant way of acting as required by society. Typical signs of stress would be that the system is not practicing what it preaches or that there are perverse processes in the system, where the means become goals (De Haan and Rotmans, 2011).
    - Secondly, when novel constellations (niches) emerge and become competitors to the existing regime, the societal system becomes subjected to pressure. It is also possible that the novel or deviant constellations make aspects of the regime obsolete. In this case a new technology becomes available that takes over the leading role of the old technology. This competition can also emerge from different 'visions' for the future held by stakeholders in the field, some of whom are directly embedded within the regime *"e.g. the current contention over the use of 'science-based' risk assessment versus more broad-based 'precautionary' approaches to chemicals regulation"* (Berkhout et al., 2003 p.23).

In short, the existing regime can be destabilized by endogenous or exogenous critical junctures (Bakker, 2001). *Endogenous critical junctures* arise from within the sociotechnical regime, for example if the existing routines fail to meet the expectations. This includes the developments which are described above as stress-related developments. *Exogenous critical junctures* arise from outside the regime (tensions and pressure). These are for example the result of shifts in the landscape or by the linkage of several new technological developments.

4.4 Transition patterns

De Haan and Rotmans (2011) uncover three patterns in transition processes (i.e. reconstellation, empowerment and adaptation) which are related to the ways in which the societal system is challenged.

4.4.1 Reconstellation

In the case that a constellation is powerful enough to overthrow the regime, De Haan and Rotmans refer to reconstellation. Reconstellation is often a top-down initiative and the result of government-led structural change. The implementing of European policies on a national level would also fall under this signature. From literature it becomes clear that legislation and hard policy goals can have a clear stimulating effect on innovation. Lee et al., (2010) for example show that the high regulatory standards played an important role in forcing technological innovations and determining subsequent direction of technological change in the area of automobile emissions control technologies (see Lee et al., 2010).

4.4.2 Empowerment

In the case that small constellations gain power and become a competitor for the incumbent regime, De Haan and Rotmans speak of empowerment. Empowerment is a bottom-up movement in which a new constellation arises, or an existing one becomes more influential, either by itself or through interacting or merging with other constellations (niche accumulation). Empowerment can be stimulated through strategic niche management (see section 4.5.2).

4.4.3 Adaptation

A third transition pattern is that of adaptation. In this case a constellation adapts its functioning to better meet societal needs. The adaptation mechanism refers to the technological add-on and hybridization in which new technologies physically link up with established technologies, often to solve particular bottlenecks. Thus, old and new technologies do not immediately compete head on, but form some sort of symbiosis (Pistorius and Utterback, 1997). In this mechanism TT occurs *"….through a stepwise process of reconfiguration. New regimes gradually grow out of old ones"* (Van den Ende and Kemp, 1999: as cited by Geels 2002, p.1272). This is also referred to as cascade dynamics, in which *"…changes in one element of the regime trigger changes in other elements which, in turn, trigger further changes."* (Geels, 2002 p. 1272)

Adaptation through the incorporation of alternative functions is the typical regime response when it is compromised. Adopting the alternative in these cases is primarily done to continue or expand the dominant way of operating (e.g. oil companies moving into green energy sources next to their oil based activities, De Haan and Rotmans, 2011). Adaptation is a primarily evolutionary approach in which the system is gradually changed. Transition management scholars, being aware of the risks of sudden actions in redesigning policy, embrace the importance of the evolutionary change or adaptation mechanism. Kemp et al. even describe transition management as *"a new steering concept that relies on 'Darwinistic' processes of guided variation and selection instead of planning."* (Kemp et al. 2007: as cited in Meadowcroft, 2009, p.325).

Transition processes are usually the result – or a combination - of empowerment, reconstellation and adaptation. It is a heterogeneous and contingent process; "*an innovation*

journey, with setbacks and new ventures, rather than the execution of a plan" (Van de Ven et al., 1989: as cited in Rip and Kemp, 1998, p.347) in which co-evolution and mutual adaptation between the technology and the system for which it's destined is pivotal. According to Dosi (1982) the emergence of new technological paradigms stems from the interplay between scientific advances, economic factors, institutional variables and unsolved difficulties on the established technological paths. The alignment of these variables can be stimulated in the following coordinated ways.

4.5 Intervention strategies

The following section describes two important interventions strategies that can be found in the literature on Technology Transitions (TT). The technological transitions framework has led to recommendations for policy intervention and governance options through concepts such as transition management (Geels and Kemp, 2000; 2006; Kemp and Rotmans, 2005) (see Markard and Truffer, 2008) and strategic niche management (Hoogma, 2000; Hoogma et al., 2002; Smith, 2003). These strategies combine several of the previously described critical success factors in a thought through manner.

4.5.1 Transition management

An intervention strategy that offers recommendations for policy intervention and governance options at the level of the socio technical system is transition management. This is the effort to guide or facilitate sustainable transitions and influence the speed and direction of the evolution of a sociotechnical system. The objective of transition management is to steer bottom-up, niche-to-regime processes of transformation towards a pre-defined goal or 'vision' (Rotmans et al., 2001). The starting point is the articulation of the vision (Berkhout et al., 2003). Rotmans et al. define transition management as *"the sum of current policy plus long-term vision, coherence, short-term action for keeping open and exploring options and process management (development rounds)"* (Rotmans et al., 2001, p.6). Transition management aspires to create changes that are beneficial to society, while innovation policies primarily seek to strengthen the economic positions of firms. (Alkemade et al., 2011).

The central issue transition management seeks to address is breaking out of the socio-technical 'lock-in' of conventional technologies such as fossil fuel-based energy systems and the 'lock-out' of alternative technologies such as carbon-saving technologies. According to Meadowcroft (2009, p.325-326) transition management has a number of promising features to offer e.g.:

- It uses longer time frames and explores alternative trajectories;
- It develops networks of actors in a particular production/consumption nexus in which actors can come together, develop shared problem definitions, appreciate differing perspectives, and above all develop practical activities;
- It stimulates 'learning-by-doing' i.e., developing experiments with novel practices and technologies to learn about their potential and limits;

- And it encourages variation in and selection among different innovative approaches.

A common trigger for virtuous cycles in the field of sustainable technologies is the guidance of the innovation process. Transition management insists that policy makers follow two parallel tracks i.e. the track of incremental adjustments to existing practices, which is referred to as 'system improvement', and the track of experiments with fundamental adjustments to dominant designs which is referred to as 'system innovation' (Meadowcroft, 2009). Transition management therefore habitually applies the following combination of policy instruments: sector-based collective visions, collaborative and experimental projects, and state expenditure to promote networking and innovation in technologies combined with more traditional policy tools such as regulation, planning, and tax-based instruments (Meadowcroft, 2009).

Evaluation of innovation policy instruments in several OECD countries (Boekholt et al., 2001) showed that the policy instrument most used to stimulate innovations are financial instruments. Smith and Kuhlman (2004) (see Hekkert et al., 2007) however, conclude that innovation processes require much more attention for instruments that support governance structures, for example:

- The building and organizing of (innovation) systems to develop a vision and a strategy, to initiate discourse and thereby align different perspectives and work towards consensus. Furthermore, such systems are important to identify and facilitate prime movers and to involve other relevant actors;
- The management of interfaces crossing sectorial, organizational and institutional borders to overcome tunnel visions and stimulate the debate;
- The creation of platforms to exchange knowledge, to learn, stimulate demand articulation, develop a strategy and a vision;
- The development and support of an infrastructure for the production and exchange of strategic information tailored to the needs of stakeholders involved and accessible to all relevant actors.

Champions/policy entrepreneurs/entrepreneurial leaders play an important role in these governance structures (see section 4.3.2). Policy entrepreneurs (e.g. Kingdon, 1995, Baumgartner and Jones, 1993) are defined as people who seek to initiate dynamic policy change by attempting to win support for policy innovation (Mintrom, 1997, p.739) For this, they make use of the following activities: identifying problems, networking in policy circles, shaping the terms of policy debates, and building coalitions. According to Hekkert et al. (2007), a possible start for virtuous circles are entrepreneurs who lobby for better economic conditions (e.g. more R&D resources or market options) to enable the technology development. As clarified in section 4.3.2 this role often is a shared responsibility by several entrepreneurial leaders who all play their specific part (Meijer, 2013).

4.5.2 Strategic niche management

The sustainable development of new technologies requires an interrelated social and technical change (Schot and Geels, 2008). However, *…"technology actors usually focus on developing, testing and optimizing technology, but neglect the embedding in broader societal goals, or leave it to a later stage"* (Schot and Geels, 2008, p.538). That is where strategic niche management (SNM) comes in. Kemp, Schot and Hoogma (1998, p.186) define SNM as: *"The creation, development and controlled phase-out of protected spaces for the development and use of promising technologies by means of experimentation, with the aim of (1) learning about the desirability of the new technology and (2) enhancing the further development and the rate of application of the new technology."*

SNM is an example of empowerment and introduces reflection into the transition process. The assumption underlying SNM is that if such niches are constructed appropriately, they can act as building blocks for broader societal changes with a clear and flexible transition vision as the central criterion, through which: "*Competencies and new skills are built up. New markets are created and user demand promoted. The position of the new configuration is strengthened and a normatively-desirable transformation of the regime is moved on".* (Berkhout et al., 2003, p. 12).

SNM refers to a bottom-up process that requires steering from within. Steering can refer to many different actions aimed at pushing the developments into the right direction. Examples are the introduction of new (group of) actors, a new perspective, a specific learning process or a set of pilots which may redirect the status quo (Schot and Geels, 2008).

Insights from innovation studies have led to three (internal) processes that are observed to be relevant for the successful development of a technological niche i.e. the articulation of expectations and visions leading to a shared agenda; the building of social networks; and learning processes at multiple dimensions. These led to the following set of more specific hypotheses (Elzen, Hoogma, and Schot 1996, p.76–78; and Hoogma et al. 2002, p.28–29: see Schot and Geels, 2008):

- Expectations contribute to successful niche building if expectations were made:
    - more robust (shared by more actors);
    - more specific (if expectations are too general they do not give guidance);
    - and have higher quality (the content of expectations is substantiated by ongoing projects).

These expectations coordinate and motivate actors to act upon the shared agenda (Van Lente, 1993: as cited in Boon et al., 2014) and can thereby become 'performative' (Borup et al., 2006: as cited in Boon et al., 2014).

- Social networks are likely to contribute more to niche development if:
    - the networks are broad, i.e. multiple kinds of stakeholders are included to facilitate the articulation of multiple views and voices; the involvement of relative outsiders may be particularly important to broaden cognitive frames and facilitate second-order learning;

  - the networks are deep, i.e. people who represent organizations, should be able to mobilize commitment and resources within their own organizations and networks.

Actors within niches try to recruit or motivate other actors to join the niche's network by sharing the agenda of technology development and thereby aim for actors who can contribute to the formation and functioning of an innovation through the provision of resources, legitimacy or social capital (Boon et al, 2014).

- Learning processes contribute more to niche development if they are not only directed at the accumulation of facts and data, i.e. first-order learning, but also enable changes in cognitive frames and assumptions - values and norms -, i.e. second-order learning. (derived from Grin and Van de Graaf 1996)

In the context of the protection of niches Boon et al. (2014) distinguish two extreme poles, i.e. the restrictive versus accommodating protection strategy, which are each other's opposites in terms of the level of 'openness' i.e. the degree to which different views and interests are accommodated by the lead actors. The restrictive protection strategy is characterized by a concentrated and homogeneous network in which there is no connection between included and excluded actors, discounting deviant voices, fast first-order learning and little reflection as part of second-order learning. The accommodating protection strategy focuses on reflective second-order learning, slower first-order learning and capturing ideas and actors from outside the niche. It includes arbitration of differing views, creating a platform for discussion, negotiating compromises, capturing others' perspectives and broadening the scope of the niche. The advantage of restrictive strategic niche management is the fact that the small core group of included actors can learn and implement quickly. The downside is that the excluded actors are not taken into account or at a late stage, leading to lower external acceptance and less robust niche narratives (Rip et al., 1995). The major activity to strengthen the niche is the formation of a narrative that legitimizes niche protection. "*The expectations and learning about innovative safety monitoring are translated into a narrative that niche advocates use when they interact with actors outside the niche during their empowerment work"* (Boon et al., 2014).

From a series of case studies (e.g. Hoogma, 2000 and Raven, 2005) it can be concluded that unsuccessful niche developments are repeatedly associated with minimal participation of outsiders in the experiments, a deficiency in terms of second-order learning and to minimal involvement of regime actors resulting in insufficient resources and institutional embedding. Another conclusion is that networks that were broad and contained outsiders provoked more second-order learning. (Schot and Geels, 2008). From Hoogma's case studies on the introduction of electric case it became clear that it is important for niches to closely follow the dynamics in the surrounding environment. Without this the lessons learned at the experimental stage are unlikely to be adopted by a larger network (Hoogma, 2000).

4.6 Conclusions

This chapter addresses Q2: *Which theoretical perspectives are needed to comprehend the process of regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes and find suitable ways of enhancing the process?*

To understand and deal with the wicked problem of regulatory acceptance and use of 3R models the integrative/multidisciplinary approach of Technology Transitions (TT) is adopted. This approach offers the possibility to combine different perspectives needed to cover the contextuality and complexity of this issue.

TT refers to major technological transformations in the way societal functions (e.g. communication, transportation, energy supply etc.) are fulfilled. Risk assessment and efficacy testing of products and substances such as pharmaceutical/vaccines and chemicals may be defined as such a social function. In the TT approach, it is only in association with human action, social structures and organizations that technology fulfils functions and becomes meaningful. 3R models, being a group of technologies that has the potential to contribute to far reaching changes in the way products and substances are produced and assessed is defined as a group of innovations which could lead to Technology Transitions (TT). The multilevel framework of TT offers a suitable basis for the analysis of the process of regulatory acceptance and use of 3R models and the influencing factors playing a role in this process. Its value lies in the fact that it does not only involve technological aspects, but also includes changes in elements such as user practices, regulation, infrastructure, and social beliefs. This interplay of different categories of influences already became apparent in the inventory that anticipated this thesis (see Appendix I). The multilevel perspective, in which regimes are embedded within landscapes and niches within regimes, offers the basis for the analytical frame to analyze the barriers and drivers influencing regulatory acceptance and use. Each level entails specific features which enable or withhold an innovation breakthrough. Where the multilevel framework of the TT approach is very useful for a comprehensive analysis of technology transitions through the inclusion of all three levels and their interplay, the SI is useful in analyzing the roles and strategies of the different actors and their interaction and offers concrete options to stimulate progress.

In the literature on Technology Transitions (TT) and System Innovations (SI), it becomes clear that established regimes or technological pathways have a serious amount of excluding power caused by a series of factors. The following barriers are retrieved from literature: the existing technology has benefited from many enhancements that were made to it throughout the years; as a result, all kinds of values are attributed to it and misfit occur of alternative technologies with the existing regime; unpredictable and inconsistent policies withhold industries from innovating and strict regulations often have an excluding effect on new technology; the 'Not Invented Here' syndrome plays a role for those stakeholders that were not involved in the development of the innovation and manufacturers anticipate on the potential rejection of innovation by end-users; new technologies require adaptation of the existing infrastructure and expertise and may have unanticipated effects.

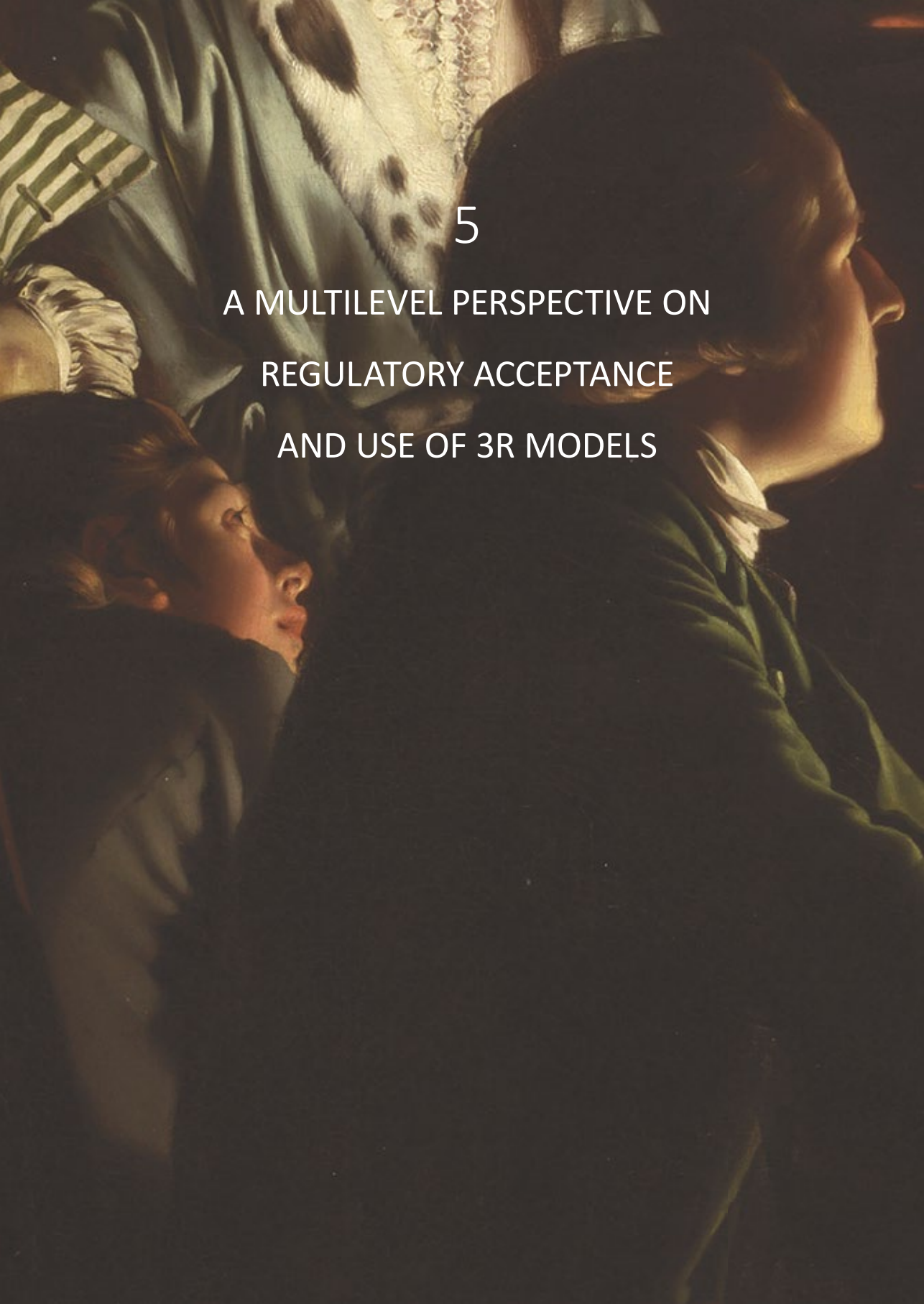
However, regimes are frequently observed to change and the following success factors and mechanisms of change can be found in the literature:

Champions and entrepreneurial leaders are strongly connected to the success of technology transfers; collaborations between advocates of the new technology are important to learn and share experiences; partnerships offer the possibility of finding and attracting the right participants; a clear technology transfer strategy is needed which formulates the mutual goals and required activities; pilot projects and demonstrations are needed to allow hands-on learning and the early involvement of the end-user is key to allow early resolution of problems and to prepare the user for the innovation. Tensions in the social system regarding the technology, user practices, symbolic meaning of technology, infrastructure, policy changes and technoscientific knowledge, are also potential drivers and may lead to periods in which linkages are weakened and new technologies become more important.

Technological change depends on the outcome of the balancing act of the drivers and barriers and the way how the sub-functions of the innovation system such as entrepreneurial activities, the generation and diffusion of knowledge through networks, the mobilization of resources and the creation of legitimacy are connected. Three patterns are distinguished in technology transitions, i.e. reconstellation, empowerment and adaptation. Elements of all three mechanisms are relevant in the context of this thesis. However, the focus in terms of optimization is on the latter two mechanisms since they offer the most practical options to stimulate the transition towards 3R models in the regulatory domain.

The intervention strategies of transition management and strategic niche management offer valuable options to enhance technology innovation/technology acceptance processes which are found to be relevant in terms of stimulating the TT towards the 3Rs. These options will be elaborated on in Chapter 10 of this thesis.

In short, the multilevel perspective on Technology Transitions (TT) offers the integrative and broad perspective needed to examine the complex and wicked problem of 3R acceptance in the regulatory domain. This perspective is therefore used as the basis for the 3R acceptance model which will be presented in Chapter 5 and which will serve to analyze the influencing forces at hand and the dynamics between them. To tailorize this perspective to the context of risk and safety assessment, in which regulatory testing is situated, Chapter 5 will adopt a risk regulation perspective in addition to the technology transition perspective.



5

A MULTILEVEL PERSPECTIVE ON REGULATORY ACCEPTANCE AND USE OF 3R MODELS

"…we have again and again encountered instances of long delay in the application of existing knowledge to the improvement of experimentation… Delays of this kind may be regarded as a sort of inertia, or rigidity, the maintenance of a habit (positive or negative) long after information is available for its correction. In the individual organism, rigidity of this kind has been shown to be associated with isolation, or lack of communication between central nervous mechanisms… It is entirely reasonable to expect a similar relationship at the sociological level…"

W. Russell and R. Burch
Authors of 'The Principles of Humane Experimental Technique'
(1959)

A short guide to Chapter 5

This chapter consists of a manuscript which was published in Altex in 2012 (for full reference see below). It serves several goals: To start with, a further analysis is made of the drivers and barriers that were retrieved in the research prior to this thesis (Schiffelers et al., 2007: see Appendix I) by using the multilevel perspective on Technology Transition (TT). This analysis resulted in the 3R acceptance model which is based on the multilevel perspective on TT as described in Chapter 4 and is presented in this chapter. It is used throughout this research as a tool to unravel the Gordian knot of influencing drivers and barriers by categorizing them into the micro-, meso- and macro level at which they exert their primary influence. Subsequently the drivers and barriers are analyzed in terms of their interaction, their dominance and their potential pliability.

In addition, Chapter 5 introduces a perspective of risk regulation to customize the general TT perspective to the context in which regulatory acceptance and use of 3R models takes place. The initial research (see Appendix I) uncovered risk aversion as a dominant feature when it comes to the acceptance and use of 3R models for regulatory purposes. A risk regulation perspective is therefore needed to examine the effects of this specific setting on technology transitions. Through these steps this chapter targets the research questions Q2, Q3a and Q3a (see section 1.2). For the sake of the completeness of the publication a certain amount of replication of the chapter 1 to 4 was necessary.

Abstract

The importance placed on risk avoidance in our society has resulted in a broad range of regulations intended to guarantee the safety of products such as pharmaceuticals and chemicals. Many of these regulations rely on animal tests. As a result, about 25% of the animal experiments in Europe are conducted for regulatory purposes. There are many initiatives that aim to replace, reduce, or refine laboratory animal use, but the regulatory acceptance and use of 3R models lags behind. The central question of this study is: *"Which variables influence the regulatory acceptance and use of 3R models and in what way?"* Regulatory acceptance is seen as one of the biggest hurdles 3R models face, but the rationale behind this is still underexplored. This study is an approach to filling that gap by combining opinions from experts in the field with literature on technology acceptance and risk regulation, resulting in a model of the variables that determine the process of the regulatory acceptance and use of 3R models.

Reference[49]
Schiffelers, M.J., Blaauboer, B., Hendriksen, C. and Bakker, W. (2012).
Regulatory acceptance and use of 3R models: a multilevel perspective. *ALTEX* 3, 287-300

49 Acknowledgements: We thank the Doerenkamp-Zbinden Foundation for sponsoring this project.

5.1 Introduction: The problem of regulatory 3R acceptance

The quote of Russel and Burch (the founding fathers of the 3R Principle) at the start of this chapter perfectly illustrates the problem 3R models[50] have faced ever since the principle was introduced in 1959, namely the slow acceptance of these 'new' technologies. Even though numerous 3R models have been developed over the past decennia the regulatory acceptance and use of these models, i.e., to prove the quality and safety of chemical compounds, pharmaceuticals, and biological products (e.g., vaccines), lags behind. Though new technologies often face a hard time in getting accepted, numerous innovations have been able to cross the existing barriers and considerable research has been dedicated to the process of innovation transfer. This study makes use of the general notices on technology transfer to better comprehend the process of regulatory acceptance and use of 3R models and the variables that drive or hamper this process. In addition, it examines ways to facilitate the process.

We live in a society that is confronted with many products that might pose a risk to our health and to the environment. To minimize the possibility of adverse effects from these products, a complex system of rules and regulations has been designed. In the EU alone, there are more than 800 laws, regulations, directives, and other documents regulating product safety and quality (De Leeuw, 2004). The requirements determine the endpoints that products must be tested on by manufacturers and, occasionally, by regulatory authorities before they are released for commercial purposes (e.g., toxicity, efficacy, etc.). Animal testing is a 'traditional' element of these product assessment procedures and accounts for at least 25% of the animal experiments conducted within the European Union (Anon., 2010). Currently, however, many developments focus on alternative testing strategies and aim to move away from animal testing. These developments include the many *in vitro* models for Research and Development (R&D) purposes of new drugs or chemical compounds.

The increasing number of 3R models available for regulatory purposes offer the possibility for manufacturers and regulators to choose the method they perceive as most suitable in those cases where the regulatory requirements offer discretionary space.[51] Furthermore, there are many 3R partnerships, both at the national and the supranational level, in which regulatory testing is part of the agenda. In addition, there is horizontal European legislation to protect animals used for scientific purposes (Directive 2010/63/EU), which states that 3R models shall be used wherever possible (European Commission, 2010).[52] [53]

50 All procedures which can completely replace the need for animal experiments, reduce the number of animals required, or diminish the pain or distress suffered by animals.

51 The legal dictionary defines "discretion" as: The power of a judge, public official or a private party (under authority given by contract, trust or will) to make decisions on various matters based on his/her opinion within general legal guidelines.

52 Article 13 of European Directive 2010/63/EU for the Protection of Laboratory Animals stipulates that: *"Member States shall ensure that, wherever possible, a scientifically satisfactory method or testing strategy, not entailing the use of live animals, shall be used instead of an animal based procedure."*

53 Being horizontal legislation (see section 2.5.1), this directive applies to all product sectors that are involved in animal testing. This means that scientists, manufacturers, and regulatory authorities within these sectors have the duty to ensure that animal usage is kept to a minimum and animal health and welfare legislation is upheld.

So in those cases where regulatory product requirements describe a 3R model or a battery of 3R models and the conventional animal model and offer discretion to regulators and industry to choose a testing method, they should, in light of this Directive, decide to use the 3R option. Finally, the objections to regulatory animal testing from an animal welfare, an economic, and a scientific perspective are on the increase, as will be elucidated below. These objections make regulatory animal testing an important area to evaluate in terms of the 3R principles. All in all, there are many reasons to move away from regulatory animal testing and toward the use of alternative models. However, despite all these initiatives, the heavy reliance on animal models for safety and quality testing purposes is very persistent. Many stakeholders in the field of alternatives to animal testing know of 3R models that were developed up to decades ago and still are not accepted by regulatory authorities or used by manufacturers for regulatory purposes.

With regard to the 3Rs, most studies focus on the technical possibilities and limitations of specific methods. Research concerning the regulatory process and the ultimate use of 3R models is limited (Freriks et al., 2005).[54] This study aims at filling that gap by answering the following central question: *"Which variables influence the acceptance and use of 3R models for product regulation purposes and in what way?"* It offers an overview of the variables influencing this process. To demarcate the research area, we start by defining regulatory acceptance and use:[55]

Regulatory acceptance: refers to the written or unwritten adoption of testing strategies by regulatory authorities. Regulatory acceptance in this context is defined as the formal adoption of a (validated) test method by a regulatory agency/authority. Depending on the product sector, regulatory acceptance can be accomplished at a national, a European, and/or a global level.

Regulatory use: refers to the actual uptake of a method by a regulatory authority or a manufacturer for quality and/or safety testing purposes. This step is often also referred to as implementation. In the field of policy science, however, implementation would cover the whole process from the initial intention to work towards alternatives to the actual use. For this reason the term regulatory use is preferred in the context of this article. Regulatory use for the purpose of this paper is looked upon as a function of regulatory acceptance in which the level of acceptance strongly determines the level of regulatory use. Although there are some cases in which regulatory use has anticipated regulatory acceptance or occurred without formal regulatory acceptance, these situations are seen as isolated cases.

---

54 Some examples of studies that focus on this process are: NIEHS, 1997; Garthoff, 2005; Schiffelers et al., 2005; Bottini et al., 2008.

55 In the subsequent steps of this research a somewhat different distinction has been made. This resulted in the substages of Formal Incorporation (FI: which refers to formal adoption of a 3R model into regulatory requirements), the Actual Regulatory Acceptance (ARA: which refers to the acceptance of a 3R model by regulatory authorities for daily assessment practices) and the Use by Industry (UI: which refers to the use of a 3R model by manufacturers to meet regulatory requirements).

5.2 Investigative approach

In order to answer the central question, this chapter elucidates the process of acceptance and use of new technologies such as 3R models and offers an overview of variables that influence this process. It builds on earlier work (Schiffelers et al., 2007: see Appendix I) and intends to bring the discussion a step further by systematically describing the process of acceptance and use and by presenting the '3R Acceptance Model', which offers an overview of the variables that are observed to influence the process of acceptance and use. For this purpose a combination of theoretical and empirical factors is described. The theoretical variables derive from theoretical perspectives on risk regulation and the acceptance of innovations (see Chapter 4). The empirical findings partly derive from the earlier work conducted in this field (Schiffelers et al., 2005, 2007). These findings were tested and updated by an additional series of ca. twenty interviews with representatives from European and US regulatory authorities, industry, and academia familiar with the subject of regulatory testing and the acceptance and use of 3Rs models. Furthermore, regulatory testing was part of the agenda in a series of international meetings attended throughout the period of 2009-2011.[56]

The empirical findings, presented in this study are the ones the researchers came across most frequently throughout the different interviews and meetings. It must be emphasized that the type and weight of the factors might differ between geographical regions, product groups, industries, and agencies. To clarify the process of acceptance and use of 3R methods for regulatory purposes the following steps are taken. First, the multilevel approach is described. Second, this multilevel perspective is applied to the field of regulatory acceptance and use of 3R models. This results in a description of the relevant variables that are considered to influence the transition towards 3R acceptance and use in the area of product regulation. Third, these variables will be recapitulated in the '3R Acceptance Model'. This model consists of the variables that are perceived to play a relevant role in the regulatory acceptance and use of 3R models. To conclude, a distinction is made between the so-called rigid variables (variables that are relatively difficult to manipulate) and the more pliable variables (Ellemers, 1976). This division is important when discussing ways to optimize the process of acceptance and use.

5.2.1 The multilevel approach: an inclusive model to comprehend regulatory 3R acceptance

The process of regulatory acceptance and use of 3R models is determined by a broad set of factors. In order to cover and comprehend the complex reality, the multilevel model of innovations and technology transformations is used (see also Chapter 4 for an in-depth

56 3R symposium: Looking into the crystal ball (January, 2009, Utrecht, The Netherlands); 7th and 8th World Congress on Alternatives and Animal Use in the Life Sciences (August, 2009, Rome, Italy; and August, 2011, Montreal, Canada); ECVAM/EPAA workshop on The Consistency Approach for Quality Control of Vaccines- a 3Rs Opportunity (January, 2010, Brussels, Belgium); Workshop ICCVAM/NICEATM: 3Rs in Vaccine Potency Testing (September, 2010, Bethesda, USA); Conference EDQM: Quality of Medicines in a Globalized World: Dream or Reality? (October, 2010, Prague, Czech Republic).

description of the multilevel perspective). This model offers valuable concepts for the analysis of long-term technological transitions by integrating insights from several disciplines (Geels, 2006). Such an integrative perspective is important in creating a complete picture of the dynamics of system innovations such as the transition from animal models to the 3Rs. Furthermore, it addresses the strong interlinkage between social and technical aspects in such technology transitions. A visual representation of the multilevel model was introduced by Rip and Kemp in 1996 (see Figure 9).
The model specifies three different levels important to look at when analyzing system innovations, viz.:

- the micro- or niche level in which novelties are developed;
- the meso- or sociotechnical regime level, which consists of the patchwork of rules and regulations, available expertise, current practices, and connected institutions;
- the macro- or sociotechnical landscape level, which relates to elements such as the material infrastructure, existing political culture and coalitions, social values, world views, the macro economy, demography, and the natural environment (Kemp, 2010).

Transitions come about through the interaction of processes at the three different levels. At stage 1 (see Fig. 9) innovations, like alternative test approaches, commence in niches, isolated from the existing regime. There is not yet one dominant design of the new technology, and experiments are conducted to work out the best design and find out what users want. The networks that support the innovations are small and unstable and *"the innovations do not (yet) form a threat (i.e., competing model) to the existing regime"* (Kemp, 2010: 293). At stage 2, the system begins to shift and the process of change starts. *"The new technology develops a technical trajectory of its own and rules begin to stabilize (e.g. a dominant design). But the innovation still forms no major threat to the regime, because it is used in specialized market niches"* (Kemp, 2010: 293). Phase 3 is that of the 'wider breakthrough' a result of *"an accumulation of socio-cultural, economic, ecological and institutional changes that react to each other."* The new technology, at this stage, is in competition with the existing regime. According to this approach, system innovations come about because the developments at the different levels *"link together and reinforce each other"* (Geels, 2006: 176). This means that system innovations are hardly ever the result of one single factor or actor, but are the *"result of the interplay between many processes and actors"* (Geels, 2006: 176).

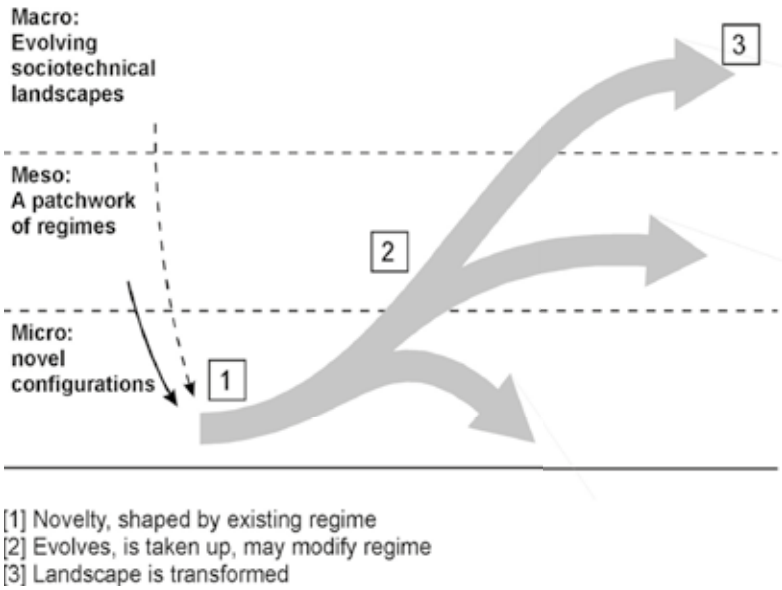


Figure 9. The multilevel model of innovation and transformation (Source: Rip and Kemp, 1996)

5.3 Results: 3R acceptance from a multilevel perspective

We will now go into each of the three levels, starting with the micro level, and define the variables that are observed to play a role at each level.

5.3.1 Influences at the niche level

At the micro level of technological niches, 3R models are novelties designed against the background of existing (well developed and established) product regulation regimes. The niches function as 'incubation rooms' that protect the new technologies from the market selection mechanism. Successful innovations will eventually break through to the regime level, but: *"New technologies may remain stuck in these niches for a long time (decades), when they face a mismatch with the existing regime and landscape."* (Kemp, 2010: 293) And this is exactly what can be observed for many 3R models, as will be described in this section. At the micro level of niches, several factors can be observed that influence the process of regulatory acceptance and use of 3R models, such as the potential of 3R models (see 5.3.1.1) versus their limitations (see 5.3.1.2), the education and experience of stakeholders involved with these models (see 5.3.1.3), and the validation process (see 5.3.1.4).

5.3.1.1 The promising potential of 3R models…

The 3R Principle was first introduced by Russell and Burch in 1959 in their revolutionary work The Principles of Humane Experimental Technique. Since then, many 3R models

have been developed and the 3R Principle is still gaining ground. *"The motive for developing and incorporating the 3Rs is usually neither altruism nor public relations. Rather, methodological improvements are sought as a means of overcoming the technical limitations inherent in current animal models. To practicing scientists, these more elegant and relevant methods represent technical progress and are considered to be additional or advanced, rather than alternative, methods… Such methods are often more valid and more reliable than those traditionally used in regulatory testing."* (Richmond, 2002) This quotation of Richmond, former head of the Home Office's Scientific Procedure division, illustrates the potential 3R models have from a scientific perspective. The models tend to have a more solid scientific base than the conventional animal models.[57] Most of the animal models were developed decades ago and are the product of a process of trial and error. They often lack formal validation and some show problems of high variability in test results and an extrapolation gap between the animal model and the human being. In addition, they frequently are a concern in terms of animal welfare and are often time consuming in comparison to the alternative model.

So even in cases where the scientific value of the animal model is questioned and 3R options have been available for quite some time, 3R models might face difficulties in breaking through, as is illustrated by the rabies case (see Chapter 6 for a full description of this case study).

5.3.1.2 …versus the refractory reality

Many 3R models are already quite extensively used for R&D of (new) products but, as mentioned, they often face a hard time in getting accepted at the regime level of existing rules, regulations, and testing practices, and so they fail to become genuine rivals to the respective conventional animal models, let alone become the leading paradigm in risk assessment procedures. According to Geels, a radically new technology has *"…a hard time to break through since it does not solely involve a change in the technology but because regulations, infrastructure, user practices and maintenance networks are aligned to the existing technology"* (Geels, 2002, p.1258).

This means that the more drastic the change when compared to the animal model, the more difficulties the 3R model is likely to face in terms of regulatory acceptance. Replacement models, in fact, embody a more radical change to the existing regime than reduction and refinement models, which are generally still based on the design of the conventional animal model. However, even refinement methods such as social housing of rats can face stumbling blocks when it comes to regulatory acceptance, as is the case in some OECD guidelines (Verwer et al., 2007).

The difficult process of acceptance is, among other things, the result of the fact that 3R models (especially replacement models) lack the ability of animal tests to mimic the entire organism. This means that a 3R model generally is not a stand-alone model but just one

57 Here it should be noticed that for the sake of reducing complexity the 3R models are being referred to as one package of innovations. In reality, however, 3R models consist of a wide range of techniques, varying from techniques that still rely on animal models but to a lesser extent to models that fully replace the animal model.

part of the puzzle, and so a combination of tests is needed to replace an animal method. On top of that, the research and development-base for novel approaches becomes smaller. It seems that the low hanging fruit have been picked, and scientists now face the challenge of developing models for more complex endpoints like carcinogenicity (ability of a compound or product to cause cancer), systemic toxicity (ability of a compound to induce organ toxicity), and reproductive toxicity (ability to harm the developing fetus or organs of reproduction). The problem here is that the development of these more complex alternatives *"…is bound up with the progress of science in developing a deeper understanding of fundamental biological processes"* (Rudacille, 1999). On the other hand, science is proceeding and thereby offering new possible approaches such as omics technologies in toxicity testing and physico-chemical methods in vaccine quality control.

5.3.1.3 Past education and former experiences

An important variable of influence on the level of acceptance of 3R models lies in the past education of the stakeholders that have to work with the models. In large part, the current generation of regulators was educated some 20-30 years ago when the credo still was '*in vivo veritas'* (Schiffelers et al., 2007, p.274: see Appendix I). This education remains very influential in terms of the level to which stakeholders feel comfortable with certain models. To gain trust in new techniques it is important to be able to work with them and gain experience with the way they function.[58] Education and training are therefore important aspects in the acceptance and use of 3Rs models. Only positive experiences with the new techniques can bend the convictions towards the 3Rs.

Due to their education and greater exposure to 3R models, the new generation of regulators most likely will incline more towards *in vitro* methods (Schiffelers et al., 2007: see Appendix I). This development can, for example, already be observed in Europe in the area of biologicals, where several Official Medicine Control Laboratories (OMCL's) have developed models to replace, reduce, or refine conventional animal models.

5.3.1.4 The validation challenge

A crucial step on the way to the implementation of alternative methodologies is the need to validate these tests. In general, regulators will accept alternative toxicity testing methods only after they have been scientifically validated. This means that they have been shown to be reliable (reproducible) and relevant for their intended purpose.[59] If the test is going to be used for regulatory purposes its robustness has to be demonstrated to the regulatory authorities (EPAA, 2007). Validation, therefore, can be regarded as a sort of 'gate keeper' to prevent 'immature' tests from entering the regime level. According to the OECD, formal validation *"contributes strongly to the international acceptance of any proposed test method"* (Spielmann, 2000). As a result, the OECD has indicated that *in vitro* toxicity studies can be accepted for regulatory purposes only after a successful experimental validation study.

58 It must be noted here that risk assessors can also be developers of 3R models themselves, as can be observed in within the European Official Medicine Control Laboratories (OMCL's). In these cases this argument can be discarded.

59 http://alttox.org/ttrc/tox-test-overview/ : consulted December 2011

But demonstrating the validity of a method to regulatory authorities is a challenging process. And every single step of a formal validation study brings about many challenges in terms of time, costs, and motivation[60] (Spielmann, 2000). One of the main challenges is that the validity of a 3R method often is evaluated by comparing it with the conventional *in vivo* model, even though the 3R model is in most cases incomparable to the conventional model, and the animal model might generate highly variable results or might be of questionable relevance. This makes it almost impossible to demonstrate correlation between the conventional and the new model, thereby creating a major hurdle in getting the new model through the validation process. On top of this, validation is a confusing concept, meaning different things to different people and under different circumstances. This has led to some situations where a validation study raised more questions than it answered Metz et al., 2002).

Although validation is an important step, it is not indispensable for regulatory acceptance. In some cases 3R tests have been accepted by regulators without being formally validated. This happened, for example, with an *in vitro* dermal absorption test. European industry in this case submitted in-house validation data to the OECD, and after peer review and international discussions, an OECD Technical Guideline for *in vitro* dermal absorption testing was adopted (Liebsch and Spielmann, 2002).

Conversely, validation does not automatically lead to regulatory acceptance. This often is a consequence of insufficient consulting of regulatory authorities in the phase of validation and failing to take the right criteria on board to validate the test for regulatory purposes. Therefore, an early involvement of regulatory authorities when validating a method is often considered a critical success factor (Bottini et al., 2008).

5.3.2 Influences at the meso level of risk regulation regimes

The meso level is formed by the sociotechnical regime, which is at the heart of the transition scheme. The term 'regime' refers to the deeply rooted collective memory of stakeholders of dominant practices. It consists of a semi-coherent set of rules, search heuristics, or paradigms relevant to that domain, giving it stability, orientation, and guidance in the decision-making (Kemp, 2010, Geels, 2002, Geels, 2006). Stakeholders within these regimes, such as regulatory authorities and industry, act according the 'logic of appropriateness', meaning that they do what they think is expected, legitimate, and rightful in the role they fulfill (Bakker, 2001). And since regimes are subject to pressure from both the macro- and the micro level, stakeholders within these regimes will have to cope with these pressures in an appropriate manner. *"Faced with these pressures, regime actors will typically opt for change that is non-disruptive, … , which leads them to focus their attention to system improvement instead of system innovation*" (Kemp, 2010, p.293).

At the meso level, several fundamental aspects are identified that are important to understanding the process of acceptance and use of 3R models for regulatory purposes, such as the leading technological paradigm (see 5.3.2.1), diverse risk regulation regimes (see 5.3.2.3), the informational asymmetry between regulators and industry (see 5.3.2.3), and transition costs (see 5.3.2.4).

60 Validation is often perceived as applied science by scientists.

5.3.2.1 The animal model is the technological paradigm

As mentioned, the regime level refers to the dominant practices consisting of a semi-coherent set of rules. Giovanni Dosi in this context refers to the technological paradigm. Dosi's paradigm is defined as a set of pieces of knowledge, both practical and theoretical, know-how, methods, procedures, physical devices, and equipment, as well as experience of successes and failures. It includes *"the 'perception' of a limited set of possible technological alternatives and of notional future developments"* (Dosi, 1982, p.152). Dosi states that technological paradigms embody strong decrees of technological changes that need to be followed or neglected. As such, technological paradigms have a powerful exclusion effect. When a technological trajectory[61] is powerful, it might be especially difficult to change from one trajectory to another (Dosi, 1982). Regulatory animal testing is such a powerful technological trajectory. Organizations around the world have been using animal models for many decades to prove the quality or safety of products. This deeply rooted experience with animal tests has provided these models with the status of 'gold standard'. Most regulators will only accept an alternative test if it will allow them to assess a compound or product in a way similar to the gold standard. As a result, animal-based regulatory testing remains, as researcher Thomas Hartung wrote, *"frozen in time, using and accepting the same old animal models again and again, often without stringent examination of their validity"* (Leist et al., 2008).

According to Dosi (1982) changing the technological paradigm means starting at the base of the problem solving phase. And that is exactly what the community of regulators often pleads for when referring to the need to first unravel the underlying biological mechanism before switching to a new testing model. The animal model incorporates this biological mechanism, even though the modus operandi remains a black box. Nevertheless, it is often felt to be essential for designing a scientifically sound 3R model to fully understand the underlying biological mechanism.

5.3.2.2 Small varieties, big consequences: the problem of diverse risk regulation regimes

Hood et al. (2001) refer to the meso level in terms of 'risk regulation regimes'. These regimes consist of the complex of institutional geography, rules, practices, and ideas associated with the regulation of a particular risk or hazard. Regulatory decision-making is a core activity in these regimes. Instead of defining one regime, the theory on risk regulation regimes emphasizes the existence of diversity in risk regulation, which is a result of the different pressures that lead to regulating the risk. The authors define three main shapers of regulatory content: 'the type of risk', 'public attitudes', and 'organized interests'. Hood et al. (2001) state that there is a relationship between public preferences and the regime content, meaning that the way in which society perceives different risks (see macro level) is reflected in the way these risks are regulated. This might, for example, explain why some risks are highly regulated and others are hardly regulated.[62]

61 Possible technological directions within the boundaries of a technology paradigm.

62 UK vehicle emissions, for example, are highly regulated, while smoking tends to be less heavily regulated although it is assumed to be a much bigger killer. And the regulation of pesticide residues in drinking water in the UK is highly risk-averse while the regulation of cancer risks of the emission of radon gas in houses shows a high level of risk tolerance (Hood et al., 2001).

On top of that, risk perception has a cultural element to it, which explains why different countries might regulate the same product in different ways.

These differences can also be observed in the area of safety and quality testing. The political and public demand for safety depends on cultural values leading to different expectations with regard to the levels of safety. The regulatory systems must accommodate these expectations (Richmond, 2002). As a result, every regulatory level developed requirements to deal with these demands, and in the past this frequently happened in a fairly isolated manner. This has led to situations in which requirements for one product differ from one region to the next. In the case of rabies vaccines, for example, it has resulted in multiple varieties of the procedure for potency testing of inactivated vaccines between the different regulatory levels. [63] [64]

Even if the regulatory requirements are the same, as is the case for most pharmaceuticals and biologicals (e.g., vaccines) within Europe, regulatory variability can occur for those products that follow the decentral route, which means that they are assessed at the level of individual Member States. Countries can then make use of the discretionary space that is offered by the regulatory requirements. The discretion may well lead to divergent ways of interpretation and implementation of the requirements by regulatory agencies (Bakker and van Waarden, 1999).

Dissimilar regulatory requirements and a diverse interpretation of requirements exact a heavy toll on innovations like 3R models and their acceptance. Manufacturers increasingly operate in international markets and face the difficult task of complying with this diversity. The use of 3R models by manufacturers for regulatory purposes depends to a large extent on the level of regulatory acceptance achieved. In case of diversity, manufacturers will either anticipate the strictest set of requirements or will even execute all the different tests requested by the different regulatory regimes. One manufacturer of biologicals pointed out that it is not uncommon for them to conduct five or even more different test protocols for the same product to comply with all these different regulations, with all due effects in terms of time, costs, and numbers of animals used.

Regulatory acceptance of a 3R model at one specific geographical level is therefore insufficient. Industries will hang on to the conventional animal model as long as they still have to conduct these tests for one region. The most favorable situation is regulatory acceptance of a 3R method at the highest possible geographical level. As a result, industry is lobbying very actively for harmonization of legislation.

At the same time, harmonization is a very lengthy and difficult process. To give an example: to change an OECD test protocol, all 30 Member States must agree to the alterations. This consensus approach means that rapid and dramatic alterations in the recommended

63 i.e., the European Pharmacopoeia, the World Health Organization, the World Organization for Animal Health (OIE), and the US Code of Federal Regulations.

64 These varieties concern, among other things, the number of tests required, the number of mice to be used, the number of dilutions to be administered, and the different criteria for evaluation of the test, i.e., death or signs of rabies and the survival rate (Bruckner et al., 2003).

OECD policy are unlikely to occur (Rudacille, 1999). *"At least for regulatory toxicity testing, the global frame and network are given by institutions such as OECD, ICH, and alike. Due to the necessity of global consent of states, organizations, and stakeholders, the time gap between availability of a novel alternative test method and its acceptance by authorities and implementation thereafter is widening."* (Garthoff, 2005)

Furthermore, harmonization is said to suffer from the 'Not Invented Here syndrome', which means that the parties involved are willing to harmonize as long as their own criteria are accepted as the standard. And finally, frontrunners – industry, regulatory authorities, or academia – that are already using high standards in the area of the 3Rs fear that, by harmonizing the requirements internationally, standards might be lowered to the common denominator (Busfield, 2006).

All in all, harmonization is considered to be very important for the acceptance and use of 3R models, but is at the same time a very difficult process, dominated by cultural differences, psychological barriers, and competition, both scientific and economic. Nevertheless, harmonization is high on the political agenda, and efforts such as the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) have already booked their first successes (Majone, 2010).

5.3.2.3 Informational asymmetry between regulators and industry

Another factor at the regime level is the information exchange between regulator and regulatee (regulated industry). The regulatee often has an information advantage, having in depth knowledge of the technical aspects regarding the product and the variety of relevant legal requirements (Abraham, 1995; De Bruijn and Koopmans, 2005; Dupree et al., 2007). Regulators may often have less 3R knowledge in comparison to industry, since industry already often uses these methods in the R&D phase or in the production process. On top of that, a 3R method is sometimes custom-made for a particular production process, meaning that only the manufacturer of that product is acquainted with the specificities of the model used. Consequently, regulatory authorities face difficulties judging the alternative method on its merits, leaving them to a certain extent dependent on the information provided by the manufacturer. This phenomenon is also referred to as 'informational asymmetry' (Heritier, 2001).

The area of risk assessment of products is very complex, and technical expertise is a crucial factor in the decision-making process of whether or not to implement a 3R method in safety and efficacy testing. The informational asymmetry therefore makes regulators cautious in adopting test models they are not completely familiar with.

On the other hand, regulatees depend on decisions made by regulators regarding their product. Manufacturers often feel they are being left in the dark regarding the precise criteria regulators will use in judging whether a 3R model will be accepted for safety or quality testing purposes. This might be the result of limited communication between regulators and developers of alternative methods, which leads to the development of 3R models that fail to take regulatory needs sufficiently into account. To solve this problem,

it is often recommended that regulators should be involved at the various stages of the validation process of a 3R model to discuss the regulatory criteria the model has to meet (EPAA, 2007, Bottini et al., 2008). All in all, the regulator-regulatee interaction is one of close interdependence (Dupree et al., 2007).

A consequence of the informational asymmetry and limited communication is that regulators lean towards relying on their existing knowledge and on the level of scientific consensus concerning animal experiments and alternatives. Without the scientific backing it is a precarious decision to incorporate an alternative into the product assessment procedure. This process of reaching scientific consensus, by the very nature of scientific methodologies, is difficult to achieve and takes a long time. And so are the changes in favor of 3R models.

An important development in this respect is the increasing number of efforts to erect forums where stakeholders from both regulatory authorities and industry can discuss in a neutral setting the pros and cons of the available models – both 3R and animal models. This type of interaction and communication is important to pave the way towards regulatory acceptance and use of 3R models.

A possible way of dealing with the informational asymmetry is by sharing research data. As one respondent put it: *"There are databases full of information, but these are not accessible because the industry owns them. A lot of the data concerning newly developed methods stay within the walls of the company."* Sharing data can help regulators build up experience with and thereby gain trust in the 3R models used by the regulatees. A further step would be to share data with other manufacturers. This could have a major effect on reducing the number of duplicated tests.

Data sharing is already considered one of the core principles in the REACH Regulation[65], and it allows companies to reduce costs and avoid unnecessary testing on vertebrate animals. To meet this requirement the chemical industry has made a start in setting up a network to share knowledge and data. It must be noted, however, that in practice this element needs further improvement, since registrants repeatedly fail to consider their obligations for sharing data or do not come to an agreement with other potential registrants on the sharing of these data (ECHA, 2010).

5.3.2.4 Transition costs

Decisions made by manufacturers are the product of continuous cost-benefit analyses to weigh what the costs or profit of an investment/innovation will be. Profit here does not solely refer to economic profit, but might also mean reputational or scientific profit. MacLachlan (1994) argues that product safety is very important to industry due to the fact that this "*responsible behaviour is vital for continued business success."*

Olson stipulates the importance of the speed and the costs of the process from R&D to market approval. The average time required, for instance, for the pharmaceutical industry

65 European Community Regulation on chemicals which deals with the Registration, Evaluation, Authorization and Restriction of Chemical substances. The new law entered into force on June 1, 2007.

to develop, test, and gain approval of a new prescription drug in the US is about 12 to 15 years (Olson, 1997; MacLachlan, 1994). Estimates about the cost of developing a new drug vary widely, from $ 800 million to nearly $ 2 billion per drug (DiMasi et al., 2003).[66] After the development and thorough testing of the product, the Food and Drug Administration generally takes another one to four years to review the application and grant market approval. The length of this regulatory procedure erodes the patent protection and is said to be an important disincentive to innovation, since every change in the production process requires a new regulatory approval (Jaffe, 1994). In short, speed of research, development, and registration is a very important factor for industries such as pharmaceutical companies. And every innovation that can speed up this process will be embraced, whereas every innovation that slows it down will be discarded.

Time and costs are therefore important arguments for the industry in the choice of a testing model. If industry foresees any economic or regulatory hurdle in using a new model it will most likely stick to the conventional method. But the cost aspect can also work in favor of 3R methods, since these methods are often quicker and less expensive than the methods they supersede (Richmond, 2002). In the best scenario, cost efficiency and reducing animal testing converge. In this respect the large-scale industrial lobby to influence REACH is frequently referred to. Here the industry's lobby gave counterweight to the call for more testing, since the chemical industry has no interest in regulations that further increase the number of required tests.

According to several respondents, another reason for stakeholders to adhere to existing test methodology is to protect their 'return on investment'. This can make research laboratories within different institutions reluctant to disrupt the existing testing infrastructure, which often still relies on animal models. And finally, respondents pointed out that research departments of regulatory authorities, industry, and academia do not want to run the risk of losing their existing knowledge of/and experience with animal testing and thereby fall behind or even become dependent of others in the field. Therefore, they anxiously hold on to existing practices.

In short, several economic barriers can be identified that are perceived to influence the acceptance and use of 3R models. However, it can be questioned whether these motives play a decisive role in the actual process of acceptance and use of 3R models for regulatory purposes. Vermeulen (2011) argues that innovations that are interesting from an economic perspective also face difficulties in breaking through due to the existing interests and the strong convictions of the stakeholders within a sector.

5.3.3 Developments at the macro level of the sociotechnical landscape

The sociotechnical landscape relates to material and immaterial elements at the macro level, such as the political culture, social values, world views, the macro-economy, demography, and the natural environment (Kemp, 2010). The sociotechnical regimes and niches are both influenced by developments at the macro level. The landscape is the hardest element of the three to change (Geels, 2002).

66 It must be noted that these high estimations are criticized by some as being in the interest of industry to keep its estimations as high as possible (Light and Warburton, 2011).

Three societal developments are distinguished here that influence the transitions towards the regulatory acceptance of 3R models, i.e., the risk society (see 5.3.3.1), the concern for animal welfare (see 5.3.3.2), and the culture of litigation (see 5.3.3.3).

5.3.3.1 The risk society…

A first and very important influence on regulatory acceptance and use of 3R models at the landscape level is this striving of modern society for risk minimization, with the precautionary principle as leitmotiv. The precautionary principle, which recommends to *"err on the side of preservation"* (Barrieu and Sinclair-Desgagne, 2003), has clear consequences for the way in which new technologies/products are adopted.

Many innovations offer high scientific and societal potential on the one hand and scientific uncertainties and health and welfare concerns on the other. The society that has to cope with such technologies is also referred to as the 'risk society'. This concept was introduced into sociology by Ulrich Beck (1992) and was later adopted by sociologist Anthony Giddens. The latter has described the risk society as follows: *"…a society where we increasingly live on a high technological frontier which no one completely understands… It is a society that is increasingly preoccupied with the future and with safety, which generates the notion of risk"* (Giddens, 1999, p.3). The problem of 'manufactured risks' caused by new products and technologies (Giddens, 1999), is that society has relatively little experience with them and thereby little knowledge of the actual risks they pose (Giddens, 1999). As a result, the risks have to be assessed by experts.

Societies' response to the unpredictability of manufactured risks is to try to prevent, minimize, and channel them, for example, by delegating this task to regulatory authorities charged with controlling possible negative side-effects of industrial activity (Malyshev, 2006). As a result, regulators have to deal with a great number of responsibilities on the one hand and uncertainties on the other. The regulators' reaction to this thorny combination is likely to be one of sticking to the routines they are familiar with. Or as Breyer put it: *"Rules – or procedures – become frozen in place and cannot readily adapt to changing scientific knowledge"*. (Breyer, 1993, p.49)

According to Breyer, the public perception of a certain risk influences the politician's action and subsequently the regulatory reaction to it (Breyer, 1993). Both the public's perception and the politician's response influence the regulator's decisions in dealing with certain risks, despite the fact that the public and politicians are unlikely to understand the complexity of the matter. The higher the presumed risk of a new technique or product, the more stringently it will be regulated, even though the perceived risk does not always correspond with the actual hazard.

Risk perception is related to several factors, such as dread, controllability, voluntariness, and observability, with the dread factor as the most influential one. The higher the dread factor of a product, the higher its perceived risk is, and the louder the call for strict regulations will be (Slovic et al., 1984). Pharmaceuticals and vaccines, for example, are categorized as belonging to the class of 'unavoidable, unsafe products', which offer

desired benefits but are not without risk (Jaffe, 1994). A certain level of risk is accepted by the public when it comes to pharmaceuticals, but this level is much lower for vaccines, which are administrated to young and healthy children.[67] Industrial chemicals are another product group relevant in terms of regulatory testing. The accepted risk for this group of products is, as a result of the high dread factor and the involuntariness of being exposed to these compounds, close to zero. This has had its effect on the regulatory requirements for different groups of chemicals that are very strict and aim for zero or negligible risk levels (Kasamatsu and Kohda, 2006).

The consequence of this societal priority of risk minimization is that the use of animals for safety and efficacy testing of new products has increased significantly over the past forty years. However, the focus on risk avoidance not only increases the number of animals used, it also is detrimental to the acceptance of alternative test models. In response to society's risk aversion, alternative methods often are not accepted by regulators. The alternative methods must be proven three times over before they are perceived to be as valid, sensitive, and specific as conventional methods. As has been discussed before, these are criteria to which conventional methods themselves often do not adhere.

5.3.3.2 …versus the concern for animal welfare

In Western society yet another development can be observed, i.e. the growing concern about the welfare of animals and the potential for animal suffering in product testing. Within Europe this concern has been translated into a legislative act, Directive 2010/63/EU, for the protection of laboratory animals for scientific purposes (EU, 2010). This horizontal legislation which states that alternative models should be used wherever possible, must be taken into account by vertical or sectorial product legislation (see Chapter 2).[68] Some vertical legislation already explicitly refers to this horizontal legislation, but a lot can still be gained in this area. For instance, there generally is little interaction between the EU committees drafting this vertical legislation and those that develop the horizontal animal welfare legislation (De Leeuw, 2004; Schiffelers et al., 2007). The result is that sectorial legislation, when revised, continues to require animal tests, even after validated alternatives have become available. Furthermore, the growing concerns about animal welfare do not outweigh the concerns related to protecting human health.[69]

5.3.3.3 Culture of litigation

Together with the development of risk minimization, the culture of litigation is gaining terrain in contemporary society. Regulatory authorities face high demands for consumer safety and they are expected to take these into account when implementing policies. This means that they bear a heavy responsibility and as a result they are particularly susceptible to a negative sense of responsibility. Or as one representative from a Dutch regulatory authority put it: *"If anything goes wrong, we will be held accountable."*

67 Due to the complex production process, vaccine lots can vary in quality and consequently in safety and efficacy.

68 Horizontal legislation pertains to animal experimentation and multilateral agreements in general. Vertical or sectorial legislation regulates the activities of a particular sector (see Chapter 2 and Appendix I).

69 An opinion poll in the Netherlands underlines this: two-thirds of the population is of the opinion that animal tests for medical purposes are acceptable (Intomart GfK, 2004).

Every change in the way substances are assessed is to some extent risky. No one can guarantee that a change can be implemented without compromising the quality of the assessment procedure of products (De Leeuw, 2004). A change in regulation and its implementation is therefore often seen as a potential liability. As a result, both industry and regulators often take a fairly passive approach to innovating product registration and release procedures. This point is illustrated by the comment of a civil servant of the European Commission (Schiffelers et al., 2005, p.37): *"It's better not to change ten times, than to make nine changes for the better and one for the worse."*

As Michael Power puts it in his inaugural speech: *"An age of 'new risk management' has dawned in corporate governance, sparked by high-profile business failures and accidents."* (Power, 1999: as cited by Hood, 2002, p.15) The potential threat of such incidents as the often described thalidomide disaster in the 1960's continuously urges manufacturers and regulators to be very cautious in the decisions they make (Olson, 1997; Carpenter, 2010). This defensive risk management is in line with what government and regulators are there for, namely protecting individuals from 'suffering'. But a strong focus on avoiding blame and liability is likely to also lead to undesirable effects (Hood, 2002). *"…The concern for blame prevention seems to be leading to protocolization and risk assessment inflation to establish procedural alibis as a form of bureaucratic insurance."* (Hood et al., 2001: 179) Politicians and regulators mainly concerned with the avoidance of (political) blame over hazard and safety might end up hardly changing anything (Hood, 2002), even if the current situation is far from optimal.[70]

This protocolization is also observed in the area of regulatory testing. Here the fear of litigation leads to adhering to trusted methods and a rigid interpretation of test protocols. This is also referred to as the 'tick box approach', a strict way of holding on to every test described in the protocols without taking a critical look at the relevance of the tests, the necessity of conducting all these tests, or the possibility of using an alternative testing model. It should be noted here, however, that the level to which this tick box approach is applied differs per product group and regulatory authority. Some product sectors, such as pesticides, are notorious for their tick box approach[71], whereas the areas of pharmaceuticals and biologicals are said to be more flexible when it comes to the interpretation of the test guidelines. But in the areas where there is more discretionary space to choose the method best suitable for the job, the existing technologies also give the highest assurance. Decisions then are based mainly on custom and practice, rather than on an informed science-driven selection of the method most likely to provide the most relevant result (Richmond, 2002). And for fear that authorities might reject certain results, regulatory affairs departments of industry are said to take pre-emptive action by anticipating the most strict registration requirements. It will be clear that this kind of risk avoidance obstructs the risk taking that is intrinsic to innovation and thereby deters the development and acceptance of new technologies, such as 3R models.

70 The phenomenon of injury litigation, for example, has become a major risk in the US and has had a chilling effect on innovation in many American industries (MacLachlan, 1994).

71 http://www.hslf.org/epa-animal-testing/ : consulted December 2011

5.4 Creating a breakthrough: towards critical junctures

As mentioned Geels states that *"radically new technologies have a hard time breaking through"* (Geels, 2002, p.1258). This is precisely what can be observed in the field of regulatory acceptance and use of 3R models where several models have been available for decades but still haven't been able to become part of the existing regulatory regime. However, history has proven that very firm configurations also can change (Geels, 2002). For this purpose, it is important to comprehend which variables influence the process of acceptance and use and in what ways.

In the previous section, a wide range of variables was defined that are considered to influence the transformation from the existing test regimes to the acceptance and use of 3R models. Some of these variables are perceived mainly to withhold, whereas others are considered to drive the process of acceptance and use. These variables and the nature of their influence are summarized in the '3R Acceptance Model' (see Figure 10).

For this model an adjusted version of the fishbone diagram of Ishikawa is used (Nathans, 1997). This diagram offers the opportunity to get to the roots of a problem that has many possible causes. The original model is adjusted in two ways. First, the model presented here not only presents the different influencing variables, it also distinguishes which are considered to be drivers and which barriers. Furthermore, a distinction is made between the niche (micro-), regime (meso-), and landscape (macro) level, as used in the multilevel perspective. At the niche level factors concerning people and methods can be found. The regime level consists of factors connected to products, organizations, institutions, and regulations. The landscape level covers the broader societal factors.

This model offers a tool to understanding the complex reality in which the acceptance and use of 3R models takes place. To stimulate a breakthrough it is important to make an additional distinction between the more pliable and the rigid factors (Ellemers, 1976).

As described, the process of technology acceptance is determined by the influence of variables at the niche, the regime and the landscape level. In general, it can be stated that the higher the level at which a variable is situated, the stronger this variable is, i.e., the more difficult it is to manipulate the variable. A variable such as the risk society exerts a substantial influence on the acceptance of 3R models, but is at the same time a factor that is very hard to counter. The more pliable factors, as a rule, can be found at the lower levels, i.e., the niche and partly at the regime level. Variables like education and training, communication and dissemination of successes, and facilitating and stimulating frontrunners are all at these lower levels and are very important ways of stimulating a breakthrough. As mentioned, the variables at the landscape and partly at the regime level are considered to be the stronger variables. However, this does not mean that change cannot occur or be stimulated at these levels, for example, by keeping animal welfare and the need for harmonization high on the political agenda.

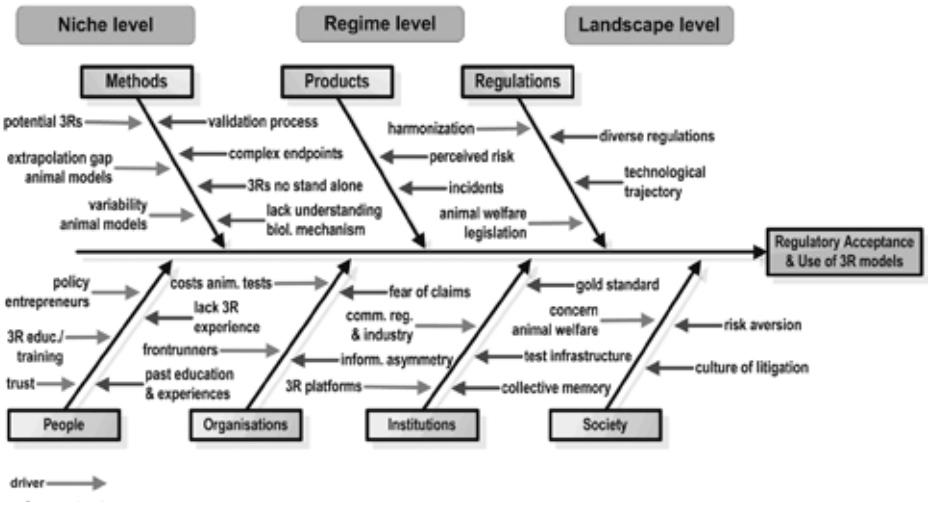


Figure 10. The 3R acceptance model

Technology acceptance is a process of shifting congregations or *"reweaving the elements",* as Geels puts it (Geels, 2002: 1259). Changes in one element can elicit changes in other elements, and alignment of these different elements can create a shift in the sociotechnical regime. This occurs, for example, if a development in a niche falls together with a change or request for change in the regime and within the sociotechnical landscape, creating a potential breakthrough or a 'critical juncture'. Critical junctures occur when existing institutions are challenged or become instable (Krapohl, 2008), and they can be of endogenous or exogenous nature (Bakker, 2001). Endogenous critical junctures start from within the sociotechnical regime and can arise if the existing routines fail to meet the expectations, for example, animal models that show high variability in test results and difficulties in terms of the extrapolation from the animal results to human beings. (This mechanism where inconsistencies arise within the regime is referred to as 'stress' in Technology Transition literature). Exogenous critical junctures are the product of developments outside the regime. These might be due to shifts in the landscape, for example, a growing concern about animal welfare, or by the linkage of several new technological developments, i.e., niche accumulation at the micro level, which can occur if a combination of 3R models proves to be a solid answer to the problems in the existing regime. (In Chapter 4 two exogenous mechanisms are described, i.e. 'pressure' from new technologies which make aspects of the regime obsolete and 'tensions' arriving from broader societal developments that as for change).

According to the theory on technology transitions, bottlenecks for new technologies can be solved more easily if they are linked with existing technologies, starting a symbiotic relation. This means that technology transition most often is an incremental process in which new regimes gradually grow out of old ones (Geels, 2002). This means that 3R models that build upon are expected to face fewer difficulties in terms of acceptance compared to models that fully replace the animal model.

According to Dosi the breakthrough of technological innovations also can be stimulated by risk taking actors that are ready to try different solutions (Dosi, 1982): an entrepreneur with so called 'Schumpeterian' features.[72] This can be an individual, an organization or a coalition of involved stakeholders that is willing to take a certain risk, keep the discussion alive, and keep the topic of the 3Rs high on the agenda (see section 4.3.2). These policy entrepreneurs, as Kingdon calls them, are skilled at coupling problems, solutions, and policies, and they can thereby respond rightly to critical junctures or policy windows in Kingdon's terminology (Kingdon, 1995, Schiffelers et al., 2005). A comparable concept is that of the boundary spanner who is characterized by his ability to connect and mediate between the parties and their different interests and who knows how to cross cultural/ organizational borders (Williams, 2002) (see section 4.3.2 for other similar concepts). These entrepreneurs facilitate mutual communication and understanding and can come from regulatory authorities, industry, academia, and NGO's. Ideally, they are trusted and fairly neutral parties with a high level of knowledge of the specific problems. They can facilitate stakeholders in recognizing and taking a proactive approach towards potential critical junctures. Alignments towards critical junctures are already taking place in some cases, in the sense that there is increasing agreement on the flaws of certain animal models and the scientific potential of the 3R models that could replace them. In these cases several boundary spanners or policy entrepreneurs can be found who put a lot of energy into bringing the parties together to take the discussion a step further.

5.5 Conclusions and recommendations

We are living in a risk-averse society, which means that our striving for risk minimization is a fundamental element of our society. A common response to the intrinsic uncertainties of new technologies is to try to minimize their potential negative side effects. This is operationalized primarily by setting up risk regimes of regulatory authorities, rules, and regulations to minimize the possible adverse effects of products like pharmaceuticals and chemicals. In terms of technology transitions, the risk-averse society is a very important feature of the landscape in which new technologies like 3R models are being developed. This risk aversion strongly influences the way stakeholders within the regulatory regime, like regulatory authorities and manufacturers of products, look upon new technologies such as 3R models. The risk aversion is amplified by public and political pressure, incidents, the culture of litigation, and the informational asymmetry between regulators and regulatees.

The sociotechnical regime of product regulation still largely depends on animal models, meaning that the current knowledge, the research infrastructure, and risk assessment practices are dominated by regulatory animal testing. However, the existing regimes face increasing pressure from social groups that ask for safe products and the reduction of animal testing simultaneously and on the other hand from niches developing 3R methods that challenge the conventional ways of testing. In this paper, an overview of barriers and

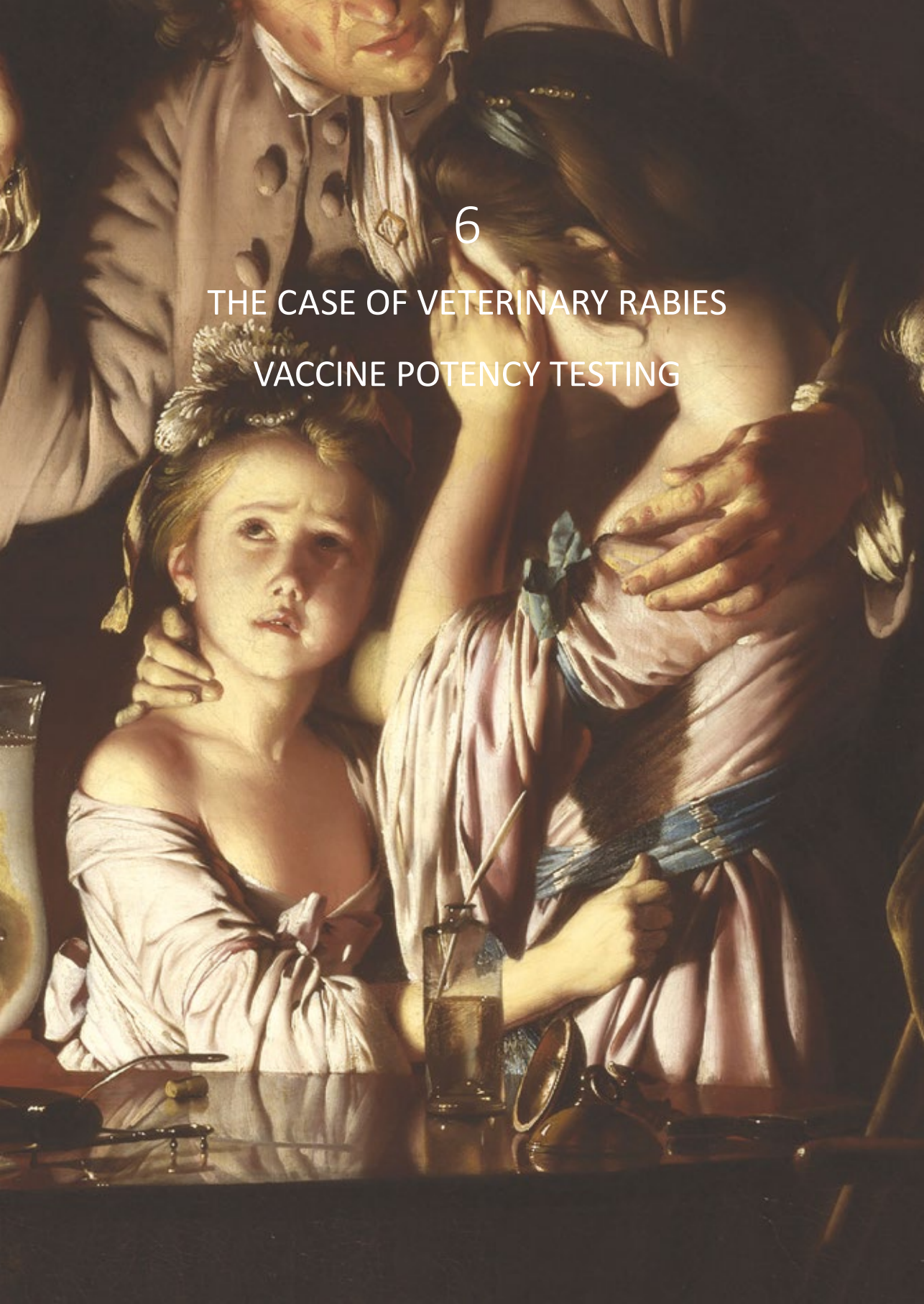
72 Joseph Alois Schumpeter was an Austrian-American economist who was probably the first to look at the important role of entrepreneurs. Schumpeter argued that the innovation and technological change of a nation comes from the entrepreneurs, or 'wild spirits'.

drivers has been presented using the multilevel perspective on technology transitions. The wide variety of drivers and barriers in the process of regulatory acceptance and use of 3R models[73] reflects the complexity of the matter. Even more so, because the combination of factors might differ per product sector and sometimes even per product.

3R methods remain relatively new compared to the standard testing routines and practices. As a result, stakeholders have less experience with them, leading to a lack of trust as to whether they can offer levels of safety comparable to the animal model. As long as 3R models suffer from this lack of trust they will have a hard time breaking through. The inertia 3R models have been confronted with ever since their introduction by Russell and Burch in 1959 is a phenomenon that can be observed in technology transitions in general and is a result of the deeply rooted collective memory of the stakeholders. By definition, innovation is uncertain and controversial until it is accepted as the norm.

Innovation starts with the willingness to accept failure (MacLachlan, 1994). Regulatory authorities and industries acknowledge that there is a lack of trust and express the need to take a 'leap of faith' in those cases where 3R methods have been thoroughly tested and validated but are still not accepted. In the area of product regulation, however, failure can have big consequences. Thus the inertia in the situation of 3R models is aggravated by the context in which these models are used, i.e., to guarantee the safety and quality of products that are looked upon as a risk to human health or the environment. For this reason, it must also be accepted that such institutional changes take their time and regulatory acceptance of 3R methods is most likely to occur as an incremental process, i.e., no change in terms of radical developments but new test regimes that gradually grow out of old ones (Geels, 2002). Many respondents even warn of discarding the animal model at too early a stage. They indicate that a 'stand-alone' situation, of either *in vivo* or *in vitro* methods, is in most situations neither feasible nor desirable (Schiffelers et al., 2007: see Appendix I). This is fully in line with Vermeulen, who states that it's not only impossible to eradicate the old institutions, it is also undesirable (Vermeulen, 2011). A well-considered combination of both types of testing, therefore, is believed to be the best feasible scenario. It is important, then, not to think in revolutions but rather in terms of evolutions. Evolutions require thinking in terms of small but on-going steps. Only intense and continuous forms of communication, dissemination, and education can help to overcome the inertia that had already been observed by Russell and Burch in 1959 (see quotation in the introduction). This means, for example, that an exhaustive approach to communication is required between stakeholders, such as regulatory authorities, industry, and academia about 3R developments and the chances they offer for regulatory testing. More specifically, communication between regulatory authorities and manufacturers should be intensified to level off the informational asymmetry between these parties. And sharing test data will help regulatory authorities to build up experience with the specific 3R models and will facilitate the process of building new experiences, rules, practices, and routines and thereby slowly change the existing institutions. And in the end, such a multitude of relatively small steps can lead to a landslide in favor of the 3Rs.

73 As mentioned earlier in this paper, regulatory use is seen as a function of regulatory acceptance, in which the level of acceptance highly determines the level of regulatory use. Without regulatory acceptance, regulatory use will only occur sporadically.



6

THE CASE OF VETERINARY RABIES VACCINE POTENCY TESTING

"I would wish to remind you all, whether you are sitting on the side of industry or on the side of regulators that, from my personal viewpoint, especially in the field of veterinary medicine, we are suffering from over-conservatism. This is really something which we all have to overcome." [74]

Jean-Marc Spieser [75]

74 Introductory speech at the Workshop of the Paul-Ehrlich-Institut (PEI): Potency testing of veterinary vaccines: the way from *in vivo* to *in vitro*, December 2010, Langen, Germany (Jungbäck, 2012).

75 Late Head of Department of Biological Standardization, OMCL Network European Directorate for the Quality of Medicines & HealthCare (EDQM), Council of Europe (CoE), Strasbourg

A short guide to Chapter 6

Chapter 6 consists of a research article which was published in 2015 in *Altex*, containing the result of a case study on the acceptance and use of the mouse antibody serum neutralization test (SNT) to replace the NIH mouse challenge test for potency testing purposes of inactivated veterinary rabies vaccines. The case study was conducted between 2011 and 2013. In anticipation of this case study a survey was done to test a series of drivers and barriers and optimizing options (Schiffelers et al, 2014a: Appendix II). The survey results served as input to the SNT case study. Through this case study the research questions Q3a and Q3b are directed (see section 1.2). For the sake of the completeness of this publication a certain amount of replication of chapters 1 to 4 was necessary.

Abstract

In April 2013 the mouse antibody serum neutralization test (SNT) was formally incorporated into European Pharmacopoeia monograph 0451 for potency testing of inactivated veterinary rabies vaccines. The SNT is designed to replace the highly variable and pain and distress causing NIH mouse rabies challenge assay. The adoption of the SNT meets the European ambition (i.e. EC and CoE) to replace, reduce and/or refine laboratory animal testing. However, regulatory acceptance and use of 3R models, such as the SNT, remains challenging. This paper aims at clarifying the process of acceptance and use of the SNT. For this purpose it reconstructs the process and reveals barriers and drivers that have been observed by involved stakeholders to have played a role. In addition it extracts lessons to stimulate regulatory acceptance in similar future processes. The incorporation of the SNT into the monographs went relatively quickly due to a thorough test development and pre-validation phase, commitment and cooperation of relevant stakeholders and a strong project coordination of the international validation study. The test was developed by the Paul Ehrlich Institut, a leading European OMCL. This facilitated its European regulatory use. The use by industry is in a critical phase. At this stage, product specific validation and the question whether the SNT will be accepted outside Europe are important influencing factors.

Reference:[76]
Schiffelers, M.J.W.A., Blaauboer, B.J, Bakker, W.E., Hendriksen C.F.M. (2015). Regulatory acceptance & use of serology for inactivated veterinary rabies vaccines. *ALTEX* 32(3)

76 Acknowledgments: We thank the Doerenkamp-Zbinden Foundation for funding this project. We thank the respondents for sharing their expertise and perspectives with us.

6.1 Introduction

Rabies – Latin for madness – is one of the oldest diseases known to mankind. It is an acute and deadly disease caused by a viral infection of the central nervous system. Once a patient develops symptoms of rabies, no treatment is available. More than 60.000 deaths are reported annually from rabies with more human deaths occurring in Asia than anywhere else in the world (30.000 deaths per annum).[77] Rabies vaccines, both human and veterinary, are crucial in fighting the disease. Every year more than 20 million people are vaccinated against rabies after being bitten. 40% of them are under the age of 15.[78]

The first rabies vaccine was developed in 1885 by Louis Pasteur and Emile Roux and was originally harvested from infected rabbit nerve tissue and inactivated subsequently. In the 1960's concentrated and purified cell-culture and embryonated egg based inactivated rabies vaccines were developed, which were much more consistent and safer in batch quality compared to the previously produced vaccine batches. Nonetheless, these inactivated rabies vaccines still derive from living organisms which may lead to variations, e.g., in the antigen amount or antigen conformation of the final product. Therefore, each vaccine batch is subjected to mandatory quality control testing. This includes safety testing[79] and potency testing.[80]

The standard protocol for rabies vaccine potency testing is the NIH mouse rabies challenge assay, which was introduced almost 60 years ago and has remained largely unchanged ever since. This assay raises serious ethical concern in terms of animal welfare. Each final batch needs potency testing (Casey et al., 2011, Krämer et al., 2009), which means infecting a group of mice with the rabies virus and immunizing half of the group with the rabies vaccine. The other 50% of these animals show signs of rabies, leading to severe suffering and death (Bruckner et al., 2003). Although the test was never officially validated the NIH mouse rabies challenge assay – from now on referred to as the NIH test – has been used by regulatory authorities for over 50 years and has thereby gained its status of gold standard. However, it has a number of serious drawbacks. First, the test parameters differ from the natural situation, e.g., the intracranial challenge route does not reflect the natural route of infection nor does the intraperitoneal vaccination reflect the normal route of immunization (Romberg et al., 2012, Wunderli et al., 2006), and test results show high variability of up to 400% (Bruckner et al., 2003, Krämer et al., 2009).

77 http://www.who.int/rabies/about/en/ :consulted on March 25, 2015

78 http://www.who.int/features/2012/world_rabies_day/en/ :consulted on March 25, 2015

79 Safety tests are designed to detect any material or property that may be harmful to the recipient, such as bacterial contamination, infectious virus or toxicity. In the case of viral vaccines in general, and especially of rabies vaccines, the specific problem of residual virulent virus is of the utmost importance. http://whqlibdoc.who.int/monograph/WHO_MONO_23_(3ed)_(part5).pdf :consulted on March 25, 2015

80 Potency is the capacity of a vaccine to protect the vaccinee against the virus, i.e., rabies. http://www.oie.int/doc/ged/D8314.PDF :consulted on March 25, 2015

These drawbacks of the NIH test have made rabies vaccine potency testing a high priority in terms of the 3Rs, i.e., models to replace, reduce or refine the conventional animal model (Russell and Burch, 1959).

Over the last decades several assays have been developed with the goal to replace, reduce or refine the NIH test (see also Schiffelers et al., 2014a: Appendix II) of which the mouse serum neutralization test (SNT) of the German Official Medicines Control Laboratory (OMCL), i.e., the Paul Ehrlich Institut (PEI) (Krämer et al., 2009, 2010), is the most recent. This serological assay was designed with the purpose to replace the NIH test for inactivated rabies vaccines for veterinary use. It was validated through an international collaborative study (Krämer et al., 2010) and became part of monograph 0451 for inactivated veterinary rabies vaccines of the European Pharmacopoeia (Ph. Eur.) in March 2012 (European Pharmacopoeia, 2013). Its regulatory acceptance was thereby largely accomplished within the European context. However, its broader international regulatory acceptance and use requires additional steps.

The SNT is in line with the ambition of the European Commission, laid down in Directive 2010/63/EU on the protection of animals used for scientific purposes, to diminish the use of laboratory animals and stimulate the acceptance and use of 3R models to replace, reduce and refine existing animal models. Directive 2010/63/EU states in article 13.2 that *"…in choosing between procedures, those which use the minimum number of animals shall be selected"* (EU, 2010). Furthermore, *"The Member States of the Council of Europe have decided that it is their aim to protect live animals used for experimental and other scientific purposes to ensure that any possible pain, suffering, distress or lasting harm inflicted as a consequence of procedures being conducted upon them, shall be kept at a minimum."*[81] To reach these goals full regulatory acceptance of 3R models is important. This means that they have to pass the following three subsequent substages, i.e.: Formal incorporation into regulatory requirements (FI); Actual regulatory acceptance by regulatory authorities (ARA); Use by industry for regulatory purposes (UI). This is often an arduous process in which many factors are seen to play a role.

In order to comprehend and augment current processes of regulatory acceptance and use and to facilitate similar imminent processes, it is important to reconstruct and analyze existing trajectories such as the SNT case.

To examine the acceptance process this chapter addresses the following key questions:

- *Which factors influence the acceptance and use of the SNT to replace the NIH test?*
- *Which additional steps are needed to replace the NIH test by the SNT or other 3R options?*
- *Which lessons can be drawn from the process of FI, ARA and UI of the SNT to stimulate regulatory acceptance for similar future processes?*

The findings derive from literature research, a series of interviews with experts in the field of rabies vaccine testing from standardization bodies, regulatory authorities (e.g., control

81 http://conventions.coe.int/Treaty/en/Treaties/Html/123-A.htm :consulted on March 25, 2015

laboratories) and industry, and international workshops that were attended by the corresponding author (see Appendix VIII for a full description of the research approach). The multilevel perspective on technology transitions (see Section 6.4) is used to capture the interrelatedness of the factors influencing the regulatory acceptance and use of the SNT.

This chapter elaborates on earlier work of the authors (Schiffelers et al., 2012: Chapter 5, Schiffelers, 2014a: Apendix II) and offers supplementary in-depth knowledge on the process of regulatory acceptance and use of 3R models through the reconstruction and analysis of the SNT case.

6.2 Process reconstruction

Below a description is given of the main developments regarding the SNT at the stages of FI, the ARA and the UI. However, we start by describing three pivotal pre-stages that have anticipated the process of FI of the SNT into the European monographs, knowingly: the stages of test development; the pre-validation and the international validation (Milne and Buchheit, 2012).

6.2.1 Pre stage I: Test development

The serological assay RFFIT (rapid fluorescent focus inhibition test), which forms the basis for the SNT, has been available for quite a while (Smith et al., 1996) and is described in the World Organization for Animal Health (OIE) Manual of Diagnostic Tests & Vaccines for Terrestrial Animals (OIE, 2012) as a standard approved technique for veterinary rabies vaccines (Krämer et al., 2009). Even though the RFFIT is faster, less painful and uses fewer animals compared to the NIH test, it has not been widely used and only little data exist concerning the comparability of the RFFIT with the NIH test. In the interest of the 3R principle to replace the NIH test and with the aim to develop an assay with better reproducibility of test results, the PEI developed the SNT serological assay, which is a modification of the RFFIT, as a refinement and reduction model (Krämer et al., 2009).[82]

The advantages of the SNT, when compared to the NIH test, are the reduction in animal use of up to 85% (Krämer et al., 2013), the reduced levels of stress and suffering of the animals involved, the better reproducibility of test results, the reduced amount of time needed to conduct this assay (less than 3 weeks instead of more than 4 weeks), and the fact that the assay is less labor-intensive and therefore less costly than the NIH test (Krämer et al., 2009, 2010). The SNT thereby offers manufacturers a higher speed of release of a vaccine batch. Furthermore, the data, evaluated by the PEI, show a good correlation between

82 *"The serological assay used, involves the immunisation of groups of 6 mice with approximately 1/5th the recommended dose volume of the test vaccine diluted appropriately, or of the reference standard vaccine preparation which is adjusted to the minimum potency allowed in the Ph. Eur. 14 days after immunisation blood samples are taken and the sera are tested individually for rabies antibody using the described virus neutralisation assay. Briefly, sera are titrated on 96 well microtitre plates and incubated for 1h with rabies virus. After adding BHK cells -baby hamster kidney cells- and incubating for 48h the presence of un-neutralised rabies virus is revealed by immunofluorescence. Dilutions of the sera that reduce the number of fluorescent cells by 50 per cent are calculated."* (Krämer et al., 2010).

the SNT and the NIH test (Krämer et al., 2009), which in itself is a remarkable outcome considering the variability of the NIH test.[83]

The test results were presented by the PEI for the first time at the meeting of the OMCL network in 2007. At that stage several OMCLs were skeptical about the test out of fear of overlooking sub-potent rabies vaccine batches. The PEI promised to provide the stakeholders with additional test results and they presented these results in May 2008 at the annual meeting of the European OMCLs involved in rabies vaccine batch release testing and at the Biological Standardization Steering Committee (SC) meeting in June 2008.[84] The latter presentation resulted in the initiative to start a collaborative study to confirm the transferability of the assay and its suitability for inactivated veterinary rabies vaccines on the European market.

6.2.2 Pre stage II: Pre-validation

In anticipation of the actual collaborative study, the decision was taken to start a small scale feasibility study to test the transferability of the serological assay to other labs. This transferability study was conducted by the PEI and the Swiss OMCL IVI (The Institute of Virology and Immunoprophylaxis). In this study 4 batches of rabies vaccines were tested using the serological assay and the NIH test in parallel. The conclusion of the pre-validation study was that the serological assay proved to be transferable to other laboratories (Krämer et al., 2010).

6.2.3 Pre stage III: International validation

Once shown to be generally applicable, an alternative method can be included into a specific Ph. Eur. monograph or into the general chapters of the Ph. Eur. However, before any candidate assay can be included, the validity of the method, in terms of its robustness and its global applicability, has to be demonstrated in a large scale collaborative study.

The aim is to demonstrate the wider transferability of the proposed assay and to confirm its suitability for the potency testing of inactivated rabies vaccines for veterinary use on the European market (Krämer et al., 2010). The less variation is observed in results of a relevant test, the more useable the test is.

In 2008 the Biological Standardization Program (BSP) of the European Directorate for the Quality of Medicines (EDQM) initiated and later on coordinated collaborative study BSP105 to broadly validate the SNT test as developed by the PEI.

---

83 This conclusion requires some additional wording in the context of the earlier comment that the NIH test in itself is highly variable. Experts increasingly criticize the way of thinking in terms of looking for correlation with a poor reference test. There is a tendency to move towards a concordance strategy allowing regulatory approval after a pass/fail correlation using potent and sub-potent batches (Stokes et al., 2012; Schiffelers et al., 2014a).

84 The SC consists of the Ph. Eur. group chairs (15, 15V, 6, 6B), the EMA, the WHO and co-opted experts. The SC determines the programme of the BSP and decides on priorities, new collaborative studies and the nomination of project leaders, in consultation with the stakeholders. http://www.edqm.eu/en/Biological-Standardisation-Programme-Committee-61.html :consulted September 2014

The Standard Operating Procedure and the reporting sheets were provided by the EDQM and vaccines were provided by several participants. The study involved 13 laboratories from 10 different countries – including Canada, the US and EU Member States. It included 8 official control laboratories of regulatory authorities and 5 manufacturer laboratories. All the laboratories were asked to test the potency of 4 different inactivated veterinary rabies vaccines – representing a range of products available on the EU market and produced by different manufacturers – using the SNT assay developed by the PEI.[85] The results were published in 2010 (Krämer et al., 2010) and were disseminated through presentations at various international congresses.

The collaborative study showed very comparable inter-laboratory results and a good comparison between the results of the serological assay and the NIH test. The sub-potent vaccine failed in both the NIH and serological test. It was therefore concluded that the SNT, as developed by the PEI, is not only a relevant assay but is also a reliable, i.e., reproducible, assay for potency testing of inactivated veterinary rabies vaccines (Krämer et al., 2010). This, combined with the advantages of saving a substantial number of mice, test time and costs, has made the SNT a serious alternative for rabies vaccine potency testing purposes.

It must be noted that the SNT is a single-dose serological assay and a semi-quantitative test that serves as a biomarker for vaccine potency. The result of the test is a pass/fail outcome, answering the core question: Is the batch tested significantly better than the minimum level of 1 international unit (IU) as specified in the monograph? This means that the test discriminates between high and low potency/quality rabies vaccine batches, but is not suitable for quantifying potency (Krämer et al., 2013).

6.2.4 Substage 1: Formal Incorporation into regulatory requirements (FI)

Monograph 0451 for inactivated rabies vaccines for veterinary use already allowed the replacement of the NIH test for batch potency by a validated alternative method, but did not refer to a specific assay. After the completion of the BSP105 the results were submitted to Ph. Eur. Expert Group 15V[86] and the assay was recommended for inclusion as an alternative batch potency assay in the Ph. Eur. monograph 0451 (Krämer et al., 2010). The formal incorporation of the SNT into the Ph. Eur. requires the steps as outlined in Figure 11. After the revision was put on the agenda a draft revision was prepared and published in 2011 in Pharmeuropa[87] for public consultation. This did not lead to any fundamental opposition and subsequently the European Pharmacopoeia

85 To anonymize the process, the samples were blinded and the results coded. One laboratory carried out the challenge test in mice to confirm the expected potencies. A sub-potent vaccine was included in the test to check if it would fail. The control sera of different levels of activity were centrally provided and the test results of the different labs were gathered and centrally evaluated through statistical evaluation using the software Combistats at the EDQM.

86 Group 15V veterinary sera and vaccines is responsible for the evaluation and approval of veterinary vaccine monographs.

87 Pharmeuropa contains draft pharmacopoeial texts for which the European Pharmacopoeia Commission is seeking comment. https://www.edqm.eu/en/Pharmeuropa-Texts-for-comments-1587.html?mbID=56 :consulted March 2015

Commission adopted the revised draft of monograph 0451 at its 142nd session in April 2012. It came into force on April 1, 2013. The formal incorporation of the SNT into the European monographs was thereby effectuated.

6.2.5 Substage 2 and 3: Actual Regulatory Acceptance and Use by Industry (ARA & UI)

Since its revision, monograph 0451 offers the possibility to conduct the SNT instead of the NIH test. In addition the monograph states that, *"In accordance with the provisions of the European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes such an alternative validated method should preferably be used for routine testing."* (European Pharmacopoeia, 2013).

Nonetheless, the incorporation of the SNT serological assay does not directly imply the replacement of the NIH test by the SNT, let alone the deletion of the NIH test from the monograph. Both assays are specified and it is up to the OMCLs and the manufacturers to choose the most suitable method. Rabies vaccine batch release is done by several OMCLs in Europe. These results are accepted within the EU by mutual recognition of data within the OMCL network for veterinary vaccine batch release.

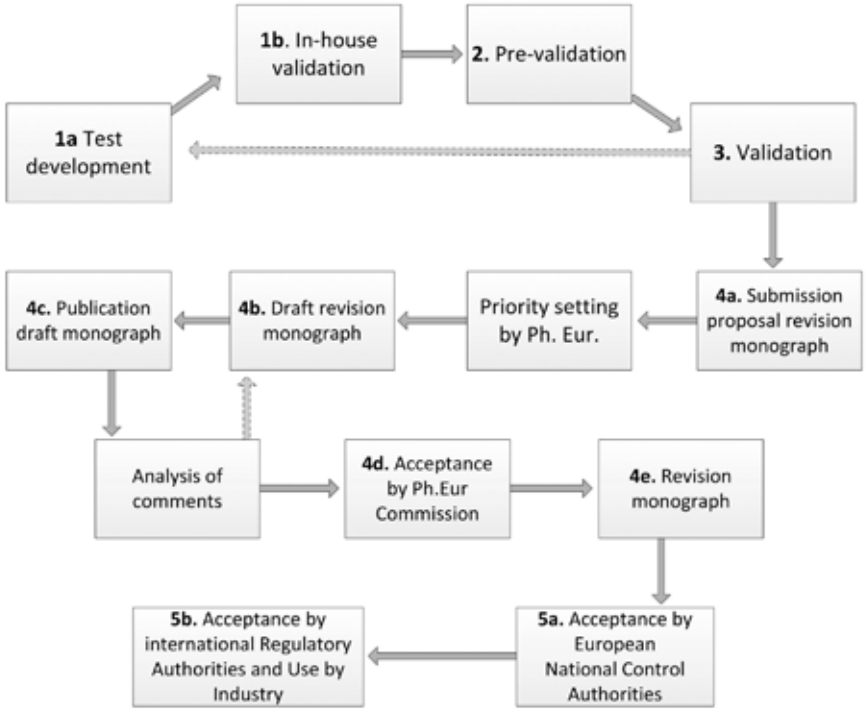


Figure 11. Ph. Eur. process from test development to test implementation.
(Based on PPT C. F. M. Hendriksen, AllChemE Seminars. Brussels, 2003)

For veterinary rabies vaccine batch control, the PEI is the leading OMCL and they encourage the SNT for batch release testing. This means that manufacturers are stimulated to use the SNT to demonstrate their vaccine's potency. Other European OMCLs involved in veterinary rabies vaccine batch release are the French ANSES, the Swiss IVI, the Hungarian NCE and the Czech ÚSKVBL. ANSES and IVI have already adopted the SNT for the quality control (QC) of inactivated veterinary rabies vaccines. The Hungarian and Czech OMCLs are in the process of implementing it.

Manufacturers however carry the final responsibility to prove the validity of the method for their specific product. At the time of writing, one manufacturer had successfully validated the SNT for their product (personal communication with employee) and declared that the NIH test will no longer be conducted for batch release purposes of veterinary rabies vaccines for the European market. Another manufacturer was in the process of validating the SNT for one of their products, but ran into some difficulties in meeting the criteria set by the PEI for the product specific validation of the serological assay (personal communication with employee).[88]

6.3 Factors influencing the FI, ARA and UI of the SNT

This section defines the drivers and barriers that are seen to have influenced the process of regulatory acceptance and use of the SNT. The process of formal incorporation of the SNT into Ph. Eur. Monograph 0451 was accomplished in about three years time when counting from the first publication of Krämer et al. in 2009. This is very fast when compared to similar processes which often take ten years or more (Cooper and Jennings, 2008). As such the process of the SNT can be seen as a best practice in terms of the FI of 3R models for regulatory purposes. This swift FI was accomplished through a mix of factors.

First of all, rabies vaccine batch release testing has been regularly defined as a hot topic and a priority in terms of the 3Rs. This was fed by the broad agreement on the weaknesses of the NIH test (Stokes et al., 2011, 2012). The PEI, as one of the leading OMCLs in Europe when it comes to rabies vaccines, harnessed the momentum to start developing the SNT and to initiate a pre-validation study which showed good scientific results in terms of reproducibility and correlation with the NIH test. Subsequently, the collaborative study was a well-organized and coordinated process. In BSP collaborative studies project leaders are assigned by the BSP Steering Committee to coordinate the projects in cooperation with the EDQM secretariat, and the subsequent steps are reviewed regularly and approved by the steering committee. The collaborative study included an international group of participants from regulatory authorities (Europe, US and Canada) and from industry. While the primary focus was Europe, non-European partners were encouraged to participate. In addition, the BSP works hand in hand with the European Pharmacopoeia Commission and the European Pharmacopoeia Groups of Experts dealing with biologicals, of which group 15V (Vaccines and sera for veterinary use) is important in this context.[89]

88 In light of this example it is interesting to note that the Expert Group on the Application of the 3Rs in Regulatory Testing of Medicinal Products of the European Medicines Agency (EMA) started a discussion on the possibilities for a lighter product specific validation for 3R models (communication with involved expert).

89 The European Pharmacopoeia Commission refers requests for revision to group 15V.

This group of experts played a critical role in encouraging communication between the stakeholders and promoting regulatory acceptance.

Nonetheless, these policy entrepreneurs also had to face several challenges at the stage of FI. The preparation of the critical reagents as a renewable resource (e.g., specific antibodies or reference antigens) is important in such collaborative studies, as is the availability of a range of vaccines from the market, including different formulations and the use of sub-potent or borderline samples to test the system (Milne and Buchheit, 2012). In this case, the availability of reagents and of appropriate samples proved challenging. In addition, the BSP has to adhere to many regulations when sending biological samples to countries outside Europe and these biological materials run the risk of being impounded at the borders (communication with BSP employee). Another challenge was convincing participants to cooperate in the collaborative study. Collaborative studies are complicated and energy consuming processes and several participants were uncertain about the benefits of taking part. Some stakeholders were anxious that participation could in the end lead to a forced use of the SNT. As a result, engaging these participants in this process took a lot of persuasion and reassurance from the side of the initiators of the collaborative study.

When it comes to the use of the SNT for regulatory purposes the monograph is clear about the preferred method, namely an alternative validated method (see Section 6.2.5). However, the monograph does not fully elucidate under what exact circumstances the 3R option will be accepted. From the side of European Regulatory Authorities the actual regulatory acceptance (ARA) of the SNT is imminent. As mentioned, the central OMCLs PEI, ANSES and the IVI already use the SNT and the other two European OMCLs involved in rabies batch release testing are in the process of implementing it. However, with the NIH test remaining part of the monograph, the Ph. Eur. leaves a certain amount of discretionary space to the regulatory authorities and manufacturers to choose the method they consider most suitable. This discretionary space is a source of uncertainty for manufacturers in terms of whether and under what conditions the 3R model will in the end be accepted. Moreover, regulatory acceptance of test results from industry, based on the SNT, requires product specific validation, which is perceived to be a significant hurdle (Casey et al., 2011). It costs manufacturers time and money and requires parallel testing, which temporarily increases the use of animals. It thus depends on the cost benefit analyses of the manufacturer whether this final critical step will actually be taken. This is even more the case for veterinary rabies vaccines due to the fact that the price margins for veterinary vaccines are usually smaller than for human vaccines. This can influence the cost-benefit analyses of whether or not to use a 3R model to the disadvantage of the new model.

Besides, manufacturers have expressed their concern about the fact that the SNT assay is a pass/fail test that offers no information on the amount of antibodies induced by the vaccine. Additionally, several respondents mentioned that the SNT, despite the good results of the collaborative study, is observed to cause some difficulties in terms of non-responders. This is connected with the immune response of mice, which tends to be somewhat unpredictable. One of the respondents remarked: *"With a lot of good*

will you can use it. However, it is the question whether there is enough good will within industry to overcome these hurdles."[90]

To circumvent the problems of non-responders and of not getting information on the amount of antibodies induced by the vaccine, several stakeholders have expressed a preference to move straight over to *in vitro* methods (i.e., antigen quantification models). Manufacturers for example have expressed a preference for glycoprotein tests like ELISA's, which they often already use for in-process control purposes when producing rabies vaccines. Glycoprotein tests quantify the amount of antigen in the vaccine and offer a more accurate control on the quality of the vaccine.[91] This point of view is shared by the US Department of Agriculture (USDA) responsible for the licensing and batch release of veterinary rabies vaccines onto the US market.

The USDA is not completely at ease with the SNT. Despite the conclusions of the collaborative study, they feel the SNT is not informative and sensitive enough. They are therefore investigating the possibilities of a direct transition towards antigen quantification tests. In response to this discussion, the PEI recently developed a multi-dose serological assay to quantify the potency of inactivated rabies vaccines for veterinary use (Krämer et al., 2013). At this stage, it is still unclear to what extent this development will influence the current discussion. For now, some manufacturers have indicated that they will work towards combining the SNT with an antigen quantification assay with the future perspective to replace the SNT altogether by antigen quantification assays, if subsequent batches of a vaccine prove to be potent (see also Section 6.5: steps ahead).[92]

Another barrier is the fact that rabies vaccine manufacturers produce their vaccines for a global market. Due to a lack of harmonization of regulatory requirements, they have to take many different regulatory requirements into account. This withholds the adoption of an alternative model for regulatory purposes. At a global level the NIH test still is the leading procedure and all regulatory frameworks entail – a variation of – this assay (Schiffelers et al., 2014a: see Appendix II). Sticking to the NIH test therefore often is the most secure option for manufacturers in terms of international regulatory acceptance. In this context, a regulator however remarked: *"The lack of harmonization is the main argument industry plays. But if manufacturers inform regulatory authorities about the possibilities of serology and its acceptance in Europe they have the potential to convince these authorities."*

90 It should be noted that the NIH test faces comparable problems, but this did not become a major hurdle due to the fact that the test was developed in a period where the acceptance criteria for a test were much lower.

91 Antigen quantification is mainly seen to be of relevance for non-adjuvanted vaccines, since the adjuvant generally interferes with the antigenicity test. For this reason manufacturers might prefer to use antigen quantification for in-process testing and serology for final batch testing to demonstrate consistency.

92 In the European context the use of antigen quantification for QC purposes could lead to difficulties for the OMCLs in terms of retesting. Article 82 of Directive 2001/82/EC enables European Member States to have specific vaccine batches – among which rabies vaccines – retested at an Official Medicines Control Laboratory (OMCL). In this process, known as "Official Control Authority Batch Release" (OCABR), tests conducted by the manufacturer are repeated (Cooper and Jennings, 2008). Repeating the antigen quantification tests however is a challenge due to the fact that these tests are custom made for a specific product. As a result the tests differ slightly per manufacturer. This means that the OMCL responsible for the retesting of a rabies vaccine must be able to conduct a large variety of specific antigen quantification tests.

In practice this effort will normally only be made if profitable in one way or the other to the manufacturer or if required by law. Nonetheless, some manufacturers are seen to have a strong company policy on the topic of the 3Rs and they have already taken substantial steps to phase out the use of the NIH test and to adapt to the changed European requirements. As such they function as frontrunners in the field.

6.4 Analyses

To analyze the influence of the different drivers and barriers on the regulatory acceptance and use of the SNT, they are positioned in the multilevel perspective on technology transitions (Schot and Rip, 1996), which was developed to better understand complex technology transition processes such as the acceptance and use of 3R models (Schiffelers et al., 2012). Such system innovations are almost always *"the result of the interplay between many factors and actors"* (Geels, 2006). Consequently an integrative approach is needed to comprehend such processes, and this is offered by this multilevel perspective. It addresses three levels of factors that influence technology transitions (Geels, 2006, Kemp, 2010; Schiffelers et al., 2012):

- The micro- or niche level where innovations are developed and validated;
- The meso- or sociotechnical regime level which includes the existing rules and regulations, expertise, dominant practices and the standing institutions;
- And the macro- or sociotechnical landscape level which comprises the material infrastructure, the existing political culture and coalitions, social values, the macroeconomy, demography and the natural environment.

Successful technology transitions require alignment of the developments at these three levels. An aggregation of the developments at these levels can only occur if an innovation (e.g., a 3R model) meets the needs of the meso- and the macro level (Schiffelers et al., 2012: Chapter 5). If a new technology does not comply with these needs, it will be incapable of escaping from the niche where it was developed (Kemp, 2010). On the other hand, the meso- and macro level have to open up to alternative ways of thinking in order to give a new technology a serious chance to break through.

Apart from playing a role in the analysis of the influences on regulatory acceptance and use, a distinction between these three levels is also helpful in defining those influences that offer better opportunities in terms of improving the acceptance and use of an innovation. The factors at the micro- and partly the meso level for example tend to offer more possibilities for change than the broader societal developments at the macro level.

The process of FI of the SNT can be defined as a success in terms of getting a 3R model incorporated into regulatory requirements. The SNT was able to escape from the niche in which it was developed and validated and to become part of the European regulatory regime, i.e., the monographs of the Ph. Eur. This is mainly the result of the solid basis that was created scientifically, in terms of the process during the pre-stages that anticipated the FI and by the legislative context that stimulates the acceptance and use of 3R models.

Additionally, its actual regulatory acceptance within Europe is largely accomplished and its use by industry in the European context is slowly progressing. However, its regulatory acceptance and use by regulators and industry at a global scale still is a big challenge.

Figure 12 summarizes the different forces (as described in Section 6.3) that are observed to have played a role in the process of regulatory acceptance and use of the SNT, using the multilevel perspective on technology transitions. It shows the opposing forces at hand.

In terms of drivers, both the pre-validation study and the collaborative study were very important in proving that the alternative method works in the hands of all the participants. The process from test development to FI was facilitated by strong policy entrepreneurs and a clear problem ownership and process management – first the PEI, then the BSP, then the European Pharmacopoeia Commission and group 15V. Additionally, the intense collaboration within the OMCL network, the early involvement of a statistician to design the study and analyze the data during the validation process and the wide dissemination of the study results have all added to the swift incorporation of the SNT into monograph 0451. However, the swiftness also came with a price. The persuasiveness of the initiators/project-coordinators was a big driver for the adoption of the SNT into the Ph. Eur. monographs, but may have partly led to restrained attention to the drawbacks of the SNT as perceived by manufacturers and regulators such as the USDA.

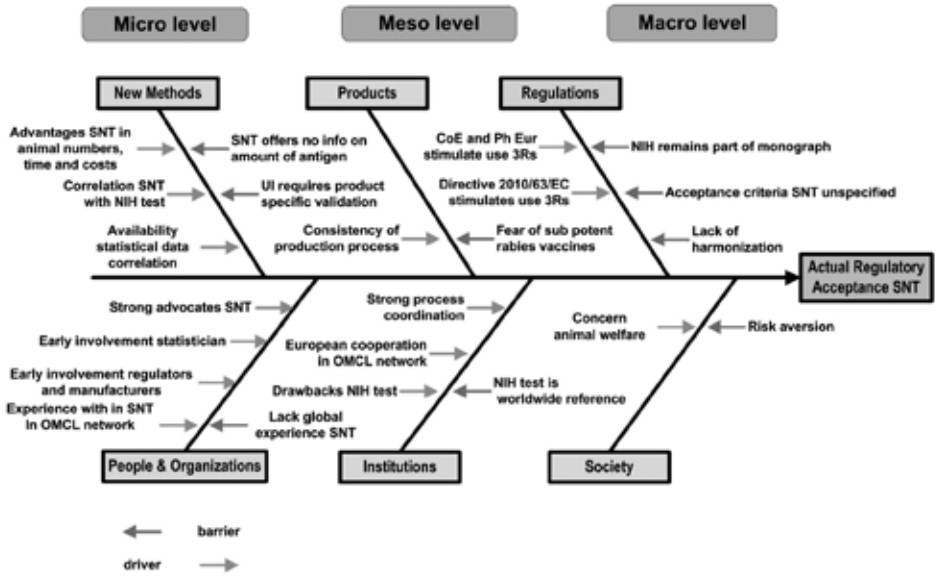


Figure 12. Multilevel perspective on drivers and barriers influencing regulatory acceptance and use of the SNT. (Source: based on Schiffelers et al. 2012: Chapter 5).

In the European situation the drivers have been able to outweigh the barriers in the substages of FI and ARA. For the broader ARA – i.e., outside Europe – and for the UI it is still highly uncertain whether the drivers will be able to outweigh the barriers. Here the previously mentioned discussion about the drawbacks of the SNT plays an important role. In the US these drawbacks are observed to outweigh the benefits of the SNT and as a result the USDA does not opt for the use of the SNT. Instead they are investing in a direct transition towards full *in vitro* methods, i.e., antigen quantification. Several manufacturers have also expressed a preference for this approach.

They question the added value of the SNT and instead would like to invest in *in vitro* methods, which most of them already use for production and in-process control purposes. The SNT for them is merely an intermediate step or even an unnecessary step from the NIH test to the use of *in vitro* methods as part of a consistency approach (see Section 6.5). In this context, several respondents from industry suggested that representatives from industry should have been involved earlier on in the process of the development and pre-validation of the SNT. According to them this could have prevented that the discussion on the drawbacks of the SNT test surfaced at the phase of UI.

6.5 Lessons learned and steps ahead

From the perspective of the European ambition to diminish the use of laboratory animals and to stimulate the use of 3R models (see Section 6.1), it is important to not only clarify the process of regulatory acceptance and use of the SNT, but also to answer the following questions:

- *Which lessons can be learned from the process of FI, ARA and UI of the SNT to stimulate regulatory acceptance for similar processes in the future?*
- *Which additional steps are needed to replace the NIH test by the SNT or other 3R options?*

The case of the SNT teaches us several important lessons regarding the regulatory acceptance and use of 3R models.

First of all it shows that there must be a firm commitment of the key stakeholders to allocate time and money to take part in such a project, to exchange method details, reagents, test samples and adhere to the specific rules of the collaborative study. The strong commitment from European stakeholders such as the PEI, the EDQM and the BSP was fed by legal texts of both the CoE[93] and the EC[94] calling for minimizing the suffering, pain and distress caused to animals used for scientific purposes.

Secondly, effective interaction between central stakeholders within industry and regulatory authorities on an international level proves to be essential. Interaction is necessary to exchange and transfer available scientific data with regard to the 3R model and to discuss criteria that have to be met for regulatory acceptance and use. It is pivotal that this is done within the regulatory framework the 3R model is destined for, which was, in this case, the

93 The European Convention for the protection of vertebrate animals used for experimental and other scientific purposes (1986, ET S 123).

94 Directive 2010/63/EU on the protection of animals used for scientific purposes.

EDQM and the OMCL network. In this way support from the regulators is stimulated from the very beginning. Early involvement is also needed from the side of industry. IFAH[95] can fulfill the network role for manufacturers – comparable to the OMCL network at the side of regulators – but it is also important to involve manufacturers individually, owing to the fact that the manufacturers are competitors. Furthermore, it is important to firmly involve regulatory authorities from other parts of the world to anticipate what is needed for broader regulatory acceptance.

Thirdly, the SNT process shows the importance of a well-designed and coordinated validation process which starts with a small scale feasibility study and moves to a large scale method transfer and validation, subsequently involving more labs and more products. Whether a 3R method will be fit for regulatory acceptance depends on the availability of reproducible test data. Regulators need proof that an assay does what it is intended to do, i.e., to make a distinction between a potent and a sub-potent vaccine batch. To be able to compare and interpret the data of the different stakeholders in a meaningful manner the early involvement of statistical knowledge to ensure a well-designed study is of great importance.

Fourthly, the validation process requires a strict process management with predefined steps to be taken and questions to be answered. This needs to be supervised by well informed and committed coordinators with the legitimacy to get the parties together.

Despite the positive experience of the SNT collaborative study, the validation process of 3R models is mostly very challenging due to the required correlation with the existing *in vivo* assay, which is very difficult if not impossible to achieve. *In vivo* assays habitually are a poor reference owing to the often highly variable test results (Schiffelers et al., 2014a: see Appendix II). Therefore, a change in the way of thinking is needed to help 3R models out of their niches and into the existing regulatory regime. Such a different way of thinking in terms of validation is the concordance strategy in which regulatory approval and implementation of an alternative method can be obtained after a pass/fail correlation using sub-potent batches instead of by seeking a full correlation with the conventional animal model (Stokes, 2012).

The next challenge, as mentioned, is regulatory acceptance of the SNT outside Europe and its use by manufacturers for batch release purposes. For this, several actions are needed, i.e., ongoing international communication among regulatory authorities and between regulatory authorities and manufacturers and continuing harmonization efforts. A first step in this direction, although non-binding, was taken in 2013 by the World Organization for Animal Health (OIE) through the adoption of the SNT potency testing of inactivated veterinary rabies vaccines in their Manual of Diagnostic Tests and Vaccines for Terrestrial Animals.[96]

95 The International Federation for Animal Health (IFAH) is the global representative body of companies engaged in research, development, manufacturing and commercialization of veterinary medicines, vaccines and other animal health products across the world.

96 http://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/2.01.13_RABIES.pdf

The acceptance and use of the SNT for human inactivated rabies vaccines is an additional step to be taken. Human rabies vaccines are generally less complex in composition as they are non-adjuvanted. Moreover, the RFFIT (the predecessor of the SNT) is already recognized by the WHO as a valid alternative method for human rabies vaccines (WHO, 2007). Nonetheless, it was agreed that further research and validation is needed for serological assays such as the RFFIT and the SNT to gain broader acceptance for human rabies vaccine potency testing (Casey et al., 2011).

In terms of QC testing of vaccines the consistency approach is an important alteration in the way of thinking. This approach, which has already gained its merits in the area of well characterized vaccines, is winning terrain in the discussions about batch release testing of classical vaccines, like rabies vaccines. The consistency approach is based upon the principle that the quality of a vaccine is the result of the strict application of a quality system and consistent production (Hendriksen et al., 2008, Kulpa-Eddy and Dusek, 2011). With this approach, the focus is changed from batch release testing to in-process control. It implies that a consistent production process is key to the quality of a vaccine. The approach allows replacing animal bioassays like the NIH test on the final batch by a battery of meaningful non-animal tests with enhanced capacity to compare new batches with batches of proven quality (Hendriksen et al., 2008). This approach requires a combination of immuno-chemical and physico-chemical tests performed in-process and on the final product. Such a combination of tests, together with adherence to the guidelines of Good Manufacturing Practice (GMP) (De Mattia et al., 2011), shall ensure that all produced batches are of the same quality as the batches that have proven to be safe and efficacious during licensing (Kulpa-Eddy and Dusek, 2011). For conventional products – like rabies vaccines – it is believed that the consistency approach will lead to a substantial reduction in animal use for potency testing purposes in the final batch (Hendriksen et al., 2008), even though some animal testing may still be required during prelicensing or validating manufacturing changes (Kulpa-Eddy and Dusek, 2011).

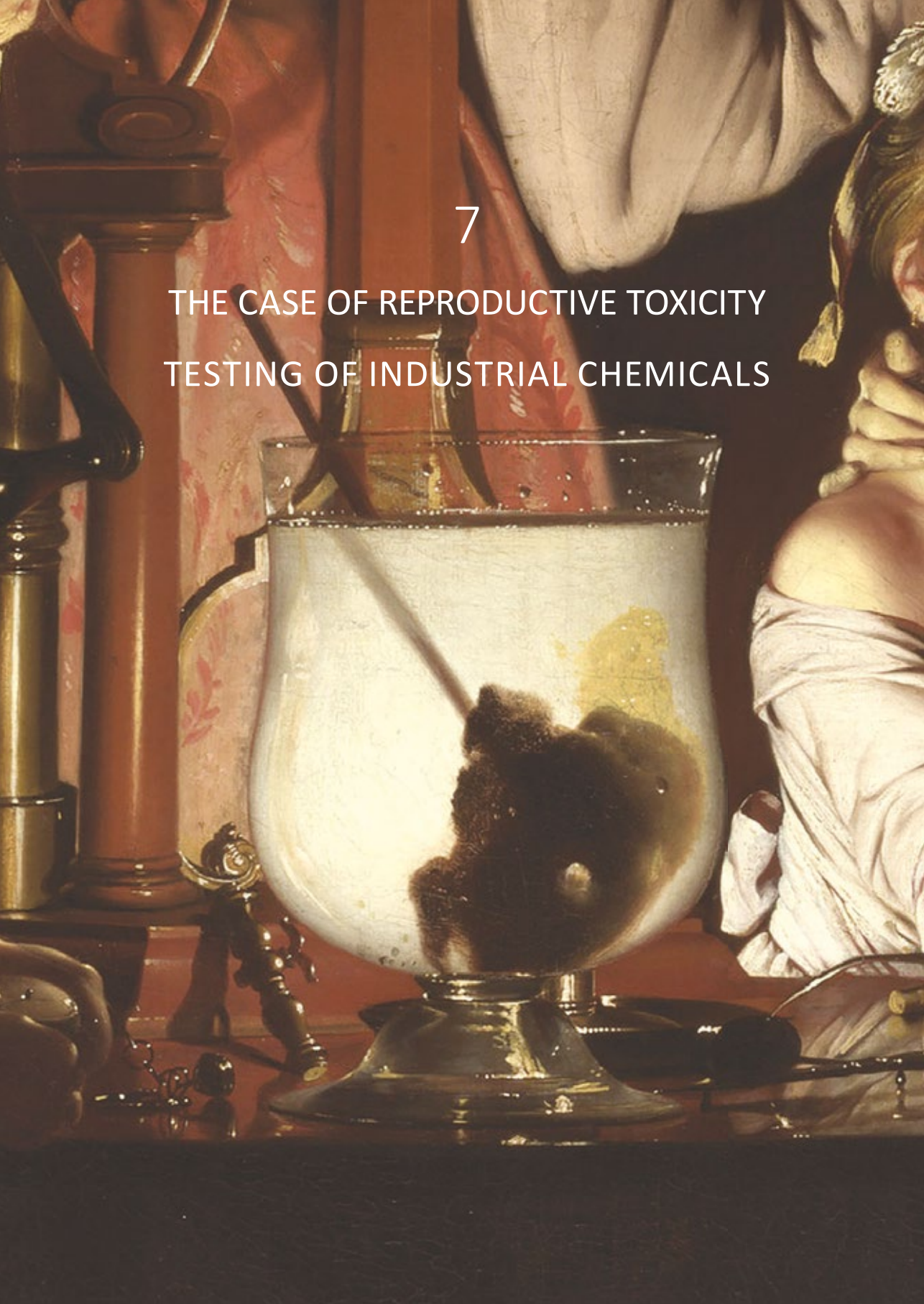
The Ph. Eur. and EMA (European Medicines Agency) underscore the potential of the consistency approach and are taking gradual steps in this direction. The general notices of the Ph. Eur. (Supplement 8.2) for example were revised in June 2013 in order to address the consistency of production approach in the context of reduction of animal testing. This change entered into force on July 1, 2014. The monograph on veterinary rabies vaccines was being adapted to these general notices at the time of writing.

For the process of actual regulatory acceptance, regulators depend on the data on the production process of a specific vaccine from manufacturers. This means that industries need to share their *in vivo* and *in vitro* data with regulators. Industries however tend to be very cautious in sharing test information with regulators as long as they are uncertain how this might influence the regulatory decision regarding their product. Manufacturers therefore ask for more clarity on the acceptance criteria when it comes to the evaluation of their product. Regulators in turn are very cautious with regard to specifying precise acceptance criteria as long as they do not know what kind of data they can expect.

This leads to a "catch 22" in which both parties are waiting for the other to take the first step (Schiffelers et al., 2014a: see Appendix II). *"To enable the submission of results of screening tests outside the drug specific regulatory decision processes, the EMA is now working on the so called safe harbor concept. In this way we can compare these data with the results of the conventional tests as described in the current requirements within a neutral context."* (personal communication with involved expert). Within such safe harbors manufacturers can discuss their test results with regulators without this having direct consequences on the evaluation of their product (see also Chapter 8).

The process of regulatory acceptance and use of 3R models to replace the NIH test for rabies vaccine potency testing has long been characterized by inertia. This inertia is fed by the dreadfulness of the disease and the fear of missing sub-potent rabies vaccine batches. Moving away from the well-known NIH test entails taking a risk in this highly risk-averse context. The biggest chance for successful change in such high risk areas is created by thinking in terms of evolutions rather than in terms of revolutions (Schiffelers et al., 2012: Chapter 5). For this reason the acceptance and use of 3R methods in this area requires an incremental process, i.e., no change in terms of radical developments but new test regimes that gradually grow out of old ones (Geels, 2002). The regulatory acceptance and use of the SNT has to be placed in this context. Even though *in vitro* methods might be preferable in several respects, the use of antigen quantification for potency testing purposes of adjuvanted vaccines still needs to overcome significant technical problems. Optimization/validation of these *in vitro* models is therefore likely to require a considerable additional amount of time. For the purpose of moving away from the NIH test within a limited timeframe, the use of the SNT is a recommendable intermediate step in the transition towards in-process control using *in vitro* methods. So, even though there is a preference among several stakeholders to make a direct move to *in vitro* methods, a more gradual approach which starts with combining serology with *in vitro* methods, is in our view a very sensible approach.

To conclude, the process that led to the formal incorporation of the SNT into Ph. Eur. monographs demonstrates the importance of the four C's of Commitment, Communication, Collaboration and Coordination (Schiffelers et al., 2014b: Chapter 8). However, the case of rabies vaccine potency testing reveals a fifth important C, namely the C of Continuity. To gradually replace the NIH test, continuity is very important. Many stakeholders have worked for decades to replace the NIH test. Their enduring effort can be seen as one of the central drivers in phasing out the NIH test and to ultimately change the collective mindset in favor of the 3Rs.



7

THE CASE OF REPRODUCTIVE TOXICITY TESTING OF INDUSTRIAL CHEMICALS

"For good ideas and true innovation, you need human interaction, conflict, argument, debate"

Margaret Heffernan
born 1955
Businesswoman and writer

A short guide to Chapter 7

Chapter 7 consists of a research article published in 2015 in *Regulatory Toxicology and Pharmacology*. It comprises the results of a case study on the acceptance and use of the Extended One-Generation Reproductive Toxicity Study (EOGRTS) to test reproductive toxicity effects of chemicals. The case study was conducted between 2012 and 2014. Through this case study the research questions Q3a and Q3b are directed (see section 1.2). For the sake of the completeness of this publication a certain amount of replication of chapters 1 to 4 was necessary.

Abstract

The two-generation study (TG 416: OECD, 2001) is the standard requirement within REACH to test reproductive toxicity effects of chemicals with production volumes >100 tonnes. This test is criticized in terms of scientific relevance and animal welfare. The Extended One-Generation Reproductive Toxicity Study (EOGRTS), incorporated into the OECD test guidelines in 2011 (TG 443: OECD 2011a) has the potential to replace TG 416, while using only one generation of rats and being more informative. However, its regulatory acceptance proved challenging. This article reconstructs the process of regulatory acceptance and use of the EOGRTS and describes drivers and barriers influencing the process. The findings derive from literature research and expert interviews. A distinction is made between three substages: the stage of Formal Incorporation (FI) of the EOGRTS into OECD test guidelines was stimulated by retrospective analyses on the value of the second generation (F2), strong EOGRTS advocates, animal welfare concern and changing US and EU chemicals legislation; the stage of Actual Regulatory Acceptance (ARA) within REACH was challenged by legal factors and ongoing scientific disputes; while the stage of Use by Industry (UI) is influenced by uncertainty of registrants about regulatory acceptance, high costs, the risk of false positives and the manageability of the EOGRTS.

Reference:[97]
Schiffelers, M.J., Blaauboer, B., Bakker, W., Hendriksen C. and Krul, C. (2015). Regulatory acceptance and use of the Extended One Generation Reproductive Toxicity Study within Europe. *Regul Toxicol Pharmacol* 71, 114-124.

97 Acknowledgement: We thank the Doerenkamp-Zbinden Foundation and the Dutch Ministry of Economic Affairs for funding this project. We thank the respondents for sharing their expertise and perspectives with us.

7.1 Introduction

Chemicals are subjected to a broad range of requirements to guarantee safety for humans, animals and the environment. The requirements describe the endpoints for which chemical substances have to be assessed and generally also the test procedures that need to be performed for a particular endpoint. Reproductive and developmental toxicity are two of the main endpoints in the assessment of industrial- and agrochemicals. These endpoints include the toxic effects of a substance on an organism's reproduction and development of its offspring. The reproduction cycle of mammals, being a highly complex process, is very difficult to investigate *in vitro*. For this reason regulatory reproductive and developmental toxicity tests are still conducted in laboratory animals with a prenatal developmental study in rodents and a non-rodent species and a one- or two generation reproduction toxicity study in rats (Janer et al., 2007b).

Since the 1980s the OECD 416 two-generation study has been the most comprehensive reproductive toxicity study (OECD, 2001). Up to 30% of the reproductive toxicity tests conducted are two-generation studies (Spielmann and Vogel, 2006), requiring around 2600 animals per study (Lilienblum et al., 2008). The two generation test is estimated to use nearly 40% of the laboratory animals under REACH (Janer et al., 2007a) and thereby is one of the major users of rodents in safety test programs.

In anticipation of the introduction of the European Directive for the Registration, Evaluation, Authorization and restriction of CHemicals – REACH (EU, 2006) concern was expressed that reproductive toxicity testing would lead to a significant increase in numbers of animals needed. Reproductive and developmental toxicity were even estimated to become the largest animal user for safety testing within REACH (Pedersen et al., 2003; Van der Jagt et al., 2004) since approximately 10,000 chemicals with an annual volume of >100 tonnes would have to be tested on reproductive toxicity. The estimates ranged from 40% to 90% of the total number of animals to comply with REACH that would be needed for reproductive toxicity testing purposes (Van der Jagt et al., 2004; Spielmann and Vogel, 2006; Hartung and Rovida, 2009; Martin et al., 2011). At about the same time, several studies became available that questioned the added value of the second generation (Cooper et al., 2006; Janer et al., 2007a,b; Martin et al., 2009; Piersma et al., 2011) and criticized the limited predictive value of the OECD TG 416 for developmental immunotoxic and neurotoxic parameters (See Section 7.2.1.).

In 2006 the Agricultural Chemical Safety Assessment (ACSA) Technical Committee of the ILSI Health and Environment Sciences Institute (HESI) proposed a whole new testing paradigm, which constituted a tiered approach of toxicity testing. Part of this paradigm was a proposal for an alternative protocol for OECD TG 416 which required only one generation of animals while being more informative in data obtained (Cooper et al., 2006). This protocol became the basis for the Extended One Generation Reproductive Toxicity Study – EOGRTS – with a reduction of up to 40% in animal use – i.e. a total of 1200 animals per study – compared to the two-generation study. In addition the EOGRTS protocol includes parameters for developmental neurotoxicity-DNT – and developmental immunotoxicity-DIT –. The Cooper protocol was proposed to the OECD secretariat for incorporation

into the OECD guidelines in 2007 and accepted in 2011 as OECD TG 443 (OECD, 2011a) after a process in which many amendments were made, as will be described in Section 7.2.1 of this chapter.[98]

The EOGRTS matches with the ambition of the European Commission to diminish the use of laboratory animals and to stimulate the acceptance and use of models to replace, reduce and refine (3Rs) existing animal models (Russell and Burch, 1959). This ambition is laid down in Directive 2010/63/EU on the protection of animals used for scientific purposes and in REACH. Directive 2010/63/EU states in article 13.2 that in choosing between procedures, those which use the minimum number of animals shall be selected (EU, 2010). Furthermore, REACH states in article 25 (1) that in order *''. . .to avoid unnecessary animal testing, testing on vertebrate animals for the purpose of this Regulation shall be undertaken only as a last resort''.* (EU, 2006) (See also Section 7.2.2). Despite these legislative stimulants and the incorporation of the EOGRTS into the OECD test guidelines, the regulatory acceptance and use of the EOGRTS within Europe has been a point of strong disparity. This raises the following key questions which will be addressed in this chapter:

- *Which factors influence the regulatory acceptance and use of the EOGRTS within Europe?* [99]
- *What is needed to augment the current process?*
- *Which lessons can be drawn from the case of the EOGRTS for future processes?*

To improve the use of the 3Rs congruent with the EC's ambition an exhaustive comprehension of the process of regulatory acceptance and use and its drivers and barriers is needed. In order to understand and examine the regulatory process, we made a distinction between the following three successive stages:

- Formal Incorporation into the OECD test guidelines (FI)
- Actual Regulatory Acceptance by regulatory authorities (ARA)
- Use for regulatory purposes by Industry (UI)

98 ''This Test Guideline is designed to provide an evaluation of reproductive and developmental effects that may occur as a result of pre- and postnatal chemical exposure as well as an evaluation of systemic toxicity in pregnant and lactating females and young and adult offspring. In the assay, sexually-mature males and females rodents (parental (P) generation) are exposed to graduated doses of the test substance starting 2 weeks before mating and continuously through mating, gestation and weaning of their pups (F1 generation). At weaning, pups are selected and assigned to cohorts of animals for reproductive/developmental toxicity testing (cohort 1), developmental neurotoxicity testing (cohort 2) and developmental immunotoxicity testing (cohort 3). The F1 offspring receive further treatment with the test substance from weaning to adulthood. Clinical observations and pathology examinations are performed on all animals for signs of toxicity, with special emphasis on the integrity and performance of the male and female reproductive systems and the health, growth, development and function of the offspring. Part of cohort 1 (cohort 1B) may be extended to include an F2 generation; in this case, procedures for F1 animals will be similar to those for the P animals''. http://www.oecdilibrary.org/environment/test-no-443-extended-one-generation-reproductive-toxicity-study_9789264122550-en. :consulted May 2014

99 Although this paper focusses on the European situation, major parts of the discussion in the US are also addressed in this chapter.

Full regulatory acceptance and use means that a 3R model has passed all three stages.This chapter builds on earlier work of the authors (Schiffelers et al., 2012: Chapter 5, Schiffelers et al., 2014a: Appendix II) which examined the process of regulatory acceptance and use of 3R models from a technology acceptance perspective (see also Section 7.3). The reconstruction of the EOGRTS case offers additional in depth knowledge of this process.

The EOGRTS is, at the time of writing, in a critical phase. Although there is agreement on the inclusion of the EOGRTS in the fifth adaptation of the REACH test methods regulation, the discussion on the actual regulatory acceptance (ARA) and the use of the EOGRTS by industry (UI) for the release of chemicals was still taking place within Europe. Disentangling the process from a more general perspective of technology acceptance offers input for this discussion and lessons for future processes.

7.2 Results

This section reconstructs the process of the acceptance and use of the EOGRTS and gives an overview of the barriers and drivers on this process throughout the three substages of Formal Incorporation (FI) of the EOGRTS in the OECD Test Guidelines (Section 7.2.1); the Actual Regulatory Acceptance (ARA) by European regulatory authorities for chemical registration and authorization purposes under REACH (Section 7.2.2); and the Use by Industry (UI) for chemical registration and authorization purposes under REACH (Section 7.2.3). The findings derive from the examination of available documents connected to the acceptance process (e.g. meeting- and workshop reports) and a series of interviews with experts involved in this process (see Appendix IX for a description of the methodology used for this case study). Several quotes from respondents are inserted to elucidate the description of drivers and barriers.

7.2.1 The Formal Incorporation (FI) of the EOGRTS in the OECD test guidelines

From 2006 onwards several developments have led to the development of the EOGRTS and its incorporation into the OECD test guidelines, with the ILSI HESI strategy for agricultural chemical safety assessment (ACSA) (Cooper et al., 2006) as the starting point. Shortly after the publication of the Cooper report, several groups of experts examined the possibilities of leaving out the second generation for other products such as industrial chemicals.[100] This discussion was especially important in the light of the new REACH regulation that was estimated to lead to a significant increase in reproductive toxicity studies (see Section 7.1). The EOGRTS was seen as a possible answer to some of the future questions in terms of the risk assessment of chemical substances, such as meeting the REACH deadlines for CMR (Carcinogenic, Mutagenic and Reprotoxic) substances, the animal welfare concern related to reproductive toxicity testing under REACH and mounting questions concerning endocrine disruptors. The conclusion was that the EOGRTS, as proposed by the ACSA

100 For instance, both the European Centre for the Validation of Alternative Methods (ECVAM) and the European Partnership for Alternative Approaches to Animal Testing (EPAA) organized workshops in 2006 with experts from industry and regulatory authorities to evaluate the applicability of this approach for industrial chemicals.

project, was applicable for industrial chemicals, if handled in a flexible way and modified to the existing requirements.[101]

Next to this European sense of urgency there was pressure from the side of the US EPA due to the aspiration to incorporate the EOGRTS into the OCSPP (Office of Chemical Safety and Pollution Prevention) guidelines that needed to be revised. The attention for the topic on both sides of the Atlantic led to a shared initiative in 2007 by the US, Germany and the Netherlands to submit a proposal to the OECD secretariat to draft an OECD test guideline (TG) based on the EOGRTS. An expert group was formed which drafted the guideline through a series of teleconferences and meetings chaired by the US and the Netherlands. In 2009, the OECD member states, except for Sweden, agreed on a draft of the TG (Gilbert, 2011).

To scientifically booster the process, four retrospective reviews were conducted (Janer et al., 2007a; Martin et al., 2009; Piersma et al., 2011; Rorije et al., 2011).[102] The retrospective analyses of Janer et al., Martin et al. and Piersma et al. conclude that the second generation has very limited added value and is not essential to establish the lowest effect level (LEL)/lowest observed adverse effect level (LOAEL) of the substances under examination. The study of Rorije et al. concludes that the second generation does not play a crucial role in the classification decision of 50 classified reproductive toxicants in Europe. Moreover, these studies underline that the EOGRTS includes a more comprehensive evaluation, offering more information than the current two generation study, while using far less animals. The studies therefore concluded that the existing testing strategies for reproductive toxicity testing needed revision in favor of the EOGRTS.

The national authorities that submitted the proposal to the OECD, were strong advocates of the new protocol and proved successful in involving frontrunners within the American/ European agrochemical industry, ICAPO-International Council on Animal Protection- and

101 The EOGRTS, as proposed by the ACSA project was developed for agrochemicals. It therefore addresses much more endpoints as requested under REACH for industrial chemicals. This means that triggering and/or waiving criteria were needed for these additional endpoints. (ECETOC, 2008).

102 The Janer study evaluated 176 multi-generation studies on 148 substances to assess the potential differences between the first and the second generation. The study concluded that the F2 did not affect the overall NOAEL (no observed adverse effect level). No critical effect was observed in de second generation and the F2 had no impact on ensuing risk assessment nor on classification and labelling (Janer et al., 2007a). In 2009 the US Environmental Protection Agency provided the OECD EOGRTS expert panel with a report on a retrospective analysis of 350 multi-generation reproductive toxicity rat studies mostly conducted for pesticides (Martin et al., 2009), concluding that . . .''There is a great deal of redundancy in the second generation tests.'' (Gilbert, 2011). The OECD expert meeting concluded in 2008 that all available two generation studies should be combined into one database to be able to fully judge the value of the second generation. This analysis was executed in the Netherlands based on a US database structure (US EPA ToxRefDB). In 2011 Piersma et al. published a retrospective analysis of 498 multi-generation reproductive toxicity study reports performed since the 1980s to review the added value of the second generation (Piersma et al., 2011). The manuscript was officially published in 2011. However, the authors have made their intermediate results continuously available to the delegates in the OECD expert group, through a secured website, monthly teleconferences and two conferences at the Dutch airport Schiphol. The retrospective analysis was finalized in October 2010 and got accepted for publication in November 2010, just a week before the OECD Joint Meeting accepted the new test guideline. Additionally, in 2011 a study of Rorije et al. became available in which the impact of the 2nd gen on classification & labelling was studied in all 50 reproductive toxicants classified at the time (Rorije et al., 2011).

other OECD Member States. The extensive animal use, the time consumed by the two generation study and the limited added value of the second generation, were the dominant drivers for these stakeholders to support the FI of the EOGRTS in the OECD TG's. Subsequently some countries took the lead (US and NL) in keeping the EOGRTS high on the OECD agenda and guiding the protocol through the OECD process. Five years after the proposal was sent to the OECD secretariat, the EOGRTS was formally incorporated into the OECD TG's (OECD, 2011a).

As a result, the process of FI of the EOGRTS is often viewed upon as a success in terms of formal regulatory acceptance. However, this does not mean that the process went effortless. Or as one of the respondents said: *''This dream became reality in 4 years' time, but it was a very rough ride through the OECD expert committee''*. Four of the points of discussion throughout this process of FI are elaborated on below.

Firstly, the issue of the added value of the second generation. This issue gave rise to the dominant discussion throughout the process of FI. The studies of Janer et al. (2007a,b), Martin et al. (2009). Piersma et al. (2011) and Rorije et al. (2011) did not convince all parties involved. Some stakeholders (e.g. the European Chemicals Agency (ECHA), Sweden and France) questioned whether the retrospective analyses covered all categories of compounds and/or whether the involved studies did meet the right criteria to reflect the reality of risk assessment and classification and labelling (e.g. Ruden and Hansson, 2008). To meet the concerns of those stakeholders who did not feel at ease with leaving out the second generation the OECD Joint Meeting (JM) agreed not to specify the EOGRTS as a replacement for the two-generation test, TG 416 (OECD, 2001). The guideline leaves it to the competent authority when and how the EOGRTS can be accepted as a replacement for TG 416.

Secondly, a discussion took place on the incorporation of clear triggers for a second generation. This discussion was held in the context of the US legal requirement to perform a two-generation reproduction and fertility study for food-use pesticides (Cooper et al., 2006). The triggers (i.e. effects that can be found in the first generation of animals) were to be interpreted as signals that a second generation of animals was needed to monitor the full effect of a substance. The US developed a broad range of triggers for this purpose, but the evaluation of these triggers revealed that they would lead to a second generation study in about 48% of the examined studies, while in only 1% of the cases there was a scientific justification for doing so. This way of operating would lead to an unjustifiable increase in the use of laboratory animals and therefore resulted in the decision to address this issue in a guidance document [103] - GD 117-supporting OECD TG 443 (OECD, 2011b).

A third discussion point concerned the reduction of the premating exposure period from 10 weeks in the two generation study to 2 weeks in the EOGRTS.[104] This approach would, according to the proponents of the EOGRTS, have sufficed in all known cases to affect fertility, but this was heavily debated by several stakeholders.

103 A guidance document serves to specify the guideline and is far easier to adapt compared to a TG.

104 The 10 weeks originate from the duration of one complete spermatogenic cycle. However the EOGRTS prescribes 2 weeks premating exposure, plus extended exposure of males after mating up to a total of 10 weeks exposure followed by testis histopathology to check for testis effects.

A fourth point of discussion, initiated in 2009 by a coalition of animal welfare organizations and industry, concerned the necessity of including immunotoxic parameters into the TG. The concern of these stakeholders was that the inclusion of immunotoxic parameters would lead to an unjustified increase of animal use and additional costs. The discussion was part of a broader debate between OECD experts on how to incorporate the extra cohorts DNT and DIT in the guideline. In 2008 the DNT and DIT cohorts were still optional but at the second Joint Meeting in 2009, Canada and the US insisted to make both cohorts obligatory. According to several respondents this resulted in the demand of other OECD member states to make additional -optional-parameters (e.g. clinical chemistry, haematology, pathology and necropsy in extra organs) obligatory. By doing so the discretionary space, that was originally offered by the Cooper protocol, gradually disappeared.

The decision to require these parameters was, according to several respondents, not based on scientific arguments. It was part of the negotiation process that was needed to get the guideline adopted. These discussion points illustrate that the process of FI was a delicate process that took a lot of sensitivity and diplomacy. For the sake of the process numerous concessions had to be made. One of the respondents described the process as both encouraging and frustrating: *"On the one hand this guideline was established in a short time span and had an unprecedented pack of data to back the proposed change. On the other hand a small minority of OECD members remained withholding and several political decisions were made that were superfluous from a scientific perspective."*

7.2.2 The Actual Regulatory Acceptance (ARA) of the EOGRTS within the context of REACH

Ever since the establishment of TG 443 there has been discussion about its implementation in the context of REACH. The regulatory acceptance of the EOGRTS for risk assessment purposes proved to be highly challenging. One of the respondents remarked; *''writing the guideline was quite easy; it is the implementation which is the difficult part''*. This is the result of a combination of factors as described below.

First of all, OECD test guidelines are non-binding guidance documents leaving the implementation of the test up to the relevant competent authority. Furthermore, TG 443 was not defined as a direct replacement for the two-generation test and is flexible in the way it can be operationalized. This offers regulatory authorities and registrants the possibility to adapt the protocol to the needs of a specific situation, but at the same time leads to an ongoing discussion regarding the use of the EOGRTS within the context of REACH.

In August 2011 an Expert group-EOGRTS EG – was established within CARACAL (the Competent Authorities for REACH and CLP) on how to operationalize TG 443 within REACH. The preliminary conclusion of this expert group was to use the EOGRTS as the preferred method under REACH. This conclusion was supported by CARACAL. Subsequently the European Commission (EC) was urged to initiate the inclusion of OECD TG 443 in the Test Method Regulation (TMR) and to modify the relevant REACH Annexes accordingly. However, the legal, procedural and financial analysis on the application of TG 443 was not covered by the mandate of the EOGRTS EG and needed to be resolved by the

Commission. During the November meeting of 2011 of the Member States Committee of ECHA (MSC), the secretariat of the European Chemicals Agency (ECHA) gave a presentation on the legal considerations and procedural aspects of the EOGRTS. The MSC was informed that in accordance with Article 13(3) of REACH Regulation, ECHA can in principle recognize OECD TG 443 as an appropriate study guideline, but to meet the requirements in Annex IX/X, 8.7.3 of REACH there are legal considerations that have to be taken into account. This means that as long as the EOGRTS is not part of the REACH TMR its legal position remains under discussion.

During the ECHA MSC meeting of September 2011 it became apparent that the MSC for the first time could not come to a unanimous agreement on a draft testing proposal for reproductive toxicity for scientific and technical reasons. Some members of the MSC were in favor of the EOGRTS, while others preferred to leave the choice for the EOGRTS or the two-generation study to the registrant. A third group supported the latter argumentation but wanted to include the second generation in the EOGRTS. To deal with this disagreement, the procedure foreseen in Article 51(7) of the REACH Regulation was used for the first time. This procedure offers the possibility to split the draft decision on the registration of chemicals in those cases where the MSC fails to reach unanimous agreement. The part on which no agreement is reached, is sent to the EC to take the final decision. ECHA received around 230 testing proposals for reproductive toxicity (Annex IX/X, 8.7.3.) for phase-in substances registered by the December 2010 deadline. These should have been examined and decided upon by December 2012. However, due to the fact that in 2012 the MSC was not able to find unanimous agreement on the use of the EOGRTS and the need for a second generation, no decisions were taken on these dossiers. Instead the parts of the dossier containing the reproductive toxicity endpoint were sent to the EC for decision making in the REACH Committee.

In November 2012 the Commission outlined their proposed approach to introduce the EOGRTS into the REACH regulatory framework. The approach involved inclusion of TG 443 in the Test Method Regulation (TMR) via the 5th Adaptation to Technical Progress (ATP) to the CLP Regulation, meaning *". . .a modification of the REACH Annexes IX and X to include a 'core' EOGRTS as the standard information requirement under point 8.7.3, and a 5 year review phase"*.[105] (see end of Section 7.2.2).

When looking at the discussion on the acceptance of the EOGRTS within Europe, two frames (lines of argumentation) can be found that dominate the discourse, i.e. the line of precaution and the line of innovation.

7.2.2.1 The precaution frame

The precaution frame is the result of a strict interpretation of the Precautionary Principle (as laid down in EU Commission Communication 2000)[106] and is driven by a perceived risk of chemicals as being involuntary and dreadful. This category of risks requires a high

105 http://ec.europa.eu/enterprise/sectors/chemicals/files/caracal/minutes-121128-29_en.pdf:consulted February 2014.

106 http://ec.europa.eu/dgs/health_consumer/library/pub/pub07_en.pdf :consulted February 2014.

level of protection from the government. In response regulators in the area of chemicals are stimulated to be very conservative in the decisions they take (Schiffelers et al., 2012: Chapter 5). Any change in the existing way of testing has an additional risk to it, often leading to a preference for *''the devil we know''* (Storer et al., 2010). In this case the risk avoidance is increased by the complexity of the reproductive cycle and the fear of calamities such as the thalidomide disaster in the 1960s. The precautionary line of argumentation is in this case mainly represented by ECHA, Sweden, Finland and to a lesser extent France. The main objection of these stakeholders when it comes to the EOGRTS is connected to skipping the second generation and thereby running the risk of missing out relevant information. According to an officer from the European Commission these stakeholders "*were very firm that the EOGRTS is not offering the information sought"* to safeguard the high level of protection on Human Health and Environment that REACH aims at in article 1.1 of REACH (EU, 2006). This was affirmed by ECHA in a statement to Nature: "*The two-generation study is the only study that covers effects on reproduction after exposure during all life-stages"* (Gilbert, 2011). An additional concern of these stakeholders is that the EOGRTS was originally designed for pesticides and not for industrial chemicals, while the retrospective analyses (see Section 7.2.1) are believed to cover insufficient industrial chemicals to offer a clear frame of reference for industrial chemicals. ECHA c.s. therefore emphasized that the two-generation study is the formal requirement in the REACH TMR (EU, 2008a). This was formulated as follows in the meeting minutes of April 2011 of ECHA's MSC: "*The current legal requirement under REACH is not the EOGRTS but the two generation reproductive toxicity study. ECHA needs to ensure with its decisions on testing proposals that this requirement is covered and the information expected from a two-generation reproductive toxicity study is available and adequate for the purposes of risk assessment and classification and labelling. Therefore, ECHA currently can accept the EOGRTS as a testing proposal of a registrant only if it is modified/specifically designed to cover the key parameters of a two-generation reproductive toxicity study (EU Test Method B.35)."*[107] The US and Canada, having to deal with a comparable legal situation for agrochemicals in which the second generation is legally required, solved this through guidance document 117 (the waiving guidance) (OECD, 2011b). It sets out internal triggers for those cases where a second generation is required by the Canadian Pest Management Regulatory Agency and by the United States Environmental Protection Agency, Office of Chemical Safety and Pollution Prevention. If these triggers are not found, the second generation can be waived. This transition regulation will be used for the intermediate period until new regulation including the EOGRTS is at force. The European OECD member states did not support guidance 117, according to one of the respondents, "*out of fear by some European member states that the goal of animal reduction would not be achieved."*

Additionally, the European regulation on classification, labelling and packaging of substances and mixtures requires that substances are classified as reprotoxic when proven that they interfere with the sexual function and fertility (EU, 2008b). Fertility is defined as the ability to mate and create offspring.

107 echa.europa.eu/.../meet_minutes_msc_17_en.pdf :consulted February 2014.

Rorije et al., argue that fertility can also be examined through parameters such as the sperm/follicle quality of the first generation (Rorije et al., 2011), but according to ECHA examining fertility requires an F2.[108]

Another point of concern of the representatives of the precautionary frame relates to the interpretation of the additional parameters that are generated by the EOGRTS. The additional information is according to ECHA only valuable for risk assessment purposes if it can be interpreted in a useful way. However, for many of these parameters no frame of reference exists, making it highly challenging from a regulators perspective to interpret the information. In addition, TG 443 is a flexible protocol which can be tailor made. This is on the one hand an advantage but on the other hand it also offers additional uncertainties in terms of how to conduct the test and how to deal with the additional data obtained. Holding on to the existing way of testing is the safest option for the parties doubting the usability of the EOGRTS. This is intensified by the fact that two generation test (TG 416) is a one-size-fits-all protocol with a long history of application and extensive experience in interpreting the results.

7.2.2.2 The innovation frame

The innovation frame in this case is represented within Europe by a group of EU MS (NL, UK, DE, DK and AU) together with several animal welfare organizations and industries. These stakeholders are advocates of the EOGRTS for a combination of scientific and animal welfare reasons which have been elaborated on in the Sections 7.1 and 7.2.1.

The EOGRTS advocates point to the fact that REACH offers discretion to regulatory authorities and manufacturers to choose the testing method best suitable for the job. Moreover, REACH favors alternative methods to conventional *in vivo* testing and article 25 (1) states that: *"In order to avoid unnecessary animal testing, testing on vertebrate animals for the purpose of this Regulation shall be undertaken only as a last resort"* (EU, 2006). Some of the member states have therefore already rejected draft registration dossiers through the MSC that still required a second generation. REACH in addition states that: "*the methods shall be regularly reviewed and improved with a view to reducing testing on vertebrate animals and the number of animals involved. The Commission, following consultation with relevant stakeholders, shall, as soon as possible, make a proposal, if appropriate, to amend the Commission Regulation on test methods. . .. so as to replace, reduce or refine animal testing-art. 13.2 REACH"* (EU, 2006).

It must be noted though that the group of advocates is not homogenous. Several MS for example are very much in favor of the EOGRTS procedure with the DIT/DNT cohorts (e.g. DE, DK and NL) while others like the UK are not.

108 To attest their point of view ECHA presented several substances that were perceived to cause unique effects in the second generation and therefore would only be classified and labeled as reprotoxic after examination of the F2. However, the RAC discussed some of these examples and concluded that there was no need for the C&L of these substances. http://echa.europa.eu/meetings-of-the-rac :consulted February 2014.

When it comes to the ARA of the EOGRTS, the representatives of the two lines of argumentation each have their own legal and scientific arguments which are brought forward repetitively. The trench between precaution and innovation lies at the center of the existing controversy and has led to fundamental disagreement in which various institutions and particularly their experts have taken firm positions on the issue. This resulted in rigid opinions on all aspects of the discussion and moving away from the existing positions, held the risk of lose face. All in all, the two lines of argumentation led to an impasse in the discussion in which *''They agreed to disagree''*.

The ball war thereby in the European Commission's court. The way the EC dealt with the impasse was by generating more information to reassure those stakeholders who do not feel at ease with skipping the second generation. This is done through an additional 5 year review on the value of the second generation for a selection of substances and through a period of parallel testing. *''Once enough EOGRTS information has been collected, it will be necessary to re-assess the test design.''*[109] There may not be a clear scientific base for the additional review phase but science alone was unable to build the bridge to escape from the existing deadlock. Gaining trust in and experience with the procedure became crucial at this stage.

Early 2014 the Commission decided that the basic EOGRTS, without an F2 study, is the preferred test method to achieve the standard information requirement under REACH. However, it recognizes the *''scientific uncertainty''* over the *''added value''* of the F2 generation and states that EOGRTS should include F2 tests for a *''certain number of substances for which significant exposure of relevant populations (consumers, professional users) occurs.''*[110]

With regards to the DNT and DIT cohorts the Commission proposed thatthe DNT/DIT tests should only be carried out in *''certain cases''*, due to *''technical, economic and practicality reasons.''*[111] The decision to leave out the DNT and DIT cohorts initially led to opposition from several stakeholders such as the Netherlands. It was seen as a missed opportunity in terms of protecting the population from potential harmful effect of chemicals.[112] The Commission therefore asked member states to suggest conditions (triggers) for inclusion of DNT/DIT cohorts. Once decided on, such conditions will be part of the amendment to the REACH annexes.[113]

109 http://chemicalwatch.com/18058/caracal-discusses-extended-one-generationstudy-in-reach :consult. June '14

110 http://chemicalwatch.com/18058/caracal-discusses-extended-one-generationstudy-in-reach :consult. June '14.

111 http://chemicalwatch.com/20219/eu-commission-notifies-wto-of-reach-amendment-for-eogrts:consult.June '14.

112 http://chemicalwatch.com/20219/eu-commission-notifies-wto-of-reach-amendment-for-eogrts:consult.June '14.

113 http://chemicalwatch.com/18534/eu-test-method-regulation-update-disappoints-animal-groups:consult.June '14.

7.2.3. The Use by Industry (UI) of the EOGRTS to comply with REACH

The adoption of the OECD Test Guideline 443 has given registrants the possibility to choose between the two-generation reproductive toxicity study (test method: EU TM B.35/OECD TG 416) and the extended one-generation reproductive toxicity study (OECD TG 443) to meet the REACH requirements in Annex IX or X 8.7.3. ECHA's position was that for the EOGRTS to meet the REACH information requirements (EU TM B.35), it had to include an extension of Cohort 1B to mate the F1 animals to produce the F2 generation, which were kept until weaning. Nonetheless, ECHA also stated that the EOGRTS can, under certain conditions, be suitable for a higher-tier study on a registered substance to fulfil the current information requirement in Annexes IX and X 8.7.3. There may be cases where registrants have specific information on properties of a substance explaining that the F2 in the EOGRTS is not necessary. The underlying scientific arguments can be used in a weight of evidence approach according to Annex XI, 1.2. of REACH to legitimize the adaptation of the standard information requirement (EU, 2006).

It is the responsibility of the registrant to present these arguments in their testing proposal. They can modify the registration dossier to include the test method they prefer to use for reproductive toxicity before they receive a draft decision from ECHA. These arguments will be deliberated on in the analysis of the testing proposal by ECHA and the successive decision making. In the next stage the registrants can comment on the draft decision for the testing proposal. In this case ECHA expects that registrants express their preference on the method they want to use, so that their preference can be considered during the decision making procedure. Some registrants actually already have proposed TG 443 in their registration dossier to meet the REACH information requirements on reproductive toxicity. However, due to the fact that the MSC could not come to an agreement, most of the test proposals from registrants for reprotoxic endpoints have been stacked up for an undefined period of time.

Furthermore, the Use by Industry of the EOGRTS in the context of REACH is influenced by the following factors: First of all, the EOGRTS protocol was originally developed for pesticides/agrochemicals. Agrochemicals need to comply to more test parameters than industrial chemicals, since they are designed to disrupt biological processes in the target species they intend to control. To meet these requirements the Cooper protocol included several endpoints (e.g. DNT and DIT) to make it an all-inclusive test protocol for agrochemical compounds. As a result the EOGRTS produces more, and more accurate, information when compared to the TG 416. This is on the one hand a scientific advantage but may at the same time be an economic drawback due to the fact that such new models are oversensitive and prone to false positives (Storer et al., 2010). Consequently, more compounds will not pass the safety criteria, which is an entrepreneurial risk in terms of the development/use of compounds. Consequently, a group of companies is not supportive of an EOGRTS that includes the DNT and DIT parameters, since it has a higher risk of hindering products from getting onto the market.

Secondly, the costs for the EOGRTS are estimated to be 2.5 times higher than the traditional two generation study. This estimation is based on information provided by industry and CRO's who refer to the costs of EOGRTS including the 2nd generation and the DIT/DNT cohorts (Cehtra, 2012). A survey conducted by Cehtra in request of ECHA speaks about an increase in costs of 41% if the EOGRTS is conducted with the second generation (Cehtra, 2012). Despite the difference in estimations the conclusion is that the EOGRTS leads to a substantial increase of costs considering the fact that a TG 416 costs about 500,000 Euro.[114] It should be notified however that the most expensive part of the EOGRTS testing concerns the performance of DIT and DNT cohorts.[115] [116] And the recent proposal of the EC excludes the DNT and DIT cohorts from the standard EOGRTS protocol.[117] For agrochemicals the EOGRTS normally is cost effective since both TG 416 and the Developmental Neurotoxicity study (TG 426) are often required for these substances. For industrial chemicals, however, it will results in increased test costs if all the additional parameters are required (Cehtra, 2012).

Thirdly, the EOGRTS is quite a complex and labor intensive testing procedure. Several labs have run trial studies and the results show that the procedure is feasible but complex (Schneider et al., 2010; Fegert et al., 2012). Moreover, the laboratory capacity to undertake the EOGRTS is still limited (Cehtra, 2012).

According to one of the respondents these practical/economical drawbacks were insufficiently taken into account at the stage of FI where *''for the sake of a swift adoption of the protocol, parameters were stacked on each other''*. From the side of agrochemicals the situation is more straightforward. The additional parameters targeted by the EOGRTS are required anyhow and the costs are comparable to or even lower than the required combination of TG 416 and TG 426. Moreover, several agrochemical companies have been involved in setting up the Cooper test protocol from the start and have thereby gained abundant experience with the EOGRTS. This combined with the higher standard of data and the use of substantially less animals has led to the fact that several agro-chemical companies have already taken the step of executing the EOGRTS to meet with US/EU test requirements. The US EPA and Health Canada in the meanwhile have already accepted two studies on pesticides submitted by Dow conducted with the EOGRTS.

However, the situation for industrial chemicals within Europe under REACH remains uncertain as has been described in Section 7.2.2 And as long as there is no clarity on how TG 443 needs to be conducted to meet the REACH information requirements, registrants will be cautious in using the EOGRTS protocol for the registration and authorization of their compounds, especially in those cases where higher costs are involved and the risk of false positives increases.

114 http://www.oecd.org/chemicalsafety/testing/46436593.pdf; page 487 :consulted June 2014.

115 http://echa.europa.eu/documents/10162/13578/meet_minutes_msc_23_en.pdf :consulted June 2014.

116 Especially the DNT is very expensive due to the embedding and histology of tissues and behavioral assessments.

117 http://chemicalwatch.com/20219/eu-commission-notifies-wto-of-reach-amendment-for-eogrts: consult.June '14.

7.3 Analyses

Section 7.2 reveals that the substage of ARA has been the most challenging part in the process of regulatory acceptance and use of the EOGRTS. To comprehend this situation it is important to disclose the interrelatedness of the drivers and barriers described above. For this purpose two analytic steps are taken.
Firstly, the influencing factors on regulatory acceptance and use are placed in the multi-level perspective on technology transitions which covers three main levels of influence on technology transitions (see text box for an explanation of this perspective). Secondly the connection between the substages of FI, ARA, and UI is analyzed to better understand the recent impasse.

7.3.1 The drivers and barriers from the multilevel perspective on technology transitions

As long as an innovation's value is disputed, it will face difficulties in getting accepted at the meso level to become part of the existing regulatory regime (see text box on multi-level perspective of technology transitions). Figure 13 illustrates this dispute with regards to the EOGRTS by displaying the contradicting pressures using the multilevel perspective on technology transitions.

THE MULTILEVEL PERSPECTIVE ON TECHNOLOGY TRANSITIONS

Regulatory acceptance and use of 3R models, such as the EOGRTS, is influenced by a combination of scientific, political, institutional, economic, ethical and social factors (Schiffelers et al., 2012: Chapter 5). Such system innovations are hardly ever the effect of a single cause but the ''result of the interplay between many factors and actors''(Geels, 2006). Therefore, an integrative approach is needed to understand such processes. The multilevel perspective on technology transitions, developed by Schot and Rip (1996) offers such an integrative approach. This perspective addresses three levels of influences which play a role in technology transitions (Kemp, 2010):

- The micro- or niche level where innovations are developed and validated;
- The meso- or sociotechnical regime level covering the existing rules and regulations, expertise, dominant practices and the standing institutions;
- And the macro- or sociotechnical landscape level covering the material infrastructure, existing political culture and coalitions, social values, the macro-economy, demography and the natural environment.

For a new technology to become accepted and used, developments at these three levels need to reinforce each other before a shift in favor of the new technology can occur. In other words, an alignment or conjunction of the three levels can only take place if an innovation (e.g. a 3R model) fulfils the needs of the meso- and the macro level. As long as the innovation does not fulfil these needs it will be unable to escape from the niche where it was developed (Kemp, 2010) as has been the case for several 3R models. Furthermore, the distinction between the levels can assist in defining the factors that are more suitable to start working on in order to improve the acceptance of the 3R model, since factors at the micro- and partly the meso level usually offer better possibilities for change than the broader societal developments at the macro level.

The drivers at the micro level proved strong enough for the EOGRTS to become accepted at the OECD level. Yet the actual regulatory acceptance needs to be effectuated at the European level and at the level of individual countries such as the US and Canada. In the context of REACH, the drivers at the micro level were not convincing enough for all parties. The underlying cause is the difference in the frames of reflection embraced by the two opposing groups. The advocates of the EOGRTS focussed on the advantages of the innovation while the advocates of TG 416 underlined the uncertainties connected to the switch to new way of testing. And since consensus is needed within ECHA's MSC, the EOGRTS repeatedly bounces back to the micro level where it's suitability was disputed and put to the test. This for example, is illustrated by the decision of the Commission to conduct an additional review phase performing the 2nd generation for a limited number of substances.

The colliding factors at the micro level have led to a policy controversy. "*Such disputes are resistant to resolution by appeal to facts or reasoned argumentation because the parties' conflicting frames –i.e. the innovation versus the precautionary frame – determine what counts as a fact and what arguments are taken to be relevant and compelling"* (Schön and Rein, 1994). While in policy disagreements, *"the parties to contention are able to resolve the questions at the heart of their disputes by examining the facts of the situation"* the parties are no longer able to do so when it comes to policy controversies (Schön and Rein, 1994).

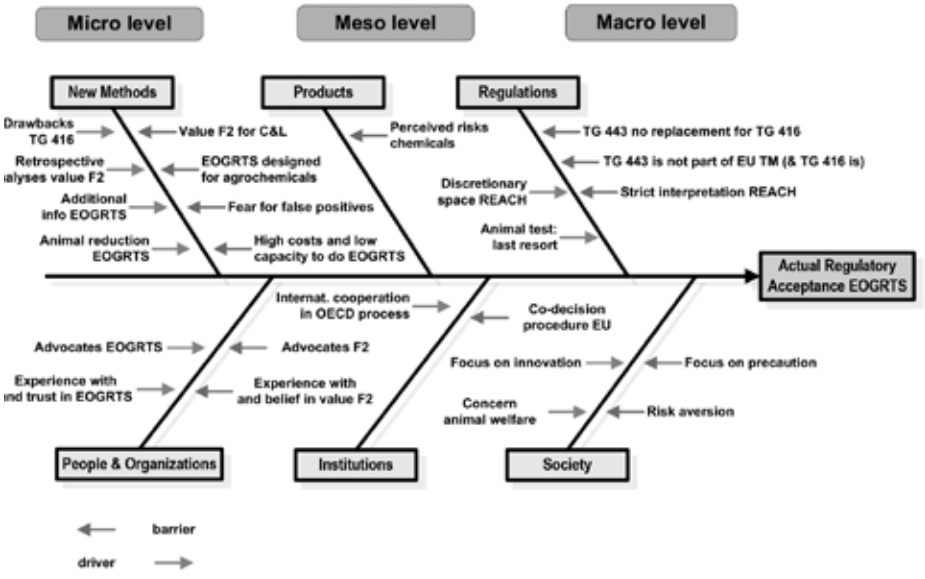


Figure 13. Multilevel perspective on drivers and barriers influencing regulatory acceptance and use of the EOGRTS within the context of REACH (Source: based on Schiffelers et al., 2012)

The formal discussions are primarily of a scientific and/or legal nature. However, the scientific and legal arguments used by both parties tend to disguise the underlying interests, uncertainties, sympathies and antipathies. From the interviews it became clear that the controversy is fed by dissimilar interests and values of the stakeholders involved. Moreover, the different individual experts represent unalike levels of influence, expertise and thereby trust in the EOGRTS and different levels of responsibility in case anything goes wrong. Off the record, respondents were sometimes willing to refer to aspects such as perceived lack of expertise and experience, reputational issues, the fear of lose face and clashing personalities. Such psychological and institutional aspects strongly influence the existing impasse but are normally only addressed in a private manner.

7.3.2 The connectedness of the substages FI, ARA and UI

Although the substages of FI, ARA and UI have their own specific drivers and barriers, the stages are strongly connected and barriers and drivers at one stage also influence the other stages. If we want to understand why consensus was hard to reach at the stage of ARA it is essential to take the drivers and barriers at the stage of FI into account. Several respondents specified that the swiftness of the process of FI slowed down progress at the stage of ARA. The doubts that played a role at the stage of FI resurfaced at the stage of ARA. Several issues were unsolved at the stage of FI and were transferred to the stage of ARA to be decided upon. Some respondents suggested that parties that had doubts about the EOGRTS only accepted TG 443 because they knew that the actual acceptance needed to be effectuated at the European level where they would have chance to reopen the discussion.

The discussion during the OECD process was predominantly a scientific one. The influence of other aspects such as feelings of discomfort and lack of experience with the proposed protocol and impracticalities of the model in terms of costs and operationalization, may have been partially underestimated. The heritage of the stage of FI is perceived to have had a big influence on the stage of ARA, while the substage of ARA on its turn influences the sub-stage of UI. As long as there is uncertainty about the ARA and the way the EOGRTS needs to be conducted, UI will be delayed. However, frontrunners within industry that proposed the EOGRTS despite these uncertainties, have on their turn also influenced the ARA by pressurizing the system to come to a decision. In other words the substages of regulatory acceptance and use are largely connected. For a 3R model to become accepted and used for regulatory purposes it is therefore very important that the involved stakeholders anticipate on the criteria/obstacles a 3R model might face in the other substages.

7.4 Discussion

Despite the fact that the OECD adopted TG 443 in July 2011… "a *quick implementation of the EOGRTS has been hampered by disagreement among experts on the value of information obtained from the 2nd generation (F2) and the extended debate whether the DNT/DIT cohorts should be included in the default study design in REACH and its relevance for*

the assessment when the F2 generation is not performed."[118] The challenging implementation process within the European context is thought-provoking since directive 2010/63/EU states in article 13.2 that: *". . .in choosing between procedures, those which use the minimum number of animals shall be selected"*. In addition recital 11 of the directive states that *''Where no alternative method is recognized by the legislation of the Union, the numbers of animals used may be reduced by resorting to other methods and by implementing testing strategies, such as the use of in vitro and other methods that would reduce and refine the use of animals.''* (EU, 2010). Furthermore, ECHA's board of appeal declared in a recent decision to annul an animal study that; *". . .Directive 2010/63/EU on the protection of animals used for scientific purposes cannot be treated as directly imposing any obligations on ECHA, but the latter's actions should not run counter to the principles – 3R – laid down therein. . .."*.[119] In this light the challenging acceptance of 3R models such as TG 443 begs the following final question; *'What is needed to augment the current process and what can be learned from this case study for future processes?'*

For 3R models such as the EOGRTS to really enter the area of risk assessment it is important that the conflicting parties reflect on the existing frames and the connected disputes. Reconsideration is needed whether these are really opposing or to a certain extent also appending arguments. Such a reflection on the existing frames (Schön and Rein, 1994) may well reveal that the parties share many mutual interests. All of the involved stakeholders for example adhere to the general principle that science should be at the basis of the decision making process and they all aim for risk minimization with regards to chemical substances. The precautionary frame and the innovation frame may therefore be less worlds apart as they at first might appear. New protocols such as the EOGRTS offer possibilities to simultaneously improve the level of innovation and of precaution. This however, requires that all parties get the chance to build experience with the new protocol and investigate its -dis-advantages for a certain period. This way the seemingly conflicting frames can slowly merge into one shared frame. Reframing though requires neutral policy entrepreneurs (Bryson and Crosby, 1992) with sufficient mandate to bring the conflicting parties together to reconsider the present situation and formulate a strategy to blend the diverging interests. When it comes to the EOGRTS this role best fits the EC.
Furthermore, it requires shared Commitment of all parties involved to a joint policy goal, intense Communication between the parties on diverging and mutual interests -which took place during the OECD process, but some interests were parked for the sake of speed of the process, enduring Cooperation between all stakeholders and a strong process Coordination, which was well arranged during the OECD process but was taken up quite late in the European context. These 4C's (Schiffelers et al., 2014b: Chapter 8) are pivotal to transform slow acceptance processes, in which non-decision making lurks, into a proactive process in which all parties work towards a clear defined policy goal.

118 http://chemicalwatch.com/18058/caracal-discusses-extended-one-generationstudy-in-reach : consulted June 2014.

119 http://echa.europa.eu/documents/10162/13575/a_005_2011_boa_decision_en.pdf :consulted June 2014.

Buying some additional time through an additional 5 year review (see Section 7.2.2) might in our opinion in this case well be a wise move from the EC to let all parties get used to the new procedure and reframe the discussion. In this process it can become clear how and when the EOGRTS should be used. In the meantime the Commission can evaluate the existing positions and work towards a stepwise strategy to overcome the remaining controversies.

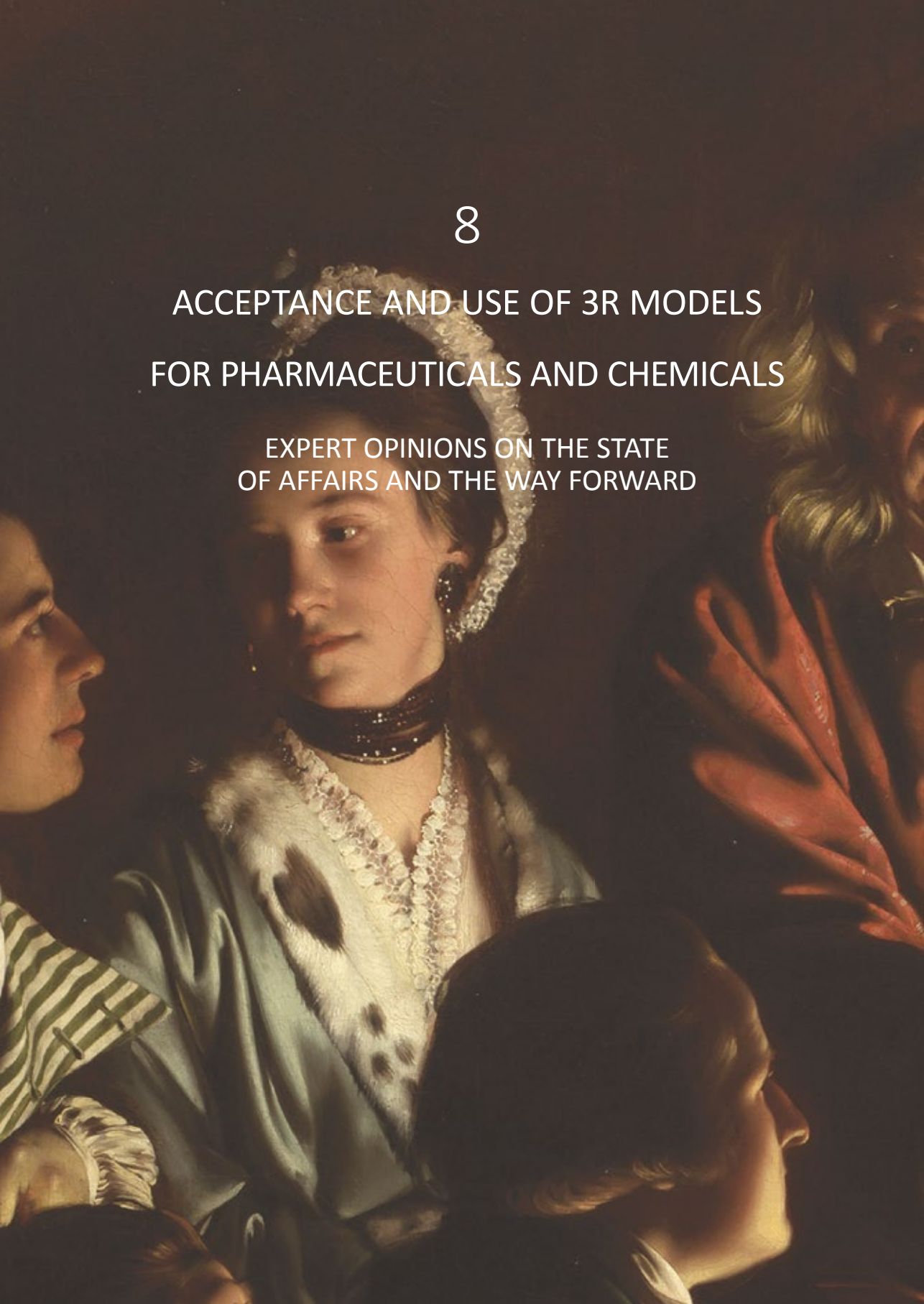
In short, the following recommendations are made to enhance the process of regulatory acceptance and use of the EOGRTS and similar future processes:

- The EOGRTS case reveals that the combination of profuse scientific information and strong advocates are important ingredients for the regulatory acceptance of a 3R model. However, it also reveals that regulatory acceptance is often a highly politicized process in which science can become part of the existing disagreement and in which other arguments e.g. the lack of experience with and trust in the new model, institutional agenda's and political realities are seen to regularly outweigh the scientific 'facts'.
- These 'other' arguments therefore continuously have to be taken into consideration. This means that a permanent anticipation is needed on the legal, practical and psychological requirements for the 3R model to proceed to the following substages of ARA and UI (e.g. legal issues regarding the existing regulatory frame, costs connected to the new protocol, capacity needed to perform it, trust in the innovation and knowledge required to work with it). A regular meta-communication is important to discuss the remaining doubts and existing interests and to identify the criteria needed to merge the remaining diverging interests.
- All stakeholders need to be aware of the fact that the three substages of FI, ARA and UI are largely connected. The heritage of a previous stage is likely to remain of influence at the subsequent stage.
- The current implementation process needs reflection on the existing frames and on the mutual and conflicting aspects within these frames.
- To guide a 3R model through the substages of regulatory acceptance and use, a neutral process manager is needed to set out a clear strategy, steer the process, keep all parties involved and be sensitive to the different interests at stake. Clearly, there is a role for the EC in bringing together the views of the different European stakeholders on such issues.
- Commitment needs to be asked by the EC from all European stakeholders in light of Directive 2010/63/EU in terms of stimulating the use of 3R models.

In short, mutual trust, shared commitment, a strong process coordination and close cooperation and communication are needed to enhance the process of regulatory acceptance and use of 3R models in general and the EOGRTS in particular (Schiffelers et al., 2014b: Chapter 8).

Finally, when it comes to reproductive toxicity testing, the EOGRTS should not be considered as a stand-alone procedure. The EOGRTS is an important step in terms of animal welfare and scientific progress. However, since many chemicals need to be tested for reproductive toxicity the step from the two generation study to the EOGRTS will not solve

the testing bottleneck for the many chemical substances in commerce. For this, a paradigm shift is needed from extensive animal testing to efficient and focused animal and *in vitro* testing. This other way of thinking combines clever ways of chemical prioritization with intelligent testing strategies in which different reproductive toxicity testing methods are combined, as for example proposed by Schaafsma et al. (2009), Spielmann (2009) and Martin et al. (2011). Such intelligent testing strategies in the end offer the best potential to minimize *in vivo* reproductive toxicity testing.



8

ACCEPTANCE AND USE OF 3R MODELS FOR PHARMACEUTICALS AND CHEMICALS

EXPERT OPINIONS ON THE STATE OF AFFAIRS AND THE WAY FORWARD

“Ideas are like rabbits. You get a couple and learn how to handle them, and pretty soon you have a dozen.”

John Steinbeck
American author and Nobelprice winner
1902-1968

A short guide to Chapter 8

This chapter consists of a workshop report published in 2014 in *Regulatory Toxicology and Pharmacology.* Through Chapter 8 the research question Q3a, Q3b and Q4 are directed (see section 1.2). To answer these questions, two expert panels were organized, one with 20 experts from the pharmaceuticals field and one with 20 experts from the field of chemicals. For the sake of the completeness of this publication a certain amount of replication of chapters 1 to 4 was necessary.

Abstract

Pharmaceuticals and chemicals are subjected to regulatory safety testing accounting for approximately 25% of laboratory animal use in Europe. This testing meets various objections and has led to the development of a range of 3R models to Replace, Reduce or Refine the animal models. However, these models must overcome many barriers before being accepted for regulatory risk management purposes. This paper describes the barriers and drivers and options to optimize this acceptance process as identified by two expert panels, one on pharmaceuticals and one on chemicals. To untangle the complex acceptance process, the multilevel perspective on technology transitions is applied. This perspective defines influences at the micro-, meso- and macro level which need alignment to induce regulatory acceptance of a 3R model. This paper displays that there are many similar mechanisms within both sectors that prevent 3R models from becoming accepted for regulatory risk assessment and management. Shared barriers include the uncertainty about the value of the new 3R models (micro level), the lack of harmonization of regulatory requirements and acceptance criteria (meso level) and the high levels of risk aversion (macro level). In optimizing the process commitment, communication, cooperation and coordination are identified as critical drivers.

Reference:[120]
Schiffelers, M.J., Blaauboer, B., Bakker, W., Beken, S., Hendriksen C., Koëter, H. and Krul, C. (2014). Regulatory acceptance and use of 3R models for pharmaceuticals and chemicals: expert opinions on the state of affairs and the way forward.
***Regul Toxicol Pharmacol* 69, 41-8.**

120 Acknowledgements: We would like to thank in alphabetic order the workshop participants: Jeffrey Bajramovic, Jurgen van Belle, Joep Bergers, Peter Bertens, Jos Bessems, Johan Descamps, Jackie van Gompel, Betty Hakkert, Ingrid Hartgers-Pools, Carla Herberts, Erik Houthoff, Etje Hulzebosch, Jaap Keijer, Dinant Kroese, Kees van Leeuwen, Dick Lindhout, Fabrizio di Mattia, Aldert Piersma, Raymond Pieters, Henk Reinen, Ivonne Rietjens, Frans Russel, Kris Siezen, Gerard Stijntjes, Chantal Smulders, Rob Vandebriel, Frank Verheijen, Marieke Verheijden, Miriam Verwei, Irma Vijn, Gerrit Jan Wennink, Joop van den Wijngaard, Lonneke Wilms, Ben van der Zeijst and Maaike van Zijverden. We thank the Doerenkamp-Zbinden Foundation and the Dutch Ministry of Economic Affairs for funding this project. We thank the Dutch Ministry of Health, Welfare and Sport for facilitating the panel discussions.

8.1 Introduction

Test methods used for risk assessment purposes depend heavily on animal models which were developed over the last 50–60 years (Scholtz et al., 2013) and the animal model in this field is often still perceived as the ''gold standard''. This holds true for both regulators and industry (Scheel and Brekelmans, 2007). Nonetheless, a growing number of models to Replace, Reduce or Refine animal tests (3R models) (Russell and Burch, 1959) has become available. Every so often, these models are scientifically more robust, economically advantageous and ethically preferable in comparison to the existing animal model. Still, regulatory acceptance and use is one of the main challenges 3R models face (Richmond, 2002; Garthoff, 2005; Bottini et al., 2008; Schiffelers et al., 2012: Chapter 5). And until now alternative approaches have only rarely been used in regulatory settings (Scholtz et al., 2013).

There is growing international awareness of the slow regulatory acceptance of 3R models. In this context two ad hoc expert panels (see also Section 8.3) were set up in a combined initiative of TNO (Netherlands Organization of Applied Scientific Research), USBO (Utrecht University School of Governance), the NKCA (Netherlands Knowledge Centre on Alternatives) and the Dutch Ministry of Health to address the following key questions:

- *What are the main factors influencing the regulatory acceptance and use of 3R models for the safety/efficacy testing of pharmaceuticals and the safety testing of chemicals?*
- *How can the involved stakeholder groups contribute to optimizing this process?*

The experts were invited based on their affiliation with the product sectors pharmaceuticals and/or chemicals and because of their familiarity with the subject of the 3Rs. The distinction between this product sectors was made, based on the assumption that the influences on regulatory acceptance and use of 3R models potentially differ between these sectors. For process optimization purposes, representatives of public and private partners of the development chain from R&D to regulatory approval of pharmaceuticals and chemicals, were invited (for more information on the expert panels see Section 8.3). Section 8.4 of this chapter is a reflection of the opinions and ideas that were brought up during the panel discussions. Section 8.5 consists of an analysis of these findings and of identifiable actions per stakeholder group.

With this report the authors intend to offer a constructive contribution to the international discussion on regulatory acceptance and to stimulate this process where possible.

8.2 The multilevel perspective on technology transitions

To understand the overall acceptance process of innovations like 3R models, the multilevel perspective on technology transitions was presented to the experts of both panels (Schiffelers et al., 2012: Chapter 5). This multidisciplinary approach offers valuable concepts for the analysis of long-term technological transitions (Schot and Rip, 1996, Geels, 2006). For innovations to break through the following three levels need alignment (Schiffelers et al., 2012: Chapter 5):

- The micro level consists of the niche in which innovations such as new test methods are developed and tested. Here drivers and barriers are found relating to the development and validation of 3R models;
- The meso level entails a mix of existing rules and regulations, expertise, practices and institutions that strongly influence the acceptance of innovations like 3R models;
- The macro level where broader societal features, like the existing material infrastructure, the political culture and coalitions, broad social values, world views, the macro-economy, demography and the natural environment, can be found (Kemp, 2010).

Every level offers a part of the explanation why innovations like 3R models face difficulties in becoming accepted. Additionally, the technology transition approach unveils the-inter-dependencies between the three levels and thereby acknowledges the importance of combining societal and technical factors when examining and/or stimulating the acceptance of 3R models. Finally, the categorization of drivers and barriers into these three levels is significant because the level also gives information on the possibility to control a certain driver or barrier. Generally speaking the factors at the micro level offer more control possibilities than those at the meso- or the macro level.

8.3 Methodology

In spring 2012 two expert panels were organized which involved a total of 40 Dutch and Belgian experts[121] from the fields of safety assessment, regulatory testing and 3R models. Both the pharmaceuticals and the chemicals panel included a total of 20 experts. The participants derived from the following three stakeholder groups:[122]

- Regulatory authorities, legislators & policy makers
- Industry
- Academia & research organizations

The panel members have contributed to the discussion 'in a private capacity' making use of their expertise and experience as a professional within their specific stakeholder group. The choice for these three stakeholder groups is based on the assumption that these are the central chain partners for regulatory acceptance and use of 3R models. Both panels aimed at the clarification of the process of regulatory acceptance and use of 3R models and at the examination of possibilities to enhance this process. These goals were targeted through the following three subsequent steps:

121 The selection of Dutch and Belgian experts might have led to a certain level of bias since these countries are known to be relatively open to the 3Rs in comparison to certain other countries. However, the fact that most of the experts operate in an international context and the fact that all experts were invited to bring forward the dominant drivers and barriers on the acceptance and use of 3Rs for their sector from an international perspective, reduces in our opinion the risk of bias.

122 For an overview of experts see acknowledgements in short guide to chapter 8. In addition several of the authors of this manuscript (i.e. Blaaubloer, Beken, Hendriksen en Koeter) were involved in the panel discussions as experts. Krul and Schiffelers facilitated the panels.

Firstly, an inventory of barriers and drivers was made. For this purpose each participant was asked to write down the three barriers and drivers on regulatory acceptance and use of 3R models which they perceived to be most influential. This resulted in a broad range of factors which were grouped in about 25 clusters of comparable factors. The clusters of factors were checked and discussed in plenary and divided into factors at the micro-, meso- and macro level.

Secondly, a further prioritization was made of the factors in terms of their influence on regulatory acceptance and use. Each participant was asked to score the clusters of factors in terms of their perceived influence in the process of regulatory acceptance and use of 3R models within their product sector. For this purpose, each participant was asked to divide a total of 5 points between the factors they perceived to be most influential on the process of regulatory acceptance and use. This exercise resulted in an overview of those factors with the highest panel scores. In other words, these drivers and barriers were perceived by the panel to be most influential in that particular product sector (see Table 2).

Thirdly, actions were identified that can be pursued by the stakeholder groups in order to optimize the process of regulatory acceptance. This identification took place through a discussion within and between the stakeholder groups on the following 3 questions:

1. *Which factors can be influenced by the own stakeholder group?*
2. *How can these factors be influenced/what are possible actions?*
3. *What can a particular stakeholder group offer to chain partners and what is needed from chain partners in terms of optimizing the process?*

To ensure the representativeness of the findings, the results of previous research in this field was taken onboard and tested during the panels (Schiffelers et al., 2007, 2012: see appendix I and Chapter 5).

8.4 Results

This section starts with an overview of factors influencing regulatory acceptance and use of 3R models in the pharmaceuticals and chemicals sector (Table 2). This summary is followed by a description of the barriers and drivers per transition level (see Section 8.2) and a brief comparison between the two sectors involved.

8.4.1 Main influencing factors

Table 2 displays the top 7 of influencing factors on regulatory acceptance and use as perceived by the two panels. The table shows both the barriers and drivers with the highest scores in terms of their impact on regulatory acceptance and use of 3R models.

Table 2. Overview of perceived dominant factors influencing regulatory acceptance and use of 3R models in the pharmaceuticals and chemicals sector

Factor	Pharmaceuticals panel	level	← barrier → driver	Factor	Chemicals panel	level	← barrier → driver
1	*Insufficient harmonization of legislation*	*meso*	←	1	*Challenging in vitro - in vivo extrapolation*	*micro*	←
2	*Uncertain predictability 3R model / Lack of validated 3R models*	*micro*	←	2	*Lack of global harmonization & Mutual Acceptance of Data*	*meso*	←
3	*Cooperation (including data sharing) & communication between stakeholders*	*all 3*	→	3	*Lack of concrete policy goals to stimulate the 3Rs*	*meso*	←
4	*3R models provide more mechanistic information*	*micro*	→	4	*Insufficient attention for probabilistic design in entire chain*	*all 3*	←
5	*Early involvement regulators to discuss acceptance criteria*	*micro meso*	→	5	*Current thinking in terms of hazard instead of risk*	*meso*	←
6	*Implementation of directive 2010/63/EU -on the protection of animals used for scientific purposes-.*	*meso*	→	6	*Cooperation & communication between stakeholders (including data sharing)*	*all 3*	→
7	*Risk-averse society*	*macro*	←	7	*Difficult accessibility regulators to discuss acceptance criteria*	*micro meso*	←

The top 7 factors in the pharmaceuticals panel received 76% of the available points. The top 7 factors in the chemicals panel received 51% of the available points. This means that within the pharmaceuticals panel there was a higher level of consensus about the main factors influencing the process of regulatory acceptance and use. In the chemicals panel many factors received just one or two points illustrating that the opinions in this panel were more dispersed. Furthermore, the drivers in the pharmaceuticals panel are seen to outweigh the barriers, while the barriers are seen to outweigh the drivers in the chemicals panel. This observation is in line with the observed difference between the two sessions in terms of perceived chances to optimize the process.

8.4.2 Cross-sectorial barriers

In the overview the following barriers can be found in both expert panels.

- Macro level: the striving for risk minimization
- Meso level: the lack of harmonization of legislation and test requirements
- Micro level: the existing uncertainties of 3R models

At the macro level societies' striving for risk minimization is identified by both panels as an important barrier for the acceptance and use of 3R models for assessment purposes of pharmaceuticals and chemicals (even though the level of risk aversion is observed to differ between these sectors: see Section 8.4.4). Societies' response to unpredictable risks is to try to prevent, minimize and channel them, e.g. by trusting this task to regulatory authorities (Malyshev, 2006). The regulators have to deal with a high level of responsibility in terms of general health protection on the one hand and uncertainties regarding products and test methods on the other. A noticeable response to this challenging combination of tasks is sticking to familiar routines. As a result "*Rules–or procedures-become frozen in place and cannot readily adapt to changing scientific knowledge"* (Breyer, 1993; Schiffelers et al., 2012: Chapter 5). This reaction blocks 3R models from getting accepted for regulatory purposes.

At the meso level the lack of global harmonization of regulatory requirements is identified as a central barrier. The regulatory arena across sectors and countries is very complex, which is highly challenging for those who seek to reduce the numbers of animals used for regulatory purposes (Seidle et al., 2010). Harmonization of international legislation is therefore a prerequisite for regulatory use of 3R models by industry. Without global harmonization, manufacturers, being global players, are not stimulated to invest in the transition from an existing animal model to a 3R model. However, harmonization is a complex process since it crosses geographical, cultural and institutional borders. In the field of pharmaceuticals increasing efforts are made to harmonize the requirements for pharmaceuticals within the context of the ICH and the VICH, but many regulatory requirements, i.e. for the batch release testing of biologicals, are still set by European/national pharmacopoeias. Within the chemicals sector harmonization is driven by the OECD guidelines which provide a collection of the most relevant internationally agreed test methods to determine the safety of chemicals and chemical preparations. These guidelines are very important in terms of international standard setting. However, it must be kept in mind that the OECD guidelines are non-binding and it remains under the jurisdiction of European and national governments whether and how these guidelines are transposed into their legislation. And even though the OECD is a frontrunner in the acceptance of non-animal tests and the reduction of animal testing through Mutual Acceptance of Data (MAD), the technical guidance documents still predominantly rely on animal assays (Arts et al., 2008).

Apart from harmonization within a sector, cross-sectorial harmonization is important, since compounds are used often within different product sectors e.g. pharmaceuticals, agrochemicals, biocides and chemicals (Seidle et al., 2010).

At the micro level experts in both panels stipulate that the remaining uncertainty about the performance of 3R models is an important barrier for many 3R models when trying to enter the existing regulatory regime (meso level). The uncertainty, according to the experts, is mainly connected to the uncertain predictability of 3R models and the challenging *in vitro–in vivo* extrapolation. Both aspects refer to the difficulty to translate test results of especially *in vitro* models to the biological effects in humans or animals for which they need to provide information. According to several experts the paradox here however is that *in vitro* models are often expected to meet criteria that were never met by most animal tests, as is also described by Spielmann (2000). While many animal models are observed to face the same translational problems when it comes to the extrapolation of the test results to humans (Arts et al., 2008; Langley, 2009; Martic´-Kehl et al., 2012).

8.4.3 Cross-sectorial drivers

The following factors were identified by both panels as important drivers for regulatory acceptance and use of 3R models.

1. Macro level: The ethical concern about animal testing.
2. Meso level: Concrete policy goals/legislation to stimulate the 3Rs.
3. Micro level: The informative and mechanism-based character of 3R models.

The ethical concern about animal testing, did not end up in the top 7 (Table 2) but formed a relevant part of the discussion in both panels. For this reason this driver is also taken onboard in this subsection. In addition, both panels referred to the crucial role of communication and cooperation which play a role at all three levels of the transition process. These factors will be elaborated on in Section 8.5 where the focus lies on enhancing the process.

A central driver at the macro level is the ethical concern of society for laboratory animals. This factor is referred to as a catalyst for the process of regulatory acceptance and use. Even though society's attention for animal welfare is observed to fluctuate, and the ethical argument on its own is not enough, the combination of ethics with the striving for better science has the potential to form a powerful motor for change (Punt et al., 2011). However, the power of this factor is observed to remain secondary to the power of risk minimization. It then depends on the level of risk aversion in a sector to which extent the ethical arguments can become a driving force (see also Section 8.4.4 sectorial differences).

At the meso level, the driving force of horizontal legislation[123], such as European Directive 2010/63/EU on the protection of animals used for scientific purposes (EU, 2010) is brought forward by experts in the pharmaceuticals panel. Article 13 of Directive 2010/63/EU states that: "*Member States shall ensure that a procedure is not carried out if another method or testing strategy for obtaining the result sought, not entailing the use of a live animal, is recognized under the legislation of the Union.*"

123 Horizontal legislation pertains to animal experimentation and multilateral agreements in general. Vertical or sectorial legislation regulates the activities of a particular sector. Horizontal legislation needs to be taken into account by sectorial/ product legislation (Schiffelers et al., 2007: Appendix I).

The experts in the pharmaceuticals panel state that the directive has a stimulating effect on the incorporation of 3R models into guidance documents such as the monographs of the European Pharmacopoeia. The European Pharmacopoeia Commission (EPC) for example decided at its 141st Session – November 2011 – that the Groups of Experts should review all monographs and chapters prescribing animal testing in line with Directive 2010/63/EU.[124] This decision has had a stimulating effect on the formal incorporation of 3R models into the monographs and the first results in this direction are starting to become visible (see press release 142nd session of EPC).[125]

In the chemicals panel the legislative discussion regarding animal testing and the use of 3R models is dominated by REACH. Article 13(4) of REACH stipulates that toxicological tests shall be carried out in compliance with EU Directive 2010/63/EU and animal testing should only be used as a last resort (EU, 2006). And where possible, 3R approaches that are already specified in the REACH Regulation, should be applied.

At the micro level, the scientific value of 3R models is stipulated, especially by the experts in the pharmaceuticals panel. 3R models are valued for their informative character about the mechanism of action, their reproducibility and their scientific robustness. This makes them attractive to work with, particularly in those cases where the existing animal model is more or less a black box and shows considerable variability in test results.

8.4.4 Sectorial differences

The following sectorial differences, in terms of factors influencing regulatory acceptance and use, are identified:

At the macro level the striving for risk minimization is observed to play a more dominant role for chemicals than it does for pharmaceuticals. The higher the level of risk aversion connected to a product group, the more cautious the involved stakeholders will be towards implementation of 3R models that involve a level of uncertainty. The accepted risk for industrial chemicals is close to zero as a result of the involuntariness of being exposed to these compounds (Schiffelers et al., 2012: Chapter 5).

At the meso level this has its effect on the regulatory requirements for chemicals that are very strict and aim for zero or negligible risk levels (Kasamatsu and Kohda, 2006). This is underlined by the chemicals experts and is substantially different when compared to pharmaceuticals. Pharmaceuticals are generally looked at from a risk–benefit perspective leading to a certain level of risk tolerance. The zero risk tolerance in the chemicals sector often leads to the so called 'tick box approach', which refers to conducting or requiring every test described in the regulatory requirements. This approach is detrimental to the acceptance and use of 3R models. The experts in the chemicals panel expressed the need for a more realistic and functional risk based approach, which takes the level of relevant exposure levels into account. An important example of this more sophisticated

124 http://www.edqm.eu/en/Report_3R_meeting-1550.html :consulted March 2013.

125 http://www.edqm.eu/en/Achievements-of-the-PhEur-Commission-for-3Rs-1533.html :consulted March 2013

way of safety assessment is described in the trend breaking document *Toxicity Testing in the 21st century: a vision and a strategy* which builds on high-throughput screening methods and computational toxicology to disclose 'toxicity pathways' (NRC, 2007). This way of thinking is also reflected in the desire to adopt a more probabilistic approach throughout the risk assessment process which would *"enable variation and uncertainty to be quantified, mainly by using distributions instead of fixed values in risk assessment..."*[126] The existing way of risk assessment will need serious reconsideration before it will be ready to make use of this type of test data. And legislators need to update the existing legislation or have to build in flexibility in the interpretation of the existing legislation to make them suitable for this paradigm shift (Krewski et al., 2009).

At the micro level the profit margins for chemicals and for veterinary medicines are smaller compared to those for human use. The profit margins influence the level of innovation, meaning that for chemicals and veterinary medicines there is less room to invest in new tests such as 3R models. The difference in investment in research and development (R&D) between the pharmaceuticals and the chemicals sector is underlined by the 2012 EU Industrial R&D Scoreboard. In the EU alone the pharmaceutical and biotechnological industry's R&D share is 15% of the total European investment in R&D, whereas the chemicals industry accounts for 5% of the total. In the US the difference is even bigger; 23% versus 3%.[127]

8.5 Enhancing the process

Enhancing the process of regulatory acceptance and use of 3R models requires the alignment of the three levels of the multilevel perspective to get a 3R model out of the micro level where it was developed and into the existing regulatory regime (meso level). This alignment necessitates an intelligent process management, which involves a smart combination of the core tools of the 4C's as will be described in Section 8.5.1. The recommendations presented in Section 8.5.1 mainly derive from literature and earlier research of the authors (Hendriksen, 2000; Schiffelers et al., 2007, 2012: see Appendix I and Chapter 5). Where available, additional considerations are included that stem from the expert panels. In Section 8.5.2 the actions per stakeholder group are summarized as defined by the experts to lift a 3R model out of the micro level. These actions are based on inter-subjectiveness, i.e. mutual understanding between the experts from the different stakeholder groups during the panels.

8.5.1 4C's to align the micro-, meso- and macro level

Where scientific excellence is seen as the basis for the acceptance of a 3R model, several process principles are critical to get a 3R model out of its niche and into the existing regulatory regime. Hendriksen earlier on already introduced the "3Cs" of common sense, commitment, and communication as the basic process principles (Hendriksen, 2000). We have adjusted and elaborated on this concept and introduce the concept of the "the 4C's" which stands for commitment, communication, cooperation and coordi-

126 http://www.eufram.com/probablistic.cfm :consulted March 2013.

127 http://ec.europa.eu/dsg/jrc/ : consulted March 2013.

nation. These 4C's are considered of key importance to align the developments at the macro-, meso-, and micro level and thereby enable this 3R model to become accepted for regulatory purposes.

8.5.1.1 Commitment

Although commitment to 3R models shows regional differences, there seems to be a growing international commitment to 3R models for ethical, economical and/or scientific reasons, as is refected in Directive 2010/63/EU. Commitment is on the one hand an enduring positive attitude towards the 3Rs but it also implies the allocation of resources, in terms of money and people, to develop, validate and implement 3R models.

8.5.1.2 Communication

This principle refers to the need to establish and maintain a constructive dialog between all stakeholders (Richmond, 2002). Communication should start at an early stage between model developers and model users and should continue throughout the chain in order to facilitate the process of validation and implementation of 3R models. According to the participants communication is progressively getting more attention at several levels. Nonetheless, it can be intensified or ameliorated in different ways.
Current communication still shows a fairly dispersed pattern, whereas frequent and sequential communication throughout the process is needed between academia, industry and regulators to:

- Identify the needs of end users in terms of model development and improvement;
- Identify the regulatory criteria that a 3R model has to meet;
- Share experience and expertise on working with 3R models.

During the expert panels, industry urged for a timely communication and interaction with regulatory authorities to deal with the uncertainty whether a 3R model will be accepted for licensing- or market authorization purposes. With regards to pharmaceuticals, an adapted drug development program is often discussed within the frame of the voluntary procedure of scientific advice at either national or the EMA level. Experts from regulatory authorities in the chemicals sector indicated that there are possibilities for dialog however not on single substances, due to the high amount of chemical substances that need to be assessed (i.e. in the context of REACH).

8.5.1.3 Cooperation

Cooperation or collaboration is the process where two or more parties interact to create a shared understanding. It is not only about exchanging information. It is also about the mutual use of this information (Denise, 1999) and about education. Cooperation is important in the phases of development, validation and implementation to gradually create a shared idea of the potential of 3R models. Here the early involvement of regulators, which was also brought up in both panels, is important (Bottini et al., 2008). The PARERE network, which stands for Preliminary Assessment of Regulatory Relevance, is a good example in this respect. EURL-ECVAM (the European Union Reference Laboratory: European Centre for the Validation of Alternative Methods) sends validation proposals

for 3R models to the PARERE network. Through this network all regulatory sectors, including those for pharmaceuticals and chemicals, can provide input on the potential regulatory relevance and suitability of proposed test methods and testing strategies.

Furthermore, this network facilitates the information flow between EURL-ECVAM and regulators regarding the development and validation of methods in identifying areas that need specific attention and in identifying regulatory experts to participate in specific EURL-ECVAM project groups. To involve also other relevant stakeholder groups the ECVAM Stakeholder Forum (ESTAF) was established in 2010. ESTAF brings together 15 European stakeholder organizations from academia, industry and civil society and animal welfare to help EURL/ECVAM in prioritizing those methods which are considered most promising for regulatory purposes. ESTAF fulfills the functions of representing specific interests, the voicing of societal concerns (e.g. animal welfare issues, sustainable testing) and the mutual sharing and dissemination of information on activities of ECVAM and its stakeholders.[128]

A second element of cooperation is data sharing. Data sharing between industry, regulatory authorities and academia is seen by both panels as paramount to obtain an overview of the information that is already available and the information that is still needed to fill the remaining knowledge gaps. Furthermore, data sharing is considered very important to increase trust in the existing 3R models. However, data sharing has its own specific barriers. Industry, research organizations and academia for example are very cautious when it comes to exchanging data out of fear of revealing commercially sensitive information and/or losing intellectual property rights.

Under REACH it is obligatory to share studies involving vertebrate animals in order to avoid unnecessary animal testing. Furthermore, data sharing is strongly stimulated by the mandatory formation of so-called Substance Information Exchange Forums (SIEFs). For non-animal tests, REACH only encourages the sharing of data to reduce costs for companies (ECHA, 2012).

Experts from the pharmaceutical industry declared that they are prepared to share *in vivo* and *in vitro* data with regulators, when given the guarantee that these will be solely used for the purpose of discussing the value of a 3R model for regulatory acceptance. In the pharmaceuticals panel the suggestion was therefore made by both regulators and manufacturers to create a procedure of 'safe harbors' in which 3R test data are brought in by industry with the purpose of evaluation for future regulatory acceptance, without the results of these tests being used for current regulatory decision making.

A third element of cooperation is the formation of public private partnerships. These partnerships fulfill a multifunctional role (Holmes et al., 2010, Seidle et al., 2010) in that they stimulate commitment to the 3Rs, facilitate information exchange, offer possibilities for the funding of projects and contribute to strengthening the network and thereby to the mutual trust between stakeholders.

128 http://ihcp.jrc.ec.europa.eu/glossary/estaf :consulted April 2013.

8.5.1.4 Coordination

Coordination is needed to guide a 3R models through the chain from model development to the ultimate use for regulatory purposes. Here it is important that each part of the chain is informed on how and when it must act to efficiently achieve the overall goal (Denise, 1999). Furthermore, cooperation between international stakeholders should be stimulated based on a clear roadmap. This kind of coordination throughout the entire chain is still largely missing. Parts of the process may be well managed, but guiding a model from one phase to another requires long term involvement and commitment, a strong problem ownership and leadership of several actors and the allocation of facilities and money for this specific target.

8.5.2 Suggested actions per stakeholder group

To target the main cross-sectorial barriers (see Section 8.4.2) the experts discussed:

- *Which factors can be influenced by their own stakeholder group to facilitate the process of regulatory acceptance and use and in what ways?*
- *What is needed in this respect from other the stakeholder groups?*

These discussions resulted in a series of required actions, some of which have to be initiated by one stakeholder group: the unilateral actions. Others actions require a combined action of two or three parties: the bilateral and tripartite actions.

8.5.2.1 Unilateral actions

- Regulatory authorities, legislators and policy makers [129]
    - Policy makers and legislators are asked to formulate concrete policy goals, at both a national and international level, that stimulate the use of 3R models for regulatory purposes.
    - Since 3R models -and *in vivo* models- often confront assessors with many questions due to the existing data gaps, policy makers (and industry: see below) are asked to allocate money in order to deal with these data gaps.
    - Regulatory authorities are asked to discuss about, decide upon and communicate about clear acceptance criteria for 3R models regulatory risk management purposes.
    - Regulatory authorities are asked to provide incentives for the use of 3R models, such as fee waivers for adjustments to the product license in favor of the 3Rs and shorter or prioritized assessment procedures in cases where manufacturers use the preferred 3R model. Although this is considered to be a difficult issue, it is valuable to progressively think in this direction.
    - All three sub groups are asked to continuously invest in harmonization. Investment here means allocating highly skilled experts and money for harmonization purposes and keeping the subject high on the political agenda.
    - The assessment of products and/or substances requires specified expertise and skills which are mostly not offered as part of the existing education programs. To prepare a group of students for this specified task, regulatory authorities can

129 For this paper regulatory authorities, legislators and policy makers are joined into one stakeholder group. This section specifies what action should be taken by which sub group.

provide internships for high potential students.
    - A shift of minds in all three subgroups is needed from the current hazard based to a more risk based approach which incorporates exposure to specify the actual risk.
- Academia & research organizations
    - Educators should integrate a more risk based way of thinking into their curricula.
    - Furthermore, educators should investigate the regulatory and industrial educational needs and take these needs into account.
    - Academia and research organizations are asked to prioritize research that can help in tackling the uncertainties of existing 3R models (see also Punt et al., 2011). Academia and research organizations express their interest in doing so, but this needs to be -co-funded by the other stakeholder groups.
    - Academia and research organizations have state of the art knowledge on 3R models. They can therefore play an important role in gathering and disseminating this knowledge to the other stakeholder groups.
- Industry
    - Industry is asked to share their available in vitro and in vivo data comparisons with regulatory authorities to fill a part of the existing data gaps and tackle the uncertainties of existing 3R models. Industry expresses willingness to do so when given the guarantee that these data will be used confidentially and for the sole purpose of qualification of these new technologies. For this purpose the suggestion is made to create safe harbors where this data exchange can take place in a protected environment.
    - Industry is asked to fund relevant academic research to fill the existing data gaps 3R models face.
    - Industries need to prioritize the 3Rs in their internal policies.

8.5.2.2 Bilateral actions

- Regulatory authorities & industry
    - Bilateral communication between regulatory authorities and industry is a strong wish from industry to discuss the way to handle the discretionary space - i.e. room for interpretation that exists in the regulatory requirements - and to clarify the acceptance criteria a 3R model has to meet.
- Regulatory authorities & academia / industry & academia.
    - Continued bilateral communication between regulatory authorities and industry on the one hand and industry and academia on the other is needed to discuss the educational and research needs regulatory authorities and industries have.

8.5.2.3 Tripartite actions

    - The 4C's of commitment, communication, cooperation and coordination are the shared domain of all involved stakeholder groups. Continuous investment into the 4C's is thereby needed by all involved stakeholder groups.

Figure 14, as presentend below, summarizes the recommended unilateral, bilateral and tripartite actions.

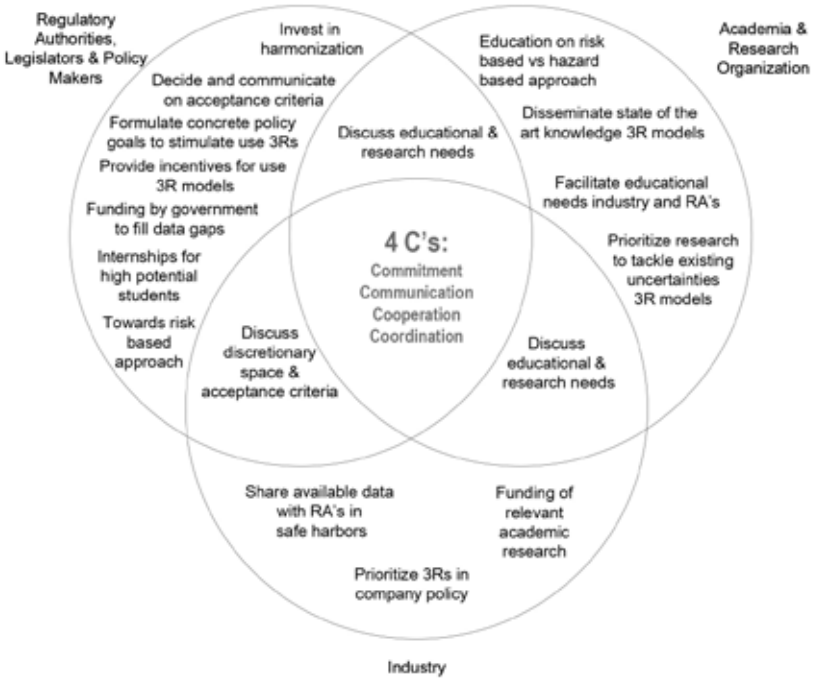


Figure 14. Actions per stakeholder group to facilitate regulatory acceptance & use of 3R models

8.6 Conclusion and discussion

Regulatory acceptance and use of 3R models is observed to be a challenging and multi-faceted process. Developments in this process are influenced by a combination of drivers and barriers at three levels:

- The micro level where 3R models are developed and validated: Barriers here are the challenging validation process and the remaining uncertainties of 3R models. The informative character of many 3R models is an important driver at this level.
- The meso level of the existing regulatory regime: Barriers here are the lack of harmonization and mutual acceptance of data and the uncertainty about regulatory acceptance of 3R models due to unclear acceptance criteria and discretionary space in the existing regulatory requirements. Legislation that explicitly stimulates the use of 3R models is a driver at this level. This also counts for the scientific flaws of the existing animal models which are a solid part of the existing regulatory framework.
- The macro level of the societal context: Here, one of the most persistent obstructing factors can be found, i.e. the high level of risk aversion within society. On the other hand the ethical concern for the welfare of laboratory animals is observed to drive regulatory acceptance and use at this level.

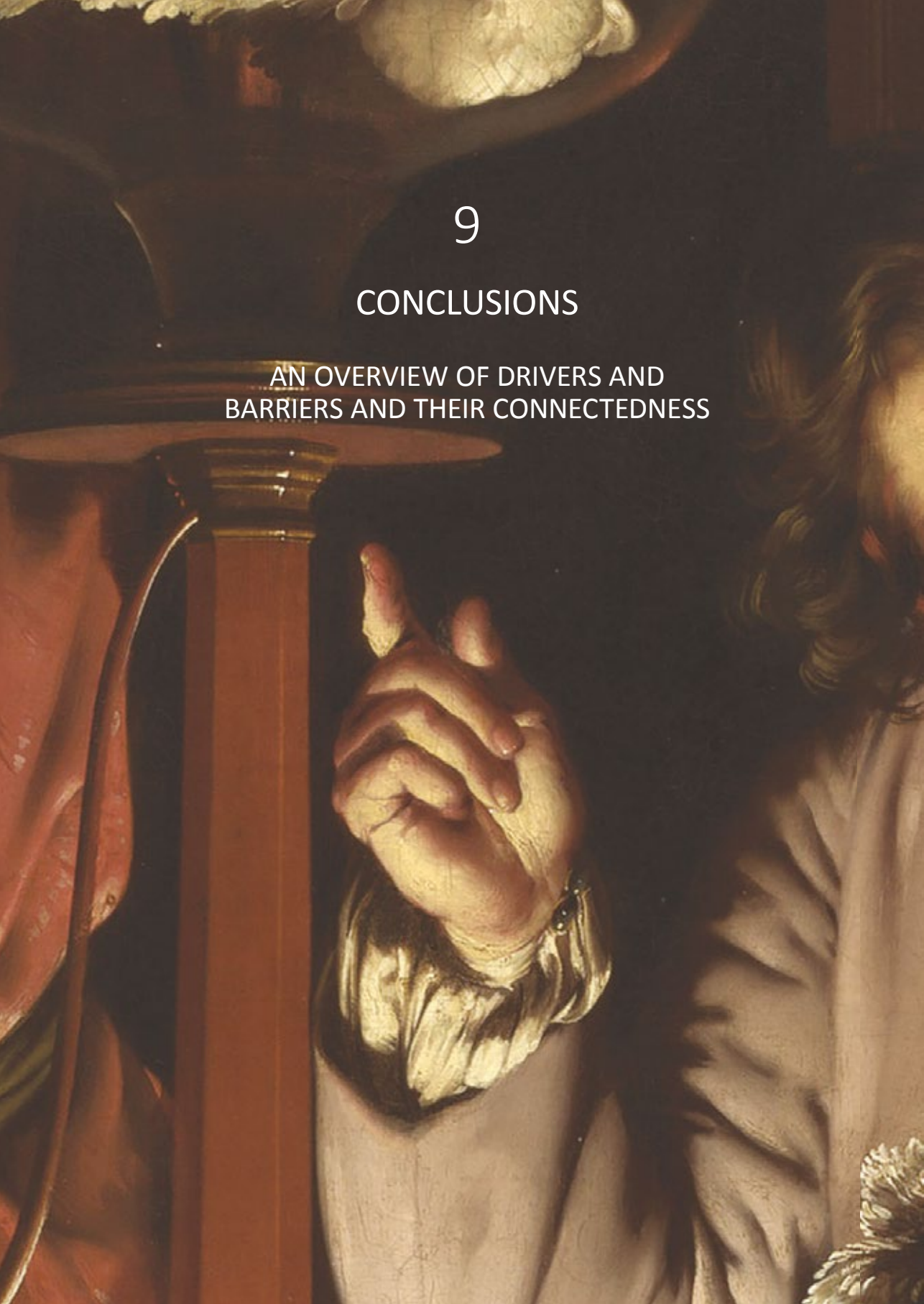
The panels showed that the influencing factors on regulatory acceptance and use are fairly comparable in both sectors. Nevertheless some sectorial differences were observed. Safety assessment of pharmaceuticals is generally based on a risk benefit analysis. For chemicals the level of risk aversion is far higher as it is for pharmaceuticals. This generally leads to a stronger risk avoiding approach in the sector of chemicals which is also reflected in the safety assessment. Regulators of chemicals are observed to strive for a zero risk level. This includes being highly cautious when it comes to shifting to a new test regime such as a 3R model. Furthermore, the profit margins are bigger for pharmaceuticals which leaves more room for investing in innovations such as 3R models. Also more knowledge is generated for pharmaceuticals (e.g. pharmacological mode of action, pharmacokinetic data -animal and human-) which allow for mechanistically and mode-of-action based risk assessments. Such sector specificities are important to take into account when trying to improve the process of regulatory acceptance and use of 3R models.

To alleviate the barriers and strengthen the drivers for regulatory acceptance and use of 3R models, it is important to make a distinction between the pliable and the more rigid factors and to target those drivers and barriers which hold the highest potential for change. Habitually the more pliable factors can be found at the micro and partly the meso level while the most persistent influences can be found at macro and partly the meso level such as the high level of risk aversion and the lack of harmonization (Schiffelers et al., 2012). Risk aversion is deeply engrained in western culture and harmonization is a problem with needs to addressed in a global manner. In this light the authors recommend prioritizing the following actions at the micro- and meso level to get a 3R model out of its niche and into the existing regulatory regime.

- Data sharing of both *in vivo* and *in vitro* data to diminish the existing uncertainties of 3R models (Industry);
- Creation of safe harbors for data sharing (Regulators);
- Prioritization of research that adds to reducing the existing uncertainties of 3R models (Academia);
- Prioritization of validation of 3R models that are most useful for regulatory purposes (Regulators and Academia). Industry should be involved to discuss the usability of the method for regulatory purposes;
- Timely involvement of regulators and other involved parties to share expertise and experience and thereby gain trust in 3R models (Regulators, Academia and Industry);
- Creation of -economic- incentives to stimulate the use of 3R models for regulatory purposes (Policy makers and Regulators);
- Formulation of concrete policy goals to stimulate the development and use of 3R models for regulatory purposes (Legislators and Policy makers).

These actions are important for both sectors but as mentioned sectorial specificities have to be taken into account when it comes to the accomplishment of these actions.

All in all, the 4C's are the central catalyzing forces for all these actions. 3R acceptance and use starts with a strong involvement and cooperation of all three stakeholder groups and shared problem definition. To then connect the micro-, meso- and macro level and help a 3R model to escape from its niche, a coordinated roadmap with a clear problem ownership and leadership is needed (see Chapter 10 for an elaboration on this).



9

CONCLUSIONS

AN OVERVIEW OF DRIVERS AND BARRIERS AND THEIR CONNECTEDNESS

"If a paradigm itself becomes the subject of discussion, this is often a sign of its decline."

Byung-Chul Han
Born 1959
South Korean-German philosopher and writer

9.1 Introduction

Regulatory acceptance and use of 3R models is in many respects a wicked problem crossing geographical, institutional and sectorial borders, involving many stakeholders and thereby diverging perspectives. Dealing with wicked problems requires a broad integrative perspective and an explorative approach that allows for complexity to exist. This is important to investigate the interrelatedness among the broad range of causal factors and to create a realistic picture of the issue under examination (APS, 2007) (See Chapter 1). Such an integrative perspective was found in the multilevel perspective on technology transitions (as described in the Chapters 4 and 5). The needed contextuality and in-depth examination was achieved through the case study approach (see Chapter 3 for an explanation of this research approach and Chapters 6 and 7 for the case studies executed in the context of this thesis). Furthermore, the drivers and barriers were discussed and particularized for the two product sectors (i.e. the sectors of pharmaceuticals - including vaccines - and chemicals) on which this thesis focuses. This was done through the performance of two expert panels, each with about 20 experts operating in the specific domain of 3Rs (see Chapter 8).

This chapter offers an overall analysis of the empirical findings as described in the chapters 6-8 in order to answer the research questions Q3a and Q3b.

Q3a. Which factors influence the regulatory acceptance and use of 3R models risk assessment and efficacy testing purposes?

Q3b. How do these factors influence the regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes?

It thereby reflects the context of discovery which was the focal point of this research. This chapter starts by offering a comprehensive analysis of the influencing factors and the underlying mechanisms as retrieved from the case studies and the expert panels. Next, an inductive approach is adopted in section 9.3 to reflect upon and come to a deeper understanding of the empirical findings, through the use of relevant notions from the literature on innovation- and risk regulation in search for dominant factors and mechanisms and their interaction. Through this context of justification, we come to an analysis of the factors influencing regulatory acceptance and use of 3R models and of possible ways to stimulate this process.

9.2 Analysis

This section evaluates the context-related factors that derive from the case studies and the expert panels (9.2.1). Next, an overview is given of the main barriers and drivers we came across in the total course of this research (9.2.2). This section answers Q3a: *Which factors influence the regulatory acceptance and use of 3R models?*

9.2.1 Context-related drivers and barriers

The barriers and drivers influencing the process of regulatory acceptance and use are looked at as social constructions which are formed by the involved actors in the specific context in which they are situated (see section 3.2). From the social science perspective which is used to examine the field of regulatory acceptance and use, there is no single reaity and social phenomena are shaped by their context. This makes it important to examine a phenomenon within this context. The case studies offered the possibility to analyze the drivers and barriers in a clear demarcated context.

The case study on veterinary rabies vaccine potency testing (SNT case) refers to process to replace the NIH mouse challenge assay in Europe by a serological assay referred to as the Serum Neutralization Test (SNT) for potency testing purposes of inactivated veterinary rabies vaccines. The SNT is a reduction and refinement model. The use of the SNT reduces laboratory animal use by up to 85% compared with the NIH test. Also, the levels of pain and distress inflicted to animals and the time and costs to conduct the test decrease significantly, while the test results are more reproducible and less variable.

The case study on reproductive toxicity testing (EOGRTS case) refers to the process to replace the two-generation reproductive toxicity study (OECD TG 416: OECD 2011a) by the Extended One-Generation Reproductive Toxicity Study (EOGRTS) for the assessment of chemical substances on reproductive toxicity under REACH. The EOGRTS is a reduction model. Using the EOGRTS can reduce the number of laboratory animals used by about 40% - a total of 1200 rats per study -. Furthermore, the tests are far more informative due to additional endpoints which are taken into account.

The values that are attributed to a certain technology (e.g. animal test or 3R models) strongly depend on the interaction between the technology and the connected individual stakeholders, involved organizations and the broader sociotechnical structures. By means of expert panels an analyses of the drivers and barriers in the broader context of the pharmaceuticals/chemicals sector was made.

9.2.1.1 Drivers in the SNT and EOGRTS case studies

The SNT and the EOGRTS case study reveal drivers at all three levels of the multilevel perspective of technology transitions.

At the macro level, both processes were driven by a combination of moral concerns and the motivation to innovate the current risk assessment approach and work towards ‘better science’. This is reflected at the meso level in legislation stimulating the use of 3R models. The use of 3R models such as the SNT and the EOGRTS is in line with the ambition of the European Commission (as described both the 2010/63/EU directive and REACH) and the EDQM (as described in the Ph. Eur. monographs) to diminish the use of laboratory animals and accept and use validated 3R models to replace, reduce and/or refine existing animal models.

Moreover, both cases revealed a clear sense of urgency placing the issue of replacing, reducing or refining the conventional model high on the agenda of the involved stakeholders. In the SNT case, there was strong international agreement on the need to replace the NIH test for a variety of scientific, ethical and economic reasons (Bruckner et al., 2003, Krämer et al., 2009 & 2010, Stokes at al., 2011 & 2012). In the EOGRTS case, the discussion was timely due to the fact that there was a European sense of urgency as a result of the REACH deadlines that had to be met. At the same time, there was pressure from the side of the US EPA, due to the objective of incorporating the EOGRTS into the OCSPP (Office of Chemical Safety and Pollution Prevention) to revise the guidelines.

As became apparent in both case studies, a very important additional driving force in the two case studies is shaped by the institutional champions/innovation entrepreneurs. In both cases change was initiated by institutionally embedded stakeholders which are part of the existing sociotechnical regime - in the SNT case the PEI, the IVI and the EDQM and in the EOGRTS case the US and Dutch OECD coordinators -. They formed the link between the sociotechnical regime (meso level) and the niche (micro level). As such, the needs of - a part of - the end users were represented from an early stage on, which is considered to be a critical success factor in technology transition (TT) literature (see section 4.3.2). User-producer interface and the level of institutional overlap between the niche and the regime are seen as an important driver in technology transitions. In both cases, the innovation entrepreneurs were able to realize this overlap and in involving a broader network of stakeholders (i.e. other regulatory authorities, manufacturers, academia and in the EOGRTS case; NGO's). Furthermore, the innovation entrepreneurs were able in keeping the innovation on the agenda. This enduring ownership is yet another driver in terms of guiding the 3R models out of its niche and into the existing regulatory regime.

The actual entrepreneurship was shared amongst various actors and is thereby a clear example of distributed heroism (Meijer, 2013). This was especially detectable in the SNT case where the five roles as described by Meijer become visible. To start with, the German Paul Ehrlich Institute (PEI) was the *creator* and *innovation entrepreneur* in terms of developing and connecting the innovation to the existing problem. Then the PEI and the Swiss IVI (The Institute of Virology and Immunoprophylaxis) started a small scale feasibility study and thereby took the following step in testing the assay. This role of *test manager* was later on adopted by the Biological Standardisation Programme (BSP) of the European Directorate for the Quality of Medicines (EDQM) who coordinated the international collaborative study. The EDQM and the Ph. Eur. Expert Group 15V were responsible for the evaluation and approval of veterinary vaccine monographs and later on became the *innovation packagers* who embedded the innovation into the Ph. Eur. monographs. The *innovation diffusers'* role was picked up again by the PEI and other Official Medicines Control Laboratories (OMCLs). However, it is still uncertain to what extent this role will be picked up by manufacturers. These roles are less easy to distinguish for the EOGRTS, but also here it became clear that one needs to think in terms of distributed heroism. In this case the US, German and Dutch authorities were observed to play an initiating role. This group of advocates later on expanded.

At the micro level, an important driving force is created by the advantages that both 3R models are observed to offer in terms of scientific progress, animal welfare and, in the SNT case, testing time and costs (see textbox above). Furthermore, both the SNT and the EOGRTS were first studied in a niche-based setting. In the SNT case, the innovation entrepreneurs started to experiment with the assay in a small scale context, which is another important driver. The potential and limitations of the test became visible through this experimental phase. The experimental stage also offered the possibility to collect and share data. The small scale feasibility study and the collaborative study thereby added to the experience with and trust in the SNT amongst the involved stakeholders. In the EOGRTS case, this experimental stage had mainly taken place through the US study conducted by Cooper et al. (2006). However, this did not involve the stakeholders in the European setting. Later on, a series of literature reviews was conducted to scrutinize the added value of using the second generation of rats (F2). This led to the conclusion that the F2 has a limited added value and the recommendation to leave out the F2 and switch to the EOGRTS. However, the antagonists, who did not get the chance to be involved in an experimental phase, were not convinced by these data and remained hesitant on replacing the two-generation test with the EOGRTS. They later on requested for an experimental phase to fill this lacuna. This was granted by the EC through the proposal to organize an additional review phase on the value of the second generation for a selection of substances and through a period of parallel testing.

9.2.1.2 Barriers in the SNT and EOGRTS case studies

Both the SNT and the EOGRTS faced difficulties in overthrowing the existing regime and in competing with the conventional model. This is due to the following combination of barriers. At the macro level, the high level of risk aversion that both rabies vaccines and chemicals face strongly influences the openness to innovations. This is again reflected at the meso level in the way that some actors are observed to hold on to the traditional way of testing i.e. the NIH test and the two-generation test. These tests have long served as the gold standard in these areas and were or still are, in the case of the NIH test, firmly embedded in the European regulatory requirements. The potency of rabies vaccines has been guaranteed through the use of the mouse challenge test for more than 60 years, elevating it to a global collective framework which has proven its value. In the EOGRTS case the frame of reference has been the two-generation reproductive toxicity test since the 1980ies. New technologies often do not match with the existing ideas about the required way of working.

The level of UI is still indefinite in both cases as a result of the uncertainty manufacturers face in terms of user demands (which in this context refers to the regulatory criteria that these 3R models will have to meet). Manufacturers still have to find out under which precise conditions the 3R model will be accepted. In the EOGRTS case, ECHA offers specific guidance on the endpoint reproductive toxicity to facilitate manufacturers. Currently though, it is unclear to what extent manufacturers have already submitted dossiers to ECHA using the EOGRTS. The UI in the rabies vaccine case is uncertain due to the fact that the Ph. Eur. monographs (still) include the NIH challenge test.

In this case, clear guidance is needed with regard to the preferred method, particularly in the context of European Directive 2010/63/EU that requires a 3Rs method to be used if available.

Changing the way of testing requires changing the test infrastructure and the product license, leading to additional costs for manufacturers. As long as the vertical legislation is ambiguous about the preferred method and the criteria under which a 3R method will be accepted the UI is confronted with extra barriers. The PEI is understood to stimulate manufacturers to adopt the SNT. However, it is currently unclear to which extent this already resulted in a broader UI.

Uncertainty regarding the level of UI is amplified by the fact that globally operating manufacturers not only have to adhere to European requirements but also to those in other regions around the world. Manufacturers see this as one of the main barriers (at the meso level) to switching to an alternative way of testing.

At the micro level, the image of what kind of information the SNT and the EOGRTS should be able to provide is largely shaped by the conventional animal models (i.e. the NIH challenge test and the two-generation reproductive toxicity test) that both offer a one-size-fits-all protocol. As long as the conventional model forms the frame of reference in validation studies, the 3R model is facing a huge validation hurdle, as described by Spielmann (2000) because these animal models usually offer a poor frame of reference (e.g. due to their high level of variability as is the case for the NIH test).

Furthermore, new technologies such as 3R models confront end-users with many uncertainties. In the SNT case for example, uncertainty was created by the fact that the SNT is not suitable for quantifying potency, while in the EOGRTS case uncertainty arose by the proposal to leave out the second generation of rats and by including new parameters which might cause difficulties in interpreting the test results. These uncertainties have led to resistance towards the use of the assay amongst several stakeholders. This resistance is amplified by the risk-averse context in which the models have to be used and the transition costs needed. Every change leads to additional uncertainties, involving extra risks. In these cases, this comes down to the fear of releasing sub-potent rabies vaccines or potentially reprotoxic chemicals. These dreads are at the basis of the discourse surrounding the SNT and the EOGRTS and lead to extended discussions regarding the possibilities and limitations of these assays. This in itself is a perfectly reasonable reaction. However, it becomes less rational when put into the perspective of the conventional model, which often includes uncertainties that are largely accepted.

9.2.1.3 Barriers and drivers from the expert panels

The barriers and drivers obtained through the case studies and those found in earlier research (Schiffelers et al., 2007: Appendix I, Schiffelers et al., 2012; Chapter 5 and Schiffelers et al., 2014a: Appendix II) were verified and specified by both expert panels for their validity in the sectors of pharmaceuticals/vaccines and chemicals. Here again the striving for risk minimization was identified by the experts as a central cross-sectorial

barrier at the macro level, while the ethical concern about animal testing was found to be an important driver. At the meso level, the lack of harmonization of legislation and test requirements recurs whereas the formulation of concrete policy goals/legislation is observed as a potential driver to stimulate the 3Rs in both sectors. At the micro level, the existing uncertainties connected to 3R models are perceived to be a central barrier whereas the informative and mechanism-based character of 3R models is seen as a driver.

A major sectorial difference was found at the macro level in terms of the level of risk aversion. The accepted risk for industrial chemicals is close to zero as a result of the involuntary exposure of being exposed to these compounds and the dread connected to the risk of chemicals (see Chapter 5). For pharmaceuticals, the level of risk acceptance very much depends on the targeted disease and the intended patient group. Vaccines, which are destined to protect broad and vulnerable patient groups like children from possible illnesses, are subjected to a far higher level of risk aversion compared to medicines which offer individual treatments and a possible last resort in severe ailments. The level of risk aversion is reflected in the regulatory requirements at the meso level. Regulatory requirements for chemicals - and certain vaccines, such as rabies - aim for zero or negligible risk levels (Kasamatsu and Kohda, 2006) which form an extra barrier to the acceptance and use of 3R models, whereas the requirements connected to many pharmaceuticals offer a certain level of risk tolerance.

In addition, the profit margins for chemicals and for veterinary medicines at the micro level are smaller compared to those for medicines for human use. These profit margins potentially influence the level of innovation, since there is less room to invest in new tests such as 3R models in the field of chemicals and veterinary medicines.

9.2.2 Overview of drivers and barriers

This section summarizes the recurring influences we came across during the empirical steps taken in anticipation of (Chapter 5) and during this research (Chapter 6-8). Table 3 offers an overview of the barriers and drivers that are observed to play a role at the micro-, meso- and macro level of the 3R acceptance model (see Chapter 5). The drivers and barriers which were identified in two or more empirical steps of this research are presented in Table 3.

Table 3. Summary of drivers and barriers

DRIVERS	Chapters/ Appendices	BARRIERS	Chapters/ Appendices
Macro level			
Animal welfare concerns	**5,6,7,8,II**	**Striving for risk minimization**	**5,6,7,8,II**
Innovation frame	**5,6,7,8,II**	**Precautionary frame**	**5,6,7,8,II**
Meso level			
Legislation stimulating the use of 3R's: Horizontal (Directive 2010/63/ EU) & vertical (e.g. REACH, Ph. Eur.)	**5,6,7,8,II**	**Strict interpretation product req. despite discretionary space** • High perceived risk products e.g. vaccines & chemicals • Fear of incidents/ releasing unsafe products • Thinking in terms of hazard instead of risk	**5,6,7,8,II** **5,6,7,8,II** **5,6,7,8,II** **5,7,8**
Profound and ongoing stakeholder interaction • 3R platforms • International collaboration and communication	**5,6,7,8, II** **5,6** **5,6,7,8,II**	**Conventional test as golden standard:** • Part of the existing infrastructure • Trust as result of long experience with it/collective memory • Remains part of regulatory requirements even if 3R model is formally incorporated	**5,6,7,8,II** **6,7** **5,6,7** **6,7,II**
Regulatory authorities committed to 3Rs	**6,7**	**Lack of harmonization**	**5,6,7,8,II**
		Unclear acceptance criteria	**5,6,7,8,II**
Micro level			
Drawbacks animal models • **Animal welfare problems** • Translational problems • Variability animal model	**5,6,7,8,II** **5,6,7,8,II** 6,7 5,6,II	**Challenging validation process** • Correlation with conventional animal model required	**5,6,8,II** **5,6,II**
Perceived advantages 3R models • **Better science: more information, less variability** • Higher animal welfare • Time reduction • Cost reduction	**5,6,7,8,II** **5,6,7,8,II** 5,6,7,II 5,6,8 5,6,II	**Perceived limitation 3R model** • Lack of understanding biological mechanism • Challenging in vitro-in vivo extrapolation.	**5,6,7,8,II** 5,6 5,8
Availability and sharing of data	**5,6,7,8,II**	Fear for false positives	7,8
Policy entrepreneurs/3R advocates	**5,6,7,8,II**	Lack of experience with/trust in new test	5,6,7,II
Early involvement end users i.e. reg. authorities and manufacturers	**5,6,7,8,II**	**Transition costs**	**5,6,7,8,II**
Experience with new test method	**5,6,7,8,II**		
Early involvement statistician	**6,8,II**		

9.3 Dominant and pliable factors, interdependencies and connections

This section addresses Q3b of this thesis: *How do these factors influence the regulatory acceptance and use of 3R models?*

To start with, conclusions can be drawn in terms of the dominance of a specific driver or barrier in the discourse on the acceptance and use of 3R models in the regulatory domain, based on the level of recurrence of drivers and barriers in the separate empirical steps (see Table 3). The drivers and barriers that recurred in every empirical step are presented using a bold font.

Secondly, it is important to note that not every recurring barrier or driver plays an equal role at the different substages of the process of regulatory acceptance and use of 3R models (FI, ARA and UI). This became clear from both case studies and the survey conducted in anticipation of the rabies vaccine case study (see Appendix II). These empirical steps for example, showed that the challenging validation process mainly plays a role at the stage of FI and that uncertainties regarding 3R models and the importance of involvement of regulatory authorities are mainly important at the stages of FI and ARA. While the lack of harmonization and the uncertainty created by unclear acceptance criteria, are especially important at the stage of UI. The dichotomy between the concern for animal welfare on the one hand and the high level of risk aversion on the other, is observed to play a role in all three of the substages. So does the need for profound and ongoing stakeholder involvement and interaction and the relevance of data sharing.

Apart from the distinction between dominant and less dominant factors in terms of their perceived influence on the process of regulatory acceptance and use of 3R models, a distinction is needed between powerful and manipulable/pliable variables (Ellemers, 1976). Pliable variables are short term modifiable factors, while powerful variables (explanatory factors) offer important explanations for the process of non-acceptance, but are much harder to influence in the short term. Preferably, the variables targeted are both dominant and pliable. Some aspects, such as the striving for risk minimization and the lack of harmonization are seen to be dominant barriers in the process of acceptance and use, are hard to modify. Targeting these factors is a long-term affair. Variables which are observed to be dominant and pliable, are the profound and ongoing stakeholder interactions, the availability and sharing of data, the early involvement of statisticians to interpret these data, innovational entrepreneurs putting and keeping the issue on the agenda, the early involvement of end users and building experience with the new test method in a niche-based setting.

Moreover, in the variety of influencing factors, a clear interdependency can be found between different developments at the macro-, meso- and micro level. When considering the recurring themes (see Table 3), it becomes clear that at the macro level the concern for animal welfare and innovation frame on the one hand, and the high level of risk aversion and precautionary frame on the other hand are very important opposing forces. Up until now the striving for risk minimization is observed to outweigh the animal welfare

concern in most of the product sectors, apart from the cosmetics sector in which the Cosmetics Directive has led to phasing out the use of laboratory animals. This certainly is the case in the area of industrial chemicals where striving for risk minimization is observed to be leading. However, also in the sector of pharmaceuticals, and especially vaccines, there is a high level of risk aversion (see Chapter 5). This is reflected at the meso level in the broad variety of product requirements to protect human beings, animals and the environment from potential adverse effects and in the choice of preferred test methods to reach this goal.

Although there are clear legislative impetuses to use 3R models, the actual decision is a balancing act between the advantages of the 3R model and the risks connected to altering the trusted routines. This balancing act often results in a precautionary approach and causing stakeholders to err on the side of caution (Barrieu and Sinclair-Desgagne, 2003), especially in cases where there is little experience and trust in the 3R model. In the area of regulatory testing, this results in holding on to the conventional way of testing, which in most cases remains the animal model.

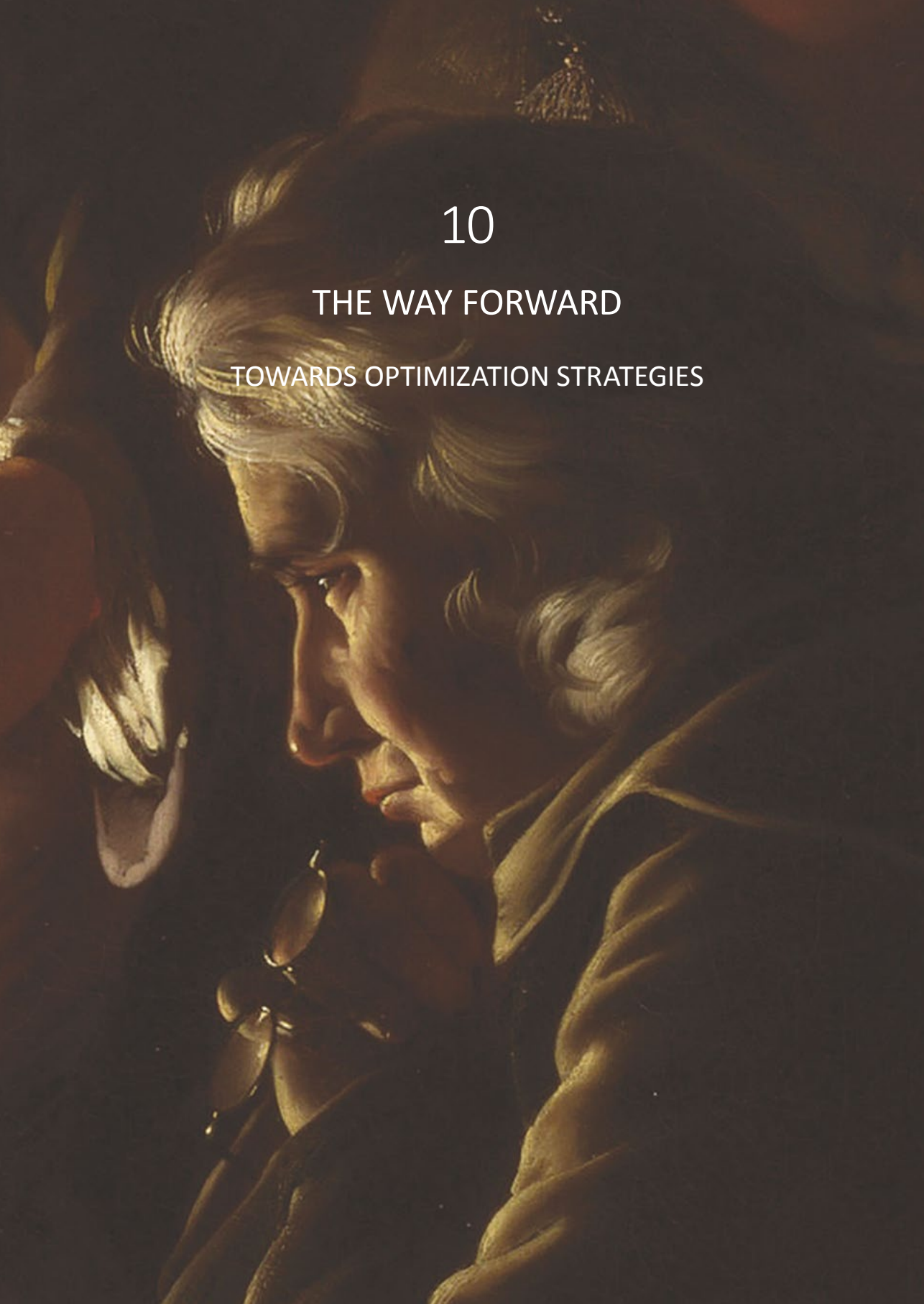
The incorporation of the animal test in the regulatory requirements decades ago has given them a firm institutionalized status. They have given guidance for many years as to which route to follow when releasing potential hazardous substances on the market. And as such, they have become a firm part of the norms and infrastructures of organizations working with them, as well as of the legitimization of their power. Institutional theory states that many of the models that give rise to organizations are based on *"rule-like systems that depend for their efficacy-for their reality, on the fact that they are widely shared, or have been promulgated by individuals or groups that have been granted the right to determine such matters"* (Scott, 1983: as cited by Scott, 2004).

However, all institutional arrangements are subject to decay and decline (Zucker, 1988). This also counts for regulatory animal testing, which is increasingly contested by new developments in risk assessment and quality control. Change in the status quo in both cases occurred through the connections between the developments at the meso level and the micro level. The replacement of a certain animal model came high on the agenda of institutional players, which were then observed to initiate strong entrepreneurial activities at the micro level by developing, testing and or pushing the new technology and by creating a strong advocacy coalition or technological innovation system (TIS) (Hekkert et al., 2007). In other words, the developments at the niches of the SNT and the EOGRTS niche were strongly influenced by the developments at the macro- and meso level, leading to the occurrence of the right momentum for change, the subsequent formation of strong advocacy coalitions and the effort of these coalitions to strengthen the scientific bases of both protocols. These developments have given rise to 'cumulative causation' (Jacobsson and Bergek, 2004) or virtuous cycles (e.g. Hekkert et al., 2007: see Markard and Truffer, 2008) which in turn led to formal changes to the existing regulatory regime Formal Incorporation (FI).

The Actual Regulatory Acceptance (ARA) of the SNT went fairly smoothly because the developer of the test is also one of the main OMCLs when it comes to releasing veterinary rabies vaccines on the European market. In the EOGRTS case, the ARA had more hurdles to overcome. The ARA at the EU level was in fact again a process, of FI, this time of the EOGRTS into the EU Test Methods Regulation (TMR). For this process, ECHA and several other stakeholders needed to be convinced. Through the FI of the EOGRTS into the Test Methods Regulation its ARA was also largely accomplished. Both cases now arrived at the final stage of Use by Industry (UI). Progress is also being made at this stage at least when it comes to the European context. The European OMCLs stimulate the use of the SNT for inactivated veterinary rabies vaccines and ECHA is stimulating manufacturers to use the EOGRTS to replace the two-generation reproductive toxicity test. Industries can subsequently play an important role in the innovation diffusion in other areas of the world. Whether they will take on this role greatly depends on the benefits they will obtain from using the SNT/EOGRTS.

Most of the energy in both case studies was directed at first-order learning i.e. generating, checking and spreading facts and data, which is important but in itself is no guarantee for the successful development of a niche. It may even withhold progress in certain cases. In the EOGRTS case, stakeholders from both sides (i.e. the innovation and the precautionary frame) mainly adopted technical and scientific arguments to make their case and cope with psychological uncertainties. They questioned each other using the same ammunition i.e. more scientific data to stress their argumentation. As a result, the scientific data became part of the controversy.

Even though first-order learning is a necessary aspect of niche development, learning processes are known to contribute more if they go beyond the level of gathering data and enable second-order learning, by stimulating changes in cognitive frames and assumptions - values and norms - within these frames (Grin and Van de Graaf, 1996). It is difficult to reconstruct whether discussions on values and norms have taken place in both cases. However, from the international meetings that were attended by the researcher in the context of the SNT case and on the 3Rs in general, it can be concluded that most discussions skipped the meta-level of norms, values and underlying beliefs and directly jumped into the discussions on technical and scientific facts and data. The highly relevant topic of discussing what the mutual and the diverging beliefs are, when it comes to animal models and 3R models and which ultimate goal they should serve, was thereby omitted. This is often observed in technology transitions (TT) …"*technology actors usually focus on developing, testing and optimizing technology, but neglect the embedding in broader societal goals, or leave it to a later stage"* (Schot and Geels, 2008, p. 538). As a result, doubts and insecurities resurface at a later stage. It is therefore very important for stakeholders to facilitate second- order learning at an early stage of the acceptance process, an aspect which will be elaborated on in the final chapter.



10

THE WAY FORWARD

TOWARDS OPTIMIZATION STRATEGIES

"The single biggest problem in communication is the illusion that it has taken place."

George Bernard Shaw
1856-1950
Irish writer, critic and political activist

10.1 Introduction

The debate on the acceptance and use of 3R models in the regulatory domain is increasingly characterized by an ambiguous environment, in which two opposing discourses are found: the innovation frame which stipulates the advantages of the 3R model and exposes the downsides of the animal model, and the precautionary frame which focuses on the uncertainties connected to the innovation (i.e. the 3R model) and the potential loss of replacing the conventional way of testing (i.e. the animal model) (see Chapter 7).

As long as the precautionary frame prevails or as both sides are equally strong, the 3R model remains stuck in its niche, unable to compete with the existing regulatory regime. The case studies, however, offer examples in which regulatory acceptance and use, at least to a certain extent, was accomplished. This success is strongly connected to the actions of committed innovation entrepreneurs who are embedded in the existing regime and put a lot of energy in strengthening the niche and simultaneously building the bridge to the existing regime.

The niche in both cases was empowered through the combined action of a diverse group of stakeholders that collected, shared and distributed data and expertise connected to the innovation, i.e. first-order learning. However, focussing on first-order learning entails the risk of overlooking important criteria that an innovation has to meet further on in the process. Both advocates and opponents of the innovation are often observed to focus on the generation and interpretation of – partly the same – scientific data to underline their argumentation. The generation and distribution of new scientific data however is no guarantee that the existing dichotomy will be solved. It may even widen the gap as these data also become politicized. In other words, first-order learning and technical discussions are very important, but insufficient to fully guide a 3R model through the 3 substages of regulatory acceptance and use.

The current technocratic perspective overlooks the fact that the introduction and acceptance of new technologies often depends more on social, psychological, cultural and historical factors than on technological merit (NRC, 2004). The slow regulatory acceptance and use of 3R models is thereby often the result of conflicting social constructions which are built around the different types of testing. These social constructions are formed by entrenched values and beliefs, narratives, images and perceptions regarding animal models and alternative ways of testing. These building blocks form the basis for the collective actions of and interactions between stakeholders, when it comes to 3R acceptance or non-acceptance. Therefore, the fact that the psychological, cultural and communicational problems are not structurally addressed is a lacuna that needs to be filled. As long as stakeholders are not enabled to mutually reflect upon their underlying presuppositions, inertia is more likely to persist. Therefore, dealing with the remaining inertia also requires second-order learning. This means that the discourse should no longer be restricted to discussing the technological and legal barriers, but should also bestow time and attention upon the values that are attributed to different techniques and the underlying uncertainties, beliefs and social dynamics that foster these values.

Although social constructions aren't changed easily, they can be reconstructed in due time. The question now is: *what is needed to stimulate and facilitate such change?* To address this query, the focus in this final chapter shifts from description of the existing situation to prescription by answering Q4: *How can the process of regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes be optimized?* It thereby focuses on the phronetic question (see Chapter 3): "*What, if anything, should and can be done to alter the current situation."*

In Chapter 9, an approximation of the influencing factors in terms of their dominance in the field was given. Moreover, a specification was given of the factors that can be identified as pliable (Ellemers, 1976) when trying to alter the current tardiness 3R models face in the regulatory domain. Targeting the factors that are both dominant and pliable offers the best prospects for change. However, looking at individual variables easily results in *kurieren am Symptom* (treating the symptom), a social engineering type of approach that focuses on how to deal with which variable. This leads to the ad hoc treatment of symptoms without offering routes to ponder the underlying causes. Instead, to overcome the existing tardiness that 3R models face in the regulatory domain, thinking in terms of enduring and coordinated strategies is needed. This requires an elevation of the current discourse out of its technocratic track and towards a more integrated and multidimensional level.

Such an integrated approach necessitates broadening the technocratic perspective and aligning the developments at the micro-, meso- and macro levels. It is a process of co-evolution and mutual adaptation between the technology and the system for which it is destined. As was already described by Dosi (1982), the emergence of new technological paradigms stems from the interplay between scientific advances, economic factors, institutional variables and unsolved difficulties on the established technological paths. This interplay has become visible throughout this research. The concept of the gold standard for example is a clear instance of a social construction which is at the stage of being re/deconstructed. The existing regime is contested by a combination of endogenous forces from within the specific regulatory regime (meso level) and exogenous forces from outside the regulatory regimes such as developments in terms of other 3R models (micro level), developments in other regulatory regimes (meso level) or broader societal developments (macro level). These forces may reinforce each other, if steered in an orchestrated manner, as will be described in this chapter.

This chapter works towards a customized integrated roadmap which is presented in section 10.3. To this end, subsection 10.2 describes the design principles which are important to keep in mind when thinking in terms of the 'the way forward'. These are the need to recognize and align the developments at the micro, meso and macro level, by making use of critical junctures and the need to think in terms of evolutionary change. Next, two strategic approaches to stimulate the regulatory acceptance and use of 3R models are described in section 10.3 that derive from technology transition literature and in which the individual drivers and barriers are incorporated. The overall policy perspective of transitions management is adopted to stimulate change at the macro- and meso level,

while strategic niche management offers ways to stimulate favorable developments at the micro level, for example by facilitating innovation entrepreneurs, the search for successful connections to the meso level and the broadening of the network. The design principles and the strategic approaches form the basis for the roadmap to change in section 10.4. This roadmap is based on a combination of these strategies and consists of a resourceful combination of theoretical and empirical elements. It thereby summarizes the lessons learned from the empirical and the theoretical findings of this research. Lastly, section 10.5 reflects upon this research by discussing the contributions which have been made to the scientific and social debate and by detecting the limitation of this research.

10.2 Design principles

10.2.1. Recognizing and aligning developments initiating change

The multilevel perspective on technology transitions helps in defining the developments at the macro-, meso- and micro level and in tracking down stress from within the system and tensions and pressures from outside the system. These are relevant developments to appreciate and combine when shaping a suitable strategy for change. As described in Chapter 4 and Chapter 5, sociotechnical change or a regime shift requires *"the alignment of developments (successful processes within the niche reinforced by changes at regime level and at the level of the sociotechnical landscape)" (Kemp et al., 2001, p. 277)* and alterations at one part of the network can activate alterations at other elements (Geels, 2002). Alterations can therefore be initiated at all three levels, e.g. pressure from political, cultural, economic developments at the macro level, the amendment of regulatory requirements and the increasing discussions connected to new visions on risk assessment and quality control of chemicals and pharmaceuticals/vaccines at the meso level and new technological developments, new products and new stakeholders or stakeholder cooperation at the micro level. The alignment of these factors can be facilitated through various proactive ways of connecting.

To start with, user-producer interaction and the level of institutional overlap between the micro level (niche/technological innovation system - TIS -[130]) and the meso level (sociotechnical regime) are important indicators for successful technological regime shifts (Kemp et al., 1998). *"It depends on the institutional networks and the interaction – mutual set of actors – of a TIS with a regime, what the level of opposition will be to transform in the way as proposed by the niche".* Both case studies presented in this thesis underscore the relevance of this interaction. The progress booked in terms of the formal incorporation (FI) of these 3R models was largely due to the fact that the initiators were important players in the existing regulatory regime. The group of stakeholders is preferably broad or needs to be broadened with relevant stakeholders. This involvement is needed to investigate the potential attitude towards the innovation. Uneasiness with a certain 3R

130 Set of networks of actors and institutions that jointly interact in a specific technological field and contribute to the generation, diffusion and utilization of variants of a new technology and/or a new product (Markard and Truffer, 2008)

model needs to be targeted at an early stage or it will be transferred to later stages of the process. It requires the investigation of and discussion about vested interests and joining the diverging points of view.

Alignment can also be initiated by stress within the existing regulatory regimes (meso level) which offers opportunities in breaking open the status quo. Stress is for example caused by inconsistencies in the system leading to critical junctures – rare events in the development of institutions.[131] The normal state of an institution is either one of stability or one of constrained, adaptive change. During critical junctures, change is substantially less constrained than it is during the phases of path dependence that precede and follow them. (Capoccia and Kelemen, 2007). Such critical junctures were for example created by the development of the REACH directive, the revision of Directive 86/609/EEC into Directive 2010/63/EU, the development of the new Cosmetics Directive and on a smaller scale, the amendments of European directives and monographs of Ph. Eur. Also the discussions on the need to change the current testing paradigm are a source of stress. The European Citizens Initiative (ECI), with its proposal to put a stop to vivisection throughout the European Union is an expression of this and may lead to a critical juncture. Institutional entrepreneurs (DiMaggio, 1988) can adopt such critical junctures to create new institutions e.g. new ways of risk assessment and quality control. It is therefore important for innovational entrepreneurs to be aware of such opportunities, since they can offer opportunities for innovations to escape from their niche.

10.2.2 Thinking in terms of evolution rather than revolution

However, transformational change is not necessarily the result of critical junctures; it is often the result of an incremental process (Capoccia and Kelemen, 2007) in which new technologies physically link up with established technologies. Thinking in terms of evolution rather than revolution (Rotmans et al., 2001) is very useful for two reasons. Firstly, critical junctures are relatively rare events and as a result there are limited opportunities for revolutionary change. Secondly, an evolutionary mechanism of technological breakthrough is likely to be a more suitable approach in this risk-averse context of regulatory testing where change invokes high levels of uncertainty and where many respondents warn for discarding the animal model too early. They indicate that a 'stand-alone' situation, of either *in vivo* or *in vitro* methods, is in most situations neither feasible nor desirable. "*The respective advantages and limitations of both animal models and in vitro models argue much more for a combined approach than either approach alone.*" (Hartung and Daston, 2009). A well-considered combination of both types of testing is therefore, for now, believed to be the best feasible scenario. In this light, reduction and refinement models are more likely to become adopted as they are less radical in changing the existing situation, when compared to full replacement models.

Over time, 3R models have already become a stronger rival to the conventional animal models. A growing amount of validated technologies has become available that have

131 Institutions are formed by the vested set of rules regulations, the connected infrastructure and related stakeholders.

the potential to compete with the conventional animal models. As a result, the conventional models are under increasing pressure. This slowly but surely leads to a shift in the symbolic meaning of both technologies in which *in vivo* testing is increasingly disputed and *in vitro* testing is on the increase. In line with this evolutionary approach, Stephens and Mak (2014) define four phases which the 3R concept in toxicology displays: the incubation phase (1959-1979); the phase of increasing acceptance and spread (1980s to early 1990s); the maturation phase (early 1990s to 2007) with an initial goal of a one-to-one replacement of an animal model by an alternative test. This goal is gradually replaced by increased attention to integrated testing strategies, particularly to the more complex endpoints. This is the paradigm shift phase (2007 until present) with the US National Research Council (NRC) report, which was proposed as a long-term transformation as the tipping point. This NRC report entitled 'Toxicity testing in the 21st Century, A vision and a Strategy' (NRC, 2007) predicted the near elimination – if not the total replacement – of animal use in toxicity testing through the development of 21st Century Toxicology (Stephens and Mak, 2014). A similar development can be observed in the area of biologicals, where the consistency approach is gaining terrain; a principle in which the quality of vaccines is based on a standardized production process (e.g. De Mattia et al., 2011).This evolutionary view is important to take into account when working towards a strategy to enhance 3R acceptance and use.

10.3 Optimization strategies

Technology transition literature defines two important optimization strategies (see also Chapter 4) which offer useful input for the roadmap to change as described in section 10.4.

10.3.1 Stimulating the use of 3Rs through transition management

Transition management is based on a process-orientated philosophy in which *"both objectives and final visions are determined socially, not just by expert scientific knowledge"* (Rotmans et al., 2001, p. 22). It aims at engaging a wide range of stakeholders over the multiple levels to create shared visions and goals in order to encompass societal values and beliefs, and requires a long-term perspective. Experiments are used to identify how successful a particular pathway may be and to stimulate learning by doing. Transition management insists that policy makers follow two parallel tracks i.e. the track of incremental adjustments to existing practices which is referred to as 'system improvement' and the track of experiments with fundamental adjustments to 'dominant designs' which is referred to as 'system innovation'. It embraces long-term thinking as a base for short-term policy with the transition objective as a crucial component. Stakeholders in the field often stress the importance of clear European policy goals to stimulate the commitment towards 3R models (See Chapter 8). Directive 2010/63/EU is a first step in this direction. However, a stronger and more concrete European vision is needed with regard to the acceptance and use of 3R models to replace, reduce and refine regulatory animal testing to deal with the remaining barriers and to influence those collectively held beliefs that restrain 3R acceptance.

10.3.2 Developing and protecting niches through strategic niche management

Strategic management of niches is promising in terms of reinforcing niches to weaken the dominant regimes and thereby facilitating 3R models to compete with the regime. The creation and protection of niches is important to build networks, stimulate learning processes, articulate expectations, promises and visions and lead to a shared agenda (Boon et al., 2014). Boon et al. define several strategies in the context of niche protection, such as the creation of a platform for discussion, the arbitration of different views, the negotiation between and capturing of different perspectives and the forming of compromises. Especially these aspects of second-order learning, i.e. discussing, negotiating and compromising, need to be further developed in 3R niches.

Furthermore, experts in the field of the 3Rs often refer to the necessity of broadening the scope of the niche through cross-sectorial/inter-niche learning. Many of the developments in one sector or niche are observed to be relevant to other sectors or niches, and even though 3R models form a diverse group of technologies, their overarching goal is to offer an alternative way of risk assessment and quality control, which is currently still largely based on animal testing. Different 3R models may aim for different sociotechnical regimes (e.g. chemicals, pharmaceuticals, cosmetics etc.), however as a group of technologies they are observed to slowly (but surely) influence the overall way of thinking about the dominant way of testing. To come to an accumulation of initiatives it is important to think in terms of advocacy coalitions (Sabatier, 1988) or Technological Innovation Systems (TIS) (Markard and Truffer, 2008). Both concepts, in the context of this research, refer to groups of committed stakeholders with a shared goal to work towards certain 3R models in the regulatory domain, within a specified timeframe. An advocacy coalition or a TIS can put a certain issue or new technology on the agenda, share knowledge and lobby for resources.

Changes occur because social constructions are contested by groups that actively try to renegotiate existing frames/meanings. Through the permutation of different groups of stakeholders with different motivations, attitudes and perceptions, new social realities may be created. Such networks are therefore preferably broad – i.e. multidisciplinary, crossing geographical, organizational and sectorial borders – to represent the different views and to provide for first-order learning (generating facts and data) and second-order learning – i.e. broadening cognitive frames and assumptions (Schot and Geels, 2008). Such reconstruction processes require intensive interaction between different stakeholders, negotiation (Stone, 2002, Sabatier 1988) and rational bargaining (Hajer, 1995) between diverging interests and frames/discourses. In the field of the 3Rs various partnerships and platforms, such as EPAA, IVTIP, ESTIV, and NC3Rs. AltTox[132] are available to facilitate the exchange of technical expertise, bring together the lessons learned and combine initiatives. However, once again the focus mostly lies on the technical aspects and far less attention is given to the underlying social constructions and the governance structures that are needed to stimulate progress.

132 European Partnership for Alternative Approaches to Animal Testing (EPAA), In Vitro Testing Industrial Platform (IVTIP), European Society of Toxicology in Vitro (ESTIV), UK National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), AltTox internet platform on Non-Animal Methods for Toxicity Testing

10.4 Roadmap to change

In this final section, a roadmap is presented based on a combination of the empirical findings, the theoretical notions and the optimization strategies presented above. It offers a general process-oriented approach by combining the central elements that are needed for a well-designed and coordinated process and covers the crucial necessary steps in the transition from regulatory animal testing to the acceptance and use of 3R models. This process-oriented approach is important to enhance the quality of the discussions that are taking place in this domain and to balance the technocratic bias which holds the involved stakeholders in a firm grip.

Three tracks are distinguished: the niche-based, the regime-based and the society-based track. These tracks are based on respectively the micro-, meso- and macro level of the multilevel perspective on TT, which has been used throughout this thesis.

It must be noted that there is no one-size-fits-all strategy. To do justice to the complex multilevel and multi-stakeholder reality, adaptation of the strategy is required for individual cases. However, each track encloses pivotal components that individuals and organizations, in their determination to optimize the process of regulatory acceptance and use of 3R models and to discuss and target the existing social constructions, should combine in composing a suitable strategy.

10.4.1 Track 1: The niche-based track (micro level)

The niche-based track starts from the innovation and focusses on the empowerment of the niche through niche development, niche accumulation and connections with the existing regime. It commences with entrepreneurial activities (see Figure 15) which are essential in innovation systems. Innovation entrepreneurs can exercise a great deal of power and have the capacity to make a real difference (DiMaggio, 1988).

When it comes to entrepreneurial activities, one needs to think in terms of distributed heroism (Meijer et al., 2013), as was already described in Chapter 9. This means that the responsibilities are shared between a core group of involved actors, each with their own competences such as test development (e.g. by regulatory authorities, manufacturers or academia), test validation (by formal validation bodies such as EURL ECVAM or the Biological Standardization Programme of the EDQM), distribution of test results (by test developers and validators through different fora) and the coordination of the communication within the advocacy coalition, and also with external stakeholders (preferably by a neutral actor in the field, with the status to bring the stakeholders together, such as the European Commission or the EDQM). These roles will be elaborated on below.

Regulatory acceptance requires that the specificities of a particular regulatory context are taken into account from an early stage (e.g. the stage of development and validation of a 3R model). Meaningful action towards the use of 3R models can only be taken in a specified domain involving the main stakeholders. This calls for the involvement of

political expertise and knowledge on regulatory domain. Subsequently, it is important to make a force field analysis defining the players in the field and their interests.

After the force field analysis is completed, the innovational entrepreneurs need to build an advocacy coalition of stakeholders committed to the proposed innovation. This coalition preferably consists of a broad group of stakeholders in which end users of the 3R model, i.e. regulatory authorities and manufacturers, are well represented. This coalition has the goal to empower the niche through the construction of a solid shared narrative. This narrative includes the sense of urgency to switch to the 3R model and offers information about the downsides of holding on to the status quo. Furthermore, it takes in the arguments which form the basis of the anti-narratives and offers options of dealing with the remaining uncertainties, e.g. through the construction of an experimental stage in which stakeholders can gain experience and trust in the innovation.

The advocacy coalition needs to be formed in an early stage, i.e. around the test development/ pre-validation. It requires broadening the circle of committed stakeholders and involving institutionalized stakeholders. This is essential to diminish the risk that the 3R model becomes the hobby horse of the advocacy coalition. If the advocacy coalition/ innovational entrepreneurs take too much responsibility for the innovation, other stakeholders may become discouraged to bring their expertise and knowledge to the idea (Kessler and Chakrabarti, 1999) or may resist in cooperating or committing to the idea. This aspect became noticeable in both case studies in which the protagonists of the 3R model under study have done a lot to put and keep the issue on the agenda of the involved regulatory level (EDQM and OECD). In their optimism they may have overlooked or even pulled rank over some of the remaining doubts of other stakeholders.[133]

To reduce the risk of actors blocking further progress, it is very important to keep an open mind to divergent arguments and to find out what the underlying assumptions are on which their arguments are based. For this, it is important that participants in the discussion *"must be able to put themselves in the shoes of other actors in the environment, and they must have a complementary ability to reflect on their own ways of framing the policy situation"* (Rhein and Schön, 1994 p. 187).

To take the divergent opinions into account, innovation entrepreneurs should be accompanied by process coordinators who are able to keep all of the involved stakeholders actively involved and to keep an open mind to opposing arguments. The process coordinator needs to dispose of excellent communicational and negotiation skills and must take a neutral position. His/her primary goal is to uncover potential controversies, facilitate the discussions about the underlying arguments and find ways to work towards consensus.

Additionally, a strategy with a clear mutual goal is developed, which includes the steps to be taken to get there. Next to these activities within the advocacy coalition, ample room must be created to broaden the network in order to discuss the perspectives of

133 In the EOGRTS case several veto players (Tsebelis, 2002) were subsequently observed to block progress at a certain stage. A veto player is a political actor with the possibility to decline a choice being made and thereby stop a change from the status quo.

those stakeholders that remain hesitant about the 3R model. Moreover, broadening of the network also includes connecting with other comparable niches to share experiences and learn from each other in terms of the process of empowerment.

In short, the following actors and actions are distinguished in this niche-based track:

- The test developers preferable are institutional players, i.e. stakeholders that already have a firm status within the regulatory regime (e.g. European regulatory authorities or manufacturers).
- If the test is developed by academia, they should stay in close contact with regulatory authorities and manufacturers to discuss the needs in terms of test development and education (see Chapter 8).
- A process coordinator with connections in and a clear overview of the specific regulatory domain. Such a process manager can be found at the European Commission, the EDQM, EURL-ECVAM or a national regulatory authority within Europe. The process coordinator must have a neutral position and excellent communication and negation skills. Furthermore, the process coordinator must have knowledge of governance structures which enable thorough discussions between opposing points of view.
- An advocacy coalition needs to be build which is composed of national/European regulatory authorities, European/global standard setting organizations and manufacturers with a clear view on the criteria which a 3R model has to meet in the regulatory domain. Possible and relevant other regions around the world need to be involved in this niche-based track.
- Manufacturers should share their expertise with 3R models as much as possible with regulatory authorities (if needed within safe harbors, see Chapter 8) to feed the discussion on the potential of 3R models.

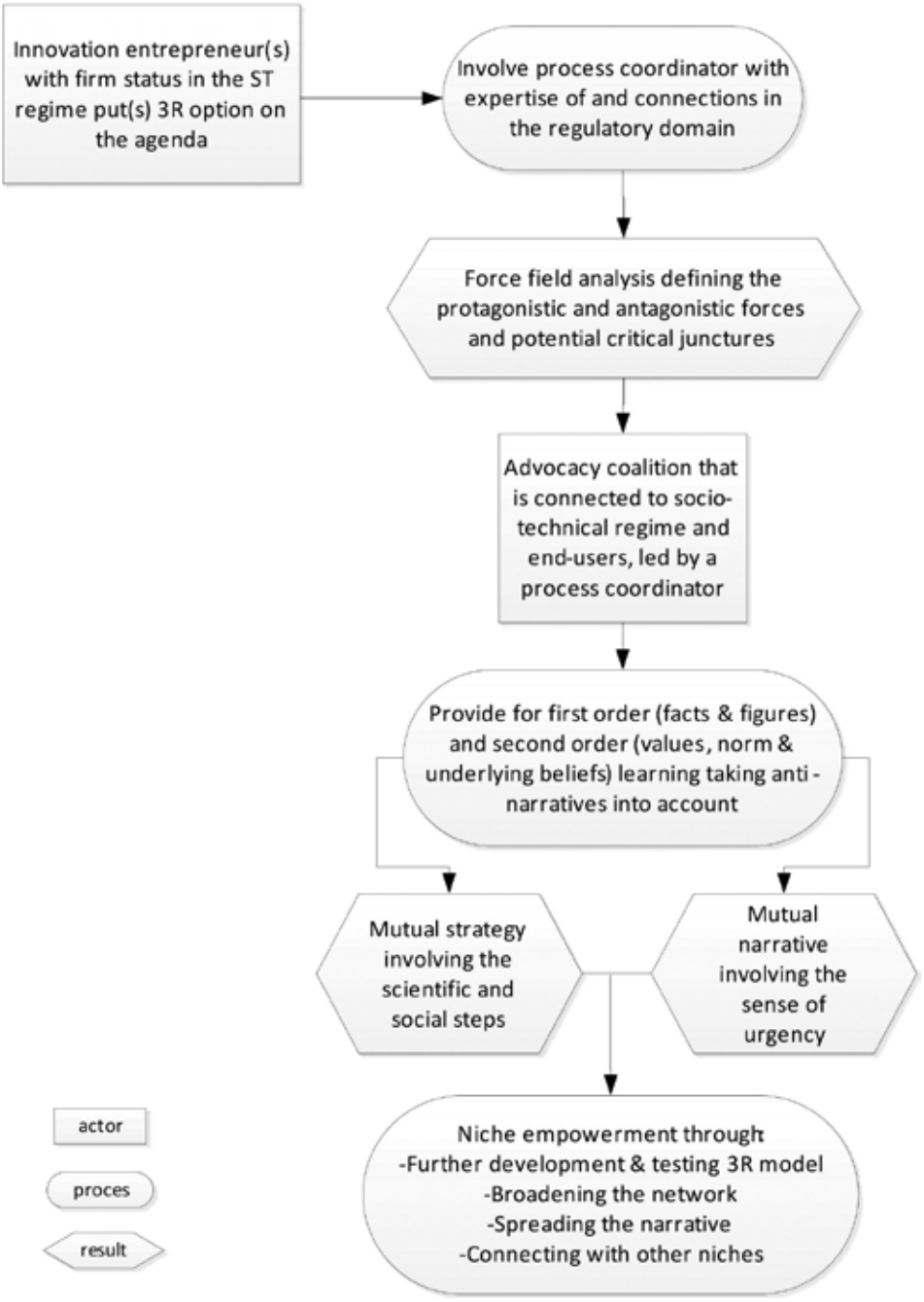


Figure 15. The niche-based track to 3R acceptance and use

10.4.2 Track 2: The regime-based track (meso level)

According to Kotter, change first requires a shared sense of urgency (Kotter, 2006). As long as stakeholders are not convinced about the need to change, no real conversion can be expected. Such a sense of urgency in the field of the 3Rs is at best instigated by clear policy goals, as was the case with the Cosmetics Directive. Although Directive 2010/63/EU sends out a clear signal in terms of the required use of 3R models, it does not set concrete goals. To initiate progress it is important that clear policy goals are set in terms of: What do we want to reach in terms of the 3Rs, through which actions and starting when? Subsequently, the actions can be specified in terms of required stakeholder involvement, existing barriers and drivers, required facilities, a division of tasks and a specified timeline.

In the absence of clear policy goals a solid debate is needed within each product sector involved in regulatory animal testing on the way that risk assessment and efficacy testing is conducted and on the added value of animal models in this domain. Through such a debate, priorities need to be set regarding the animal models that should be designated for phasing out and the 3R models that have to be prioritized for regulatory acceptance. Too often the 3R discussions take place without any frame of what they should result in. This leads to meetings in which the importance of change in favor of the 3Rs is stipulated but no actual steps are defined /agreed upon to reach this goal. The European Commission (including EMA, ECHA etc.), the EDQM and the OECD are important institutions that should take a leading role in organizing and facilitating such meetings and in setting clear mid-term and long-term goals with regards to the acceptance and use of 3R models in the regulatory domain and facilitating an ongoing discussion on altering the current test paradigm. This involves discussions on the diverging beliefs and values (the prevailing social constructions) both connected to *in vivo* and *in vitro* testing and working towards mutual ideas and goals.

Furthermore, the issue of harmonization needs to stay high on the international agenda within Europe and on a global level. Both regulatory authorities and manufacturers are observed to be aware of the importance of this issue. At the same time, this is one of the hardest elements to change since it involves many regional actors, interests and cultural differences. Manufacturers can play a role in informing regulatory authorities on the potential of alternative strategies which they for example already use for risk assessment or quality control purposes elsewhere in the world.
In short, the following actors and actions are distinguished in this regime-based track:

- The European Commission (including EMA, ECHA etc.), the EDQM and the OECD are important institutions to take a leading role in setting clear mid-term and long-term goals with regard to the acceptance and use of 3R models in the regulatory domain.
- The European Commission (including EMA, ECHA etc.), the EDQM and the OECD are key players in keeping the issue of harmonization on the agenda.
- National member states of the European Commission and the OECD should give clear signals to these organizations in terms of the need to stimulate the acceptance and use of 3R models in the regulatory domain and in terms of the ongoing need to harmonize regulatory requirements.

- Discussions have to be organized by the European Commission, the OECD and the EDQM to discuss the required changes towards a risk-based/ consistency approach.
- Regulatory authorities should provide incentives to manufacturers to stimulate the use of those 3R models that are formally accepted (see chapter 8).
- In the light of Directive 2010/63/EU manufacturers should prioritize 3R models within their company policy (see Chapter 8).

10.4.3 Track 3: The society-based track (macro level)

Animal testing generates many strong emotions, however it remains a relatively minor issue in the public policy discourse (Pijnappel, 2016) when for example compared to safety issues The issue of animal testing used to be a primarily moral issue to the broad public. It is therefore important that this discussion is also broadened by including the scientific questions connected to animal testing. For example, a public discussion is needed on the value of animal-based risk assessment and quality control. This means that the translational problems connected to animal testing in terms of the value of the test results for human beings should be addressed in the public domain. Moreover, the general public needs to be involved in the discussions about what such a test paradigm should offer. This means that the public needs to be informed about the advantages and disadvantages of the current paradigm, as well as about the possibilities and limitations of new test options. This is a difficult task since animal testing and the use of 3R models is a complex and highly expertised domain. Nonetheless, the basic aspects of the conventional way of testing and the 3R options can be explained to a broader public. In this discussion scientific dilemmas connected to this issue should be addressed. The European Civil Initiative (ECI) already took a clear step in this direction. Through the initiative 'Stop Vivisection' the ECI collected more than 1.150.000 signatures in support of a paradigm shift in the way biomedical and toxicological research are being conducted.[134] Even though the initiative did not lead to the requested abrogation of Directive 2010/63/EU and the prohibition of animal testing, the ECI could be used to encourage further (European) discussion on this issue. Several actors such as regulatory authorities, academia and non-governmental organizations (NGOs), such as animal welfare organizations, can play a role in continuing and broadening this discussion.

The following actors and actions can be distinguished in this society-based track:

- NGOs, such as animal welfare organizations, play an important role in keeping the issues of animal testing and the required paradigm shift on the agenda.
- Also, regulatory authorities, standard-setting organizations, such as the EDQM and the OECD, manufacturers and academia should play a role in keeping the issue of a required paradigm shift on the agenda.

134 "*Considering clear ethical objections to animal experiments and solid scientific principles that invalidate the "animal model" for predicting human response, we urge the European Commission to abrogate directive 2010/63/EU on the protection of animals used for scientific purposes and to present a new proposal that does away with animal experimentation and instead makes compulsory the use - in biomedical and toxicological research - of data directly relevant for the human species."* http://www.stopvivisection.eu/ : consulted December 2015

- Politicians should be informed of the limitations and possibilities of both test paradigms. NGO's as well as (national) regulatory authorities and academia can play an important role in informing politicians on the limitations and progress in this domain.

10.4.4 Connecting the three tracks

The process-oriented approach offered through these three tracks requires circularity and interconnectedness between the elements at all three levels. Focussing on just a few of these elements will slow down potential progress. Ideally, the three tracks are connected, which strengthens the developments of each individual track. This requires continuous Commitment, Communication, Collaboration or Cooperation and Coordination. These four C's which are strongly connected to each other were already introduced as core drivers to facilitate the process of regulatory acceptance and use of 3R models (see Chapter 8). They form the cement between the three tracks and are pivotal to enable the connection between them and to align the elements within these tracks. Below, a further specification is given to clarify what is needed in terms of the 4 C's. Furthermore, a fifth C for Continuity is added.

10.4.4.1 Commitment

Commitment is the fundament needed for initiating change. Without commitment, advances within or between the three tracks may occur randomly but will be difficult to realize. Numerous studies have shown that innovational entrepreneurs need solidarity around the idea (e.g., Howell and Higgins, 1990, Markham and Griffin, 1998) in order to convert the idea into innovation success. This means that there needs to be shared commitment and consensus about the central idea and the way it is going to be reached. Commitment is no status quo and requires a continuous process of reciprocal communication, collaboration and coordination.

10.4.4.2 Communication

Many of the hurdles encountered by 3R models are connected to communication problems. Most of the involved actors possess a high level of expertise in their specific domain and tend to communicate with other stakeholders from this specific perspective. This produces a technocratic focus which tends to cover up and even aggravate the underlying sentiments which drive stakeholders' perseverance. Staying on this technocratic track holds the risk of ending up in a 'dialogue of the deaf' in which controversies are deepened and actors do not comprehend each other any longer (Koppenjan and Klijn, 2004). This has been observed in the EOGRTS case, where one of the respondents concluded that the different parties in the end *"agreed to disagree"*.

Thorough discussions are needed about the course to follow. A mutual agreement on the course of action is very important to steer the energy in the desired direction. Crossing technical, institutional and sectorial boundaries and a relentless endeavor to understand each other's beliefs and considerations is desirable to set a mutual course of action and to overcome the existing controversies. Communication and meta-communication – i.e.

communication about the way of communicating – are therefore the central engines to deal with these barriers. Such insights are needed to discuss the philosophical differences at the regulatory level and to assure that the differences in social constructions, on values attributed to the different testing models and the connected underlying moralities are adequately addressed. In this process it is also important to specify which actors are enforcing the institutionalized norms (Scott, 1983). This information is pivotal to decide upon the eventual desirable steps to bridge the differences in points of view. This requires a neutral process leader, with the legitimate position to bring the parties together and commit them to a thorough communicational process. Through this communication process, various interests are examined, balanced and negotiated. (See the element of Coordination below). In the SNT and the EOGRTS process respectively, the EDQM and the European Commission were the most appropriate candidates to fulfil this role.

10.4.4.3 Cooperation

Wicked problems, such as the regulatory acceptance of 3R models, involve many actors working from different institutional backgrounds (Koppenjan and Klijn, 2004). Cooperation between these actors is very important but at the same time highly challenging. Nevertheless, a collaborative approach is one of the main strategy options to deal with wicked problems (Koppenjan and Klijn, 2004, APS, 2007) and user-producer interaction is seen as a key item in explaining innovation success (Rip and Kemp, 1998). The level of institutional overlap between the niche, the technological innovation system (TIS) and the regime is a main indicator for successful technology transitions (Markard and Truffer, 2008 and Geels, 2002). In other words, the breakthrough of an innovation requires strong cooperation within, and between niches, and between the niche and existing sociotechnical regime. To make these connections, the development of a shared frame of reference is important. This can be achieved through collaboration in a collective pilot project, which stimulates the negotiation process between these actors regarding central research questions, the research methodology and the assessment criteria. Consensus on these topics at an early stage supports agreement at later stages. This is also referred to as 'negotiated knowledge' (Koppenjan and Klijn, 2004). Moreover, it contributes to gaining trust in each other, which is an important lubricant in the contact between the different actors.

10.4.4.4 Coordination

Innovation requires a coordinated process to successfully connect the elements at the micro-, meso- and macro level, as well as the phases of FI, ARA and UI. The heritage of previous stages remains of great importance to subsequent stages. This is especially true if uncertainties in previous stages remain undiscussed, as to a certain extent happened in the cases of the EOGRTS and the SNT. Neutral process coordination is currently often missing. Developers and/or advocates of alternative tests are observed to put a lot of energy in testing and promoting the 3R model and thereby run the risk of going too fast and leaving those in doubt behind (see niche-based track). Disagreements in this domain are almost always discussed in a technical way discussing scientific data. However, in cases of controversy, scientific data show their subjective face. In other words, data often become part of the conflict and are frequently interpreted in such a way that it under-

scores the stakeholder's own perspective. For this reason, process coordination is needed to broaden the perspective and to stimulate an open and fruitful discussion on – potential - controversies. In addition, the process coordinator should supervise the process to scan for potential barriers and drivers along the way and to make sure that the planned actions are taken in a timely fashion.

10.4.4.5 Continuity

In addition to the previous four C's the fifth C of Continuity is added. The case studies of this thesis have shown that what has been achieved is largely related to the commitment and determination of innovational entrepreneurs. This is an important but vulnerable way of operating. Many projects are started and reach good results but fade away again after the funding is stopped or the committed individuals have left the scene. Therefore continuity is needed. Continuity not only requires long-term policy goals and the ongoing investment in the development and validation of 3R models.

It also requires a long-term planning, broadening the involved group of stakeholders and ongoing dissemination of results from the development of the 3R model to its regulatory acceptance and use.[135]

This chapter highlighted the importance of involving the social dynamics between the diverging actors and of targeting the psychological and communication factors that underlie existing controversies. It also discussed several options through which existing controversies can be bridged. It is important to note that this roadmap is meant to offer guidance in adopting a process-oriented approach. However, it does not pretend to be complete, nor does it intend to be a 'one-size-fits-all' solution.

The roadmap is a general design, which needs to be adapted to each case. Different cases may require different or additional steps or a different step sequence. The eventual approach should be contingent with the involved force field and the specificities of the regulatory frame for which the 3R model is destined.

By adopting a broader approach in dealing with the existing tardiness in 3R acceptance, and by targeting the social constructions connected to the conventional methods and to the 3R models, the existing social constructions can be reconstructed. Such a reconstruction process requires consorted action and substantive discussions at a meta-level, which clearly surpass the pros and cons on a scientific and technological level while helping stakeholders to reflect upon their assumptions. A bottom-up niche-based process may transform the dominant sociotechnical configuration, or it may fail to do so; while a top-down process may only prompt incremental changes that relieve tensions in the incumbent sociotechnical regime

135 The EPAA has for example shown that the dissemination of information on 3Rs has a direct impact on moving 3Rs methods from R&D to validation, acceptance and implementation. However, the absence of a process and/or organization/institution for post validation implementation support and dissemination of new methods that was identified as a gap hampering the adoption of 3 R models after their validation. http://ec.europa.eu/enterprise/epaa/3_activities/3_2_progress_reports/epaa_report_final_2007.pdf : consulted December 2015

(Berkhout et al, 2003). Therefore, a combination of a bottom-up and a top-down strategy is aimed for. Clear policy goals are needed to provide a solid frame to steer in the desired direction, and within it, a strong niche management is required to facilitate innovations. Furthermore, the discussion regarding the current test paradigm has to stay high on the political and institutional agendas. Such a consorted and coordinated approach, with a clear mutual goal and in which communication is the central engine, offers the best route to guide 3R models out of their niche and into the existing socio-technical regime.

10.5 Reflection

In this paragraph a reflection is given on the contributions of this thesis to the scientific and social debate and on its limitations.

10.5.1 Contributions of this research to the scientific and social debate

To start with, this research largely underlines what the dominant drivers and barriers are according to the literature on Technology Transitions (TT). The case studies conducted in the context of this thesis provided a detailed examination of single examples of TT and thereby provided insight knowledge on the dynamics of TT. Furthermore, they offered in-depth information on the process of regulatory acceptance and use and illustrated the developments in two product sectors which are seen to be the dominant sectors in terms of regulatory testing (see Chapter 2). Thomas Kuhn has argued that a discipline without a large number of thoroughly executed case studies is a discipline without systematic production of exemplars, and that a discipline without exemplars is an ineffective one. (Flyvbjerg, 2006, p. 27) This study has contributed to the creation of in-depth exemplars.

The case studies reflect the inherent wickedness of the central issue and the interrelatedness of the forces at hand. They offer valuable insights for these specific situations but they are also believed to offer valuable understandings for other cases of 3R regulatory acceptance and even for the movement towards alternate approaches of risk assessment and quality control. This belief is based on the fact that the factors we came across during the case studies are largely in line with earlier research conducted in this field (see Appendix I) and with the results of the panels conducted in the context of this research. Hereby, the case studies are likely to also provide reliable information about the broader class of comparable cases. For example, based on the empirical findings one can conclude that there is indeed a strong connectedness between the developments at the micro, meso and macro levels, and between the subsequent substages of FI, ARA and UI. Furthermore, the role of institutional entrepreneurs was emphasized by the two case studies. However, the weight of specific factors is likely to vary per sector and even per case.

A specification to the TT literature was made through the incorporation of the risk-averse context in which 3R models have to demonstrate their utility. Through this thesis a bridge was built between TT and risk regulation literature. Innovations normally already face difficulties in entering the existing regime. These are intensified by the risk-averse characteristics of the regulatory regime for the risk assessment of pharmaceuticals and chemicals. These regimes are even characterized as over-conservative, thus influencing user demands and amplifying barriers, such as the remaining uncertainties connected to 3R models. In addition, the position of the conventional model is fortified as a consequence of the resistance to change. It also means that sudden changes to the existing system are unlikely to happen. Transition patterns in a risk-averse environment are more likely to show a combination of empowerment and adaptation patterns, with new systems gradually growing out of the old system. This also means that of the 3R models, reduction and refinement models stand a better chance of being adopted.

A epistemological presupposition that was uncovered by this research is the stakeholders' strong focus on the generation, validation and dissemination of scientific data to stimulate progress. Most of the energy is put in dealing with the remaining scientific uncertainties and working towards the generation of additional data. This is driven by an underlying relentless belief in scientific progress to solve the remaining problems. It however neglects the fact that technologies are social constructions which are formed by a broad set of norms, values and underlying beliefs. This research provides information to balance this view. It illustrates that scientific data are often observed to exacerbate the existing controversies instead of solving them. Solving controversies requires alternative strategies, such as a discussion and negotiation between different points of view, as examined in this chapter. In other words, the process of acceptance and use of 3R models for regulatory purposes can only be understood and targeted by broadening the perspective. This means that stakeholders within this domain will have to broaden their view and try to understand and engage other perspectives (anti-narratives) in the field in order to overcome the remaining barriers and bridge the existing controversies.

As such, this thesis is a plea for natural sciences to open up to other disciplines (e.g. innovation science, sociology and policy sciences) to broaden the dominant technocratic perspective and to balance the instrumental rationality with value rationality, which is important to ensure that scientific and technical development does not take place without ethical checks. As described in Chapter 3, a scientist who conducts research based on value rationality sees it as his task to analyze the values of stakeholders regarding a certain theme in order to comprehend the dominant perspectives and their consequences in a specific domain and to offer alternative views where needed. This brings Flybjerg's argumentation (2006) as described in Chapter 3 into a new dimension. Not only should the social sciences adopt a phronetic approach, but the natural sciences should also be more susceptible to this perspective in order to decrease the risk of becoming too technocratic, and thereby overlooking the underlying aspects of remaining challenges. Such a value-based perspective is for example needed to reflect on the intended role that safety assessment and quality control testing should take in society and the role of animal testing and 3R models therein. Hence, there is

a need for a more realistic and functional risk-based approach, which takes the level of relevant exposure levels into account, or for a consistency approach which takes new ways of production of vaccines into account.

10.5.2 Limitations and suggestions for future research

When employing the findings of this research, one has to be aware of the following limitations. The first limitation is connected to the use of case studies as central research method. Case studies are very informative about these particular examples. However, due to the specificity of the particular context one cannot automatically aggregate the lessons to other cases. The pluriformity found in practice regarding the types of tests, the types of products, the involved stakeholders and the regulatory requirements makes it impossible to derive general conclusions and one-size-fits-all solutions. This means that every case will require an initial inventory of the forces at hand to explore the specificities that need to be taken into account when defining a transition strategy. Further in-depth knowledge is necessary to get a more complete understanding of the existing inertia in the field of regulatory acceptance and use of 3R models as an example of TT (technology transition).

This thesis provided the research area with relevant contextualized in-depth information. However, no direct solutions are offered for the powerful barriers such as the lack of harmonization and the level of risk aversion. Additionally, the barriers and the critical success factors were primarily collected amongst stakeholders that are already connected/committed to issue of the 3Rs. This means that no clear picture can be given as to what extent this issue is relevant and known by less involved stakeholders. This may well implicate that the barriers are higher in reality since to many stakeholders the issue of the 3Rs is simply not a priority. Additionally, the weighing of the factors was done by the involved stakeholders. Although via triangulation several factors were detected that are very likely to play a more dominant role than others, the weight given to these factors remains subjective.

Moreover, the case studies examined are examples of TT towards a specific 3R model. However, scientific knowledge is rapidly increasing and new test paradigms are arising and frequently brought into the debate on the transition from animal models to alternative ways of testing. These visions entail ways of shifting from animal experimentation to new testing strategies such as omics, high-throughput screening, Integrated Assessment and Testing Approaches (IATAs), Adverse Outcome Pathways (AOPs) and Modes of Action (MoA).[136]

136 *"Integrated testing strategies based on combinations of advanced in vitro and in silico methods that model the mechanism(s) of action and address the adverse outcome pathway are now the focus of research and development in several prototype initiatives".* (e.g. Blaauboer et al., 1999, Rovida et al., 2015).

In the field of biologicals, the consistency approach is a promising concept in moving away from *in vivo* testing (e.g. De Mattia et al., 2011).[137] These integrated approaches provide valuable concepts for the understanding of the mechanism(s) of toxicity, the assessment of specific endpoints and the quality control of vaccines. The contradictory forces at the macro level (i.e. the risk aversion versus concern for animal welfare) may be targeted through these more intelligent ways of risk assessment and quality control. This thesis touches upon these new and more intelligent ways of working, however the main focus was the acceptance and use of specific 3R models. Therefore, supplementary research is needed to specify what is needed to effectuate this additional paradigm shift.

137 This approach is *"based upon the principle that the quality of vaccines is a consequence of a quality system and of consistent production of lots with similar characteristics to those lots that have been shown to be safe and effective in humans or the target species"* (Hendriksen et al., 2008, p. 73).

APPENDICES

I

INITIAL STUDY

(SCHIFFELERS ET AL., 2007)

Factors Stimulating or Obstructing the Implementation of the 3Rs in the Regulatory Process

Marie-Jeanne W. A. Schiffelers[1], Bas J. Blaauboer[2], J. Martje Fentener van Vlissingen[3], Janne Kuil[4], René Remie[5], Joop W .G. M. Thuring[6], Manon A. Vaal[7] and Coenraad F. M. Hendriksen[8,9]

[1] Utrecht University, Utrecht School of Governance (USG), Utrecht; [2] Utrecht University, Institute for Risk Assessment Sciences (IRAS), Utrecht; [3] Erasmus University Medical Center, Erasmus Laboratory Animal Science Centre (EDC), Rotterdam; [4] Dutch Society for the Protecion of Animals (DSPA), The Hague; [5] Solvay Pharmaceuticals, Weesp; [6] Notox B.V., 's Hertogenbosch; [7] Utrecht University, Science Shop for Biology, Utrecht; [8] Netherlands Vaccine Institute (NVI), Bilthoven; [9] Netherlands Centre Alternatives to Animal Use (NCA), Utrecht University, Utrecht; all institutions are located in the Netherlands

Summary

Approximately 30% of animal use within the European Union (EU) is done to meet regulatory requirements. The tests are often repetitive in nature and may cause severe suffering, due to the procedures used and to rigidly predefined end points. In addition, product evaluation procedures often take long and are very expensive. Over the last decades the heavy reliance on animal experimentation in this area has met serious objections, both ethical and economical in nature. This study describes obstacles and opportunities to implement the 3Rs in regulatory animal testing. The findings are based primarily on interviews with legislators, regulators, industry, science and animal welfare organisations and reflect shared perceptions of these respondents. In order to increase the application of the 3Rs in regulatory testing a number of technical, political and social obstacles must be overcome. This study offers insight into the persistent character of regulatory animal testing and can function as a starting point for further discussion on how to tackle these problems. To this end, several recommendations are made ranging from strategic test approaches and data sharing to strengthening the policy network and improving communication between 3Rs experts and regulators. The study is an initiative of the national project group "Regulatory Animal Testing", which consists of a group of Dutch experts on animal testing working for a variety of organisations in the field.[1] They felt the need for cooperation to initiate a discussion at relevant levels and to identify possible solutions in order to implement the objectives of the three R's in testing for regulatory purposes without loss of scrutiny in safety and/or efficacy evaluation needed for product release.

Zusammenfassung: Was fördert und was hemmt die Einführung der 3R bei amtlichen Zulassungsverfahren?

Etwa 30% der Versuchstiere in der Europäischen Union (EU) werden in behördlichen Zulassungsverfahren verwendet. Die Tests wiederholen sich naturgemäss oft und können schweres Leiden verursachen, abhängig vom Versuch und von den strikt vorgegebenen Endpunkten. Darüber hinaus dauert die Klassifizierung von Produkten oft lange und ist sehr teuer. In den letzten Jahrzehnten wurde das grosse Vertrauen, das bei diesen Verfahren in die Tierversuche gestzt wurde, ernsthaft erschüttert, aus ethischen, aber auch aus ökonomischen Gründen. In dieser Studie werden Hemmnisse und Möglichkeiten beschrieben, die 3R in amtlichen Zulassungsverfahren einzuführen. Die Resultate basieren in erster Linie auf Interviews mit Abgeordneten und Angehörigen von Behörden, der Industrie, der Wissenschaft und von Tierschutzorganisationen; sie reflektieren die Auffassungen der Befragten. Um der Anwendung von 3R Methoden bei amtlichen Zulassungsverfahren stärkeres Gewicht zu verleihen, müssen eine Reihe von technischen, politischen und sozialen Hindernissen beiseite geräumt werden. Diese Studie bietet einen Einblick in den gegenwärtigen Stand der Zulassunsgverfahren und könnte als Ausgangspunkt für weitere Diskussionen dienen, wie die Probleme gelöst werden könnten. Zu diesem Zweck werden einige Empfehlungen ausgesprochen, die von strategischen Testabläufen und Datenaustausch bis zur Verstärkung des politischen Netzwerks und einer verbesserten Kommunikation zwischen 3R Experten und Zulassungsbehörden reichen. Die Studie ist eine Initiative der nationalen Projektgruppe "behördlich vorgeschriebener Tierversuche", die aus einer Gruppe holländischer Experten für Tierversuche besteht, die für verschiedene Organisationen auf diesem Gebiet arbeiten[1]. Sie erachteten eine Kooperation als dringend geboten, um eine Diskussion auf relevanten Ebenen zu starten und mögliche Lösungen aufzuzeigen. Die 3R Prinzipien sollen bei amtlichen Zulassungeverfahren eingeführt werden, ohne einen Verlust an Sicherheit und Wirksamkeit bei der Produktfreigabe befürchten zu müssen.

Keywords: regulatory animal testing, registration, regulatory requirements, 3Rs, validation, alternatives, stakeholders, stream model, policy

Received 25th July 2007; received in final form and accepted for publication 15. 10. 2007

*This study was supported by grants from ZonMw (the Netherlands Organisation for Health Research and Development, program for alternatives to animal testing), Solvay Pharmaceuticals and Organon.

[1]Underlying findings of this article are available at at http://www.bio.uu.nl/scienceshop/ ; under publications.

1 Introduction

In 2002, approximately 10.7 million animals were used for experimental purposes in the European Union.

The experiments were conducted in the following areas:

About 30% of these animal tests are carried out to comply with regulatory requirements (from within the following categories: R&D for medical, veterinary and dentistry products, toxicology/safety evaluation and production and quality control)[3]. These legal requirements prescribe which experiments must be conducted in order to licence and release a compound or product onto the market for human, animal or environmental applications. This implies that these tests are mainly applied in the areas of production and quality control of human and veterinary medicine and of toxicological and safety evaluation of other compounds or products, e.g. pesticides, household products, cosmetics, food additives. Many national and international parties, often with divergent interests, are involved in setting these requirements, aiming at efficacy, consumer safety and environmental protection.

Regulatory animal testing is usually laid down in standard protocols. It is often repetitive in nature and more likely to cause severe suffering than other types of animal testing, due to the procedures used and predefined experimental end points, in contrast to most other types of animal research. Because of these characteristics, regulatory animal testing is an important area to evaluate for 3R policy opportunities.

In 1959, W. M. L. Russell and R. L. Burch proposed the implementation of the 3Rs principle to animal experimentation in "The Principles of Humane Experimental Technique" (Russell and Burch, 1959).

1. Replacement: the substitution of insentient material for conscious living higher animals.
2. Reduction: reduction in the numbers of animals used to obtain information of a given amount and precision.
3. Refinement: any decrease in the incidence or severity of inhumane procedures applied to those animals, which still have to be used.

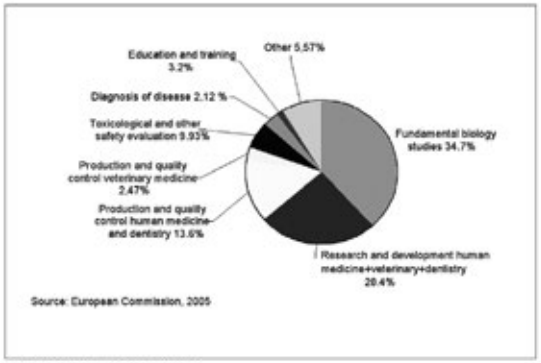


Fig. 1: Purposes of experiments

Much can be gained in testing for regulatory purposes with respect to the 3Rs (replacement, reduction, refinement) by looking critically at the approach to animal testing, the necessity of the test, the way tests are conducted, the possibilities to use alternative methods, etc. (IVTIP, 2000).

It must be emphasised that when this article mentions alternatives to animal testing, this refers to all three R's and not exclusively to the R of "replacement".

In practice however, it proves difficult to implement the 3Rs as an important concept in testing for regulatory purposes. In order to identify opportunities and obstacles in implementing the 3Rs within testing for regulatory purposes this study examines the decision-making process underlying regulatory animal testing.

2 Investigative approach

A policy-making process is typically surrounded and influenced by a variety of factors and actors. The process is permanently subject to conflicts (of interest) between the various stakeholders in society and on the administrative level itself. Both the object and the result of the policy are the outcome of a permanent political (power) struggle. In order to explain how various factors and actors influence the policy-making process, this analysis makes use of the stream model (Kingdon, 1995). Although the theory of the stream model is designed to analyse the process of agenda setting, it proved to be a very useful model in this context to describe and understand the factors influencing the implementation of the 3Rs in the regulatory process.

In the stream model, the policy-making process is regarded as an organised anarchy in which problems, parties and solutions each behave according to their own dynamics. The model postulates three streams: problems, solutions and political/administrative developments. The problem stream represents the multitude of issues in society that need to be addressed. The political/administrative stream reflects the political and administrative actors caught up in a continuous battle for votes, budget and support. Finally, the solution stream reflects the ideas, plans and pilot projects – developed by parties, lobby groups and civil servants – which may lie around unused for years. Developments within these streams determine whether opportunities arise for the streams to connect. Such opportunities for the confluence of streams are known as policy windows.

[3]In the Netherlands this percentage is 29% (Food and Consumer Product Safety Authority (VWA), "Zodoende"; Annual Report on Animal Experimentation and Laboratory Animals, 2005, The Hague, The Netherlands)

When the streams meet, it is possible to effect changes in policy or to initiate new policy (Walraven et al., 2002).

Chances to produce new policy or to modify existing policy can be created and used by a so-called policy entrepreneur (Hart't et al., 1995). The entrepreneur has capacities to function as the initiator of new policy. The personal characteristics of the entrepreneur and the social relevance of the group he or she represents are two of the factors that determine how successful a confluence of streams will be.

In terms of the stream model, this study also aims at providing insight into the influences and entrepreneurs that can ensure that the various streams successfully converge to create a policy window.

The European legislative process is long and complicated. It involves many different actors who have either a formal or an informal opportunity to contribute to policy shaping and lobbying the decision-makers. The multitude of stakeholders and the barter of issues from other policy sectors make it hard to predict how a particular initiative will fare. In view of these characteristics, the legislative process is also difficult to predict. At the European level, one can hardly speak of coherent and strong policy-making. The stakeholders of influence do not operate in a focussed formal framework but rather in a network with flexible relationships.

This survey is exploratory and descriptive in nature. The identification and description of factors influencing the policy-making process is based on the qualitative research methods of desk research and a series of interviews. Approximately thirty stakeholders were interviewed in-depth to get an overview of their views on the complex issue of regulatory animal testing. The survey is therefore mainly based on the respondents' ideas and perceptions of the factors they see as influential in the decision-making process at the European level. Many categories of putative stakeholders were considered first. From these, several categories of stakeholders were selected based on their assumed significance. The following categories were defined to select respondents for this research:

- Legislators (policymakers);
- Regulators (governmental agencies and authorities responsible for the implementation and maintenance of laws and regulations, with the authority to approve or reject the release of products on the market);
- Science (academia, research institutes);
- Industry;
- Animal interest organisations.

The respondents came from the European and Dutch context and were selected in close consultation with the project group "Regulatory Animal Testing". A complete list of respondents can be found in the acknowledgements.

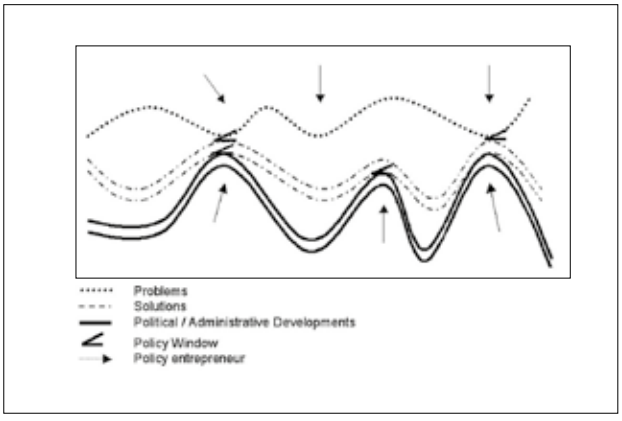


Fig. 2: Representation of the Stream Model
Sources: Walraven et al., 2002; van de Graaf and Hoppe, 1989

3 Results

Regulatory animal testing is a persistent element in assessment procedures for licensing a compound or product for release onto the market. Even though the number of alternative test methods keeps increasing, these new methods are not automatically implemented in the guidelines for assessment procedures in order to replace the more classical animal tests. This is due to a combination of technical, political/administrative and societal factors. In order to accelerate the implementation of alternative methods in regulatory testing, a number of obstacles must first be overcome. What follows below is an overview of the most relevant factors influencing the use of the 3Rs in assessment protocols. These factors have been grouped using the problem stream and the solution stream of the above mentioned stream model. The political and administrative factors are the focus of this study. Next to these factors, several technical and societal factors have been identified.

3.1 Problem Stream

The problem stream represents a combination of problems, which have to be addressed in order to reduce the extent of animal use for regulatory requirements.

3.1.1. Technical problems

Availability of alternative methods

Most alternative test models developed so far are intended to replace relatively simple test methods e.g. for local toxicity / one target organ. However, most remaining animal experiments are complex tests for which it is difficult to find alternatives. For example, many animals are needed for tests in reproductive toxicology (embryo toxicity) and systemic toxicology, which are much more complicated to entirely replace. Science is now facing the task of developing such complex alternative test methods, either by refinement, reduction of the number of animals needed per test, or replacement. Opinions on the feasibility of this task are divided.

Technical expertise

For the methods that have been developed in the area of the 3Rs, there might be more room for application and acceptance. Up till now, such methods are often insuffi-

ciently or too narrowly publicized and are known to a limited audience and therefore used to a limited extend.

Technical expertise is a crucial factor in the decision making process whether or not to implement the 3Rs in safety and efficacy testing for regulatory requirements. Since the field is very complex, only a select number of experts are able to contribute to discussions on the matter. Legislators and regulators might have limited knowledge of alternative methods when they do not have detailed and updated scientific information at their disposal. This makes it difficult for them to evaluate the merits of these test models. They are therefore strongly influenced by the extent of scientific consensus concerning animal experiments and alternatives. Without this type of scientific backing, politicians are reluctant to take a political stand. Furthermore, scientists who do have expertise on alternative test models often lack the knowledge of, and access to, the policymaking process, and therefore cannot effectively inform legislators and regulators about these possibilities (Sauerborn et al., 1999). This hinders the necessary communication between legislators and regulators on the one hand and scientists in the field of the 3Rs on the other. Much of the knowledge with regard to the 3Rs remains unused, according to the experts consulted.

Availability of data
Another important barrier with regard to implementing the 3Rs in regulatory animal testing is that, for reasons of competitiveness, industry is reluctant to make research data available. This means that a vast volume of valuable information regarding the 3Rs already exists but is not available to third parties. And although much is undertaken to harmonize the registration requirements between countries and to accommodate the mutual acceptance of test results across borders, this lack of available data still leads to unnecessary use of classical "non 3R" testing models.

"Traditional" versus "new" methods
The present generation of regulators was mainly educated some 20 to 30 years ago when the credo still was "*in vivo veritas*". The new generation of regulators will most likely incline towards *in vitro* methods. There is, however, a risk that each "school" will exclude the other methods to a certain extent. The new generation of scientists/regulators may run a risk by making the transition too quickly, thus missing the opportunity to convince others by demonstrating valid evidence, while the older generation may be too dismissive of novel *in vitro* methods. This would impede the acceptance of alternatives. It should be emphasised that a "stand-alone" position, of either *in vivo* or *in vitro* methods, is neither feasible nor desirable.

Validation process
Before newly developed alternative methods can be put into practice, these methods first have to be validated. Although the number of alternative tests developed and accepted has risen sharply in the recent past (Balls, 2002), validation is looked upon as a difficult and time consuming process dominated by a small number of interest parties. The process often leads to interpretation problems, since the scientific prestige of the various players may be at stake. Validation of alternative methods has therefore become a process that takes many years. This shows a sharp contrast to *in vivo* methods that have never been formally validated and are widely accepted.

Implementation
Furthermore, the process after validation is often even more time consuming, since alternative methods must prove themselves many times over before they are accepted by regulators and become part of legislation (Spielmann, 2000).

Why acceptance takes such a long time will be discussed in the following paragraphs.

3.1.2. Political and administrative problems

The main political and administrative factors considered to be barriers when trying to implement the 3Rs for regulatory purposes are: agenda setting, legislative/regulatory context and acceptance of validated methods.

Agenda setting
The EU concerns itself primarily with the internal market of the now 27 European states. Other issues near the top of the agenda are safety and risk limitation. Animal welfare as such has a lower priority. This also means that governments and industry have limited budgets for developing alternatives, particularly because they are aware that an alternative method, once developed and validated, will be subject to a very time consuming period of negotiations before it can be accepted for regulatory purposes. The growing concern regarding consumer safety and environmental impact is translated into more regulatory requirements for products, and will therefore result in an increase of the number of animal tests used for regulatory purposes.

Legislative/regulatory context
Two types of legislation are relevant to the use of experiments on animals. The first is "horizontal legislation" pertaining to animal experimentation and multilateral agreements. The second is "vertical" or "sectorial legislation". The latter regulates the activities of a particular sector, for example the approval of pharmaceuticals, which indirectly affects animal experimentation. In principle, vertical legislation must take horizontal legislation into account. For example, Directive 86/609/EEC, which regulates the protection of animals used for experimental and other scientific purposes, should be taken into account by vertical legislation (Article 21). Directive 86/609/EEC applies the "no, unless" principle. This directive stipulates that alternatives, if available, should be used (Article 7). Some vertical European legislation already explicitly refers to this directive. Moreover, even when this is not the case, the provisions in the horizontal legislation about animal testing must be respected in all other regulations. In practice, however, this is often done insufficiently or not at all due to a combination of factors. There is too little cooperation between the EU committees that draft "safety and efficacy regulations" and those that develop "animal welfare regulations" (de Leeuw, 2004). As a result, when directives are revised, they continue to include requirements for animal tests even after alternatives have been validated and used by corporations. This is why different directives can, and often do, conflict. There are no traffic regulations indicating which directive has priority in case of conflicting rules in different directives (de Leeuw, 2004).

In addition, Member States are given relatively great discretion to interpret European directives within the limits of national law, for instance in the area of pharmaceuticals. In the EU alone, there are more than 800 laws, regulations, directives and other documents regulating the use of laboratory animals to ensure the safety of humans, animals or environment (de Leeuw, 2004). These regulatory requirements are often used in quite a rigid manner, the so-called "tick box approach". This refers to a rigid method of quality and safety control by assessors of products and compounds. One can speak of the tick-box approach when assessors simply request every test prescribed in the protocol to be executed without having a critical look at the necessity of conducting all these tests. This results in great differences between EU Member States and the extent to which they are open to alternatives.

Finally, European legislation dealing with market related issues usually outweighs animal welfare issues. Animal welfare issues are primarily seen as the responsibility of the individual Member States, even as Article 22 of the Directive 86/609/EEC obliges the Commission to collect information about the legislative framework for regulatory purposes in each of the Member States, and to evaluate these for the protection of animals.

Acceptance of validated methods
As mentioned in paragraph 3.1.1, successful validation does not guarantee acceptance. The slow pace of acceptance is caused by a combination of factors. Firstly, legislators and regulators are facing increasing demands for consumer safety and risk minimisation. These authorities are expected to take this increasing demand for safety into account when developing and implementing policies. In the area of policy implementation, the regulators in particular are considered to be reluctant to implement the 3Rs in evaluating testing protocols and dossiers for product registration. One main reason for this reluctance is the heavy responsibility regulators bear for the safety of products they allow onto the market. In addition, regulators are often relatively unfamiliar with the properties and scientific qualities of relevant, but relatively new, alternative test methods. They therefore tend to adhere to classic (animal) models and are sceptical towards accepting new methods with different scientific end points. A number of respondents believe legislators and regulators are waiting for a scientific consensus before taking the risk of incorporating alternatives into policy. This process of reaching scientific consensus, by the very nature of scientific methodologies, is difficult to achieve and takes a long time. As a consequence, so are the changes in policies.

Along the same lines, industry is identified as a conservative force, preferring to play safe by anticipating the strict registration requirements regulators will set. As a result, industry tends to stick to the old, animal based models, even when alternative models are available.

3.1.3. Societal problems

The societal factors that are perceived as barriers when trying to reduce the extent of animal use for regulatory purposes can be divided into two categories: public opinion and risk minimisation.

Public Opinion
Although there is public resistance to animal testing in general, the growing focus on consumer safety has so far taken priority over the concern for animal interests. Public opinion on animal testing depends also on the purpose for which the animals are used. For example, animal testing for medical purposes is much more accepted than safety testing of cosmetics (Aldhous et al., 1999).

Public opinion has a powerful potential to influence the attitude of politicians and industry towards animal testing and alternative methods and is of great importance to the image of companies and products. This has prompted various companies to promote alternatives as part of a Corporate Societal Responsibility concept. Moreover, if animal testing is in the public eye, it usually becomes of more interest to politicians too. However, when product safety is a hot topic, the concern about the welfare of animals used in experiments tends to loose any priority.

The influence of public opinion is therefore ambiguous. It has the potential to encourage the development and implementation of alternative methods in case there is a high public concern regarding animal welfare. Currently it is more likely to hamper the implementation of the 3Rs due to the growing emphasis on consumer safety.

Risk minimisation
The tendency within Western society towards a so-called "zero risk" concept is a serious threat to replacing, reducing and refining regulatory animal testing. The ongoing call for extra research based on the precautionary principle is a manifestation of this. According to many respondents, the advocates of the "zero risk" concept are insufficiently aware of the consequences this has in terms of the increased use of animals for testing purposes.

3.2 Solution stream

The solution stream contains several possibilities to overcome (some of) the barriers described above. In order to effectively reduce the number of animals used for regulatory purposes a mix of technical, political/administrative and societal solutions is needed.

3.2.1 Technical solutions

Options for reduction and refinement
Although it is difficult to develop alternative methods for more complex tests, much progress has been made in reducing the use of animals and in refining methods for complex end points, such as local acute toxicity, some vaccine tests and pyrogen testing. When it comes to the more complex experiments, reduction and refinement seem to offer the best prospects for the time being. Respondents expect good results in the near future from developments aimed at reducing the number of laboratory animals used to test scientific hypotheses and refining tests to limit the suffering of laboratory animals. The most substantial gain according to interviewees is expected from strategic test approaches and data sharing.

The strategic test approach, also referred to as step-by-step approach in toxicity tests and consistency approach for biological products, offers a chance to reduce the number of animal tests (Health Council of the Netherlands, 2001). It is important to apply strategic planning before carrying out any animal experiments in an effort to ensure appropriate implementation of the 3Rs. For example, there is tremendous potential for the increased use of screening tests to assist in prioritising chemicals for further test-

ing for their pharmaceutical potential (Combes et al., 2002). This could result in cancelling subsequent *in vivo* tests in case *in vitro* tests indicate a compound to be toxic. Currently, animal experiments are often carried out regardless of such initial results.

The merit of sharing data is widely accepted. It is essential for verification purposes, secondary analyses and the information can be used for (computerised) modelling of pharmacodynamic and putative properties of compounds. In order to be able to maximise the implementation of the 3Rs, scientific inquiry must be open to the public domain. Since the advantages of data sharing are sufficient, a considerable amount of energy should be put into overcoming one of the main hurdles by finding ways in which data can be shared without jeopardising privacy or breaching confidentiality promised to data providers (Fienberg, 1985).

Broaden communication on the 3Rs
As mentioned, scientific knowledge is often underused due to the poor communication of results by scientists to other stakeholders (Sauerborn et al., 1999). In order to promote the implementation of scientific developments with regard to the 3Rs, these developments should be made public to a wider audience of legislators and regulators than is done presently.

3.2.2 Political/administrative solutions

Influencing agenda setting
Animal welfare organisations represent important actors in the struggle to get animal welfare and the 3Rs on the political agenda and keep them there. Therefore, ongoing attention from these non-governmental organisations (NGOs) to this subject is an important driving force.

Harmonisation of legislation and regulations
Even though harmonisation is an onerous and time-consuming process, it seems to be a precondition to reduce any unnecessary testing caused by insufficient coordination between various legislations and regulations. This has been achieved for pharmaceuticals through the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH)). Harmonisation can give an important boost to reduce unnecessary testing. Furthermore, Mutual Acceptance of Data (MAD) ensures that research is not needlessly duplicated. Therefore, it is of great importance to develop a Mutual Recognition Agreement, or MRA. However, as is the case with harmonisation, developing an MRA is a laborious process.

Towards acceptance of validated methods
Acceptance of validated methods is a mixture of awareness of the problem, commitment to change, availability and knowledge of alternatives and positive experiences with these methods, finding them reliable to reach the relevant scientific end points. The new generation of assessors/regulators (expert reviewers of test data in the process of registration of products) is expected to have awareness of the concept of the 3Rs because of greater exposure to the issue during education.

Combination of technical expertise and political intuition
Experts involved in elaborating safety requirements and tests have mainly technical expertise. In order to effectively bring about change, experts should also be familiar with the policy-making process and forces at play in politics and administration. In risk assessment, the public perception of product safety is also relevant, e.g. recent incidents may cause (public, political) risk perception to change for a short or longer term, and administration is bound to be sensitive to risk perception as a political fact. Adequate awareness of factors of a different nature than just scientific/technological expertise and fact-finding creates a better chance to effectively exert influence.

Ethical review committees
Practically all Western European countries have procedures for the ethical review of a proposed animal test. Although Member States vary widely in terms of the organisation of the review process (composition of review committees, their tasks, status, and the levels at which tests are reviewed), ethical review committees could play an important role in being loud and clear when they believe certain regulatory tests are outdated and could be replaced or modified, e.g. by contacting the national representative in the regulatory bodies.

In order to bring the potential contribution of alternative approaches for risk assessment to the attention of regulators, it is important to have 3Rs experts represented at expert committees of regulatory authorities, both at the EU and at the national level.

3.2.3. Societal solutions

Risk communication
To tackle the problem of the tendency towards "zero risk level", better information about the risk/benefit balance must be provided. First, it is necessary to clearly convey the message that zero risk cannot be achieved by any means. Second, open communication about potential risks creates an opportunity to bring risk acceptance back to proportions that are more realistic and fosters individual freedom of choice.

3.3. Stakeholder analysis

Four stakeholder groups are seen by the respondents as most dominant in the policy-making process and the implementation of 3 Rs for regulatory purposes:

1. *Regulators*, being the assessors of new products and compounds, have most influence on the implementation of legislation, and, therefore, on the feedback loop between applicant and authority. Regulators also exert direct influence on the policy-making process in their position as experts who advise legislators on drafting new policy or revising existing policy. Regulators also have authority over the industry regarding application and release of their compounds and products.
2. *Industry*, in turn, exerts great influence on legislators through strong lobby and expertise as well.
3. *Animal welfare organisations* influence the policy-making process directly (political lobby) and indirectly (through public opinion). This influence is aimed primarily at the initial stages of the process: the agenda-setting stage and partly the policy-making stage. It is mainly by their competence to mobilise public opinion that these organisations have power to influence the political agenda. Therefore, they

can be a driving force behind reforms that implement the 3Rs. However, animal welfare organisations have very little direct influence on regulators, respondents remarked, since regulators take their cues first and foremost from the heavy responsibility they bear.

4. *Experts*, who are shown in this study to have a great deal of influence on the development of policy, are mainly found within the stakeholder groups: regulators, scientists and industry.

Although more and more legislation is now formulated at EU level, the Member States themselves can ultimately be regarded as dominant actors in the political arena. After all, it is the Member States that provide the experts who help draft and revise regulations at every level. The Member States' influence is even greater because of the necessity to implement EU regulations into their national legislation. European legislation usually leaves Member States enough discretionary room for their own interpretation of policy. Consequently, (national) experts play a dominant role in policy-making, while regulators play a dominant role in policy interpretation and implementation.

4 Conclusions and Recommendations

Regulatory animal testing is perceived by a vast majority of the respondents as a very persistent element in the assessment procedures for licensing a compound or product for release onto the market. Even though the number of alternative test methods keeps increasing, even scientifically validated alternative methods are not easily included in assessment procedures. In order to effectively replace, reduce or refine animal testing for regulatory purposes, first of all a common understanding about the nature and importance of the problem needs to be established. Only a concise problem definition will enable tackling the problem in an effective manner. Therefore, there seems to be a fair level of consensus about the fact that there is a problem. Stakeholders however seem to have different opinions on the level of priority it must have in comparison to other problems that need to be addressed. To tackle the problem of regulatory animal testing, it is of special importance that regulators and industry grant it priority. Experts in the field of the 3Rs and NGO's concerned about animal welfare can play an important role in keeping this issue on the agenda of these stakeholders. The problem of regulatory animal testing should be addressed by combined technical, political/administrative and societal solutions.

Three categories of opportunity can be used to tackle the problem of the large number of animals used within regulatory testing.

1. Technology: There is a need for the development and validation of new, supplementary techniques that need to be accepted by a process of expert peer review to replace, reduce or refine animal experimentation.
2. Communication: There is considerable room for improvement of the exchange of knowledge between stakeholders about methodologies, results, etc. The necessary improvements would start with basic awareness, leading up to the implementation of a communication strategy.
3. Co-ordination/harmonisation: Supported by a better communication strategy, the stakeholders should coordinate their actions more closely. The desired result is harmonisation, i.e. the dovetailing of legislation and regulations in various regions and sectors.

In terms of the stream model, these three categories each contribute in their own way to a convergence and confluence of the three streams (the solution, problem, and political/administrative streams), which then can create a policy window. The technology category deals with new ways to enlarge and improve the solution stream, while the communication category offers a way to bring the political/administrative stream and the problem stream closer to the alternatives stream. Coordination and harmonisation have a great potential to reduce regulatory animal testing.

In order to create new implementation opportunities for the 3Rs within the regulatory framework, progress is needed in all three streams, and the resulting improvements must subsequently be brought together. Entrepreneurs or "advocates" who seek to improve conditions for the application of the 3Rs principles can facilitate this. The dominant actors who are leading toward implementing the 3Rs in regulatory animal testing can play the role of policy entrepreneurs. Some examples of such actors are: innovative companies in the field, experts in the field of the 3Rs, animal welfare organisations, ethical review committees, the inspectorates and committed individuals in any stakeholder group.

Conclusion

Regulatory animal testing is deeply ingrained in the procedures for evaluating compounds and products before they are allowed onto the market. Society, however, is growing increasingly critical of such animal tests required by protocol. Hence, initiatives from a variety of backgrounds are taken to reduce regulatory animal testing. This study is one of those initiatives. By analysing the technical, political and administrative and social factors that influence the use of regulatory animal testing, the researchers aimed to contribute to the quest for possible solutions and to stimulate further discussion in order to reduce regulatory animal testing. Both the respondents in the qualitative inquiries, and the discussions in different stakeholder debates, have shown that the problem is recognized and that progress in this field would be much welcomed, for various reasons (protection of animals, cost of testing, and access to a common market by harmonized regulatory processes) and is therefore viable.

In order to stimulate further reduction of regulatory animal testing this article ends with the following recommendations:

Recommendations

- Invest in data sharing, retrospective analyses and strategic test approaches;
- Use risk communication in order to influence the level of risk acceptance;
- Make the costs of conducting animal tests more visible;
- Visualise the limitations of current animal testing procedures;
- Widely publicise available (validated) alternatives;
- Improve communication between stakeholders;
- Strengthen the policy network;
- Harmonise various laws and regulations concerning product registration.

References

Aldhous, P., Coghlan, A., and Copley, J. (1999). Let the people speak. *New Scientist 162(2187)*, 26-31.

Balls, M. (ed.) (2002). Alternatives to animal experiments: Progress made and challenges ahead. The proceedings of the ECVAM Status Seminar June 2002, JRC Ispra, Italy. *ATLA 30, Suppl 2*, 1-243

Combes, R., Schechtman, L., Stokes, W. S. and Blakey, D. (2002). Recommendations on Best Scientific Practices for Subchronic/Chronic Toxicity and Carcinogenicity Testing In: proceedings of The International Symposium on Regulatory Testing and Animal Welfare:, *ILAR J. online 43 Suppl.:* Regulatory Testing and Animal Welfare, accessed 12 December 2006, 112-117.

Fienberg, S. E., Martin, M. E., Straf, M. L. (1985). Sharing Research Data, Committee on National Statistics, Commission on Behavioral and Social Sciences and Education, National Research Council. Washington D.C., USA: National Academy Press.

European Commission (2005). Fourth report on the Statistics on the number of animals used for experimental and other scientific purposes in the Member States of the European Union, Commission of the European Communities, COM (2005) 7 final, Brussels, Belgium.

Graaf, H. van de and Hoppe, R. (1989). *Policy and Politics (Beleid en politiek).* Muiderberg, The Netherlands: Coutinho publishing company.

Hart't, P., Wille, A., van der Meer, F. M. et al. (2002). *Politiek ambtelijke verhoudingen in beweging*. Amsterdam, The Netherlands: Boom publishing company.

Health Council of the Netherlands (2001). Toxicity testing: a more efficient approach. Health Council of the Netherlands, publication number 2001/24. The Hague, The Netherlands.

IVTIP (In Vitro Testing Industrial Platform) (2000). The Role of an Industrial Platform in the area of in vitro testing; Position paper October 2000, www.ivtip.org/PosPapers/pospaper2000.html, accessed 12 december 2006

Kingdon, J. W. (1995). *Agendas, Alternatives and Public Policies*. 2nd ed. New York, USA: Longman Pub. Group.

Leeuw, W. de (2004). De ethische toesting van dierproeven: wat heeft Europa ons te bieden, In J. Swart, J. Wolters, and H. Zwart (eds.), *DEC's in discussie: de beoordeling van dierproeven in Nederland.* Budel, The Netherlands: Publisher Damon.

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

Sauerborn, S., Nitayarumphong, S., Gerhardus, A. (1999). Strategies to Enhance the Use of Health Systems Research for Health Sector Reform, *Tropical Medicine and International Health 4 (12)*, 827-835.

Spielmann, H. (2000). Would Sisyphus meet the challenge of validation from test development to global regulatory acceptance? In M. Balls, A. M. van Zeller and M. E.Halder (eds.), P*rogress in the Reduction Refinement and Replacement of Animal Experimentation* (27-38). Amsterdam: Elsevier.

Walraven, G., van Erp, S., Knegtel, M. (2002). *Beleid komt niet vanzelf; gemeentelijk gezondheidsbeleid in NO-Brabant*. PON intsituut voor advies onderzoek en ontwikkeling in N-Brabant, Tilburg, The Netherlands.

Acknowledgements

We thank: I. Arendzen (Dutch Ministry of Health, Welfare and Sport / Inspectorate of the Food and Consumer

Product Safety Authority; VWA), E.C. de Bordes (Faculty of Veterinary Medicine, Utrecht University),

B.R.A. van den Bos (former Dutch MEP for the D'66 party), J. H. Fentem (Unilever- Safety and

Environmental Assurance Centre), B.J. Fernhout (Intervet), B. Garthoff (Bayer CropScience;

European Consensus Platform for Alternatives (ecopa) and the European Federation of Pharmaceutical

Industries (EFPIA), A. Gautrais (Enterprise DG, European Commission), E.R.M. Geuns (Solvay), M. van der Graaff (Nefarma), B.C. Hakkert (national coordinator OECD/EU, Directives programme, National Institute for Public Health and the Environment (RIVM), T. Hartung (European Centre for the Validation of Alternative Methods (ECVAM); Joint Research Centre of the European Commission), M. Heinen (Eurogroup for Animal Welfare), E. Honig (Intervet), H.B.W.M. Koëter (scientific director of the European Food Safety Agency (EFSA), R. Kroes(†) (former scientific director of the Interfaculty Institute for Risk Assessment Sciences (IRAS); Utrecht University) and former president of EUROTOX and the International Life Sciences Institute (ILSI), J.W. van der Laan (Medicines Evaluation Board (CBG); European Agency for the Evaluation of Medicinal Products (EMEA), W.A. de Leeuw (Inspectorate of the Food and Consumer Product Safety Authority; VWA), S. Louhimies (Environment DG, European Commission), G.J. Mulder (Leiden/Amsterdam Center for Drug Research (LADCR), I.F.H. Purchase (ICI/Zeneca, University of Manchester), T. Rijnders (Vice President Research, Organon), S.C. Schutte (Organon), R. Stolp (Intervet), P.W. van Vliet (Health Council of the Netherlands), J.M.G. Vorstenbosch (Ethics Institute, Utrecht University), D. Wagner (former principle administrator Organization for Economic Co-operation and Development (OECD), for their cooperation to this survey.

Correspondence to

Marie Jeanne W. A. Schiffelers
USBO Advies
Utrecht University
Bijlhouwerstraat 6
3511 ZC Utrecht
The Netherlands
e-mail: m.j.w.a.schiffelers@uu.nl

II

THE RABIES VACCINE SURVEY

(SCHIFFELERS ET AL., 2014a)

Biologicals 42 (2014) 205–217

Contents lists available at ScienceDirect

Biologicals

journal homepage: www.elsevier.com/locate/biologicals

Replacing the NIH test for rabies vaccine potency testing: A synopsis of drivers and barriers

Marie-Jeanne Schiffelers [a,*], Bas Blaauboer [b], Wieger Bakker [a], Coenraad Hendriksen [c,d]

[a] Utrecht University School of Governance, Bijlhouwerstraat 6, 3511 ZC Utrecht, The Netherlands
[b] Utrecht University, Institute for Risk Assessment Sciences (IRAS), P.O. Box 80.178, 3508 TD Utrecht, The Netherlands
[c] Institute for Translational Vaccinology (InTraVacc), P.O. Box 450, 3720 AL Bilthoven, The Netherlands
[d] Utrecht University, Faculty of Veterinary Medicine, Department Animals in Science and Society, The Netherlands

ARTICLE INFO

Article history:
Received 10 July 2013
Received in revised form
18 December 2013
Accepted 8 April 2014
Available online 15 May 2014

Keywords:
Rabies vaccines
Potency testing
3R models
Regulatory acceptance
Barriers and drivers
Multilevel innovation perspective

ABSTRACT

Approximately 70% of animal use is utilized to demonstrate quality control of vaccines. Especially rabies vaccine potency testing, using the NIH challenge test, involves objections in terms of scientific relevance, animal welfare concern and costs. Several 3R models have been proposed to refine, reduce or replace this test. Some are formally incorporated into regulatory requirements, but actual regulatory acceptance and use by industry lags behind, raising the question concerning which factors influence this process. This question is answered by a combination of literature review, interviews and a survey among 50 rabies vaccine experts. The findings are analyzed using the multilevel perspective on technology transition, which distinguishes 3 levels of factors influencing innovation acceptance. At the micro level (where 3R models are developed and validated) the dis-advantages of, and fractional experience with, 3R models, scarce data sharing and demanding validation processes exist. The meso level (existing regulatory regime) encloses the barriers of the 'gold standard', the lack of harmonization and the driving force of legislation stimulating 3Rs use. The macro level (the societal context) combines risk aversion and increased concern for animal welfare. Regulatory acceptance and use of 3R models requires dedicated stakeholder communication, cooperation and coordination at all three levels.

1. Introduction

Vaccines are often produced by or derived from living micro-organisms. As a result they enclose a complex molecular structure of antigens and their quality might vary from batch to batch. Other components, like preservatives and adjuvants can interfere with the quality of the product [1]. The complexity and potential variability of vaccines require that products undergo extensive characterization before being licensed. Furthermore, every vaccine batch or lot[1] has to be subjected to quality control (QC) to determine its purity, safety, efficacy and potency. The requirements for this QC testing are generally described in the monographs of the European Pharmacopoeia (Ph. Eur.) and of national Pharmacopoeias such as of the U.S. or Japan. Furthermore, the guidelines of the World Health Organization (WHO) and the World Organization for Animal Health (OIE) give guidance on how to ensure to vaccine quality.

Most of the tests used for the QC of classical vaccines, like rabies vaccines, are based on 1950s era technology relying heavily on live animal models [2,3]. These tests have seen little change since their adoption and, as a result, are often flawed from current scientific and ethical perspective.

It is estimated that approximately 70% of all animal use for vaccine related purposes is required for batch release testing.[2] The majority of the animals are used to meet potency testing requirements. The strong dependency on live animal models makes

* Corresponding author. Tel.: +31 30 2539319; fax: +31 30 2537200.
E-mail address: m.j.w.a.schiffelers@uu.nl (M.-J. Schiffelers).

[1] The term batch usually refers to veterinary vaccines while the term lot refers to human vaccines. This paper targets both the veterinary and the human rabies vaccine sector. A choice is made to use the term batch when speaking about rabies vaccines in general.

[2] http://www.google.nl/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=2&ved=0CDYQFjAB&url=http%3A%2F%2Fec.europa.eu%2Fenterprise%2Fepaa%2F3_events%2F3_3_workshops%2Fwp4_pharm_5_vaccins_pps&ei=7fMQUYuMJoKq0QX0oID4AQ&usg=AFQjCNEyfkqpbr3jWm2PwFT3cschfD%20 last accessed 1 June 2013.

http://dx.doi.org/10.1016/j.biologicals.2014.04.001
1045-1056/

batch release testing a focal point in terms of 3R policy (i.e. possibilities to replace, reduce or refine the animal models) [4].

Over the past decades, a broad series of 3R models have been developed for the QC of vaccines, including rabies biologics [2]. However, many obstacles remain in the acceptance of these models for regulatory testing.

This raises the following key questions, which will be addressed in this paper:

- *Which factors influence the acceptance and use of 3R methods for vaccine potency testing?*
- *Which possibilities are available to augment this process?*

To answer these questions the case study of rabies vaccine potency testing is used. Rabies vaccine potency testing is frequently referred to in terms of the challenging process of regulatory acceptance and use of 3R models and has been marked as an international priority in terms of the 3Rs initiatives [5–7].

To grasp complexity of the situation a multidisciplinary view is utilized in which a combination of scientific, ethical, social, economic, organizational, and institutional influences are taken into account. To profoundly analyze the process of regulatory acceptance and use of 3R models to replace the NIH test, a distinction is made between the following three stages a 3R model has to pass:

These 3 sub stages of regulatory acceptance and use are:

I. **FI**: the formal incorporation into regulatory requirements;
II. **ARA**: the actual acceptance by regulatory authorities;
III. **UI**: the use for regulatory purposes by industry.

The regulatory acceptance and use of 3R models is perceived to be very challenging. Each stage poses its own challenges to 3R models developed for QC purposes and acceptance and meeting the needs in one of these stages does not guarantee acceptance in the other two. This article adds to the comprehension of this process by creating an overview of drivers and barriers influencing regulatory acceptance. Furthermore, several suggestions are made with regards to what is needed to overcome the barriers and strengthen the drivers.

Although the findings are specified for rabies vaccine potency testing, similar mechanisms can be observed for the acceptance and use of 3R models for other vaccines. This article should offer valuable insight for both the potency testing of rabies vaccines as for the broader area of classical vaccines.

2. Methodology: case study approach

To analyze the multifaceted problem of regulatory acceptance and use of 3R models for vaccine QC purposes a case study approach is applied. The strength of this approach is its ability to combine different information sources of qualitative and quantitative data [8] thereby creating an inclusive picture of the examined situation.

2.1. Case selection

Potential case studies had to meet the following criteria:

- the existing model for regulatory testing is an animal model which is under discussion;
- there is at least one 3R model available;
- and this 3R modes is somewhere in the process of becoming regulatory accepted and/or commonly used.

The case study of rabies vaccine potency testing was selected as a result of the following combination of characteristics;

- The standard procedure for rabies vaccine potency testing, the NIH mouse potency challenge test, is questioned for reasons of scientific validity, costs and animal welfare consequences.
- The drawbacks of the NIH test have prompted regulatory authorities to actively encourage the evaluation of rabies vaccine potency testing, resulting in an international agreement about the need to work towards reduction, refinement and eventually replacement of the NIH test for both human and veterinary rabies vaccine potency testing [9–12].
- A broad selection of 3R options is already available such as humane endpoints (refinement), reducing the number of animals per dose, the number of doses or the number of dilutions (reduction), serology instead of challenge tests (combination of reduction and refinement) and antigen quantification models (replacement). This paper focuses on serology and antigen quantification models as potential replacements for the NIH test.
- Several initiatives have been taken to develop serological and antigen quantification models (see Section 4.1.1.) of which some have even reached the stage of FI. However none of the models has been able to dethrone the NIH test [13].
- Even though rabies vaccine potency testing and the use of 3R models to replace the NIH test may have several specificities in terms of regulatory acceptance and use, this case study has sufficient common features to be applicable to the acceptance and use of 3R models of classical vaccines in a broader sense.

2.2. Research methods

Case study research relies on multiple sources of evidence to examine the full range of influencing variables. For this rabies vaccine case study a combination of literature research, expert interviews, information derived from relevant scientific meetings and a survey among 50 rabies vaccine experts, was used.

The literature review provided an overview of the regulatory framework, stakeholders involved, *in vivo* and *in vitro* potency testing assays and a first inventory of variables influencing the regulatory acceptance and use of the 3R models for rabies vaccine potency testing. This information served as input for the interview and the participative observation at scientific meetings.

Between 2010 and 2012 six international meetings on 3R models for vaccine testing in general and rabies vaccines in particular, were attended.[3] The official reports of these meeting were examined for factors that potentially drive or withhold regulatory acceptance and use of 3R models for –rabies– vaccine potency testing purposes [9–12].

Furthermore, 15 interviews were conducted with representatives from European and US vaccine regulators, manufactures and academia. The population of interviewees consisted of 10 rabies

[3] **ECVAM/EPAA workshop**: The Consistency Approach for Quality Control of Vaccines – a 3Rs Opportunity (January, 2010, Brussels: Belgium); **Workshop ICCVAM/NICEATM**: 3Rs in Vaccine Potency Testing (September, 2010, Bethesda: USA); **Conference EDQM**: Quality of Medicines in a Globalized World: Dream or Reality? (October, 2010, Prague: Czech Republic); **Workshop Paul-Ehrlich-Institut (PEI)**: Potency testing of veterinary vaccines: the way from *in vivo* to *in vitro* (December 2010, Langen: Germany); **International Workshop ICCVAM/NICEATM** on Alternative Methods for Human and Veterinary Rabies Vaccine Testing: State of the Science and Planning the Way Forward (October 11–13, 201, U.S. Department of Agriculture Center for Veterinary Biologics, Ames, Iowa, USA); **EPAA Workshop** on Waiving of Human Rabies Vaccine Potency Testing: Validation Status and Implementation Strategies of *in vitro* Glycoprotein Quantification Methods (8/9 October 2012, Arcachon: France).

experts (from both the veterinary and the human vaccine field) and five 'other vaccine experts'. By interviewing these 'other vaccine experts' the specificity of the factors in the rabies vaccines case was checked. The combination of literature research, attendance of meetings and interviews resulted in a list of factors that potentially influence the acceptance and use of 3R methods for –rabies– vaccine potency testing and of suggestions to optimize this process.

2.2.1. Survey

The potential drivers and barriers and suggestions to enhance the process were tested through a digital survey. The survey consisted of hypotheses which were formulated based on the findings from the expert interviews and the literature research. A total of five series of hypotheses was tested: (I: general hypotheses; II: hypotheses on Formal Incorporation (FI); III: hypotheses on Actual Regulatory Acceptance (ARA); IV: hypotheses on Use by Industry (UI) and V: hypotheses on speeding up the process).

The survey was spread using the survey tool Survalyzer and was filled out by an average of 50 rabies vaccine experts from the veterinary and human vaccine field. Series I of the hypotheses was filled out by a total of 56 respondents. The response gradually decreased during the survey to a total of 48 respondents at series V (see Table 1).

The respondents derived from the following stakeholder groups:

- RA: Regulatory authorities (including control authorities)
- GI: Governmental Institutes
- In: Industry
- Ac: Academia
- Oth: Other (including retired officials and consultants in the field).

Table 1 shows the details of the survey response per series of hypotheses and per stakeholder group.

Of the respondents 51% are situated in a European country, 31% in the USA, 11% in Latin America, 5% in Japan or China and 2% in Canada. 43% of the respondents operate at a global level, 20% at a European level and 37% at a national level.

The survey findings are presented with a left indent and using a bold and smaller font. The individual survey results also reveal the total number of respondents using N:x and the number of respondents per stakeholder group.

Table 1
Survey response -in total numbers and percentages- per series of hypotheses divided after stakeholder group.

Survey response	I: general hypoth. $N = 56$	II: FI hypoth. $N = 50$	III & IV: ARA & UI hypoth. $N = 50$	V: process optimization hypoth. $N = 48$
Stakeholder group:				
Regulatory authority	18 = 32%	17 = 33%	17 = 34%	17 = 35%
Governmental Institute	10 = 18%	9 = 18%	8 = 16%	8 = 17%
Industry	18 = 32%	16 = 31%	16 = 32%	15 = 31%
Academia	3 = 5%	2 = 4%	2 = 4%	2 = 4%
Other	7 = 13%	7 = 14%	7 = 14%	6 = 13%
Total	56 = 100%	51 = 100%	50 = 100%	48 = 100%

2.2.2. Respondent selection

For the selection of the interview- and the survey respondents purposive sampling was used in which subjects are selected because of some specific characteristics. For this research we were mostly interested in the perception of particular people (or groups of people) connected to the process of rabies vaccine QC regarding the issue of acceptance and use of 3R models for regulatory purposes. To be able to judge the hypotheses tested by the survey, respondents with knowledge and expertise of rabies vaccine QC *and* of 3R models to replace, reduce or refine the NIH challenge test had to be found. For this reason, respondents were selected through the networks of the earlier mentioned international meeting on 3R models and vaccine QC. This might have led to a certain level of bias in the sample, since these meetings are likely to entice participants that are already open to thinking about alternative testing strategies. However, it should be mentioned that the rabies vaccine community is a fairly small community and the major players in this field were represented at these international meetings.

3. Theoretical frame: the multilevel perspective on technology transitions

To organize the case study findings the multilevel perspective on technology transitions is used (for a detailed description of this approach see Schiffelers et al., 2012) [14]. This sociotechnical transition approach, developed in the Netherlands by Schot and Rip offers valuable concepts for the analysis of long-term technological transitions by integrating insights from several disciplines [15,16]. Furthermore, it reveals the interdependencies between the socio-economical and the technical factors that influence technology acceptance. It thereby offers an inclusive tool to analyze the acceptance process and the variables that influence this process.

The model identifies three levels which influence the acceptance of a new technique [14]:

- The micro- or niche level which consists of an 'incubation room' where innovations are developed and put to the test. Here drivers and barriers are found that relate to the development and validation of innovations;
- the meso- or sociotechnical regime level which entails a mix of rules and regulations, expertise, practices and institutions;
- and the macro- or sociotechnical landscape level in which elements like the material infrastructure, existing political culture and coalitions, social values, world views, the macro-economy, demography and the natural environment can be found [17].

The three levels need convergence to create a shift in the existing sociotechnical regime (meso level) leading to the incorporation of an innovation. This convergence can occur if a new technique satisfies the demands at the regime level and/or the broader landscape level, creating a 'window of opportunity' [18,19] for change. Apart from being a useful instrument to analyze complex transition cases the model offers possibilities to determine which of the variables are more promising than others in terms of optimizing the acceptance process in favor of the 3Rs. This potential for change is connected to the level at which a certain factor can be found. Factors at the micro level, for example, are often easier to influence than the factors at the meso- or the macro level.

System innovations are rarely the effect of a single cause but are the *"result of the interplay between many factors and actors"* at these three different levels [14,16]. In those cases where a new technique faces a mismatch with the existing regime and sociotechnical landscape, the innovation is likely to remain stuck at the niche level

[17]. This scenario is frequently observed in the regulatory acceptance and use of 3R models.

4. Factors influencing regulatory acceptance and use of 3R models to replace the NIH test

In this section the drivers and barriers influencing the regulatory acceptance and use of 3R models for rabies vaccine potency testing are presented using the multilevel model on technology transitions. For every level the main drivers and barriers are described that are observed to influence the acceptance and use of 3R models. Where possible an additional distinction is made between the sub stages of Formal Incorporation –FI–, the Actual Regulatory Acceptance –ARA– and the Use by Industry –UI–.

The variables presented below are the ones that were detected in at least two of the three research methods, i.e. literature research, interviews and/or international meetings. The survey findings are used to test the level of agreement with the identified variables.

4.1. Variables at the micro level (niche)

The micro level –niche in which new technologies are developed and validated–, discloses the following influences on regulatory acceptance and use of 3R models:

- the advantages and disadvantages of the existing 3R models (4.1.1.);
- scattered 3Rs expertise and experience and the need for data sharing (4.1.2.) and;
- the difficult validation process (4.1.3.).

4.1.1. Existing serological and antigen quantification models and their advantages/disadvantages

In the past decades several serological and antigen quantification models have been developed as potential alternatives for the NIH test. Examples are the rapid fluorescent focus inhibition test (RFFIT) [20] for inactivated rabies vaccines and the fluorescent antibody virus neutralization test (FAVN) [21]. Both models were initially designed to determine the rabies virus neutralizing antibody level in the serum of immunized individuals and animals. However, there is a development towards using these models for batch potency testing purposes as is for example described in the article published in Biologicals by Krämer et al., in 2009 entitled "The rapid fluorescent focus inhibition test as a suitable method for batch potency testing of inactivated rabies vaccines." [22]. The study has led to the development and validation of the mouse antibody serum neutralization test (SNT) [23] which is based on the RFFIT. In 2012 this serological assay was incorporated into the European Pharmacopoeia monograph for veterinary inactivated rabies vaccines as a possible method for potency testing purposes [24].[4]

Antigen quantification models (full replacement models) measure the amount of antigen/immunogen content of the vaccine. These tests include the single radial diffusion (SRD) test [25] which is accepted in Austria for batch release testing of inactivated rabies vaccines for human use [5], the antibody binding test (ABT) that became a WHO protocol in 1973 [26] and several ELISA procedures [e.g. Refs. [27,28]] such as an *in vitro* ELISA potency test which is used in Japan since 1996 to release non-adjuvanted veterinary rabies vaccines [29] and an ELISA which is currently used by the French National Control Laboratory ANSM[5] to quantify the viral glycoprotein G in non-adjuvanted human rabies vaccines [12].

The appropriateness of available serological and antigen quantification models is always a major component of the discussions on regulatory acceptance and use. This paragraph discusses the main advantages and disadvantages of these 3R models.

The *animal welfare advantages* of serology and antigen quantification are undisputed. Serological models offer a significant refinement and reduction, using fewer animals when compared to the NIH challenge test and causing less pain and distress by avoiding intracranial injections as well as avoiding development of clinical disease in unprotected and control animals [29]. Antigen quantification methods like the SRD, the ABT and ELISA procedures, are full replacement models, using no animals at all.

70% of the respondents agree that vaccine batch release testing depends too strongly on animal models.

Total N: 56. Composition per stakeholder group: RA:11 of 18 respondents, GI:9 of 10, Ind:14 of 18, Ac:3 of 3, Oth:2 of 7.

89% of the respondents agree that 3R models offer significant possibilities to cut back the animal use for vaccine batch release purposes.

Total N: 56. Composition per stakeholder group: RA:17 of 18 respondents, GI:10 of 10, Ind:15 of 18, Ac:3 of 3, Oth:5 of 7.

An important *scientific advantage* is that the serological alternatives (FAVN, RFFIT and the SNT) and antigen quantification tests such as the SRD and ELISAs show strong inter-laboratory reproducibility and produce less variable results when compared to the NIH challenge assay [6]. On top of that, ELISAs are robust, quantitative and precise [6]. The advantages of antigen quantification assays like the SRD and ELISA procedures have led to the recommendation of prioritizing them for product specific validation of human and non-adjuvanted veterinary rabies vaccines. Manufactures are encouraged to consider moving directly to an antigen quantification test for batch release potency testing where feasible [12]. However, it must be noted that ELISA's are predominantly product specific assays which means that they are not standardized across products and that reagents are not universally available [5]. In addition, antigen quantification tests are not recommended for adjuvanted products, since the adjuvant might interfere with the antigen quantitation ELISA, while antigen dissociation from the adjuvant will compromise information on antigen – adjuvant interaction. For the potency testing of adjuvanted veterinary rabies vaccines the serological assay SNT was developed by the PEI. Krämer et al. [22] demonstrate that this assay correlates with the results of the NIH challenge test.

There are however also several *scientific uncertainties* seen to withhold the regulatory acceptance and use of serological and antigen quantification models. A critique that antigen quantification models is often get that they fail to provide evidence for their biological functionality, i.e. their ability to mimic the full biological system and its antigen/immunogenicity response [22]. Serology like the SNT is as a stand-alone method under discussion within industry due to its inability to detect batches with low amounts of glycoprotein (personal communication with representatives from industry).

48% agree that serological assays are no stand-alone method to replace the NIH test due to the fact that they do not allow quantifying the amount of antigen.

[4] http://www.edqm.eu/medias/fichiers/edqm_annual_report_2012.pdf.

[5] Agence national de sécurité du medicament et des produits de santé.

Total N: 50. Composition per stakeholder group: RA:11 of 17 respondents, GI:4 of 8, Ind:8 of 16, Ac:1 of 3, Oth:0 of 7.

At this time both serological and antigen quantification models have unresolved data gaps [11] which lead to hesitations concerning acceptance of the methods for regulatory purposes.

69% agree that technical limitation and scientific questions withhold the FI of 3R models.

Total N: 51. Composition per stakeholder group: RA:13 of 18 respondents, GI:5 of 9, Ind:13 of 18, Ac:2 of 3, Oth:2 of 7.

62% are of the opinion that technical limitation and scientific questions withhold the ARA of 3R models.

Total N: 50. Composition per stakeholder group: RA:14 of 17 respondents, GI:3 of 8, Ind:10 of 16, Ac:2 of 3, Oth:2 of 7.

To intercept these limitations current scientific thought suggests the combination of the serological and antigen quantification models as a panel of assays used in the evaluation of vaccines.

58% are of the opinion that the combination of serology and an antigen quantification offers the most realistic potential to replace the NIH test.

Total N: 50. Composition per stakeholder group: RA: 8 of 17 respondents, GI: 5of 8, Ind:12 of 16, Ac:2 of 3, Oth:2 of 7.

Most serological and antigen quantification assays show *economic advantages* in the sense that they reduce the costs of the procedure and time to release a vaccine batch [6,23]. This especially counts for the inexpensive ELISA procedures and the SRD [5]. Moreover, the RFFIT, the SNT and ELISA procedures offer manufacturers a higher speed of release of a vaccine batch [5,22]. The economic potential of 3R models is highly agreed upon by the respondents of the survey:

82% agree that using 3R models for rabies vaccine batch release has an economic advantage in shortening the release time of a vaccine.

Total N: 50. Composition per stakeholder group: RA:15 of 17 respondents, GI:7 of 8, Ind:15 of 16, Ac:1 of 2, Oth:3 of 7.

Nonetheless, opposing economical argumentations suggests that costs and time needed to switch to a 3R model also need to be taken into account.

4.1.2. Scattered 3R expertise and experience and the need for data sharing

The *expertise and experience* with 3R models to replace the NIH potency test is *scattered* due to the diversity of available models and their sources of origin. As a consequence there is no common understanding of how to use and interpret a particular 3R model. Many 3R models like ELISAs are custom made for a particular production process, meaning that only the manufacturer of that particular product is acquainted with the specifics of that particular model. As a result regulators run the risk of having a setback in 3R knowledge in comparison to the industry, since vaccine manufacturers have often build up experience with specific 3R models in the development and production phase of a vaccine.

48% agree that regulatory authorities have limited experience with 3R models that might replace the NIH test.

Total N: 50. Composition per stakeholder group: RA:7 of 17 respondents, GI:3 of 8, Ind:11 of 16, Ac:2 of 2,Oth:1 of 7.

Such a lack of experience is detrimental to the acceptance and use of 3R models to replace the NIH test [14].

56% of the respondents underline that limited experience withholds the regulatory acceptance of 3R models.

Total N: 50. Composition per stakeholder group: RA:11 of 17 respondents, GI:6 of 8, Ind:7 of 16, Ac:2 of 2, Oth:2 of 7.

However, in the case of rabies vaccines this picture needs to be nuanced. Several regulatory authorities have been leading in the development and validation of 3R models in an effort to replace the NIH potency test [23,22,30,31], thereby expanding their experience and expertise. Nonetheless the knowledge often remains concentrated to that particular stakeholder. To tackle the problem of scattered experience and to build trust in an alternative assay, sharing 3R test data is essential.

84% agree that data sharing is important to gain trust in 3R models for regulatory use.

Total N: 50. Composition per stakeholder group: RA:16 of 17 respondents, GI:8 of 8, Ind:11 of 16, Ac:2 of 2, Oth:5 of 7.

Yet, vaccine manufacturers are said to be reluctant to share their 3R test data with regulatory authorities as long they are uncertain about their return on investment in terms of regulatory acceptance of the model. And regulatory authorities are reluctant to assure acceptance of a 3R model in the absence of data. This situation leads to a *"catch 22, with both stakeholder groups waiting for the other to make the first move"* as it was articulated by a regulator from the US Department of Agriculture (USDA). A suggested way out of this deadlock is for industry to share *in vivo* and *in vitro* test data of batch release testing with regulators when given the guarantee that these will be solely used for the purpose of discussing the value of a 3R model for regulatory acceptance. This could be realized through using a safe harbor approach in which batch release test data are brought in by industry to be evaluated for future regulatory acceptance, without the results of these tests being used for current regulatory decision making [32].

94% agree that manufacturers should share their proprietary *in vivo* and *in vitro* data with regulators on the condition that these data will be treated confidentially.

Total N: 48. Composition per stakeholder group: RA:17 of 17 respondents, GI:8 of 8, Ind:13 of 15, Ac:2 of 2, Oth:5 of 6.

74% feel that industries are willing to share their proprietary *in vivo* and *in vitro* test data with regulators, when given the guarantee that these data will be treated confidentially.

Total N: 50. Composition per stakeholder group: RA:14 of 17 respondents, GI:4 of 8, Ind:14 of 16, Ac:2 of 2, Oth:3 of 7.

4.1.3. Difficult validation process

Validation is a prerequisite for regulatory acceptance and use of 3R models to replace the NIH potency test. At the same time, it is a major challenge to get 3R models validated due to the fact that regulatory authorities often ask for correlation of the alternative model against the NIH test. However, comparing a 3R model with this biological assay is very difficult since it proves to be a poor reference [12] with highly variable test results (see also Section 4.2.1.)

61% agree that the need for correlation of a 3R model with the NIH test is a barrier for the FI of 3R models.

Total N: 51. Composition per stakeholder group: RA:9 of 17 respondents, GI:6 of 9, Ind:10 of 16, Ac:2 of 2, Oth:4 of 7.

After validation and formal incorporation of the model into regulatory requirements (FI), the step of product specific validation must be performed. The costs and time connected to this specific validation is an additional hurdle, especially for veterinary vaccines which have smaller profit margins.

52% agree that the need for product specific validation of 3R models withholds the use of 3R models.

Total N: 50. Composition per stakeholder group: RA:12 of 17 respondents, GI:6 of 8, Ind:6 of 16, Ac:1 of 2, Oth:1 of 7.

To deal with the validation hurdles, several suggestions were made to facilitate the validation process. First of all stakeholders express the wish to see guidelines developed on criteria for the validation of *in vitro* methods.[6] And instead of a full correlation with the NIH test, the suggestion is made to use a concordance strategy in which regulatory approval and implementation of an alternative method can be obtained after a pass/fail correlation using subpotent batches [12].

On top of that, bio-statistical methods are important in facilitating the analysis of data generated during the validation process [23]. The involvement of bio-statistical expertise proved to be very helpfull in the SNT validation process.

69% agree that early involvement of a statistician in the validation process is a driver for FI of a 3R model.

Total N: 51. Composition per stakeholder group: RA:14 of 17 respondents, GI:6 of 9, Ind:9 of 16, Ac:2 of 2, Oth:4 of 7.

4.2. Variables at the meso level (sociotechnical regime)

The meso level consists of the sociotechnical regime and is at the heart of the transition system. It includes a semi-coherent set of rules, test paradigms and the existing test infrastructure. The term 'regime' refers to the engrained collective memory of the standard practices that stakeholders are familiar with [14,17]. The following variables are found to drive or withhold 3R models from becoming an integral part of the regulatory regime:

- the disadvantages of the NIH test (4.2.1.);
- the lack of harmonization of test requirements and 3R acceptance criteria (4.2.2.) and;
- legislation encouraging the use of 3R models (4.2.3.).

4.2.1. The NIH test: a seriously flawed gold standard

The NIH potency test, which was originally developed at the United States National Institutes of Health in 1953, was at first only part of the US minimum requirements for rabies vaccines [33]. The test was later adopted by the WHO Expert Committee on Rabies and is now part of most national and international requirements for inactivated rabies vaccines. It is thereby entrenched in the rabies vaccine Quality Control (QC) infrastructure (both physically and mentally) all over the world and is part of a long standing shared wisdom [6]. Being useable for both veterinary adjuvanted and human non-adjuvanted rabies vaccines it offers a frame of reference to all partners involved in the process of development, production and QC of rabies vaccines. This is a clear advantage in a complex global context involving numerous vaccine producers, and regulatory authorities and a big variety of products. In addition the NIH test has proven useful [33] and has long been recognized as a reliable assay [34]. This combination has provided it with the status of gold standard. A status which is still firmly established.

75% are of the opinion that the NIH test is still considered by regulatory authorities to be the gold standard.

Total N: 51. Composition per stakeholder group: RA:12 of 17 respondents, GI:7 of 9, Ind:13 of 16, Ac:2 of 2, Oth:4 of 7.

At the same time it is widely recognized that the NIH test is seriously flawed and there is broad international agreement to prioritize the NIH mouse potency test in terms of the 3Rs [11,12]. First of all, the actual procedure, in which groups of animals are immunized intraperitoneally (i.p.) with the vaccine and subsequently challenged intracerebrally (i.c.) with a virulent rabies virus, has remained largely unchanged ever since it was introduced almost 60 years ago. As a result of the procedure, around 50% of the animals die or show signs of rabies which involves severe suffering [5]. The route of infection (i.c.), the resulting disease, the numbers of animals used, estimated to be 50.000 to 70.000 mice per year – in the EU and the US– [5] and the repetitive nature of QC testing makes the procedure a serious concern in terms of animal welfare [3,7,22].

73%/74% agree that animal welfare concern regarding the NIH test is the major driver for FI/ARA of 3R models.

FI: Total N: 51. Composition stakeholder group: RA:14 of 17 respondents, GI:9 of 9, Ind:9 of 16, Ac:1 of 2, Oth:4 of 7.

ARA: Total N: 50. Composition stakeholder group: RA:15 of 17, respondents, GI:8 of 8, Ind:8 of 16, Ac:1 of 2, Oth:5 of 7.

52% agree that animal welfare concern drives industry away from the NIH test and towards 3R models for batch release purposes.

Total N: 50. Composition per stakeholder group: RA:7 of 17 respondents, GI:5 of 8, Ind:10 of 16, Ac:1 of 2, Oth:3 of 7.

The NIH test is also seriously flawed in scientific respect. The test has never been properly validated, the test parameters differ considerably from the natural situation (e.g. the i.c. challenge does not reproduce the natural transmission of rabies nor does the i.p. vaccination reflect the normal route of immunization) [6,32,35] and the results often show a high intra- and inter-laboratory variability [5–7,13,22,36,37] with a variation sometimes even up to 400% between different tests [5,22]. A retrospective analysis conducted by the French National Control Laboratory (ANSM) revealed that of all NIH assays conducted in two years' time, only 42% met all validity criteria set by the Ph. Eur. Monograph [13]. Consequently, tests have to be repeated frequently. Therefore, the relevance of this bioassay is highly disputed.

72% agree that the variability of NIH test results drives the ARA of 3R models to replace the NIH test.

Total N: 50. Composition per stakeholder group: RA:13 of 17 respondents, GI:7 of 8, Ind:9 of 16, Ac:2 of 2, Oth:5 of 7.

[6] http://www.edqm.eu/en/Report_3R_meeting-1550.html; last accessed 12 December 2012.

78% agree that the variability of the NIH test drives UI towards the use of 3R models for batch release purposes.

Total N: 50. Composition per stakeholder group: RA:13 of 17 respondents, GI:8 of 8, Ind:12 of 16, Ac:2 of 2, Oth:4 of 7.

Additionally, lab personnel run the risk of being exposed to and becoming infected by the virulent rabies virus [5,22]. However, this risk is not perceived to be a dominant driver towards the use of 3R models for batch release purposes.

22% agree that the safety concern of working with virulent rabies virus in the laboratory is a driver for regulatory authorities to move towards the use of 3R models.

Total N: 50. Composition per stakeholder group: RA:6 of 17 respondents, GI:2 of 8, Ind:2 of 16, Ac:1 of 2, Oth:0 of 7.

32% agree that the safety concern of working with virulent rabies virus is a driver for industry to move towards the use of 3R models.

Total N: 50. Composition per stakeholder group: RA:8 of 17 respondents, GI:3 of 8, Ind:5 of 16, Ac:0 of 2, Oth:0 of 7.

Finally, bioassays like the NIH test are often expensive and time consuming [5,13,28,33]. They take about 4 weeks [22] and they require skilled staff and sophisticated laboratory animal facilities [6,23,22,38].

All these drawbacks have prompted European and North American regulatory authorities to reconsider the use of the NIH test.

4.2.2. Lack of harmonized regulatory requirements and 3R acceptance criteria

The QC of rabies vaccines is regulated by a broad variety of guidelines and monographs such as the European Pharmacopoeia (Ph. Eur.) monograph 0451 for inactivated veterinary vaccines and Ph. Eur. monograph 0216 for rabies vaccine for human use prepared in cell cultures (Ph. Eur) [24,39], the US Code of Federal Regulations for rabies vaccines for human use [40], the USDA Code of Federal Regulations for veterinary rabies vaccines [41], the WHO technical report series vol. 94 for inactivated rabies vaccines for human use [42] and the OIE manual of standards for diagnostic tests and vaccines for veterinary rabies vaccines [43]. These reference standards differ slightly, but they all entail variations of the NIH test to ensure the potency of rabies vaccines [5]. The reference standards are translated into national legislation which may lead to additional variations.

This patchwork of regulatory requirements and the fact that most vaccine manufacturers operate on a global scale makes "……regulatory acceptance one of the most challenging aspects of alternatives implementation for manufacturers…." [44]. For products such as rabies vaccine, which have many different marketing authorizations in different countries all over the world, a variation in the test protocol is problematic since it has to be accepted by every single regulatory authority of the countries where the vaccine is registered. And to get a vaccine license altered, manufacturers often have to run *in vivo* and *in vitro* tests in parallel to fulfill the requirements of these regulatory authorities [44].

Clearly the acceptance of a 3R model by one or just a few regulatory authorities will not incentivize the industry to adopt a 3R model.

64% agree that the lack of harmonization is the main reason for industry to stick to the NIH test.

Total N: 50. Composition per stakeholder group: RA:13 of 17 respondents, GI:6 of 8, Ind:10 of 16, Ac:1 of 2, Oth:2 of 7.

Furthermore there is a lack of mutually agreed criteria for regulatory acceptance of 3R models .

60% agree that there are no shared definitions between regulators on regulatory acceptance criteria.

Total N: 50. Composition per stakeholder group: RA:9 of 17 respondents, GI:7 of 8, Ind:10 of 16, Ac:1 of 2, Oth:3 of 7.

The harmonization of acceptance criteria and the mutual recognition of test results between regulatory authorities are critical steps in obtaining broader acceptance and use of reduction and refinement approaches [7]. And for this purpose harmonization should be kept high on the political agendas, both on the international and national level. Additionally national investment is needed to prioritize the issue of harmonization of test requirements.

75% share the idea that investment at the national level is needed to harmonize regulatory requirements.

Total N: 48. Composition per stakeholder group: RA:13 of 17 respondents, GI:8 of 8, Ind:10 of 15, Ac:2 of 2, Oth:3 of 6.

4.2.3. Regulations encouraging the use of the 3Rs

Several pieces of horizontal and vertical regulation[7] actively encourage the use of existing reduction and replacement models. In the European context, the horizontal Directive 2010/63/EU on the protection of animals used for scientific purposes states in article 13 that:

> *"Member States shall ensure that a procedure is not carried out if another method or testing strategy for obtaining the result sought, not entailing the use of a live animal, is recognized under the legislation of the Union."* [45].

Being horizontal legislation, this Directive applies to all product sectors involved in animal testing [14] and consequently also to European vaccine regulation [46]. It thereby is a potential driver for the regulatory acceptance and use of 3R models.

58% agree that the existence of horizontal legislation like Directive 2010/63 drives the regulatory acceptance of 3R models to replace the NIH test.

Total N: 50. Composition per stakeholder group: RA:11 of 17 respondents, GI:7 of 8, Ind:6 of 16, Ac:2 of 2, Oth:3 of 7.[8]

In the view of the Directive 2010/63/EU, which entered into force on the 1st of January 2013, the European Pharmacopoeia Commission started to evaluate the texts of the Pharmacopoeia that recommend alternatives to animal tests, in order to make this information available to the users and thereby encouraging the use of 3R models.[9] In addition the European Pharmacopoeia Commission makes the following

[7] Horizontal regulation pertains to animal experimentation and multilateral agreements in general. Vertical or sectorial regulation regulates the activities of a particular sector see: http://www.altex.ch/All-issues/Issue.50.html?id=95&aid=3.

[8] Here it is noteworthy that a majority from RA (65%) agrees to the driving force of this horizontal legislation whereas only a minority (38%) from industry agrees.

[9] http://www.edqm.eu/en/European-Pharmacopoeia-news-43.html; last accessed 1 June 2013.

statement in the general notices with the purpose to encourage the implementation of the 3Rs (monograph number 10000) [2].[10]

> *"….This does not imply that performance of all the tests in a monograph is necessarily a prerequisite for a manufacturer in assessing compliance with the Pharmacopoeia before release of a product. The manufacturer may obtain assurance that a product is of Pharmacopoeia quality from data derived, for example, from validation studies of the manufacturing process and from in-process controls."*

> *"….With the agreement of the competent authority, alternative methods of analysis may be used for control purposes, provided that the methods used enable an unequivocal decision to be made as to whether compliance with the standards of the monographs would be achieved if the official methods were used. In the event of doubt or dispute, the methods of analysis of the Pharmacopoeia are alone authoritative."*

The general monographs on Vaccines for veterinary use (0062) & Vaccines for human use (0153): in addition state

> *"….In accordance with the General Notices, alternative test methods may be used to demonstrate compliance with the monograph and the use of such tests is particularly encouraged when this leads to replacement or reduction of animal use or reduction of suffering."*

Other parts of the world such as Australia, Canada and the USA also have pieces of horizontal legislation controlling animal use and thereby supporting the 3Rs, but the US Animal Welfare Act excludes mice, rats and birds. The US however does offer possibilities for manufacturers to submit a new test procedure based on an alternative model. This can be implemented if the test has proven to correlate directly with efficacy, which often is a complicated step to take as explained in Section 4.1.3. Worldwide horizontal legislation to stimulate the use of 3R models does not yet exist but it is seen as a potential tool to stimulate the acceptance and use of 3R models on a global level.

67% agree that international legislation is needed to stimulate the regulatory acceptance and use of 3Rs.

Total N: 48. Composition per stakeholder group: RA:12 of 17 respondents, GI:8 of 8, Ind:7 of 15, Ac:2 of 2, Oth:3 of 6.

Even though progress is being made in Europe in terms of regulation stimulating the use of 3R, the problem is that both the horizontal and vertical regulation leave room for interpretation [46]. This discretionary space leaves it to the responsibility of regulatory authorities and manufactures to choose the method they see as most suitable. The consequence is that manufacturers are uncertain whether an alternative model will eventually be accepted by the regulatory authorities for batch release purposes. In the absence of such a guarantee, combined with the existing lack of harmonization, manufacturers are likely to stick to the existing practices i.e. the NIH test. Manufacturers therefore indicate that they would prefer an explicit description of a set of 3R models which will be accepted if implemented properly. A similar signal was given at the ad hoc meeting on the 3Rs in June 2012 at the premises of the EDQM[11] in Strasbourg. In the context of the introduction of the EU Directive 2010/63/EU, the European Pharmacopoeia Commission decided at the 141st Session (November 2011) that its Groups of Experts should review all monographs and chapters prescribing animal tests since[12]:

> *"A clear need was identified for the wording of Ph. Eur. texts to be explicit when prescribing the use of non-animal tests to replace animal tests or when exemptions are given to revert back to animal testing…."*

4.2.4. The potential of the consistency approach

An additional development, that is important mentioning here, is the consistency approach. The consistency approach is not so much a driver for the regulatory acceptance and use of 3R models but is relevant in terms of moving away from the NIH test. The approach is based upon the principle that the quality of a vaccine is the result of the strict application of a quality system and consistent production [1,9,38]. With it, a movement is made from batch release testing to in process control. It implies that the production is key to the quality of a vaccine. The approach offers the opportunity to replace animal bioassays like the NIH test on the final lot with a battery of meaningful tests with enhanced capacity to compare new batches with batches of proven quality [38]. The approach requires a combination of immunochemical and physicochemical tests, performed in process and on the final product. This combination of tests, together with adhering to the guidelines of Good Manufacturing Practice (GMP) [9], shall ensure that all produced batches are of the same quality as the batches proven to be safe and efficacious during licensing [1]. For conventional products such as rabies vaccines this approach can lead to a substantial reduction in animal usage for potency testing purposes in the final batch [38]. Although the survey did not include a specific hypothesis on this approach the survey respondents made numerous remarks that referred to the potential of the consistency approach to move away from the NIH test.

4.3. Variables at the macro level (sociotechnical landscape)

The macro level consists of the broad societal developments influencing the transition from the NIH potency test to 3R alternatives. This level usually is the hardest level to change, since it involves developments and public values that are deeply engrained into society and often compete with each other. At this level the following competing variables are identified:

- the increasing animal welfare apprehension (4.3.1.); and
- the striving for risk minimization (4.3.2.).

4.3.1. Animal welfare apprehension

Animal welfare concern is seen as an important driver to move away from the conventional way of rabies vaccine potency testing and towards the use of 3R models (see also Section 4.2.1). In the area of vaccine potency testing, rabies vaccines are one of the top priorities in terms of the 3Rs to reduce animal use and suffering. However the level of concern for animal welfare is very much rooted in cultural values. In Western society, where a high level of human health and welfare is fairly common, the public and political attention for animal welfare is relatively high as is reflected in Directive 2010/63/EU. In countries where the human health and welfare are often still at stake, animal welfare is not a first priority. This means that they are predominantly incentivized by economic and scientific arguments.

[10] http://online6.edqm.eu/ep708/NetisUtils/srvrutil_getdoc.aspx?0L3OpDZ4uDrmnC30mC4KkQ7Hj/10000E.pdf?res=true; last accessed 17 December 2013.

[11] European Directorate for the Quality of Medicines & Health Care of the Council of Europe.

[12] http://www.edqm.eu/en/Report_3R_meeting-1550.html; last accessed 1 June 2013.

Table 2a
Main barriers and drivers for 3R models to replace the NIH test at the stage of **Formal incorporation into the regulatory requirements –FI–.**

Main barriers	% of survey agreement	Main drivers	% of survey agreement
Fear for release sub potent rabies vaccine	75%	Taking regulatory needs onboard when developing 3R models	81%
NIH test considered to be gold standard	74%	Animal welfare concern	73%
Technical limitations and remaining scientific questions of 3R models	69%	Early involvement statistician	69%
Regulators require correlation 3R model with NIH test	61%	Variability NIH test	57%

4.3.2. Risk averse society

Though animal welfare is a public value that has the potential to drive the use of 3Rs for rabies vaccines potency testing, the next public value is observed to outweigh this variable.

Modern society is characterized by a continuous striving for risk minimization. This is translated in a dense patchwork of regulations to minimize potential adverse effects of human activities and products. This is certainly applicable for vaccines. This group of products is looked at with great prudence due to their biological and thereby variable nature and the fact that they are often administered to vulnerable groups of people like children. Rabies vaccines are especially looked at with great caution, although from a somewhat different angle. Rabies is a deadly zoonosis which kills over 70,000[13] people worldwide each year, with most casualties in developing countries. Rabies vaccines are a vital tool in the prevention of rabies infections and the treatment of exposed individuals. Each year, more than 15 million people are estimated to receive post-exposure rabies prophylaxis treatment after being exposed to rabies [12]. The serious progression and the lethality of the illness make the availability of a high quality vaccine paramount.

Even though true rabies vaccine failures are believed to be rare,[14] the automated reaction to deal with the combination of uncertainty and responsibility, is sticking to the procedure regulators are familiar with [14] which is the NIH potency test.

75% agree that the FI of 3R models is withheld by the regulators fear for releasing subpotent rabies vaccines.

Total N: 51. Composition per stakeholder group: RA:14 of 17 respondents, GI:5 of 8, Ind:13 of 16, Ac:2 of 2, Oth:4 of 7.

68% believe that the fear of subpotent rabies vaccines withholds the ARA of 3R models even after their FI.

Total N: 50. Composition per stakeholder group: RA:11 of 17 respondents, GI:7 of 8, Ind:11 of 16, Ac:2 of 2, Oth:3 of 7.

Manufacturers on their turn are also risk averse. Their risk aversion is mainly caused by the fear that a change in potency testing in favor of the 3Rs may in the end not be accepted by one or the other regulatory authority. Their way of anticipating this possible non-acceptance resembles the above described reaction of the regulatory authorities on uncertainty, namely sticking to the old routines.

56% feel that industry's regulatory affairs departments anticipate the most conservative requirements.

Total N: 50. Composition per stakeholder group: RA:9 of 17 respondents, GI:6 of 8, Ind:8 of 16, Ac:2 of 2, Oth:3 of 7.

Such conservatism is partly functional since it reflects societies' demand for risk minimization, but it becomes dysfunctional when it retains scientifically valuable developments from entering into the regulatory regime.

5. Discussion and conclusion

Regulatory acceptance and use of 3R models for potency testing purposes of rabies vaccines is high on the international agenda driven by ethical, scientific and economical motives. Progress has been made in the sense that several 3R models have been incorporated into monographs or guidelines (see Section 4.1.1). Nonetheless, the NIH test still is the dominant procedure for rabies vaccine potency testing purposes and 70% of the respondents of the survey agree that vaccine batch release depends too strongly on animal models (see Section 4.1.1.).

Every sub stage of acceptance and use of a 3R model (i.e. FI, ARA and UI) can delay or withhold the regulatory acceptance of a 3R model to replace the NIH test. And acceptance in one stage is no guarantee for acceptance further on in the process.

64% agree that FI into regulatory requirements does not automatically lead to ARA.

Total N: 50. Composition per stakeholder group: RA:12 of 17 respondents, GI:4 of 8, Ind:12 of 16, Ac:1 of 2, Oth:3 of 7.

70% agree that FI of a 3R model to replace the NIH test does not automatically lead to UI.

Total N: 50. Composition per stakeholder group: RA:15 of 17 respondents, GI:6 of 8, Ind:11 of 16, Ac:0 of 2, Oth:3 of 7.

Therefore, the survey made a distinction between drivers and barriers at these three sub stages of regulatory acceptance and use of 3R models for rabies vaccine potency testing purposes. The barriers and drivers that obtained the highest levels of agreement during the survey are summarized per sub stage in Table 2a, b and c. This summary is followed by a short analysis of the findings. Apart from analyzing the differences per sub stage, the level of agreement per stakeholder group is evaluated. Here we have focused on the stakeholder groups of regulatory authorities and industry as two of the main players in the field and the two biggest groups of respondents during the survey. Only the cases with considerable differences in levels of agreement (of 25% or more) have been discussed.[15]

The fear of releasing sub potent vaccines, the technical limitations of 3R models and the NIH test as gold standard are perceived to be the dominant barriers for both the FI and the ARA of 3R

[13] Due to the fact that in many cases the symptoms are not recognized as being rabies the informal numbers are estimated much –at least 10 times– higher (communication with Nico Visser former employee Intervet, July 2010).

[14] http://iccvam.niehs.nih.gov/meetings/Rabies/Vacc/Wksp-2011/present/Rupprecht-508.pdf; last accessed 5 February, 2013.

[15] We are aware of the somewhat deviant way of responding of the "Other" group. However from the results of the survey it is impossible to give a definite clarification for this difference since no underlying motives were given in addition to the level of agreement with the hypotheses.

Table 2b
Main barriers and drivers for 3R models to replace the NIH test at the stage of **Actual Regulatory Acceptance** –**ARA**–.

Main barriers	% of survey agreement	Main drivers	% of survey agreement
Fear for release sub potent rabies vaccine	68%	Data sharing of test results between industry and regulators	84%
Technical limitations and remaining scientific questions of 3R models	62%	Taking regulatory needs into account when developing 3R model	82%
NIH test considered gold standard & regulators require correlation 3R model with NIH test	58%	Variability NIH test	78%
Limited experience with a 3R model	58%	Animal welfare concern	74%

Table 2c
Main barriers and drivers for 3R models to replace the NIH test at stage of UI.

Main barriers	% of survey agreement	Main drivers	% of survey agreement
Lack of harmonization	64%	Economic advantage by shortening time of release	82%
No shared definitions between regulators on acceptance criteria	62%	Variability NIH test	78%
Unclear acceptance criteria	58%	Combination of 3R methods is promising	58%
Product specific validation needed	52%	Animal welfare concern	52%

models. In addition the limited experience with 3R models is observed to be a barrier for the ARA. The UI is perceived to be mainly withheld by the lack of harmonized regulatory requirements, the lack of shared and clear acceptance criteria and the requirement to conduct a product specific validation (although product specific validation is mainly perceived as a barrier by respondents from regulatory authorities (71%) and to a far lesser extent by respondents from industry (38%): see Section 4.1.3).

On the other hand several drivers are identified to stimulate the sub stage of FI, ARA and UI. The concern for animal welfare and the variability of the NIH test are seen as important drivers in all three sub stages. It must be noted that the level of agreement about the driving force of animal welfare differs. 82% and 88% of respondents from regulatory authorities agree that animal welfare drives the stages of FI and ARA respectively, while only 56% and 50%, respectively of respondent of industry is of that same opinion. The timely consideration of regulatory needs is observed to drive the FI and the ARA. The advantages of 3R models in terms of reduction of costs and time of release and the potential of a combined use of 3R models are seen to be important drivers for the UI (although regulatory authorities seem to be more skeptic about the driving potential of combining of methods (47%) in comparison to respondents from industry (75%)). Furthermore data sharing of test results between industry and regulators is seen as an important driver for ARA. This is however mainly the case according to regulatory authorities (94% agree) whereas a smaller percentage of respondents from industry agree (68%). Both stakeholder groups report that industries are willing to share these data (82% of RA and 88% of industry (see Section 4.1.2.)

And finally, the early involvement of a statistician is perceived to be an important driver for the FI of 3R models. Here again a difference can be observed between regulatory authorities (82% agree) and industry (56% agree).

Regulatory acceptance and use of 3R models for –rabies– vaccine batch release purposes proves to be a multifaceted problem. The following figures, using the 3R Acceptance Model [14] provide a synopsis of influencing forces at hand. This model is based on the multilevel approach on technology transitions (see Section 3). It displays the barriers (Fig. 1a) and drivers (Fig. 1b) as described in Section 4.

This overview not only summarizes the variables that where described but also offers the possibility to determine the more pliable from the rigid variables [47]. Generally speaking the more workable or pliable factors can be found at the lower levels (micro and partly meso level) of this model, while the variables at higher levels (partly meso and the macro level) are usually of a more rigid nature. Making this distinction is relevant in terms of choosing a strategy to enhance the acceptance process in favor of 3R models.

6. Optimizing the process of regulatory acceptance and use

In this final section suggestions are made to speed up the process of regulatory acceptance of 3R models for rabies vaccine QC. Although the rabies vaccine case is specific in some respects (e.g. when it comes to its long history of developing 3R models, its high level of agreement about the drawbacks of the existing animal model, the high animal welfare concern due to the severity of the procedure, an even higher risk aversion due to the characteristics of the disease and the presence of both human and veterinarian rabies vaccines) it is also quite comparable to other classical vaccines cases in several respects. The regulatory acceptance and use of 3R models for many other classical vaccines also face the problem of flawed gold standards, the lack of harmonized test requirements and uncertainties connected to developed 3R models -see for example the Canadian case study on diphtheria/tetanus potency testing of Long and Griffin- [48]. Therefore the following suggestions to optimize the process are also useful for a broader group of classical vaccines that show comparable features as the rabies vaccine case.

The different stages of regulatory acceptance and use (FI, ARA and UI) each require their own specific catalyzers. To stimulate the process the following recommendations are made to facilitate promising 3R models in getting out of the niche and become part of the existing regulatory regime:

6.1. Recommendations to enhance the process of formal incorporation

- To guide a 3R model towards FI it is important that the model is developed and validated in close communication/cooperation with the stakeholders involved in regulatory acceptance and use

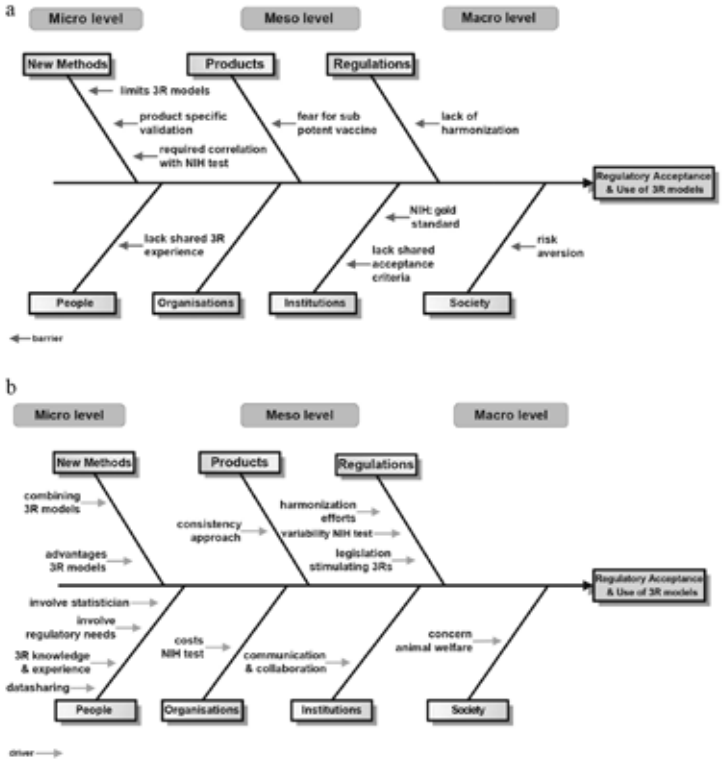


Fig. 1. (a) **Barriers** to acceptance and use of 3R models to replace the NIH test for rabies vaccines potency testing purposes using a multilevel perspective on technology transitions [14]. (b) **Drivers** for acceptance and use of 3R models to replace the NIH test for rabies vaccines potency testing purposes using a multilevel perspective on technology transitions [14].

(i.e. regulatory authorities and manufacturers) to incorporate the criteria an alternative potency assay has to meet to get accepted for batch release purposes [7,23]. The development and validation of the SNT is a good example of this. The test was developed by the German OMCL and validated under the auspices of the Biological Standardization Program (BSP) of the EDQM. This has contributed to speedy incorporation of the validated test into Ph. Eur. monograph 0451 for inactivated rabies vaccines for veterinary use.

88% of the respondents embrace the recommendation of timely involvement of regulatory authorities into the process of development /validation of 3R models.

Total N:48; Composition per stakeholder group: RA:17 of 17 respondents, GI:7 of 8, Ind:11 of 15, Ac:2 of 2, Oth:5 of 7.

➢ Next, **a clear comprehension of the regulatory framework in which the 3R model should get incorporated** is crucial. This can be stimulated by the timely involvement of regulators. However developers/validators of 3R models should also become better informed about the regulatory process and the requirements the 3R model will have to face when trying to enter the regulatory regime.

According to 90 % of the respondents developers and validators of 3R models should be educated in the regulatory process and regulatory needs.

Total N: 48. Composition per stakeholder group: RA:16 of 17 respondents, GI:8 of 8, Ind:12 of 15, Ac:2 of 2, Oth:5 of 6.

➢ Furthermore, the **early involvement of statisticians** is important to design the validation study and to analyze the data generated during the validation of the model. This involvement was one of the success factors in the collaborative study for the SNT (see section 4.1.3.).

- To get a validated 3R model on the agenda for formal incorporation into the regulatory requirements of one or more standard setting organizations, **policy entrepreneurs are of vital importance** [18]. Policy entrepreneurs in this context are defined as individuals or organizations that strive for the initiation of a policy change in favor of the 3Rs. They must have a good reputation, state of the art knowledge, a solid network and financial support to organize international meetings where experience can be shared and data exchanged. Several committed policy entrepreneurs can be identified in the case of rabies vaccine potency testing that have put a lot of energy in bringing the parties together and stimulating the discussion. These policy entrepreneurs are very much needed to facilitate such a difficult process and to keep the issue high on the international agenda.

  Respondents identify both regulatory authorities (63%)/industry (63%) as suitable candidates to take the lead in stimulating the use of 3Rs.

  Total N: 48. Composition question regulatory. authorities per stakeholder group: RA:11 of 17 respondents, GI:6 of 8, Ind:7 of 15, Ac:2 of 2, Oth:4 of 6. / Question industry Total N: 48. Composition per stakeholder group: RA:10 of 17 respondents, GI:4 of 8, Ind:10 of 15, Ac:2 of 2, Oth:4 of 6.

- Subsequently, a **process manager needs to be appointed** to guide a 3R model through the process of validation and into the process of FI. Such a process manager should stimulate the collaboration between international stakeholders and should set out a clear road map of the steps that need to be taken and by whom [32]. The process manager does not necessarily need to come from a regulatory authority but for the FI of a 3R model into the regulatory requirements the firm involvement of regulatory authorities is is pivotal.

6.2. Recommendations to enhance the process of actual regulatory acceptance

- **Regular communication and committed collaboration** between developers, validators, regulators and end users are needed both during and after the collaborative study/validation phase for regulatory authorities to gain collective trust in and experience with the model [32].

  96% of the respondents see frequent communication on regulatory needs between regulatory authorities and industry as an important prerequisite for the acceptance and use of 3R models for regulatory purposes

  Total N: 48. Composition per stakeholder group: RA:17 of 17 respondents, GI:8 of 8, Ind:14 of 15, Ac:2 of 2, Oth:5 of 6.

- Furthermore, ARA of a 3R models needs to be facilitated by **data sharing of both *in vitro* and *in vivo* data.** Regulators need these data to build trust in the 3R model and in the consistency of the production process. This means that regulatory authorities should also be informed about these data in case of subpotent batches. To stimulate data sharing safe harbors should be created in which data can be examined and confidentiality is guaranteed [32].

6.3. Recommendations to enhance the process of use by industry

- **Legal requirements should firmly motivate the replacement of *in vivo* by *in vitro* or consistency testing** [6] and should clarify how to handle in case there is room for interpretation or choice of methods.
- To tackle the correlation problem for product specific validation both manufacturers and regulatory authorities need to **invest in developing a sound concordance strategy** (see Section 4.1.3.)
- **Organizational incentives** such as economic advantages to start using 3R models should be created for example by speeding up the process of marketing authorization when using 3R models to replace the NIH test [23] or by fee waiving for adjustments in the vaccine license in favor of a 3R model that is incorporated into the Ph. Eur. monographs.
- And finally, both **national and international authorities need to invest in harmonization** of test requirements and acceptance criteria for 3R models. This investment requires long term commitment, funding and expertise.

In short: the issue of acceptance and use of 3R models to replace the NIH test for potency testing purposes of rabies vaccines is gaining positive momentum. It is timely for regulators and manufacturers to move away from the NIH test and towards 3R models. To overcome the barriers to replace the NIH test (see Fig. 1a) for rabies vaccine potency testing in particular and of similar assays for classical vaccines in general, a well-designed process is required. This process starts at the micro level where early involvement of all stakeholders is needed during the test development and validation of promising 3R models. And it then goes to the meso level where frequent communication and international collaboration are needed to create a shared understanding of the potential of the 3R model at hand and to formulate clear acceptance criteria. Only an intelligent process design that makes use of a combination of the drivers at the micro, meso level and macro level (see Fig. 1b) will help overcoming the barriers that 3R models in vaccine quality control still face.

Acknowledgments

We thank the Doerenkamp-Zbinden Foundation for funding this project.

We thank Dr. Dieter Lütticken for sharing his expertise with the corresponding author.

We thank Dr. Marlies Halder for proof reading of this research paper.

We thank all the experts that have shared their opinion with us by granting an interview and/or by taking part in the survey.

References

[1] Kulpa-Eddy J, Dusek D. Application of the consistency approach to reduce animal use in vaccine potency testing. Proc Vaccinol 2011;5:232–5.
[2] Hoonakker ME. Vaccines, animal experiments and alternatives: a survey of reduction, refinement and replacement strategies in human and veterinary development, production and quality control. The Netherlands: Netherlands Vaccine Institute and Utrecht University; 2011. Project of the program Dierproeven begrensd II, ZonMW.
[3] Dozier S, Brown J, Currie A. Bridging the gap between validation and implementation of non-animal veterinary vaccine potency testing methods. Animals 2011;1:414–32.
[4] Russell WMS, Burch RL. The principles of humane experimental technique. London: Methuen; 1959 [Reprinted by UFAW, 1992].
[5] Bruckner L, Cussler K, Halder M, Barrat J, Castle P, Duchow K, et al. Three Rs approaches in the quality control of inactivated rabies vaccines. ATLA 2003;31:429–54.
[6] Romberg J, Lang S, Balks E, Kamphuis E, Duchow K, Loos D, et al. Potency testing of veterinary vaccines: the way from *in vivo* to *in vitro*. Biologicals 2012;40:100–6.
[7] Casey W, Schmitt M, McFarland R, Isbrucker R, Levis R, Arciniega J, et al. Improving animal welfare and reducing animal use for human vaccine potency testing: state of the science and future directions. Proc Vaccinol 2011;5:33–46.

[8] Yin RK. Case study research: design and methodsIn Applied social research methods series, vol. 5. Thousand Oaks, Inc: Sage Publications; 2003.

[9] De Mattia F, Chapsal JM, Descamps J, Halder M, Jarrett N, Kross I, et al. The consistency approach for quality control of vaccines – a strategy to improve quality control and implement 3Rs. Biologicals 2011;39:59–65.

[10] Potency testing of veterinary vaccines for animals: the way from *in vivo* to *in vitro*. In: Jungbäck C, editor. Developments in biologicals, vol. 134; 2012.

[11] Stokes WS, Kulpa-Eddy J, McFarland RM. Introduction and summary of the international workshop on alternative methods to reduce, refine, and replace the use of animals in vaccine potency and safety testing: state of the science and future directions. Proc Vaccinol 2011;5:1–15.

[12] Stokes W, McFarland R, Kulpa-Eddy J, Gatewood D, Levis R, Halder M, et al. Report on the international workshop on alternative methods for human and veterinary rabies vaccine testing: state of the science and planning the way forward. Biologicals 2012;40:369–81.

[13] Servat A, Kempff S, Labadie A, Scherrefer JL, Bouè F, Cliquet F. In vivo potency tests of rabies inactivated vaccines for veterinary use: a 2-year retrospective analysis of data according to the criteria of the European Pharmacopoeia. Pharmeuropa 2008;20:4.

[14] **Schiffelers MJWA, Blaauboer BJ, Hendriksen CFM, Bakker WE. Regulatory acceptance and use of 3R models: a multilevel perspective. ALTEX 2012;3: 287–300. Available at:** http://www.altex.ch/en/index.html?id=168&aid=4.

[15] Schot J, Rip A. The past and future of constructive technology assessment. Technol Forecast Soc Change 1996;54:251–68.

[16] Geels FW. Multi-level perspective on system innovation: relevance for industrial transformation. In: Olsthoorn X, Wieczorek AJ, editors. Understanding industrial transformation: views from different disciplines. Dordrecht The Netherlands: Springer; 2006. pp. 163–86.

[17] Kemp R. The Dutch energy transition approach. Int Econ Econ Policy 2010;7: 291–316.

[18] Kingdon JW. Agendas, alternatives and public policies. 2nd ed. New York: Longman Pub. Group; 1995.

[19] Geels FW. Technological transitions as evolutionary configuration processes: a multi-level perspective and a case study. Res Policy 2002;31:1257–74.

[20] Smith JS, Yager PA, Baer GM. A rapid tissue culture test for determining rabies neutralizing antibody. In: Kaplan MM, Koprowski H, editors. Laboratory techniques in rabies. 4th ed. Geneva: WHO; 1996. pp. 354–7.

[21] Cliquet F, Aubert M, Sagne L. Development of a fluorescent antibody virus neutralization test (FAVN test) for the quantitation of rabies-neutralizing antibody. J Immunol 1998;212:79–87.

[22] Krämer B, Schlédger H, Behremdorf-Nicol HA, Hanschmann KM, Duchow K. The rapid fluorescent focus inhibition test is a suitable method for batch potency testing of inactivated rabies vaccines. Biologicals 2009;37:119–26.

[23] Krämer B, Bruckner L, Daas A, Milne C. Collaborative study for validation of a serological potency assay for rabies vaccine (inactivated) for veterinary use. Pharmeur Biol Sci Notes 2010;2:37–55.

[24] European Pharmacopoeia Monograph 04/2013:0451. Rabies vaccine (inactivated) for veterinary use. Ph. Eur. 8th ed. Strasbourg, France: European Department for the Quality of Medicines within the Council of Europe; 2013.

[25] Ferguson M, Seagroatt V, Schild GC. A collaborative study on the use of single radial immunodiffusion for the assay of rabies virus glycoprotein. J Biol Stand 1984;13:283–94.

[26] Arko RJ, Wiktor TJ, Sikes RK. The antibody binding test for vaccine potency. In: Kaplan MM, Koprowski H, editors. Laboratory techniques in rabies. Geneva: WHO; 1973. pp. 292–4.

[27] Perrin P, Morgeaux S, Sureau P. In vitro rabies vaccine potency appraisal by ELISA: advantages of the immunocapture method with a neutralizing anti-glycoprotein monoclonal antibody. Biologicals 1990;18:321–30.

[28] Rooijakkers E, Groen J, Uittenbogaard J, van Herwijnen J, Osterhaus A. Development and evaluation of alternative testing methods for the in vivo NIH potency test used for the quality control of inactivated rabies vaccines. Dev Biol Stand 1996;86:137–45.

[29] Gamoh K, Senda M, Itoh O, Muramatsu M, Hirayama N, Kotke R, et al. Use of ELISA for in vitro potency test of rabies vaccines for animal use. Biologicals 1996;24:95–101.

[30] Ferguson M, Heath A. Report of a collaborative study to assess the determination of glycoprotein antigen content of rabies vaccines for human use. Biologicals 1992;20:143–54.

[31] Fournier-Caruana J, Poirier B, Haond G, Jallet C, Fuchs F, Tordo N, et al. Inactivated rabies vaccine control and release: use of an ELISA method. Biologicals 2003;31:9–16.

[32] Schiffelers MJWA, Blaauboer BJ, Bakker WE, Beken S, Hendriksen CFM, Koëter HBWM, et al. Regulatory acceptance and use of 3R models for pharmaceuticals and chemicals: expert opinions on the state of affairs and the way forward. Regul Toxicol Pharmacol 2014;69:41–8.

[33] Barth R, Diderrich G, Weinmann E. NIH test, a problematic method for testing potency of inactivated rabies vaccine. Vaccine 1988;6:369–77.

[34] Shin J, Lei D, Conrad C, Knezevic I, Wood D. International regulatory requirements for vaccine safety and potency testing: a WHO perspective. Proc Vaccinol 2011;5:164–70.

[35] Wonderli PS, Dreesen DW, Miller TJ, Baer GM. The rabies peripheral challenge test: more accurate determination of vaccine potency. Vaccine 2006;24: 7115–23.

[36] Knezevic I. Stability evaluation of vaccines: WHO approach. Biologicals 2009;37:357–9.

[37] Jivapaisarnpong T, Schofield T, Krause PR. A vaccine measured with a highly variable assay: rabies. Biologicals 2009;37:412–5.

[38] Hendriksen C, Arciniega J, Bruckner L, Chevalier M, Coppens E, Descamps J, et al. The consistency approach for the quality control of vaccines. Biologicals 2008;36:73–7.

[39] European Pharmacopoeia Monograph 04/2008:0216. Rabies vaccine for human use prepared in cell cultures. Ph. Eur. 7th ed. Strasbourg, France: European Department for the Quality of Medicines within the Council of Europe; 2011.

[40] **U.S. Code of Federal Regulations, Title 9, Chapter I: animal and plant health inspection service, Department of Agriculture, Subchapter E – viruses, serums, toxins, and analogous products; organisms and vectors, Part 113.209,rabies vaccine. killed virus. Washington D.C., USA: U.S. Government Printing Office; updated 2010.**

[41] **United States Department of Agriculture. Center for veterinary biologics, code of Federal regulations, Title 9, animals and animal products. Part 113.209-standard requirements. SAM 308: supplemental assay method for potency testing of inactivated rabies vaccine in mice using the National Institutes of Health test. Available at:** http://www.aphis.usda.gov/animal_health/vet_biologics/vb_sams_300_series.shtml; **2007.**

[42] WHO. Annex 2: recommendations for inactivated rabies vaccine for human use produced in cell substrates and embryonated eggs. In: WHO expert committee on biological standardization, fifty-sixth report, technical report series, vol. 941. Geneva: World Health Organization; 2007. pp. 83–132.

[43] **OIE. Rabies. In: Manual of standards for diagnostic tests and vaccines for terrestrial animals. 6th ed. World Organization for Animal Health; 2012. pp. 276–91 Available at:** http://www.oie.int/eng/normes/mmanual/A_summry.htm.

[44] Descamps J, Giffroy D, Remy E, Mortiaux F, Mareschal JC, Ponsar C, et al. A case study of development, validation, and acceptance of a non-animal method for assessing human vaccine potency. Proc Vaccinol 2011;5:184–91.

[45] European Commission. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. Off J Eur Union L 2010;276:33–79.

[46] Jennings M, Morton DB, Charton E, Cooper J, Hendriksen C, Martin S, et al. Application of the three Rs to challenge assays used in vaccine testing: tenth report of the BVAAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement. Biologicals 2010;38:684–95.

[47] Ellemers JE. Vreê kunnen verklaren of iets kunnen veranderen: krachtige versus manipuleerbare variabelen [Explaining a lot or changing a little: rigid versus pliable variables: not available in English]. Beleid en Maatschappij 1979;3:179–82.

[48] Long ME, Griffin G. Challenges and opportunities for the implementation of the 3Rs in Canadian vaccine quality control. Regul Toxicol Pharmacol 2012;63:418–25.

III

OVERVIEW OF VALIDATED AND ACCEPTED 3R MODELS PER ENDPOINT

Overview of Validated and Accepted 3R Models per Endpoint

Total Number of Validated and Accepted Alternative Methods[1]	
Acute mammalian toxicity	N=8
Biologicals and vaccines	N=10[2]
Carcinogenicity	N=2
Chronic Toxicity	N=1
Dermal absorption/penetration	N=1
Ecotoxicity	N=3
Endocrine active substances	N=7
Eye corrosion and/or irritation	N=12
Genotoxicity	N=9
Phototoxicity	N=3
Pre-and nonclinical safety studies drug dev.	N=2
Pyrogenicity	N=6
Reprod. & developmental toxicity	N=4
Skin corrosion and/or irritation	N=12
Skin sensitization	N=9
Total number of methods	N=89

1 http://alttox.org/validation-and-acceptance-status-of-alternatives-2/ :consulted May 2015 This table offers a good indication of available methods but it is not complete.

2 The number of 3R models approved by the Ph. Eur. in the domain of biologicals and vaccines however is much higher, as is described in the report of Hoonakker et al. 2011. The tables at the pages 87-108 reveal that the Ph. Eur. has already approved 60 3R models for potency testing purposes of vaccines. These are mostly serological methods (reduction models) but also antigen determination methods (replacement models) and the possibility to use single dilutions and clinical endpoints (refinement options).

IV

OVERVIEW OF REGULATORY AUTHORITIES IN THE DIFFERENT PRODUCT SECTORS

Overview of Regulatory Authorities in the Different Product Sectors

Regulatory level/ sector	Pharmaceuticals	Biologicals	Chemicals
Global	ICH/ VICH[1]	ICH/ VICH	OECD[2]
European	European Medicines Agency (EMA) (licensing) EDQM[3]; Ph Eur.[4] (requirements for production and Quality Control)	European Medicines Agency (EMA) (licensing) EDQM; Ph. Eur. (requirements for production and Quality Control)	European Chemicals Agency (ECHA) (Registration, evaluation, and authorization of chemicals) Scientific Comm. on Health and Environ. Risk (SCHER)
National	e.g. US Food and Drug Agency (FDA) Health Canada College ter Beoordeling Geneesmiddelen/ Medicines Evaluation Board CBG/MEB (NL)	e.g. FDA/US Department of Agriculture (USDA) (human/vet. use) Health Canada National Institute for Public Health and the Environment, RIVM (NL)	e.g. Environmental Protection Agency (EPA) (US) Environment Canada National Institute for Public Health and the Environment, RIVM (NL)

1 The International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use& the International Cooperation on Harmonization of Technical Requirements for Registration of Veterinary Products

2 Organisation for Economic Co-operation and Development

3 The European Directorate for the Quality of Medicines & HealthCare

4 European Pharmacopoeia

V

THE INDIAN TALE OF THE BLIND MAN AND THE ELEPHANT

The Indian Tale of the Blind Men and the Elephant

The legend of the blind men and the elephant
John Godfrey Saxe's (1816-1887) version of the famous Indian legend[1]

It was six men of Indostan
To learning much inclined,
Who went to see the Elephant
(Though all of them were blind),
That each by observation
Might satisfy his mind.

1. The *First* approach'd the Elephant,
And happening to fall
Against his broad and sturdy side,
At once began to bawl:
"God bless me! but the Elephant
Is very like a wall!"

2. The *Second*, feeling of the tusk,
Cried, -"Ho! what have we here
So very round and smooth and sharp?
To me 'tis mighty clear
This wonder of an Elephant
Is very like a spear!"

3. The *Third* approached the animal,
And happening to take
The squirming trunk within his hands,
Thus boldly up and spake:
"I see," quoth he, "the Elephant
Is very like a snake!"

4. The *Fourth* reached out his eager hand,
And felt about the knee.
"What most this wondrous beast is like Is
mighty plain," quoth he,
"'Tis clear enough the Elephant
Is very like a tree!"

5. The *Fifth*, who chanced to touch the ear,
Said: "E'en the blindest man
Can tell what this resembles most;
Deny the fact who can,
This marvel of an Elephant
Is very like a fan!"

6. The *Sixth* no sooner had begun
About the beast to grope,
Then, seizing on the swinging tail
That fell within his scope,
"I see," quoth he, "the Elephant
Is very like a rope!"

And so these men of Indostan
Disputed loud and long,
Each in his own opinion
Exceeding stiff and strong,
Though each was partly in the right,
And all were in the wrong!

MORAL.

So oft in theologic wars,[2]
The disputants, I ween,
Rail on in utter ignorance
Of what each other mean,
And prate about an Elephant
Not one of them has seen!

1 http://www.noogenesis.com/pineapple/blind_men_elephant.html
2 This can also be read as scientific disparities

VI

SENSITIZING CONCEPTS

Sensitizing Concepts

Sensitizing Concepts Case Studies

0. **General**
   a. Stakeholders
   b. Product charateristics
   c. Regulatory requirements
   d. Animal use
   e. Available 3R models
   f. Chronology process SNT/EOGRTS

1. **Macro level drivers**
   a. Concern animal welfare society

2. **Macro level barriers**
   a. Level of risk aversion
   b. Ethical concerns

3. **Meso level drivers**
   a. Drawbacks conventional test
   b. Additional information 3R model
   c. Stimulating force Directive 2010/63/EU, REACH, Ph. Eur monographies
   d. Discretionary space regulatory requirements

4. **Meso level barriers**
   a. Tick box approach regulatory requirements
   b. Advantages conventional model
   c. Informational assymetrie
   d. Risks connected to assessed product/ substance
   e. No clear acceptance criteria
   f. Transition costs

5. **Micro level drivers**
   a. Advocats /advocacy coalitions
   b. Process management
   c. Scientific advantages 3R model
   d. Economic advantages 3R model
   e. Ethical advantages 3R model

6. **Micro level barriers**
   a. Challenging validation process
   b. Remaining questions / uncertainties 3R model
   c. Costs connected to development and testing 3R model

VII

DISSEMINATION OF RESEARCH RESULTS

Dissemination of Research Results

Presentations at conferences, congresses and International meetings since start of PhD project

Januari 2009	3R symposium: looking into the Chrystal ball, Utrecht, The Netherlands Presentation results Schiffelers et al, 2007
Augustus 2009	7th World Congress on alternatives and animal use in the life sciences, Rome, Italy Presentation results Schiffelers et al, 2007
November 2009	NIG annual conference, Leiden, The Netherlands Paper presentation regulatory behavior (part of article Schiffelers et al., 2012)
Juni 2010	ECPR conference: risk regulation in an age of crisis, Dublin, Ireland Paper presentation industry decision making (part of article Schiffelers et al., 2012)
September 2010	Workshop ICCVAM/NICEATM, Washington DC, USA Poster presentation draft results Rabies case study (basis for article Schiffelers et al. 2013 and 2015b)
Augustus 2011	8th World Congress on alternatives and animal use in the life sciences, Montreal, Canada Presentation draft results Rabies case study (basis for article Schiffelers et al. 2013 and 2015b) Presentation draft manuscript Schiffelers et al., 2012
June 2013	SLIM symposium, Hilversum, The Netherlands Invited speaker: presentation articles Schiffelers et al, 2012 and 2014b
September 2013	EUSAAT meeting, Linz, Austria Presentation draft results EOGRTS case study (basis for article Schiffelers et al., 2015b)
June 2014	NVT congress, Veldhoven, The Netherlands Invited speaker: presentation results panels (Schiffelers et al., 2014b)
August 2014	9th World Congress on alternatives and animal Use in the life sciences, Prague, Czech Republique Presentation draft results EOGRTS case study (Schiffelers et al., 2015a) Presentation draft results SNT case study (Schiffelers et al., 2015b)
March 2015	Symposium retirement Lukas Bruckner regulator IVI, Bern, Switzerland Invited speaker: presentation articles Schiffelers et al., 2013 and 2015b
April 2015	Workshop ECVAM EURLM, Ispra, Italy Invited speaker: presentation articles Schiffelers et al., 2012 and 2015a
September 2015	Congres IABS Egmond aan Zee, Netherlands Invited speaker: presentation articles Schiffelers et al, 2003 and 2015b Panellist in session on regulatory acceptance and use of 3R models for QC biologicals

VIII

RESEARCH APPROACH

CASE STUDY VETERINARY RABIES VACCINE POTENCY TESTING (SNT CASE: CHAPTER 6)

Research Approach

Case Study Veterinary Rabies Vaccine Potency Testing (SNT Case: Chapter 6)

This appendix defines the research approach, the methodology, the selection criteria for the SNT case and the respondent selection that are used for the SNT case study as described in Chapter 6.

Case study approach: Causal process tracing

Regulatory acceptance and use is a process that is influenced by a broad variety of drivers and barriers (Schiffelers, 2007, 2012, 2014a,b). With this manuscript the authors aim at creating a clarification of the underlying mechanism of regulatory non-acceptance of the SNT (and other 3R models) through the examination of the variables influencing the acceptance process. This means creating an in-depth picture to unravel causal mechanisms by reconstructing events and situations that have unfolded over time.
The case study approach of causal process-tracing is used for this purpose (George and Bennett, 2005; Blatter and Haverland, 2012). The goal of process-tracing is to obtain information about specific events and steps within a process through the analysis of available documents and interviewing the central actors within this process (Tansey, 2007). With this qualitative analysis technique the intervening causal process between a dependent variable (i.e., regulatory acceptance and use of 3R models) and various independent variables (e.g., scientific information, level of risk aversion, concern about animal welfare, regulatory frame, etc.) is studied.

Research methods

Case study research relies on multiple sources of evidence. Different research methods are combined in order to arrive at a comprehensive representation of the examined situation (Yin, 2003). The variables influencing the process of regulatory acceptance and use of the SNT were identified through a combination of literature review and expert interviews. The literature research provided an overview of the regulatory framework, stakeholders involved, existing testing practices and variables influencing the regulatory acceptance and use. The examined sources consisted of scientific publications, meeting reports, websites – e.g., EDQM, EMA, EPAA, PEI – and press releases.
Between 2010 and 2012 six international meetings on 3R models for vaccine testing in general and rabies vaccines in particular, were attended. (see also Schiffelers et al., 2014a). The official reports of these meeting were examined for factors that potentially drive or withhold regulatory acceptance and use of 3R models for (rabies) vaccine potency testing purposes (De Mattia et al., 2011; Jungbäck, 2012; Stokes et al., 2011, 2012).
Furthermore, a series of 15 interviews was conducted with representatives from European and US vaccine regulators, European standardization bodies and manufacturers in 2010-2012 to collect the respondents' perspectives on the variables influencing the process of acceptance and use of the SNT (see section on respondent selection). To update this information, six of these experts were interviewed once more in 2014 to collect the most recent developments in this field. The interviews were audiorecorded and transcribed. Next the transcripts were analyzed to make an inventory of drivers and barriers per stub

stage and of the optimization possibilities.
The combination of literature research, attendance of meetings and interviews resulted in an overview of factors that have influenced the acceptance and use of the SNT for the potency testing of inactivated veterinary rabies vaccines and of suggestions to optimize this process.
The interviews were semi-structured, asking open-ended questions designed to reconstruct the process and identify the drivers and barriers per subsequent sub-stage. The interviews began with the question of the involvement of the respondent in the process, a short chronology of this involvement and the position of his/her organization regarding the SNT. Next, a series of questions was asked concerning the barriers and drivers per substage of the process. Lastly, interviewees were asked to give their views on optimizing the current process of regulatory acceptance and use of the SNT within – and where possible outside – Europe. The main questions were the same for every respondent, but the focus differed depending on the respondent's involvement in the process.

Case selection
To be able to illustrate the process of regulatory acceptance and use, the case study had to meet the following criteria:
- The existing regulatory test is an animal model which is under discussion;
- there is a model available to reduce, replace or refine the existing animal model (3R model);
- this 3R model is in the process of becoming accepted/used for regulatory purposes.

The SNT case formally meets all three criteria. Furthermore, the fact that the SNT is formally accepted for regulatory purposes within the European context, offers the chance to give an in-depth description of the process and of the influence of the different causes (independent variables) on the outcome of regulatory acceptance (dependent variable).

Respondent selection
The respondent selection was done through a combination of criterion and snowball sampling (Patton, 2001). Through criterion sampling a small group of relevant respondents was selected beforehand using the selection criteria of being a (scientific, legal and/or political) expert with experience in/or knowledge of the SNT case study and with former or current involvement in this case study. Involvement means having been able to closely follow or take part in (parts of) the process of acceptance and use of the SNT. The first sample consisted of a group of 5 experts involved in the process of regulatory acceptance and use. Next, the population was broadened though snowball sampling by asking each respondent for other suitable candidates. Suitability was defined as being directly or indirectly involved in one or more of the substages of FI, ARA or UI. This might have led to a certain level of bias in the sample, since people directly or indirectly involved might be more positive about the 3R model under discussion and the strategy that was followed. However, it should be mentioned that we explicitly observed the diverging opinions during the meetings, to be able to select respondents with potentially different perspectives concerning this case study. The respondents in the first round of interviews originated from the following stakeholder groups: European standardization body (4), national regulatory authority (4) and industry (7). In the second round this division was as follows: European standardization body (2), national regulatory authority (2) and industry (2).

IX

RESEARCH APPROACH

CASE STUDY REPRODUCTIVE TOXICITY TESTING (EOGRTS CASE: CHAPTER 7)

Research Approach

Case Study Reproductive Toxicity Testing (EOGRTS Case: Chapter 7)

This appendix describes the research approach, the research methods, the criteria for the case selection and the process of respondent selection that are used for the EOGRTS case study as described in Chapter 7.

Case study approach: causal process tracing

Regulatory acceptance and use is a process that is influenced by a broad variety of drivers and barriers (Schiffelers et al., 2012, 2014 b). With this manuscript the authors aim at creating a profound clarification of the underlying mechanism of regulatory (non)acceptance of available 3R models through the examination of the variables influencing this complex problem. This means creating an in depth picture to unravel causal mechanisms, by reconstructing events and situations that have unfolded over time. For this purpose the case study approach of causal process-tracing is used (George and Bennett, 2005; Blatter and Haverland, 2012). Through this qualitative analysis technique the intervening causal process between an dependent variable (i.e. regulatory acceptance and use of 3R models) and various independent variables (e.g. scientific information, level of risk aversion, concern about animal welfare, regulatory frame, etc.) is scrutinized.

Research methods

Case study research relies on multiple sources of evidence. Different research methods are combined in order to get to a comprehensive representation of the situation examined (Yin, 2003). The variables influencing the process of regulatory acceptance and use of the EOGRTS were identified through a combination of literature review and expert interviews. The literature research provided an overview of the regulatory framework, stakeholders involved, existing testing practices and variables influencing the regulatory acceptance and use. The examined sources consisted of scientific publications, meeting reports –e.g. minutes between 2010 and 2013 of the Member States Committee of ECHA (MSC), the Risk Assessment Committee of ECHA (RAC) and the Competent Authorities for REACH and GLP (CARACAL)-, websites of involved stakeholders –e.g. OECD, EC, ECHA-, press releases and correspondence between stakeholders. In addition, a series of 18 in depth interviews was conducted between 2012 and 2013 to collect the respondents' perspectives on the process of acceptance and use of the EOGRTS (see also section on respondent selection). The interviews were semistructured, asking open-ended questions designed to reconstruct the process and identify the drivers and barriers per subsequent substage, i.e. FI, ARA and UI. The interviews began with the question of the involvement of the respondent in the process, a short chronology of this involvement and the position of his/her organization regarding the EOGRTS. Next, a series of questions were asked regarding the barriers and drivers per substage of the process. Lastly, interviewees were asked to give their views on optimizing the current process of regulatory acceptance and use of the EOGRTS within Europe. The main questions were the same for every respondent but the focus differed depending on the respondents involvement in the process.

Most interviews were audio recorded and subsequently transcribed. In those cases where interviews were not recorded (N = 4) the interviews were transcribed and made available to

the respondent for validation of the findings. Next the transcripts were analyzed to make an inventory of drivers and barriers per stub stage and of the optimization possibilities.

Case selection

In order to fully depict the issue of regulatory acceptance and use, the case study had to meet the following criteria:

- The existing regulatory test is an animal model which is under discussion;
- there is a model available to reduce, replace or refine the existing animal model (3R model);
- and this 3R model is in the process of becoming regulatory accepted/used.

The EOGRTS case meets all three criteria. The fact that the EOGRTS is already quite far in the process of becoming accepted/ used for regulatory purposes, offers the possibility to depict the full process of different causes (independent variables) influencing the outcome of regulatory acceptance and use (dependent variable).

Respondent selection

The respondents selection was done through a combination of criterion and snowball sampling (Patton, 2001). Through criterion sampling a small group of relevant respondents was selected beforehand using the selection criteria of being a scientific, legal and/or political expert with experience in/or knowledge of the EOGRTS case study and with former or current involvement in this case study. Involvement means having been able to closely follow or take part in – parts of – the process of acceptance and use of the EOGRTS. The first sample existed of a group of 5 experts involved in the process of regulatory acceptance and use of the EOGRTS. Next, the population was broadened through snowball sampling asking each respondent for other suitable candidates. Suitability was defined as direct or indirect involvement in one or more of the substages of FI, ARA or UI. We have explicitly looked for respondents with different perspectives on the case both in terms of stakeholder groups as in terms of opinion regarding the EOGRTS. The respondents came from the following stakeholder groups: European legislators (5) European and national regulatory authorities (4), industry/contract research organizations (5), academia (2) and animal welfare organizations (2).

REFERENCES

Abraham, J. (1995). *Science, Politics and the Pharmaceutical Industry: Controversy and Bias in Drug Regulation.* London, UK: Routledge.

Anon. (2010). *Sixth Report from the Commission to the Council and the European Parliament on the Statistics on the number of animals used for experimental and other scientific purposes in the member states of the European Union COM* (2010) 511 final. Brussels, Belgium: EC.

APS (Australian Public Service Commission) (2007). *Tackling Wicked Problems: A Public Policy Perspective*. Commonwealth of Australia. Available at: http://www.apsc.gov.au/publications-and-media/archive/publications-archive/tackling-wicked-problems

Arts, J.H.E., Muijser, H., Jonker, D., van de Sandt, J.J.M., Bos, P.M.J., Feron, V.J. (2008). Inhalation toxicity studies: OECD guidelines in relation to REACH and scientific developments. *Exp. Toxicol. Pathol.* 60: 125–133.

Asplund, M. & Sandin, R. (1999). The survival of new products. *Review of Industrial Organization* 15: 219–236.

Bailey J. (2005). Non-human primates in medical research and drug development: a critical review. *Biogenic Amines.* 19: 235–55.

Bakker, W. and van Waarden, F. (eds.) (1999). *Ruimte rond regels: stijlen van regulering en beleidsuitvoering vergeleken.* Amsterdam, The Netherlands: Boom.

Bakker, W. (2001). *Sturen op de tijstroom: Onderwijs voor werkende jongeren en beleid tussen economie en ontplooiing*, 1945-1995. The Netherlands: Thela Thesis.

Balls, M. (2009). *The Three Rs and the Humanity Criterion; An abridged version of the Principles of Humane Experimental Technique by W.M.S. Russell and R.L. Burch*, Frame Nottingham, UK.

Barrieu, P. and Sinclair-Desgagne, B. (2003). *The paradox of precaution*. Scientific series, Cirano, Centre interuniversitaire de recherché en analyse des organisations. Montreal, Canada.

Basketter, D.A., York, M., McFadden, J. P., and Robinson, M. K. (2004). Determination of skin irritation potential in the human 4-h patch test. *Contact Dermatitis* 51: 1–4.

Beck, U. (1992). *The Risk Society, Towards a New Modernity.* London, UK: Sage Publications Ltd.

Berkhout, F., Smith A. and Stirling A. (2003). Socio-technological regimes and transition contexts. SPRU Electronic Working Paper. SPRU Paper No. 106, *Science & Technology Policy Research.* Brighton, United Kingdom http://www.sussex.ac.uk/Units/spru/publications/imprint/sewps/sewp106/sewp106.pdf

Bernard, C. (1865) *Introduction à l'étude de la medicine experimentale.* Available at : http://www.gutenberg.org/cache/epub/16234/pg16234-images.html

Bergek, A., Jacobsson, S., Carlsson, B., Lindmarki, S., Rickne, A. (2005). *Analysing the dynamics and functionality of sectoral innovation systems – a manual*. In: 10 Year Anniversary DRUID Summer Conference, Copenhagen, June 27–29.

Blaauboer B.J., Barratt M.D., Houston J.B. (1999). *The integrated use of alternative methods in toxicological risk evaluation.* ECVAM Integrated Testing Strategies Task Force Report 1. *ATLA* 27: 229-237.

Blaauboer BJ, Andersen ME. (2007) The need for a new toxicity testing and risk analysis paradigm to implement REACH or any other large scale testing initiative. *Arch. Toxicol*. 81: 385-387.

Blaauboer, B.J. (2015) The long and winding road of progress in the use of *in vitro* data for risk assessment purposes: from "carnation test" to integrated testing strategies. *Toxicology* 332: 4-7.

Blatter, J.K. and T. Blume (2008). In Search of Co-variance, Causal Mechanisms or Congruence? Towards a Plural Understanding of Case Studies. *Swiss Political Science Review* 14: 315-356.

Blatter, J. and Haverland, M. (2012). *Designing Case Studies: Explanatory Approaches in Small-N Research.* Houndsmills Basingstoke: Palgrave Macmillan.

Boeije, H. (2005). *Analyseren in kwalitatief onderzoek*. Amsterdam: Boom.

Boekholt, P., Lankhuizen, M., Arnold, E., Clark, J., Kuusisto, J, de Laat, B., Simmonds, P., Cozzens, S., Kingsley, G., Johnston, R. (2001). *An International Review of Methods to Measure Relative Effectiveness of Technology Policy Instruments.* Brighton/ Amsterdam: Technopolis.

Boon, W.P.C., Moors, E.H.M. and Meijer A.J. (2014) Exploring dynamics and strategies of niche protection. *Research Policy* 43 (4): 792-803.

Bottini, A.A., Alepee, A., Phillips, B., Gribaldo, L., De Silva, O., Hartung, T., Hendriksen, C.F.M., Kuil, J., Pazos, P., Rhein, C., Schiffelers, M.J.W.A., Stokes, W., Theobald, A., Vidal, J.M., van de Sandt, H., Breier, S., Sintes, J.R., Blaauboer, B.J., (2008). Optimization of the post-validation process, the report and recommendations of ECVAM workshop 67a. *ATLA* 36: 353–366.

Boulter, L. and Bendell, T. (2002). Managing the technology transfer process, *IEEE Transaction* 643-648.

Bowen, G.A. (2006). Grounded Theory and Sensitizing Concepts. *International Journal of Qualitative Methods* 5 (3). http://ejournals.library.ualberta.ca/index.php/IJQM/index

Breyer, S. G. (1993). *Breaking the Vicious Circle: toward effective risk regulation*. London, UK: Harvard University Press.

Bremer, S., Balduzzi, D., Cortvrindt, R., Daston, G., Eletti, B., Galli, A.,Huhtaniemi, I., Laws, S., Lazzari, G., Liminga, U., et al. (2005). *The effects of chemicals on mammalian fertility*. The report and recommendations of ECVAM Workshop 53—The first strategic workshop of the EU ReProTect Project. *ATLA* 33: 391–416.

Bruckner, L., Cussler, K., Halder, M., et al. (2003). Three Rs approaches in the quality control of inactivated rabies vaccines. *ATLA* 31: 429-454.

Bryson, J.M., Crosby, B.C. (1992). *Leadership for the Common Good.* San Francisco, CA: Jossey-Bass Publishers.

Busfield, J. (2006). Pills, power, people: sociological understanding of the pharmaceutical industry. *Sociology* 40: 297-314.

Capoccia G., and Kelemen, R.D. (2007). The Study of Critical Junctures: Theory, Narrative, and Counterfactuals in Historical Institutionalism *World Politics* 59 (3): 341-369.

Casey, W., Schmitt, M., McFarland, R. et al. (2011). Improving animal welfare and reducing animal use for human vaccine potency testing: State of the science and future directions. *Proc Vaccinol.* 5: 33-46. http://dx.doi.org/10.1016/j.provac.2011.10.003

Castle P. (1996). Alternatives to animal testing: achievements and recent developments in the European Pharmacopoeia. *Dev.Biol.Stand.* 86: 21-29.

Carpenter, D. (2010). *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton, USA: Princeton University Press.

Chakrabarti, A. K., and J. Hauschildt. 1989. The division of labor in innovation management. *R&D Management* 19 (2): 161–71.

Charmaz, K. (2003). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.). *Strategies for qualitative inquiry* (2nd ed.: 249-291). Thousand Oaks, CA: Sage.

Charton, E. (2008). Current regulations on alternatives and review of progress European. Pharmacopoeia activities - an overview. https://www.edqm.eu/medias/fichiers/e_charton_article_new_approaches.pdf

Cehtra (2012). Report on Survey of Worldwide CROs: Costs and Practicalities of Two New OECD Guidelines for Testing Chemical Substances OECD 443, Extended One-Generation Reproductive Toxicity Study, and OECD 488, Transgenic Rodent Somatic and Germ Cell Mutation Assay. https://echa.europa.eu/documents/10162/13628/survey_report_worldwide_cros_en.pdf

Cohen, M., Marsh, J., and Olsen, J. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly* 17: 1–25.

Cooper, R.L., Lamb, J.C., Barlow, S.M., Bentley, K., Brady, A.M., Doerrer, N.G., et al., (2006). A tiered approach to life stages testing for agricultural chemical safety assessment. *Crit. Rev. Toxicol.* 36 (1): 69–98.

Cooper, J. and Jennings, M. (2008). *Advancing animal welfare and the 3Rs in the batch testing of veterinary vaccines.* RSPCA, United Kingdom.

Council of Europe (2010). *European Pharmacopoeia* 7.0. Vol.1. Strasbourg, France: Council of Europe.

Cozijnsen, A.J., Vrakking, W.J. & van IJzerloo, M. (2000) Success and failure of 50 innovation projects in Dutch companies. *European Journal of Innovation Management* 3: 150–159.

Curren, R.D., Southee, J.A., Spielmann, H., Liebsch, M., Fentem, J., and Balls. M. (1995). The role of prevalidation in the development, validation and acceptance of alternative methods. *ATLA* 23: 211-217.

Darwin, C. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* John Murray, London.

De Bruijn, H. and Koopmans, M. (2005). Enforcing the law: Strategies used by regulates and enforcement officials. Paper ECPR. http://regulation.upf.edu/index.php?id=budapest_2005

De Haan, J, and Rotmans, J. (2011). Patterns in transitions: Understanding complex chains of change. *Technological Forecasting & Social Change* 78: 90–102.

De Leeuw, W. (2004). De ethische toetsing van dierproeven: wat heeft Europa ons te bieden. In J. Swart, J. Wolters, and H. Zwart (eds.), *DEC's in Discussie: de Beoordeling van Dierproeven in Nederland.* Budel, The Netherlands: Damon.

De Mattia F., Chapsal J.M., Descamps J., Halder, M., Jarrett, N., Kross, I., Mortiaux, F., Ponsar, C., Redhead, K., McKelvie. J., Hendriksen, C. (2011). The consistency approach for quality control of vaccines - a strategy to improve quality control and implement 3Rs. *Biologicals* 39 (1): 59-65. http://dx.doi.org/10.1016/j.biologicals.2010.12.001

DiMaggio, P. (1988). Interest and agency in institutional theory. In L. Zucker (ed.): *Institutional patterns and culture.* Cambridge, MA: Ballinger Publishing Company: 3-22.

DiMasi, J.A., Hansen, R.W., and Grabowski, H. (2003). The price of innovation: new estimates of drug development costs. *J Health Econ.* 22: 151-185.

Denise, L. (1999). Collaboration vs. C-Three (cooperation, coordination, and communication). *Innovating Reprint* 7 (3) The Rensselaerville Institute. http://www.seattle.gov/neighborhoods/education/documents/Collaborationvs.C-Three.pdf.

Doig, J. W. and Hargrove, E. C. (1987). Leadership and Innovation. Entrepreneurs in Government, Baltimore, MD: The John Hopkins University Press.

Dosi, G., (1982). Technological paradigms and technological trajectories. A suggested interpretation of determinants and directions of technological change. *Research Policy* 11: 147-162.

Dupree, M., Etienne, J., and Lecoze, J. C. (2007). The regulator-regulatee interaction: insights taken from a high risk business firm. 2nd Annual Cambridge Conference on regulation, Inspection & Improvement, Cambridge. http://www.cbr.cam.ac.uk/pdf/Dupre_et_al_Paper.pdf

EC (2013). Seventh Report on the Statistics on the Number of Animals used for Experimental and other Scientific Purposes in the Member States of the European Union. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52013DC0859

ECETOC (2009). Workshop on Triggering and Waiving Criteria for the Extended One-generation Reproduction Toxicity Study: 14–15 April 2008, Barza d'Ispra. Workshop Report No. 12 Brussels, Belgium: European Centre for Ecotoxicology & Toxicology of Chemicals. *ATLA*. 37(2): 1-7.

ECHA (2010). Evaluation under REACH, Progress Report 2010. ECHA-11-R-001-EN, Helsinki, Finland: European Chemicals Agency. https://echa.europa.eu/documents/10162/13628/evaluation_under_reach_progress_report_2010_en.pdf

ECHA (2012). Factsheet guidance on data sharing ECHA-12-GF-01-EN. Available at: http://echa.europa.eu/documents/10162/13631/data_sharing_en.pdf.

ECHA (2014). The Use of Alternatives to Testing on Animals for the REACH Regulation. Second report under Article 117(3) of the REACH Regulation. http://echa.europa.eu/documents/10162/13639/alternatives_test_animals_2014_en.pdf

Edquist, C. (1997). Systems of innovation approaches – their emergence and characteristics. In: Edquist, C., McKelvey,M. (Eds.), *Systems of Innovation: Technologies. Institutions and Organizations* (1–35). London: Pinter.

Edquist, C., 2005. Systems of innovation: perspectives and challenges. In: Fagerberg, J., Mowery, D.C., Nelson, R.R. (Eds.) *The Oxford Handbook of Innovation* (181–208). New York: Oxford University Press.

Ekwall, B., Clemedson, C., Crafoord, B., Ekwall, B., Hallander, S., Walum, E., and Bondesson, I. (1998). MEIC evaluation of acute systematic toxicity Part V. Rodent and human toxicity data for the 50 reference chemicals. *ATLA*. 26: 571–616.

Ellemers, J. E. (1976). Veel kunnen verklaren of iets kunnen veranderen: krachtige versus manipuleerbare variabelen. In: Ellemers, J.E. (1995). *Modernisering, macht, migratie.* Opstellen over maatschappij en beleid. Amsterdam Meppel: Boom.

EPAA, (2007). Regulatory acceptance of alternative approaches, EPAA workshop, 18-19 June, Centre Brochette, Brussels. http://www.epaa.eu.com

EU (2006). Regulation (EC) No 1907/2006 of the European parliament and of the council of 18 December 2006 concerning the registration, evaluation, authorization and restriction of chemicals (REACH), establishing a European chemicals agency. http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:396:0001:0849:EN:PDF/

EU (2008a). Regulation (EC) No 440/2008 of 30 May 2008 Laying Down Test Methods Pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:142:0001:0001:EN:PDF/

EU (2008b). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 december 2008 on Classification, Labelling and Packaging of Substances and Mixtures, Amending and Repealing Directives 67/548/EEC and 1999/45/EC, and Amending Regulation (EC) No 1907/2006. http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:353:0001:1355:en:PDF/

EU (2010). Directive of the European parliament and of the council of 22 September 2010 on the protection of animals used for scientific purposes. OJ Eur. Union L276, 33–79. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:276:0033:0079:en:PDF/

European Pharmacopoeia (2013). Monograph 04/2013:0451. Rabies vaccine (inactivated) for veterinary use. Ph. Eur. 8th edition. Strasbourg, France: European Department for the Quality of Medicines within the Council of Europe.

Fagerberg, J., (2006). Innovation: A guide to the Literature. In Fagerberg, J. and Mowery. D. (Eds): *The Oxford Handbook of Innovation*. DOI: 10.1093/oxfordhb/9780199286805.003.0001

Fegert, I., Billington, R., Botham, P., Carney, E., FitzGerald, R.E., Hanley, T., Lewis, R., Marty, M.S., Schneider, S., Sheets, L.P., Stahl, B., van Ravenzwaay, B., (2012). Feasibility of the extended one-generation reproductive toxicity study (OECD 443). *Reprod. Toxicol*. 34 (3): 331–339.

Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research Qualitative Inquiry. 12 (2): 219-245. http://qix.sagepub.com/content/12/2/219.abstract

Flyvbjerg, B. (2012). Making social science matter In: Papanagnou G. (ed.) *Social Science and Policy Challenges Democracy, Values and Capacities* (25-56). Paris: UNESCO Publishing.

Franco, N.H. (2013). Animal Experiments in Biomedical Research: A Historical Perspective. *Animals* 3: 238-273; doi:10.3390/ani3010238 http://www.navs.org/file/Animals-Experiments-in-Biomedical-Research.pdf

Freeman, C. (1987) *Technology Policy and Economic Performance: Lessons from Japan*. London: Pinter Publishers Ltd.

Freriks, A., van der Meulen, B., van den Belt, H., et al. (2005). *Noodzakelijk kwaad, Evaluatie Wet op de dierproeven.* http://www.nvdec.nl/page8/page1/files/Evaluatie%20WOD.pdf

Garthoff, B., 2005. Alternatives to animal experimentation: the regulatory background. *Toxicol. Appl. Pharmacol.* 207 (Suppl 2): 388–392.

Geels, F. and Kemp, R. (2000). Transities vanuit sociotechnisch perspectief, report for the study "Transities en Transitiemanagement" of ICIS and MERIT for the Department of Environment for the NMP-4, Okt 2000, UT, Enschede en MERIT, Maastricht. http://kemp.unu-merit.nl/pdf/geelskemp.pdf

Geels, F.W. (2002). Technological transitions as evolutionary configuration processes: a multi-level perspective and a case study. *Research Policy* 31: 1257-1274. http://dx.doi.org/10.1016/S0048-7333(02)00062-8

Geels, F.W. (2006). Multi-level perspective on system innovation: relevance for industrial transformation. In X. Olshoorn and A. J. Wieczorek (Eds.) *Understanding Industrial Transformation: Views from Different Disciplines* (163-186). The Netherlands: Springer.

Geels, F.W. and Schot, J. (2007) Typology of sociotechnical transition pathways. *Research Policy* 36: 399–417. doi:10.1016/j.respol.2007.01.003

George, A. L. and Bennett, A. B. (2005). *Case Studies and Theory Development in the Social Sciences.* Cambridge, Massachusetts: The Mitt Press.

Gerde, P. (2005). Animal models and their limitations: On the problem of high-to-low dose extrapolations following inhalation exposures. *Experimental and Toxicologic Pathology* 57: 143–146. doi:10.1016/j.etp.2005.05.016

Giddens, A. (1999). Risk and responsibility. *The Modern Law Review* 62: 1-10.

Gilbert, N. (2011). Data gaps threaten chemical safety law: European companies are not providing robust information to regulators or alternatives to animal experiments. *Nature* 475, 150–151. Available at: http://www.nature.com/news/2011/110712/full/475150a.html/

Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research.The Qualitative Report, 8(4): 597-606. Avaliable at: http://nsuworks.nova.edu/tqr/vol8/iss4/6

Gold, R. (1958). Roles in sociological field observation. *Social Forces* 36: 217-213.

Grin, J., and Van de Graaf, H. (1996). Implementation as communicative action – an interpretative understanding of interactions between policy actors and target groups. *Policy Sciences* 29: 291–319.

Guy, A., Gauthier, C., and Griffin, G. (2008a). Adopting alternative methods for regulatory testing in Canada. Proc. 6th World Congress on Alternatives & Animal Use in the Life Sciences August 21-25, 2007, Tokyo, Japan. *AATEX* 14, Special Issue: 323-327.

Guy, A and Griffin, G. (2008b) Incentives and impediments to adopting alternative shellfish testing methods in Canada. Proc. 6th World Congress on Alternatives & Animal Use in the Life Sciences August 21-25, 2007, Tokyo, Japan. *AATEX* 14, Special Issue: 763-767.

Hackam, D.G, and Redelmeier D.A. (2006) Translation of research evidence from animals to humans. *JAMA* 296: 1731–2.

Hajer, M.A. (1995). *The Politics of Environmental Discourse: Ecological Modernization and the Policy Process.* New York, NY: Oxford University Press.

Harder, B.T. and Benke, R. (2006). Transportation Technology Transfer: Successes, Challenges, and Needs: A Synthesis of Highway Practice National Cooperative Highway Research Program. NCHRP Synthesis 355. Available at: http://www.nap.edu/catalog/13923/transportation-technology-transfer-successes-challenges-and-needs

Hartung, T. and Daston, G. (2009). Are In Vitro Tests Suitable for Regulatory Use? *Toxicological Sciences* 111(2): 233–237. doi:10.1093/toxsci/kfp149

Hartung, T. and Rovida, C. (2009). Opinion: chemical regulators have overreached. *Nature* 460: 1080–1081.

Hendriksen, C. and Van der Gun, J. (1995). Animal models and alternatives in the quality control of vaccines: are *in vitro* methods or *in vivo* methods the scientific equivalent of the emperor's clothes. *ATLA*. 23: 61-73.

Hendriksen, C.F.M. (2000). Replacement, reduction and refinement and biologicals: about facts, fiction and frustration. In: Balls, M., van Zeller, A.M., Halder, M.E. (Eds.), *Progress in the Reduction, Refinement and Replacement of Animal Experimentation* (51–63). Amsterdam: Elsevier Science B.V.

Hendriksen, C.F.M. (2006). Replacement, reduction and refinement in the production and quality control of immunobiologicals, *AATEX* 11(3): 155-161.

Hendriksen, C., Arciniega, J.L., Bruckner, L., Chevalier, M., Coppens, E., Descamps, J., Duchêne, M., Dusek, D.M., Halder, M., Kreeftenberg, H., Maes, A., Redhead, K., Ravetkar, S.D., Spieser, J-M., Swam, H. (2008), The consistency approach for the quality control of vaccines. *Biologicals* 36: 73-77.

Hendriksen, C.F. (2009). Replacement, reduction and refinement alternatives to animal use in vaccine potency measurement. *Expert Rev.Vaccines* 8(3): 313-322.

Hekkert, M., Suurs, R.A.A., Negro, S., Kuhlmann, S., Smits, R. (2007). Functions of Innovation Systems: A new approach for analysing technological change. *Technological Forecasting and Social Change* 74 (4): 413–432.

Heritier, A. (2001). Regulator-regulatee interaction in the Liberalized Utilities: Access and Contract Compliance in the Rail Sector Max-Planck-Projectgruppe, Recht der Gemeinschaftsgüter, Bonn 2001/12. http://www.coll.mpg.de/publications/regulator-regulatee-interaction-liberalized-utilities

Holmes, A.M., Creton, S., Chapman, K., (2010). Working in partnerships to advance the 3Rs in toxicity testing. *Toxicology* 267: 14–19.

Hood, C., Rothstein, H., and Baldwin, R. (2001). *The Government of Risk: Understanding Risk Regulation Regimes.* New York, USA: Oxford University Press.

Hood, C. (2002). The risk game and the blame game. *Government and Opposition* 37: 15-37.

Hoogma, R. (2000). *Exploiting Technological Niches*., Enschede, The Netherlands: Twente University Press.

Hoogma, R., Kemp, R., Schot, J., Truffer, B. (2002). *Experimenting for Sustainable Transport. The approach of Strategic Niche Management.* London/New York: Spon Press.

Hoonakker, M. (2011). Vaccines, animal experiments and alternatives: A survey of reduction, refinement and replacement strategies in human and veterinary vaccine development, production and quality control Projectreport as part of the programme Dierproeven begrensd II ZonMw.

Intomart GfK (2004). Publieke opinie over dierproeven in Nederland, a study commissioned by the Dutch Society for the Protection of Animals, Hilversum, The Netherlands.

Jacobsson, S., and Bergek, A. (2004). Transforming the energy sector: the evolution of technological systems in renewable energy technology. *Industrial and Corporate Change* 13 (5): 815–849.

Jaffe, M. E. (1994). Regulation, litigation and innovation in the pharmaceutical industry: An equation for safety. In Hunziker, J. R. and Jones, T. O. (Eds.) *Product Liability and Innovation; Managing Risk in an Uncertain Environment* (120-128). Washington D.C., USA: The National Academies Press.

Janer, G., Hakkert, B.C., Slob, W., Vermeire, T., Piersma, A.H., (2007a). A retrospective analysis of the two-generation study: what is the added value of the second generation? *Reprod. Toxicol*. 24: 97–102.

Janer, G., Hakkert, B.C., Piersma, A.H., Vermeire, T., Slob, W., (2007b). A retrospective analysis of the added value of the rat two-generation reproductive toxicity study versus the rat subchronic toxicity study. *Reprod. Toxicol*. 24: 103–113.

Jungbäck, C. (ed.) (2012). Potency testing of veterinary vaccines for animals: The way from *in vivo* to *in vitro*. *Dev Biol.* 134: 149-152.

Kasamatsu, T. and Kohda, K. (2006). Commentary: Balancing risks. *Regul Toxicol Pharmacol.* 46: 100-104.

Kemp, R. (1994). Technology and the transition to environmental sustainability—the problem of technological regime shifts. *Futures* 26 (10): 1023–1046.

Kemp, R., Schot, J., Hoogma, R. (1998). Regime Shifts to Sustainability Through Processes of Niche Formation: The approach of Strategic Niche Management. *Technology Analysis & Strategic Management* 10 (2): 175-195.

Kemp, R., Rip, A., Schot, J.W. (2001). Constructing transition paths through the management of niches. In: Garud, R., Karnoe, P. (Eds.) *Path Dependence and Creation.* (269–299) NY: Lawrence Erlbaum, Mahwah.

Kemp, R. and Rotmans, J. (2005). Transition management: managing the co-evolution of technical, environmental and social systems. In: Weber, K.M., Hemmelskamp, J. (Eds.) *Towards Environmental Innovation Systems.* (33–55) Heidelberg: Springer.

Kemp, R. (2010). The Dutch energy transition approach. *International Economics and Economical Policy* 7: 291-316. http://dx.doi.org/10.1007/s10368-010-0163-y

Kern, F. and Howlett, M. (2009). Implementing transition management as policy reforms: a case study of the Dutch energy sector. *Policy Sci.* 42: 391–408.

Kingdon, J. W. (1995). *Agenda's, Alternatives and Public Policies* (2nd ed). New York, USA: Longman Pub. Group.

Klein, K.L., Scott, W.J. and Wilson, J.G . (1981). Aspirin-induced teratogenesis: a unique pattern of cell death and subsequent polydactyly in the rat. *J Exp Zool.* 216: 107 – 12.

Knight A. (2007). Systematic reviews of animal experiments demonstrate poor human clinical and toxicological utility. *ALTEX* 14: 125–30.

Knight, A. (2008). Estimates of Worldwide Laboratory Animal Use. Letters, *ATLA* 36: 494-495.

Kooijman, M. (2013). *Why animal studies are still being used in drug development: An innovation system perspective.* (PhD thesis) Utrecht University, Utrecht.

Kooijman, M., van Meer, P.J.K., Gispen-de Wied, C.C., Moors, E.H.M., Hekkert, M.P., Schellekens, H. (2013). The risk-based approach to ATMP development – Generally acceptedby regulators but infrequently used by companies. *Regulatory Toxicology and Pharmacology* 67: 221–225.

Koppenjan, J.F.M. and Klijn, E.H. (2004). *Managing uncertainties in networks: a network approach to problem solving and decision making.* NY: Routledge

Kotter, J. and Rathgeber, H. (2006). *Our Iceberg is Melting: Changing and Succeeding Under Any Conditions.* London: Pan Macmillan.

Krämer, B., Schildger, H., Behrensdorf-Nicol. H. A. et al. (2009). The rapid fluorescent focus inhibition test is a suitable method for batch potency testing of inactivated rabies vaccines. *Biologicals* 37: 119-126. http://dx.doi.org/10.1016/j.biologicals.2009.01.001

Krämer, B., Bruckner, L., Daas, A., and Milne, C. (2010). Collaborative study for validation of a serological potency assay for rabies vaccine (inactivated) for veterinary use. *Pharmeur Bio Sci Notes*. 37-55.

Krämer, B., Kamphuis, E., Hanschmann, K. M. et al. (2013). A multi-dose serological assay suitable to quantify the potency of inactivated rabies vaccines for veterinary use. *Biologicals* 41: 400-406. http://dx.doi.org/10.1016/j.biologicals.2013.08.003

Krapohl, S. (2008). *Risk Regulation in the Single Market; The Governance of Pharmaceuticals and Foodstuffs in the European Union*. Palgrave Studies in European Union Politics, Hampshire, UK: Palgrave Macmillan.

Krewski, D., Andersen, M.E., Mantus, E., Zeise, L., (2009). Toxicity testing in the 21st century: implications for human health risk assessment. *Risk Anal.* 29 (4): 474–479.

Krul, C., Freidig, A., Van de Sandt, H. (2006). In: Swart, J., Groothuis, G., Horbach, J., Van der Valk, J. (Eds.), *Kan het ook anders?: Beschouwingen over alternatieven voor dierproeven.* Budel, The Netherlands: Damon.

Kulpa-Eddy, J. and Dusek, D. (2011). Application of the consistency approach to reduce animal use in vaccine potency testing. *Proc Vaccinol.* 5: 232-235. http://dx.doi.org/10.1016/j.provac.2011.10.024

Milne, C. and Buchheit, K.H. (2012). EDQM's 3R activities in the field of quality control of vaccines. *ALTEX* Proc 1/12, Proceedings of WC8: 65-69.

Kuschel, G.B. (2012). Rationality and Phronesis in Economics: A Rhetorical Moment. Thesis at Erasmus University Rotterdam.

Langley, G. (2009). The validity of animal experiments in medical research. RSDA 1, 161–168 http://www.drhadwentrust.org/downloads/publications/LangleyValidityofAnimalResearchEnglish09__2_.pdf.

Lee, Y.C., Li, M.L., Yen, T.M. and Huang, T.H. (2010). Analysis of adopting an integrated desion making trail and evaluation laboratory on a technology acceptance model. *Expert Systems with Applications* 37: 1745-1754.

De Leeuw, W. (2004). De ethische toetsing van dierproeven: wat heeft Europa ons te bieden? In: Swart, J., Wolters, J. and Zwart, H. (Eds.), *DEC's in Discussie: de Beoordeling van Dierproeven in Nederland.* Budel, The Netherlands: Damon.

Leist, M., Hartung, T., and Nicotera, P. (2008). The dawning of a new age of toxicology. *ALTEX* 25: 103–114.

Liebsch, M. and Spielmann, H. (2002). Currently available *in vitro* methods used in the regulatory toxicology. *Toxicol. Lett.* 127: 127-134.

Light, D. W. and Warburton, R. (2011). Demythologizing the high costs of pharmaceutical research. *BioSocieties* 6: 34-50. doi:10.1057/biosoc.2010.40

Lilienblum, W., Dekant, W., Foth, H., Gebel, T., Hengstler, J.G., Kahl, R., Kramer, P.J., Schweinfurth, H., Wollin, K.M., (2008). Alternative methods to safety studies in experimental animals: role in the risk assessment of chemicals under the new European chemicals legislation (REACH). *Reprod. Toxicol.* 82: 211–236.

Lindblom, C. E. (1959). The science of ''Muddling through''. In A. Etzioni (Ed.), *Readings on modern organizations.* Englewood Cliffs, NJ: Prentice Hall.

Long, M.E. and Griffin, G. (2012). Challenges and opportunities for the implementation of the Three Rs in Canadian vaccine quality control. *Regulatory Toxicology and Pharmacology* 63: 418–425.

Lundvall, B.-Å.(1988). Innovation as an interactive process: from user-producer interaction to the national system of innovation, In: Dosi, G. Freeman, C., Nelson R., Silverberg, G.. and Soete, L. (Eds.), *Technical Change and Economic Theory Innovation as an interactive process: from user-producer interaction to the national system of innovation*. London: Pinter.

MacLachlan, A. (1994). The chemical industry: Risk management in Today's Product liability Environment. In Hunziker, J. R. and Jones, T. O. (Eds.), *Product Liability and innovation; Managing Risk in an Uncertain Environment.* Washington DC, USA: The National Academy Press.

Majone, G., & Wildavsky, A. (1978). Implementation as evolution. In H. Freeman (Ed.), *Policy studies annual review* Vol. 2: 103–117. Beverly Hills, CA: Sage Publications.

Majone, G. (2010). Foundations of risk regulation: Science, decision-making, policy learning and institutional reform. *EJRR.* 1: 5-19.

Malyshev, N. (2006). Regulatory policy: OECD experience and evidence. *Oxf Rev Econ Pol.* 22: 274-299.

Mamat B.B. and Roslan, S.B. (2012). Critical success factors (csfs) on technology transfer effectiveness in manufacturing industry: a critical review. *International Journal of Business, Economics and Law* 1: 173-170.

Markard, J., Truffer, B. (2008). Technological innovation systems and the multi-level perspective: Towards an integrated framework, *Res Policy* 37: 596–615 doi:10.1016/j.respol.2008.01.004

Martic´-Kehl, M.I., Schibli, R., Schubiger, P.A., (2012). Editorial: Can animal data predict human outcome? Problems and pitfalls of translational animal research. *Eur. J. Nucl.Med. Mol. Imaging* 39: 1492–1496. DOI 10.1007/s00259-012-2175-z

Martin, M.T., Mendez, E., Corum, D.G., Judson, R.S., Kavlock, R.J., Rotroff, D.M., et al., (2009). Profiling the reproductive toxicity if chemicals from multi-generation studies in the toxicity reference database. *Toxicol. Sci.* 110 (1): 181–190.

Martin, M.T., Knudsen, T.B., Reif, D.M., Houck, K.A., Judson, R.S., Kavlock, R.J., Dix, D.J., (2011). Predictive model of rat reproductive toxicity from ToxCast high throughput screening. *Biol. Reprod.* 85: 327–339.

Matthews R.A.J. (2008). Medical progress depends on animal models—doesn't it? *J R Soc Med.* 101: 95–8.

Mathison, S. (1988). Why triangulate? *Educational Researcher*, 17(2): 13-17.

Meadowcroft, J. (2009). What about the politics? Sustainable development, transition management, and long term energy transitions. *Policy Sci.* 42: 323–340. DOI 10.1007/s11077-009-9097-z

Meijer, A. J. (2013). From Hero-Innovators to Distributed Heroism: An in-depth analysis of the role of individuals in public sector innovation, *Public Management Review*. Available at: http://dx.doi.org/10.1080/14719037.2013.806575

Metz. B., Hendriksen, C., Jiskoot, W. and Kersten, G. (2002). Reduction of animal use in human vaccine quality control : opportunities and problems, *Vaccine* 20: 2411-2430

Nathans, H. (1997). Adviseren als tweede beroep, resultaat bereiken als adviseur, herziene druk. Deventer, The Netherlands: Kluwer bedrijfsinformatie.

Nelson R.R.and Winter, S.G. (1977). In search of a useful theory of innovations. *Research Policy*. 6: 36-76.

Nelson, R.R., and Winter, S.G. (1982). *An Evolutionary Theory of Economic Change*. Cambridge: Harvard University Press.

Nelson, R., (2006). *Economic Development From the Perspective of Evolutionary Economic Theory.* Columbia University Working Papers in Technology Governance and Economic Dynamics no. 2 http://technologygovernance.eu/files/main/2006013112494141.pdf

NIEHS (1997*). Validation and regulatory acceptance of toxicological test methods*: A Report of the ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publication No. 97-3981, NIEHS, USA.

NRC (National Research Council), (2004). *Accelerating Technology Transitions: Bridging the Valley of Death for Materials and Processes in Defense Systems.* Washington, D.C.: The National Academies Press.

NRC (National Research Council), (2007). Committee on toxicity testing and assessment of environmental agents. *Toxicity Testing in the 21st Century: A Vision and a Strategy.* Washington, D.C.: The National Academies Press,

Nuffield Council on Bioethics (2005). The ethics of research involving animals. Available at: http://nuffieldbioethics.org/wp-content/uploads/The-ethics-of-research-involving-animals-full-report.pdf

NVWA (Nederlandse Voedsel- en Warenautoritiet) (2014). *Zodoende 2013: Jaaroverzicht van de Nederlandse Voedsel- en Warenautoriteit over dierproeven en proefdieren*. November 2014. The Hague, The Netherlands.

OECD (2001). *TG 416. OECD Guideline for Testing of Chemicals. Two-Generation Reproduction Toxicity Study.* Organisation for Economic Co-operation and Development, Paris, France.

OECD (2005). *Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment*. Series on Testing and Assessment, No. 34. OECD, Paris, France.

OECD (2011a). *TG 443. OECD Guideline for Testing of Chemicals. Extended One- Generation Reproductive Toxicity Study*. Organisation for Economic Cooperation and Development, Paris, France.

OECD (2011b). *Guidance Document 117 on the Current Implementation of Internal Triggers in Test Guideline 443 for an Extended one Generation Reproductive Toxicity Study, in the United States and Canada.* Organisation for Economic Cooperation and Development, Paris, France.

OIE (2012). Rabies. In Manual of standards for diagnostic tests and vaccines for terrestrial animals (276-291). 6th edition. World Organization for Animal Health. http://www.oie.int/eng/normes/mmanual/A_summry.htm

Olson, M. K. (1997). Firm characteristics and the speed of FDA approval. *JEMS.* 6: 377-401.

Olson, D.L., Birge, J.R., and Linton, J.D. (2014) Introduction to risk and uncertainty management in technological innovation. *Technovation* 34 (8): 395–398.

Osborne, S. P. and Brown, L. (2005). *Managing Change and Innovation in Public Service Organizations*. Milton Park: Routledge.

Patton, M. Q. (1980). *Qualitative evaluation methods*. Beverly Hills, CA: Sage.

Patton, M. Q. (2001). *Qualitative Research and Evaluation Methods.* (2nd ed.). Thousand Oaks, CA, USA: Sage Publications.

Pedersen, F., De Bruijn, J., Munn, S., Van Leeuwen, K., (2003). *Assessment of Additional Testing Needs Under REACH. Effects of (Q)SARS, Risk Based Testing and Voluntary Industry Initiatives.* JRC Report EUR 20863.

Piersma, A.H., Rorije, E., Beekhuijzen, M.E.W., Cooper, R., Dix, D.J., Heinrich Hirsch, B., Martin, M.T., Mendez, E., Muller, A., Paparella, M., Ramsingh, D., Reaves, E., Ridgway, P., Schenk, E., Stachiw, L., Ulbrich, B., Hakkert, B.C., (2011). Combined retrospective analyses of 498 rat multi-generation reproductive toxicity studies: on the impact of parameters related to F1 mating and F2 offspring. *Reprod. Toxicol.* 31: 392–401.

Piersma, A.H., Ezendam, J., Luijten, M., Muller, J.J.A., Rorije, E. Van der Ven, L.T.M.,Van Benthem, J. (2014). A critical appraisal of the process of regulatory implementation of novel *in vivo* and *in vitro* methods for chemical hazard and risk assessment. *Crit Rev Toxicol.* 44 (10): 876–894.

Pijnappel, M.C. (2016). *Lost in Technification: uncovering the latent clash of societal values in Dutch public policy discourse on animal-testing alternatives.* (PhD thesis). Radboud University, Nijmegen.

Pistorius, C. and Utterback, J. (1997). Multi-Mode Interaction Among Technologies. *Research Policy.* 26 (1): 67–84.

Pound, P., Ebrahim, S., Sandercock, P., Bracken, M.B., Roberts, I. (2004). Reviewing Animal Trials Systematically (RATS) Group. Where is the evidence that animal research benefits humans? *BMJ.* 328: 514–7.

Pound, P. and Bracken, M.B. (2014). Is animal research sufficiently evidence based to be a cornerstone of biomedical research? *BMJ* 348. doi: http://dx.doi.org/10.1136/bmj.g3387

Power, M.K, 'The New Risk Management', inaugural lecture of P. D. Leake, Professor of Accounting and Director of CARR, December 1999.

Pressman, J. L., & Wildavsky, A. (1973). *Implementation. How great expectations in Washington are dashed in Oakland.* Berkeley, CA: University of California Press.

Punt, A., Schiffelers, M.J.W.A., Horbach, J., van de Sandt, J.J.M., Groothuis, G.G.M., Rietjens, I.M.C.M., Blaauboer, B.J., (2011). Evaluation of research activities andresearch needs to increase the impact and applicability of alternative testing strategies in risk assessment practice. *Regul. Toxicol. Pharmacol*. 61: 105–114.

Raven, R.P.J.M. (2005). *Strategic niche management for biomass.* (PhD thesis) Eindhoven, The Netherlands: University of Technology.

Richmond, J., (2002). Refinement, reduction, and replacement of animal use for regulatory testing: future improvements and implementation within the regulatory framework. *ILAR J.* 43 (Suppl. 1): S63–S68 http://ilarjournal.oxfordjournals.org/content/43/Suppl_1/S63.full.

Rip, A. and Kemp, R. (1996). 'Towards a Theory of Sociotechnical Change', mimeo University of Twente, report prepared for Batelle Pacific Northwest Laboratories, Washington, D.C. An edited version has been published as book chapter, 'Technological Change' In: Rayner, S.and Malone, E.L. (Eds.) (1998), *Human Choice and Climate Change. An International Assessment* (327-400, Vol. 2). Washington D.C., USA: Battelle Press.

Rip, A. and Kemp, R. (1998). Technological change. In: Rayner, S. and Malone, E.L. (Eds.), *Human Choice and Climate Change.* Vol. II, Resources and technology. Battelle Press, Columbus, OH, pp. 327-399. ISBN 9781574770469 Avaliable at: http://doc.utwente.nl/34706/

Rittel, H.W.J.,and Webber, M.M. (1973). Dilemmas in a General Theory of Planning. *Policy Sciences* 4: 155-169.

Rogers, E. (2003). *The Diffusion of Innovations*. Fifth Edition. The Free Press, New York.

Romberg, J., Lang, S., Balks, E. et al. (2012). Potency testing of veterinary vaccines: The way from *in vivo* to *in vitro*. *Biologicals* 40: 100-106. http://dx.doi.org/10.1016/j.biologicals.2011.10.004

Rorije, E., Muller, A., Beekhuijzen, M.E.W., Hass, U., Heinrich-Hirsch, B., Paparella, M., Schenk, E., Ulbrich, B., Hakkert, B.C., Piersma, A.H. (2011). On the impact of second generation mating and offspring in multi-generation reproductive toxicity studies on classification and labelling of substances in Europe. *Regul.Toxicol. Pharmacol*. 61 (2): 251–260.

Rotmans, J., Kemp, R., & van Asselt, M. (2001). *Transitions & transition management. The case for a low emission energy supply.* ICIS working paper: I01-E001. Maastricht

Rouach, D. (2003). Technology transfer and management, *Special feature: new technology transfer practices.* 21-27.

Rovida, C., Alépée, N., Api, A.M. et al. (2014). Integrated Testing Strategies (ITS) for Safety Assessment. T4 Workshop report. *Altex* 32 (1): 25-40.

Rudacille, D. (1999). Development of alternatives to animal use for safety testing and hazard assessment. http://www.solutions-site.org/node/84

Ruden, C. and Hansson, S.O. (2008). Letter to the editor. Comment on: ''A retrospective analysis of the two-generation study: what is the added value of the second generation?'' by G. Janer, B.C. Hakkert, W. Slob, T. Vermeire, A.H. Piersma (*Reprod. Toxicol*. 2007. 24: 97-102.) *Reprod. Toxicol.* 25: 397–405.

Russell, W. and Burch, R. (1959). *The Principles of Humane Experimental Technique*. London, UK: Methuen & Co.

Sabatier, P.A. (1988). An advocacy coalition framework of policy change and the role of policy-oriented learning therein, *Policy Sci*. 21 (2–3): 129–168.

Schaafsma, G., Kroese, E.D., Tielemans, E.L.J.P., Van de Sandt, J.J.M., Van Leeuwen, C.J. (2009). REACH, non-testing approaches and the urgent need for a change in mind set. *Regul. Toxicol. Pharmacol.* 53: 70–80.

Scheel, J., and Brekelmans, C. (2007). Implementation of the 3Rs in European regulation – activities of working group 4 of the European partnership for alternative approaches to animal testing I. Impact of liability issues and the precautionary principle, II, Evaluation of statistical reporting for measuring the uptake of 3Rs in regulatory testing, *AATEX* 14, Special Issue: 775–778 http://altweb.jhsph.edu/wc6/paper775.pdf

Schiffelers, M. J. W. A., Hagelstein, G., Harreman, A., and van der Spek, M. (2005). Regulatory animal testing, Report number: P-UB-2005-10, August 2005.

Schiffelers, M. J. W. A., Blaauboer, B. J., Fentener van Vlissingen, J. M. et al. (2007). Factors stimulating or obstructing the implementation of the 3Rs in the regulatory process. *ALTEX* 24: 271-278. Available at: http://www.altex.ch/resources/n4_07S271278Schiffelers6.pdf

Schiffelers, M.J.W.A., Blaauboer, B.J., Hendriksen, C.F.M., Bakker, W.E., (2012). Regulatory acceptance and use of 3R models: a multilevel perspective. ALTEX 29 (3): 287–300. Available at: http://www.altex.ch/All-issues/Issue.50.html?iid=133&aid=4/

Schiffelers, M. J., Blaauboer, B., Bakker, W. and Hendriksen, C. (2014a). Replacing the NIH test for rabies vaccine potency testing: A synopsis of drivers and barriers. *Biologicals* 42: 205-217. Available at: http://dx.doi.org/10.1016/j.biologicals.2014.04.001

Schiffelers, M. J., Blaauboer, B., Bakker, W. et al. (2014b). Regulatory acceptance and use of 3R models for pharmaceuticals and chemicals: Expert opinions on the state of affairs and the way forward. *Regul Toxicol Pharmacol* 69: 41-48. Available at: http://dx.doi.org/10.1016/j.yrtph.2014.02.007

Schiffelers, M.J.W.A., Blaauboer, B.J, Bakker, W.E., Hendriksen C.F.M. (2015a). Regulatory acceptance & use of serology for inactivated veterinary rabies vaccines. *ALTEX* 32 (3): 211-221 Available at: http://dx.doi.org/10.14573/altex.1501261

Schiffelers, M. J., Blaauboer, B., Bakker, W. et al. (2015b). Regulatory acceptance and use of the extended one generation reproductive toxicity study within Europe. *Regul Toxicol Pharmacol* 71: 114-124. Available at: http://dx.doi.org/10.1016/j.yrtph.2014.10.012

Schmidt M, Raghavan B, Müller V , Vogl T , Fejer G, Tchaptchet S, et al . (2010) . Crucial role for human Toll-like receptor 4 in the development of contact allergy to nickel . *Nat Immunol.* 11: 814-9.

Schneider, S., Kaufmann, W., Strauss, V., van Ravenzwaay, B., (2010). Vincloszolin: a feasibility and sensitivity study of the ILSI–HESI F1-extended one-generation rat reproduction protocol. *Regul. Toxicol. Pharmacol*. 59: 91–100.

Scholtz, S., Sela, E., Blaha, L., et al. (2013). A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. *Regul. Toxicol. Pharmacol.* 67: 506–530.

Schön, D.A. and Rhein, M. (1994). *Frame Reflection Toward the Resolution of Intractable Policy Controversies.* New York: Basic Books.

Schot, J., Hoogma, R. and Elzen, B. (1994). Strategies for shifting technological systems. The case of the automobile system. *Futures* 26 (10): 1060–76.

Schot, J. and Rip, A. (1996). The past and future of constructive technology assessment. *Technol Forecast Soc Change* 54, 251-268. http://dx.doi.org/10.1016/S0040-1625(96)00180-1

Schot, J. and Geels, F. (2008). Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. *Technology Analysis & Strategic Management* 20 (5): 537–554.

Scott, W. R. and Meyer, J.W. (1983). The organization of societal sectors. In: Meyer, J.W. and Scott, W.R. (Eds). *Organizational Environments: Ritual and Rationality* (129-53), Beverly Hills, CA: Sage.

Scott, W. R. (2004). Institutional Theory: Contributing to a Theoretical Research Program. In: Smith, K. G. and Hitt, M.A (eds.) *Great Minds in Management: The Process of Theory Development.* Oxford UK: Oxford University Press. https://www.researchgate.net/profile/W_Scott/publication/265348080_Institutional_Theory_Contributing_to_a_Theoretical_Research_Program/links/54de42450cf2966637857c60.pdf

Schumpeter, J. (1942). *Capitalism, Socialism and Democracy.* New York: Harper and Row.

Seidle, T., Robinson, S., Holmes, T., Creton, S., Prieto, P., Scheel, J. (2010). Cross-sector review of drivers and available 3Rs approaches for acute systemic toxicity testing. *Toxicol. Sci*. 116: 382–396.

Slovic, P., Fischhoff, B., and Lichtenstein, S. (1984). Behavioural decision theory perspectives on risk and safety. *Acta Psychol.* 56: 183-203.

Smith, J., Yager, P. and Baer, G. (1996). A rapid tissue culture test for determining rabies neutralizing antibody. In Kaplan, M. and Koprowski, H. (Eds.), *Laboratory Techniques in Rabies* (354-357). 4th edition. Geneva, Switzerland: WHO.

Smith, A. (2003). Transforming technological regimes for sustainable development: a role for appropriate technology niches? *Science and Public Policy* 30 (2): 127–135.

Spielmann, H. (2000). Would Sisyphus meet the challenges of validation from test development to global regulatory acceptance, In: Balls, M., van Zeller, A. M, and Halder M. (Eds.*) Progress in the Reduction Refinement and Replacement of Animal Experimentation.* New York, USA: Elsevier.

Spielmann, H. and Vogel, R. (2006). Editorial: REACH testing requirements must not be driven by reproductive toxicity testing in animals. *ATLA* 34: 365–366.

Spielmann, H., (2009). The way forward in reproductive/developmental toxicity testing. *ATLA* 37: 1–16.

Stephens, M.L. and Mak, N.S. (2014). History of the 3Rs in Toxicity Testing: From Russell and Burch to 21st Century Toxicology. In: Allen, D. & Waters, M.D. (Eds.) *Reducing, Refining and Replacing the Use of Animals in Toxicity Testing.* RSC Publishing.

Stokes, W., Kulpa-Eddy, J., McFarland, R., et al. (2011). Introduction and summary of the international workshop on alternative methods to reduce, refine, and replace the use of animals in vaccine potency and safety testing: State of the science and future directions. *Proc Vaccinol.* 5: 1-15. http://dx.doi.org/10.1016/j.provac.2011.10.001

Stokes, W., McFarland, R., Kulpa-Eddy, J., et al. (2012). Report on the international workshop on alternative methods for human and veterinary rabies vaccine testing: State of the science and planning the way forward. *Biologicals* 40: 369-381. http://dx.doi.org/10.1016/j.biologicals.2012.07.005

Stone , D. (2002). *Policy Paradox: The Art of Political Decision Making.* New York: Norton

Storer, R.D., Sistare, F.D., Reddy, M.V., DeGeorge, J.J., (2010). Toxicological pathology an industry perspective on the utility of short-term carcinogenicity testing in transgenic mice in pharmaceutical development. *Toxicol. Pathol*. 38, 51–61. http://intl-tpx.sagepub.com/content/38/1/51.full.pdf+html/

Swanborn, P.G. (2003). *Case-study's, Wat, wanneer en hoe?* Amsterdam – Meppel: Boom.

Tansey, O. (2007). Process tracing and elite interviewing: A case for non-probability sampling. PS: *Political Science and Politics* 40: 765-772.

Taylor, K., Gordon, N., Langley, G., Higgins, W. (2008). Estimates for Worldwide Laboratory Animal Use in 2005. *ATLA* 36: 327–342

Taylor, M.R., Rubin, E.S. and Hounshell, D.A. (2005). Regulation as the Mother of Innovation: The Case of SO2 Control. *Law & policy* 27(2): 348-378.

Tijmstra, J. & Boeije, H.R. (2011). *Wetenschapsfilosofie in de context van de sociale wetenschappen.* Amsterdam: Boom/Lemma.

Tsebelis, G. (2002). *Veto Players: How Political Institutions Work.* Princeton, New Jersey: Princeton University Press.

Tushman, M. L., and Nadler. D. (1986). Organizing for innovation. *California Management Review* 28 (3): 74–92.

Vandebriel, R. J. and Opperhuizen, A. (2011). Knelpunten bij de ontwikkeling, validatie en implementatie van Alternatieven voor Dierproeven. RIVM Rapport 340720005/2011

Van den Ende, J. and Kemp, R. (1999). Technological transformations in history: how the computer regime grew out of existing computing regimes. *Research Policy* 28: 833–851

Van den Berg, B. (2011). Alternatieven voor dierproeven in de geneesmiddelensector. Een inventarisatie van belemmeringen bij het implementeren van alternatieve methoden voor dierproeven in de geneesmiddelensector en formulering van aanbevelingen voor toekomstig beleid. Onderzoeksrapport in het kader van de master: Management, Policy-Analysis & Entrepreneurship in Health and Life Sciences, Vrije Universiteit Amsterdam.

Van der Jagt, K., Munn, S., Tørsløv, J., de Bruijn, J., (2004). Alternative Approaches can Reduce the Use of Animals Under REACH. http://publications.jrc.ec.europa.eu/repository/bitstream/JRC29111/EUR%2021405%20EN.pdf

Van der Panne, G. Van Beers, C., Kleinknecht, A. (2003). Success and failure of innovation: a literature review. *International Journal of Innovation Management* 7 (3): 1–30.

Van de Ven, A., Angle, H. L. and Poole M. S. (1989). *Research on the management of innovation: the Minnesota studies*. New York: Harper & Row.

Van Meer, P.J.K. (2013). *The Scientific Value of Non-Clinical Animal Studies in Drug Development.* (PhD Thesis). Utrecht University.

Van Thiel, S. (2014). *Research Methods in Public Administration and Public Management: An Introduction.* Oxford, UK: Routledge.

Vermeulen, P. (2011). *De verankerende organisatie: Een institutioneel perspectief op veranderen en vernieuwen.* Den Haag, The Netherlands: Boom Lemma Publishers.

Verwer, C. M., van der Ven, L. T. M., van den Bos, R., and Hendriksen, C. F. M. (2007). Effects of housing conditions on experimental outcome in a reproduction toxicity study. *Regul Toxicol Pharmacol.* 48: 184-193.

Wagner, K., Fach, B. and Kolar, R. (2012). Inconsistencies in Data Requirements of EU Legislation Involving Tests on Animals. *Altex* 29: 302-332.

Wallace, D. (1995). *Environmental Policy and Industrial Innovation. Strategies in Europe, the U.S., and Japan*. London: Earthscan.

Wallace, Ruth A. and Alison Wolf. (1999). *Contemporary Sociological Theory*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Walter, A., Praveen Parboteeah, K., Riesenhuber, F. and Hoegl, M. (2011). Championship Behaviors and Innovations Success: An Empirical Investigation of University Spin-Offs. *J Prod Innov Manag*, 28: 586–598

WHO (2007). WHO Technical Report Series No 941, Annex 2: Recommendations for inactivated rabies vaccine for human use produced in cell substrates and embryonated eggs. Available at: http://whqlibdoc.who.int/trs/WHO_TRS_941.pdf?ua=1

Williams, P. (2002). The competent boundary spanner, *Public Administration* 80: 103-124.

Wunderli, P., Dreesen, D., Miller, T. and Baer, G. (2006). The rabies peripheral challenge test: More accurate determination of vaccine potency. *Vaccine* 24: 7115-7123. http://dx.doi.org/10.1016/j.vaccine.2006.06.078

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Yin, R. K. (2008). *Case Study Research: Design and Methods* (4th ed). Thousand Oaks, CA: Sage,

York, M., and Steiling, W. (1998). A critical review of the assessment of eye-irritation potential using the Draize rabbit eye test. *J. Appl. Toxicol*. 18; 233–240.

Zucker, L. (Ed.) (1988). *Institutional patterns and culture.* Cambridge, MA: Ballinger Publishing Company.

SUMMARY

1. Introducing the problem of 3R non-acceptance in the regulatory domain

Annually, approximately 11.5 million laboratory animals are used in Europe for a variety of purposes (EC, 2013). About 25% of these animals are used to meet the regulatory requirements for the assessment of substances and products on safety and/or efficacy. Regulatory animal testing raises concerns for scientific, ethical and economic reasons. Most of these animal models were developed many decades ago and have never been formally validated (Spielmann, 2000). Numerous tests produce variable results (e.g. Bruckner et al., 2003) and/or face extrapolation problems due to profound differences between the laboratory animal and the target animal species (e.g. Piersma et al., 2014, Martić-Kehl, Schibli & August, 2012). As a result there is increasing doubt about the scientific value of animal models for human application (Pound and Bracken, 2014, Van Meer, 2013). Moreover, regulatory animal testing raises ethical concerns and the procedures frequently cause serious pain and distress to the animals involved. Russell and Burch, in 1959, introduced the 3Rs principle to 'replace, reduce and refine' animal models. The 3R approaches have the potential of combining better science with fewer (or even no) animals, less animal suffering and faster and cheaper test results. A broad range of 3R models is now available and their use is being instigated through Directive 2010/63/EU on the protection of animals used for scientific purposes (EU, 2010). However, the acceptance and use of 3R models in the regulatory domain is a highly challenging process. This thesis provides an in-depth analysis to understand why it is a challenging process and examines possibilities to enhance 3R acceptance and use in the regulatory domain. As described in Chapter 1 this study addresses the following questions:

Q1. How can regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes be defined?

Q2. Which theoretical perspectives are needed to comprehend and enhance this process?

Q3a. Which factors influence the regulatory acceptance and use of 3R models?

Q3b. How do these factors influence the regulatory acceptance and use of 3R models ?

Q4. How can the process of regulatory acceptance and use of 3R models be optimized?

Regulatory animal testing is a wicked problem. It crosses geographical, institutional and sectorial borders, and involves many different stakeholders – both public and private – often with diverging perspectives. Furthermore, it is characterized by a multilevel playing field and a highly risk-averse context (Schiffelers et al., 2007). To deal with this complex problem it is important to obtain an in-depth understanding of the process through investigating examples of regulatory – non - acceptance.

Where most studies in this field adopt a technical perspective (i.e. by scrutinizing the technical aspects of a 3R model), this thesis uses a multidisciplinary and integrative approach to analyze the process of acceptance and use of 3R models in the regulatory domain and its underlying dynamics. A Technology Transition (TT) perspective is adopted to analyze the broad range of variables that are observed to influence this process. The initial study conducted by Schiffelers et al. in 2007 summarized general

categories of influencing factors. However, it was concluded that these variables are likely to differ per sector and even per case. No sector/case- specific information was obtained at that stage. This thesis provides such case- and sector-specific information, focusing on the pharmaceutical (including vaccines) and chemical sectors within Europe. It thereby contributes to a better understanding of this process and provides ways to overcome existing barriers.

2. Defining 3R acceptance for regulatory purposes

Chapter 2 addresses Q1: *How can regulatory acceptance and use of 3R models be defined?* Regulatory testing refers to risk assessment of substances and to safety and efficacy testing of products for human or animal use. Many of the regulatory tests are based on animal models that were developed in the first half of the 20th century. These tests thereby became, and currently still are, part of the requirements to which these products are subjected. Russell and Burch's publication (1959) was a starting point in terms of the 3Rs principle. Particularly from the 1970ies on the recognition for 3R approaches increased and many 3R models have been developed ever since. Regulatory requirements specify which tests need to be conducted to assess a certain product or substance. Most requirements offer discretionary space to choose the most suitable test options and to introduce alternative ways of testing and Directive 2010/63/EU on the protection of laboratory animals requires the use of 3R models where available. This means that both regulatory authorities and manufacturers have the possibility, and are encouraged by the directive, to make use of this discretionary space in favor of 3R models, i.e. to accept and use 3R models for regulatory purposes. To analyze the process of regulatory acceptance and use it is divided into three substages:

1. The Formal Incorporation of 3R models into regulatory product requirements (FI);
2. The Actual Regulatory Acceptance of these models by regulatory authorities[1] (ARA) for licensing, safety assessment and quality control purposes of regulated substances; and
3. The Use by Industry of these models to meet regulatory product requirements (UI).
4. The acceptance and use of 3R models for regulatory purposes proves to be a challenging process in which the innovations have to compete with institutionalized animal procedures, also referred to as the 'gold standard'. The way this process is examined, is described in the following section.

3. Research approach

This thesis uses a social science perspective to examine the developments in a field that is driven by the natural sciences. Earlier research (Schiffelers et al., 2007) revealed that a broad perspective is needed to understand the phenomenon of the slow regulatory acceptance of 3R models. The 2007 study also concluded that the process of regulatory acceptance and use of 3R models needs to be examined in the specific context in which it is situated. The assumptions about the nature of reality and the sources of knowledge (ontology), which are at the basis of this research, including the methodological

1 Licensing organizations, standard setting bodies and product assessors.

choices (epistemology), are a reflection of Bent Flyvbjerg's philosophy described in his book *Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again* (Flyvbjerg, 2001). Flyvbjerg underlines the relevance of the specific context in which a certain phenomenon takes place. Due to the required contextuality, the social sciences will not succeed in producing general, predictive, context-independent theories. The value of social sciences lies in the aspect of *phronesis,* balancing instrumental rationality by value rationality and ensuring that scientific and technical developments do not take place without ethical checks and balances. This requires narrative inquiries that describe and interpret the central problem from the perspective of involved stakeholder groups (social constructionism). It entails the analyses and interpretation of the values and interests as a point of departure for managed action with the goal of social change. Well managed action requires contextualism (Flyvbjerg, 2012). This thesis therefore embraces the 'power of example' and makes use of a research design that allows for nuances to exist. To collect detailed information about the process of regulatory acceptance and use, while taking its destined context into account a case study approach was used examining two concrete examples of 3R acceptance in the area of vaccines (pharmaceuticals, Chapter 6) and in the area of (industrial) chemicals (Chapter 7). Subsequently, the thesis zooms out to put the findings into the broader perspective of the two sectors of pharmaceuticals -including vaccines - and chemicals (Chapter 8). Through two expert panels the broader usability of the findings was tested and optimizing options were collected. The model of causal mechanisms used for analyzing the empirical findings was offered by the multilevel perspective of technology transitions. It enables the analysis of the multi-causality and interdependencies between the forces at hand (see Chapter 9) and offers potential strategies to enhance the current process (see Chapter 10).

4. Technology transition perspective to analyze regulatory acceptance of 3Rs

Chapter 4 addresses Q2: Which theoretical perspectives are needed to comprehend and enhance the process of regulatory acceptance and use of 3R models? To deal with the wicked/multi-faceted problem of regulatory acceptance and use of 3R models the integrative/multidisciplinary approach of Technology Transitions was adopted. TT refers to major technological transformations in the way societal functions (e.g. communication, transportation, energy supply etc.) are fulfilled. The risk assessment of chemicals and quality control of pharmaceuticals is such a social function. In the TT approach, technology can only fulfill its social function and become meaningful if it is able to associate with human action, social structures and connected organizations. The multilevel perspective on TT offers the analytical frame to examine the barriers and drivers influencing regulatory acceptance and use. This perspective defines three levels entailing specific features that enable - or withhold - innovation breakthroughs: innovations are developed and tested in niches (the micro level). These niches are embedded in and influenced by the meso level of sociotechnical regimes which covers the existing ways of operating which are fixed in institutions, rules and regulations. The meso level on its turn is embedded in the macro level of the societal landscape, in which broader cultural, demographical, technological, political and economic developments deter-

mine the developments at the meso and the micro level. The value of the multilevel perspective is that it goes beyond the technological aspects of innovations and also includes required changes in elements such as user practices, regulations, infrastructure, and social beliefs. From TT literature, it becomes clear that established regimes or technological pathways have a serious amount of excluding power caused by a broad series of factors at the macro, meso and micro levels. All kinds of values are attributed to the existing technologies and alternative technologies often do not match with these expectations. Nonetheless, regimes are frequently observed to change and the following success factors and mechanisms of change can be found in literature:

- Champions and innovation entrepreneurs that are strongly connected to an innovation are of critical importance in technology transfers;
- Collaborations between advocates of the new technology are important to learn and share experiences. Public private partnerships offer the possibility of finding and attracting the right participants;
- A clear technology-transfer strategy directs the formulation of mutual goals and required activities;
- Pilot projects and demonstrations allow hands-on learning and the early involvement of the end-user is key to allow early resolution of problems and to prepare the user for the innovation.

Tensions in the sociotechnical regime in terms of criticism about the existing technology, changing user practices, obsolete test infrastructure, policy changes and changing scientific knowledge, are also potential drivers and may lead to periods in which the existing regimes are weakened and new technologies become more important. Technological change depends on the outcome of the balancing act of the drivers and barriers.

5. The multilevel perspective as the basis for the 3R acceptance model

The multilevel perspective on Technology Transitions was used to develop the 3R acceptance model which is presented in Chapter 5. This 3R acceptance model serves several goals. Firstly, it enables the categorization of the broad range of drivers and barriers after the level at which they are observed to exert influence (i.e. the micro, meso or macro level). Secondly, it facilitates the analysis of the dynamics between these variables both within and between the three levels. Thirdly, it assists in specifying the strong and the pliable variables, in the sense that the factors at the micro and partly the meso level tend to offer more possibilities for change than the broader societal developments at the macro level. TT is almost always *"the result of the interplay between many factors and actors"* (Geels, 2006). Successful technology TT requires alignment of the developments at these three levels. An aggregation of the developments can only occur if an innovation (e.g., a 3R model) meets the needs of the meso- and the macro level. Through the 3R acceptance model the remaining parts of the puzzle to answer Q2 are offered. In addition, Chapter 5 adopts a risk regulation perspective to tailorize this perspective to the risk averse context in which regulatory testing is situated. The striving for risk-minimization is a dominant feature of the societal context for which 3R models are intended. This striving is operationalized by setting up risk regimes of regulatory authorities, rules, and regula-

tions to minimize the possible adverse effects of products like pharmaceuticals and chemicals. The risk-averse context strongly influences the way product assessors look at new technologies. It entails an intrinsic aversion to uncertainty and thereby to change (Breyer, 1993) which is amplified by public and political pressure to minimize risks the occurrence of incidents. Innovation starts with the willingness to accept failure (MacLachlan, 1994), which in the area of product regulation can have big consequences. Regulatory authorities and industries acknowledge that there often is a lack of trust to take this 'leap of faith'. For this reason, it must be accepted that such institutional changes take time and that regulatory acceptance of 3R methods is most likely to occur as an incremental process, i.e., no change in terms of radical developments, but in terms of new test regimes that gradually grow out of old ones (Geels, 2002).

6. The case of veterinary rabies vaccine potency testing (SNT case)

Chapter 6 clarifies the process of acceptance and use of the mouse antibody serum neutralization test (SNT) and thereby addresses the research question Q3a and Q3b: *Which factors influence the regulatory acceptance and use of 3R models?* and *How do these factors influence the regulatory acceptance and use of 3R models?* The case study reconstructs the process and reveals the barriers and drivers that have played a role in it, as observed by involved stakeholders. The findings derive from literature, expert interviews, attendance to a series of international meetings and from a survey conducted in anticipation of this case study (Appendix II). The SNT was designed to replace the highly variable NIH mouse rabies challenge assay and was formally incorporated into European Pharmacopoeia monograph 0451 for potency testing of inactivated veterinary rabies vaccines in April 2013. The Formal Incorporation (FI) into the monographs of the European Pharmacopoeia can be defined as a success in that the SNT relatively quickly escaped from the niche in which it was developed and (pre-)validated. (Pre-)validation through a collaborative study was key in proving that the alternative method works in the hands of the participants. The process from test development to FI was facilitated by strong and committed innovation entrepreneurs and a clear problem ownership and process management of institutional players.[2] Additionally, the intense collaboration within the OMCL network, the early involvement of a statistician to design the study and analyze the data during the validation process and the wide dissemination of the study results added to the swift incorporation of the SNT into monograph 0451. The fact that the test was developed by the Paul-Ehrlich-Institut in Germany, one of the leading European OMCLs when it comes to rabies vaccines, facilitated its Actual Regulatory Acceptance (ARA) within the European context. Moreover, the legislative context of the European Pharmacopoeia, which encourages the acceptance and use of 3R models, also stimulated this process. However, the swiftness also came with a price. The persuasiveness of the initiators/project-coordinators was a big driver for the adoption of the SNT into the monographs of the European Pharmacopoeia, but partly

2 First the German Official Medicines Control Laboratory (OMCL) of the the Paul-Ehrlich-Institut, then the Biological Standardization Program (BSP) of the European Directorate for the Quality of Medicines (EDQM), then the European Pharmacopoeia Commission and Group 15V on veterinary sera and vaccines, responsible for the evaluation and approval of veterinary vaccine monographs.

led to restrained attention to the drawbacks of the SNT, as perceived by manufacturers and regulators (such as the US Department of Agriculture - USDA -, responsible for the regulation of rabies vaccines). According to the USDA, the SNT causes concern because it does not offer the required information on the amount of antibodies induced by the vaccine. Both the USDA and several manufactures questioned the value of the SNT and instead would like to invest in *in vitro* methods, which most of the manufacturers already use for in-process control purposes. In other words, the drivers have been able to outweigh the barriers in the substages of FI and ARA in the European situation. However, for the broader regulatory acceptance the perceived drawbacks of the SNT are seen to withhold progress.

The following lessons can be drawn from the SNT case for similar future processes. First of all, the case study shows that there must be a firm commitment of the key stakeholders to allocate time and money to take part in such a project, exchange method details and testing material and adhere to the specific rules of the collaborative study. The strong commitment was fed by legal texts of both the Council of Europe[3] and the European Commission[4], stimulating the use of 3R models. Secondly, effective interaction between central stakeholders within the regulatory framework, for which the 3R model is destined, proves to be essential. Thirdly, the SNT process shows the importance of a well-designed and coordinated validation process in which mutual trust in the new method can be build. Fourthly, the validation process requires a strict process management with predefined steps to be taken, questions to be answered and goals to be reached. However more attention should have been granted to the hesitations of potential end-users to prevent resistance to the innovation at the final stage.

7. The case of reproductive toxicity testing (EOGRTS case)

Chapter 7 addresses the case study of the acceptance and use of the Extended One Generation Reproductive Toxicity Study (EOGRTS) to replace the two-generation study (OECD TG 416)[5] for risk assessment purposes of industrial chemicals in the context of REACH.[6] As the SNT case study, it defines the drivers and barriers influencing the process at the three substages of regulatory acceptance and use and thereby also addresses the research questions Q3a and Q3b (see section 1 of the summary). The findings derive from document analysis and expert interviews. OECD TG 416 (2001) is the standard requirement within REACH to test reproductive toxicity effects of chemicals with production volumes >100 tonnes. It is criticized for scientific and ethical reasons. TG 416 is estimated to use nearly 40% of the laboratory animals under REACH (Janer et al., 2007a) and thereby is one of the major users of rodents in safety-test programs. In anticipation of the introduction of REACH, reproductive and developmental toxicity were even estimated to become the largest animal user for safety testing within REACH

3 Treaty 123: European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes

4 Directive 201/63/EC on the protection of animals used for scientific purposes

5 The Organisation for Economic Co-operation and Development Test Guideline

6 European Directive for the Registration, Evaluation, Authorization and restriction of CHemicals – REACH (EU, 2006)

(Pedersen et al., 2003; Van der Jagt et al., 2004). REACH states in article 25 (1) that in order *". . .to avoid unnecessary animal testing, testing on vertebrate animals for the purpose of this Regulation shall be undertaken only as a last resort."* (EU, 2006). Reproductive toxicity testing thereby became a serious point of concern. The EOGRTS, incorporated into the OECD test guidelines in 2011 (OECD TG 443), has the potential to replace TG 416 while using only one generation of rats (with a reduction of up to 40% in animal use – i.e. a total of 1200 animals per study –) and being more informative. However, its regulatory acceptance in the context of REACH proved challenging.

The stage of Formal Incorporation (FI) of the EOGRTS into OECD test guidelines went swiftly. It was stimulated by a series of retrospective analyses on the value of the second generation of research animals[7], by strong EOGRTS advocates[8] that were connected to the existing regulatory regime, by animal welfare concern and by changing US and EU chemicals legislation which led to the search for alternative ways of testing. However, the FI did not go without effort. There was a series of tough discussion points that had to be settled, such as the disagreement on the added value of the second generation and whether or not to incorporate parameters for developmental neurotoxicity (DNT) and developmental immunotoxicity (DIT) in the standard EOGRTS protocol. These discussions were not fully settled at the stage of FI and were transferred to the stage of Actual Regulatory Acceptance (ARA) within REACH. The process at this stage was furthermore challenged by the fact that TG 416 is the required method in the Test Methods Regulation (TMR) of REACH, which was stipulated by those in doubt about the EOGRTS. The advocates of the EOGRTS on the other hand stated that regulatory authorities and manufacturers have the obligation to choose 3R approaches where suitable and available. This legal discussion went hand in hand with an ongoing scientific dispute on the added value of the second generation and the adoption of additional parameters. At the stage of Use by Industry (UI) this led to uncertainty among registrants about the acceptance of the EOGRTS for the safety assessment of their chemical substances. The uncertainty was intensified by the high costs of the procedure, the risk of false positives due to the additional parameters that chemicals are tested for, and the complexity and thereby manageability of the EOGRTS.

The drivers at the micro level proved strong enough for the EOGRTS to become accepted at the OECD level. In the context of REACH however, these drivers could not convince all parties. Two frames (lines of argumentation) were found to dominate the discourse on the acceptance and use of the EOGRTS, i.e. the line of precaution and the line of innovation. The advocates of the EOGRTS focused on the advantages of the innovation while the advocates of TG 416 underlined the uncertainties connected to the EOGRTS and leaving out the second generation which resulted in a controversy between these two frames. New scientific data were not able to bridge the gap and even fueled the disagreement. The parties at a certain stage *"agreed to disagree"* and since consensus is needed within ECHA's[9] Member States Committee, the EOGRTS bounced back to the micro level where its suitability was discussed anew.

7 Janer et al., 2007a; Martin et al., 2009; Piersma et al., 2011; Rorije et al., 2011

8 Starting with the US, Germany and the Netherlands submitting a proposal to the OECD secretariat to draft an OECD test guideline (TG) based on the EOGRTS

9 European Chemicals Agency

From the EOGRTS case, the following lessons can be drawn. It reveals that profuse scientific information, strong advocates and critical junctures (e.g. changing legislation) offer important ingredients for the regulatory acceptance of a 3R model. However, it also reveals that regulatory acceptance is often a highly politicized process in which science can become part of the existing disagreement and in which other arguments (e.g. the lack of experience with and trust in the new model, institutional agendas and political realities) are seen to outweigh the scientific 'facts'. These 'other' arguments therefore continuously have to be taken into consideration. This requires ongoing communication about and anticipation on potential legal, practical and psychological drivers and barriers. In addition, stakeholders need to be aware of the fact that the three substages of FI, ARA and UI are strongly connected and the heritage of a previous stage is likely to remain of influence at the subsequent stage.

8. Expert opinions on 3R regulatory acceptance for pharmaceuticals and chemicals

In Chapter 8 the barriers and drivers in the pharmaceutical (including vaccines) and chemical sectors are described and compared. Moreover, options to optimize this acceptance process are given. This chapter thereby not only targets the research questions Q3a and Q3b but also offers a preliminary answer to Q4 *How can the process of regulatory acceptance and use of 3R models for risk assessment and efficacy testing purposes be optimized?* The questions were addressed in two expert panels, one with 20 experts from the pharmaceutical field and one with 20 chemical experts.[10] The panel results revealed that there are many similar mechanisms within both sectors that prevent 3R models from becoming accepted. A shared barrier at the micro level is the uncertainty connected to the new 3R models. This uncertainty is mainly connected to the undefined predictability of 3R models and the challenging *in vitro–in vivo* extrapolation. Both aspects refer to the challenges of translating test results of 3R models (especially *in vitro* models) to the biological effects in humans or target-animals. At the meso level, the lack of global harmonization of regulatory requirements was identified as a central barrier, as was the lack of clear acceptance criteria when using a 3R model. The macro level was perceived to be largely shaped by the high levels of risk aversion. A shared driver at the micro level is the scientific value of 3R models. 3R models are valued for their informative character about the mechanism of action, their reproducibility and their scientific robustness. At the meso level, the driving force of horizontal legislation, such as European Directive 2010/63/EU on the protection of animals used for scientific purposes, is brought forward. The ambition to decrease the dependence on animal tests is reflected in vertical regulatory requirements such as the monographs of the European Pharmacopoeia and REACH. A central driver at the macro level is the ethical concern of society for laboratory animals. This factor is referred to as a catalyst for the process of regulatory acceptance and use.

10 Firstly, the panel members were asked to make an individual inventory of barriers and drivers. The factors were clustered around themes and in terms of exerting a driving or obstructing influence on the process of regulatory acceptance and use. Secondly, a prioritization was made in plenary in terms of the perceived dominance of the drivers and barriers on the process of regulatory acceptance and use. Thirdly, actions were identified that can be pursued by the stakeholder groups to optimize the process of regulatory acceptance.

Subsequently, the following sectorial differences were identified. Safety assessment of pharmaceuticals is generally based on a risk-benefit analysis. The level of risk aversion is far higher for chemicals and vaccines than it is for pharmaceuticals in general. This normally leads to a stronger risk avoiding approach in the sector of chemicals and vaccines, which is also reflected in the safety assessment. Regulators of both chemicals and vaccines are observed to strive for a zero risk level. This includes being highly cautious when it comes to changing the existing test regimes, e.g. by accepting a 3R model. Furthermore, the profit margins are bigger for pharmaceuticals, leaving more room to invest in 3R models. Also more knowledge is generated for pharmaceuticals[11] which allow risk assessments based on a more mechanistic mode of action. Such sector specificities are important to take into account when trying to improve the process of regulatory acceptance and use of 3R models.

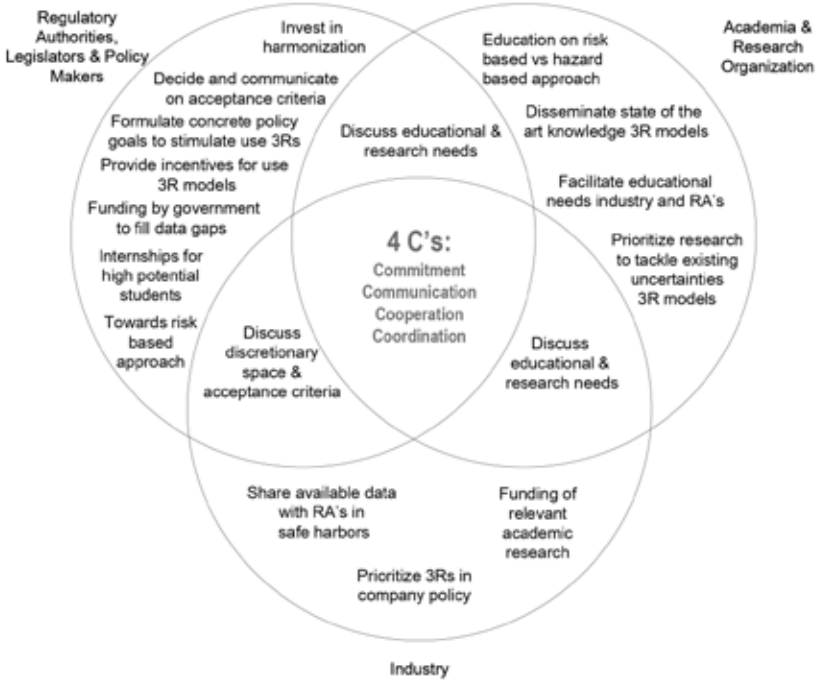


Figure 1. Actions per stakeholder group to facilitate regulatory acceptance & use of 3R models

To enhance the process, the panel experts identified a series of required actions, some of which need to be initiated by one stakeholder group: the unilateral actions. Others actions require a combined action of two or three parties: the bilateral and tripartite actions. An overview of these actions is presented in Figure 1. The 4Cs of Commitment, Communication, Cooperation and Coordination are the central catalyzing forces for these actions (Schiffelers et al., 2014b: see Chapter 8).

11 e.g. pharmacological mode of action, animal and human pharmacokinetic data

9. Conclusions

Chapter 9 offers the concluding analysis of empirical findings from the case studies and the expert panels and thereby answers the research questions Q3a and Q3b. Chapter 9 discloses the following dominant drivers (based on the level of recurrence of drivers and barriers in the separate empirical steps):

- Animal welfare concerns and the striving for innovative approaches (macro level);
- Legislation stimulating the use of 3R's and strong stakeholder interaction (meso level); and
- Drawbacks of animal models, advantages of 3R models, the availability and sharing of data, the presence of innovation entrepreneurs, the early involvement of regulatory authorities and manufacturers and gaining experience with and thereby trust in the new test method (micro level).

The following dominant barriers are identified:

- The striving for risk minimization and precaution (macro level);
- A strict interpretation of product requirements, the high perceived risk of products like vaccines and chemicals, the fear of incidents and releasing unsafe products, the firm institutionalized status of the conventional test, the lack of harmonization of regulatory requirements and unclear acceptance criteria for 3R models (meso level); and
- The perceived limitations of 3R models and the transition costs (micro level).

Apart from the difference between dominant and less dominant factors, a distinction was made between the pliable, short term modifiable factors, and the powerful (explanatory) factors which offer important explanations for the process of non-acceptance but are much harder to influence on the short term (Ellemers, 1976). Preferably, the targeted variables are both dominant and pliable. Some aspects, such as the striving for risk minimization and the lack of harmonization are dominant and hard to modify at the same time. Targeting these factors is a long-term affair. Variables which are observed to be dominant and pliable are the profound and ongoing stakeholder interactions, the availability and sharing of data, the early involvement of statisticians to interpret these data, innovation entrepreneurs who address and keep the issue on the agenda, the early involvement of end users and building experience with the new test method in a niche-based setting. When considering the recurring themes it becomes clear that the concern for animal welfare and innovation frame on the one hand, and the high level of risk aversion and precautionary frame on the other hand are centrally opposing forces at all three levels. Up until now the striving for risk minimization was observed to outweigh the animal welfare concerns in most of the product sectors, apart from the cosmetics sector in which the European Cosmetics Regulation (No 1223/2009) has led to phasing out the use of laboratory animals.

Transitions in both the SNT and the EOGRTS case were realized through the connections made between the developments at the meso level and the micro level, through which the need for replacement of a certain animal model came high on the agenda of institutional players. They initiated strong entrepreneurial activities at the micro level by developing, testing and/or pushing the new technology and by creating a

strong advocacy coalition. Most of the energy in both case studies however was directed at first-order learning, i.e. generating, checking and distributing facts and data. Both the advocates and opponents of the innovation were observed to focus on the generation and interpretation of – partly the same – scientific data to underline their argumentation. Such first-order learning is important, but it entails the risk of overlooking important criteria for innovation later on in the process. Learning processes are known to contribute more if they go beyond the level of gathering data and enable second-order learning, by stimulating changes in cognitive frames and assumptions (values and norms) within these frames (Grin and Van de Graaf, 1996), bringing us to Chapter 10 on optimizing the process.

10. The way forward

As described in the previous chapters, first-order learning and technical discussions are very important, but insufficient to fully guide a 3R model through the three substages of regulatory acceptance and use. The current technocratic perspective overlooks the fact that the introduction and acceptance of new technologies often depends more on social, psychological, cultural and historical factors than on technological merit (NRC, 2004). Scientific data may even widen the gap as they become part of the controversy. Stakeholders that have been involved in processes of regulatory acceptance are well aware that these processes clearly exceed the rules of natural sciences in which the numbers tell the tale. The process of regulatory acceptance and use is to a large extent a political and a psychological process and slow evolvement is often the result of conflicting social constructions formed by deep-rooted values and beliefs, narratives, images and perceptions regarding animal models and alternative ways of testing. These underlying elements are at the basis of the actions of, and the interactions between stakeholders when it comes to 3R –non-acceptance.

Although social constructions aren't easily changed, they can be reconstructed in due time. To contribute to the reconstruction process, this final chapter addresses Q4. *How can the process of regulatory acceptance and use be optimized?*

First of all, it is important to realize that alterations can be initiated at all three levels. Innovation entrepreneurs need to be conscious of potential opportunities at all three levels and of impending critical junctures (e.g. changes in policy or legislation). However, transformational change is not necessarily the result of critical junctures; it is often the result of an incremental process (Capoccia and Kelemen, 2007) in which new technologies physically link up with established technologies. Thinking in terms of evolution rather than revolution (Rotmans et al., 2001) is very useful for two reasons. Firstly, critical junctures are relatively rare events. Secondly, an evolutionary mechanism of technological breakthrough is likely to be a more suitable approach in this risk-averse context where change invokes high levels of uncertainty.

Transformation can be facilitated through various proactive ways of connecting the drivers at the three levels of the multilevel perspective. Transition management aims

at involving a wide range of stakeholders over the multiple levels and at creating a long-term perspective with shared visions and goals to encompass societal values and beliefs. This is in line with the call of experts in the field who stipulate the need for European policy goals to stimulate the use of 3R models. Through strategic niche management the creation and protection of niches is targeted. It aims at building networks, stimulating learning processes, the articulation of expectations, promises and visions and the creation of coalitions with a shared agenda (Boon et al., 2014). The roadmap to change as presented in this chapter combines elements of both strategies and as such combines a bottom-up and a top-down strategy anddistinguishes three tracks: The roadmap distinguishes three tracks:

1. The niche-based track (Figure 2) starts from the innovation and focusses on niche empowerment through niche development and accumulation and connections with the existing regime.

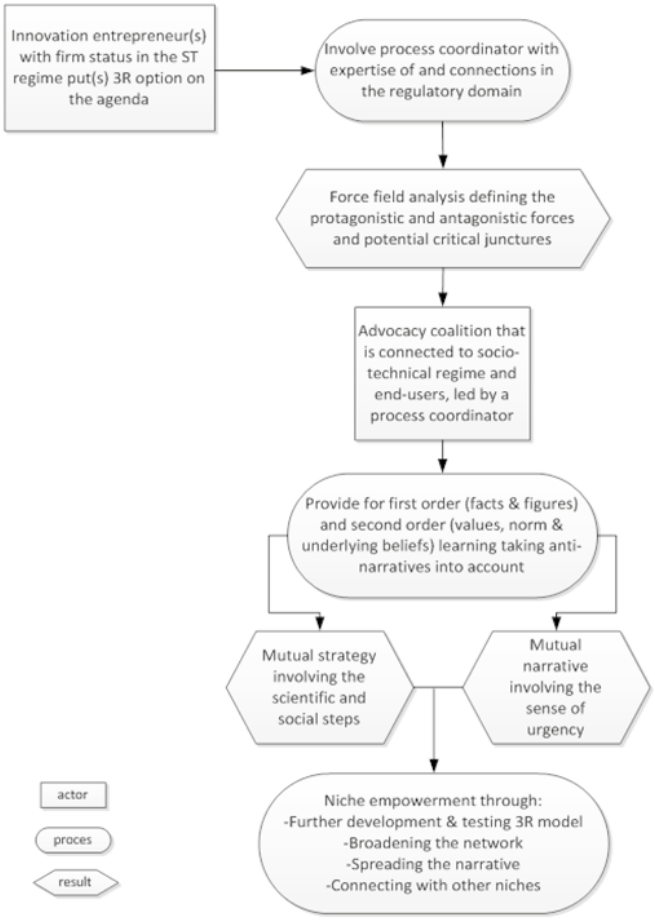


Figure 2. The niche-based track to 3R acceptance and use

2. It begins with entrepreneurial activities in which responsibilities are shared between a core group of involved actors taking the specificities of a particular regulatory context into account and connecting it with that regulatory regime. This advocacy coalition develops a strategy with a clear mutual goal and steps that have to be taken to get there. A force field analysis is made to define the players in the field and their interests. The coalition empowers the niche through the construction of a solid shared narrative including the sense of urgency to switch to the 3R model. Furthermore, it takes in the arguments which form the basis of the anti-narratives and offers discussion options on how to deal with remaining uncertainties. To this end, the innovation entrepreneurs should be accompanied by neutral process coordinators with specific skills to deal with existing disagreements and controversies and who are able to keep all of the stakeholders actively involved.
3. The regime-based track aims at defining a broader shared sense of urgency. Such a sense of urgency in the field of the 3Rs is preferably instigated by clear policy goals, as was the case with the Cosmetics Directive. It requires a specification of W*hat do we want to reach in terms of the 3Rs, through which actions and starting when?* The European Commission, the EDQM and the OECD are important institutions that can take a leading role in setting clear mid-term and long-term goals with regard to the acceptance and use of 3R models in the regulatory domain.
4. The society-based track aims at keeping the issues of animal testing and the required paradigm shift high on the agenda. Even though the concern for animal welfare is observed to fluctuate in society and ethical arguments are insufficient on their own, the combination of ethics with the striving for better science has the potential to form a powerful motor for change (Punt et al., 2011). Several actors like regulatory authorities, academia and animal welfare organizations, can play a role in continuing and broadening this discussion.

The process-oriented approach offered through these three tracks requires interconnectedness of the elements at all three levels. Focusing on just a few of these elements entails the risk of 'Kurieren am Symptom' (treating the symptom), a social engineering type of approach that focuses on how to deal with which variable. This leads to the treatment of symptoms without offering routes to ponder the underlying causes. Instead, to overcome the existing tardiness 3R models face in the regulatory domain, thinking in terms of enduring and coordinated strategies is needed. This requires an elevation of the current discourse out of its technocratic track and towards a more integrated and multidimensional level taking political, sociological and psychological powers into account. The 4Cs of Commitment, Communication, Cooperation and Coordination are the central motor to facilitate the development within and between the three tracks. Furthermore, a fifth C for Continuity is added since progress in this area requires a long term effort.

An epistemological presupposition that was uncovered by this research is the stakeholders' strong focus on the generation, validation and dissemination of scientific data to stimulate progress. Most of the energy is put in dealing with scientific uncertainties and working towards the generation of test data. However, this approach has proven to fall short. This thesis should therefore also be read as a plea for natural

sciences to open up to other disciplines (e.g. sociology, policy science and innovation science) to broaden the dominant technocratic perspective and to balance instrumental rationality with value rationality. This thesis makes the connections between these disciplines. It is highly recommended to proceed in the directions indicated here to ensure that scientific and technical developments do not take place without ethical and social checks, and that stakeholders in this field become more aware that non-technical/ non-scientific motives play a pivotal role in the slow transition towards 3R models in the regulatory domain. Only when these underlying beliefs, values and uncertainties are seriously taken into account can progress be made in this area.

SAMENVATTING

(summary in dutch)

1. Inleiding op regulatoire acceptatie en gebruik van 3V-modellen

In Europa worden jaarlijks ongeveer 11,5 miljoen proefdieren gebruikt voor verschillende biomedische doeleinden en onderwijs (EG, 2013). Naar schatting 25% van deze dieren zijn nodig in het kader van de regelgeving ter beoordeling van stoffen, producten en geneesmiddelen op veiligheid en/of werkzaamheid. Deze dierproeven leiden tot wetenschappelijke, ethische en economische bezwaren. De diermodellen zijn vaak decennia geleden ontwikkeld en daarmee gebaseerd op de kennis van toen en - veelal - nooit formeel gevalideerd (Spielmann, 2000). Menig diermodel levert sterk wisselende resultaten op (bijv. Bruckner et al., 2003) en/of leidt tot extrapolatieproblemen als gevolg van de verschillen tussen het proefdier en het doeldier (veelal de mens) waarvoor de testresultaten dienen (Piersma et al., 2014, Martić-Kehl, Schibli & August, 2012). Als gevolg hiervan rijzen er steeds meer twijfels over de wetenschappelijke waarde van diermodellen voor de mens (Pond en Bracken, 2014, Van Meer, 2013). Bovendien leiden dierproeven tot veel ethische bezwaren. Russell en Burch introduceerden in 1959 het 3V-principe met als doel diermodellen te 'vervangen, verminderen en verfijnen' (zie paragraaf 1.1 van dit proefschrift). Deze 3V-benadering biedt de mogelijkheid betere wetenschap te combineren met minder (of geen) diergebruik, minder ongerief voor de dieren en snellere en goedkopere testen. Inmiddels is een breed scala aan 3V-modellen beschikbaar. Het gebruik ervan wordt gestimuleerd door de Europese Richtlijn 2010/63/EU betreffende de bescherming van dieren die voor wetenschappelijke doeleinden worden gebruikt (EU, 2010). De acceptatie en het gebruik van 3V-modellen voor regelgevingsdoeleinden blijkt echter een moeizaam en tijdrovend proces. Dit proefschrift biedt inzicht in dit proces en biedt handelingsstrategieën om de acceptatie en toepassing van 3V-modellen in het regulatoire domein te bevorderen. Het onderzoek richt zich daartoe op de volgende onderzoeksvragen:

Q1. Hoe is het proces van acceptatie en gebruik van 3V-modellen voor de risico- en effectiviteitsbeoordeling van stoffen, producten en geneesmiddelen (het regulatoire domein) te definiëren?

Q2. Welke theoretische perspectieven zijn nodig om dit proces te analyseren?

Q3a. Welke factoren beïnvloeden de acceptatie en het gebruik van 3V-modellen in het regulatoire domein?

Q3b. Hoe beïnvloeden deze factoren de acceptatie en het gebruik van 3V-modellen in het regulatoire domein?

Q 4. Hoe kan het proces van acceptatie en gebruik van 3V-modellen worden bevorderd?

Proefdiergebruik in het regulatoire domein is een complex vraagstuk. Het doorsnijdt geografische-, institutionele- en sectorale grenzen, en omvat een groot aantal verschillende belanghebbenden - zowel publieke als private - vaak met uiteenlopende invalshoeken en belangen. Bovendien wordt het veld waarin dit vraagstuk speelt, gekenmerkt door een sterke mate van risicomijding. (Schiffelers et al., 2007). Omgaan met dit 'wicked problem' vraagt om inzicht in het acceptatieproces. Om zicht te krijgen op factoren die van invloed zijn op het acceptatie- en toepassingsproces is het van belang praktijkvoorbeelden van dergelijke processen te bestuderen.

Veruit de meeste aandacht in het regelgevende veld gaat uit naar de technische ontwikkeling en validatie van 3V-modellen. Uit eerder onderzoek werd duidelijk dat een technische benadering van het probleem onvoldoende is om de kenmerken van en de ontwikkelingen in dit veld te verklaren (Schiffelers et al., 2007: zie Appendix I). Daarom wordt in dit proefschrift een breder multidisciplinair perspectief gehanteerd om het acceptatieproces en de bijbehorende dynamieken te analyseren. Dit perspectief wordt geboden door het probleem te benaderen vanuit een *'Technology Transition'* (TT) benadering (zie paragraaf 4 van deze samenvatting). Deze benadering biedt een integrale kijk op dit proces en daarmee de mogelijkheid het grote aantal factoren, dat invloed uitoefent op de acceptatie en het gebruik van 3V-modellen, in ogenschouw te nemen. In de eerdere studie uit 2007 werd verder geconcludeerd dat de stimulerende en belemmerende factoren naar alle waarschijnlijkheid verschillen per sector en zelfs per casus. Sector-/casus-specifieke informatie werd in de genoemde studie echter niet verkregen. Dit proefschrift gaat door waar het onderzoek uit 2007 ophield. Het biedt zowel sectorspecifieke- als casus-specifieke informatie. De focus ligt daarbij op de farmaceutische- (inclusief vaccins) en de chemische sector in Europa. Dit proefschrift draagt op deze wijze bij aan het verkrijgen van een diepgaander inzicht in en begrip van het acceptatieproces en biedt handvatten om knelpunten in dit proces aan te pakken.

2. Definitie van het acceptatieproces van 3V-modellen in het regulatoire domein

Hoofdstuk 2 geeft antwoord op Q1: *Hoe is het proces van acceptatie en gebruik van 3V-modellen in het regulatoire domein te definiëren?* Regulatoir testen verwijst naar het testen van stoffen, producten en geneesmiddelen voor menselijk of dierlijk gebruik om tegemoet te komen aan wettelijke eisen rond veiligheid en /of effectiviteit. Veel van deze testen zijn gebaseerd op diermodellen die in de eerste helft van de 20e eeuw zijn ontwikkeld en vervolgens ingebed werden in de wettelijke eisen waaraan de producten/ stoffen dienen te voldoen. De publicatie van Russel en Burch (1959) introduceerde het 3V-principe. Het streven was dierproeven zoveel mogelijk te vervangen, verminderen en verfijnen. Vanaf de jaren 70 nam de aandacht voor het 3V-principe toe en sindsdien zijn veel 3V-modellen ontwikkeld en beschikbaar gekomen. De wettelijke eisen waaraan producten dienen te voldoen specificeren vaak welke testen uitgevoerd kunnen/moeten worden om een product te beoordelen. Veelal bieden deze productvereisten ook discretionaire ruimte aan productbeoordelaars en producenten om te bepalen wat de meest geschikte test is om aan te tonen dat het product aan de wettelijke eisen voldoet. Daarmee bestaat dus ook de mogelijkheid een 3V-methode te gebruiken. Richtlijn 2010/63/ EU geeft bovendien aan dat de Europese lidstaten het gebruik van 3V-modellen dienen te waarborgen zodra deze modellen beschikbaar zijn en worden toegestaan door Europese wetgeving. Dit betekent dat regelgevende instanties en fabrikanten gestimuleerd worden de beschikbare discretionaire ruimte te benutten ten gunste van 3V-modellen.
Om het proces van de acceptatie en gebruik van 3V-modellen in het regulatoire domein te bestuderen, is het opgedeeld in de volgende drie deelfasen:

1. De fase van *'Formal Incorporation'* (FI). Dit betreft de wettelijke erkenning van 3V-modellen, oftewel de opname ervan in de wettelijke eisen waaraan producten/ stoffen onderworpen zijn.

2. De fase van *'Actual Regulatory Acceptance'* (ARA) verwijst naar de daarop volgende fase waarin 3V-modellen daadwerkelijk benut worden door regelgevende instanties[1] voor de veiligheids- en werkzaamheidsbeoordeling van producten/stoffen/geneesmiddelen.
3. De fase van *'Use by Industry'* (UI) oftewel het gebruik van deze modellen door de industrie om aan te tonen dat hun product aan de veiligheids- en/of werkzaamheidseisen voldoet.

De acceptatie en het gebruik van 3V-modellen in het regulatoire domein is een uitdagend proces waarin een innovatie (3V-model) moet concurreren met een geïnstitutionaliseerde dierproef, voorheen vaak aangeduid als de 'gouden standaard'.

3. Onderzoeksaanpak

In dit proefschrift wordt een sociaalwetenschappelijk perspectief gebruikt om de ontwikkelingen in het door de natuurwetenschappen gedomineerde onderzoeksgebied te analyseren. Eerder onderzoek (Schiffelers et al., 2007) toonde aan dat dit perspectief bijdraagt aan het verkrijgen van inzicht in het moeizame proces van acceptatie en gebruik van 3V-modellen. De veronderstellingen, die ten grondslag liggen aan dit onderzoek, bouwen voort op Bent Flyvbjerg's boek "*Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again*" (Flyvbjerg, 2001). De waarde van de sociale wetenschappen ligt in de *'phronesis'*, waarin er aandacht is voor de balans tussen instrumentele/technische rationaliteit en waardenrationaliteit. De afweging tussen beide rationaliteiten is van belang om ervoor te zorgen dat wetenschappelijke en technische ontwikkelingen niet plaatsvinden zonder dat rekening gehouden wordt met ethische en maatschappelijke overwegingen. Deze wijze van kijken vergt dat het centrale probleem beschreven en geïnterpreteerd wordt vanuit het perspectief van de betrokken partijen. Flyvbjerg onderstreept het belang van het betrekken van de specifieke context waarin een bepaald fenomeen plaatsvindt. De analyse en interpretatie van waarden en belangen binnen deze specifieke context is naar zijn idee nodig voor het toewerken naar passende oplossingen en 'well managed action´.
In dit proefschrift wordt het belang van de 'the power of example' onderstreept. De onderzoeksopzet biedt ruimte voor contextspecifieke informatie en nuances. Aan de hand van de case study benadering is een tweetal concrete voorbeelden van 3V-acceptatie onderzocht; één op het gebied van vaccins als voorbeeld van farmaceutica (Hoofdstuk 6) en één op het gebied van chemicaliën (Hoofdstuk 7). Vervolgens zijn de bevindingen in het bredere perspectief van de farmaceutische- en de chemische sector geplaatst (Hoofdstuk 8). Via twee expertpanels is de bredere geldigheid van de bevindingen getest en zijn mogelijkheden om het acceptatieproces te stimuleren onderzocht. Het theoretisch model dat gebruikt is voor de analyse van de empirische bevindingen is afkomstig uit Technology Transition (TT) literatuur. Dit integrale perspectief biedt de mogelijkheid de invloed van het brede scala aan beïnvloedende factoren te bestuderen en te analyseren (Hoofdstuk 9). Ook biedt het inzicht in mogelijke strategieën om het acceptatieproces te versnellen (Hoofdstuk 10).

1 Vergunningverlenende organisaties, normstellende organisaties en productbeoordelaars.

4. Een 'Technology Transition' (TT) perspectief voor de analyse van het acceptatieproces

Hoofdstuk 4 heeft betrekking op Q2: *Welke theoretische perspectieven zijn nodig om het proces van acceptatie en gebruik van 3V-modellen te analyseren?* Om het probleem van de acceptatie en het gebruik van 3V-modellen in het regulatoire domein te kunnen bestuderen is gebruik gemaakt van TT literatuur. TT verwijst naar grote technologische wijzigingen in de wijze waarop sociale functies (zoals communicatie, transport, energievoorziening) worden vervuld. De risicobeoordeling van chemische stoffen en kwaliteitscontrole van geneesmiddelen is zo'n sociale functie. In de TT-benadering kan een technologie alleen op een zinvolle wijze haar sociale functie vervullen als zij is afgestemd op de wensen en verwachtingen van betrokken partijen en aansluit bij de ontwikkelingen in het systeem waarin het gebruikt moet gaan worden. Het TT perspectief biedt een analytisch kader om de stimulerende en belemmerende factoren categorisch te onderzoeken. Het omvat drie niveaus, met niveauspecifieke factoren die de doorbraak van een innovatie kunnen stimuleren danwel tegenhouden. Zo worden innovaties ontwikkeld en getest in niches (het microniveau). Deze niches zijn ingebed in en worden beïnvloed door het mesoniveau van het socio-technisch regime. Dit bestaat uit een samenspel van geïnstitutionaliseerde werkwijzen, instanties, regels en voorschriften. Het mesoniveau is op zijn beurt ingebed in het macroniveau bestaande uit het maatschappelijke landschap, waarin de bredere culturele, demografische, technologische, politieke en economische ontwikkelingen de gebeurtenissen op mesoniveau en microniveau beïnvloeden.

De waarde van dit multi-level perspectief is dat het zich niet alleen richt op technologische aspecten van innovaties maar ook aandacht schenkt aan de benodigde veranderingen op het vlak van de huidige testpraktijken, de bestaande infrastructuur, het wettelijk kader en de bestaande overtuigingen van betrokkenen ten aanzien van de bestaande en de nieuwe technologie. Uit TT-literatuur blijkt dat gevestigde regimes innovaties vaak weren als gevolg van een reeks van factoren op het macro-, meso- en microniveau. Zo worden er allerlei waarden toegeschreven aan de bestaande technologieën en voldoen de nieuwe technologieën vaak (nog) niet aan deze waarden. Toch vinden TTs met regelmaat plaats. De literatuur onderscheidt de volgende succesfactoren en strategieën die van belang zijn voor het realiseren van een TT:

- *'Champions'* of innovation-entrepreneurs met zowel een sterke connectie met de innovatie als met het bestaande regime zijn van cruciaal belang voor Technology Transitions. Zij kunnen de ontwikkelingen in de niche verbinden met de ontwikkelingen binnen het socio-technische regime;
- Samenwerkingsverbanden tussen voorstanders van de nieuwe technologie zijn belangrijk om de innovatie te ontwikkelen, van elkaar te leren en ervaringen uit te wisselen. Publiek private partnerschappen zijn van belang om deze samenwerkingsverbanden te vergroten en te verstevigen;
- Verder is een duidelijke transitiestrategie van belang. Door middel van het formuleren van heldere doelen en het vastleggen van bijbehorende activiteiten kan een TT gerichter gerealiseerd worden;

- Proefprojecten en demonstraties bieden ruimte voor praktijkgericht leren en voor het tijdig betrekken van de behoeften en verwachtingen van eindgebruikers. Op deze wijze kunnen potentiele problemen tijdig gesignaleerd worden en wordt de gebruiker voorbereid op de innovatie.

Spanningen binnen het bestaande socio-technisch regime zijn ook potentieel stimulerende factoren en kunnen leiden tot fasen waarin de bestaande werkwijze wordt verzwakt en nieuwe technologieën terrein winnen. Dit kan het gevolg zijn van kritiek op de bestaande technologie, het veranderen van bestaande gebruikerspraktijken, een verouderde (test)infrastructuur, veranderingen in het beleid en het beschikbaar komen van nieuwe wetenschappelijke inzichten. Technologische verandering is het resultaat van de mate waarin stimulerende factoren de bestaande barrières weten te overtreffen.

5. Het multilevel-perspectief als basis voor het 3V-acceptatiemodel

Het multilevel-perspectief vormde de basis voor de ontwikkeling van het 3V-acceptatiemodel dat gepresenteerd wordt in Hoofdstuk 5. Dit model dient verschillende doelen. Ten eerste faciliteert het de categorisering van de brede reeks aan stimulerende en belemmerende factoren naar het niveau waarop zij zich bevinden; dat wil zeggen het micro-, meso- en macroniveau. Ten tweede maakt het de analyse van de dynamiek tussen deze variabelen zowel binnen als tussen de drie niveaus mogelijk en inzichtelijk. Ten derde helpt het bij het specificeren van de sterke en de manipuleerbare variabelen, in de zin dat de factoren op micro- en deels het mesoniveau normaliter beter manipuleerbaar zijn dan de factoren op het macroniveau. Succesvolle TT vergt afstemming tussen de ontwikkelingen op de drie niveaus en zijn bijna altijd *"het resultaat van de wisselwerking tussen vele factoren en actoren"* (Geels, 2006). Een TT kan alleen plaatsvinden als een innovatie, bijvoorbeeld een 3V-model, voldoet aan de behoeften, wensen en eisen die eraan gesteld worden op het meso- en macroniveau. Het 3V-acceptatiemodel biedt een aanvullend puzzelstuk voor de beantwoording van Q2. Verder wordt in dit hoofdstuk een risico-regelingsperspectief geïntroduceerd om de specificiteit van het onderzoeksterrein in ogenschouw te nemen. Het streven naar risico-minimalisatie is namelijk een dominant kenmerk van de regulatoire context waarvoor 3V-modellen zijn bedoeld. Het streven naar risico-minimalisatie is terug te vinden in wettelijke vereisten waar geneesmiddelen en chemicaliën aan onderworpen zijn en heeft invloed op de manier waarop productbeoordelaars kijken naar nieuwe technologieën. Het leidt bijvoorbeeld tot een intrinsieke afkeer van onzekerheid en daarmee tot terughoudendheid ten aanzien van verandering (Breyer, 1993). Dit wordt versterkt door de publieke en politieke druk om risico's te beperken. Innovatie begint met de bereidheid tot falen (MacLachlan, 1994) maar falen op het terrein van productregulering kan grote gevolgen hebben. Regelgevende instanties en bedrijven hebben vaak nog onvoldoende vertrouwen in de 3V-modellen om deze sprong te wagen. Het is van belang te realiseren dat de transitie naar 3V-modellen in het regulatoire domein gebaat is bij een geleidelijk proces. Dat wil zeggen: geen radicale veranderingen maar nieuwe '*regulatory regimes'* die geleidelijk uit oude regimes ontstaan (Geels, 2002).

6. De SNT casus betreffende werkzaamheidstesten voor veterinaire rabiës vaccins

In Hoofdstuk 6 wordt het proces beschreven van de acceptatie en het gebruik van de muis antilichaam test *'Serum Neutralisation Test'* (SNT) ter vervanging van de *'NIH mouse challenge test'* voor het onderzoek naar de werkzaamheid van rabiës (hondsdolheid) vaccins.' De SNT is een verminderings- en verfijningsmodel. Het gebruikt aanzienlijk minder dieren (een reductie tot wel 85%) en veroorzaakt aanmerkelijk minder stress en leed bij de betrokken dieren. De casus richt zich op de onderzoeksvragen Q3a en Q3b. De casestudy reconstrueert het acceptatieproces en beschrijft de remmende en stimulerende factoren die daarin een rol hebben gespeeld. De bevindingen zijn afkomstig van een combinatie van beschikbare literatuur op het vlak van deze casus, interviews met experts, deelname aan een reeks internationale bijeenkomsten en een enquête uitgevoerd voorafgaand aan deze casestudy (Appendix II). De SNT is in april 2013 opgenomen in Monografie 0451 van Europese Farmacopee betreffende de werkzaamheid van geïnactiveerde veterinaire rabiësvaccins. Deze casus is een voorbeeld van succesvolle Formal Incorporation (FI) omdat de SNT relatief snel uit de niche waarin ze werd ontwikkeld en gevalideerd, wist te ontsnappen. Dit was mede het resultaat van de succesvolle (pre)validatie van de SNT in een *'collaborative study'*.[2] Door middel van deze gezamenlijke studie kon aangetoond worden dat de test vergelijkbare resultaten oplevert wanneer uitgevoerd door verschillende partijen. Het proces van testontwikkeling, (pre)validatie en de FI werd gefaciliteerd door toegewijde innovatie-entrepreneurs en een duidelijk probleemeigenaarschap van en procesmanagement door institutionele spelers.[3] Daarnaast droegen de volgende factoren bij aan dit succes: de intensieve samenwerking binnen het *'Official Medicines Control Laboratories'* netwerk[4], de vroege betrokkenheid van een statisticus om de validatiestudie te ontwerpen en de gegevens te analyseren en de brede verspreiding van de resultaten van het onderzoek. Het feit dat de test werd ontwikkeld door het Paul Ehrlich Instituut in Duitsland, één van de toonaangevende Europese beoordelaars van rabiësvaccins, vergemakkelijkte de Actual Regulatory Acceptance (ARA) ervan binnen Europa voor de kwaliteitscontrole van deze veterinaire rabiësvaccines. Verder stimuleert de Europese Farmacopee de acceptatie en het gebruik van 3V-modellen. De snelheid in de fasen van FI en ARA had echter een mogelijk negatief effect op de fase van ARA (buiten Europa) en de UI (Use by Industry). Hoewel de innovatie-entrepreneurs en coördinatoren van de collaborative study een belangrijke succesfactor vormden in de FI van de SNT, was er minder aandacht vanuit deze initiatiefnemers voor de resterende minpunten van de SNT. Het Amerikaanse Ministerie van Landbouw (USDA), verantwoordelijk voor de regulering van rabiësvaccins in de Verenigde Staten, en verschillende producenten

2 Een onderzoek ten behoeve van de validatie van een testmodel waarbij een breed scala aan stakeholders samenwerkt binnen vooraf vastgestelde toetskaders.

3 Als eerste is het Duitse 'Official Medicines Control Laboratory' (OMCL) het Paul Ehrlich Instituut, vervolgens het 'Biological Standardization Program' (BSP) van het European Directorate for the Quality of Medicines & HealthCare (EDQM), vervolgens de Europese Farmacopee Commissie en "Group 15V" inzake veterinaire sera en vaccins verantwoordelijk voor de evaluatie en goedkeuring van veterinaire vaccins monografieën.

4 Netwerk van overheidslaboratoria die iedere partij van een vaccin toetsen aan de specificaties zoals vastgelegd in het registratiedossier, waarna de partij kan worden vrijgegeven voor de markt.

van dit vaccin hadden duidelijke bedenkingen bij de SNT. Dit omdat de hoeveelheid antistoffen die de SNT bepaalt, onvoldoende informatie biedt over de werkzaamheid van het vaccin. Zowel de USDA als verschillende fabrikanten zetten daarom vraagtekens bij de waarde van de SNT en wilden in plaats daarvan investeren in *in vitro*-methoden, die het merendeel van de fabrikanten al gebruikt voor productiebewakingsdoeleinden. Met andere woorden, de bevorderende factoren waren dominant in de deelfasen FI en ARA. Voor de ruimere acceptatie (ARA buiten Europa) en het gebruik (UI) wegen de beperkingen van de SNT echter zwaarder dan de voordelen.[5]

De SNT casus biedt de volgende leerpunten voor soortgelijke toekomstige processen. Ten eerste toont de casestudy aan dat er een stevige betrokkenheid van belanghebbende partijen nodig is. Het vraagt commitment en het alloceren van geld en tijd om de innovatie gezamenlijk te ontwikkelen en te valideren. Deze betrokkenheid wordt gestimuleerd door wetsteksten van zowel de Raad van Europa[6] als de Europese Commissie[7], die het gebruik van 3V-modellen stimuleren. Daarnaast was nauwe samenwerking nodig voor de uitwisseling van methoden, gegevens en testmateriaal. Ook was er nauwkeurige afstemming om te zorgen dat de partijen voldoen aan de specifieke regels van een collaborative study. Het SNT-proces toont daarmee het belang aan van een helder en gecoördineerd validatieproces en een strak procesmanagement, van vooraf gedefinieerde stappen, van vastgestelde vragen die beantwoord dienen te worden en van te behalen doelstellingen. Het feit dat de initiatiefnemers afkomstig waren uit het bestaande '*regulatory regime'* was van groot belang om het model te laten landen ín dat betreffende regime.

7. De EOGRTS casus betreffende het testen van chemicaliën op reproductie toxiciteit

Hoofdstuk 7 beschrijft het proces van de acceptatie en het gebruik van de *'Extended One Generation Reproductivity Toxicity Study'* (EOGRTS) ter vervanging van de twee-generatie test (OECD TG 416)[8] voor de risicobeoordeling van industriële chemicaliën in het kader van REACH.[9] Evenals bij de SNT casus is het proces van acceptatie gereconstrueerd en zijn de stimulerende en belemmerende factoren, die een rol spelen in de subfasen van FI, ARA en UI, beschreven. De casus geeft daarmee eveneens antwoord op de onderzoeksvragen Q3a en Q3b. De bevindingen zijn afkomstig van een analyse van beschikbare literatuur en beleidsdocumenten op het vlak van deze casus en interviews met betrokken experts. OECD TG 416 is de standaardtest

5 In dit verband hebben verscheidene respondenten uit de industrie gesuggereerd dat zij in een eerder stadium betrokken zouden moeten zijn geweest in het proces van de ontwikkeling en pre-validatie van SNT om tijdig te spreken over eventuele beperkingen van het 3V-model.

6 Verdrag 123: Europees Verdrag inzake de bescherming van gewervelde dieren die voor experimentele en andere wetenschappelijke doeleinden worden gebruikt

7 Richtlijn 2010/63/EU betreffende de bescherming van dieren die voor wetenschappelijke doeleinden worden gebruikt

8 The Organisation for Economic Co-operation and Development (Organisatie voor Economische Samenwerking en Ontwikkeling –OESO-) Test Guideline (Test richtlijn).

9 Europese richtlijn voor de Registratie, Evaluatie, Autorisatie en beperking van chemische stoffen - REACH (EU, 2006)

binnen de *'Test Methods Regulation'* (TMR) van REACH om de reproductietoxiciteit van chemische stoffen met productievolumes >100 ton te testen. De test wordt bekritiseerd om een combinatie van wetenschappelijke en ethische redenen. Geschat werd dat het onderzoek naar reproductietoxiciteit verantwoordelijk zou zijn voor bijna 40% van het totale proefdiergebruik in de context van REACH (Janer et al., 2007a) en als potentieel grootste bron van diergebruik binnen REACH (Pedersen et al, 2003;. Van der Jagt et al., 2004). Bovendien stelt REACH artikel 25 (1) het volgende: *"Om dierproeven te voorkomen worden ….. slechts in laatste instantie proeven op gewervelde dieren uitgevoerd. Tevens moeten maatregelen worden genomen om te voorkomen dat proeven meerdere malen worden uitgevoerd."* (EU, 2006). De EOGRTS, in 2011 opgenomen in de OECD-testrichtlijnen (OECD TG 443), biedt de mogelijkheid TG 416 te vervangen, gebruik makend van slechts één generatie van ratten. Dit leidt tot een vermindering in het aantal proefdieren per test van ongeveer 40% (een totaal van zo'n 1.200 proefdieren per test). Bovendien biedt de EOGRTS meer informatie doordat meer onderzoeksparameters onderzocht worden. De acceptatie van de EOGRTS in de context van REACH bleek echter een lastig proces.

Het stadium van de formele acceptatie (FI) van de EOGRTS in de OECD-testrichtlijnen verliep evenals bij de SNT casus relatief vlot. Dit was eveneens te danken aan sterke voorstanders die verbonden waren aan het bestaande regulatoire regime[10] en daarnaast aan een reeks retrospectieve analyses over de toegevoegde waarde van het gebruik van de tweede generatie proefdieren (F2) (Janer et al, 2007a; Martin et al, 2009; Piersma et al, 2011; Rorije et al, 2011), de zorg over het dierenwelzijn en het veranderend Europees en Amerikaans wettelijk kader rondom chemische stoffen dat bijdroeg aan de zoektocht naar alternatieve testwijzen. Het proces voorafgaand aan de FI bestond echter ook uit diverse stevige discussies over onder meer de toegevoegde waarde van de tweede generatie en het al dan niet opnemen van parameters voor ontwikkelingsneurotoxiciteit (DNT) en ontwikkelingsimmunotoxiciteit (DIT) in het standaard EOGRTS-protocol. Deze discussies zijn niet volledig afgehecht tijdens de fase van FI omdat de precieze invulling van de EOGRTS besproken diende te worden binnen de specifieke context waarin de test gebruikt zou gaan worden. Dit maakte dat in de fase van ARA in het kader van REACH, deze discussies wederom ter tafel kwamen. Daarbij benadrukten de stakeholders, die twijfelden over de geschiktheid van de EOGRTS, de wettelijke inbedding van TG 416 in de TMR van REACH en het niet wettelijk verankerd zijn van de EOGRTS in deze TMR. Voorstanders van de EOGRTS benadrukten daartegenover de bredere verplichting vanuit REACH en Directive 2010/63/EC om 3V-modellen te gebruiken wanneer beschikbaar en toepasbaar. Deze juridische discussie ging hand in hand met de eerder beschreven wetenschappelijk discussie. Deze voortdurende discussies leidden tot onzekerheid onder de producenten van chemicaliën over de acceptatie van de EOGRTS voor de registratie en risicobeoordeling van hun chemische stoffen. Deze onzekerheid werd versterkt door de hoge kosten die verbonden zijn aan het uitvoeren van de EOGRTS, het toegenomen risico van vals-positieven als gevolg

10 Te beginnen met de Verenigde Staten, Duitsland en Nederland die een voorstel indienden bij het OECD-secretariaat om een OECD-testrichtlijn op te stellen (TG) gebaseerd op de EOGRTS

van de extra parameters waar de chemicaliën op getest worden en de complexiteit en daarmee de uitvoerbaarheid van de EOGRTS. De stimulerende factoren op microniveau (beschikbaarheid van diverse retrospectieve analyses met grote hoeveelheden data en de betrokkenheid van sterk gecommitteerde innovatie-entrepreneurs en de voordelen van de EOGRTS in termen van minder diergebruik en extra informatie) bleken sterk genoeg voor de FI van de EOGRTS binnen de OECD TGs. In de context van REACH konden deze factoren echter niet alle partijen overtuigen.

Twee schijnbaar tegengestelde lijnen van argumentatie domineerden het discourse in de fase van ARA, te weten de argumentatielijn waarin het voorzorgsprincipe domineerde versus de argumentatielijn waarin het innovatieperspectief voorop stond. De voorstanders van de EOGRTS richtten zich op de voordelen van de innovatie, terwijl de voorstanders van TG 416 de onzekerheden onderstreepten die verbonden zijn aan de EOGRTS en aan het weglaten van de tweede generatie dieren. Het aandragen van aanvullende wetenschappelijke data bleek niet afdoende om de ontstane kloof te overbruggen en leek op enig moment zelfs de onenigheid te voeden. Dit leidde ertoe dat de partijen uiteindelijk *"agreed to disagree".* De EOGRTS zakte daarmee in eerste instantie terug naar het microniveau om haar geschiktheid opnieuw te bespreken/ duidelijker aan te tonen. In 2014 is de EOGRTS alsnog onderdeel geworden van de TMR. Onduidelijk blijft echter in hoeverre de UI gerealiseerd is.

Uit deze casus kunnen de volgende lessen worden getrokken. Gedegen wetenschappelijke informatie, sterke en gecommitteerde voorstanders en veranderingen in het wettelijk kader zijn belangrijke ingrediënten om een 3V-model de kans te bieden een plek te verwerven in het regulatoire regime. Maar de acceptatie van 3V-modellen blijkt vaak meer voeten in aarde te hebben. Zo is het acceptatieproces vaak een sterk gepolitiseerd proces, waarin psychologische-, politieke- en sociale argumenten (zoals het gebrek aan ervaring met en vertrouwen in het nieuwe model, institutionele agenda's en maatschappelijke wensen en verwachtingen) vaak zwaarder wegen dan de wetenschappelijke 'feiten'. Deze 'andere' argumenten moeten dus constant meegenomen worden. Dit vergt voortdurende communicatie over en anticipatie op mogelijk stimulerende en remmende factoren. Bovendien dienen belanghebbenden zich bewust te zijn van het feit dat de drie subfasen van FI, ARA en UI sterk met elkaar verbonden zijn en dat de erfenis van een eerdere fase van invloed blijft op daarop volgende fases.

8. Expert panels over de acceptatie het gebruik van 3V modellen voor farmaceutische en chemische producten

In Hoofdstuk 8 worden de factoren die een rol spelen bij de acceptatie van 3V-modellen in de farmaceutische sector (waaronder vaccins) en chemische sector, beschreven en vergeleken. Daarnaast zijn opties ter verbetering van het acceptatieproces geïnventariseerd. Dit hoofdstuk levert daarmee antwoord op onderzoeksvragen Q3a en Q3b en biedt tevens een voorlopig antwoord op Q4: *Hoe kan het proces van acceptatie van de regelgeving en het gebruik van 3V-modellen worden bevorderd?*

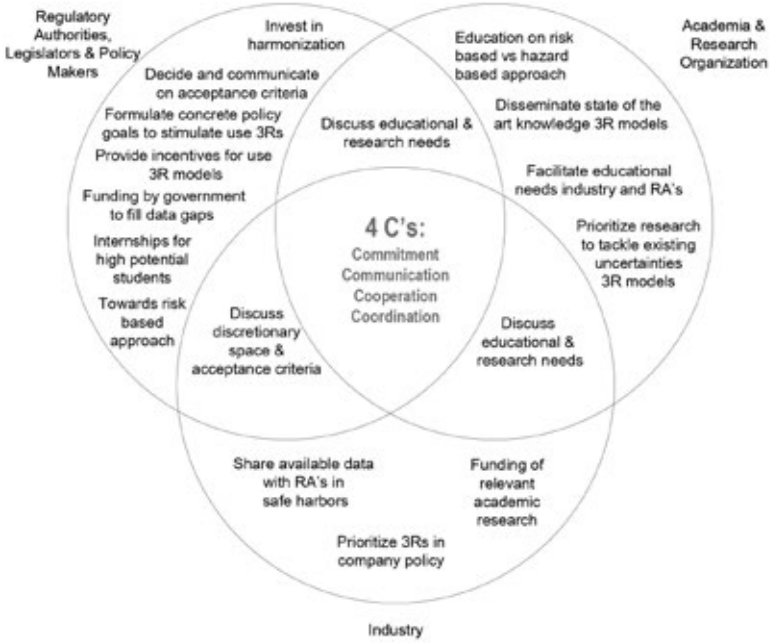
Deze vragen zijn voorgelegd aan twee expertpanels, één met experts afkomstig uit de farmaceutische sector en één met experts uit de chemische sector.[11] Uit de resultaten komt naar voren dat in beide sectoren vergelijkbare mechanismen voorkomen die invloed uitoefenen op de acceptatie van 3V-modellen. Zo spelen op het microniveau de onzekerheden, die 3V-modellen vaak nog met zich meebrengen, in beide sectoren een belangrijke rol. Er is veel discussie over de voorspellende waarde van 3V-modellen, evenals over de vaak lastige *in vitro-in vivo* extrapolatie. Op het mesoniveau vormt het gebrek aan harmonisatie van wet- en regelgeving een gedeelde centrale barrière, evenals het ontbreken van duidelijke acceptatiecriteria voor 3V-modellen. Op het macroniveau speelt de sterke afkeer van risico's in beide sectoren een belangrijke rol. Daartegenover staan de volgende gedeelde stimulerende factoren. 3V-modellen worden door beide panels gewaardeerd vanwege hun informatieve karakter in termen van werkingsmechanisme, reproduceerbaarheid en wetenschappelijke robuustheid. Op het mesoniveau speelt de aanjagende werking van horizontale wetgeving, zoals de Europese Richtlijn 2010/63/EU, een rol. De ambitie van deze richtlijn om de afhankelijkheid van dierproeven te verminderen is terug te vinden in verticale wet- en regelgeving, zoals de monografieën van de Europese Farmacopee en REACH. Een centrale stimulerende factor op macroniveau is de zorg van de samenleving over dierproeven vanuit dierwelzijnsperspectief.

Tevens zijn sectorale verschillen geïdentificeerd. De veiligheidsbeoordeling van farmaceutische producten hanteert risico-baten analyse. Voor chemicaliën en vaccins geldt dat de afkeer van risico's gemiddeld genomen veel hoger is dan voor farmaceutische producten. Dit komt ook tot uiting in de veiligheidsbeoordeling. Beoordelaars van chemische stoffen en vaccins streven naar een 'zero risk level' en zijn mede daardoor zeer terughoudend in het veranderen van het bestaande testregime, bijvoorbeeld door over te stappen op een 3V-model. Bovendien zijn de winstmarges bij geneesmiddelen groter, waardoor er meer ruimte is te investeren in innovaties zoals 3V-modellen. Tot slot wordt voor farmaceutische producten meer data gegenereerd[12] wat risicobeoordeling op basis van de werking van een product mogelijk maakt, een mogelijkheid die ontbreekt bij chemicaliën. Dergelijke sector kenmerken zijn belangrijk om mee te nemen bij het overwegen van mogelijkheden om het acceptatieproces te bevorderen.

Ter bevordering van dit proces hebben de experts verschillende unilaterale, bilaterale en tripartite acties geïdentificeerd. Een overzicht van deze acties is weergegeven in Figuur 1. De 4C's van 'Commitment, Communication, Coöperation and Coördination' vormen de centrale motor voor de aandrijving van deze acties (Schiffelers et al., 2008: zie Hoofdstuk 8).

11 Ten eerste is de panelleden gevraagd om een individuele inventarisatie van remmende en stimulerende factoren te maken. De factoren werden gegroepeerd naar thema. Ten tweede is een prioritering gemaakt in termen van mate van invloed op het acceptatie proces. Ten derde zijn acties geïdentificeerd die door de drie betrokken stakeholder groepen kunnen worden opgepakt ter verbetering/versnelling van het acceptatieproces.

12 Bijvoorbeeld farmacologisch werkingsmechanisme, dierlijke en menselijke farmacokinetische gegevens



Figuur 1. Acties per stakeholdergroep ter bevordering van de acceptatie en het gebruik van 3V-modellen in het regulatoire domein

9. Conclusies

In Hoofdstuk 9 worden de empirische bevindingen afkomstig van de casestudies en de expertpanels aan elkaar gekoppeld. Daarmee wordt antwoord gegeven op de onderzoeksvragen Q3a en Q3b.

De volgende stimulerende factoren zijn terug te vinden in zowel de casestudies als de expertpanels en tonen zich daarmee dominant in het discourse over de acceptatie en het gebruik van 3V-modellen:

- De ethische bezwaren ten aanzien van proefdiergebruik en de behoefte aan proefdiervrije benaderingen om tegemoet te komen deze bezwaren (macroniveau);
- Wetgeving die het gebruik van de 3V's bevordert en voortdurende afstemming en interactie tussen de centrale stakeholders over de te bereiken doelen en de te volgen strategie (mesoniveau); en
- De nadelen van dierproeven, de voordelen van 3V-modellen, de beschikbaarheid en uitwisseling van testresultaten, de aanwezigheid van innovatie-entrepreneurs, de vroegtijdige betrokkenheid van zowel regelgevende instanties als fabrikanten en het opdoen van ervaring met en vertrouwen winnen in de nieuwe testmethode middels (pre)validatie-studies (microniveau).

De volgende barrières komen naar voren in iedere empirische stap:

- Het streven naar het minimaliseren van risico's en het bijbehorende voorzorgsprincipe (macroniveau);
- De gepercipieerde risico's van vaccins en chemische stoffen, de angst voor mogelijke

incidenten en het vrijgeven van onveilige producten, de daarop volgende strikte interpretatie van producteisen en de geïnstitutionaliseerde status van de conventionele test. Het gebruik van 3V-modellen door de industrie wordt vooral belemmerd door het gebrek aan harmonisatie van de regelgeving en de onduidelijke acceptatiecriteria voor 3V-modellen (mesoniveau); en

- De beperkingen en onzekerheden die 3V-modellen vaak nog met zich meebrengen, kosten die gemaakt moeten worden en de inspanning die geleverd moet worden voor de overgang naar de nieuwe wijze van testen (microniveau) terwijl de feitelijke acceptatie ervan tot op het laatst onzeker blijft.

Sommige van deze factoren vormen krachtige (verklarende) variabelen voor het moeizame acceptatieproces maar zijn tegelijkertijd lastig te beïnvloeden (Ellemers, 1976). Zo vormen het streven naar risicominimalisatie en het gebrek aan harmonisatie krachtige verklaringen, maar deze zijn tegelijkertijd lastig te beïnvloeden. De aanpak van deze factoren vergt een lange adem. Andere factoren zijn beduidend beter te beïnvloeden. Te denken valt aan het stimuleren van interactie tussen betrokken stakeholders, het beschikbaar maken en uitwisselen van gegevens, de aanwezigheid van innovatie- entrepreneurs, de betrokkenheid van statistici om testresultaten te interpreteren en de tijdige betrokkenheid van eindgebruikers om de gebruikerseisen te kunnen inventariseren en ervaring op te kunnen doen met de nieuwe testmethode.

Uit de analyse blijkt dat op alle drie de niveaus twee tegengestelde krachten aanwezig zijn. Aan de ene kant is er een innovatieve benadering, die voortkomt uit een combinatie van ethische en wetenschappelijke drijfveren. Aan de andere kant is er een meer behoudende voorzorgbenadering, die tegemoet komt aan de hoge mate van risicoaversie. De uitkomst van de strijd tussen beide krachten bepaalt in hoeverre er sprake zal zijn van een TT. Tot nu blijkt de voorzorgsbenadering het vaak nog te winnen van de innovatiebenadering (afgezien van de cosmeticasector waarin de *'European Cosmetics Regulation'* (No 1223/2009) heeft geleid tot de geleidelijke afschaffing van het gebruik van proefdieren en daarmee de ontwikkeling en het gebruik van veel nieuwe 3V-modellen). Maar de innovatiebenadering wint in dit onderzoeksveld langzaam maar zeker aan terrein.

De gerealiseerde transities in zowel de SNT als de EOGRTS casus zijn terug te voeren op de verbindingen die gelegd konden worden tussen de ontwikkelingen op meso- en microniveau, waardoor de noodzaak tot vervanging, vermindering of verfijning van een bepaald diermodel hoog op de agenda kwam ter staan van de institutionele spelers. De meeste energie in beide casus was echter gericht op het genereren, controleren en distribueren van onderzoeksdata (zogenaamd *'first order learning'*). Dit type informatie is van groot belang, maar een focus op dit type leren houdt het risico in dat andersoortige aspecten die een rol spelen bij de acceptatie van 3V-modellen onvoldoende meegenomen worden. Van leerprocessen is bekend dat ze meer bijdragen als ze verder gaan dan het niveau van het verzamelen van gegevens en voorzien in *'second order learning'*. Hierbij gaat het om het stimuleren van veranderingen in cognitieve kaders en het gezamenlijk reflecteren op de veronderstellingen (waarden en normen) binnen deze kaders (Grin en Van de Graaf, 1996). Dit brengt ons bij het laatste hoofdstuk van dit proefschrift waarbij de vraag is hoe het proces van acceptatie en toepassing bevorderd kan worden.

10. De weg vooruit

Uit de voorgaande hoofdstukken is gebleken dat het genereren, valideren en delen van onderzoeksdata erg belangrijk is, maar niet afdoende om een 3V-model door de drie subfasen van regulatoire acceptatie en gebruik te leiden. Met het huidige dominante technocratische perspectief van de stakeholders TT richting de 3Vs beogen, wordt te snel over het hoofd gezien dat de invoering en acceptatie van nieuwe technologieën vaak meer afhangt van sociale, psychologische, culturele en historische factoren, dan van de technologische waarde van een innovatie (NRC, 2004). Een eenzijdige focus op technische data, kan de kloof tussen de verschillende meningen en belangen zelfs vergroten wanneer de data onderdeel worden van de controverse. Stakeholders die actief betrokken waren bij processen van acceptatie van 3V-modellen, zijn zich vaak pijnlijk bewust van het feit dat deze onderliggende processen met regelmaat een grotere rol spelen dan men zou wensen vanuit het natuurwetenschappelijke adagium 'meten is weten'. Het acceptatieproces evolueert vaak langzaam als gevolg van, vaak tegenstrijdige, sociale constructies. Deze sociale constructies zijn gevormd door diepgewortelde waarden en overtuigingen, verhalen, beelden en percepties ten aanzien van zowel het diermodel als de alternatieve wijze van testen. Dit zijn elementen die weinig zichtbaar zijn maar grote invloed uitoefenen op de wijze waarop stakeholders acteren en met elkaar interacteren als het gaat om de acceptatie van 3V-modellen. Hoewel sociale constructies niet gemakkelijk te veranderen zijn, kunnen ze te zijner tijd worden gereconstrueerd. Dit laatste hoofdstuk beschrijft wat er gedaan kan worden om de huidige situatie te veranderen en richt zich daarmee op onderzoeksvraag 4: *Hoe kan het proces van acceptatie van de regelgeving en het gebruik van 3V-modellen worden bevorderd?*

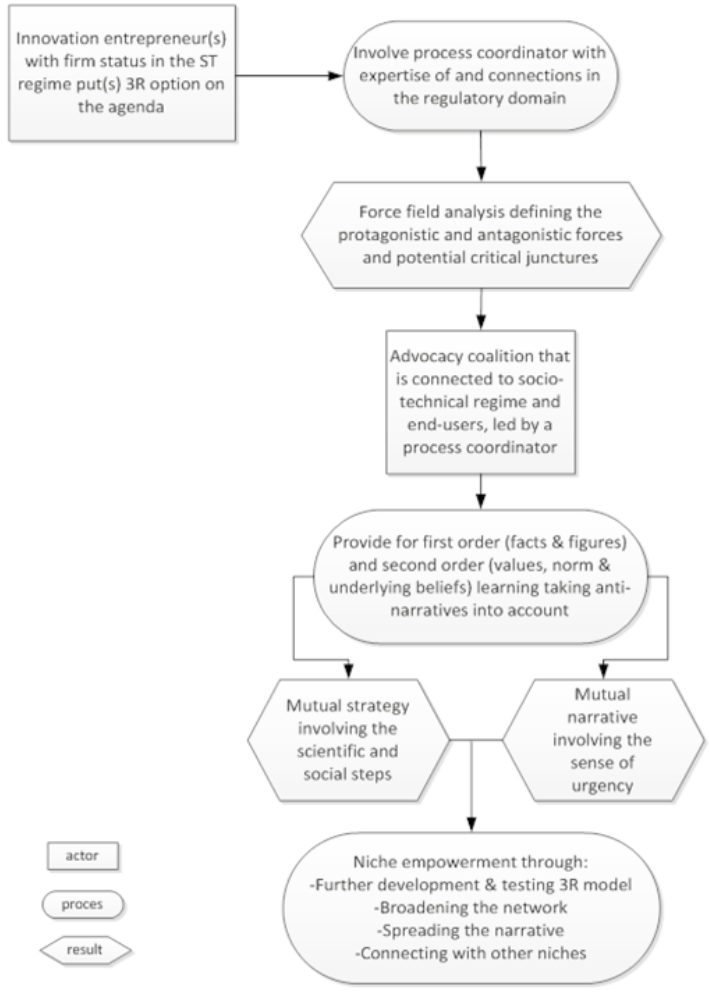
Allereerst is het van belang te beseffen dat het 3V-acceptatiemodel op alle drie de niveaus mogelijkheden biedt tot verandering. Wel vergen veranderingen op het macro- en deels het mesoniveau doorgaans meer volhardendheid. Innovatie-entrepreneurs dienen zich bewust te zijn van de potentiële mogelijkheden op alle drie de niveaus en van mogelijke kruispunten (*'critical junctures'*) waarbij de ontwikkelingen op verschillende niveaus bij elkaar kunnen komen; bijvoorbeeld rondom veranderingen in beleid of wetgeving. Dit vergt een goed gevoel voor timing, Transformaties zijn echter lang niet altijd het gevolg van critical junctures. Vaker nog is transformatie het gevolg van een cumulatief en evoluatief proces (Capoccia en Kelemen, 2007) waarin nieuwe technologieën fysiek aansluiten bij gevestigde technologieën. Dit denken in termen van evolutie in plaats van revolutie (Rotmans et al., 2001) is om twee redenen belangrijk. Ten eerste zijn critical junctures relatief zeldzame gebeurtenissen. Ten tweede is een evolutionaire benadering van technologische transformatie waarschijnlijk een meer passende aanpak in deze risicomijdende context waarin plotselinge verandering vooral onzekerheid betekent.

TT kan op verschillende manieren bevorderd worden door verbindingen te maken tussen de stimulerende factoren op het micro-, meso- en macro niveau. Transitiemanagement is erop gericht een breed scala aan belanghebbenden op het micro-, meso- en macro niveau te betrekken om een gezamenlijke lange-termijn visie en middellange termijn doelstellingen vast te stellen ten aanzien van het gebruik van dierproeven in het regulatoire domein. Dit is in lijn met de geuite behoefte van veel betrokkenen in dit veld om toe te werken naar heldere (Europese) beleidsdoelstellingen om het gebruik van 3V-modellen te

stimuleren en toe te werken naar minder (of geen) proefdiergebruik. Richtlijn 2010/63/EU biedt brede kaders, maar biedt geen concrete beleidsdoelstellingen. Verder is strategisch niche-management van belang voor de versterking van de niches waarin nieuwe 3V-modellen ontwikkeld en gevalideerd worden. Strategisch niche-management is erop gericht de netwerken van de niches te vergroten, de link met het regulatoire regime te versterken, de leerprocessen te stimuleren, gezamenlijke verwachtingen en visies te formuleren en coalities met een gedeelde agenda te vormen (Boon et al., 2014).
Het in hoofdstuk 10 gepresenteerde stappenplan combineert elementen van beide strategieën en combineert daarmee een *'top-down'* en een *'bottom-up'* strategie.
Het stappenplan onderscheidt drie routes:

1. De *'niche-based track'* (Figuur 2) start bij de innovatie en richt zich op nicheversterking door nicheontwikkeling en - accumulatie (verbindingen met andere niches waarin eveneens 3V-modellen worden ontwikkeld) en het leggen van verbindingen met het bestaande regulatiore regime. Het begint met innovatie-entrepreneurs die nauw verbonden zijn met de regulatoire context waarin het 3V-model zijn plek moet zien te verwerven. Ter versterking van de niche is het van belang samen te werken met een kerngroep van betrokken stakeholders; de zogenaamde *'advocacy coalition'*. Dit samenwerkingsverband ontwikkelt vervolgens een gezamenlijke strategie met duidelijke gemeenschappelijke doelen en een concreet stappenplan om deze doelen te bereiken. Een krachtenveld-analyse wordt gemaakt ter identificatie van de spelers in het veld en hun belangen. Binnen de niche wordt gewerkt aan een gedeeld verhaal met een heldere argumentatie over het belang van en de urgentie tot overschakelen op het 3V-model. Dit narratief houdt rekening met mogelijke tegenwerpingen van buitenaf en biedt veel ruimte tot discussies over andere opvattingen en resterende onzekerheden. Daartoe is het van belang dat de innovatie-entrepreneurs (doorgaans inhoudelijke experts met een sterke drive om de technologie zo snel mogelijk geïmplementeerd te krijgen en daarmee het risico te lopen tegengeluiden te negeren) begeleid worden door neutrale procescoördinatoren. Deze procescoördinatoren dienen over de vaardigheden te beschikken om te kunnen omgaan met de bestaande controverses. Dit vergt het zoeken naar gemeenschappelijke belangen die bijvoorbeeld gevormd kunnen worden door heldere beleidsdoelstellingen op het meso niveau, naar mogelijkheden tot consensus en naar manieren om alle belanghebbenden actief betrokken te houden.
2. De *'regime-based track'* is gericht op het creëren van een bredere *'sense of urgency'* om het gebruik van 3V-modellen en alternatieve testbenaderingen te stimuleren. Een dergelijk gevoel van urgentie op het gebied van de 3V's kan gecreëerd worden door heldere beleidsdoelstellingen, zoals het geval was met de Richtlijn voor Cosmetica. Het vereist een specificatie van *"Wat willen wij bereiken in termen van de 3V's, door middel van welke acties en wanneer?"* De Europese Commissie, de EDQM en de OESO kunnen een leidende rol spelen in het vaststellen van dergelijke doelen met betrekking tot de acceptatie en het gebruik van 3V-modellen in het regulatoire domein.
3. Bij de *'society-based track'* ligt de focus op het hoog op de agenda houden van de beperkingen die dierproeven met zich meebrengen en op het verhelderen van de noodzaak tot een paradigmaverschuiving. De combinatie van ethische argumenten en het streven naar betere wetenschap is daarbij in potentie een krachtige motor

voor verandering (Punt et al., 2011). Er zijn al veel stakeholders die de noodzaak tot een 'paradigmashift' continu adresseren. Het vergt echter doorlopende aandacht van een breed scala aan stakeholders vanuit de wetenschap, industrie, regelgevende instanties en non-gouvernementele organisaties zoals dierenwelzijnsorganisaties, om ervoor te zorgen dat de noodzaak tot een paradigmaverschuiving breder postvat.



Figuur 2. De *'niche based track'* ter bevordering van de acceptatie en het gebruik van 3V-modellen

De procesmatige aanpak die via deze drie routes wordt geboden vereist een nauwe verwevenheid tussen de elementen op de drie niveaus. Focussen op slechts een paar van deze elementen houdt het risico in van *'Kurieren am Symptom'*. Dit leidt tot de ad hoc behandeling van de symptomen, zonder dat gewerkt wordt aan de onderliggende

oorzaken. In plaats daarvan is het van belang te denken in termen van een duurzame en gecoördineerde veranderstrategie om het moeizame acceptatieproces aan te pakken. De 4 C's van *'Commitment, Communication, Coöperation and Coördination'* (zie Schiffelers et al., 2014b: Hoofstuk 8) vormen de hoofdmotor om de afstemming binnen en tussen de drie routes mogelijk te maken. Bovendien wordt een vijfde C van *'Continuity'* toegevoegd aangezien 3V-acceptatie een lange adem vergt.

Een epistemologische vooronderstelling van de centrale stakeholders in dit onderzoeksterrein, die in dit onderzoek is blootgelegd, is dat de innovatie-entrepreneurs vooral aandacht schenken aan de productie, validatie en verspreiding van wetenschappelijke gegevens om de acceptatie van 3V-modellen te stimuleren. Veruit de meeste energie wordt gestoken in het omgaan met wetenschappelijke onzekerheden en het genereren en interpreteren van nieuwe testgegevens. Deze aanpak schiet echter duidelijk tekort als het gaat om de acceptatie en het gebruik van 3V-modellen in het regulatoire domein. Veel van de barrières zijn immers gelegen in aspecten die veeleer psychologisch, cultureel of politiek van aard zijn. Dit proefschrift bevat dan ook een pleidooi gericht aan de natuurwetenschappers om zich bij complexe besluitvormingsprocessen open te stellen voor andere disciplines, zoals de sociologie, de beleids- en innovatie-wetenschappen. Dit is van belang om het overheersende technocratische perspectief te verbreden en instrumentele rationaliteit in evenwicht te brengen met een waarderationaliteit. Met dit proefschrift wordt een inspanning geleverd om de verbinding tussen deze disciplines te maken voor het specifieke veld van 3V-acceptatie. Het verdient de aanbeveling om verder te gaan in de hier aangegeven richting en ervoor te zorgen dat de technologische ontwikkelingen niet plaatsvinden zonder deze in een breder sociaal wetenschappelijk perspectief te plaatsen.

Stakeholders met ervaring op het terrein van 3V-acceptatieprocessen zijn zich reeds sterk bewust van het feit dat deze sociaalwetenschappelijke aspecten een belangrijke rol spelen in het acceptatieproces. Zij worstelen echter met de vraag hoe met dit complexe vraagstuk, met zijn veelheid aan beïnvloedende factoren, om te gaan. In dit proefschrift zijn daartoe verschillende handvatten geboden. Daarbij is het cruciaal dat er sprake is van voortdurende communicatie tussen betrokkenen, rekening houdend met de diversiteit aan meningen, waarden en ervaren onzekerheden. Door stelselmatig aandacht te schenken aan de verschillende percepties ten aanzien van specifieke 3V-modellen, kunnen mogelijke controverses blootgelegd en aangepakt worden. Pas dan kunnen potentiële barrières tijdig opgespoord en aangekaart worden en kan progressie in de acceptatie en het gebruik van 3V-modellen in het regulatoire terrein bevorderd worden.

DANKWOORD

Dankwoord

Na de afronding van mijn milieukundescriptie over internationaal walvisvaartbeleid vroeg mijn begeleider, Professor Pieter Glasbergen, of ik het onderzoek naar common pool resources verder wilde uitwerking tot een proefschrift. Het onderwerp was heel interessant maar ik wilde eerst de wijde wereld in trekken en samen met Albert naar Zuid-Amerika te gaan om mijn onderzoeksobjecten in het wild te observeren. En toen ik vervolgens de kans kreeg bij Greenpeace walvissen te gaan beschermen, was het idee om te gaan promoveren van de baan. Maar de optie om te promoveren bleef in mijn hoofd bestaan 'voor als zich een mooi onderwerp zou aandienen'. Die kans deed zich voor toen in 2004 *'The Regulatory Animal Testing (RAT) Project Group'* met daarin Bas Blaauboer, Martje Fentener van Vlissingen, Coenraad Hendriksen, Marianne Kuil, René Remie, Joop Thuring en Manon Vaal, USBO advies vroeg onderzoek te doen naar de acceptatie en het gebruik van 3V modellen. Samen met Gerrit Hagelstein, Annemiek Harreman en Martijn van der Spek werkte ik aan het rapport *Regulatory Animal Testing* dat in 2007 leidde tot een wetenschappelijke publicatie in Altex (Schiffelers et al., 2007). Het artikel ontving in 2008 de Altex prijs voor publicatie van het jaar en vormde de opstap naar het promotietraject.

Op een zomerse dag in 2008 nodigde Coenraad Hendriksen me uit voor een drankje op het Ledig Erf en opperde de mogelijkheid het onderzoek uit te werken naar een promotietraject. Ik wist inmiddels genoeg van het onderwerp om te weten dat het een zeer interessant onderzoeksterrein is met genoeg complexiteit en uitdagingen om me meerdere jaren mee bezig te houden. En de gedachte een klein steentje bij te kunnen dragen aan het vervangen, verminderen en verfijnen van dierproeven maakte de motivatie compleet. Er volgde een periode van zoeken naar financiering. De oplossing diende zich aan toen Bas Blaauboer de Doerenkamp-Zbinden leerstoel "Alternatieven voor Dierproeven in de Toxicologische Risicobeoordeling" ging bekleden, waaruit een groot deel mijn onderzoek gefinancierd kon worden. Toen Wieger Bakker bereid bleek om de dagelijkse begeleiding vanuit USBO te verzorgen, stond er niets meer in de weg en kon het onderzoek begin 2009 starten.

Wat volgde was een mooi, leerzaam en uitdagend onderzoeksproces waarin ik als beleidswetenschapper in aanraking kwam met de wereld van de immunologie, toxicologie en risicobeoordeling. Een wondere wereld waarin ik me tijdens menig congres 'Alice in Wonderland' waande temidden van andere gebruiken, culturen en taal. Dit proefschrift is het product van de reis door deze wereld.

Ik wil iedereen die ik op deze reis heb leren kennen en die aan dit onderzoek, op welke wijze dan ook, een bijdrage heeft geleverd heel erg bedanken. De volgende mensen wil ik in het bijzonder bedanken. Te beginnen bij Coenraad, Bas en Wieger.

Coenraad, jouw vraag bracht me op het pad om me verder in dit onderwerp te verdiepen. Ik heb er nooit spijt van gehad. En ik ben erg blij met jou als één van mijn promotoren Je hebt me, met je jarenlange 3V-ervaring, telkens van zeer steekhoudende en opbouwende feedback voorzien en hield me scherp. Daarnaast kon ik iedere verjaardag rekenen op een e-card en heb ik vaak moeten lachen om je relativerende, no-nonsense dierenartsen humor.

Bas, door jou werd het financieel mogelijk dit project daadwerkelijk uit te voeren. Daarbij was je altijd een promotor in de letterlijke zin van het woord. Je hebt me menigmaal het podium geboden door me voor te dragen voor een presentatie of me te betrekken bij een radio-interview. Ook bracht je me in contact met belangrijke stakeholders in dit veld en sprak je telkens met veel enthousiasme over het belang van mijn onderzoek.

Wieger, aan jou heb ik veel te danken, zowel in termen van mentale steun, als in het aanscherpen van mijn betoog. Je scherpe analytische blik en je opbouwende feedback hebben me door menig lastig moment geholpen. Je was altijd bereid om met me van gedachten te wisselen over voorliggende vragen. Verder heb ik me binnen USBO, ondanks mijn exotische onderzoeksthema, altijd gesteund gevoeld. En daar heb jij grotendeels toe bijgedragen.

De leden van de leescommissie: Dr. Karl-Heinz Buchheit, Prof. dr. Wim Kremer, Prof. dr. Albert Meijer, Prof. dr. Aldert Piersma en Prof. dr. Ben van der Zeijst wil ik heel hartelijk danken voor het kritisch lezen en beoordelen van mijn manuscript.

De geïnterviewden, surveyrespondenten en panelleden wil ik hierbij nogmaals hartelijk danken voor het delen van hun kennis en ervaring, het geven van input en het meedenken over mogelijke oplossingen. Zonder jullie medewerking was dit proefschrift er niet gekomen.

Mijnheer Lütticken, u wil ik heel hartelijk bedanken voor de middagen die u vrij hebt willen maken om mij de basisprincipes van de immunologie bij te brengen. Het waren heel leerzame middagen, die me geholpen hebben de meer technische aspecten binnen de rabiës casus afdoende te kunnen begrijpen.

Marlies Halder en Lukas Bruckner wil ik bedanken voor het mij wegwijs maken in de rabiës vaccin casus, het toeleiden naar betrokken partijen en voor het kritisch lezen van conceptartikelen.

Aldert Piersma, jij hebt me wegwijs gemaakt in de EOGRTS casus en hebt ervoor gezorgd dat deuren geopend werden die anders gesloten waren gebleven. Ook voor je feedback op het conceptartikel ben ik je zeer erkentelijk.

Albert Meijer, nogmaals heel erg bedankt voor de tijd die je vrij wilde maken om het conceptmanuscript kritisch te bekijken en mij van gedegen en opbouwende feedback te voorzien. Het is de kwaliteit van het proefschrift zeker ten goede gekomen.

Cyrille Krul, ik heb het heel erg fijn gevonden om samen met jou aan de panels en aan het EOGRTS artikel te kunnen werken. Je was altijd een heel betrokken en behulpzame sparringpartner. En daarbij was het altijd leuk en motiverend om met je te werken.

Marianne Kuil, dank je voor je steun vanaf het begin van dit project. Je bent als één van de initiatiefnemers altijd heel erg geïnteresseerd geweest in de voortgang van het onderzoek.

Margo, Meggie, Marieke, Nynke, Sebastiaan, Aline en Eeke bedankt dat jullie tijd hebben willen vrijmaken om me voor te bereiden op mijn verdediging.

Paul, bedankt voor je geduld en steun. Hoewel ik altijd eerst en vooral adviseur ben gebleven, heeft het promotietraject natuurlijk veel tijd gevraagd. En jij hebt daar vanuit jouw rol als manager advies nooit moeilijk over gedaan. Daar ben ik je zeer dankbaar voor.

Mijn lieve kamergenoten Martijn, Gerolf, Mariska en Aline: Vooral in de laatste weken hebben jullie af en toe mijn stressuitingen moeten verduren. Door het aan te horen en mee te denken, hebben jullie me echt geholpen deze laatste loodjes te kunnen dragen.

Lieve vriendinnen Alexli, Brenda, Mieke, Lisette, Sjoer, Sas, Nicole en Jacqueline. Ik hoop jullie in de toekomst weer wat vaker te kunnen zien om samen van het leven te genieten.

Lieve Lientje, het was heel erg fijn om regelmatig samen in de bieb te werken aan onze proefschriften en tussendoor samen koffie te drinken. Het helpt om dat proces met een collega/vriendin te kunnen delen, die weet wat het inhoudt om te promoveren. Gelukkig blijven we als collega adviseurs veel met elkaar werken en kunnen we als vriendinnen regelmatig het leven blijven doornemen. Ik ben erg blij dat jij mijn paranimf wilt zijn.

Lieve Aleid, jij zou er eveneens als paranimf bij zijn. Je had je hier al heel lang op verheugd en was altijd enorm betrokken bij mijn onderzoek. Ik kon je bij nacht en ontij bellen.... Ik had nog heel veel vragen over het promoveren aan je willen voorleggen. Jij was immers als gepensioneerd dierenarts, doctor en ervaren paranimf een veteraan op dit vlak. Ik zal je missen op 17 juni. Je bent er in mijn gedachten bij en zoals beloofd: ik zal je ring dragen!

Lieve Eekie, ik had me natuurlijk geen betere 'nieuwe' paranimf kunnen wensen. Mijn grote wijze zus en mijn mantelzorg-maatje. Je hebt me in de laatste maanden van topdrukte heel veel (mantel)zorg uit handen genomen en ik heb je heel regelmatig 'lastiggevallen' om stoom af te blazen als zaken niet liepen zoals gewenst. Ik ben er trots op mijn grote zus achter me te hebben staan op 17 juni!

Mijn lieve broer Jan en zussen Eeke, Cara, Siene en Stans. Ik wil jullie bedanken voor de betrokkenheid bij het onderzoek van jullie kleine zusje. Ik ben intens blij dat jullie in mijn leven zijn!

Lieve Roosje en Finn. Jullie hebben regelmatig gevraagd wanneer ik klaar zou zijn. Nu is het dan echt zover! Ik hoop dat jullie niet al te veel last hebben gehad van het feit dat ik “zo graag” wilde promoveren. Ik beloof jullie om hierna niet meer altijd gelijk in slaap te vallen zodra ik met jullie een film ga kijken. En ik wil weer heel veel leuke dingen samen met jullie en met papa gaan doen.

Mijn grote liefde Albert. ‘Frau boktor’ is nu echt doctor. Je was al tijdens mijn scriptieperiode mijn steun en toeverlaat. En nu weer bij deze uit de hand gelopen scriptie. Promoveren gaat een stuk makkelijker als er thuis zoveel steun voor is. Ik ben je hier heel erg dankbaar voor en ik kijk ernaar uit meer tijd te hebben om samen te genieten van onze rozen, kikkers en vogeltjes. En om mooie dingen te ondernemen met Rosa en Finn.

Lieve papa en mama. Ik ben jullie enorm dankbaar voor het liefdevolle nest waar ik in ben opgegroeid en de stabiliteit die ik daardoor heb gekregen.
Ik draag dit proefschrift aan jullie op!

ABOUT THE AUTHOR

ABOUT THE AUTHOR

Name:	Marie-Jeanne Schiffelers
Date of birth:	02-10-1970
Place of birth:	Hoensbroek
Nationality:	Dutch
Function:	Senior consultant and researcher
Institution:	Utrecht University School of Governance
e-mail:	m.j.w.a.schiffelers@uu.nl

Marie-Jeanne Schiffelers studied Environmental Policy Sciences at Utrecht University. After finishing her masterthesis in 1995 she worked for three years as a biodiversity campaigner at Greenpeace Netherlands in Amsterdam. In 1998 she returned to her Alma mater and started to work for Utrecht University School of Governance as a consultant in the public domain. Currently, she works as a senior consultant and researcher in the field of policy science, organizational change and technology transitions and is contracted by a broad range of clients.[1]

She has been involved in several projects in the field of animal testing and 3R models e.g.:

- Project manager research 'Regulatory Animal Testing' by order of the Science Shop for Biology of Utrecht University (2004-2005)
- Member research team 'Evaluation Decree Biotechnology in Animals' (2005-2006)
- Member research team 'Impact Assessment for the revision of EU Directive 86/609/EEC on the protection of animals used for experimental and other scientific purposes' (2006-2007)

In 2008 she was rewarded with the ALTEX prize 2008 for the article: *Factors that Stimulate or Obstruct the Implementation of 3Rs in the Regulatory Process* (Schiffelers et al., 2007) which was the result of the researchproject on Regulatory Animal Testing. This project anticipated the PhD project on the Acceptance and Use of 3R Methods for Regulatory Purposes that she worked on from 2009 until 2016, next to her job as senior consultant.

Her work has been published in several international refereed journals such as Biologicals, Altex, Regulatory Toxicology and Pharmacology and the European Journal on Risk Regulation.

1 e.g.: The European Commission, the Dutch Ministry of Health, The Dutch Ministry of Security and Justice, The Dutch Minstry of Agriculture/Economic Affairs, The Dutch Council for the Judiciary, The Court of Law Midden Nederland, The Netherlands School of Public & Occupational Health and a range of Dutch municipalities and provinces.