

An Overview of the Development of the Welfare Quality[®] Project Assessment Systems

edited by
Linda Keeling



An Overview of the Development of the Welfare Quality[®] Assessment Systems

edited by

Linda Keeling



WELFARE QUALITY REPORTS No. 12

Welfare Quality Reports
Edited by Mara Miele and Joek Roex

School of City and Regional Planning
Cardiff University
contact: LeoR@cardiff.ac.uk

September 2009

The present study is part of the Welfare Quality® research project, which has been co-financed by the European Commission, within the 6th Framework Programme, contract No. FOOD-CT-2004-506508. The text represents the authors' views and does not necessarily represent a position of the Commission who will not be liable for the use made of such information.

© Copyright is held by the authors of the individual contributions.

Typeset by Joek Roex, Cardiff.
Printed in Sweden by SLU Service/Reproenheten, Uppsala.

ISBN 1-902647-82-3
ISSN 1749-5164

CONTENTS

	<i>Preface</i>	<i>iii</i>
	<i>List of Contributors</i>	<i>v</i>
1	Definition of a Framework for Developing Assessment Systems	1
2	Investigating Possible Measures to Include in the Assessment Systems	9
3	Scoring Animal Welfare: Difficulties and Welfare Quality® Solutions	15
4	Testing the Assessment System: Refinement and Definition	33
5	Investigating Societal Values on Farm Animal Welfare: The Example of Welfare Quality®	43
6	Practical Welfare Improvement Strategies	57
7	Preparing the Assessment System for Implementation	67
8	An Overview of the Application of the Animal Welfare Assessment System in Latin America	79
	Bibliography	91
	Appendix	95

PREFACE

Nowadays, the perception of food quality is not only determined by its overall nature and safety but also by the welfare status of the animal from which it was produced. In other words, animal welfare is an important attribute of an overall ‘food quality concept’.

A main thrust of Welfare Quality® is to underpin the animal welfare attribute of food quality and to allow transparent communication on this issue within the supply chain and to the consumer. Thus Welfare Quality® set out to develop an integrated standardised methodology for the assessment of welfare in cattle, pigs and poultry from farm to slaughter and to determine the way in which the measurements should best be integrated.

Sub Project 2 ‘Monitoring welfare’ was central in this approach. In SP2, the most appropriate specialists and experts were gathered and the task was approached in a truly international collaborative effort. In collaboration with social scientists analyses of consumer/citizen perceptions and attitudes were combined with existing knowledge from animal welfare science and thereby 12 areas of concern were identified that should be adequately covered in the measurement systems. Within SP2 scientists then developed, refined, standardised and validated welfare measures under these areas of concern.

It was generally agreed to concentrate on so-called performance measures that are based on measuring the actual welfare state of the animals in terms of, for instance, their behaviour, fearfulness, health or physical condition. Such animal-based measures include the effects of variations in the way the farming system is managed (role of the farmer) as well as specific system-animal interactions. The practical application and feasibility of the measures was also studied in SP2 and they are described in Welfare Quality® Technical Documents for the different species.

SP2 also developed a standardised procedure to summarize the results of the measurements into an overall animal welfare score, which reflects the welfare quality of a particular farm or slaughter facility. Again in a multidisciplinary and collaborative effort, this evaluation model was tuned according to experts from animal and social sciences, and stakeholders in the agricultural sector. This procedure is also described in the above mentioned Technical Documents.

The underlying science and the research outcomes of SP2 are being published in international scientific journals. However, by nature such publications are covering only separate aspects or measures and it was felt an integral overview of SP2’s activities was essential to grasp the whole picture. Moreover, the accessibility of scientific literature for a more general readership is not always optimal.

I am therefore very happy with the publication of this issue in the Welfare Quality® series, as it summarizes not only the science and the outcomes but also the underlying process and reasoning and is accessible for a broad audience.

I would like to thank all contributors for this additional effort!

Harry Blokhuis
Coordinator Welfare Quality®

LIST OF CONTRIBUTORS

R. Botreau

INRA, Saint-Genès Champanelle, France

A. Butterworth

University of Bristol, United Kingdom

B. Engel

Wageningen University and Research Centre, Netherlands

B. Forkman

University of Copenhagen, Denmark

B. Jones

Roslin Institute, United Kingdom

L. Keeling

Swedish University of Agricultural Sciences, Uppsala, Sweden

U. Kjærnes

SIFO, Oslo, Norway

X. Manteca

Universitat Autònoma de Barcelona, Spain

M. Miele

Cardiff University, United Kingdom

P. Perny

Pierre et Marie Curie University, Paris, France

C.G. van Reenen

Wageningen University and Research Centre, the Netherlands

Isabelle Veissier

INRA, Saint-Genès Champanelle, France

DEFINING A FRAMEWORK FOR DEVELOPING ASSESSMENT SYSTEMS

Linda Keeling

1.1 INTRODUCTION

It is no small task to decide on the framework to be used when developing a monitoring system, and certainly not when that monitoring system is to be one of the main deliverables of the Welfare Quality® project. Of course it was not a blank slate at the start of this first Work Package in Sub Project 2. The Welfare Quality® project followed on closely from an earlier EU funded network called ‘Measuring and monitoring farm animal welfare’. Many partners had been involved in that network and many had experience of other monitoring systems. Nevertheless the difference now was that such a large number of researchers needed to reach consensus on decisions regarding the framework. Meeting this challenge has become one of the strengths of the Welfare Quality® project. A preliminary version of the framework developed in this Work Package was later evaluated by social scientists in Sub Project 1. It was also presented to the Advisory Committee, a collection of representatives from various stakeholder groups.

Formally there were two objectives to Work Package 2.1. The first was to propose a rationale for developing monitoring systems. In practice this meant reaching consensus on the basic underlying principles regarding the assessment of animal welfare on farms and identifying potential measures to be included. The second was to decide on which measures were most promising to be developed and tested in the following Work Package (WP2.2) and to decide on the general methodological principles to be applied when these measures are being refined.

1.2 SCOPE OF THE MONITORING SYSTEM

It was already decided that the project would focus on cattle, pigs and poultry, and one of the first decisions was to concentrate on the productive phase of the most commonly occurring categories of these species. That is to say, the decision was taken to focus on dairy cattle, beef cattle and veal calves, sows with piglets, fattening pigs, laying hens and broilers. Thus the original 3 species became 7 categories of animals. It was decided at this stage not to include dairy calves or heifers and not to include laying hens during the rearing period. Likewise, it was decided that welfare at slaughter should concentrate on beef cattle, slaughter pigs and broiler chickens i.e. those animal categories where slaughter is a key part of the food chain process. Practical considerations lay behind these decisions to limit the scope, based on the view that once the framework and principles were established, these areas were those that could most easily be addressed using knowledge developed in the other animal categories. For example, many of the measures developed for veal calves could be used for dairy calves, and measures for the rearing of broilers could be used also for rearing laying hens. At slaughter, measures developed for assessing welfare during transport and slaughter for beef cattle could also be used for dairy cows and so on.

Another aspect of the scope is the type of variables that can be used in welfare assessment. They can generally be divided into resource-based measures (e.g. space allowance, type of floor etc), management-based measures (biosecurity measures, feeding times etc.) and animal-based measures (e.g. injuries, fear etc). The first two are generally regarded as 'input' measures and the last as an 'output' measure. The decision to use as much possible animal-based measures for welfare assessment had already been taken, but management- and resource-based measures were still regarded as important for identifying causes of poor welfare, risk to welfare or as substitute measures in cases where no valid animal-based measures were available. A further consideration was the desired ability to compare the welfare of animals kept in different systems; which again argued strongly for the use of animal-based measures.

This emphasis on animal-based measures was counter to most common assessment schemes at the time which used mainly resource-based measures. These are parameters that can easily be observed in that they require comparatively little training of the assessor and usually have high inter and intra-observer repeatability. They are also typically parameters used for legislation. While resources determine the physical situation for the animal, management-based parameters are also important. Major decisions regarding the animal's life; how and when the animal is fed, moved and mixed with other animals are management decisions. But even apparently minor differences in the way the animal is handled are known to affect welfare. Some management factors can be determined from farm records (if available) or interviews, but best would be to observe or video film the farmer, something that is obviously difficult to do in a feasible way. Nevertheless, welfare is a characteristic of the animal and animals differ in their genetics, early experience and temperament and therefore may experience the same environment in different ways. This

view and the decision to base the welfare assessment on animal-based measures as much as possible was confirmed and even strengthened in these early discussions.

Parallel to the discussion on which categories of animals and what types of measure should be included, there was a discussion on the philosophical background of what constitutes a good life for animals. Although one might imagine that this was the most problematic issue – it was not. In the scientific literature the emphasis has been on the mental states of animals, the biological functioning of animals and the naturalness of the environment and behaviour. It was soon decided that Welfare Quality® did not need to decide upon a single ethical approach or a single definition of welfare since if the final monitoring system was to be widely accepted it was important to take account of all these different approaches. Moreover there was general agreement that welfare is a multidimensional concept comprising both physical and mental health and that it includes a wide range of aspects such as physical comfort, absence of hunger and disease, possibilities to perform motivated behaviour and so on.

1.3 IDENTIFICATION OF MEASURES

Tables of potential measures were produced for each category of animal for use on farm and/or at slaughter and classified according to whether they were animal-based, management-based or resource-based. The tables included references to any literature as well as a classification according to information on the validity of the measure, its reliability and its feasibility. Explanatory information was given for what additional research was thought to be necessary and a suggestion of whether this measure had large, moderate or no merits for potential inclusion in future monitoring systems. These tables therefore document all measures that were discussed, even those considered to have little potential, as a form of ‘memory’ of the selection process. In addition literature reviews summarised the previous work in the area as further documentation of the selection process.

Although initially the total numbers of measures identified and discussed was large, following the evaluations of the validity, reliability and feasibility of these measures, it soon became apparent that there was an insufficient number of well developed animal-based measures for the different species to address the different welfare issues. These tables were the basis for identifying where more research was needed in WP2.2. This evaluation process therefore not only identified the absence of any measure for certain areas of welfare concern, but also highlighted that some measures were being used despite the fact that they had been neither validated, nor tested for reliability under commercial conditions.

Most noticeable was the lack of measures reflecting the mental states of the animals. Two other examples of problems related to the identification of measures are presented in the following paragraphs.

Abnormal behaviour, such as stereotyped behaviour or injurious behaviour, is generally agreed to have high validity as an indicator of poor welfare. There are several different measures that can potentially be used to monitor them e.g. number of animals performing the abnormal behaviour, the proportion of time spent performing them etc. But these measures can be difficult to take reliably, since animals often interrupt their abnormal behaviour when disturbed, by for example a person coming to observe them, unless the animal is habituated over a period of time and that reduces the feasibility of the measure. One way forward therefore was to find other indicators of abnormal behaviour that did not have these methodological problems. For injurious behaviour such an alternative can be to look for indicators that an animal is the recipient/target of the injurious behaviour. For example, evidence of feather pecking damage on the plumage of birds is an indicator that feather pecking is present in that flock. Likewise, evidence of bitten tails is evidence of tail biting behaviour in the group. But this is not so easy for other types of abnormal behaviour, such as stereotypic behaviour.

The opposite situation was also found. For some welfare problems there were several different measures considered valid and reliable and reported in the scientific literature, although not necessarily developed for in all species. For example, there are several lameness scoring systems for broilers and cattle, although few for pigs. For scoring of scratches and wounds on the body however, despite a wealth of measures reported in the literature, closer examination reviewed that none had been tested for validity or reliability.

1.4 DEVELOPING A FRAMEWORK

The best known operational list of welfare dimensions is the Five Freedoms (proposed by the Farm Animal Welfare Council in the UK) although other definitions have also formed the basis of monitoring systems. The difficulty in developing a framework is, as discussed previously, that many individuals and groups of individuals have developed their own personal views on what is and what is not important. Examples of areas of concern highlighted by different groups of stakeholders are that the animal should: not suffer, be healthy, be physically comfortable, and be able to show natural behaviour. However, on closer examination it is clear that within each area of concern there are many detailed aspects that need to be addressed. For example, issues around health obviously include physical injuries, the presence of disease and pathologies and perhaps even what management steps are taken to reduce the risks. Suffering on the other hand is a complex concept embracing an animal's varied emotional states that are often difficult to measure. Even the possibility to perform natural behaviour can be interpreted in different ways;

indeed the question of what is natural in an artificial, commercial farm environment is open to debate.

The work on defining the framework for Welfare Quality® was developed in collaboration with WP2.3 and the final list of 12 areas of concern resulting from the Welfare Quality® project is presented in more detail in chapter 3. The final list of 12 criteria was not however the first list developed in the Welfare Quality® project. In WP2.1 the original list consisted of 10 areas of concern and is mentioned here because it was this list that was used by the social scientists in Sub Project 1 in their focus groups discussions referred to in chapter 5 of this report. The measures identified in WP2.1 were allocated to these different area of concern/ criteria and separate tables were made according whether the measure was intended to be taken on farm or at slaughter. These tables were further discussed with key people in other Sub Projects, not least Sub Project1 and they were also sent to the stakeholder representatives who were members of the Welfare Quality® Advisory Committee at the time.

The 10 areas of concern, originally identified were:

1. hunger, thirst or malnutrition;
2. physical comfort and security;
3. health: injuries;
4. health: disease;
5. pain (not related to injuries or disease);
6. normal / natural social behaviours;
7. normal / natural other behaviours;
8. human–animal relationship;
9. negative emotions (apart from pain);
10. positive emotions.

The 12 criteria that are the final output of Welfare Quality® are:

1. absence of prolonged hunger;
2. absence of prolonged thirst;
3. comfort around resting;
4. thermal comfort;
5. ease of movement;
6. absence of injuries;
7. absence of disease;
8. absence of pain induced by management procedures;
9. expression of social behaviours;
10. expression of other behaviours;
11. good human–animal relationship;
12. positive emotional state.

The two lists are very similar and the revision from 10 areas of concern to 12 criteria is the consequence of the ‘bottom up’ approach taken in WP2.1 (where measures were

identified) meeting the ‘top down’ approach taken in WP2.3 (where integration methods were developed). The changes were to maximise the likelihood of there being at least one animal-based measure in each criterion.

1.5 METHODOLOGICAL CONSIDERATION

A central aspect of the Welfare Quality® monitoring system that it is science based. A final part of WP2.1 was to decide on the general methodological principles to be applied when the potential measures were to be refined in WP2.2.

Validity is the extent to which a measure actually measures what it is suppose to measure. It is obvious that measures included in the monitoring system have to say something about the welfare of the animal. A large step towards improving validity had already been taken with the decision to use animal-based measures as much as possible. The main focus for this work on general methodological principles therefore focussed on the reliability of the measures. Reliability was divided into intra-observer reliability (how similar are observation by the same observer on the same animals) and inter-observer reliability (how similar are observation by different observers on the same animals). Recommendations of how to test this, under farm conditions, under experimental conditions using real animals, or using photographs and video clips were formulated, along with the advantages and disadvantages of each method. For example, comparisons of live animals under commercial conditions would be the best, but there can be difficulties in getting enough animals representing the full range of scores. Plus if the animals are to be in the same welfare state, they should ideally be evaluated twice on the same day and there is the risk that the observer may ‘remember’ the previous score given to a specific animal. Looking at video clips or photos provides more control but, depending on the measure, it may make the evaluation more difficult or easier than it would be in practice. An instruction document was developed outlining how to check for reliability so that measures were tested in a comparable way. More information on how the measures were validated and tested for reliability is given in Chapters 2 and 4.

1.6 CONCLUSION

This was a small Work Package compared to the other ones in this Sub project since it consisted only of reviews of the literature and meetings. However as the name suggests it aimed to define the framework, that is to say the basic structure and methodology that was

to be used in the remainder of the Work Package. It also made a significant contribution to one of the other main goals of Welfare Quality®, that is ‘to integrate and interrelate the most appropriate specialist expertise in the multidisciplinary field of animal welfare in Europe and also link to expertise elsewhere.’ The meeting and discussions organised by the partners were attended by over 40 animal and social scientists in the project as well as by selected external experts.

ACKNOWLEDGEMENTS

Thanks are given to Björn Forkman, Kees van Reenen, Cecile Arnould, Knut Bøe, Antonio Velarde and Christoph Winckler who were responsible for tasks within this Work Package and on whose deliverables this chapter is based.

INVESTIGATING POSSIBLE MEASURES TO INCLUDE IN THE ASSESSMENT SYSTEMS

Björn Forkman

2.1 BACKGROUND

The goal of Work Package 2.2. was to investigate possible measures to include in the welfare monitoring scheme of Welfare Quality[®]. The candidate measures were first proposed in the COST Action 486 and later refined and further discussed in Work Package 2.1.

2.2 MEASURING ANIMAL WELFARE

Any system of monitoring animal welfare eventually comes up against the problem of how to define welfare. In some cases this definition is not explicitly stated but can be deduced from the variables measured, or alternatively the definition used may be so broad as to be of little use (Ingemann et al., in press).

In the case of Welfare Quality[®] an operational approach was taken with four areas of concern being identified (good feeding, good housing, good health and appropriate behaviour). These four areas were further divided into 12 criteria in all, good feeding was e.g. divided into absence of prolonged hunger and absence of thirst etc. During the introductory discussions of the project it became apparent that most scientists of the project had a view of animal welfare that is closely related to that proposed by Duncan and Petherick (1991). In this view of animal welfare the emotions, or the subjective experiences of the animal, are considered to be of prime importance. This accorded well with the reasoning in Cost 846, which placed a heavy emphasis on animal based measures. Alternative concepts such as the naturalness of the behaviour and/or environment were

accorded less importance. It was argued that no variable that could not be detected through the behaviour of the animal should be accorded any importance and thus should not be included in the monitoring scheme.

In most monitoring systems of animal welfare the emphasis is on the avoidance of bad welfare. Welfare Quality® is one of the first assessment schemes to try and include measures of positive welfare, i.e. measures of positive emotions.

Concurrently with WP2.2 social scientists within Welfare Quality® made a number of studies of how members of the public viewed the definitions used by the scientists of Welfare Quality® (see books 9–11 in this series). While they were mostly favourable to the definition used the one area of concern which they did not feel that the scientists had taken seriously enough was the one of naturalness. Luckily many of the aspects of naturalness can be captured by resource based measures (e.g. access to daylight, access to outdoor run etc.) and it was therefore possible to relatively easily include this aspect.

2.3 CHOOSING THE MEASURES

The aim of the project was to develop an assessment scheme for the three most important farm animal species; pigs, poultry and cattle. However each of these species exist in very different systems under different conditions. In the current project we decided to include, in poultry: both layers and broilers, in pigs: both sows and fattening pigs, and finally in cattle: dairy, fattening cattle and veal calves. It is clear that this is not a complete list, even for these species (missing are e.g. layer pullets, broiler breeders, boars, dairy calves etc).

However, our aim was to evaluate the welfare of these animals as well as possible and therefore include both measures on farm and at slaughter. With only one measure per criteria this would mean that we would need to develop approximately 170 different measures. There was some discussion on whether it was possible to develop measures for use on transport as well, however the decision was taken that it in many cases would be possible to gather information concerning the transport partly by resource and management measures and partly by evaluation of the welfare of the animals when they reached the slaughter plant.

Because of the huge number of measures it was decided to focus on the largest animal groups, alternatively those with arguably the greatest animal welfare problems (e.g. older rather than younger animals). A decision was also taken, as far as possible, to focus on measures that were well known and had been used previously. The most important exception to only using well known measures was for positive emotions. No such validated measures for on farm animal assessment were known at this point, and the groups investigating these measures had to develop and validate them ‘from the beginning’.

As previously stated it was decided to use animal based measures wherever possible, partly because of the definition of animal welfare, but also because this could help overcome one of the major problems of animal welfare assessment when using resource based measures; how to compare good and bad farmers within the same system. An animal based approach means that the measure of animal welfare is relatively independent of the farming system.

In practice, however, different approaches to e.g. assessing the human-animal relation is needed dependent on the system. A forced approach test to a crated sow will need to use different measurements than that of a forced choice approach test to a loose housed sow, likewise for e.g. tethered and loose housed dairy or loose housed layers and layers in cages. If each test or measure was 'perfect', i.e. it only tested what it was developed for and was not affected by anything else it should be possible to titrate the responses in one situation to that of another so that it would be possible to say e.g. whether the human animal relation is the same in two different systems. An attempt to do this titration was done for layers in cage and loose house systems (Graml et al., 2009), but for most measures this ended up as a theoretical possibility.

An alternative and easier but more subjective, approach to this problem is that instead of titrating the behaviours of the animals, the scientists themselves can do the "translation", e.g. by stating what an acceptable avoidance distance is in caged birds, and what an acceptable avoidance distance in loose housed birds is. However when more "cases" are included and evaluated separately, the system becomes less stringent and it is harder to compare the different production systems. Fortunately this problem is less for some measurements than for others, a panting pig is probably too warm independently of the system it is in.

2.4 TYPES OF STUDIES

There were three main types of studies within this work package. Validation studies were done for the comparatively few criteria for which there was no applicable measure (e.g. measures of positive emotions), repeatability studies were done for those measurements that were considered valid (e.g. abnormal behaviour), and finally standardisation studies were done for well known measures, and for which there already existed one or several protocols (e.g. lameness/gait scoring in broilers).

No formal criteria was employed when deciding whether a measure had previously been validated or not. A combination of face validity and knowledge of the literature decided which measures to be assigned to each type of study.

2.4.1 Validation

Various different approaches were used to study the validity of a given measure. Some studies used knowledge of the effect of environmental variables. In investigating the resting behaviour of sows (for the criterion "comfort around resting") Scott et al. (2009) investigated the resting behaviour of sows on different surfaces. The assumption was that sows on fully slatted floors have less comfort around resting than those on part slatted floors, or those on deep straw flooring (this assumption rests on previous studies reported in the literature). Scott et al. then proceeded to see what the effect of these floor types was on the frequency of different resting postures and posture changes, as well as on the frequency of bursitits.

Other studies correlated measures within the same situation to validate the new measures. So was e.g. vocalizations investigated as a possible measure of hunger and/or aggression around feeding in sows (Špinka et al, 2009).

Other ways of validating a test, by using e.g. pharmaca were much less commonly used in Welfare Quality. An exception is the study by Velarde et al. (2009), who investigated measures of general fear. This approach was based on the use of anxiolytics to validate a given behavioural reaction as a measure of fear.

2.4.2 Repeatability

Measures not only need be valid, they also need to show a high degree of repeatability if they are to be used in different welfare assessment systems. It soon became apparent however that there are problems even with such seemingly simple aspects. There was a general agreement that interobserver repeatability, as well as short term repeatability (e.g. repeatability within the same day), was very important, not least because behavioural measures are often seen as being subjective and non-repeatable.

The issue of long-term repeatability is more controversial. Some measures will clearly change over time and as the animals grow older. Lameness scores in broilers is a good example, here attempting to do a long term repeatability study will show changes in lameness scores as the birds grow older. In this case the repeatability issue may be solved by measuring several consecutive batches at the same age. (Always supposing that there is very little variation between the batches when they are hatched.)

Some other measures vary over time in a more unpredictable manner. Most would consider temperature to be an "objective" measure with a high repeatability, but this clearly changes both over the day and months. This might appear self-evident but other cases are less so, e.g. the level of cleanliness of the animals will change depending on the temperature, or when animals have access to an outdoor area depending on the precipitation/muddiness of the outdoor area.

The result has been that short-term repeatability has been taken into account in all measures, whereas the importance of long term repeatability varies between measures.

2.4.3 Standardisation

The final type of animal based measure was one for which we believed that there existed one or several generally accepted measures, and the question for the scientists of Welfare Quality was to decide on which measure to use. Examples for this type of measure are condition scoring in cattle and lameness in broilers. Because of the nature of the welfare assessment scheme the practicality of the individual measures was given a higher importance than would be the case in a scientific study.

2.4.4 Resource and Management

Apart from the animal based measures some resource based measures and management based measures were used. The selection of these was based on the scientific literature and the knowledge of the participants in the groups. No attempt was made to investigate these measures further.

2.5 THE RESEARCH TEAMS

An application procedure in which each scientist could apply for a given a research project for a measure or cluster of measures was employed. The experts in the field were assigned as project leaders, however these same experts have often previously investigated and developed measures within their fields of expertise. To get a second point of view each project also had a second expert from another research institution.

The money was allocated depending on the type of project (validation, repeatability or standardisation). Because of the large number of measures and the ambition to have a second opinion of each measure the amount of money allocated to each project was comparatively small. Something that posed problems for especially those researchers that were asked to develop new methods.

2.6 THE RESULTS

The results of the various projects have been presented elsewhere (see Books 9–11 in this series).

2.7 LESSONS WE HAVE LEARNED

Although a large number of different measures were originally classified as well known and assigned to only be standardized it became apparent that there was a lack of good repeatability studies for many of them. Another general finding was that no good measures (valid and repeatable) could be found for the positive emotions. Clearly this is an area that needs further research if positive emotions are to be part of future animal welfare assessment schemes.

2.8 ACKNOWLEDGEMENT

I would like to acknowledge all the help I have had in coordinating this work package from the respective species leaders: cattle – Christoph Winckler, pigs – Antonio Velarde, poultry – Andy Butterworth, and last but not least the help of Linda Keeling.

SCORING ANIMAL WELFARE: DIFFICULTIES AND WELFARE QUALITY[®] SOLUTIONS

Isabelle Veissier, Raphaelle Botreau and Patrice Perny

One objective of the Welfare Quality[®] project is to propose a standardised method for the information on animal welfare to customers or consumers. The need for such unified scoring system came from the evidence that:

1. Animal welfare remains an important concern for EU citizens (European Commission, 2005, 2007a, 2007b; Evans and Miele, 2007) and although you can't fatten a pig by weighing it, an information system, possibly translated into a label, may contribute to raising welfare standard (Polten, 2007);
2. EU consumers feel not enough informed on the welfare of animal and cannot take animal welfare into account appropriately when purchasing food (European Commission, 2007a, 2007b);
3. Across Europe, welfare claims are often used on products from animals and several schemes have been put in place to guarantee animal welfare to consumers (e.g. Freedom food scheme in the UK, IKB by the meat industry in the Netherlands). These schemes can differ in the measures used to check animal welfare, in the thresholds set to differentiate high vs. poor welfare, or in the way the information is integrated to form an overall judgement. A unified scoring system seems necessary to ensure credibility of welfare claims;
4. The EU commission, in its White Paper to the parliament (European Commission, 2002) launched the idea of a unified labelling system that could be used for bilateral negotiations between countries. Again this requires a unified scoring system.

Welfare Quality[®] is building a welfare assessment system incorporating numerous potential measures based on the animals, the resources or the management of animal units (farms or slaughter plants). The assessment system could then be used to determine the welfare status of cattle, pigs and poultry on animal units. This would generate a substantial amount of data as well as the need to integrate these data into an overall evaluation of the animal unit. Therefore, in order to facilitate the implementation of the assessment system, Welfare Quality[®] designed a scoring model to integrate these data and to translate value judgements into refined and easily understandable information that could help the decisions made by stakeholders with regard to animal welfare.

Within the time constraints of Welfare Quality®, scoring models were designed for the welfare of dairy cows, fattening cattle, veal calves, fattening pigs, and broilers during the time they are on farms (i.e. excluding transport and slaughter except when measures taken at slaughter can be used to assess the welfare on the farm).

3.1 DIFFICULTIES

The first difficulty came from the fact that welfare is a multidimensional concept. It comprises both physical and mental health and includes several aspects such as physical comfort, absence of hunger, disease, possibilities to perform motivated behaviour, etc. Making an overall assessment of a farm from numerous measures taken on that farm can be considered as adding apples and pears, that is things which are very different in nature are put together. If not performed appropriately, such integration may well serve to hide welfare problems.

The second main difficulty we faced was that on the one hand data were produced to describe farms (e.g. on farm A, 2% cows are very lean, 5% are lame, 50% cows flee when approached etc.) and on the other hand a judgement had to be made from such data. And a judgement cannot be value free. This problem becomes even more crucial when not only a judgement has to be made from a single measure (e.g. is 2% lean cows a good or a bad result?) but also on a group of measures (e.g. is 2% lean cows + 5% lame cows + 15% fearful cows etc. a good result?). We definitively faced ethical dilemmas and technical difficulties when building a model for the overall assessment of the welfare of animals.

3.1.1 Ethical Dilemmas

The welfare of an animal is a matter of how an animal experiences its life. When a single aspect of welfare is considered, the animal's point of view could be obtained, at least when short term preferences are considered (e.g. demand curves to know what food is preferred). It seems merely impossible to determine how an animal would rank very different aspects of welfare (e.g. being afraid of something and being sick). The assessment of an animal's welfare is thus to some extent done from a human point of view.

Already at the level of the individual animals, ethical decisions have to be taken as to how to 'add-up' the different aspects of welfare into a single value. First, some people may find some aspects more important than others while others would do the opposite, e.g. some may rank health higher than behaviour while others can consider that behaviour is more important (Fraser, 1995). Second, some may allow compensation between welfare aspects while others may not. In addition, such compensation may not be permitted when one aspect of welfare is extremely low while it is permitted above a certain threshold.

A similar but even greater problem arises when one wants to consider the welfare of groups of animals. Should we concentrate on the animals that are in the poorest conditions or consider the 'average' welfare across animals? At that stage it is unrealistic to ask a group of animals to judge their overall welfare, using e.g. demand curves. Again, the assessment of the overall welfare of a group of animals is inevitably based on human points of views.

At the same time, when doing an overall assessment of a farm, some may consider that the assessment must be based on what corresponds to normal practices (where normal means the most common). The results obtained by a farm may thus be compared to those of farms from the same population (Whay et al., 2003). By contrast, some fight against the idea that simply doing a little bit better is enough to obtain good welfare (Bekoff, 2008). According to such views, the assessment of welfare shall be based on what is theoretically considered excellent, good etc.

3.1.2 Technical Difficulties

The main technical problem we faced was that data could be collected on very different scales:

- Some data were expressed on a nominal scale, e.g. the method used to dehorn young cattle (cauterisation by hot iron or by caustic paste);
- Some data were expressed on an ordinal scale, that is they were expressed as ordered categories, e.g. piglets can have a normal weight, be lean, or even very lean. In that case one does not know if the difference between normal and lean is similar to the difference between lean and very lean;
- Some data were expressed on cardinal scales, e.g. the distance to which an animal can be approached.

Nevertheless, even in the case of data recorded on a cardinal scale, the interpretation may not be linear. For instance, it is likely that one would not consider that a cow that flees when approached by 3.5 m differs from a cow which can be approached by 3 m while this person can value very differently a cow approached by 0.5 m and a cow touched by the observer (i.e. approached by 0 m), although the difference in distance is exactly the same (0.5 m).

All observations need thus recoding on a common value scale, expressing the level of welfare of the animals.

Ethical decisions need be reflected in the scoring system so we had to design a model that would be flexible enough to model accurately such decisions. Mathematical models and methods have been developed in decision theory to deal with multicriteria evaluation problems. The great variability of problems encountered in practice has led scientists with different backgrounds (management sciences, mathematical psychology, economics, operational research and computer sciences) to develop a variety of formal models and methodologies to support evaluation tasks and decision making activities, e.g. outranking

approach based on ordinal aggregation methods, Multiattribute Utility Theory (MAUT) based on cardinal aggregation methods (additive or non-additive, e.g. Choquet Integrals). These works concern different important issues relevant to multidimensional evaluation such as measurement problems (how to represent preferences, perceptions or performances by numerical information), aggregation problems (how to derive overall evaluations from multidimensional and possibly conflicting viewpoints), uncertainty and imprecision modelling (Bouyssou, 1990; Roy, 1993; Bouyssou et al., 2001). All this material aims at facilitating the scientific preparation of decisions and can be useful at any stage of a decision process. In particular, the contribution of decision theory to real practice is to provide:

- a formal background for problem structuring tasks;
- a body of theoretical results for the mathematical justification of evaluation procedures;
- a set of procedures for the elaboration of recommendations in different types of problems (description, choice, sorting, or ranking).

Welfare Quality® investigated these methods in order to produce reliable methods for the overall evaluation of animal welfare in order to support decisions in this domain.

3.2 ARCHITECTURE OF THE WELFARE QUALITY® SCORING MODEL

We built on the results of the stakeholder surveys (consumers) and discussions among Welfare Quality® scientists to define a set of 4 main principles that must be safeguarded in order to ensure good welfare:

- Good feeding: Are the animals properly fed and supplied with water?
- Good housing: Are the animals appropriately housed?
- Good health: Are the animals maintained in good health?
- Appropriate behaviour: Does the behaviour of the animals reflect optimised emotional states?

Each of these principles is subdivided into 2 to 4 criteria which in turn must be met to ensure good welfare (Table 3.1).

In turn, each criterion is assessed by recording one or more measures that reflect various aspects of that criterion and/or the consequences of the associated problem. For instance, the absence of disease is checked by recording the incidence and severity of symptoms

TABLE 3.1 Welfare principles and criteria.

Welfare principles	Welfare criteria
Good feeding	1. Absence of prolonged hunger 2. Absence of prolonged thirst
Good housing	3. Comfort around resting 4. Thermal comfort 5. Ease of movement
Good health	6. Absence of injuries 7. Absence of disease 8. Absence of pain induced by management procedures
Appropriate behaviour	9. Expression of social behaviours 10. Expression of other behaviours 11. Good human-animal relationship 12. Positive emotional state

while the comfort of the resting area for dairy cows is checked by observing the ease of lying down, postures adopted when lying down, and cleanliness of the animals.

While welfare principles and criteria have been chosen thanks to a top-bottom approach (defining first principles, then criteria, and finally deciding on measures to be used to check the compliance to criteria), the achievement of an overall assessment of animal welfare on a particular animal unit follows a bottom-up approach (Figure 3.1). First, criterion-scores are calculated from the results of an animal unit for the various measures, then criterion-scores are calculated and finally the overall assessment is produced.

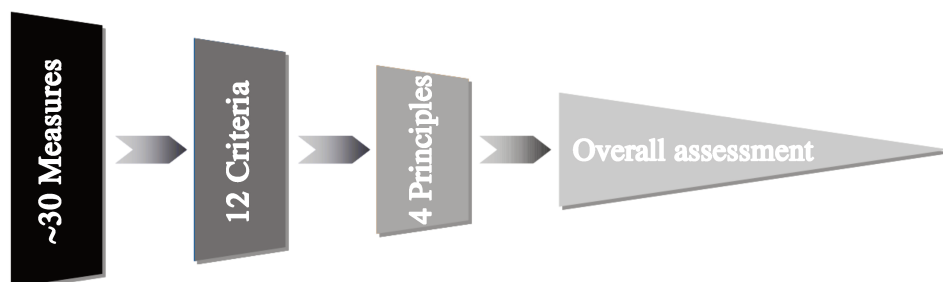


Figure 3.1 Bottom-up approach chosen in Welfare Quality® to produce an overall assessment of animal welfare.

3.3 THE CONSULTATION PROCESS

We, Welfare Quality® partners responsible of the development of the scoring models, felt that we did not have the right to take ourselves ethical decisions such as the ones listed above. We thus consulted people outside our group to guide the construction of the scoring model.

3.3.1 Consultation during the Construction of the Scoring Model

We developed a model using methods that left open the possibility to take one ethical decision or its opposite (e.g. allowing or not compensation between principles or criteria, taking into account the animals in the worst conditions or all animals). During the construction of the model, we consulted a number of people and adjusted the model according to their opinion. Several groups of people were consulted:

- for the first step of the scoring model, that goes from the data obtained by measures to criterion-scores, we consulted the animal scientists Welfare Quality® who had worked on the measures because we considered they were the people who knew the most about the meaning of these measures in terms of welfare for the animals and also they had some experience of what kind of results can be found in practice; between four to eight scientists were consulted depending on the criteria and the animal species;
- for the second step, which goes from criterion-scores to principle-scores, we consulted animal and social scientists that had leading roles in Welfare Quality® (work package or task leaders). The social scientists brought in the point of view of the population they studied (citizens, retailers, or producers) while the animal scientists were considered to bring the animal point of view;
- for the final step, from principle-scores to overall assessment, animal and social scientists were again consulted together with stakeholders (members of the Advisory Committee of Welfare Quality® composed of representatives of producers organisations, breeders, retailers, veterinarians, NGOs, the European Society for Agricultural Ethics, and institutions (International Organisation for Animal Health – OIE). Stakeholders were essential at that stage because the production of a final overall assessment very much depends on how the system is going to be used.

At no time did we ask these people to choose between ethical dilemmas. We rather showed them datasets and asked them to react on them (examples provided in Tables 3.2 and 3.3), by assigning values on a 0-100 scale where:

- 0 corresponds to the worst situation one can find on an animal unit (i.e. the situation below which it is considered there cannot be further decrements in welfare);
- 50 corresponds to a neutral situation (i.e. level of welfare is not bad but not good);

TABLE 3.2 Dataset shown to animal scientists for them to assign a score for absence of injuries according the percentage of calves affected by moderate injuries (such as hairless spots) or severe injuries (such as a joint lesions).

	% calves with no injury	% calves with moderate injuries	% calves with severe injuries	Rank	Score
Farm 1	100	0	0	?	?
Farm 2	0	100	0	?	?
Farm 3	0	0	100	?	?
Farm 4	50	50	0	?	?
Farm 5	0	50	50	?	?
Farm 6	50	0	50	?	?
Farm 7	50	25	25	?	?
Farm 8	20	50	30	?	?
Farm 9	0	25	75	?	?
Farm 10	75	25	0	?	?
Farm 11	75	0	25	?	?

Notes: the scientists were requested to rank the (virtual) farms from worst to best and then to assign a score to each of them.

TABLE 3.3 Combinations of criterion-scores shown to animal and social scientists for them to assign resulting principle-scores.

Score for criterion 'Absence of prolonged hunger'	Score for criterion 'Absence of prolonged thirst'	Score for principle 'Good feeding'
25	75	?
40	60	?
50	50	?
60	40	?
75	25	?

- 100 corresponds to the best situation one can find on a farm (i.e. the situation above which it is considered there cannot be further improvements in welfare).

In addition, we specified that a farm that scores below 20 has high risk to be excluded from any welfare scheme.

For the last step of the scoring model, the consultation was run by discussion among scientists and stakeholders, during meetings and by e-mail. Potential uses of the assessment system were discussed as well as ways to aggregate the four principle-scores while taking into account both theoretical requirements and the distribution of results across farms visited in Welfare Quality®.

3.3.2 Consultation after the Construction of the Scoring Model

Once the scoring model was developed, it was discussed among citizens and farmers juries in the United-Kingdom, Italy, the Netherlands, and Norway. The results are not yet known at the time we are writing this text.

3.4 THE MODEL

3.4.1 From Measures to Welfare Criteria

Although this is not generally the case, some measures may be related to several criteria (e.g. low body condition score can originate from hunger or disease, or both). In order to avoid double counting, measures have been allocated to only one criterion, except in very few cases where we could distinguish the way they were interpreted (e.g. access of cattle to pasture is used to check the ease of movement (especially for animals which are tethered in winter) and the expression of behaviour).

The data produced by the measures relevant to a given criterion are interpreted and synthesized to produce a criterion-score – on the 0-100 scale described above – that reflects the compliance of the animal unit to this criterion.

Because the total number of measures, the scale on which they are expressed, and the relative importance of measures varies between and within criteria and also between animal types, the calculation of scores varies accordingly. Three main types of calculation are used:

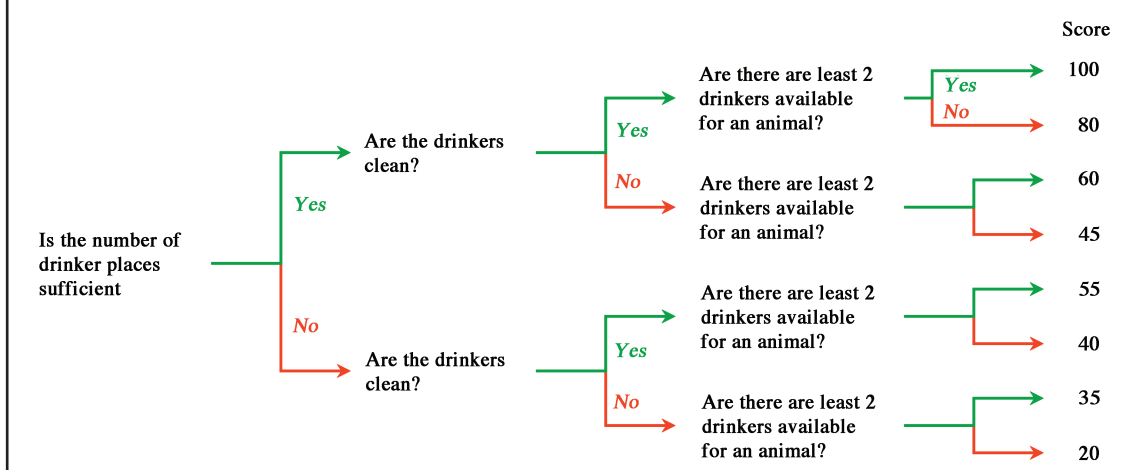
- When all measures used to check a criterion are taken at farm level and are expressed in a limited number of categories, a decision tree is designed. An example is provided in Box 3.1.
- When a criterion is checked by only one measure taken at individual level but with several degrees (e.g. moderate vs. severe problems), the proportion of animals observed can be calculated (e.g. percentage of animals walking normally, percentage of moderately lame animals, percentage of severely lame animals). In that case a weighted sum is calculated, with weights increasing with severity. An example is provided in Box 3.2.
- When the measures used to check a criterion lead to data expressed on different scales (e.g. percentage of animals lying outside the lying area, average latency to lie down expressed in seconds), data are compared to an alarm threshold that represents the limit between what is considered abnormal vs. normal. Then the number of alarms is valued. An example is provided in Box 3.3.

The animal scientists consulted had to interpret the raw data in terms of welfare. When necessary, alarm thresholds were defined with them. Then experts were asked to score virtual farms. In case weighted sums were to be calculated, this consultation was used to define weights that produce the same ranking of farms as the one given by experts. Then the scores they assigned were used to define functions to transform data into scores.

This exercise showed that experts do not follow a linear reasoning, e.g. for a given disorder a 10 % increase does not yield the same decrement in expert scores at the bottom of the

Box 3.1 Decision tree as applied to absence of prolonged thirst in fattening pigs.

Thirst is not assessed directly on animals because signs of dehydration can be detected only in extreme cases. The number of drinking places, their functioning and their cleanliness are assessed. The recommended number of pigs is calculated (10 pigs per functioning drinking place and 5 for a drinking place of reduced capacity). If there are more pigs in the pen than recommended then the number of drinking places is considered insufficient. Thereafter, cleanliness of drinkers and whether pigs have access to two drinkers in the same pen is considered. The following decision tree is applied:



Box 3.2 Weighted sum and I-spline functions as applied to lameness in dairy cows.

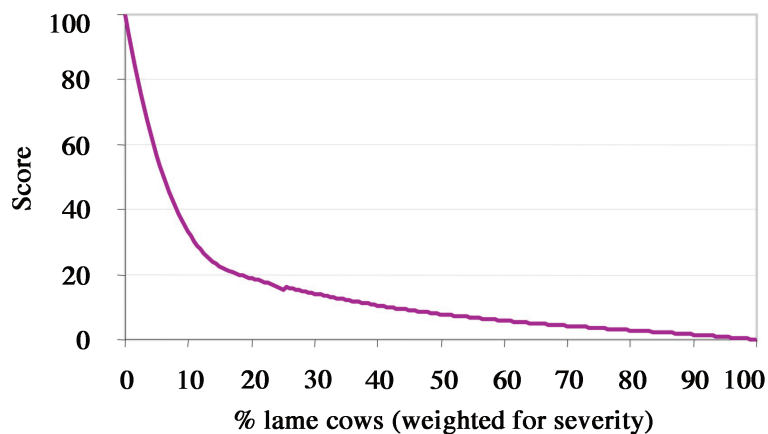
The % of animals moderately lame and the % of animals severely lame are combined in a weighted sum, with a weight of 2 for mild lameness and 7 for severe lameness. This sum is then transformed into an index that varies from 0 to 100:

$$\text{Index for lameness } I = \left(100 - \frac{2 (\% \text{ mild}) + 7(\% \text{ severe})}{7} \right)$$

This index is computed into a score using I-spline functions:

$$\text{When } I \leq 65 \text{ then Score} = (0.0988 \times I) - (0.000955 \times I^2) - (5.34E^{-05} \times I^3)$$

$$\text{When } I \geq 65 \text{ then Score} = 29.9 - (0.944 \times I) - (0.0145 \times I^2) + (1.92E^{-05} \times I^3)$$



BOX 3.3 Use of alarm thresholds applied to absence of diseases in broilers.

In broiler chicken the following disorders are checked on the farm or at slaughter: ascites, dehydration, septicaemia, hepatitis, pericarditis, subcutaneous abscesses, mortality and culling. The incidence of each disorder is compared to an alarm threshold, defined as the incidence above which a health plan is required at the farm level.

Disorder	Alarm Threshold (%)
Ascites	1
Dehydration	1
Scepticaemia	1.5
Hepatitis	1.5
Pericarditis	1.5
Subcutaneous abscess	1

When the incidence observed on a farm reaches half the alarm threshold, a warning is assigned. The number of alarms and warnings detected on a farm are calculated. They are used to calculate a weighted sum finally transformed into a score using I-spline functions (as in the example shown in Box 3.2).

0–100 scale (where most animals get this disorder) than at the top of the scale (when most animals are normal). It is therefore necessary to resort to non-linear functions to produce criterion-scores, in this case I-spline functions. Briefly, I-spline functions allow calculation of portions of curves so as to obtain a resulting smooth monotonic curve. They are expressed in the form of cubic functions (Box 3.2).

When a criterion was composed of very different items that experts find difficult to consider together at first glance, blocks of measures were first considered and they were further aggregated using Choquet integrals (see next section).

3.4.2 From Welfare Criteria to Welfare Principles

Criterion-scores are synthesized to calculate principle-scores. For instance, the scores obtained by an animal unit for absence of injuries, absence of disease, and absence of pain due to management procedures are combined to reflect compliance of this unit with the principle ‘Good health’. The consultation of animal and social scientists showed that they consider some criteria more important than others (e.g. in most animal types, ‘Absence of disease’ is more important than ‘Absence of injuries’ which in turn is more important than ‘Absence of pain induced by management procedures’). Nevertheless they do not allow compensation between scores (e.g. absence of disease does not compensate injuries and vice versa). A specific operator (Choquet integral) was used to take into account these two lines of reasoning. Briefly, the Choquet integral is one of the most expressive decision criteria. It is an aggregation function that generalizes the notion of weighted average when weights are not only attached to each criterion but also possibly to any subset of criteria (Grabisch and Roubens, 2000). The Choquet integral calculates the differences between the minimum score and the next minimum score and assigns a weight (called ‘capacity’) to that difference. This process is repeated until the highest score. An example of calculation of principle-scores is provided in Box 3.4.

3.4.3 From Welfare Principles to Overall Assessment of Farms

The scores obtained by an animal unit on all welfare principles are used to assign that farm to a welfare category. How many and what welfare categories are necessary depend on the purposes for which the welfare assessment will be used. After consultation of the Advisory Committee of Welfare Quality®, various uses of the assessment were identified. These uses and their implications for welfare categories are summarised in Table 3.4. According to the range of potential uses of the assessment, four welfare categories are to be distinguished:

- Excellent: the welfare of the animals is of the highest level,
- Enhanced: the welfare of animals is good,
- Acceptable: the welfare of animals is above minimal requirements,
- Not classified: the welfare of animals is low and considered unacceptable.

We decided to use an outranking method to assign farms to these categories because it is especially designed to assign items (here farms or slaughterhouses) to predefined categories, while allowing adjustment to the level of compensation between principle-scores. Briefly, aspiration values are defined for each category. They represent the goal that the farm should try to achieve to be assigned to a given category. Then membership rules have to be set that describe how far a farm has to reach these aspiration values to be assigned to the given category defined by such value.

Box 3.4 Use of a Choquet integral to calculate the principle-scores for ‘good health’.

‘Good health’ integrates 3 criteria; ‘Absence of injuries’, ‘Absence of disease’, and ‘Absence of pain due to management procedures’. First the scores obtained by a farm for the 3 criteria are sorted in increasing order. The first criterion-score is considered, and then the difference between that score and the next criterion-score is multiplied by the ‘capacity’ of the group made of all criteria except the one that brings the lowest score. Following this, the difference between the last but one score and the next score is multiplied by the ‘capacity’ of the group made by the combined criteria except those that bring the two lowest scores. This can be written as follows:

$$\text{Principle-score} = \begin{cases} S_6 + (S_7 - S_6)\mu_{78} + (S_8 - S_7)\mu_8 & \text{if } S_6 \leq S_7 \leq S_8 \\ S_6 + (S_8 - S_6)\mu_{78} + (S_7 - S_8)\mu_7 & \text{if } S_6 \leq S_8 \leq S_7 \\ S_7 + (S_6 - S_7)\mu_{68} + (S_8 - S_6)\mu_8 & \text{if } S_7 \leq S_6 \leq S_8 \\ S_7 + (S_8 - S_7)\mu_{68} + (S_6 - S_8)\mu_6 & \text{if } S_7 \leq S_8 \leq S_6 \\ S_8 + (S_6 - S_8)\mu_{67} + (S_7 - S_6)\mu_7 & \text{if } S_8 \leq S_6 \leq S_7 \\ S_8 + (S_7 - S_8)\mu_{67} + (S_6 - S_7)\mu_6 & \text{if } S_8 \leq S_7 \leq S_6 \end{cases}$$

Where S_6 , S_7 , and S_8 are the scores obtained by a given farm for Criterion 6 (‘Absence of injuries’), 7 (‘Absence of disease’), and 8 (‘Absence of pain due to procedures’),
 μ_6 , μ_7 , μ_8 are the capacities of Criterion 6, 7 and 8,
 μ_{67} is the capacity of the group made of criteria 6 and 7,
 etc.

TABLE 3.4 Potential uses of the welfare assessment.

Potential uses of the welfare assessment	Welfare categories needed	
	No.	Description
Scenario 1 Standard for cross-compliance and future definition of a minimum animal welfare standard Compulsory labelling defining several levels of welfare	2	Below / Above minimum legal requirements or equivalent (if no legislation)
	4	Poor / Normal / Good / Excellent
Scenario 2 Setting welfare targets for farm certification schemes - Voluntary labelling	3	Very high welfare / High welfare / Other*
Scenario 3 Feedback to producers for monitoring the results of welfare improvement strategies Self assessment tool	Several	Very poor ----- Very high level of welfare**
Scenario 4 Assessing new animal farming systems/breeds Furthering research on animal welfare	Several	Very poor ----- Very high level of welfare**

Notes: * ie insufficient to enter a certification scheme; ** with intermediate categories.

We first discussed whether the aspiration values should be absolute or relative. Absolute values mean that they are set whatever the results obtained by farms. This might be dangerous since absolute aspiration values may be either too difficult to achieve or too easy so not discriminating among farms. Relative values means that they correspond to a certain % of farms from a given population in which the system is used; e.g. one could decide that farms that are in the top 10% of their population are considered excellent, whatever the exact level of welfare they have achieved. In turn such relative values seem difficult to handle for three reasons. Firstly, excellent may in fact correspond to a low level of welfare. Secondly, the classification of farms depends on the population observed, in other words a farm may be classified excellent in a given population whereas the same farm could be enhanced or even just acceptable in another population. Thirdly, relative aspiration values prevent monitoring progresses on farms, e.g. an enhanced farm may remain enhanced even though the welfare of animals improves. We decided to use absolute aspiration values and to check that these were realistic according to the farms visited in Welfare Quality® and that they would discriminate farms. In addition, we adjust membership rules according to the results obtained by farms (see section ‘Tests on datasets’).

Because (1) the 0–100 scale on which principle-scores are expressed has the same meaning whatever the principle, and (2) principles were considered to have all the same importance for the animals, we chose to set flat reference profiles, i.e. only one aspiration value for a given category, valid for the four principles. The aspiration value for the category ‘Acceptable’ was set at 20 because 20 had a specific meaning on the 0-100 scale use for expressing criterion- and principle-scores: when a farm scores below 20 on one criterion or principle, this farm has high risk to be excluded from any welfare scheme. Again 50 had a specific meaning: a farm that scores 50 is not good and not bad. So we decided to set the aspiration value for the category ‘Enhanced’ just above, that is at 55. Finally the aspiration value for ‘Excellent’ was set symmetrically to the one for ‘Acceptability’, that is at 80.

3.5 TEST ON DATASETS

3.5.1 From Measures to Welfare Criteria

For the calculation of criterion-scores for dairy cows, two consultations were run: one before and one after data had been collected on farms. When the initial evaluations by scientists were confronted with the actual distribution of data across the dairy farms visited in Welfare Quality[®], it appeared that the scientists had been rather severe. For instance regarding lameness, 50% of farms had more than 12% lame cows and so would score 30 or below. We presented the distribution of lameness to the scientists consulted (Figure 3.2) and asked them if they were willing to change their initial evaluations. Most of them were less severe during this second consultation (Table 3.4). However, they considered lameness to be a serious problem for cows and so 50% farms still score lower than 44. For the final scoring of farms, we used expert scores from the second consultation.

The calculation of scores for fattening cattle, veal calves, fattening pigs, and broilers were defined following only one consultation because this was run after data had been collected on farms.

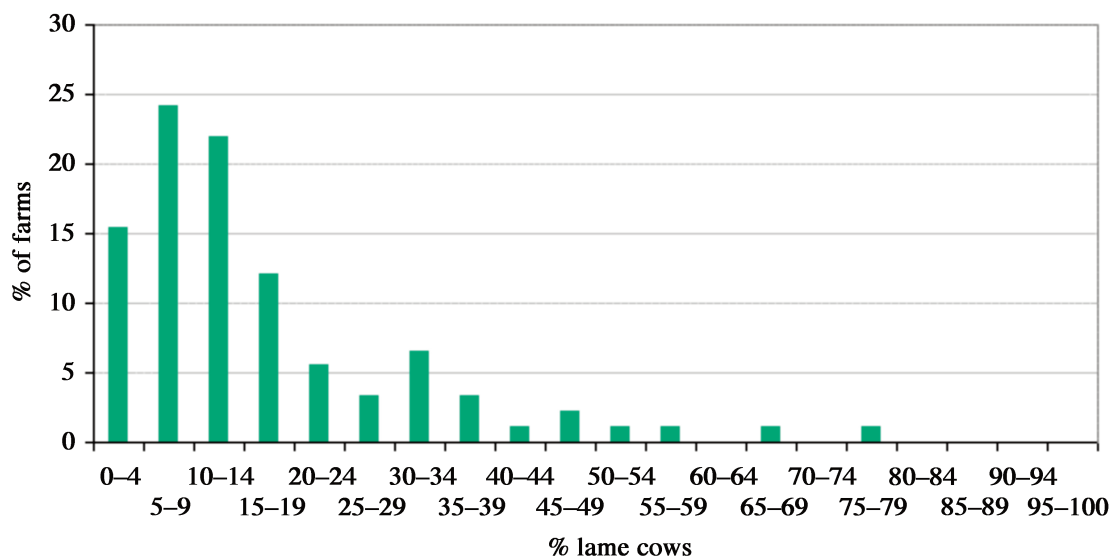


FIGURE 3.2 Distribution of lameness across dairy farms visited in Welfare Quality[®]
Notes: the % lame cows is weighted for the severity of lameness: 0.14 for moderate and 1 for severe lameness.

TABLE 3.4 Comparison between scores assigned for lameness by animal scientists before and after being informed of the distribution of lameness across the dairy farms visited.

% lame cows ¹	Mean expert scores	
	1st consultation	2nd consultation
10	34	46
30	13	20
60	6	6

Notes: ¹ weighted for severity: 0.14 for moderate and 1 for severe lameness.

3.5.2 From Welfare Criteria to Welfare Principles

For the calculation of principle-scores for dairy cows, we again ran two consultations: one before and one after data had been collected on farms. In that case, the evaluations were almost the same. We thus decided to proceed with only one consultation, without confronting the experts to the results obtained by farms.

3.5.3 From Welfare Principles to Overall Assessment

We analysed the distribution of dairy farms across the aspiration values for the welfare categories ‘Acceptable’, ‘Enhanced’, and ‘Excellent’ (respectively 20, 55, and 80) for each principle. It turned out that farms could hardly achieve 80 in any welfare principle. Actually, such a value was obtained only for the principle ‘Good feeding’. Nevertheless, a wide range of results was obtained for the other principles, from 0 to 70. We thus decided to keep the three aspiration values as defined (20, 55, and 80), having in mind that if the system is put in place improvements should be encouraged and scores of 80 or more should be obtained.

Then to decide that a farm performs enough to be considered to fulfil the aspiration value of a category, several membership rules were tested. A first very intuitive rule is ‘unanimity’ which means that a farm needs to reach the aspiration value of a given category for all welfare principles to be assigned to that category. For instance, to be enhanced a farm would need to score at least 55 on all principles. Such a rule seemed not realistic since half the farms visited in Welfare Quality® would be ‘Not classified’ and the other half would be only ‘Acceptable’. We thus investigated other rules where a farm needs to score higher than the aspiration value of a category only on some principles (3 or 2 out of the 4) while not falling below the aspiration value of the next lower category for any principle. We observed the likely distribution of farms using such rules (Figure 3.3) and we also compared the classification obtained by farms according to the various rules tested, to the general impression of observers who visited the farms (expressed on a Visual Analogue Scale). The most appropriate rule appeared to be the following:

A farm is considered ‘Excellent’ if it scores more than 55 on all principles and more than 80 on two of them. It is considered ‘Enhanced’ if it scores more than 20 on all principles and more than 55 on two of them. Farms with ‘Acceptable’ levels of animal welfare score more than 10 on all principles and more than 20 on three of them. Farms that do not reach these minimum standards are not classified (Figure

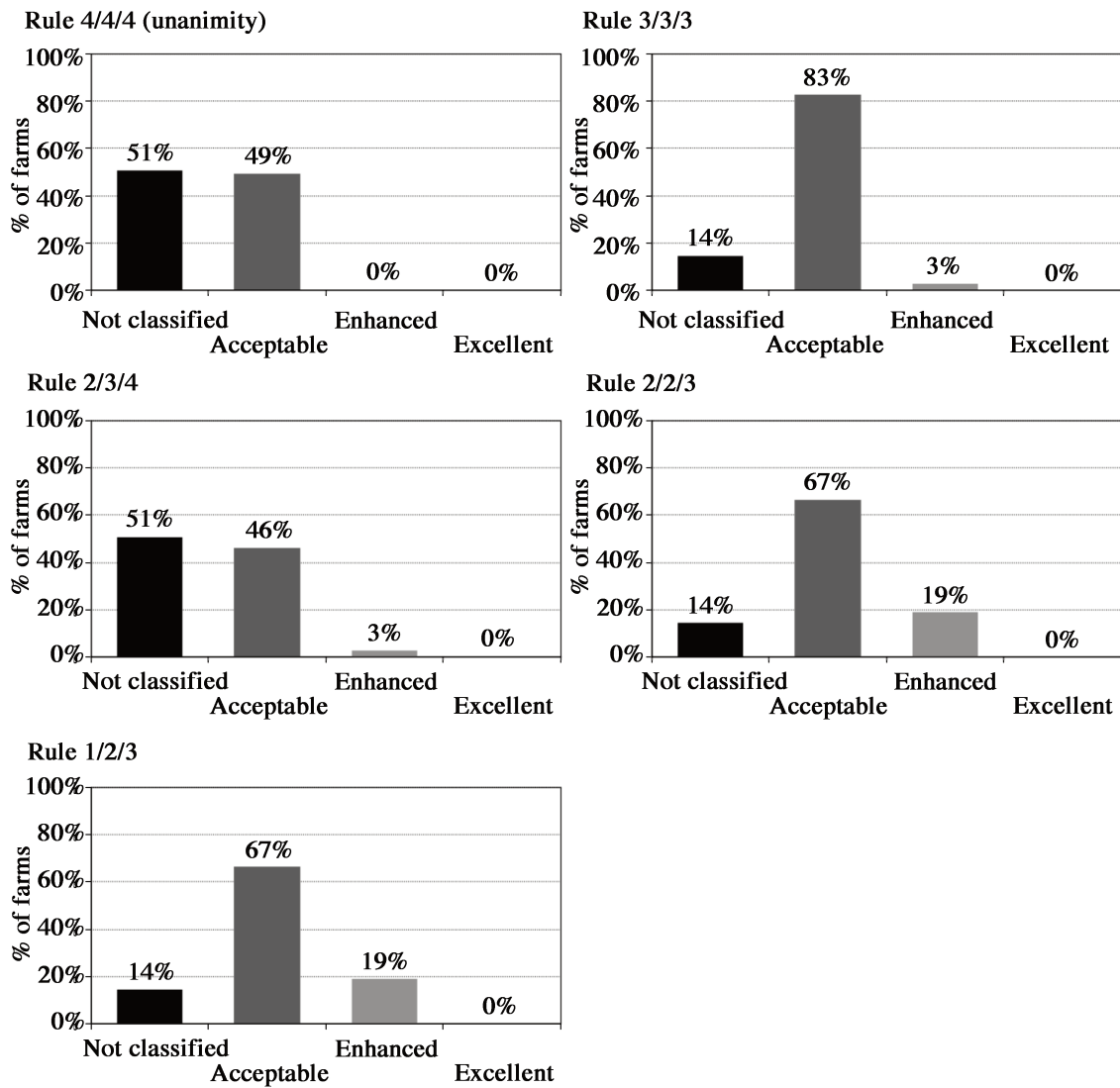


FIGURE 3.3 Distributions of 69 European dairy farms across the four welfare categories, according to several rules

Notes: For each rule, the numbers indicate how much a farm needs to reach the aspiration value of a given category to be classified in that category. For instance, with Rule 2/3/4 a farm needs to reach 80 on 2 welfare principles to be excellent, it needs to reach 55 on 3 principles to be enhanced and it needs to reach 20 on all 4 principles to be acceptable. In any case, a farm shall not fall below the aspiration value of the next category on any principle (e.g. a farm cannot be excellent if it falls below 55 on at least one principle).

3.4). In addition, an indifference threshold equal to 5 is applied: For instance, 50 is not considered significantly different from 55.

The value of the indifference threshold was decided according to the standard error between experts' answers on the datasets used to define the construction of the criteria. This error equals 5.13. We thus decided of an indifference threshold of 5.

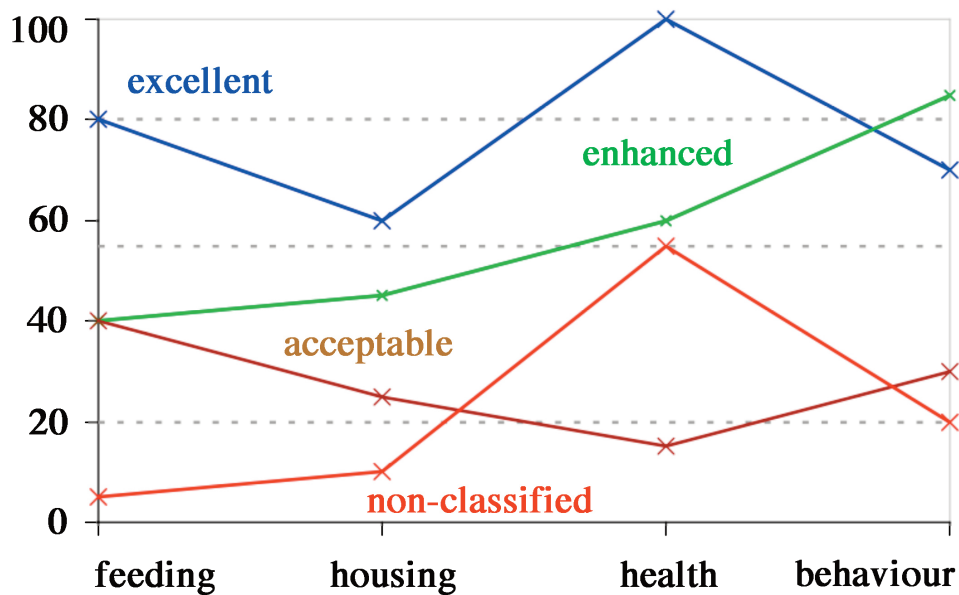


FIGURE 3.4 Examples of farms in the four welfare categories.

3.6 BACK TO ETHICAL ISSUES

Although we did not address ethical issues directly, the people we consulted made themselves ethical decisions and these are reflected in the scoring model.

First, it is clear that more emphasis is put on the animals in very poor conditions than on the best ones. This is reflected in the calculation of criterion-scores. For instance, as shown in the figure from Box 3.2, 9% lame cows results in a score of 50 whereas if the same importance were assigned to animals in bad and good condition then a farm would obtain a score of 50 when 50% animals are lame. Nevertheless, the overall welfare of a group of animals also matters. In the example of lameness, a farm with 5% severely lame cows and 50% moderately lame cows scores lower than a farm with 10% severely lame cows and no cow moderately lame. Similar results are found whatever the animal types and the criteria. Therefore, the scoring model reflects a balance between priority given to animals in the poorest conditions and the overall welfare of the whole herd of flock.

As mentioned earlier, compensation is very limited between criteria. However some compensation is still possible. As a consequence, a farm that scores less than 20 on one criterion has some chance to finally be considered acceptable if it scores high for the other criteria. For example, a farm that scores 42 for housing (with 60 for hunger and 40 for thirst), 39 for housing (50 for resting, 25 for thermal comfort, 75 for ease of movement), 20 for health (14 for injuries, 38 for disease, and 52 for pain), and 43 for behaviour (35 for social and other behaviours, 65 for human-animal relation and positive emotional state) would still be considered acceptable although it obtains only a 14 for Criterion Absence of injuries. After discussion with the scientists consulted and also from the first citizens'

juries run, this seems not in line with most people perception of overall welfare (where all criteria need to reach a certain minimal level). To meet this perception, an additional rule should be applied whereby a farm shall not be considered acceptable when it falls below 20 on one criterion. Nevertheless, at present, such a rule appears premature as it would lead to the rejection of half the farms.

Finally the scoring model results both from theoretical expectations of what is poor vs. good welfare, as reflected in the definition of aspiration values for the various welfare categories (20 for Acceptable, 55 for Enhanced, and 80 for Excellent) and what can realistically be achieved in practice, as reflected in the rules chosen to assign farms to a welfare category (e.g. a farm needs to be at least excellent on two principles and only enhanced on the two others to be considered excellent). We hope that a satisfactory balance between theory and pragmatism has been achieved in that scoring model.

Finally, most systems used in certification or control are based on a number of points that the farm must comply with. To our knowledge, these systems generally produce a pass/fail answer and it might be difficult for a farmer that fails to figure out how far from passing he was. The Welfare Quality[®] scoring system allows producing more than a pass/fail answer by allowing four grades and may thus encourage farmers to improve their status in regard to animal welfare.

3.7 THE LESSONS WE HAVE LEARNED

When starting Welfare Quality[®] we identified the need to produce a model for the overall assessment of animal welfare. We knew that such an exercise would not be an easy one because animal welfare is multidimensional and that intuitive aggregation methods, such as weighted sums, allow full compensation between dimensions, which did not seem appropriate to evaluate animal welfare. This is where operational research entered into the scene and helped us going from a mere description of animal welfare to its evaluation. And we learned a lot from this exercise.

First, we learned that describing is easy but assessing is tricky!

Second, assessing means making a judgement and so a clear strategy for the consultation of people that will make this judgement is necessary. In Welfare Quality[®], several groups of people were consulted: (1) animal scientists in order to give a precise interpretation of the measures taken on animals or their environment in terms of what can be perceived by animals, (2) social scientists to bring the opinion of the stakeholder groups they studied (consumers-citizens, retailers, producers), and (3) stakeholders themselves. Stakeholders were involved only in a late stage to avoid them being both judges and judged.

Third, adequate methods are to be chosen to model appropriately people opinions. We avoided methods that would allow full compensation between welfare aspects. And we finally used an outranking method to assign animal units in four welfare categories: not classified, acceptable, enhanced, and excellent, each category likely to correspond to a decision by stakeholder (reject/accept/reward).

Looking backwards on the work we did, we would like to warn those who may develop such models in the future e.g. for other animal types that it is important that they work in close connection with the researchers who develop measures. Indeed having a clear idea of the architecture of the model is essential for the development of measures, ensuring that any criterion can be checked by relevant measures, and that the measures can be easily interpreted in terms of welfare.

The aggregation process followed in Welfare Quality®, although rather tricky in order to match people opinions, can still be explained in simple terms. For instance curves can be shown to simulate results and ethical positions are made clear. It is absolutely necessary to communicate on the model so that it is not perceived as a black box but as something that reflects correctly animal welfare and can help to form opinions about animal units and even to take adequate and fair decisions about them.

TESTING THE ASSESSMENT SYSTEM: REFINEMENT AND DEFINITION

Kees van Reenen and Bas Engel

4.1 INTRODUCTION

The monitoring systems developed must enable differentiation between living conditions of animals and their sensitivity must correspond to actual variations between farms. To our knowledge, such tests of the sensitivity of whole monitoring systems of animal welfare have not been carried out. This needs to be done through surveys on representative samples of different rearing systems and local conditions (e.g. climate) found in Europe. In addition, practicability must be checked in a wide range of conditions.

Practical (including economic) feasibility dictates that a farm animal welfare monitoring system must be as concise and easy to carry out as possible. Therefore, there is a need to examine alternative monitoring systems, differing in the level of detail (e.g. with respect to the number of different welfare measures included). To this end, the so-called ‘full monitoring system’, encompassing the full range of measures that are used in the aggregation process to provide an overall welfare assessment, should be tested on a large number of farms, in order to obtain a data set which is sufficiently large to allow for useful epidemiological and other statistical analyses. A large scale approach is also necessary in order to be able to test the full monitoring systems in a representative range of conditions across the EU.

Work Package 2.4 (WP2.4) was established to define the ‘full monitoring systems’ (Task 2.4.1.), which could then be applied on a sample of farms across Europe in each species. Meaningful data could only be obtained when observers in each country have the same level of training. Therefore, Task 2.4.2 was devoted to the training of future observers. Finally, within Task 2.4.3, the application of ‘full monitoring systems’ in practice was carried out, and the appropriate methodological approach was developed for the statistical analyses of the data obtained.

4.2 STARTING POINT OF WORK PACKAGE 2.4

Unlike individual welfare measures (of disease, of fear, etc.), a welfare monitoring system comprising a range of different aspects and measures of farm animal welfare and intended to provide an estimation of the overall level of animal welfare, cannot be properly (i.e. experimentally) validated. Firstly, there is no true gold standard for the overall level of welfare (e.g. on a particular farm) and secondly, controlled differences in the overall level of welfare (e.g. between groups of farms) do not exist. What can be done, however, is the definition of an approximate gold standard for a welfare monitoring system, based on scientific consensus as well as common sense. An important question in this respect concerns the range of different measures and aspects that should be included in the monitoring system (so-called content validity). Moreover, individual measures to be used in the monitoring system should be scientifically valid, reliable and feasible. These aspects of individual measures of farm animal welfare are examined experimentally in Work Package 2.2 (WP2.2).

4.3 DEFINITION OF FULL MONITORING SYSTEMS

Results of the studies performed in WP2.2 were evaluated in Task 2.4.1. The combined expertise in Sub Project 2 was used to compare and evaluate the validity, reliability and feasibility of the many individual measures. In addition, remaining issues around data collection on farm and at slaughter were evaluated and decided upon. Finally, it was checked whether the draft ‘full monitoring systems’ met the expectations and requirements of consumers, retailers and producers. If necessary, drafts were adjusted following this latter check.

4.4 TRAINING THE PEOPLE INVOLVED IN ON-FARM OBSERVATIONS

In this task, standardized methods were proposed for training observers to be involved in the on-farm observations in Task 2.4.3. A second objective was to define clear criteria observers have to meet in order to be allowed to participate in the on-farm observations. The third objective was to hold training sessions with groups of partners working on the same animal species and type.

In Task 2.4.3, measures were to be obtained by observers and/or partners who were not necessarily involved in the development of that particular measure. Therefore, proper training prior to the start of Task 2.4.3 was vital, in order to ensure that: (1) the practical execution of the respective full monitoring systems was similar across partners and countries, and (2) there was a sufficient level of reliability across observers.

Task 2.4.2 had two sub-tasks, i.e. sub-task 2.4.2.1 and sub-task 2.4.2.2.

4.4.1 The Development of Standardized Methods to Train Observers (Sub-task 2.4.2.1)

A method for training was developed that included two successive phases: (1) a laboratory phase, where future observers ('trainees') are trained with the use of audio-visual material (i.e., photos or video clips), and (2) an on-farm phase where trainees visit actual farms and the training involved the assessment of live animals. During the laboratory phase, it is assumed that for each photo or video clip there is a 'true' measurement available, preferably established by a panel of experts – this true measurement or true score acts as the so-called 'gold standard', against which trainees are trained. After an instruction period, trainees have to score a certain number of photos and/or video clips, and achieve a minimum number of 'correct answers', i.e., in accordance with the 'gold standard' previously established by the experts. This minimum number of correct answers is determined by statistical considerations, and the decision whether or not a trainee passed the critical threshold is the result of a statistical analysis of the data. In WP2.4, a range of specific selection criteria were suggested, depending on the type of measurement. Photos and video clips in this respect should be representative and a sufficient number of different photos and/or video clips should be available; otherwise no sufficient accuracy can be achieved in a statistical evaluation. Moreover, results may be biased because of memory effects when a modest number of photos and/or video clips are used.

It is suggested that a trainee is ready to enter the on-farm phase of the training after the successful completion of the laboratory phase, for example by passing an exam. During the on-farm phase of the training, rather than looking at photos and/or video clips, the trainee examines individual animals or pens. Similar to the laboratory phase, during the on-farm phase there is a need for a sufficient number of animals or pens (depending on the measure that is observed) in order to yield sufficiently accurate results. Among the trainers (i.e. those who instruct and supervise the trainees on-farm) there should be at least one expert who will be involved in the statistical evaluation of the on-farm data. This expert can be referred as the 'silver standard', indicating that the expert's scores are close to, or may be even equivalent to, the 'gold standard'. Since the expert can be expected to be less repeatable on-farm (quick decisions have to be made, under less than ideal conditions, etc.) than in the laboratory with photos and/or video clips (no time constraints, ideal conditions, there is the possibility for discussion with other experts, etc.), the term 'silver' rather than 'gold' standard is used in the context of on-farm training. Upon completion of the on-farm training, the successful candidate should have scored a sufficient number of animals or pens in accordance with the expert evaluation.

During this sub-task, training material (photos and video clips, as well as clear descriptions of the protocols) was produced and obtained in the various farm animal species, and put on the Welfare Quality® webtool.

4.4.2 The Realization of Training Sessions with Groups of Partners Working on the Same Animal Species and Type (Sub-task 2.4.2.2)

For the realization of the work a number of so-called ‘species groups’ were established. In each group, participants from different European countries involved in the same farm animal species worked together with regard to the definition of the protocols, the preparation and the organization of training sessions, and the application of the protocols on-farm by trained observers. There were basically four species groups: one in dairy and beef cattle (with participants from Austria, Germany, the Czech Republic, Italy, the UK and Sweden), one in veal calves (with participants from France, Italy and The Netherlands), one in pigs (sows and finisher pigs, with participants from Spain, France, Belgium, The UK and The Netherlands), and one in poultry (laying hens and broilers, with participants from Sweden, France, The UK and The Netherlands). Each species group organized at least one plenary training session at one location, which lasted between 5 and 7 days.

Considering the importance of reliable data, it was decided that the training should be ongoing during the period of data collection. Therefore, observers from all partners, in each species group, regularly met during data collection to check (and if necessary to correct) any deviation from the required standard.

4.5 REFINEMENT OF MONITORING SYSTEMS AND DEFINITION OF FINAL MONITORING SYSTEMS

In order to fulfil the objectives of WP2.4, it was important to apply the ‘full monitoring systems’ on farms across Europe, thereby creating a data-base that would allow for comprehensive quantitative analyses. The full monitoring systems were run on farms for cattle (dairy cows, beef cattle and veal calves), pigs (sows and finisher pigs) and poultry (laying hens and broilers). An overall overview of the numbers of farms that were visited is given in Table 4.1.

4.5.1 Running Surveys in Cattle

In dairy cows, a total of 91 farms were visited (see Table 4.1): 25 in Austria, 15 in the Czech Republic, 26 in Germany and 25 in Italy. Farms differed with respect to various

TABLE 4.1 Overview of numbers of farms visited in Work Package 2.4, for the different farm animals species and distributed across countries.

Species	Country ¹									Total
	AT	DE	CZ	IT	ES	FR	NL	UK	SE	
Dairy cows	25	26	15	25						91
Beef cattle	30			25				30		85
Veal calves				24		50	150			224
Sows							45	45		90
Pigs					41	30				71
Laying hens							29		48	77
Broilers						26		32		58
Total	55	26	15	74	41	106	224	107	48	696

Notes: ¹ AT: Austria, DE: Germany, CZ: Czech Republic, IT: Italy, FR: France, NL: The Netherlands, UK: United Kingdom, SE: Sweden.

husbandry aspects, including, for example, the type of housing (loose housing and tethered housing), the presence of an outdoor ‘loafing area’, or providing cows access to pasture.

In beef cattle, a total of 85 farms were visited, 30 in Austria, 30 in the UK and 25 in Italy. On some farms cattle were grazed whereas on other farms cattle were kept in intensive systems indoors. There were also notable differences between farms in the use of litter (straw). In Sweden, a number of slaughterhouses were visited for the specific assessment of the welfare of beef cattle before and during slaughtering.

In veal calves, a total of 224 farms were visited, 50 in France, 24 in Italy and 150 in The Netherlands. The sample of farms included farms for the production of both so-called ‘white’ veal and ‘pink’ veal. Farms also varied according to the type and origin of calves, group size, size of the farm, diet (amount of milk replacer and amount and type of solid feed), climate control, day light intensity, and management.

4.5.2 Running Surveys in Pigs

In sows, a total of 90 farms were visited, 45 in The UK, and another 45 in The Netherlands. Some of the farms examined in The UK used an extensive outdoors system. Indoor husbandry systems used different types of floor, bedding material, and feeding system.

In finishing pigs, a total of 71 farms were visited, 41 in Spain and 30 in France. Similar to sow farms, the sample of farms with finisher pigs included both extensive (outdoors) and intensive farms, as well as farms with and without bedding material on the floor.

In addition to farms, nine slaughterhouses were visited in Spain, for the collection of data with regard to the welfare status of finishing pigs before (while being unloaded), and during slaughtering (stunning).

4.5.3 Running Surveys in Poultry

In laying hens a total of 77 farms were visited, 29 in The Netherlands, and 48 in Sweden. A range of various housing and husbandry systems was represented in the sample of farms: free range systems with different systems indoors, as well as housing systems with conventional cages, furnished cages, and aviaries.

In broilers a total of 58 farms were visited, 32 in The UK and 26 in France. On a limited number of farms examined in The UK, broilers had the opportunity to go outside.

4.6 PROCESSING OF THE DATA AND DEFINITION OF FINAL MONITORING SYSTEMS

Within this sub-task, the appropriate statistical methodological approach for the analyses of data collected on-farm was developed, allowing for the refinement of monitoring systems, and the ultimate definition of final monitoring systems. This approach consists of three parts: (1) the analysis of correlations and associations between different animal-based measures, (2) the calibration of simplified versions of the monitoring system against the full version, and (3) a risk factor analysis.

4.6.1 Analyses of Associations between Different Animal-based Measures

Over the course of the Welfare Quality® project, the ‘full monitoring systems’ were applied on samples of farms across Europe in seven farm animal species (dairy cows, beef cattle, veal calves, fattening pigs, sows, laying hen and broilers). These data are used as input for the analyses. The first part of these analyses concerns the examination of correlations and associations between different animal-based measures, for example between different health measures or behaviours. Several statistical techniques are used here, including the calculation of (residual) correlations and principal component analysis. The outcome of the correlational analysis is of interest because (1) some variables, when highly associated with others, might be omitted, for example at an initial screening stage, to identify farms at risk, (2) before a risk factor analysis is undertaken, choices have to be made among highly correlated variables, in order to avoid confusion in the search for relevant risk factors of animal welfare, and (3) marked association may lead to simplification of the monitoring system. More specifically, for those measures that are significantly and highly correlated, regression analysis could be used to predict one measure from the other, and these results, in turn, may be used in the calibration of simplified versions of the monitoring system against the full version.

This is an example of an outcome of correlational analysis in dairy cows. In dairy cows, the final data set contained 27 animal-based variables, related to basically four fields of interest: (1) the body condition (% lean cows), the cleanliness of the animals (cleanliness of udder, legs and flank), health (skin lesions, respiratory disease, diarrhoea, lameness), behaviour (lying down, and social, in particular agonistic behaviours including head butting), and the human-animal relationship, reflected in the behavioural response of individual cows to being approached by an unfamiliar person (i.e., the observer applying the Welfare Quality® protocol on-farm). The correlational analysis yielded two significant correlations of moderate size: the % of cows with ocular discharge was positively correlated with the % of cows with severe skin injuries, i.e. cows with one or more large hairless patches and/or severe local swellings of the skin ($r = + 0.53$, $P < 0.001$), and the % of non-lame cows in a herd was positively correlated with the % cows that could not be approached by the observer in the human approach test ($r = + 0.52$, $P < 0.001$). The first significant correlation may possibly reflect the existence of a common factor (or common factors) on-farm that influence both the condition of the skin and the health status of the eyes. The latter correlation is particularly interesting, since it indicates that on farms with many non-lame cows it was more difficult to approach the animals than on farms with many lame ones, and, conversely, that on farms with many lame cows, a larger % of cows could be approached than on farms with many non-lame ones. Although such a correlation calculated on data collected on farms in practice by definition not necessarily refers to an underlying causal relationship, a likely biological explanation of the relationship between lameness and the response of cows to a human might be that lame animals are physically less able to respond to an approaching human (for example, by showing an active avoidance response) even when they would have a high level of fear of humans. If this is true, the behavioural test that was used to measure fear of humans in dairy cows would require some additional attention in terms of validation. Correspondingly, it would seem premature at this point to predict one of these correlated measures (i.e., the % of non-lame cows, or the % of cows that could not be approached by a human) from the other for simplification purposes.

4.6.2 Calibration of Simplified Versions of the Monitoring System against the Full Version

The calibration part of the data analysis comprises two steps: (a) the examination of animal-based measures separately, and (b) the calibration of simplified versions of the monitoring system against the gold standard.

When examining animal-based measures separately, the question is whether a protocol for an individual measures (for example, a health measure or a behavioural measure) can be simplified without losing too much information. A simplification may involve, for example, a smaller sample size or a reduced observation time in the case of a behavioural measure. Simulation techniques are used to model simplifications. In this context, it is also important to consider the sample sizes that are required to detect a welfare problem, for example to detect with some confidence (for instance, 90% or 95%) that the proportion of “defectives” exceeds a (pre-specified) critical level. Although a reduced sample size

may enhance the feasibility of the protocol, it may also compromise the accuracy of the protocol.

The essential principle underlying the calibration of simplified versions of the monitoring system against the gold standard is the notion that the output of the multicriteria evaluation in WP2.3, using the full set of measures, is the true gold standard against which simplified versions of the full monitoring system should be calibrated. Removal of entire measures, for the sake of simplification and reduction, is not a feasible option since every modification would require the subsequent re-construction of the monitoring system, including the multicriteria evaluation and the development of algorithms for the aggregation of measures into an overall welfare assessment. Second best is mimicking a simplified, reduced system in some acceptable way. To that end, rather than fully removing a measure, this measure could be replaced by a prediction based on another correlated measure. Note that a true re-construction of the monitoring system would involve consideration of the impact of the remaining measures, insofar as these measures are associated with the measure that is removed. This is mimicked by replacing a measure by a prediction based on another measure with the use of regression analysis. In this way the full monitoring system can be employed. Similarly, the construction of welfare evaluation criteria is fully preserved when simplified versions of certain measures are used, for example in terms of the numbers of animals observed or the total observation time. The calibration will, therefore, involve the examination of correlations and associations between, on the one hand, welfare scores – according to the multicriteria evaluation developed in WP2.3 – based on the full set of measures (i.e., the gold standard) – and, on the other, welfare scores based on simplified versions of the full monitoring system, consisting of a combination of: (1) original measures, (2) original measures that were adapted in terms of, for example, the number of animals or the observation time, and (2) predictions of certain measures based on other measures. After calibration of simplified systems against the full version, final welfare monitoring systems can be chosen from several alternatives based on a cost-benefit analysis, where cost represents the practical and economic feasibility of a monitoring system, and benefit refers to the precision compared to the full monitoring system.

4.6.3 Risk Factor Analysis

Within each of the 7 species that were studied in Welfare Quality®, a risk factor analysis is carried out where the associations are examined between animal-based measures on the one hand, and environment-based ones on the other. This will provide information on potential risk factors for animal welfare. This may help us, for example, to understand possible causes of reduced welfare in practice, and to suggest ways to improve farm animal welfare.

A suite of statistical programs has been made, that enables the user to select important risk factors for different response variables (animal-based measures). Two programs are used for the identification of single potential risk factors (both qualitative factors, i.e. class variables, and quantitative factors, i.e. covariables); these are input for a third program

where multiple risk factors (class variables and covariables) can be modelled simultaneously. The third program comprises forward, backward and stepwise selection and all subsets selection. Interactions among class variables and between class variables and covariables are also included in the third program. A fourth program inspects possible interactions between covariables. A fifth program, based on a generalized linear model (GLM) framework, evaluates any given model, for example a final model obtained after the application of model selection with programs 1-3, in terms of test results, predictions, pairwise comparisons and diagnostic plots.

The suite of selection programs is intended for data that are grouped at farm level. If there is a need to split farms into sub-units (for example, different age groups), these sub-units are first treated as quasi-independent units, and analysed with the suite of programs described earlier. To accommodate for different numbers of basic units, such as pens, in the sub-units, statistical weights can be specified. After final model selection, the selected risk factors can then be more reliably inspected in a mixed model analysis. In this analysis the measurement units could be, for example, pens within farms. Dependence between pens of the same farms is accounted for by inclusion of random farm effects. In addition, an alpha-beat version of a mixed model selection program can be applied to the data as well.

The suite of selection procedures allows for different types of response variables. For example proportions or percentages are handled by logistic regression, employing maximum quasi-likelihood. In the case of quasi-independent units, these proportions or percentages are subsequently analyzed by a generalized linear mixed model, employing penalized quasi-likelihood (PQL), a method motivated by an approximation of maximum likelihood by Laplace integration.

The suite of selection programs primarily uses the adjusted R square as a (lenient) criterion for selection. The programs also offer an overview of significance levels (P-values). Subsequent mixed model analyses focus on significance levels only.

This is an example of outcome of a risk factor analysis in dairy cows. The animal-based variables recorded on-farm in the dairy cow study were related to environment-based measures that represented potential risk factors for cow welfare. Relevant environment-based measures that were recorded on-farm during the application of the protocol included: the type of housing (i.e., loose housing versus tethered), the water flow (sufficient, partly sufficient or not sufficient), the cleanliness of the drinkers (sufficient, partly sufficient or not sufficient), and the efficiency of the drinkers (sufficient, partly sufficient or not sufficient), the presence of an outside loafing area (present versus not present), and access to pasture (access versus no access). The measure ‘% of lean cows’ was significantly affected by the type of housing ($P < 0.01$), and by the presence of an outdoor loafing area ($P < 0.05$). On farms with loose housing, the % of lean cows was lower than on farms where cows were tethered (predicted means with SE: $12.8\% \pm 1.30$ and $22.8\% \pm 5.01$ for loose housed and tethered cows, respectively). On farms with an outdoor loafing area loose housing, the % of lean cows was lower than on farms without an outdoor loafing area (predicted means with SE: $11.0\% \pm 1.58$ and $19.0\% \pm 2.57$ for farms with and those

without an outdoor loafing area, respectively). The type of housing also had a significant effect ($P < 0.001$) on the % of cows with severe skin injuries (predicted means with SE: $26.6.8\% \pm 3.12$ and $72.98\% \pm 5.81$ for loose housed and tethered cows, respectively). The % of cows with severe skin injuries was also significantly affected by the cleanliness of the drinkers ($P < 0.001$), and the water flow of the drinkers ($P < 0.001$). On farms with clean drinkers the % of cows with severe skin injuries was lower than on farms with unclean drinkers (predicted means with SE: $85.2\% \pm 10.55$, $54.7\% \pm 7.03$, and $23.7\% \pm 3.90$ for farms with unclean, partly clean or clean drinkers, respectively). Similarly, on farms with a sufficient water flow, the % of cows with severe skin injuries was lower than on farms with an insufficient water flow (predicted means with SE: $71.2\% \pm 5.64$, $29.5\% \pm 22.51$, and $25.0\% \pm 4.16$ for farms with an insufficient, a partly sufficient or a sufficient water flow of the drinkers, respectively). However, these associations should be treated with care (as should significant correlations between different animal-based measures), since they do not necessarily reflect causal relationships. There might be other intervening or confounding factors involved in these effects, and these should be considered as well. Nevertheless, the demonstration of potential risk factors in a dataset obtained on farms in practice represents a first step in the identification of strategies to improve farm animal welfare in terms of housing and management.

INVESTIGATING SOCIETAL VALUES ON FARM ANIMAL WELFARE: THE EXAMPLE OF WELFARE QUALITY®

Mara Miele and Unni Kjærnes

5.1 EUROPEAN CONSUMER CONCERNS AND VALUES ABOUT THE WELFARE OF FARM ANIMALS

The problems facing the farm animal population in Europe over the last decades (from BSE to FMD, salmonella and dioxin) have led to a ‘crisis of consumer confidence’ in the European animal farming industry. Two recent Eurobarometer surveys of public attitudes towards farm animal welfare found that a large majority of the respondents was concerned, especially for the treatment of laying hens in battery cages, broiler chickens, and pigs, while the welfare of dairy cows was not considered a problem. A significant minority of respondents thought about animal welfare while shopping for meat and other animal products and a high percentage of them declared that they were willing to pay a (small) premium price for animal products obtained in an animal friendly way. 62% would be willing to change the place they shop to find animal friendly products. The lack of availability of animal friendly products in ordinary shops is a problem and many lamented the lack of information of welfare friendly labels. According to Eurobarometer 2007, about 70% of European citizens believe that buying welfare friendly products can improve animal welfare (Miele and Evans, forthcoming).

In recent years animal welfare NGOs and actors in the animal supply chains have developed market initiatives to improve animal welfare, communicated to consumers via food labels: In Great Britain, we find the pioneering initiative of the Freedom Food label, organic and free range rearing systems are developed, and, more recently, McDonald’s Europe and several quality retailers have produced guidelines which include animal welfare conditions for their suppliers. However, to date, there is no reference method for assessing the actual welfare achieved on farms and for checking the various welfare claims. Welfare Quality® aims to fill this gap at a European level by assessing all aspects of animal welfare on farms and at the time of slaughter.

Animal welfare is a multidimensional concept, embracing health, physical comfort, fitness, expression of behaviour, and absence of stress (Fraser, 2008). One dimension is unlikely to compensate for another, e.g. better health does not justify that animals are prevented from expressing their behaviour. Hence, assessing welfare requires that all aspects are checked. But animal welfare is also a very malleable and contested concept (Miele and Bock 2007) and levels of health, physical comfort, etc. deemed unacceptable tend to change over time (Fraser, 2008). Views differ among various actors in the supply chains, with higher welfare expectations among consumers than farmers or other supply chain actors (Miele and Parisi, 2001; Bock et al. 2007, 2008; Miele and Evans, forthcoming). Therefore technical definitions of animal welfare are always intertwined with ethical judgements. In developing the WQ protocol we tried to address this multidimensionality as well as the complexities of ethical judgements in European societies, in order to ensure the robustness and legitimacy of a protocol that should become the European ‘animal welfare standard’.

A variety of mechanisms have been built into the project to grasp the opinions of ‘the European public’ on the construction of an animal welfare standard, to promote a sustained dialogue between animal scientists and social scientists, and bring in the voice of future users (consumers, farmers, retailers, NGOs and others) early enough to have an impact on the research design and outcome. Investigations of the public were conducted through focus group interviews of citizens, population surveys, and a large number of in-depth interviews with farmers, retailers, industry representatives, certifying bodies, and NGOs. At regular intervals the animal scientists and the social scientists have met at ‘integration meetings’ to discuss the results from their respective research conducted in parallel. From these interactions a cross-fertilisation of ideas and a common language was developed. In the final stage of the project a new research method was adopted to gauge the opinion of the public about the WQ standard: citizen juries. The task was designed by the social scientists in the project but was conducted by social scientists and animal scientists together. A parallel investigation was carried out with farmers: farmers’ workshops.

By collating this information, we have produced a detailed and dynamic picture of opinions and practices regarding farm animal welfare across Europe (for a brief overview, see Kjærnes et al., 2007b).

In this chapter we concentrate on analyses of European citizen’s opinions on farm animal welfare, presenting some central findings from the focus group discussions on welfare principles and criteria and population surveys, both studies carried out in the seven European countries that have been studied in SP1: France, Great Britain, Hungary, Italy, the Netherlands, Norway and Sweden..

5.2 FOCUS GROUP DISCUSSIONS: CITIZENS' VALUES AND POSSIBLE GAPS BETWEEN PUBLIC UNDERSTANDING OF ANIMAL WELFARE AND SCIENTIFIC DEFINITIONS

Focus group discussions were conducted in 2005 when the first working protocol of the WQ assessment and monitoring scheme was developed. A broad interview guide allowed the participants to speak for themselves, to identify, define and prioritize their concerns, and relate them to the list of concerns and parameters (Evans and Miele, 2007; Miele and Evans, forthcoming). This method offers qualitative information that can be deeper than that obtained in single interviews. The trade-off is that in allowing the group to take its own course we risk some members conforming to group pressure and remaining silent when their experience is not common. We were eager to recruit 'ordinary' consumers rather than people already highly motivated by animal welfare issues. The overall threshold for participation therefore included people with 'a bare minimum level of interest in either farming or animal welfare'. The selection criteria were meant to ensure that people from different socio-demographic and lifestyle backgrounds were included, thus allowing us to explore the full range of discourses associated with animal welfare and welfare-friendly food products. Each group was relatively homogeneous; urban mothers, rural women, couples without children, seniors, young singles, political and vegetarian consumers, respectively (Evans and Miele 2007).

The discussions dealt with the consumption of animal foods and issues of human-animal relationships in the context of food consumption practices (Table 5.1). We addressed the participants' knowledge about animal farming, what kind of information is considered relevant for assessing the 'animal friendliness' of available products, sources of information, and their opinions about public institutions and private organisations in providing such information. Moreover, we explored how people define the welfare of farm animals and their opinions on the list of areas of concerns developed by the WQ scientists.

5.2.1 Consumers' General Knowledge about Farming Practices and Farm Animal Welfare

Despite clear cultural differences between the seven countries, some observations regarding consumer knowledge about farming practices and farm animal welfare seem to apply across the majority of our study countries. First and foremost, there is a general lack of knowledge about these subjects. In the Netherlands, knowledge about farm animal welfare and farming practices was found to be fragmentary, ambivalent and tainted by negative emotions. In Italy, the majority of participants seemed to have little knowledge about practices currently utilized in modern Italian farming systems and they tended to assume that a small scale/ traditional type of animal farming is still prevalent. In Hungary, consumer knowledge about farming and farm animal welfare was limited mainly to an understanding of how these activities/issues impact upon public health. In Great Britain,

TABLE 5.1 Themes addressed in the focus group interviews.

Themes addressed	Topics within each theme
1. Culinary practices	<ul style="list-style-type: none"> • Food consumption cultures and habits in different countries • The consumption, preparation and purchase of meat, dairy and egg products
2. Consumers' general knowledge about farming practices and animal welfare	<ul style="list-style-type: none"> • Sources of information • Evaluation of available information • Gaps in the provision of information
3. Consumers' knowledge of welfare-friendly food products	<ul style="list-style-type: none"> • Consumers' familiarity with welfare-friendly products • Consumers' familiarity with welfare-friendly certification/assurance schemes (and the criteria behind them) • Perceived pros and cons of different products and schemes
4. Consumers' evaluation of the provision of information about welfare-friendly products	<ul style="list-style-type: none"> • The level of consumer demand for information about animal welfare • Consumer preferences regarding product labelling
5. Consumers' interactions with and perceptions of welfare-friendly products	<ul style="list-style-type: none"> • Perceived positive and negative attributes of welfare-friendly foodstuffs • Barriers to purchasing welfare-friendly foodstuffs • Ethical dilemmas related to the purchase of welfare-friendly foodstuffs
6. Responsibility	<ul style="list-style-type: none"> • Consumer perception of who should be taking responsibility for animal welfare • The perceived roles of consumers in relation to the state with regard to animal welfare • Consumer perception of who is actually taking responsibility for animal welfare • The interconnections between consumer practices, consumer knowledge and notions of responsibility
7. Agency	<ul style="list-style-type: none"> • Strategies of political mobilisation adopted by consumers in relation to animal welfare • Consumers' perception of their ability to influence animal welfare • Consumer boycotts and 'buycotts' of specific meat or animal products
8. Trust	<ul style="list-style-type: none"> • Reflexive and non-reflexive consumption practices • Who consumers trust/distrust to provide reliable information about animal welfare • Why consumers trust some organisations but distrust others • Do levels of trust vary in relation to the specific issue under consideration (e.g. labelling, monitoring)?
9. Consumers' evaluations of a proposed scientifically based standard for farm animal welfare	<ul style="list-style-type: none"> • Participants' spontaneous animal welfare concerns and priorities • Participants' reactions to the list of ten welfare concerns developed by Welfare Quality scientists

too there was a lack of knowledge, apart from specific issues relating to poultry and calves for veal production. However, we must qualify this general conclusion by adding that both the level and type of knowledge varied depending on the social and cultural background of the participants. Thus, for example (as one might reasonably expect), the politically active and vegetarian participants in general possessed a more detailed level of

understanding about these issues. Furthermore, whilst the majority of participants seemed to derive their information on these subjects from secondary or indirect sources, on the whole, those living in rural areas were more likely to have had more direct experience of farms and farming practices than their urban counterparts, although this did not always shape their perceptions in a positive direction.

Second, across most of the countries focus group participants seemed to derive their (indirect) knowledge mainly from the mass media, to a large extent fashioned by scandal driven media that focused predominantly on negative issues. In France, knowledge seemed to be shaped by striking and highly emotive media images and opinions tended to be linked to topical questions. In Great Britain, people's understandings were shaped by mass media that seemed to be more concerned with issues of animal transport and slaughter than animal rearing. The media portrayal of recent food scandals, such as BSE and salmonella, has increased Italian consumer interest in (and knowledge about) farming and certain food products, such as battery produced eggs and foie gras. The mass media seemed to exert a significant influence over consumer knowledge also in Hungary and the Netherlands.

However, a significant number of participants in each country simply did not want to be reminded about the death of animals associated with consuming meat. Moreover, quite a few did not want to be reminded about the animal welfare conditions behind the foods at the time of shopping. These latter ones expected that existing institutions (national and supranational governments, retailers) take care that animal farming is thoroughly regulated and controlled. For example, an Italian participant stated that:

‘Anyway, as to buying products with information about animal welfare, I buy Esselunga products, even though they contain little information, because I trust them and I expect that they control and take care of these issues’ (Italy, Group 1: Urban Mothers).

To summarise, there is a general lack of public knowledge about farming practices and animal welfare problems in particular. The knowledge that people do possess is to a large extent shaped by mass media discourses. Some participants deliberately avoided expanding their knowledge on these subjects in a desire to distance themselves from ethically challenging issues (the death of animals) with an expectation that existing institutions should address these problems. However, the vast majority of participants considered animal welfare an issue that mattered to them and a significant number believed that more information would have an impact on their behaviour (e.g. place of shopping, forms of protest, boycott).

5.2.2 Consumers' Spontaneous Concerns about Farm Animal Welfare

All focus group participants were asked to make a list of their own spontaneous concerns about animal welfare and then to discuss together what issues were more relevant in order to ensure a good level of farm animal welfare (Table 5.2). There is clearly a common understanding of animal welfare among the participants in all countries, with only some

TABLE 5.2 Spontaneous concerns regarding farm animal welfare among focus group participants.

Spontaneous Concerns	France	Italy	Nether-lands	Great Britain	Sweden	Norway	Hungary
Outdoor access, free range, extensive production		X	X	X	X	X	X
Possibility to choose between indoors and outdoors							
Space, natural space	X	X			X		
Natural type of feed	X	X	X	X	X	X	X
No artificial growth stimulants							
Lifespan							
Time for normal growth							
Humane slaughter	X	X	X	X	X		
Transport (limited or avoided)	X	X	X	X	X	X	
Respect	X	X		X	X	X	
Care							
Physical comfort and security							
Good hygiene		X	X				X
Good quality of life	X	X		X		X	
Small scale production		X			X	X	
Breeding, genetic modification				X			
Products with someone 'accountable for' (farmer, vet.)		X				X	
No mutilations, no pains			X		X		
Natural light, fresh air		X	X				
Distractions (playing)			X				
Animals as individuals (name)						X	
Natural reproduction		X					
No use of routine medicines				X			
Wild animals						X	
Company, love, happiness					X		

culturally specific variations. Concerns about animal welfare are clustered into three main sets: animal-environment, animal farming practices, and human-animal relations. The aspect most often mentioned is 'Outdoor access', a general definition encompassing concerns addressing the need for animals not to be confined in closed environments, at least for part of their lives. Many consider 'outdoor access' for farm animals the best compromise between the ideal life for animals (the wild) and the need to rear animals for human consumption. For Norwegian participants, but also Italian, the best life for animals is in the wild, and only for 'wild animals' can we talk about real animal welfare. Other important environmental aspects in the definition of farm animal welfare are 'Space', 'Fresh air', 'Natural light', 'Cleanliness and hygiene of stables'. Feeding practices (especially feed additives, growth additives, hormones, force feeding, unnatural foods) are at the forefront of people's minds in every country, with most emphasis in Italy and France.

For the majority of participants in all countries the quality of life of animals is most important. But the slaughter of animals is problematic and for some a reason for feeling guilty (or uneasy) about eating meat: 'I need to know how these animals are killed...' and

‘To me the way in which animals are slaughtered is the most import issue’ (Italy). In Norway, negative emotions were associated mainly with meat from other countries:

‘Animals from here, they can walk outside in freedom, especially in the summer, and have generally a decent life. There is a big difference in (farming) practices in Germany or other countries, for example England, where these huge herds of livestock are kept inside all the time. To give you an example when I see that ‘Belgium’ or ‘Belgium Blue’, I do not have much desire to eat that!’ (Norway).

Human-animal relationships are mostly defined in terms of farmers’ duty to care for farm animals; a need to ensure respect for animals. Small-scale production is most often associated with better care, better chances for animals to be considered ‘individuals’, identified with a name, and not part of an undistinguished ‘mass’, as in industrial production (factory farming): ‘Respect, space to roam, good feeding ... a good life’ and ‘While it was alive I would like it live how it was suppose to live’ (GB). In more rare cases human-animal relationships encompass intimate emotions: ‘I have never seen a farmer caressing a cow’ (Italy).

5.2.3 Consumers’ Reactions to the Scientist’s List Welfare Concerns

After having discussed their own definition of animal welfare the participants were asked to comment on the ‘Ten areas of concerns’ identified by the WQ animal scientists. The participants received a document describing these areas of concerns and were invited to point to differences and/or commonalities with their own understanding of animal welfare and, more specifically, with the list of concerns that they previously agreed on. Reactions to the ‘scientific’ list of concerns were generally positive. In Italy, all areas were considered important and the language used in the description of the ten areas, especially the last five, gained their consensus and favour:

‘The categories identified by the scientists are better than ours because there is more than what we have said’ (Italy).

‘I dare to say that this list is complete!’ (Italy).

‘The 10 categories selected... represent what we have said but in a deeper way’ (Italy).

In France, however, only the first five areas of concern were accepted as appropriate (‘basic’, ‘common’, ‘realistic’), the second five areas were considered ‘unrealistic’, ‘utopian’ for farm animals for technical and financial reasons, considering the current industrialisation of farming. Some French participants saw the use of terms such as ‘emotions’ or ‘frustration’ to be inappropriate for farm animals because those terms were ‘too human’.

In Great Britain, nearly all groups thought the concerns very comprehensive and addressed most of their concerns, except for ‘genetic modification’ and ‘slaughter’. In Norway the discussion around the ‘ten areas of concerns’ was very limited, most participants felt that the list was covering all important aspects. The only point raised and to some extent debated, was the language used to talk about animal welfare which was ‘too close to human nature’. Some participants said that the list seemed to concern children instead of animals. The emotional life of animals was a more contested issue. In the Netherlands as well there was not much discussion about the list of concerns; the participants did not particularly agree or disagree with these principles. However, a number of Dutch participants expressed worries, and even shock, that areas like ‘prolonged hunger’, ‘thirst’ and ‘malnutrition’ were listed in a future monitoring scheme, thinking that such basic animal needs were already met in contemporary modern animal farming. For Dutch participants all areas of concerns were considered important; when asked, they prioritised hunger, pain, human-animal relations, and injuries, viewing natural social behaviours and positive emotions the least important. In Sweden the general response on the experts’ list was positive (it is important and relevant). The list did not generate much discussion in any of the groups and the comments were mostly very general: ‘it is good and comprehensive’, ‘well formulated, one can understand exactly what they mean. However, comparing their reactions to the expert list with their own list reveals some important differences: ‘Outdoor access’ ‘small scale farming’ and ‘GMO restriction’, ‘quality of animal feed’ ‘natural life circle’ were not mentioned in the expert list while they were relevant in the spontaneous lists of participants’ concerns.

In general, the participants reacted favourably to the experts’ list of areas of concerns for a monitoring scheme and they were positive towards science in general. Most participants identified more commonalities than differences between their understanding and a scientific approach to defining the welfare of animals. However, analysis shows that for the vast majority of consumers the concept of ‘animal welfare’ implies more positive aspects of the life of animals. Hunger, thirst, and mutilations are very important issues but they are considered to be basic needs of animals and should not be listed as welfare. For most participants across the seven countries ‘outdoor accesses’, ‘natural life cycle’, ‘natural feed’ and ‘small scale production’ are viewed as particularly relevant conditions for ensuring the welfare of farm animals, whilst in the ‘expert list’ these aspects are considered opportunities for welfare.

5.3 POPULATION SURVEY: CONSIDERATION OF ANIMAL WELFARE, TRUST, AND FOOD PURCHASES

The focus group interviews have given us in-depth, qualitative insight into people’s understanding of farm animal welfare, but we also needed information on the extent of popular engagement and how that is reflected across Europe. For that purpose

representative population surveys were carried out in the same seven countries in September 2005. Data were collected through Computer-Assisted Telephone Interviews (CATI) (Kjærnes and Lavik 2008). The survey was based on probability samples, 1500 in each country.

The survey confirms generally very positive responses to questions about interest in farm animal welfare. Across the seven countries, a clear majority appears to be interested or very interested, with some national variation, ranging from 65 to 87%. The French and the Dutch seem to be somewhat less interested, the British are in the middle, while Hungarian, Swedish, Norwegian, and, ranking highest, Italian respondents are very enthusiastic. This interest is expressed even in wide-spread demands for information (Kjærnes and Lavik 2008, 61). Many are worried about how animals are treated on the farm, during transport and at the slaughterhouse (ibid, 33). The general association between bad welfare and intensive forms of production is reflected in most worry for poultry (except Hungary, where poultry is often kept in backyards and on small farms), pig farming coming next in most of the countries. Norwegian and Swedish respondents are consistently less worried for all aspects of farm animal welfare.

Still, interest is not the same as action. The questionnaire was constructed so as to find out more about links between concern and welfare friendly purchases. Two major points can be made; first, consumer action is considerably lower than general interest might indicate and, second, what people say about own involvement is much higher than any market share for welfare friendly labels might indicate. Does this prove that people are not morally consistent in their actions, that positive answers are more about self-presentation than actual commitment? The analyses indicate that the answers are much more complicated than that.

5.3.1 Animal-friendly Purchases of Meat and Eggs

While a majority is interested in animal welfare, fewer think about it while shopping. Figure 5.1 compares the proportions being generally interested with those thinking about animal welfare while buying meat. Italians rank highest, together with the Swedes, while the Dutch have here been joined by the Norwegians at the lower end. The biggest 'discrepancy' between interest and purchases is found in Norway, the least in Sweden and France.

The most developed market for animal friendly options is for eggs, including free-range, barn eggs and organic eggs. Eggs can therefore serve as a better indicator of links between interest and (self-reported) purchases. Table 5.3 shows the proportions of respondents emphasising organic production and the welfare of hens, respectively, when buying eggs. Together, these give a positive impression; a majority emphasizes animal welfare when buying eggs. Again the most enthusiastic are in Italy, followed by the French, while the Norwegians remain at the bottom of the list. The national response patterns are therefore quite similar for meat and eggs. If the answers are to be believed, three out of four Italians think about welfare and more than half of them buy organic eggs. The lower engagement

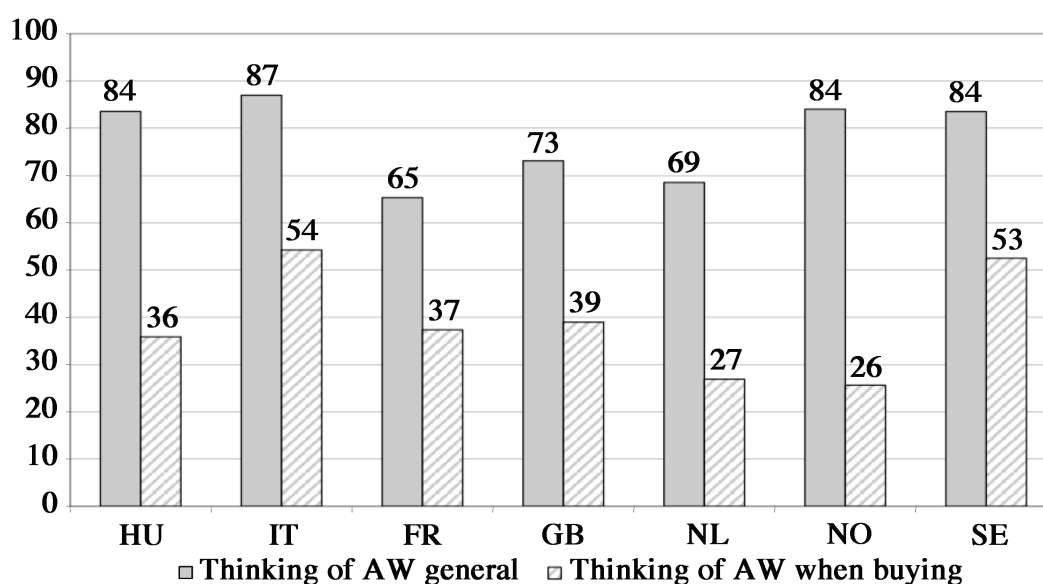


FIGURE 5.1 General interest in animal welfare (among all)* and thinking of animal welfare when buying meat (among those who have bought meat).**

Notes: % proportion 4+5 on a scale 1–5; * ‘Thinking of farm animal welfare in general, how important is this issue for you on a scale of 1–5, where 1 is not at all important and 5 is very important?’, weighted, ‘don’t know’ excluded, N = HU: 1462, IT: 1478, FR: 1497, GB: 1490, NL: 1489, NO: 1493, SE: 1496; ** ‘When you purchase meat or meat products, how often do you think about the welfare of the animals from which the meat has come, on a scale of 1–5 where 1 is never and 5 is always? (Among those who have bought meat)’, weighted, ‘don’t know’ excluded, N = HU: 1249, IT: 1173, FR: 1337, GB: 1330, NL: 1248, NO: 1364, SE: 1334.

TABLE 5.3 The importance of organic and treatment of hens when buying eggs.*

	HU	IT	FR	GB	NL	NO	SE
Organic	39	45	40	23	35	12	25
Treatment of the hens	58	77	73	64	62	41	59

Notes: % proportion ‘very important’ among those who have bought eggs; * ‘Continuing with eggs, are the following factors very important, fairly important or not important at all? A. Organic. B. Treatment of the hens’, weighted, N=1500 in each country, ‘don’t know’ excluded.

among Norwegians might indicate that they are less concerned with animal welfare in their own shopping.

So far we have interpreted animal friendly actions as purchases of specific animal products. A more resolute consumer response is to become a vegetarian or, going further, a vegan. By that they turn down the whole system of animal production. While the extent of veganism in Europe is very limited, there are quite a few vegetarians. These are not clear-cut categories and the figures from our survey as well as from other sources are uncertain. But national differences are quite consistent (Kjørstad 2005). Most vegetarians are found in Great Britain (about 8%), fewest in France (1%), with the other countries in between. There is no clear quantitative support indicating that vegetarianism and veganism have grown in recent years, but symbolically and politically they are important and may serve as a hidden threat in mobilisation against contemporary production methods.

5.3.2 Why Not?

Supply is a strongly limiting factor when it comes to ‘welfare friendly consumer choices’ and there are also problems with transparency and information (Roe and Marsden, 2007). Alone market conditions may explain a lot of the lack of action. However, another reason for not taking personal action may be that, in spite of interest and worry, they see no need for it. A question about whether welfare conditions in their own country have improved or deteriorated showed that, overall, most respondents believe that farm animal welfare has improved over the last ten years in their own country, the only exception being Hungary (Figure 5.2). Only minor proportions think conditions have deteriorated. Many respondents seem to trust that actors involved in meat production are doing their job. This optimism about trends in farm animal welfare is in accordance with earlier findings on food safety and nutrition, where many Europeans tend to be optimistic about changes in modern food provisioning (Kjærnes et al. 2007a). There are, however, strong and consistent differences in which actors they trust; experts and NGOs are generally highly trusted to tell the truth about animal welfare problems, followed by the media and public authorities, while much fewer believe in the truth-telling of market actors (Kjærnes and Lavik 2008, 40). When producers and retailers take action, monitoring by civil society groups, independent experts, and public authorities is crucial. For some people and in some places general mistrust in institutional actors constitutes an important cause for people not to take action as consumers. A person not trusting a label or the intent behind it to improve welfare conditions is not likely to spend money by buying these products.

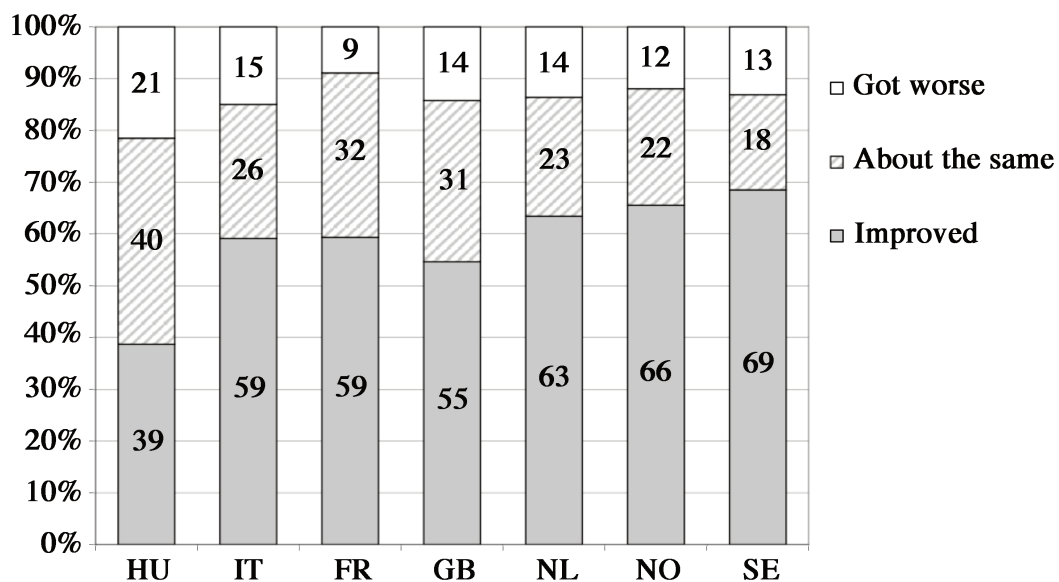


FIGURE 5.2 ‘In general, over the past 10 years, do you think that farm animal welfare in your country has improved, is about the same or has got worse?’

Notes: weighted, ‘don’t know’ excluded, N = HU: 1310, IT: 1365, FR: 1483, GB: 1384, NL: 1434, NO: 1463, SE: 1393.

5.3.3 Packaging Farm Animal Welfare

Considering the optimism and the lack of availability, surprisingly many still claim that they make efforts to improve animal welfare through their food purchases. Everywhere, these proportions are much larger than the market shares for labelled welfare friendly products, including organics (Roex and Miele, 2005). National variations do not reflect variations in availability. Wishing to present oneself as morally ‘better’ than one’s practices reflect is a well known and very human phenomenon. However, positive responses indicate that a norm has been established about what is the right thing to do (there are ‘correct’ answers). This survey, together with those conducted in Eurobarometer, suggests that norms seeing animal welfare as a consumer responsibility are widespread in Europe, even though with some national variations.

The survey analyses as well as the focus group interviews bring several examples of practices which the participants regard as animal friendly (Evans and Miele 2008). Some associate good welfare to particular species, like sheep in Norway and Mangalica pigs in Hungary. More common in some countries is to associate animal welfare with alternative forms of production, often conceptualised as organic, but not necessarily certified as such.

Most influential, however, are references to domestic production systems. Meat produced in their own country, within systems and forms of regulation that they know, is by many seen as more animal friendly. To Swedes, buying ‘Swedish’ meat represents concern for the animals. In a similar way, British respondents seem to trust the quality assurance systems and labelling schemes introduced by the big retailers. Italians, on their part, prefer meat from local vendors in the belief that sticking to local (small) producers also brings better welfare. The ways animal welfare is linked to food consumption are thus framed by specific local/national food cultures, market structures and public discourses. In France many share a broad, inclusive understanding of food quality (Evans and Miele 2007; Kjørstad 2005). High quality food (fresh, authentic and with an exquisite taste) is often regarded as the outcome of production with higher standards of animal welfare. The conception that what is good for the animal is also good for humans is widespread, in particular in France, Italy, and Hungary (Kjærnes and Lavik 2008). In these ways farm animal welfare is mobilized and included in larger stories about nation (like Sweden), place (like Toscana) or brand (like Tesco) – imagined communities and identities. Consumer actions are thus linked to cultural identity as well as market power.

5.4 CONCLUSIONS

Several points can be made on the basis of the focus group discussions and the survey. First, the participants in the seven countries did not prioritise animal welfare considerations while shopping for food but when asked about it they showed high interest and

engagement. Second, in most countries the majority of participants had limited knowledge of animal farming practices, but associate negative welfare with industrial-intensive methods of production (factory farming), large scale production (mass production) and positive welfare with small scale production and extensive systems (free range, typical). ‘Organic’ is unanimously perceived as the most welfare friendly system of production across the seven study countries. Third, product labels and brands are considered useful sources of information for assessing the animal friendliness of products. Their importance varies across countries, consistent with availability: most important in Great Britain, Sweden and the Netherlands, least important in Norway and Hungary. In Italy and France brands are considered important for assessing the welfare friendliness even though most brands have little or no explicit reference to animal welfare. Fourth, many state that they as consumers do make an effort for farm animals, many more than actual market shares indicate. They do so with reference to larger stories, like belief in special breeds or species, domestic production, or local farms, stories that emerge from specific national and local public discourses, market structures and culinary cultures. Fifth, while the consistently positive answers lend support to the view that a norm has been developed in Europe about consumer responsibility for animal welfare, consumer action is strongly dependent on trust in experts, food suppliers and regulatory institutions.

The aim of these investigations, focus group discussions and population survey, has been one step to (re)introduce democratically mandated preferences into the framing and conduct of research activities that take place. The voice of the ‘users’ was introduced at an early stage of the development of the WQ monitoring scheme, in a debate that, up till now, has been mostly limited to actors of the supply chain. In this way WQ has attempted to establish a real science–society dialogue and to engender opportunities for mutual learning and cross-fertilization of ideas and values between scientists in the project and the lay public. This dialogue has been then taken forward in another investigation in Subproject 4, the citizens’ and farmers’ juries in three EU countries on the proposed protocol for the WQ standard.

6

DEVELOPING PRACTICAL WELFARE IMPROVEMENT STRATEGIES

Bryan Jones and Xavier Manteca

6.1 INTRODUCTION

The primary objective of Sub Project 3 (SP3) of the Welfare Quality® Project was to develop and test practical ways of improving the welfare of farmed pigs, chickens and cattle. The research in SP3 was multidisciplinary and its search for remedial strategies embraced environmental, genetic and management approaches. SP3 was divided into six work packages (WPs). Each of these WPs addressed a key welfare issue that was perceived as particularly important by farmers, retailers, academics, legislators, consumers and other stakeholders. The six WPs focused on the following topics:

- WP 3.1. Minimizing handling stress: improving stockmanship;
- WP 3.2. Genetic solutions to welfare problem: a) safeguarding leg conformation and longevity in pigs, and b) psychobiological characteristics and adaptation in dairy cattle;
- WP 3.3. Eliminating injurious behaviour: a) feather pecking and cannibalism in laying hens, and b) tail biting in pigs;
- WP 3.4. Reducing lameness in a) broiler chickens and b) dairy cattle;
- WP 3.5. Minimising neonatal mortality in pigs;
- WP 3.6 Alleviating social stress in a) pigs (genetics of aggression and dietary changes to reduce aggression in group-housed pregnant sows), and b) intensively kept beef cattle.

In Sub Project 2, Welfare Quality® researchers drew together the views and concerns of consumers, representatives of the farm animal industry, retailers, biologists, social scientists and legislators. The resulting information was used to establish the following four principles which are considered essential to safeguard and improve farm animal welfare: good feeding, good housing, good health and appropriate behaviour. Twelve clear criteria were also defined within the 4 principles (Kjaernes and Keeling, 2006; Botreau et al., 2007). These principles and criteria of good welfare formed the basis for the development of the Welfare Quality® welfare assessment systems in each of three main species: cattle, pigs and poultry.

There are at least four important links between the work carried out in SP3 and SP2. Firstly for example, several researchers were involved in both SP2 and SP3, this not only ensured the rapid and effective transfer of knowledge but it also contributed to the parallel development of both Sub Projects. Secondly, the principles and criteria of good welfare developed in SP2 provided a framework for the development of practical welfare improvement strategies in SP3. In other words, the SP3 work packages focused on developing practical strategies that would help fulfil the requirements of several of the SP2 welfare criteria. This linkage is clearly evident in the structure of the Technical Information Resource on Practical Strategies for Improving Farm Animal Welfare. This resource has been developed as part of the SP4 effort and it features many of the findings from SP3 (see below). Thirdly, many of the tests and measures that were used and/or refined in SP2 to assess a farm's compliance with the Welfare Quality® principles and criteria of good welfare were also used to determine the effectiveness of the practical strategies developed in SP3. In turn, this dual-purpose approach also provided an effective examination of the reliability and validity of the SP2 measures used. The fourth link looks to the future implementation of the Welfare Quality® welfare assessment systems devised in SP2. Once the welfare status of a farm has been determined using the above system(s) the feedback of results and the provision of practical advice on remedial strategies will help the farmer to deal successfully with any problems or risk factors that were identified. In this context, the welfare improvement strategies developed in SP3, as well as the Information Resource, will contribute significantly to the advisory component of the cyclical process of farm assessment – feedback and advise – welfare improvement – re-assessment etc.

In the following pages we provide a summary of the Information Resource as well as some concrete examples of the links between principles and criteria developed in SP2 and the practical strategies developed and tested in SP3.

6.2 INFORMATION RESOURCE PRACTICAL STRATEGIES FOR IMPROVING FARM ANIMAL WELFARE

As mentioned above the Information Resource developed in SP4 was built on the results of SP3 and the existing literature. It is closely linked to the principles and criteria of good welfare developed by Welfare Quality® researchers in SP2. This web-based resource will be freely available and it is intended to enable farmers, advisors, veterinarians, researchers and policy groups to easily identify practical strategies that can help to solve specific welfare problems or at least to minimize their occurrence or intensity. The resource is subdivided into two chapters. The first chapter provides the background, rationale and brief descriptions of each of the principles and criteria which form the basis of the Welfare Quality® assessment systems developed in SP2. It also very briefly summarises the welfare

implications of failures to satisfy the criteria as well as the general causes of such failures and potential remedial measures.

The second chapter covers the scientific background and benefits of all the viable welfare improvement strategies identified in SP3 together with additional relevant information on the many remedial methods that have been generated outside the Welfare Quality® project. The six main sections of Chapter 2 respectively reflect the aims of the six work packages in SP3 (see above). Each of these sections is subdivided into two parts. The first part provides a state of the art scientific background to the particular issue under investigation in that section by describing the main environmental, genetic and management variables underpinning its development as well as its welfare and economic consequences, and the advantages and disadvantages of the remedial measures that have been developed by external researchers. The second part focuses on the viable welfare improvement strategies that were identified in SP3. In order to qualify as viable these strategies / recommendations not only needed to satisfy both welfare and economic requirements but they also had to be practicable, i.e. safe, affordable and easy to implement by the farmer and/or breeding company. Instructions for their implementation are also provided in the resource.

Thus, the resource web site and associated material offers different levels of information for different degrees of interest, with clear links to other more detailed sources of information such as the scientific papers cited, relevant reports, regulatory documents, websites etc.

6.3 NEONATAL MORTALITY IN PIGS

Neonatal mortality clearly falls within the scope of Criterion 6 (Absence of injuries other than those due to disease or voluntary interventions) that was identified by SP2 researchers. Pigs show a high prevalence of neonatal mortality; indeed data from the UK indicates that 11.85% of all live-born pigs die within the 72h post-parturition period (Meat and Livestock Commission, 2006). Besides constituting an important economic problem, piglet mortality is also becoming an increasingly significant welfare concern. The availability of reliable on-farm measures of welfare, such as those developed in SP2, is essential in order to determine the efficacy of any proposed remedial strategy.

Neonatal mortality in pigs is a complex multi-factorial problem that involves elements of piglet health status and behaviour, the sow's behaviour and the physical environment (Baxter et al., 2008). The most common cause of death is crushing or overlying by the sow; the traditional way of dealing with this is the farrowing crate which effectively reduces the likelihood of crushing by restricting the movement of the sow (Edwards and Fraser, 1997). However, confinement at the time around parturition seems very stressful for the sow. This may be particularly distressing for sows that are accustomed to being

loose-housed during gestation, a practice that will become mandatory for all farms in the EU in 2013. Additionally, the use of farrowing crates might compromise aspects of maternal behaviour that could promote offspring survival (Baxter et al., 2008). In any case, public concerns regarding animal welfare may limit the use of restraint systems in the future, thereby highlighting the pressing need to develop alternative welfare-friendly strategies for preventing crushing (Edwards, 2002).

Although crushing is often viewed as the ultimate cause of most piglet mortality it may be just the end point of a chain of events that began even before parturition. In fact, current research places increasing emphasis on the biological characteristics of both the piglet and the sow. The piglet's development and physical condition at birth has a major impact on survival. Stillborn mortality is correlated with reduced body weight, a disproportionately long and thin body, abnormal shape proportions, and being late in the farrowing birth order (Baxter et al., 2008). Live-born mortality is also highly dependent on the vigorousness of the piglet, irrespective of body weight. Less active individuals face a higher risk of being crushed through a variety of interplaying factors: they take longer to locate the udder and suck the colostrum, which in turn retards their growth and increases the risk of hypothermia and starvation. Piglets that are cold seek closer contact with the sow, thus raising the likelihood of crushing. Indeed, crushing is more prevalent in outdoor (colder conditions) than in indoor herds (Edwards, 2002). A lack of vigorousness and a poor physical condition of newborn piglets are correlated with a number of physiological, e.g., rectal temperature, and laboratory measures, e.g. reduced plasma concentrations of urea, phosphor, calcium and a poorer index of *in vitro* cellular immune function (Tuchscherer et al., 2000).

Encouragingly, many aspects of piglet survival are heritable and there seems to be sufficient genetic variance to allow economically viable selection for welfare-friendly characteristics (Knol, 2002). Since neonatal mortality in pigs can also be reduced through improved husbandry one of the objectives of WP3.5 in SP3 was to consider the effects of genetic selection for survival in farrowing systems other than the conventional farrowing crate. Piglets and sows from two genotypes (High Survival and Average lines) were farrowed in an outdoor system and an indoor loose-housed system. The selection strategy had its greatest impact outdoors with only 12% total mortality in the High Survival line compared with 18% in the Average line. Regardless of farrowing environment, maternal behaviour was influenced by genotype and High Survival sows were better mothers; displaying more careful behaviours than Average sows. They also performed significantly less crushing behaviour during farrowing than the Average sows.

6.4 LAMENESS IN BROILERS

Lameness is included in the definition of Criterion 7 (Absence of disease) in the Welfare Quality® principles and criteria of good welfare. Lameness resulting from leg disorders is one of the main welfare problems in broiler chickens (FAWC 1992; FAWC 1998; European Commission 2000). Indeed, because intensive broiler chicken production now exceeds 2 x 10¹⁰ birds worldwide (Dawkins et al., 2004), lameness is probably one of the most widespread welfare problems in modern agriculture. Lameness also has economic costs, some birds may have to be culled and the surviving lame birds are likely to lose weight and to be downgraded at slaughter (Kestin et al., 1999). As many as 90% of birds in some flocks show at least some degree of lameness by slaughter age (Kestin et al., 1992), and some studies report up to 30% of birds moderately to severely lame (Sanotra et al., 2001). However, the incidence of lameness varies considerably between farms, e.g. Dawkins et al. (2004) found a mean percentage of severely lame birds of 9%, with a range of 0 to 20.

Researchers in WP3.4 asked if broiler lameness could be alleviated by sequential feeding (S) of two diets. The novel regime they developed involved ten 48-H sequential-feeding cycles from 8 to 28 days of age. Three treatments were compared: the normal complete diet (C) and two alternations of diets varying in protein and energy contents (S1: E+P- followed by E-P+; S2: E-P+ followed by E+P-). All chickens received the same feed during the starter and finisher periods (0-7 and 29-38 d of age). Body weight, feed intake, general activity and gait score, bone quality and carcass conformation were measured to evaluate leg condition and general performance. These measures feature strongly in the SP2 assessment system.

Gait score was improved in birds receiving sequential feeding (mean GS = 2.41 vs. 2.61 in controls) although this only reached significance in birds fed with the poor-energy/high-protein diet during the first day of each cycle. The reduced lameness in the sequential group could reflect the increased motor activity observed during the sequential feeding phase which in turn reflected the longer times spent feeding and foraging. Feed conversion and carcass conformation were not impaired by sequential feeding. Clearly, lameness in broilers can be reduced by slowing down their early growth rate and speeding it up once their bones have developed.

In short, this novel sequential feeding regime not only reduced instances of lameness but also brought the broilers up to standard slaughter weight without the need of any additional feeding days. This innovative feeding method could improve the birds' welfare by reducing lameness while safeguarding the farmers' profits at the same time.

6.5 AGGRESSION IN PIGS

Aggression features in SP2's Criterion 9 (Expression of social behaviours - balance between negative, such as aggression, and positive aspects, such as social licking). Post-mixing aggression in commercial pig production is common, compromises welfare and profitability and cannot be significantly reduced by low-cost changes to the environment. A genetic component to individual aggressiveness has been described in pigs so selective breeding against this characteristic should be possible if an easily measured indicator trait can be shown to be genetically associated with aggressive behaviour. However, selection aimed at reducing aggressive behaviour might also exert negative effects if genetic correlations exist between aggressiveness and other traits including practical (e.g. reduced handling ease) and/or ethical and welfare (e.g. reduced responsiveness or inactivity) ones. Therefore, researchers in WP3.6 estimated the genetic contribution to individual aggressiveness and, through the analysis of genetic correlations, they validated a method of predicting a pig's likely involvement in aggressive encounters based on a count of skin lesions (lesion score, LS) suffered after mixing. They also established the genetic correlations between aggressive behaviour and other traits.

Aggressive behaviour was recorded continuously for 24h after mixing and lesion scores (LS) were recorded at 24h and 3 weeks post-mixing in 895 purebred Yorkshire pigs and 765 Yorkshire x Landrace of both sexes. Pigs' behaviour during handling as well as their general activity was also scored in the same population. All pigs were housed in partially slatted pens with straw bedding.

Two behavioural traits had moderate to high heritability similar to that of growth traits; these were the duration of involvement in reciprocal fighting and the delivery of non-reciprocated aggression (NRA). Conversely, receipt of NRA had a lower heritability. Genetic correlations suggested that lesions to the anterior region of the body apparent 24h after mixing were associated with reciprocal fighting, receipt of NRA and, to a lesser extent, delivery of NRA. Lesions to the midsection and rear were associated primarily with receipt of NRA. Pigs which engaged in reciprocal fighting delivered NRA to other animals but rarely received it themselves. A genetic merit index using lesions to the anterior, central and rear regions recorded 24h post-mixing as separate traits should allow selection against animals that participate in reciprocal fighting and in NRA. Positive correlations between LS observed 24h and 3 weeks after mixing, especially for lesions to the centre and rear of the body, indicate that this score is predictive of those received under stable group conditions. Selective breeding for reduced post-mixing LS should have a long-term ameliorative effect on aggression-related injuries even after dominance relationships are initially established; in other words the pigs will be generally less aggressive.

Inactivity was weakly heritable and negatively associated with bullying, suggesting that pigs selected for reduced aggression might also be slightly less active. The ease with which

pigs entered and exited a weigh crate had low positive genetic correlations with aggressive behaviours, although aggressive pigs were also more active during weighing.

In conclusion, genetic selection for reduced aggression is feasible. Fighting and bullying post-mixing were moderately heritable, and skin lesion counts 24hrs after mixing could be used as a proxy trait. Because of the low genetic correlations, selection for reduced aggression is likely to have only a small negative impact on the ease of handling at weighing.

6.6 STOCKMANSHIP

Good stockmanship is essential for satisfying the requirements of the Criterion 11 (Good human-animal relationship - no fear of humans). The term 'stockmanship' covers the way animals are handled, the quality of their daily management and health care, and how well problems other than disease are recognised and solved (Waiblinger and Spoolder, 2007). The quality of stockmanship has a profound effect on the animals' welfare and productivity (Hemsworth and Coleman, 1988; Boivin et al., 2003; Waiblinger et al., 2006). For instance, despite centuries of domestication exposure to human beings remains one of the most potentially alarming experiences for many farm animals. More specifically, unless they have become accustomed to human contact of either a neutral or positive nature the predominant reaction to people is one of fear (Duncan, 1990; Jones, 1997). Chronic fear is a major welfare problem that can lead to handling difficulties, injury, and stress as well as impaired growth, reproductive performance and product quality (Hemsworth and Coleman, 1988; Jones, 1997; 2004).

Despite the experimental evidence accumulated over the last 20 years, we still need to improve our understanding of the variables underpinning human-animal interactions and the resultant human-animal relationships. These include: the ability of farm animals to discriminate between humans, the best methods for selecting and training stockpersons, and potential interactions between the effects of handling and housing systems (Hemsworth, 2003; Raussi, 2003).

Welfare Quality® researchers studied the attitudes & behaviour of beef and dairy cattle farmers to human-animal contact and the human-cow relationship in WP3.1. A questionnaire covering the perceived ease of handling cattle, the husbandry conditions, the farmers' attitudes towards cattle, and their behaviour during husbandry and handling procedures was sent to 300 beef cattle farmers / breeders in France. Some of these were then visited and interviewed regarding husbandry and handling. Their calves' behaviour was also observed in a crush test in the presence of a human.

Researchers also observed transfers of beef bulls from commercial farms to a slaughter plant, (1,202 bulls from 108 farms). Questionnaires on attitudes towards bulls and to working with them were completed by 88 farmers, and plasma cortisol concentration and meat pH were measured (i.e., more SP2 measures). Similar surveys of farmer attitudes and behaviour were conducted at dairy farms in Austria and Italy together with direct assessment of the cows' reactions to humans in standardised tests (with or without prior handling) and of the stockpersons' behaviour when milking. The tests, measures and analyses used in these studies all feature in the assessment systems developed in SP2.

Both beef and dairy farmers emphasised the importance of good human contact (quality and frequency) in increasing the ease of handling. The fact that 28% of them failed to recognise genetic background as important in determining the ease of handling is particularly surprising since the temperament of heifers or cows was the first trait they considered in decisions on culling. Farmers sometimes showed negative behaviours (hitting, shouting), but most of them agreed that calm, gentle and patient handling is very important. Calves were generally much calmer if the farmers enjoyed contact with their animals than if they had little interest in them. Encouragingly, calves that were gently handled during the first weeks of age were consistently and durably (up to 40 weeks) less fearful in human approach tests (described in SP2 Technical Documents) than non-handled ones. Furthermore, beef bulls from farms where the farmers had positive attitudes towards them showed lower cortisol and better meat quality after transport.

Researchers in this Work Package also built on the data and material (photographs, videos, technical reports etc) generated in SP2 studies as well as information in the existing literature to develop a multi-media training programme for stockpersons working with cattle, pigs and poultry. The training package utilises cognitive-behavioural intervention techniques to specifically target those key attitudes and behaviours of stockpersons that are known to have a direct effect on farm animals' fear of humans. A 'Quality Handling' training programme (incorporating software, trainers' manuals, newsletters etc) has been developed through intercontinental collaboration and then tested in the various species. It describes how animals' fear responses to humans vary between farms, how fear of humans can damage welfare, productivity and ease of handling, how to improve the stockpersons' attitudes and behaviour and thereby build a positive human-animal relationship etc. This package will soon be made commercially available in a number of languages for the cattle, pig and poultry industries.

6.7 CONCLUDING REMARKS

The fact that several Welfare Quality® researchers were involved in various parts of both Sub Projects 2 and 3 undoubtedly contributed to their effective (and often parallel) development by facilitating the rapid cross fertilisation of expertise, techniques, measures

and emerging knowledge. We hope that the results of such synergy are evident in our descriptions of the paths followed in the development of the remedial strategies and the technical information resource described above. Furthermore, the effective uptake and implementation of the welfare assessment systems built in SP2 demand a cyclical process of a) assessment, b) identification of problems and risks, c) feedback of results and of advice regarding potential solutions, followed by d) future reassessment etc. The practical welfare improvement strategies identified in SP3 provide an extremely important contribution to the advisory component of that process.

6.8 ACKNOWLEDGEMENTS

We are grateful to the many researchers who contributed to the overall research carried out in Sub Project 3. More specifically, we would like to thank those whose work is especially featured in this chapter; they include Emma Baxter, Alistair Lawrence, Rick D'Eath, Simon Turner, Xavier Boivin, Suzanne Waiblinger, Christine Leterrier, Marko Ruis and Hans Spolder.

PREPARING THE ASSESSMENT SYSTEM FOR IMPLEMENTATION

Andy Butterworth

The main area in which the large amount of preparation of welfare assessment systems and methods carried out in Sub Project 2 (SP2) are brought together is SP4, which brings together many strands from different parts of the Welfare Quality® project – social science and focus group activities from SP1, welfare assessment tools from SP2, welfare improvement strategy tools from SP3 and implementation studies on Pig, Cattle and Poultry farms. This chapter deals with six specific areas of work; intervention studies, knowledge resources, implementation scenarios, economic and socio economic effects of implementation, citizen and farmer juries and, finally, technical documentation. Many of the projects in SP4 link with one another, and in the attached flow chart (Appendix 1 and 2), you will be able to see the inter relationships between the projects in SP4 and the previous strands of the Welfare Quality® project.

7.1 PREPARING FOR IMPLEMENTATION – THE INTERVENTION STUDIES

The Work Packages in SP4 which have the clearest direct link with the work carried out in SP2, and which are refinements of SP2 work and steps toward implementation, are the studies in WP4.6.1 Pigs, WP4.6.2 Cattle and WP4.6.3 Poultry. In these we have continued to use and to refine the animal assessment tools first developed in SP2 on farms in Austria, Germany, Italy, Sweden, UK, Spain and the Netherlands. These studies ask a seemingly simple question – if the farm animals are assessed for lameness, tail biting, body condition etc. and we provide the producer with information on how he is doing compared to other farmers – can this information be successfully used to support management decisions, and lead to improvements in both animal welfare and in farm production.

The structure for all of the three species implementation studies was similar. At the start of the farm visits, a farm questionnaire was completed with the assistance of a representative of the farm, which provides a description of the farm, house and herd or flock (i.e. resource-based measures). The areas of resource-based measures required to

give a good background and risk factors for, this example, might include pig, cattle and poultry implementation studies.

7.1.1 Pig Implementation Studies

The Welfare Quality® protocols have been used to assess farms in Sweden, Spain and UK. Farmer perception and attitudes to the values and costs of assessment were gathered in three stages from farms in each country. This task links closely with the activities in WP4.4.2, where complementary analysis of farmers' expectation and levels of acceptability of the outcomes and grading of the on farm assessment and monitoring scheme through different farming systems is being carried out. Firstly, an initial structured, in-depth interview was carried out (before an assessment had taken place). Following an initial farm assessment using the Welfare Quality® protocols, the results of individual assessment categories, were fed back to the farmer. An example of an extract from a farmer report is given in Figure 7.1.

The farmer's perception of the outcome was then assessed a few months after receiving their score through a second structured interview to identify whether the assessment score has led them to make changes on farm, altered their commercial practices, and whether their perception and attitude to the assessment had changed. In particular farmer responses to the following were sought:

- How can the assessment system be used to maximum effect?
- What are the impressions of the assessment system protocol – its limitations/potential?
- Do farmers foresee any commercial benefits/disadvantages from the assessment scheme?
- How have farmers found the welfare assessment process?
- The potential of the assessment scheme for raising awareness of individual farm performance.
- Do farmers see assessment as a tool that could provide information about techniques to improve animal welfare including stock-handling skills, to improved design of animal housing?
- What have farmers learnt from the assessment?
- How could improved performance be supported through the provision of advice or training?
- What do farmers perceive as a reward for improved performance?

The results of the initial (benchmarking), and subsequent farm assessments are then used to assess whether the process of assessment (and feedback to the farmer of assessment results) results in detectable changes in the measures made – the aim being to determine whether the information provided in an inspection report can be of value to the farmer to allow him to make plans (for example in health plans), or make management or housing changes which alter the outcome of the animal welfare parameters measured. For example – if the assessment shows that lameness in pigs is a specific problem, can information

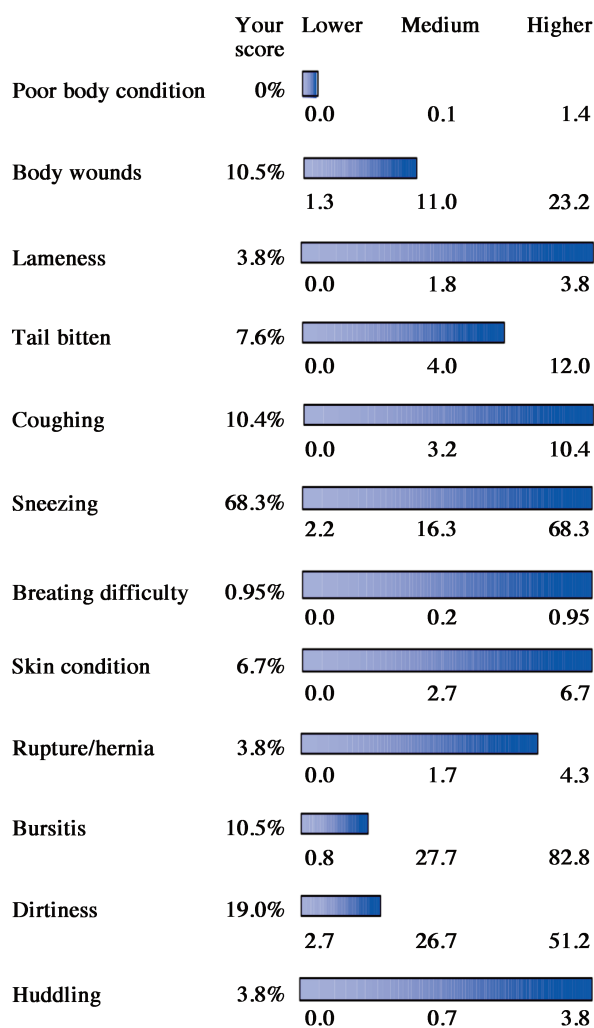


Figure 7.1 Extract of information provided to a pig farmer participating in the study. The figures allow the farmer to see his ‘individual’ score for each measure and to compare it with the overall range of scores seen within the entire study group.

Notes: for example, this farmer has no animals with a poor body condition score, so his score is lower than average, he scored just below average for the number of body wounds on his animals, and among those farmers with the highest levels of lameness.

provided by the assessment be used by the farmer to alter the conditions which affect lameness in his animals.

7.1.2 Cattle Implementation Studies

The perceptions and attitudes towards the Welfare Quality[®] assessment system of beef farmers in Italy, Austria and Germany were studied using interviews. A number of these farms were then subjected to a pre-, interim- and post-assessment using the Welfare Quality[®] tool as derived from SP2. The contribution of the system to welfare improvement was investigated using different levels of feedback of information (i.e. control, written report, additional advice). The feasibility of the practical application of the measures and

the combined 'tool' was evaluated by its application in the field in 'close to commercial' use.

During this process, adjustments to the checklists already trialled in WP2.4 were made, and also adjustments and refinements to the assessment protocol and the training material used to prepare assessors. Pre-, interim- and post-assessment of farms using the Welfare Quality® assessment tool was then carried out in order to determine whether the provision of information based on these assessments or additional advice contribute to improving welfare on the farms.

From the farms visited in the first part of this study, a number in each country were selected for 'post-assessment' using the Welfare Quality® assessment tool. The selection aimed to create a sample of farms with large deviation in terms of welfare levels and also based on the decisions taken in WP2.3 regarding acceptable welfare levels for inclusion into a Welfare Quality® system. The post-assessments were made after approximately 6 months, each time in the same way with farms allocated to three groups:

- Control: Only assessments are carried out.
- Feedback: At the second farm visit (interim) the farmer received written information about the farm performance.
- Feedback and advice: At the second farm visit (interim) the farmer received written information about the farm performance. This was discussed with him or her and advice given on how improvements may be achieved. (Note – it is clear that, in practice, it is necessary to keep advice and assessment separate, here, for feasibility reasons, it was carried out by the same persons.)

7.1.3 Poultry Implementation Studies

This study focused on broilers chicken for two main reasons. First of all the production cycle of broilers is relatively short and so it is possible of assess the welfare of different flocks and check any improvement of animal welfare as a result of farmers' interventions. A second reason for focusing on broilers is that it allowed us to link this project to another project where the use of camera systems installed in the slaughterhouse allowed automated checking for foot pad lesions. The study was carried out in three countries; Netherlands, Italy and the UK, in order to include countries that have well established broiler production industries, but which may differ in how farmers perceive animal welfare.

Building on and developing the existing work carried out in WP2.2 and WP2.4, the expertise, protocols, and training materials developed by these Work Packages was used to ensure that trained assessors with access to the operational protocols were able to carry out the assessments. Training of the teams took place in the Netherlands, and the protocols and assessment documents to be used were agreed between the three country teams.

In order to understand farmers' perception and attitude of the implementation strategies, the participating farmers were interviewed on the farm to inquire about the following issues:

- The conditions under which to best implement the farm assessment tool
- The farmers interest in the information that the tool delivers and the comparative information that this may provide to enable 'benchmarking' with other farms
- The relevance of the information in terms of knowledge and awareness but also practical use
- Their interest in specific advice and training
- The relevance and benefit of a 'good welfare performance' reward.

Three groups of farmers were interviewed:

- Farmers who received information about the assessment but experience no on-farm assessment (control group)
- Farmers who experienced on-farm assessment plus feedback on individual and benchmark results (written report)
- Farmers who experienced on-farm assessment and individual advice (written report, plus personal explanation).

In general, the studies explore the extent to which it is possible to inform the farmer about the extent (and perhaps 'type') of different welfare conditions on his farm. With time, and after analysis, it may be possible to determine whether a pattern of risk factors emerges which could allow the producer to make decisions which can reduce these welfare impacts (for example lameness, skin lesions, bullying of certain animals). After a farm assessment, the results for individual measures can be fed back to the farmer, and aggregated scores can then be created using combined weighted sums. By combining resource-based information with animal-based measures, the potential exists to provide a powerful tool for informing the farmer of the welfare status of his animals, to enable him to see how he compares to other farms, and also to support improvements and management decisions. The farmer can receive a result for each single assessment measure, for example, how many thin cows he has. He could also be given the information collected together into grouped data, for example health measures, measures linked with environmental factors such as skin lesions, hock and foot damage, or behavioural measures. Finally, it is possible to combine all measurement results to give an aggregated overall score and it may be possible in the future that retailers and consumers would make purchasing decisions for some animal products based on combined welfare assessment scores.

By tackling welfare problems through a management support route, not only may welfare be improved, but it is also possible that there will be increased farmer productivity through reduced losses from disease, lameness, damage during slaughter etc. In these 'implementation projects' (4.6.1, 4.6.2, 4.6.3), we have been building links with commercial farm partners to provide animal welfare information which we hope is of real use to the business. An example is the use of automatic cameras in slaughter houses in UK, Netherlands and Italy, which score and record the levels of foot pad problems in

broiler chicken. When this information is added to the farm management systems operated by the companies, it is possible for the company and its producers to see the direct link between ‘welfare and quality’.

Good reference material has been required for the practical assessment protocols (see later) and in the implementation studies described earlier in this chapter. Examples of reference material have included photographic scoring scales for skin lesions, skeletal abnormalities, disease conditions – and also reference material for scoring locomotion, behaviours, and responses of the animals to the test protocols to be used in the assessment scheme.

Sequences of videos and reference photos may be required in the future to enable training of individuals who might chose to adopt the Welfare Quality® protocols. This is likely to take the form of ‘example’ training material where the techniques used in the assessment protocols are demonstrated using filmed examples of the assessor carrying out the technique. Some of this has been filmed in real farm situations on farm in SP2 and SP4. This material includes standardised training material which can be used to assess pre, interim and post training performance for personnel being trained.

7.2 KNOWLEDGE RESOURCES

In WP4.3.1 a resource of places to find existing and new knowledge on improvement strategies suitable for solving welfare concerns is being collated, alongside production of a resource of information that will enable advisors and farmers to identify appropriate improvement strategies for welfare. This material is derived from the studies carried out in SP3 and also from specific work produced in SP2. The intention is that this material will be available as a web based resource to allow producers, advisors and people involved in on farm assessment to explore practical improvement strategies and advice.

In WP4.4.1 teams in France and the UK have carried out work with certification and inspection bodies to examine the mechanisms by which evaluation criteria and requirements are introduced and implemented within food certification and assurance companies. This has thrown up a lot of interesting findings, questions and challenges.

7.3 IMPLEMENTATION SCENARIOS

In WP4.4.3 possible scenarios were created - ways in which the Welfare Quality® scheme may be used in the future, how it could be incorporated into farm assessment systems, and how a support system to maintain the technical methods and documents required could be used. The information used to explore these scenarios is derived from the social science findings of SP1, the assessment scope and practicality information derived from SP2 and the work with the Advisory Committee. This task has focussed on practical considerations regarding the management and control of the assessment system and the infrastructures and tools required for its support. These have been analysed in the context of different possible 'scenarios' for how the assessment system can be used and the implementation of the various outputs of both the assessment and information systems.

Scenario planning is a strategic management tool that organizations can use to create a better understanding about the future. Scenarios are 'focused descriptions of fundamentally different futures presented in coherent script-like or narrative fashion'. The use of scenarios in the scenario planning method is therefore different from scenario's that are based on statistical simulations of uncertainties. The focus is not on single-line forecasting, or making precise predictions, but on bounding, and better understanding future uncertainties. The method becomes therefore more useful as phenomena become more complex and more uncertain. Compared to the Delphi-method, scenario development is, for example, better equipped to identify strategic issues and to deal with broad problem scopes. In general, the method helps to learn how the future could be, given all its uncertainties and complexities. It thus helps to identify strategic issues that deserve attention in strategy or policy development regarding the Welfare Quality® assessment and information systems.

In the first round, two so-called, forced scenarios have been developed. One scenario describes an 'all positive' pattern of events and developments taking place in the future, the other an 'all negative' pattern. They are thus two extreme values from a wide range of possible scenarios. The scenarios were subsequently discussed with stakeholders. From these discussions it became apparent that uncertainty regarding the implementation of the Welfare Quality® system focuses on the different roles that a body would fulfil supporting the implementation. The second round of scenario development described three scenarios on the basis of these different roles. Although several important questions remain, new insights have been drawn from the scenarios. From these insights one may determine different strategic options that would stimulate the adoption of the Welfare Quality® assessment system.

The most important lesson that can be drawn from the forced scenarios which were created in WP4.4.3 is that a positive impact on animal welfare from the finalization of the Welfare Quality® project cannot be taken for granted. Whether the project will eventually have a positive impact on animal welfare will depend on a number of uncertainties, as who will

pay for animal welfare and who will gain, what role the WTO will play, whether a dominant leader will influence the course of events, etcetera.

The subsequent set of learning scenarios shows however that not all these events are completely beyond control. The environment that is in some ways hostile or friendly to the implementation of the Welfare Quality® assessment system, can also be influenced. We have learned that a body is in this respect inevitable. Overall, the aim of such a body would perhaps be described best as identifying and seizing the opportunities to improve animal welfare on the basis of the existing systems and players in animal-based production chains across Europe. The body would have:

- a scientific role of updating the system with the latest scientific insights;
- an ethical role of turning the system into a measuring rule against which farms, farming systems and perhaps brands and products can be benchmarked;
- a legitimizing role both in ensuring that the system has a solid basis among stakeholders in society, both within animal interest groups and beyond, in a wider group of stakeholders concerned with sustainable development;
- a supporting role in stimulating adoption of the system among farmers and business. Here one can think of services, trainings and packages that help individual farmers, farmer organizations, or farmers supplying to retailers or other chain members to adopt the system. Because national environments may vary within Europe, such a role is perhaps fulfilled most effectively at a national level rather than an EU level.

Several vital questions remain to be answered. The most basic one is probably ‘How much would a body to implement the Welfare Quality® assessment system cost?’ This question can’t however be answered without answering the underlying questions.

- What would be the form of the organization? A simple platform in which stakeholders meet on a periodical basis would cost nearly nothing, but an office building with staff would be much more costly (not only in terms of direct costs but also in terms of bureaucracy and administrative costs).
- What existing institutes and organizations could fulfil (some of) these roles? The Welfare Quality® assessment system will not be implemented from a clean sheet of paper. Ethical and service providing roles are already carried out by several organizations in Europe. The question how much implementation would cost would depend on the extent to which it can build on these existing structures.
- How much and from whom it could earn by providing its services? The Welfare Quality® assessment methodology may also have a monetary value, because it can provide a solid answer against shaming campaigns (as long as the legitimizing role is fulfilled well). The question remains who should pay for it and how much? Would a body charge the farmers for using the system, or the brands and retailers for providing the legitimacy, or the training and certification organizations that implement the system at the farm and in turn charge the farmer?
- Related to this question, the question remains, what the costs and benefits would be to the farmers and to other players in the chain? The answer to this question would again influence the adoption of the system.

7.4 BUILDING ON THE IMPLEMENTATION SCENARIOS

Work Package 4.4.3 has created the framework for scenario descriptions which explore potential influences on routes for implementation of Welfare Quality[®] animal based assessment systems, and Work Package 4.5.2 builds on this and studies the economic (producer, retailer, consumer) and sociological effects of potential implementation of the Welfare Quality[®] system for on farm and slaughter assessment of cattle, pigs and poultry. The combination of scenarios and a socio-economic analysis of these scenarios offers the best potential for meaningful modelling of the results of early adoption of the Welfare Quality[®] animal-based assessment scheme, and so the most effective way of anticipating and predicting possible ways in which animal based assessment could influence farm economies and consumer choice. These scenarios, created in WP4.4.3, outline the possible ways in which the farm assessment and product information resources produced in the Welfare Quality[®] project as a whole might be used by farmers, producers, certification bodies, retailers and legislators. This new WP 4.5.2 involves liaison with work partners in a number of areas of the Welfare Quality[®] project so that information required to assess the socio-economic impact of these scenarios can be assessed. The links that will have to be made will relate to:

- The information system proposed in Welfare Quality[®], which uses the animal assessment systems created in WP2.2 and refined in WP2.4, and the aggregation and interpretation system created in WP2.3 to create product information technical documents developed in WP 4.1.
- The scenarios. An outline exploration of who may use the systems, and what economic, marketing and welfare implications may their use have?

WP4.5.2 will provide valuable first insights into potential adoption patterns of the Welfare Quality[®] system through the following tasks.

1. An inventory of adoption outcome measures – i.e. to explore and list the potential measures which could be used to assess the extent of adoption of any future scheme. To enable any realistic assessment of adoption, agreed ‘adoption measure’ indices will be explored, as different measures of adoption will require different information to be collated. Examples might include; Market share, % of consumer purchasing involving Welfare Quality[®] assessed product, % of total product available involving Welfare Quality[®] assessed product, the assessed total product value, the number of farms using the Welfare Quality[®] scheme and the number of animals assessed using Welfare Quality[®] methods
2. Adoption scenario literature research of previous ‘implementation’ of assessment schemes
3. A listing of institutional or other variables, which can influence the development of the adoption pattern. By identifying the institutional variables, the influence of institutional forces on the possible development of the Welfare Quality[®] scheme can

be better explored using the model developed. The role of regulators, retailer groups, international farmer and producer groups and NGO's in shaping the potential development of the scheme will be explored.

4. A listing of variables that potentially can be influenced by the adoption patterns of the Welfare Quality® scheme. For example, farmer incomes, distribution of costs and benefits, competitive position, and social impact variables.
5. After discussion and exploration of potential influences and models in tasks 1 to 4, a decision on the model (or models) to be used to assess the adoption patterns derived from the scenarios will be made.
6. Construction of a model. Adaptation of the model will focus on key institutional variables derived from the narrative scenarios. The scope of factors will be limited, but it is considered possible to build on the findings of task 3 to select the most insightful variables.
7. A concise discussion of the potential impact of the different adoption patterns.

7.5 TESTING POSSIBLE ROUTES TO IMPLEMENTATION – CITIZEN AND FARMER JURIES

In the citizen and farmer jury activities in WP's 4.4.1 and 4.4.2, consumer and farmer acceptability of the outcomes and grading of the on farm assessment and monitoring scheme has been explored. The aim of this iterative learning process has been to identify strengths and weaknesses in the potential routes for implementation of the assessment system. Focus groups have been given presentations and information, and then asked to articulate their views on the usefulness of animal welfare information in their own fields of experience, and of their views on the assessment and scoring process. The presentations are derived from the work carried out in SP2 (and presented by researchers from SP2) and the information came from a range of other stakeholders (NGO's, certification body personnel etc).

7.6 CREATING PRACTICAL ASSESSMENT PROTOCOLS

With the assistance of NEN, the Dutch Normalisation Institute, the technical descriptions, i.e. the Welfare Quality® protocols for pigs, cattle and poultry which enable the assessment of animals on the farm and at the slaughterhouse, have been created. This process has used researchers from SP4, SP3 and SP1 who have taken the formative assessment protocols derived from the work in SP2 and shortened, refined and illustrated these to produce

working protocol documents. In these assessment protocols we have brought together many of the different work areas in the Welfare Quality® project. The protocols contain descriptions of the methods to be used along with reference photographs, technical specification for any specific methods used, and an explanation of the process used to combine the information collected on the farm and at the slaughterhouse into results and scores. The 'integrated' nature of this process is also described. The strength of the Welfare Quality® integrated animal welfare assessment system lies in using the measures together and in an agreed way. Critically, the measures must be applied following training - to ensure a harmonised, reliable, repeatable and robust outcome. The importance of agreed training for any person or organisation which may adopt the Welfare Quality® approach is discussed in the protocol documents.

Farmers in many countries are now familiar with assessment to a standard – for example the EUREPGAP standards. One of the benefits of inspection and assurance schemes are that farmers can see that they are being asked to operate to the same standards as their neighbouring farms – and also to farms in their own country and other countries. Assurance schemes help purchasers, retailers and the public to have confidence in the products they buy and use. Most existing schemes do not include direct animal-based measures, but use indirect markers for how animals are kept such as inspection of housing, feed and water provision, and disease control and medicine use. Sub Project 4 has pulled together the earlier work in the Welfare Quality® project to develop material and investigate the future implement of the welfare monitoring and information systems and the welfare improvement strategies.

8

AN OVERVIEW OF THE APPLICATION OF THE ANIMAL WELFARE ASSESSMENT SYSTEM IN LATIN AMERICA

S.M. Huertas, M. Paranhos da Costa, X. Manteca, F. Galindo and M.S. Morales

8.1 INTRODUCTION

Within Welfare Quality® leading groups of specialists were integrated to build on a network of research strengths recognizing important societal and policy objectives in relation to the area of food quality and safety. This will allow the development of pan-European standards for on-farm welfare assessment and product information systems as well as practical strategies for improving animal welfare.

Considerable effort has also been put into analyzing and addressing the perceptions and concerns of principal stakeholders (public, NGOs, industry, government and academia) and providing appropriate feedback. Educational and media initiatives, web-based platforms etc. were developed enhancing societal involvement.

Almost 50% of food from animal origin consumed in the European Union comes from Latin American countries presenting a wide variety of scenarios with around 400 million bovines, 30 millions pigs, 9.3 million tons of chicken and a wide range of production systems. For example, in relation to bovine production, most of them are extensive or semi-extensive, totally or partially based on pastures.

In recent years, experts of several Latin American countries have been working together on different areas regarding animal welfare, trying to develop practical strategies to improve farm animal welfare, as well as a sustainable rural economy improving animal husbandry and food safety levels, and integrating environmental aspects of animal production, in order to decrease poverty levels and increase health and animal welfare.

The management of this project was closely linked to the management structure of the Welfare Quality® project. Activities in Welfare Quality® were defined and clustered in the nine Sub Projects and the extension project is defined as Sub Project 10 (SP10).

The general aim of the Welfare Quality® project “Integration of animal welfare in the food quality chain: from public concern to improved welfare and transparent quality – EXTENSION” (that became SP10 in Welfare Quality®) was thought also as a way to integrate in European and Latin American groups working on those issues.

The Latin American participants in the Welfare Quality® Extension are: São Paulo State, University, UNESP (Brazil); Universidad de Chile, UCH (Chile); Universidad Nacional Autónoma de México, UNAM (México) and Universidad de la República, UDELAR-FV (Uruguay). The coordination and link is performed by the Universidad Autónoma de Barcelona, Spain.

This integration between Europe and Latin America is essential in order to respond to the globalization of the issue and to develop widely applicable welfare assessment systems and practical solutions for problems affecting animal welfare. Progress in this field in a given geographical area is affected by and should be matched to progress in other areas of the world. The cooperation between animal welfare scientists working within Welfare Quality® in Europe and relevant groups in Latin America reinforce the achievement of results that benefit human and animal quality of life.

The proposal had four objectives: 1) to study consumers’ attitudes and beliefs towards animal welfare in Brazil, Chile, Mexico and Uruguay; 2) to test and implement the animal welfare monitoring system (for pigs, cattle and poultry) developed in Sub Project 2 (WP 2.4) in the conditions encountered in Latin America; 3) to develop practical strategies to improve the welfare of farm animals (pigs, cattle and poultry), with particular emphasis on animal handling during loading and transport; and 4) to increase existing knowledge of some of the major welfare problems of extensive systems of animal production and its integration to environmental sustainability issues. It was expected that the above objectives complement and contribute to the achievement of the objectives of Welfare Quality® and their global uptake in several ways. Working together, experienced scientists from both continents in a variety of climatic, economic and social conditions have provided robustness and scientific quality of the strategies and assessment tools developed in Welfare Quality®.

This chapter deals with the second objective of the project, namely to test and implement the animal welfare monitoring system (for pigs, cattle and poultry) developed in Sub Project 2 of Welfare Quality (WP 2.4) in the conditions encountered in Latin America. The experience gained by Latin American scientists when applying the animal welfare monitoring system in their own countries has revealed a number of areas that deserve further attention.

8.2 ACTIVITIES DEVELOPED IN LA COUNTRIES WITHIN SP10

The activities of SP10 covered the following aspects: workshops in four Latin American countries, training of observers in the protocols methodology, application of animal welfare assessment protocols at farm and slaughterhouses, evaluation of practical problems observed in the application of the protocols and implementation of solutions and evaluation of other issues related to Welfare Quality®.

8.3 WORKSHOPS ON ASSESSMENT OF WELFARE QUALITY® PROTOCOLS

In order to adapt the animal welfare monitoring system developed in Europe to the conditions commonly encountered in Latin America, three workshops on cattle, pigs and poultry were held in Uruguay, Brazil And Chile. The workshops had the assistance of European and Latin-American experts and facilitators in order to transfer knowledge and experiences in the application of the animal welfare assessment protocols developed in the European countries. Travel expenses of European experts were covered by the European Commission through DG SANCO.

8.3.1 Workshop on Welfare Assessment in Cattle (Beef and Dairy)

The first workshop on Welfare Assessment in Cattle (beef and dairy) took place in Uruguay, August 2007. The main point discussed during the activities was the use of the assessment protocol taking into consideration the extensive and semi-extensive conditions in Latin American systems and the need to translate the protocol into Portuguese and Spanish.

8.3.2 Workshop on Welfare Assessment in Pigs

The workshop on Pigs' Welfare Assessment was carried out at Concordia, Santa Catarina (Brazil) in November 2007. The main objective was to introduce the protocols for pig welfare assessment, developed by the experts linked to the Welfare Quality® Project, evaluating their applications in the Latin American environment.

The activities were carried out in two private farms and in one private slaughterhouse. In general not many differences were found in the application of the protocols and the workshop was important to clarify some aspects in the protocols application.

8.3.3 Workshop on Welfare Assessment in Poultry

The Welfare Assessment in Poultry training was held in Santiago, Chile, January, 2008 and the main purpose was to train Latin American members of the project on the Welfare Quality® poultry protocol and to discuss the applicability of the protocols to the Latin American conditions.

Regarding the applicability of the protocols, the general feeling was that the monitoring system is sufficiently applicable to the Latin American conditions. The main concern was related to how objectively the Qualitative Behavioural Assessment could be measured, considering there were some words that in Spanish are used in a different way (e.g. inquisitive) and in smaller degree the participants found the interview part of the protocol too long.

8.4 TRAINING OBSERVERS

Following the workshops, more observers were trained in each participating country in order to start the visits to farms and slaughterhouses and evaluate and validate the protocols. In all cases, the practical trainings were carried out at the Universities using the information, material and support provided by Welfare Quality® instructors during the previous workshops.

A total of 48 persons were trained on protocols application. Sixteen in Brazil, 11 in Chile, 11 in Mexico and 10 in Uruguay acquired the necessary skills to apply the protocols at farms and slaughterhouses.

8.5 APPLICATION OF ANIMAL WELFARE ASSESSMENT PROTOCOLS AT FARMS AND SLAUGHTERHOUSES

8.5.1 Poultry

A total of 33 visits to evaluate poultry protocols were performed in the four countries involved, including nine intensive broiler farms; 17 laying hens farms; two free range facilities and five slaughterhouses.

Poultry production is very similar in most of the big exporting countries in Latin America (Brazil and Chile). However, most of the observers found that the application of the protocols took too much time inside the farm. Moreover, strong biosecurity measures applied in almost all the facilities adding some limitations to perform the visit.

8.5.2 Pigs

A total of seven extensive and 13 intensive farms were visited and protocols were successfully applied on almost 30,000 animals in all the countries involved in the project. Further, six slaughterhouses for pigs were visited in Brazil.

8.5.3 Beef

Due to the different characteristics of livestock between Europe and America, a total of 26 extensive and semi-extensive beef cattle farms were visited in three countries (Uruguay, Chile and Brazil) and 4,500 animals were tested. The average size of the farms varied from 250 hectares in Chile to 2,700 in Brazil, and the average number of cattle per farm ranged from 752 in Chile to 3,600 in Uruguay.

Some important problems on the applicability of the protocols were found, arising mainly from farm and herd size, extensive pastoral systems, cattle identification and, occasionally, climatic conditions. Alternative approaches to overcome these difficulties were developed during the welfare assessment protocols' application (described below).

8.5.4 Dairy

A total of 55 dairy farms of different systems were visited in the 4 countries involved in the study, and some 6,500 animals were registered. In some countries there is semi-extensive dairy production thus some alternative methodologies were developed during the application of the welfare assessment protocols (described below).

8.5.5 Slaughter

A total of 12 cattle slaughterhouses, 6 slaughterhouses for pigs were visited in BR, MX and URU and only in Brazil four broiler slaughterhouses were visited.

8.6 DIFFICULTIES WHEN APPLYING THE PROTOCOL IN LATIN AMERICA

The difficulties identified from the experience in Latin America fall into three categories:

- a. Areas of concern equally applicable to European and Latin American conditions.
- b. Methodological and logistical problems encountered when applying the animal welfare monitoring protocols in Latin-American conditions.
- c. Points that are missing or are not applicable when using the animal welfare monitoring system in Latin-American extensive farms.

8.6.1 Areas of Concern Equally Applicable to European and Latin American Conditions

- Qualitative assessment is seen as an innovative and useful methodology. One problem encountered by observers, however, is that some of the original English terms are difficult to translate into Spanish or Portuguese. We believe that this difficulty may also apply to non-English speaking European countries. Although it is recognized that a precise definition of each term may not be essential, some observers may feel that they are not qualified to do the assessment if many of the terms do not have a clear meaning to them.
- Defining the appropriate sample size (i.e. number of animals assessed per farm) appears to be one of the most difficult areas, taking into consideration the size of many farms in LA. The problem is further complicated by the fact that a given sample size will account for a larger percentage of animals on a small farm compared with a large farm. This will mean that the reliability of the data would be higher on small farms than on bigger farms.
- Related to the previous problem, defining which animals have to be assessed on a given farm is also a difficult issue.
- The sensitivity of heat stress parameters was questioned. One important welfare problem in cattle (particularly dairy) is heat stress and yet, the welfare monitoring system does not include any parameter related to this problem. Although this will be particularly important in hot climates, it is worth emphasizing that Holstein dairy cows may suffer heat stress when the temperature rises above 25-30 °C, which is not at all uncommon in summer even in temperate countries. As for fattening pigs, wallowing behaviour was not taken into account and the reliability of the parameter manure on the body was questioned.
- Breed or age differences may have an important effect on some parameters and on how to interpret them. For example, flight distance in broilers may be short not because of lack of fear but because of lameness; as the prevalence of lameness varies between grillers and roosters, interpretation of results might be confounded.

8.6.2 *Methodological and Logistical Difficulties Encountered When Applying the Animal Welfare Monitoring System in Latin America Extensive Farms*

- On large extensive farms or pastoral systems, doing a herd scan may not be possible. Some observers also commented on the fact that, even when feasible, this parameter may not yield very useful information. For example, a herd scan could be done when animals are regrouped into the handling facilities; however changes in their behaviours (e.g. increased fear) will make the behavioural measures not representative of the normal state of these animals.
- Most of the beef and dairy farms in Latin America are extensive or semi-intensive systems, where feeding troughs are used at all or only rarely. Thus avoidance distance at the feed trough is impossible or very difficult to assess.
- Another problem observed was that it was impossible to assess 70% of the cows in lactation in some farms, since the number of lactating cows is very large, in some case more than 1000.
- Also, individually identifying or even observing the animals may be difficult in some farms. For example, pens might be too large to see the animals, even with binoculars, or the animals may be out of sight if the vegetation is very high. Further, not all farms have adequate individual identification systems. The same difficulty applies to observe the same side of all individuals as required by the cattle/dairy cow protocol.
- Seeing enteric problems and hearing respiratory conditions (coughing or sneezing) is very difficult when animals can not be easily approached or lie on the mud.
- In some cases when the animals are free in the pastures, observation would be impeded by some grass species, ground's slope or presence of dangerous animals (like snakes).
- In relation to cleanness of animals in extensive system, it is unusual to see dirty animals, except for some dirt on the hocks due only to occasional bad weather conditions.

8.6.3 *Methodological and Logistical Difficulties Encountered When Applying the Animal Welfare Monitoring System in Latin American Slaughterhouses*

Differences in the size and speed of the slaughterhouse may lead to problems when trying to observe some of the parameters. For example, in some countries slaughterhouses kill more than a thousand animals per day, slaughtering an average of 100–150 animals per hour or more, the speed is so high that it is difficult to apply the protocol.

An additional problem is that in some slaughterhouses it is not possible, or at least it is very difficult, to assess the effectiveness of the stunning process in 200 animals as required by the protocol. Also, as the average number of bulls slaughtered per day may be very low, assessing the minimum number of bulls as required by the protocol may be difficult. Finally, when animals are driven from the lairage area to the stunning box in large groups, identifying individual animals becomes very difficult.

Related to unloading of animals from the trucks, the consideration of a truck as a unit is not feasible; the consideration of animals individually is more practical.

Moreover, Latin American experts found that it is extremely difficult to record each animal from the corral to the stunning box.

Some signs of inadequate stunning were not always possible to measure, such as pupil dilatation, eye rotation, corneal reflex, response to pain stimulus, only the most obvious signs can be recorded, such as extensive kicking, righting reflex or attempts to raise the head, because the facilities do not allow getting close to the animals.

Measures of pH of the same animals are very difficult to perform and imply visiting the slaughterhouse one more day. Nevertheless in exporting industries this is a routine procedure and the protocols can be used.

8.6.4 Points that Are Missing or Are Not Applicable When Using the Animal Welfare Monitoring System in Latin America Extensive Farms

Some of the main aspects that have a clear impact on welfare and are not included in the current protocol are:

- risk of predation;
- horn flies and ticks;
- sunburn;
- pasture size and condition;
- mineral supplements;
- wallowing facilities (only pigs);
- water supply from natural sources;
- shade;
- mounting;
- branding (in cattle), it is compulsory in some countries and very common in others;
- distance walked (from pasture/paddock to milking parlour) and quality of the walkway (in dairy cows).

8.7 SUGGESTED REMEDIAL STRATEGIES AND RECOMMENDATIONS

Some of the areas of concern that are equally applicable to European and Latin American conditions require careful discussion among scientists involved in SP2 and SP10. In this regard, it is worth mentioning that several of the participants attended the final SP2 meeting

and had the opportunity to discuss many of the issues raised in this report with the scientists in SP2.

On the other hand, we strongly believe that many of the problems that are encountered when applying the monitoring system to extensive farms could be solved if there is some flexibility in the methodology and a readiness to use alternatives when a particular measure is not applicable, as well as to incorporate new, relevant measures (although we do recognize that this may pose some problems in the integration phase). For example, at the slaughterhouse it would be possible to record information on the group instead of the individual animals when animals are driven in group to the stunning box and start recording individual information just after the entry into the stunning box. It could be useful to divide the driving area into several sectors (defined specifically for each plant); for large slaughterhouses we suggest at least 50 animals per sector (e.g. pen, corridor, shower, and line chute).

It is important to mention, however, that this flexibility should not fundamentally change the monitoring system. Therefore, it is very important that observers are well trained not only in the system itself but in the characteristics of the production system so that they can understand and achieve the objectives of the assessment.

8.7.1 A Suggested New Indicator: Human–Animal Interactions during Handling

The human–cattle interactions are intense during handling in the pens. Therefore, we understand that the quality of handling (and consequently of the interaction) should be assessed. We suggest that after avoidance distance and clinical scoring assessments the animals should be driven by the farm's team, simulating a routine procedure moving them from one pen to another, walk through the corral pens, squeeze chute, line chute, restrainer or scale; without carrying out any other procedure. The assessment has to be done on all the animals of the group or with at least 100 animals. We suggest the following indicators:

- use of resources to drive the animals (flags, sticks, electric prods, others);
- aggression towards the animals (number of occurrences);
- speed to drive the animals: slow (walk), fast (trot), very fast (run);
- human vocalization (nothing, talk and whistle, shouts)
- number of occurrences of falling, slipping, jumping, animals that jump on other animals (trampling), lying down, hits by gates and animal vocalizations.

Since injuries caused by a bad practice in the vaccine application are common in Latin American herds, we suggest considering the injuries due to vaccination such as lesions/swellings.

In most of the countries iron branding is compulsory by law, we suggest to include the iron branding scars as a general comment about the herd or define if should be considered as a 'hairless patches' or a lesion / swellings.

Fortunately, tail docking is not common in LA dairy farms; some finding could be caused by accidents or bad handling. Moreover, we suggest considering (when doing the clinical assessment) if the cow's tail is broken, since this observation is quite easy to carry out and it offers an idea about the quality of the handling.

In dairy farms, we also suggest taking into consideration the time spent by the cows in the waiting room and in the milking parlour; and the resources available in the waiting room (water, bed, food, shade, sprinklers, etc.).

8.8 OTHER ISSUES RELATED TO WELFARE QUALITY®

It is important to point out that the INCO-WQ Project in Latin America contributed to promoting a large variety of events in animal welfare in the region that we consider as other outcomes, besides the assessment of the protocols. Example include: developing extension material (brochures, papers, books, web pages, etc); participation in seminars and congresses; promotion of animal welfare laws; diverse research projects linked with Welfare Quality® in different countries; degree, master and doctor's theses on animal welfare topics; improving the engagement on animal welfare issues with different stakeholders in each country and growing of international recognition of Latin America's work in animal welfare.

8.8.1 Dissemination Material

Material such as brochures and books, development of web pages in Spanish and Portuguese and presentations in seminars and congresses were done during this period (see references).

8.8.2 Animal Welfare at the University Level

- In the four countries considerable effort has been made to include animal welfare in the curricula at the Veterinary and Animal Science Schools and Faculties.
- More than 20 monographies and theses (Masters and PhD) have been produced in all LA groups engaged in the INCO-Welfare Quality® project.
- Different programs through Permanent Education in the University or similar have been carried out in all the countries to capacitate truck drivers, stockmen, producers and stakeholders in Good Management Practices. Only in Uruguay more than 3,000 people have qualified.

8.8.3 Animal Welfare at Legislative Level

- In Uruguay an Animal Welfare Law was approved in 2008 at the legislative level, and it is presently passing the regulation process before being implemented.
- Members of the Animal Welfare Group have coordinated and drafted the General Animal Welfare Bill in Mexico. This initiative has been presented to the Mexican Senate and will be voted during the end of 2009. Furthermore, this group has participated in the updating of 4 regulations on different animal welfare issues.

8.8.4 Animal Welfare and the Commitment of Stakeholders

- In most of the countries, an increase in the awareness on animal welfare at different levels was observed, both within the country and between participating countries.
- Authorities of the Ministries of Agriculture or Livestock in all countries were directly involved in different animal welfare programs.
- Finally, both producers and industry collaborated in spreading extension programs of good management practices.

8.8.5 Collaboration between Countries and Other Institutions

- Interchange of different Universities of the involved countries (e.g. students, professors, etc).
- Creation of the Latin American Coalition on Animal Welfare supported by the World Society for the Protection of Animals (WSPA).
- Inter American Committee on Animal Welfare of the Americas under the support of the International Organization of Animal Health (OIE).
- Creation of the first OIE Collaborating Centre on Animal Welfare for the Americas, an enterprise between Uruguay and Chile, with the cooperation of both Ministries of Agriculture or Livestock and Universities.
- A proposal for the creation of an OIE Collaborating Centre on Animal Welfare at the Faculty of Veterinary Medicine of the National Autonomous University of Mexico (UNAM) has been presented at the OIE General Assembly in Paris during May 2009.

8.9 ACKNOWLEDGMENTS

Authors want to thank the enthusiastic group of colleagues that worked in each country on the INCO-WQ project related activities, even with difficulties but always with a high level of commitment.

BIBLIOGRAPHY

- Baxter, E.M., Jarvis, S., D'Eath, R.B., Ross, D.W., Robson, S.K., Farish, M., Nevison, I.M., Lawrence, A.B. and Edwards, S.A. (2008) 'Investigating the behavioural and physiological indicators of neonatal survival in pigs', *Theriogenology*, 69, pp. 773–783.
- Bekoff, M. (2008) 'Why “good welfare” isn't “good enough”: minding animals and increasing our compassionate footprint', *ARBS Annual Review of Biomedical Sciences*, 10, pp. T1–T14.
- Boivin, X., Lensink, B.J., Tallet, C. and Veissier, I. (2003) 'Stockmanship and farm animal welfare', *Animal Welfare*, 12, pp. 479–492.
- Botreau, R., Veissier, I., Butterworth, A., Bracke, M.B.M. and Keeling, L.J. (2007) 'Definition of criteria for overall assessment of animal welfare', *Animal Welfare*, 16, pp. 225–228.
- Bouyssou, D. (1990) 'Building criteria: a prerequisite for MCDA', in: C.A. Bana and E. Costa (eds) *Readings in Multiple Criteria Decision-Aid*, Springer Verlag, Heidelberg, pp. 58–80.
- Bouyssou, D., Marchant, T., Pirlot, M., Perny, P., Tsoukias, A. and Vincke, P. (2001) *Evaluation and Decision Models a Critical Perspective*. Kluwer Academic Publishers.
- Dawkins, M.S., Donnelly, C.A. and Jones, T.A. (2004) 'Chicken welfare is influenced more by housing conditions than by stocking density', *Nature*, 427, pp. 342–344.
- Duncan, I.J.H. (1990) 'Reactions of poultry to human beings', in: R. Zayan and R. Dantzer (eds) *Social Stress in Domestic Animals*, Kluwer, Dordrecht, pp. 121–131.
- Edwards, S.A. and Fraser, D. (1997) 'Housing systems for farrowing and lactation', *Pig Journal*, 39, pp. 77–89.
- Edwards, S.A. (2002) 'Perinatal mortality in the pig: environmental or physiological solutions?', *Livestock Production Science*, 78, pp. 3–12.
- European Commission (2000) *The Welfare of Chickens Kept for Meat Production (Broilers)*, Report of the Scientific Committee on Animal Health and Animal Welfare, published online <<http://www.cordis.lu>>.
- European Commission (2002) *Animal Welfare Legislation on Farmed Animals in Third Countries and the Implications for the EU*, COM(2002) 626. Office for Official Publications of the European Communities, Luxemburg.
- European Commission (2005) *Attitudes of Consumers towards the Welfare of Farmed Animals*, Special Eurobarometer 229/Wave 63.2, published online <http://europa.eu.int/comm/public_opinion/archives/ebs/ebs_229_en.pdf>.
- European Commission (2007a) *Attitudes of Consumers towards the Welfare of Farmed Animals – Wave 2*, Special Eurobarometer 229(2)/Wave 64.4, published online <http://ec.europa.eu/food/animal/welfare/survey/sp_barometer_fa_en.pdf>.
- European Commission (2007b) *Attitudes of EU Citizens towards Animal Welfare*, Special Eurobarometer 270/Wave 66.1, published online <<http://ec.europa.eu/food/animal>>.

[/welfare/survey/sp_barometer_fa_en.pdf](#)>.

- Evans, A. and Miele, M. (eds) (2007) *Consumers' Views about Farm Animal Welfare*. Cardiff University.
- Evans, A., and Miele, M. (2007) *Consumers' Views about Farm Animal Welfare. Part I: National Reports based on Focus Group Research*, Welfare Quality Reports No.4. Cardiff: Cardiff University.
- Evans, A., and Miele, M. (2008) *Consumers' Views about Farm Animal Welfare: European Comparative Report Based on Focus Group Research*, Welfare Quality Reports No.5. Cardiff: Cardiff University Press.
- FAWC (1992) *Report on the Welfare of Broiler Chicken*. Farm Animal Welfare Council, Surbiton.
- FAWC (1998) *Report on the Welfare of Broiler Breeders*. Farm Animal Welfare Council, Surbiton.
- Fraser, D. (1995) 'Science, values and animal welfare: exploring the "inextricable connection"', *Animal Welfare*, 4, pp. 103–117.
- Grabisch, M. and Roubens, M. (2000) 'Application of the Choquet integral in multicriteria decision making', in M. Grabisch, T. Murofushi and M. Sugeno (eds) *Fuzzy Measures and Integrals: Theory and Applications*, Physica-Verlag, Heidelberg, pp. 348–374.
- Hemsworth, P.H. and Coleman, G.J. (1998) *Human-Livestock Interactions: The Stockperson and the Productivity of Intensively Farmed Animals*. CAB International, Wallingford.
- Hemsworth, P.H. (2003) 'Human-animal interactions in livestock production', *Applied Animal Behaviour Science*, 81, pp. 185–198.
- Jones, R.B. (1997) 'Fear and distress', in: M.C. Appleby and B.O. Hughes (eds) *Animal Welfare*, CAB International, Wallingford, pp. 75–87.
- Jones, R.B. (2004) 'Environmental enrichment: the need for practical strategies to improve poultry welfare', in: G.C. Perry (ed.) *Welfare of the Laying Hen*, CAB International, Wallingford, pp. 215–225.
- Kestin, S.C., Knowles, T.G., Tinch, A.E. and Gregory, N.G. (1992) 'Prevalence of leg weakness in broiler chickens and its relationship with genotype', *Veterinary Record*, 131, pp. 190–194.
- Kestin, S.C., Su, G. and Sorensen, P. (1999) 'Different commercial broiler crosses have different susceptibilities to leg weakness', *Poultry Science*, 78, pp. 1085–1090.
- Kjaernes, U. and Keeling, L.J. (2006) *Principles and Criteria of Good Farm Animal Welfare*, Welfare Quality® Fact Sheet, Wageningen University, Lelystad.
- Kjærnes, U. and Lavik, R. (2008) 'Opinions on animal welfare and food consumption in seven European countries', in: U. Kjærnes, B. B. Bock, E. Roe and J. Roex (eds) *Consumption, Distribution and Production of Farm Animal Welfare*, Welfare Quality Reports No.7, Cardiff University, Cardiff.
- Kjærnes, U., Harvey, M. and Warde, A. (2007a) *Trust in Food: A Comparative and Institutional Analysis*. Palgrave Macmillan, London.
- Kjærnes, U., Roe, E. and Bock, B. (2007b) 'Societal concerns on farm animal welfare', Paper read at 'Assuring Animal Welfare: From Societal Concerns to Implementation', Second Welfare Quality® stakeholder conference, Berlin, 3–4 May.

- Kjørstad, I. (2005) 'Consumer concerns for food animal welfare', Part IA: Literature reviews', in: J. Roex and M. Miele (eds) *Farm Animal Welfare Concerns: Consumers, Retailers and Producers*, Welfare Quality Reports No.1, Cardiff University, Cardiff.
- Knol, E.F., Leenhouwers, J.I. and Lende, T. van der (2002) 'Genetic aspects of piglet survival', *Livestock Production Science*, 78, pp. 47–55.
- Meat and Livestock Commission (2006) *The Pig Yearbook*. Meat and Livestock Commission, London.
- Miele, M. and Evans, A.B. (forthcoming) 'When food become animals: Ruminations on ethics and responsibility in care-full practices of consumption', *Ethics, Place and Environment*.
- Miele, M. and Bock, B. (2007) 'Competing discourses of farm animal welfare and agri-food restructuring', *International Journal of Sociology of Agriculture and Food*, 15(3), pp. 1–7.
- Pickup, H.E., Cassidy, A.M., Danbury, T.C., Weeks, C.A., Waterman, A.E. and Kestin, S.C. (1997) 'Self selection of an analgesic by broiler chickens', *British Poultry Science*, 38, pp. S12–S13.
- Polten, R. (ed.) (2007) *Proceedings of the Workshop 'Animal Welfare Improving by Labelling?'*, Brussels, 28 March.
- Raussi, S. (2003) 'Human-cattle interactions in group housing', *Applied Animal Behaviour Science*, 80, pp. 245–262.
- Roe, E., and Marsden, T. (2007) 'Analysis of the retail survey of products that carry welfare-claims and of non-retailer led assurance schemes whose logos accompany welfare-claims', in: U. Kjærnes, M. Miele and J. Roex (eds) *Attitudes of Consumers, Retailers and Producers to Farm Animal Welfare*, Welfare Quality Report No.2, Cardiff University, Cardiff.
- Roex, J., and Miele, M. (2005) *Farm Animal Welfare Concerns: Consumers, Retailers and Producers*, Welfare Quality Reports No.1. Cardiff University, Cardiff.
- Roy, B. (1993) 'Decision science or decision-aid science?', *European Journal of Operation Research*, 66, pp. 184–204.
- Sanotra, G.S., Berg, C. and Lund, J.D. (2003) 'A comparison between leg problems in Danish and Swedish broiler production', *Animal Welfare*, 12, pp. 677–683.
- Tuchscherer, M., Puppe, B., Tuchscherer, A. and Tiemann, U. (2000) 'Early identification of neonates at risk: traits of newborn piglets with respect to survival', *Theriogenology*, 54, pp. 371–388.
- Waiblinger, S. and Spoolder, H. (2007) 'Quality of stockpersonship', in: A. Velarde and R. Geers (eds) *On Farm Monitoring of Pig Welfare*, Wageningen Academic Publishers, Wageningen, pp. 159–166.
- Waiblinger, S., Boivin, X., Pedersen, V., Tosi, M., Janczak, A.M., Visser, E.K., Jones, R.B. (2006) 'Assessing the human–animal relationship in farmed species: a critical review', *Applied Animal Behaviour Science*, 101, pp. 185–242.
- Weeks, C.A., Danbury, T.D., Davies, H.C., Hunt, P. and Kestin, S.C. (2000) 'The behaviour of broiler chickens and its modification by lameness', *Applied Animal Behaviour Science*, 67, pp. 111–125.

Whay, H.R., Main, D.C.J., Green, L.E. and Webster, A.J.F. (2003) 'An animal-based welfare assessment of group-housed calves on UK dairy farms', *Animal Welfare*, 12, pp. 611–617.

APPENDIX

Andy Butterworth

The links between the different work packages in SP4, and their degree of interdependency and also their dependency on work carried out in previous work package – and particularly the links developed with previous work packages in SP1, SP2 and SP3 to work toward practical implementation of the Welfare Quality® assessment scheme.

TABLE A1 The links between the different work packages in SP4 and to previous work in SP1, SP2 and SP3.

WP	Short description	High dependency links	Moderate dependency links
4.1.2	Technical document construction	SP2, 2.3, 2.4, 2.5	WP461, 462, 463, 425, 443
4.2.1	Software system for data collection on farms, web based database to store data & calculate welfare scores from measures & interactive platform to enable interactions on measures and scores with stakeholders	SP2, 2.2, 2.4, 2.5	WP461, 462, 463
4.2.5	Collection of reference and training material	SP2 2.2, 2.4	WP461, 462, 463
4.2.6	Train users in WP2.4	SP 2 2.2, 2.4, WP4.2.5	WP461, 462, 463
4.3.1	To collate existing and new knowledge on improvement strategies suitable for solving welfare concerns, and to produce a resource of information that will enable advisors and farmers to identify appropriate improvement strategies for welfare concerns.	SP3, SP2 2.2, 2.4	WP461, 462, 463
4.3.2	Active web based resource linked to the WelfareQuality web tool promoting the uptake and dissemination of the information derived from the strategies developed in SP3 and from other sources.	SP3, SP2 2.2,2.4	WP461, 462, 463
4.4.1.1	To assess consumers acceptability of the outcomes and grading of the on farm assessment and monitoring scheme through different farming systems classified as 'excellent', 'good' and 'basic' in three study countries.	SP1, SP2 2.2, 2.4	WP443
4.4.1.2	To examine the mechanisms by which specific evaluative criteria and requirements are introduced and negotiated within food certification and assurance procedures.	SP1, SP2 2.2, 2.4	WP4411, 442
4.4.2	To assess farmers acceptability of the outcomes and grading of the on farm assessment and monitoring scheme through different farming systems in three study countries.	SP1, SP2 2.2, 2.4	WP443
4.4.3	Scenario creation and consultation with stakeholders.	SP1, SP2 2.2, 2.3, 2.4, 2.5	WP4412, 461, 462, 463

TABLE A1 CONT. The links between the different work packages in SP4 and their relations to other Sub-Projects.

WP	Short description	High dependency links	Moderate dependency links
4.5.1	To assess 'on-farm' economic implications and explore marketing perspectives of improved animal welfare strategies	SP1, SP2 2.2, 2.3	
4.5.2	To quantify the likely socio-economic impact of the implementation scenarios created in 4.4.3	WP443	WP4412
4.6.1	To determine the optimum mechanism for implementing the welfare assessment system on pig farms.	SP1 SP2 2.2, 2.3, 2.4, 2.5	
4.6.2	To determine the optimum mechanism for implementing the welfare assessment system on cattle farms.	SP1 SP2 2.2, 2.3, 2.4, 2.5	
4.6.3	To determine the optimum mechanism for implementing the welfare assessment system on poultry farms.	SP1 SP2 2.2, 2.3, 2.4, 2.5	
4.7	To formulate recommendations for policy makers on the potential implementation of the animal welfare assessment system, the product information system and the improvement strategies.	All areas of project.	

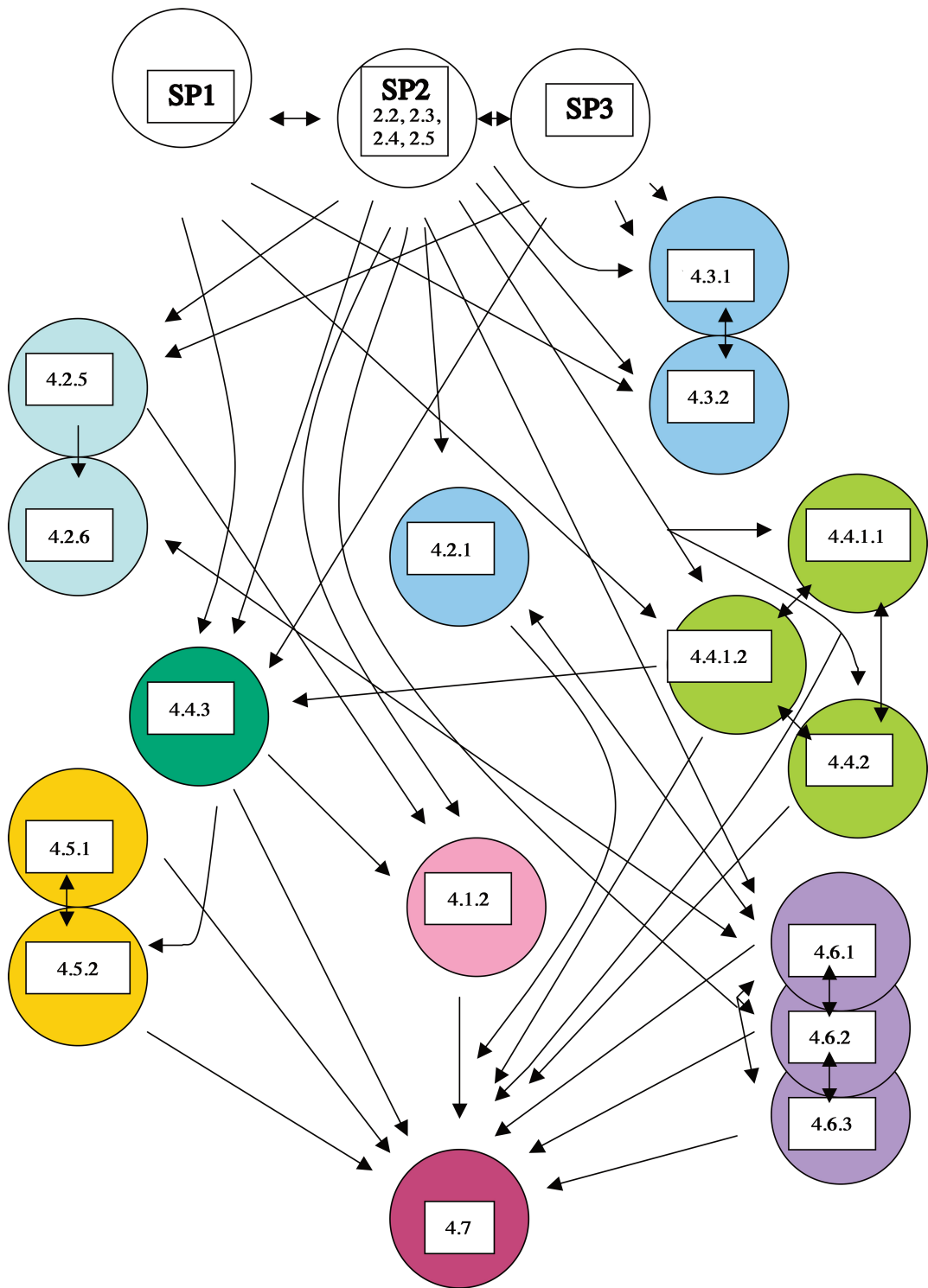


FIGURE A1 Schematic links between the implementation work packages in SP4 and their links to previous work in SP1, SP2 and SP3.



This report is an official deliverable of the Welfare Quality® project.

Welfare Quality® is a European research project focusing on the integration of animal welfare in the food quality chain: from public concern to improved welfare and transparent quality. Welfare Quality® is co-financed by the European Commission, within the 6th Framework Programme, contract No. FOOD-CT-2004-506508.

Project Office Welfare Quality
Animal Sciences Group of Wageningen UR
Edelhertweg 15
P.O Box 65
NL - 8200 AB Lelystad
The Netherlands
Phone +31 320 293503
Fax +31 320 238050
e-mail: info@welfarequality.net
Internet: www.welfarequality.net

The text of this report represents the authors' views and does not necessarily represent a position of the European Commission who will not be liable for the use made of such information.

