The Old School, Brewhouse Hill, Wheathampstead,
Hertfordshire AL4 8AN, UK

Animal Welfare 2009, 18: 363-370
ISSN 0962-7286

Overall assessment of animal welfare: strategy adopted in Welfare Quality®

R Botreau*†, I Veissier† and P Perny‡

† INRA, UR1213 Herbivores, F-63122 Saint-Genès-Champanelle, France
‡ LIP6, University of Paris 6, 104 Avenue du Président Kennedy, F-75016, Paris, France
* Contact for correspondence and requests for reprints: raphaelle.botreau@clermont.inra.fr

Abstract

Animal welfare is multidimensional; its assessment relies on complementary measures covering all dimensions. Welfare Quality® constructed a multicriteria evaluation model for its assessment at unit level (farms, slaughterhouses). Four welfare principles are distinguished ('Good feeding', 'Good housing', 'Good health', and 'Appropriate behaviour'). An animal unit receives four principle scores (expressed on a 0–100 value scale). These scores are aggregated together to form the overall assessment by sorting animal units into predefined welfare categories boundaried by reference profiles. A unit is assigned to the welfare category above the profile it is considered at least as good as. Several assignment procedures were tested on a set of 69 dairy farms and compared with observers' general impressions. The welfare categories, reference profiles and assignment procedure were defined in consultation with social scientists, animal scientists and stakeholders. Four welfare categories were defined: 'Excellent', 'Enhanced', 'Acceptable', and 'Not classified'. The reference profiles were set at 80, 55 and 20, corresponding to aspiration values for Excellent, Enhanced and Acceptable. The assignment procedure resulted from a compromise between theoretical opinion on what should be considered excellent, enhanced or acceptable, and what can realistically be achieved in practice: to be assigned to a given category, a unit must reach its aspiration value on 2 or 3 of the 4 principles, and not score below the aspiration value for the next lowest category on the other principle(s). The model can be used for several purposes, including identifying welfare problems on a farm to advise farmers, or checking compliance with certification schemes.

Keywords: *advice, animal welfare, certification, expert opinion, multicriteria evaluation, overall assessment*

Introduction

The aims of the European Welfare Quality® (2004–2009) project were to design methods for the overall assessment of cattle, pig, and poultry welfare, on-farm and at slaughter. This assessment tool is designed to help farmers and slaughterhouse managers identify welfare problems and monitor progress. It can also provide information to consumers on the welfare of the animals from which products are sold. Welfare is a multidimensional concept that requires all the component dimensions to be checked by specific indicators (eg Mason & Mendl 1993; Fraser 2003; Botreau *et al* 2007a). Therefore, the Welfare Quality® assessment incorporates numerous welfare measures focused essentially on animals and to a lesser extent on resources or animal management, whether on-farm or in slaughterhouses. For the purposes of this paper, we will use the word unit to refer to either a farm or a slaughterhouse. The Welfare Quality® assessment generates a substantial amount of data that needs be interpreted in terms of welfare and then integrated to produce an overall evaluation at unit level. For the assessment to be used routinely and consistently across Europe and between stakeholders, it is essential that this overall evaluation is based on a formal model (see Botreau *et al* 2007a). The general architecture of the evaluation model developed in Welfare Quality® was outlined in Botreau *et al* (2007b). Briefly, Welfare Quality® identified 12 key animal welfare criteria: absence of prolonged hunger; absence of prolonged thirst; comfort around resting; thermal comfort; ease of movement; absence of injuries; absence of disease; absence of pain induced by management procedures; expression of social behaviour; expression of other behaviours; good human-animal relationship and absence of general fear (Botreau *et al* 2007). Data collected on an animal unit are used to check unit compliance with the 12 welfare criteria. The scores obtained at criterion level are then collated to assess unit compliance with four main welfare principles (Good feeding, Good housing, Good health, and Appropriate behaviour). Finally, these principle scores are used to conclude on an overall evaluation (Table 1). The model was fine-tuned in response to expert views from a panel of animal scientists, social scientists, and stakeholders.

https://doi.org/10.1017/S0962728600000762 Published online by Cambridge University Press

Table 1 Sequential evaluation structure, from measures collected on-farm through the welfare criteria and principles and, finally, the overall assessment (Botreau *et al* 2007c). Criteria and principles are common to all animal types under study (pigs, cattle, and poultry). Measures are specific to animal types. The table includes the measures designed for dairy cows, on-farm (Botreau *et al* 2008).

Measurement in dairy cows	Welfare criteria (12)	Welfare principles (4)	
Body condition score	Absence of prolonged hunger	Good feeding	
Provision of water (drinker [n] and cleanliness)	Absence of prolonged thirst		
Behaviour around resting, cleanliness	Comfort around resting	Good housing	
No reliable measure available at present	Thermal comfort		
Presence of tethering, access to exercise	Ease of movement		Overall assessment
Skin alteration, lameness	Absence of injuries	Good health	
Respiratory, digestive, reproductive diseases	Absence of disease		
Dehorning, tail docking (rare)	Absence of pain induced by management procedures		
Incidence of aggressive behaviours	Expression of social behaviours	Appropriate behaviour	
Access to pasture	Expression of other behaviours		
Avoidance distance to human approach	Good human-animal relationship		
Qualitative assessment of behaviour	Positive emotional state		

The definition and construction of the criteria and principles were described in previous papers (Botreau *et al* 2007c; Botreau *et al* 2008) and will be only summarised here as background information. The present article focuses on the final stage of the evaluation model, ie producing the overall assessment from the four principle scores (expressed on a 0–100 value scale).

Background information

Initial steps in the development of the welfare assessment model: from measures to principles

The 12 welfare criteria proposed by Welfare Quality® were discussed in 55 consumer-citizen focus groups across seven European countries. They received widespread agreement, and consumer-citizens rejected the idea that some criteria may be more important than others (Veissier & Evans 2007).

Welfare Quality® partners checked the relevance, repeatability and feasibility of numerous welfare measures and finally defined 30 to 60 measures on animal units for each animal type studied (dairy cows, fattening bulls, veal calves, broilers, layers, sows and piglets, and fattening pigs; see Veissier *et al* [2007] for a description of the measures). Animal unit compliance with each welfare criterion identified in the project is checked by one or several measures. At this stage, data produced by relevant measures on each criterion are turned into welfare value scores on a common scale ranging from 0 to 100, where 0 corresponds to the worst situations, 100 to ideal situations, and 50 is for 'neither good nor bad' situations. Any animal unit scoring below 20 on any single criterion has a high risk of being excluded from any certification scheme that includes welfare concerns. The 12 criterion scores are integrated to form 4 scores at principle level using an operator that weights the different criteria while strongly limiting between-criteria compensations.

From principles to the overall assessment: what problem to answer?

To produce an overall assessment at unit level, we have to aggregate the four principle scores into a single evaluation. The choice of aggregation method depends closely on the type of issue the stakeholders intend to resolve with an overall assessment, ie the type of decision they want to be able to take based on the unit evaluations. For example, if the aim is to select the best unit, an optimisation method may be relevant, but this is not the case if the aim is to distinguish several categories of welfare offered to animals. Decision-aiding systems commonly define four main reference problems (Roy 1996):

- the description of units based on a number of unit characteristics, expressed in an appropriate language shared by all stakeholders;
- the choice of a subset of units, that is kept as narrow as possible in order to facilitate the final choice of only one unit, ie the best on offer;
- the sorting of units, achieved by assigning each unit into pre-defined categories defined in terms of norms (ie situations that stakeholders consider as typifying the different categories);
- the ranking of units, from best to worst.

The description is a prerequisite for the three other types of problems, because the description has to be formalised to make it possible to issue recommendations (which is the objective of the three other types of problems). However, the description alone is not sufficient to generate a recommendation which inevitably includes value judgements. The Welfare Quality® assessment tool has been developed as a

https://doi.org/10.1017/S0962728600000762 Published online by Cambridge University Press

decision aid for stakeholders, ie to help farmers design welfare improvement strategies, to help certification bodies decide whether or not to incorporate units into a certification scheme, to help consumers decide whether or not to buy a certain product, etc. Therefore, the assessment model shall result in neither the choice of best unit nor in the ranking of all units, but rather in an absolute evaluation of the level of welfare offered by the units to the animals, with welfare defined in several categories, from very low to excellent welfare. Consequently, in the context of Welfare Quality®, the overall assessment of animal welfare corresponds to a sorting problem aimed at assigning animal units to predefined ordered categories of welfare.

Materials and methods

The animal units are to be assigned to welfare categories according to the scores they obtained on the four welfare principles (Good feeding, Good housing, Good health, Appropriate behaviour) expressed on the 0–100 value scale. We decided to use a multicriteria analysis methodology similar to ELECTRE TRI (see, for example, Moscarola & Roy 1977) in order to respect the multidimensionality of animal welfare (Mousseau & Slowinski 1998). Briefly, the animal units are compared to reference profiles outlining the categories to be distinguished. The tool is then developed through three steps: (i) definition of the welfare categories; (ii) position of the reference profiles and (iii) definition of the rules for comparing units with reference profiles.

To go through these three steps, we consulted both scientists (in social and animal sciences) and stakeholder groups with an interest in the development of an animal welfare assessment tool. Several groups of experts were thus consulted: the Welfare Quality® Advisory Committee composed of representatives from a producer group, a retailer, two animal protection groups, the Federation of Veterinarians of Europe, and political institutions (the World Organisation for Animal Health and the European Commission Directorate General for 'Health and Consumers'), a task force composed of 5 social scientists and 9 animal scientists, all partners of Welfare Quality®, and the Management Committee of the project, made up of 2 social scientists and 6 animal scientists.

An overview of the welfare assessment method was presented at the first meeting between the Management Committee and the Advisory Committee. Together with the Management Committee, we suggested a list of potential uses stakeholders could make of an overall animal welfare assessment tool. We then asked the individual Advisory Committee members to react to and comment on the proposed scenarios, specifying their own organisations' views on potential uses of the Welfare Quality® assessment system. The task force collated these views to establish a final set of scenarios representing all the potential uses at least one organisation considered interesting, defined welfare categories so as to implement these scenarios, proposed profiles to boundary welfare categories, and initiated discussion on possible procedures for allocating each unit to one of the defined welfare categories. The Management Committee took the final decision on the reference profiles. Concerning the assignment process, several of the procedures put forward by the task force were tested on a set of real farms. We thus assessed 69 randomly chosen dairy farms in Austria, Germany and Italy with a view to obtaining farms in each welfare category, so as to check that the welfare assessment was able to detect differences between farms. The observers visiting farms in Austria and Germany (n = 44) also gave their own general impression of the farms they visited, using a 12-cm visual-analogue scale, from a very bad to an excellent level of welfare. We ran logistic regressions to check whether the farm classifications obtained using the different procedures fitted with the general impression of the observers auditing the farms. A Likelihood-Ratio test was used to estimate the strength of these links between the observer's impression of a given farm and the category reached by that farm according to each of the proposed procedures. Finally, we consulted the Advisory Committee during a second group meeting on the most appropriate procedure to be applied, taking into account theoretical aspects, the distribution of the 69 tested farms, and the agreement of the resulting farm categorisation system with observers' general impressions.

Results

Definition of welfare categories

Four scenarios were defined for overall welfare assessment tool utilisation. These scenarios are not mutually exclusive:

Scenario 1

The assessment system is used to set a minimum level of animal welfare that is acceptable; it could either replace the current EU regulations on minimum animal welfare standards or help in implementing a regulation on compulsory labelling with a yes/no label or a system using several levels (similar to the one used for energy consumption by household appliances).

Scenario 2

The assessment system is used to distinguish instances in which the welfare of the animals not only exceeds the minimum acceptable level but attains high levels, sufficient to implement a voluntary labelling system.

Scenario 3

The assessment system is used by the unit manager as a self-assessment management tool to identify welfare problems or risks, and to monitor improvements.

Scenario 4

The assessment system is used as a research tool to evaluate farming systems and practices.

In order to be implementable, each of these four scenarios requires a number of welfare categories to be distinguished. The number of categories required may differ between scenarios: from two categories for Scenario 1 to at least four categories for Scenarios 3 and 4 (Table 2). This means that all four of the following categories are needed to enable

https://doi.org/10.1017/S0962728600000762 Published online by Cambridge University Press

Table 2 Implications of potential uses on the welfare categories distinguished.

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

Table 3 Utilisation of welfare categories for each potential scenario. Depending on the use targeted, the response could be binary (eg is the level of welfare in accordance with legal requirements?) or gradual (is the farm –, +, ++ or +++).

Welfare categories	Voluntary label (Niche market: scenario 2)	Voluntary label (Quality label: scenario 2)	Compulsory label (scenario 1)	Self-assessment or research tool (scenarios 3 and 4)
Excellent	Yes	Yes	Yes	+++
Enhanced	No	Yes	Yes	++
Acceptable	No	No	Yes	+
Not classified	No	No	No	–

simultaneous implementation of all four scenarios cited above:

• Excellent, the welfare of the animals is excellent — the animal unit may correspond to a niche market, via a label guaranteeing consumers very high quality products (this label could be dedicated to animal welfare);

• Enhanced, the welfare of the animals is good but not excellent — good farming practices are applied and are sufficient to ensure a good level of welfare within a more general quality label which could encompass other product characteristics, like food quality, food safety, environmental impact;

• Acceptable, the welfare of the animals is acceptable (above minimal requirements defined for a compulsory label), but insufficient for the farm to enter a certification scheme based on respect for animal welfare; and

• Not classified, the welfare of the animals is low and considered unacceptable.

The use of the above welfare categories would differ from one scenario to another, leading to either a binary answer or a graded answer (Table 3).

Reference profile

Since four ordered welfare categories are to be distinguished, three reference profiles have to be defined corresponding to the limits between consecutive categories. Each reference profile corresponds to a set of four values expressed on the 0–100 scale, one per principle.

The 0–100 scale used to express principle-scores has the same meaning whatever the principle. Therefore, we chose to set flat reference profiles, ie identical scores on all principles (Figure 1). This also appeared consistent with the idea that all welfare dimensions are equally important (see *Background information*). The experts consulted to construct the principles were told that at below 20, a farm would not be considered acceptable. Therefore, 20 was chosen for Profile 3 (ie the lower limit for the 'Acceptable' category). The experts were also told that 50 means 'neither good nor bad'. Hence, a slighter higher score (55) was chosen for Profile 2 (lower limit for the 'Enhanced' category). Profile 1 (lower limit of the 'Excellent' category) was set at 80 to be symmetrical to Profile 1. The four scores constituting a given profile correspond to 'aspiration values' to be achieved by an animal unit to reach the category set by this profile.

Rules for comparing farms with reference profiles

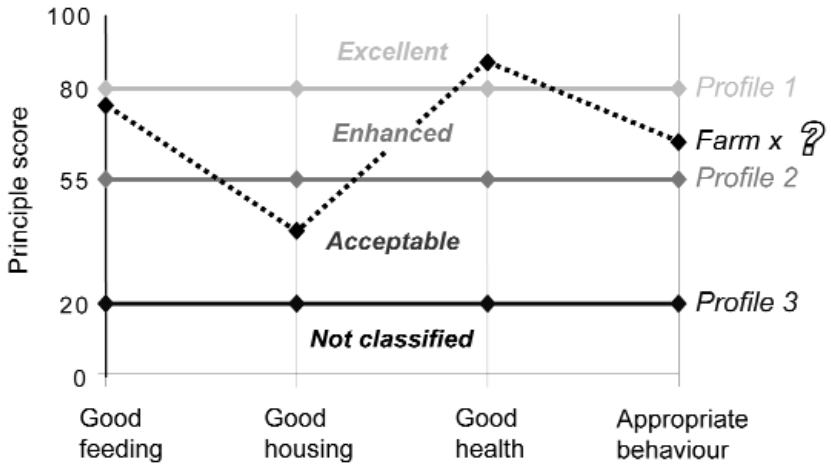
Elaboration of a set of potential procedures

The scores obtained by an animal unit on the four welfare principles are to be compared to the scores for profiles 1, 2 and 3. Rules need to be set to decide when an animal unit is considered at least as good as a profile. An intuitive rule is

https://doi.org/10.1017/S0962728600000762 Published online by Cambridge University Press

Figure 1

Reference profiles boundarying welfare categories. A Farm *x* is compared to each profile and then assigned to the welfare category above the profile that the farm is considered at least as good as.



unanimity, whereby a unit is considered better than a profile if it scores higher than that profile on all welfare principles, ie it always reaches the aspiration values set for that category. If a unanimity rule is taken, then, to be excellent, an animal unit needs to score at least 80 on every principle, to be considered enhanced it needs to score at least 55 on every principle, and to be considered acceptable it needs to score at least 20 on every principle. Less strict rules may be applied where, to be assigned to a given category, a unit only has to score better than the aspiration values of that category on a given number of principles. However, it would be inappropriate to consider a unit as excellent if it scores more than 80 on three principles but only 10 on the fourth one. Consequently, even when more lenient rules are applied, there is a fixed minimum requirement for the remaining principles. In our assessment system, this minimum generally equals the aspiration values of the next lowest category. For instance, we may decide to consider an animal unit as excellent if it scores higher than 80 on two principles (80 being the aspiration value for excellence) and higher than 55 on the two remaining principles (55 being both the minimum requirement for excellence and the aspiration value for the enhanced category). This kind of rule system makes it necessary to define a Profile 4 corresponding to the minimum requirement for the bottom-but-one category (acceptable). We decided to set Profile 4 at 10 for all principles, corresponding to half the scores of Profile 3.

We thus investigated five different procedures for assigning animal units to welfare categories (see below):

Procedure 1, Unanimity rule (4/4/4): to be assigned to a given category, an animal unit needs to reach the aspiration value of that category on all 4 principles, eg to be considered excellent, an animal unit must score at least 80 (Profile 1) on all principles.

Procedure 2, Qualified-majority rule (3/3/3): to be assigned to a given category, an animal unit must reach the minimum requirement of that category on all principles, and reach the aspiration value on 3 of them, eg to be considered excellent, an animal unit must score at least 55 (Profile 2, minimum requirement) on all principles and reach 80 (Profile 1, aspiration value) on three of them.

Procedure 3, First set of mixed rules (2/3/4):

- to be considered excellent, an animal unit must score at least 55 (Profile 2) on all principles and reach 80 (Profile 1) on 2 of them;
- to be considered enhanced, an animal unit must score at least 20 (Profile 3) on all principles and reach 55 (Profile 2) on 3 of them;
- to be considered acceptable, an animal unit must score at least 20 (Profile 3) on all 4 principles;
- otherwise, the animal unit is not classified.

Procedure 4, Second set of mixed rules (2/2/3):

- to be considered excellent, an animal unit must score at least 55 (Profile 2) on all principles and reach 80 (Profile 1) on 2 of them;
- to be considered enhanced, an animal unit must score at least 20 (Profile 3) on all principles and reach 55 (Profile 2) on 2 of them;
- to be considered acceptable, an animal unit must score at least 10 (Profile 4) on all principles and reach 20 (Profile 3) on 3 of them;
- otherwise, the animal unit is not classified.

Procedure 5, Third set of mixed rules (1/2/3):

- to be considered excellent, an animal unit must score at least 55 (Profile 2) on all principles and reach 80 (Profile 1) on 1 of them;
- to be considered enhanced, an animal unit must score at least 20 (Profile 3) on all principles and reach 55 (Profile 2) on 2 of them;

https://doi.org/10.1017/S0962728600000762 Published online by Cambridge University Press

Table 4 Consequences of the different procedures on how five virtual farms are categorised, the farms are characterised by their scores on the four principles, noted here P1, P2, P3 and P4 for Good feeding, Good housing, Good health and Appropriate behaviour, respectively. The five farms are differentiated on P2.

	Principle scores				Welfare category according to each procedure*				
Farm number	P1	P2	P3	P4	Procedure 1	Procedure 2	Procedure 3	Procedure 4	Procedure 5
1	60	90	95	70	Enhanced	Enhanced	Excellent	Excellent	Excellent
2	60	65	95	70	Enhanced	Enhanced	Enhanced	Enhanced	Excellent
3	60	45	95	70	Acceptable	Enhanced	Enhanced	Enhanced	Enhanced
4	60	14	95	70	Not classified	Acceptable	Not classified	Acceptable	Acceptable
5	60	4	95	70	Not classified	Not classified	Not classified	Not classified	Not classified

*See text for explanation on procedures.

• to be considered acceptable, an animal unit must score at least 10 (Profile 4) on all principles and reach 20 (Profile 3) on 3 of them;

• otherwise, the animal unit is not classified.

In all cases, the sorting procedure follows a descending order: for a given farm, we first check whether it can be assigned to the category 'Excellent', and if not, we check whether it can be assigned to the second-best category 'Enhanced', and so on.

To illustrate the differences between Procedures 1 to 5, we tested them on a set of five virtual farms (Table 4). These farms were characterised by the same principle-scores on P1, P3 and P4 (ie respectively, Good feeding, Good housing and Appropriate behaviour), but differed on P2 (Good housing). Table 4 clearly illustrates how the five procedures differ in severity (to the extent of a difference of one category for a given farm), the most severe being Procedure 1 (unanimity) and the most indulgent being Procedure 5.

Choice of the most appropriate procedure

Figure 2 shows the distribution of the 69 dairy farms visited across the four welfare categories according to the five assignment procedures described above. Using Procedures 1 and 3, 51% of the 69 farms fell in the lowest category (not classified), no farms or only 3% were enhanced, and no farm was deemed excellent. Procedure 2 clearly led to a non-discriminatory distribution, with 83% of farms in the 'Acceptable' category. With Procedures 4 and 5, a majority of the farms (67%) fell in the category 'Acceptable' and 14% of farms were not classified. Still no farms were considered excellent. After a close look at the data collected on these farms, it proved reasonable to consider that none of the farms should be considered excellent. Indeed, there was a high incidence of lameness, injuries and disease, and no farm could be considered excellent on two welfare principles.

Procedures 4 and 5 revealed a significant relationship between the welfare categories and the general impression of the observers who audited the farms, with χ^2 values of 5.56 and associated probabilities of 0.02, respectively. Procedures 1, 2 and 3 led to non-significant links with the observers' impressions ($\chi^2 = 2.58$, 1.28 and 3.80, $P > 0.05$).

The five procedures were presented to the Welfare Quality® Advisory Committee, together with the farm distributions and the data on concordance with observers' general impressions. Procedures 1 and 3 were excluded because of being either too severe (allowing few or no farms to reach the 'Enhanced' category) or not reflecting observers' impressions. Procedure 2 was excluded because it was not discriminatory between farms (the vast majority of farms were 'Acceptable') and not reflecting observers' impressions. Procedure 5 appeared too flexible, as only one score out of four needed be excellent for a farm to be considered as 'Excellent' all-round. Procedure 4 was finally retained as it appeared to offer a good compromise between severity, discrimination of farms, and concordance with observers' general impressions.

Discussion and conclusions

The construction of the assessment system proposed in Welfare Quality® respects the multidimensionality nature of animal welfare, relying on four main principles to be fulfilled to ensure welfare. In addition, the assessment model allows the formalisation of the reasoning followed by animal scientists, social scientists and potential users for scoring the animal units. The model relies on both theoretical aspects (eg for the definition of the reference profiles boundarying the four welfare categories) and on what can be realistically achieved in practice (eg data collected within Welfare Quality® on real farms were used to choose appropriate assignment procedures). The following procedure is proposed for assigning animal units to welfare categories based on the scores obtained on the four principles: 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; it is considered 'acceptable' if it scores more than 10 on all principles and more than 20 on three of them. If the farm fails the final condition, it is deemed

https://doi.org/10.1017/S0962728600000762 Published online by Cambridge University Press

not classified. This procedure still needs to be tested on the other animal types studied in Welfare Quality®.

Since the system formalises the reasoning of experts while describing all the calculations necessary to go from measures to overall assessment, it can easily be standardised and automated for routine use. This is an essential characteristic for an evaluation system to be used within certification schemes or other systems that require a normative procedure. To be implementable, this kind of assessment system has to be transparent, with information accessible to any interested party. To ensure standardisation and availability of information, technical documents are currently prepared to ensure such transparency with the help of the Nederlands Normalisatie-instituut (NEN). In parallel, a software tool is being developed to facilitate on-farm or in-slaughterhouse data collection and storage, and streamline the automated calculation of welfare scores and categorisation of animal units.

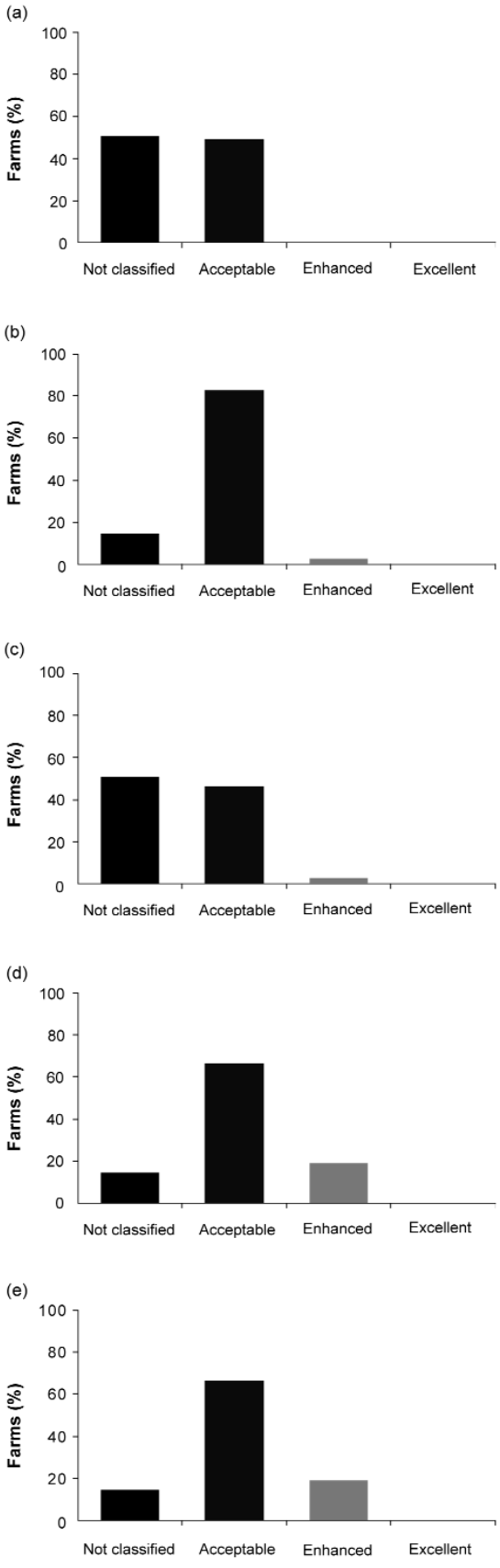
We are aware that the proposed assessment system is not perfect. Any method relying on the definition of a small number of categories will always carry threshold effects, and a farm or slaughterhouse manager may not have an interest in first improving the most problematic point detected on their farm/slaughterhouse. This could be the case of a farm with 35 on good feeding, 55 on good housing, 20 on good health and 50 on appropriate behaviour, which would be categorised as acceptable. The most problematic point of this farm is health, yet improving its health score from 20 to 54, which would be a major improvement, would still not be enough to upgrade its classification. On the contrary, to become enhanced, the farmer could opt to work on increasing its behaviour score just 5 points from 50 to 55, which would only be a relatively small improvement. Hence, there may be only limited benefits for the animals. However, this non-desired effect is minimised by the fact that it is risky for the farmer to remain at a score around 20 on health, because a slip of just one point on this principle would be sufficient to make the farm drop into the 'not classified' category.

To be fully efficient in helping farmers and slaughterhouse managers improve the level of animal welfare, the assessment system proposed in Welfare Quality® should be completed by advice provided by a person qualified to identify the root causes of the problems highlighted through the assessment process. Once these causes have been properly detected, the unit manager can be offered remedial solutions. The assessment system described may well help advisers to identify causes of poor welfare by focusing attention on specific issues. It may thus act as a useful step towards improving the global level of welfare offered to animals in Europe throughout their life, from farm to slaughter.

Animal welfare implications

The assessment system elaborated by Welfare Quality® (including the description of measures and the calculation of scores) could be seen as a reference method serving several purposes. The assessment can be used: i) to provide farmers with a broad picture of the welfare status of their animals and

Figure 2



Distributions of 69 European dairy farms between the four welfare categories, using the five proposed procedures (Figures 3a, 3b, 3c, 3d, 3e for procedures 1, 2, 3, 4, 5, respectively).

https://doi.org/10.1017/S0962728600000762 Published online by Cambridge University Press

to identify the aspects requiring attention; ii) to give policy-makers a better understanding of the welfare implications of farming systems and practices; iii) to certify animal units on welfare grounds, by implementing the assessment system on units to be certified or by certifying the system and the practices employed on these units, and iv) to facilitate informed decision by stakeholders, including consumers.

Acknowledgements

We thank the Welfare Quality® partners and members of the project's Advisory Committee who acted as experts for fine-tuning the evaluation model. We are also grateful to the Welfare Quality® partners who visited the dairy farms and provided us with their data. The present study is part of the Welfare Quality® research project 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.

References

Botreau R, Bonde M, Butterworth A, Perny P, Bracke MBM, Capdeville J and Veissier I 2007a Aggregation of measures to produce an overall assessment of animal welfare: Part 1, A review of existing methods. *Animal 1*: 1179-1187

Botreau R, Bracke MBM, Perny P, Butterworth A, Capdeville J, van Reenen CG and Veissier I 2007b Aggregation of measures to produce an overall assessment of animal welfare: Part 2, Analysis of constraints. *Animal 1*: 1188-1197

Botreau R, Veissier I, Butterworth A, Bracke MBM and Keeling LJ 2007c Definition of criteria for overall assessment of animal welfare. *Animal Welfare 16*: 225-228

Botreau R, Capdeville J, Perny P and Veissier I 2008 Multicriteria evaluation of animal welfare at farm level: an application of MCDA methodologies. *Foundations of Computing and Decision Sciences 33*: 1-18

Fraser D 2003 Assessing animal welfare at the farm and group level: the interplay of science and values. *Animal Welfare 12*: 433-443

Mason G and Mendl M 1993 Why is there no simple way of measuring animal welfare? *Animal Welfare 2*: 301-319

Moscarola J and Roy B 1977 Procédure automatique d'examen de dossiers fondée sur une segmentation trichotomique en présence de critères multiples. *RAIRO Recherche Opérationnelle 11*: 145-173. [Title translation: Automatic file examination based on trichotomic segmentation in presence of multiple criteria]

Mousseau V and Slowinski R 1998 Inferring an ELECTRE TRI model from assignment examples. *Journal of Global Optimization 12*: 157-174

Roy B 1996 *Multicriteria Methodology for Decision Aiding* p 57. Kluwer Academic: Dordrecht, The Netherlands

Veissier I and Evans A 2007 Rationale behind the Welfare Quality® assessment of animal welfare. *Proceedings of the 2nd Welfare Quality® Stakeholder Conference* pp 19-22. 3-4 May 2007, Berlin, Germany

Veissier I, Forkman B and Jones B 2007 *Proceedings of the Second Welfare Quality® Stakeholder Conference on Assuring Animal Welfare: From Societal Concerns to Implementation* pp 70-84. 3-4 May 2007, Berlin, Germany

https://doi.org/10.1017/S0962728600000762 Published online by Cambridge University Press