Applying Welfare Quality® strategy to design a welfare assessment tool for foxes and mink farms

R. Botreau[1,2], M. Gaborit[1,2] and I. Veissier[1,2]
[1]INRA, UMR1213 Herbivores, 63122 Saint-Genès-Champanelle, France; [2]Clermont Université, VetAgro Sup, UMR Herbivores, BP 10448, 63000, Clermont-Ferrand, France; raphaelle.botreau@clermont.inra.fr

Abstract

WelFur project aims to develop on-farm welfare assessment protocols for mink and foxes, using the Welfare Quality® framework. The evaluation relies on a sequential structure, where four welfare principles, subdivided into 12 criteria, are constructed and then aggregated to obtain an overall evaluation at farm level. WelFur partners developed measures to check the compliance of fox and mink farms with all criteria. We applied the multicriteria methods developed in Welfare Quality® to construct criteria, consulting two expert panels to set parameters. To aggregate criteria and then principles to produce an overall assessment, we directly used the Welfare Quality® protocol. The general framework remains the same, but several specificities emerged. First, the whole production cycle, from breeding to slaughter, occurs on the farm, implying to consider three different periods to cover the whole cycle. Depending on the period, number and types of animals, and resources used differ. Second, two different species of foxes and their hybrids are to be considered. They differ in behaviour and needs, impacting the interpretation of some measures. We thus had to run separate consultations for periods, animal types or species, and then aggregate them. This work shows that the rationale followed in Welfare Quality® can be extended to other species once taking into account the specificities linked to species or production systems.

Keywords: animal welfare, overall assessment, multicriteria evaluation, expert opinion, foxes, mink.

Introduction

This study is part of the WelFur project (2010-2012), supported by EFBA in collaboration with seven research institutes and universities. The objective of this project is to introduce a certification program at farm level that will guarantee a good level of welfare on certified farms. WelFur thus aims at applying to fox and mink farms the welfare assessment rationale developed in the Welfare Quality® project (2004-2009), initially developed for cattle, pigs and poultry.

Background information: architecture of the Welfare Quality® assessment tool

Since welfare is a multidimensional concept, covering several aspects such as physical comfort, absence of hunger, thirst, diseases, injuries and stress, and expression of positive behaviour (e.g. the five freedoms, Farm Animal Welfare Council, 1992), the Welfare Quality® project developed a multicriteria evaluation tool. This system is based on a sequential structure, where four principles of welfare (*good feeding*, *good housing*, *good health*, *appropriate behaviour*), subdivided into

12 criteria, are constructed and then aggregated to obtain an overall evaluation at farm level (Table 1). This structure was designed to be common to all animal production systems.

Farms are assessed on each criterion thanks to one to several measures relying on different aspects of the considered criterion. All the measures and the construction process are described in the Welfare Quality® protocols (Welfare Quality® consortium 2009a, b, c).

Specificities of fur farming systems in comparison with Welfare Quality®

The WelFur project aims at providing welfare assessment protocol over the whole production cycle of farmed foxes and mink. Even if the general construction procedure is the same as in Welfare Quality®, several specificities, listed below, have to be taken into consideration in the model construction for fur animals.

Each model built in Welfare Quality® was associated with only one species (e.g. cattle) and one type of animal (for cattle: fattening cattle, veal calves, or dairy cows). On the contrary, for foxes and mink systems the whole production cycle (including breeding period, lactation, weaning, fattening and finally slaughter and pelting) occurs on the same farm. As a consequence, to assess the farm it is necessary to consider all types of animals (adult males, adult females, under-weaning juveniles and growing juveniles). In addition, two different species of foxes and their hybrids are to be considered (commonly reared on the same farms): Blue fox and Silver fox. The different animal types and species differ in behaviour and needs, and thus the interpretation of a given measure can vary. Thus, the construction of some criteria needed to be adjusted according to the different animal types and/or species so that their differences could be taken into account for the interpretation of the related measure.

Table 1. Animal welfare principles and criteria defined in the Welfare Quality® project.

Principles	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 diseases
	8	absence of pain due to management procedures
Appropriate behaviour	9	expression of social behaviours
	10	expression of other behaviours
	11	good human-animal relationship
	12	positive emotional state

Moreover, to have an overall view of the whole fur farm, the entire production cycle is to be evaluated. Three periods (pelting to mating (P1) / mating to weaning (P2) / weaning to pelting (P3)) were defined and have to be assessed (Figure 1). Depending on the period, the number and types of animals (adult males, adult females, under-weaning and growing juveniles), and the resources used differ.

The construction of criteria was thus directly impacted by these specificities, e.g. at criterion level the data collected at several periods have to be integrated and this requires specific arrangements for the calculation of scores.

Material and methods

Construction process and methods

The overall welfare assessment is the result of three construction steps: (1) construction of criteria from measures, (2) construction of principles from criteria, and (3) aggregation of the four principles to produce an overall evaluation at farm level. Methods used are described in Botreau *et al.* (2008) for Steps 1 and 2, and in Botreau *et al.* (2009) for Step 3, in addition to Welfare Quality® protocols. Only Step 1 differs from one species to another, and thus was done in WelFur specifically for foxes and mink. Measures, developed by WelFur partners, related to a given criterion have to be interpreted and synthesized to produce one score at criterion level, expressed on a 0-100 scale indicating the degree of compliance of the farm with this specific criterion (from 0 the worst possible situation to 100 the ideal one). Methods used to calculate scores differ from one criterion to another in order to be adapted to the number of measures to be considered, to the type of scale on which the data is observed, and to the relative importance to be assigned to each measure within the criterion. Two main methods of calculation were

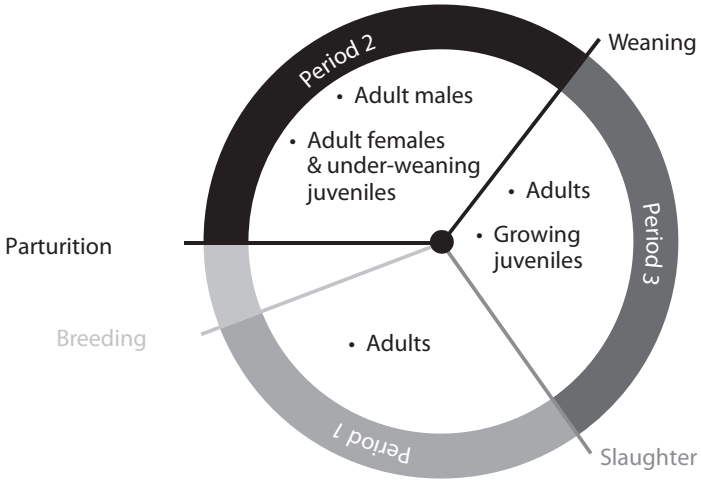

Figure 1. The three periods of the production cycle considered for foxes and mink.

thus used to construct the criteria: construction of a decision-tree, and calculation of weighted sums combined with non-linear transformation (thanks to I-spline functions). While deciding of the type of construction, we and WelFur partners considered the whole range of existing variability across the European production area (e.g. for mink from Finland to Greece). Finally, and specifically to WelFur, when a given measure is assessed at several periods of the production cycle and/or on several animal types (e.g. adult vs. growing juveniles), periods or animal types were aggregated using Choquet integrals (Grabisch and Roubens, 2000).

Consultation process

As in Welfare Quality®, the assessment tool was tuned to fit the average judgement of experts, thus modelling their opinion on (1) the interpretation in terms of welfare of the measures, (2) the relative importance of the different measures, criteria and principles, and (3) the level of compensation to be authorized between welfare issues. To do that with WelFur partners we created two expert panels (5 experts for foxes and 7 experts for mink), composed of animal scientists and vets, chosen for their knowledge on the two species in terms of physiology and behaviour and also on farming practices. We consulted each panel during a one day meeting. When necessary, we pursued the consultation by e-mail exchanges.

Data treatment

Depending on the type of construction, we treated the data using Excel (to calculate means and also to use the solver to set the optimized parameters for the construction), and/or GenStat® 11th edition, using the Biometris set of procedures, to fit I-spline functions (by least-squares minimization).

Results

Thanks to exchanges with WelFur partners in charge of the development of foxes and mink welfare measures, we identified the measures to be looked at for each criterion (Mononen *et al.*, 2012) and we proposed for each criterion a construction method. These choices led to 55 and 80 different questionnaires, for foxes and mink respectively. Only few examples are presented below in order to illustrate the different methods used.

Example 1: interpretation of the use and type of neck tongs in Period 1, as part of Criterion 8 'Absence of pain due to management procedures' in foxes

Two measures are considered to assess Criterion 8, including the handling procedure of animals using neck tongs. In Period 1, when only adult blue and/or silver foxes are present, a decision tree is designed including the frequency of use of the neck tong to handle animals, the type of material that covers the animal-end of the neck tong and the species which is handled with this use and type of neck-tong (since the behaviour of the fox in respect to the handling is different from one species to another). The resulting decision tree thus leads to seven possible situations (Figure 2).

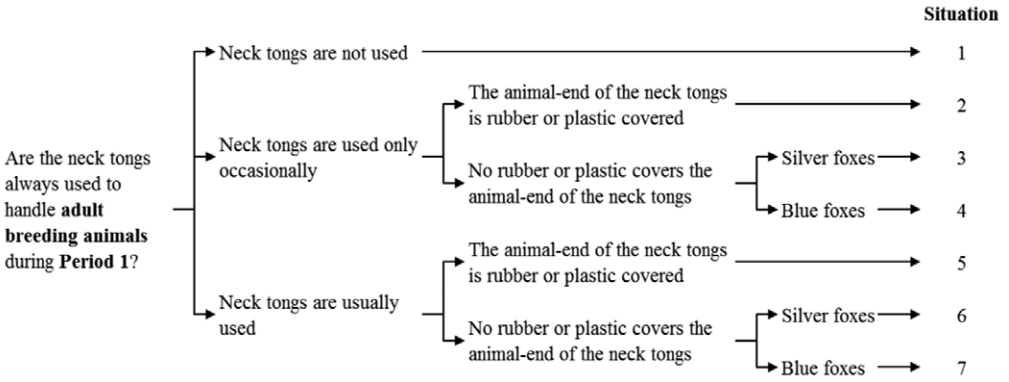

Figure 2. Decision tree used to assess the use and type of neck tongs in period 1 for foxes, as part of Criterion 8.

We asked the experts to assign a score to each of the seven possible situations from 1 to 7 (Table 2). These scores represent their own evaluation using the standardized scale [0,100]. The score to be assigned to each situation thus corresponds to the average score assigned by the experts. Since there may be several types of neck tongs and both species on a given farm, several situations may be observed simultaneously on the farm. Then, the score to be assigned to the farm is the one corresponding to the worst situation observed on at least 15% of the animals.

Example 2: interpretation of quality of the nest box in Period 3, as part of Criterion 3 'Comfort around resting' in mink

Criterion 3 (Comfort around resting) in mink is assessed through two measures observed at individual level during the three periods: the access to a nest box and the quality of the nest box/resting area. At cage level we consider four categories of quality of the nest box/resting,

Table 2. Experts' answers and final score (i.e. mean of experts' scores) assigned to each of this seven possible situations about use of the neck tong in period 1 for foxes, as part of Criterion 8.

Situation	Score assigned by experts					Score (mean)
	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	
Situation 6	0	0	0	15	0	3
Situation 7	10	0	7	15	0	6
Situation 3	30	17	60	50	0	31
Situation 4	40	17	67	50	0	35
Situation 5	20	17	33	40	100	42
Situation 2	50	100	93	95	100	88
Situation 1	100	100	100	100	100	100

corresponding to different severities of problems: good (Category 0), moderately good (Category 1), moderately bad (Category 2), and bad (Category 3) resting comfort. At farm level, the percentages of mink in each category of nest box quality (called p0, p1, p2 and p3, for categories 0, 1, 2 and 3 respectively) are combined in a weighted sum. This sum is then transformed into an index (*I*) that varies from 0 to 100, and finally this index is computed into a score using I-spline functions. To define the weights to be assigned to the categories and the parameters of the I-spline functions, we asked the experts to rank and then to score a set of virtual farms characterized by their percentages of animals in the four categories (cf. Table 3 for the questionnaire presented to experts, and Table 4 for their answers and the results of our calculations based on their answers).

Weights were defined using the solver of Excel in order to obtain the same ranking of farms with experts' mean score and Index value. Then, we defined the parameters of I-spline functions so that to minimize the sum of square errors between experts' mean scores and the calculated ones.

Example 3: aggregation of the scores obtained for periods 1, 2 and 3 for water provision, as part of Criterion 2 'absence of prolonged hunger' in mink

Table 3. Questionnaire presented to the experts to score quality of the nest box in period 3 for mink, as part of Criterion 3.

Farms	% good resting comfort (p0)	% moderately good resting comfort (p1)	% moderately bad resting comfort (p2)	% bad resting comfort (p3)	Rank	Score
Farm 1	100	0	0	0		
Farm 2	75	25	0	0		
Farm 3	75	0	25	0		
Farm 4	50	50	0	0		
Farm 5	50	0	25	25		
Farm 6	25	75	0	0		
Farm 7	25	50	25	0		
Farm 8	0	100	0	0		
Farm 9	0	75	25	0		
Farm 10	0	50	50	0		
Farm 11	0	50	25	25		
Farm 12	0	25	75	0		
Farm 13	0	0	100	0		
Farm 14	0	0	75	25		
Farm 15	0	0	50	50		
Farm 16	0	0	25	75		
Farm 17	0	0	0	100		

Table 4. Experts' answers, weights, index value and final score assigned to each of the 17 farms composing the questionnaire about quality of the nest box in period 3 for mink, as part of Criterion 3.

	Situation				Scores given by experts Ei								Index	Score
	p0	p1	p2	p3	E1	E2	E3	E4	E5	E6	E7	Mean		
Weights	**0**	**1**	**2**	**3**										
Farm 17	0	0	0	100	0	0	0	0	0	0	0	0	0	0
Farm 16	0	0	25	75	0	0	0	0	0	0	0	0	8	1
Farm 15	0	0	50	50	10	0	3	10	0	0	0	3	17	3
Farm 14	0	0	75	25	10	0	5	15	10	0	0	6	25	7
Farm 13	0	0	100	0	30	0	10	25	5	10	0	11	33	12
Farm 12	0	25	75	0	40	10	15	25	15	15	0	17	42	18
Farm 11	0	50	25	25	40	20	25	20	15	15	0	19	42	18
Farm 10	0	50	50	0	50	25	30	30	18	20	0	25	50	24
Farm 9	0	75	25	0	50	30	40	35	21	30	0	29	58	31
Farm 5	50	0	25	25	50	35	35	25	18	30	35	33	58	31
Farm 7	25	50	25	0	60	40	45	40	20	50	25	40	67	40
Farm 8	0	100	0	0	60	50	50	40	20	50	20	41	67	40
Farm 6	25	75	0	0	75	60	60	50	19	60	30	51	75	51
Farm 3	75	0	25	0	75	70	65	60	50	70	50	63	83	64
Farm 4	50	50	0	0	90	80	65	70	30	80	40	65	83	64
Farm 2	75	25	0	0	100	90	70	80	75	90	60	81	92	80
Farm 1	100	0	0	0	100	100	100	100	100	100	100	100	100	100

To assess the absence of prolonged thirst on a mink farm, several questions based on resources and management are used, which differ according to the period by taking into account the risk of freezing and the risk of overheating during each period. The Criterion 2 is assessed thanks to three decision trees specific to each period. After calculating the sub-score for each period (similar to Example 1), we need to merge them into a global score covering the three periods. Consulted experts were asked to react on combinations of sub-scores obtained on periods 1, 2 and 3, assigning a score to each of the combinations (the first 4 columns of Table 5 correspond to the questionnaire and the last 9 columns to experts' answers and calculated scores).

We used the Excel solver to find the parameters of the Choquet integral that best match the experts' mean scores by least square minimization. Two indices can be calculated to help the interpretation of the parameters of the Choquet both in terms of relative importance of the three periods and of interaction between sub-scores.

The analysis of results highlights two characteristics in experts' reasoning. First, there is a clear hierarchy between periods: Period 2 is more important than Period 3 which is more important

Table 5. Experts' answers and final criterion-score assigned to each of the 13 farms composing the questionnaire about aggregation of the three periods' sub-scores for water provision for mink, as part of Criterion 2. Grey scales represent general marking of scenarios addressing the relative importance of welfare scores between periods.

Farms	Score for period 1	Score for period 2	Score for period 3	Experts' answers								Criterion-score
				E1	E2	E3	E4	E5	E6	E7	Mean	
Farm 1	25	50	75	25	40	27	35	53	50	30	37	38
Farm 2	25	75	50	30	45	30	40	75	60	30	44	44
Farm 3	40	50	60	40	50	42	50	48	50	40	46	45
Farm 4	40	60	50	50	50	45	50	60	50	40	49	48
Farm 5	50	25	75	25	35	25	30	30	30	25	29	28
Farm 6	50	40	60	40	50	40	45	38	45	40	43	41
Farm 7	50	50	50	50	50	50	50	50	50	50	50	50
Farm 8	50	60	40	50	50	40	50	65	50	45	50	46
Farm 9	50	75	25	30	40	25	40	70	30	25	37	39
Farm 10	60	40	50	40	50	40	45	35	40	40	41	41
Farm 11	60	50	40	40	50	40	40	40	40	45	42	43
Farm 12	75	25	50	25	35	25	30	25	25	20	26	27
Farm 13	75	50	25	25	40	25	35	45	25	25	31	32

than Period 1 in respect to 'absence of prolonged thirst' (with indices of importance of 0.52, 0.30, and 0.18, respectively). This may be explained by the fact that there are much more animals in Periods 2 and 3 than in Period 1 and that in Period 2 a problem around water provision may be dramatic for lactating females and their litters. Second, all interaction indices are positive (between 0.23 and 0.45), showing that the experts want to limit compensations between the periods: to obtain a low score on one period is sufficient for the farm to be considered bad for that measure (i.e. experts are not satisfied when one of the three scores is low). This can be easily seen in Table 6: all farms (with an average score of 50) are scored below 50, even Farm 2 that scores only 44 whereas it would have scored 58.5 using a weighted sum for the calculation, showing that this farm is penalized for the bad score (25) obtained in Period 1.

Discussion

Variability among experts exists. In Welfare Quality® similar differences between experts also appeared. Since we are dealing here with an interpretation process, it is not surprising to find such variability in answers, illustrating the diversity of viewpoints.

One important step was lacking when preparing this paper: the test of the assessment protocols on commercial farms over a whole production cycle. Concerning raw data, all three periods were tested on mink and fox farms in 2011 and early 2012. However score calculations will be

possible only when the software derived from foxes and mink welfare assessment protocols will be finalized. This software is under construction and would be ready in a β version for summer 2012. It would be then possible to process the calculations and present the results to experts to check the plausibility of the results. The objective is to make the final protocols and software tool available through the web by the end of 2012.

As a conclusion, this work shows that the rationale developed in Welfare Quality® for cattle, pigs and poultry may be successfully adapted to other species, taking however into consideration the specificities of their production systems. For foxes and mink, these adaptations led to a higher complexity of the construction so that the different animal types and periods of the production-cycle can be considered. One indicator of this complexity is the number of questionnaires to be filled in by experts: 14 for dairy cows on farm in Welfare Quality® vs. 55 and 80 for foxes and mink farms respectively.

Acknowledgements

We thank all WelFur partners and members of the expert panels for fine-tuning the evaluation models. We are especially grateful to the partners in charge of the development of the measures for their great participation to our work, for the fox team: Leena Ahola, Jaakko Mononen, Tarja Koistinen, and Hanna Huuki, and for the mink team: Steen Henrik Møller, Steffen W. Hansen, and Jens Malmkvist. Finally, we thank EFBA for the financial and practical support to the WelFur research project.

References

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: 287-316.

Botreau, R., Veissier, I. and Perny, P., 2009. Overall assessment of cow welfare: strategy adopted in Welfare Quality®. Animal Welfare 18: 363-370.

Farm Animal Welfare Council, 1992. FAWC updates the five freedoms. Veterinary Record 17: 357.

Grabisch, M. and Roubens, M., 2000. Application of the Choquet integral in multicriteria decision making. In: Grabisch, M., Murofushi, T. and Sugeno, M. (eds.) Fuzzy Measures and Integrals – Theory and Applications, Studies in Fuzziness and Soft Computing. Physica Verlag: Heidelberg, Germany. pp. 348-374.

Mononen, J., Møller, S.H., Hansen, S.W., Hovland, A.L., Koistinen, T., Lidfors, L., Malmkvist, J., Vinke, C.M. and Ahola, L., 2012. The development of on-farm welfare assessment protocols for foxes and mink: the WelFur project. Animal Welfare, in press.

Welfare Quality® consortium, 2009a. Welfare Quality® Assessment Protocol for Cattle. Lelystad, the Netherlands.

Welfare Quality® consortium, 2009b. Welfare Quality® Assessment Protocol for Pigs. Lelystad, the Netherlands.

Welfare Quality® consortium, 2009c. Welfare Quality® Assessment Protocol for Poultry. Lelystad, the Netherlands.