Netherlands Journal of Agricultural Science 47 (1999) 93-104

Overall welfare assessment of pregnant sow housing systems based on interviews with experts

M.B.M. BRACKE[1,*], J.H.M. METZ[1], B.M. SPRUIJT[2] AND A.A. DIJKHUIZEN[3]

[1] DLO-Institute of Agricultural and Environmental Engineering (IMAG-DLO), P.O.Box 43, NL-6700 AA Wageningen, The Netherlands
[2] Animal Welfare Centre – Faculty of Veterinary Medicine, University of Utrecht, Yalelaan 17, NL-3584 CL, Utrecht, The Netherlands
[3] Department of Farm Management, Wageningen Agricultural University, Hollandseweg 1, NL-6706 KN Wageningen, The Netherlands
[*] Corresponding author (fax: +31-317-425670; e-mail: m.b.m.bracke@imag.dlo.nl)

Received 10 Sept. 1998; Accepted 24 February 1999

Abstract

In interviews with 11 pig experts the main housing systems for pregnant sows were identified as tethering (T), individual housing in stalls (IS), group housing with stalls (GS), trickle feeding or biofix (B), electronic sow feeding (ESF), and outdoor housing with huts (O). The family pen system (Fam) was added as a reference system.
The experts were asked to give a welfare score for each housing system. The two individual housing-systems (mean scores: T = 1.8; IS = 2.3) scored significantly lower than more intensive indoor group-housing systems (GS = 5.4; B = 5.3; ESF = 6.2), and these scored lower than the more extensive systems (O = 8.0; Fam = 9.1; ANOVA, $P < 0.001$). Furthermore, T ranked lower than IS in the Sign test ($P = 0.008$).
The most important aspects for welfare assessment were space, substrate, feeding-related agonism and social parameters such as group size and group stability. Three different models were constructed to calculate welfare scores from the arguments given by the experts. When represented graphically the results seem comparable to the expert scores, although two of the three models differed significantly from the expert scores using analysis of variance. These results indicate that pig experts are able to perform overall welfare assessment in a rational way that allows modelling and that there is a consensus underlying welfare assessment. These outcomes provide support for the further development of a decision support system to assess farm animal welfare on a scientific basis.

Keywords: animal welfare, pigs, decision support system, model

Introduction

Overall welfare assessment is a highly desired goal of policy makers in the Western countries. This has resulted in the development and political employment of practical welfare-index systems for on-farm use in certain countries, for example in Austria (Bartussek, 1988 and 1997) and Germany (Sundrum et al., 1994). However, these authors also acknowledge that the scientific basis should be improved (Bartussek, 1997; Sundrum, 1997). On the other hand, several authors have stated that the task of performing objective overall welfare assessment is complex and may be too difficult for science (Rushen & De Passillé, 1992; Fraser, 1995; Dawkins, 1997). While individual humans generally have little problems in forming private opinions about the welfare status of animals, finding an objective basis appears to be more difficult. Authors differ on how to define welfare and, while there seems to be consensus that for welfare assessment many factors should be taken into account, the problem of weighting different aspects remains open for study.

In an attempt to meet the demand for a more objective welfare-assessment tool we have started to develop a computerised decision support system for assessment of the welfare status of pigs (Bracke *et al.*, 1997; Bracke *et al.*, submitted). Such a system takes a description of a housing system as input and uses scientific knowledge collected in a database to yield output in the form of a welfare score. Such a system also requires a formalised procedure to perform overall welfare assessment in an explicit way, but no standard exists for such a task. In this study we take a group of scientific experts as the standard to collect quantitative data concerning welfare assessment which can be used in the development of a decision support system. In particular the objective of this study was to establish input-output relationships between housing and welfare; to determine which are the most important attributes of housing systems of pregnant sows with respect to welfare; and to assess the perspectives of modelling of welfare assessment. For this purpose we asked pig experts to give welfare scores for the main housing systems for pregnant sows and we modelled their arguments to assess whether subjective welfare scores can be recalculated from attributes of the housing systems.

Materials and methods

Interviews

Eleven scientists from six West-European countries, who had all been involved in the development of welfare friendly housing systems for pigs, were interviewed in independent sessions. They performed the following tasks:

1. Identify the most important housing systems for pregnant sows.
2. Compare the answers given with the following predefined list: tethering (T); individual housing in stalls (IS); group housing with stalls (GS); group housing with trickle feeding, i.e. biofix (B); group housing with transponders, i.e. electronic sow feeding (ESF); outdoor housing with arcs/huts (O); and the family pen system (Fam). Descriptions of these housing systems can be found in the booklets of the Pig Welfare Advisory Group (Anonymous, 1997b), EU Scientific Veterinary Committee Report (Anonymous, 1997a), Svendsen & Svendsen (1997) and Backus (1997). Some references for the family pen (Fam) system include Stolba

(1981) and Wechsler (1996). The Fam system, while not being a competitive production system, was included as a reference system. In relation to Fam it was pointed out to the expert that we were interested in the welfare status of the pregnant sows only. It was specified that we were not interested in the welfare of the piglets in this system, nor in the economic, labour or environmental values of any of the housing systems, unless they affected the overall welfare status of the pregnant sows.

3. Give an overall welfare score for typical-example farms of each housing system on a scale from 0 (worst) to 10 (best). A typical-example was explained as being a most typical farm rather than as one with particularly good or bad welfare conditions. In addition, the expert was asked to give scores only for systems he/she knew well. As a result each expert specified his/her own systems within the limits of the main types of housing systems. This procedure allowed highly variable attributes such as management or stockmanship to be specified and it ensured maximally realistic conditions while still allowing generalisation of the results.
4. Justify each score. Each expert was asked to explain all scores including why the best and worst scores were not 0 and 10 respectively. This was done by listing the positive and negative attributes for each system and by explaining differences between adjacent scores. In addition to the scores for the main housing systems, the experts were free to add variations of housing systems by specifying new scores, for example for the same ESF system with and without straw.

Modelling

Modelling proceeded as follows. A welfare value is assigned to each attribute level in the list of welfare relevant attributes. An example of such an attribute is ‘straw quantity’. Its levels may be ‘no straw’, ‘some straw’ and ‘deep straw’. Only one attribute level can be true for a given housing system. All levels within each attribute receive a welfare value between 0 (worst) to 10 (best) according to their welfare rank. Weighting factors, which apply across attribute levels, were not used in this study, i.e. within each model all levels weighed equally. For every housing system, overall welfare scores are calculated in an additive way as the average of all welfare values of all attributes (i.e. the sum of welfare values divided by the total number of attributes, cf. Table 4).

For example, an expert may explain the low score for one housing system by ‘lack of exercise’, as compared to ‘spacious field’ for the other, better housing system. Accordingly, pen space is being identified as a relevant attribute, with ‘lack of exercise’ being the worst and ‘spacious field’ being the best-ranked level. Corresponding welfare values are 0 and 10 respectively. Any intermediate levels, as specified by the expert, receive intermediate welfare values proportionally.

There were four methods to generate welfare scores. One consisted of the personal welfare scores as obtained directly from each expert. In addition, three models were constructed. In model A the set of attribute levels as mentioned by an expert were used to calculate scores for the housing systems as specified by the same expert. This was done for all experts to generate the set of data for model A, which is a within-expert model. Model B is an across-expert model, in which the different attribute-level lists of all experts were summarised and organised by removing doubles and by regrouping levels in a logical way. For example, when one expert makes a distinction between the levels 'straw is present' and 'no straw', and another expert between 'some straw' and 'deep straw', the integrated list contains 'no straw', 'some straw', and 'deep straw' (with welfare values of 0, 5 and 10 respectively). In this way all logical distinctions made by the experts were retained in model B. Model C consists of a selection of the more important attributes in model B as indicated by the degree these had been used by the experts to explain their welfare scores.

To test effects of housing, experts and expert-housing interactions on the expert scores analysis of variance (ANOVA) was used (Anonymous, 1993). The models were also tested using ANOVA to see whether any of the scores generated by the three models differed significantly from the scores given by the experts. ANOVA requires the data to be normally distributed. This was confirmed by evaluation of normal plots of residuals. The null hypothesis was that there is no difference, implying that all four methods (the expert scores and the models) generate equivalent welfare scores for the main housing systems for pregnant sows. A one-sided Sign test was used to compare ranks of the housing systems that did not differ significantly in the ANOVA test.

Results

The experts generally agreed that the main housing systems for pregnant sows are tethered (T), individual housing in stalls (IS), group housing with feeding and/or lying stalls (GS), trickle feeding or biofix (B), transponder or electronic sow feeder systems (ESF) and outdoor housing with arcs/huts (O). Other variants include indoor floor feeding and various trough feeding systems. The biofix system, where food is delivered into the trough at a controlled rate (e.g. Olsson *et al.*, 1986), was the least well-known of the 6 major housing systems in that three experts felt insufficiently familiar with the biofix system to give a welfare score.

Significant differences (ANOVA, $P < 0.001$) were found between the mean expert-scores for individual, confined housing (T, IS) versus loose, group housing (all other systems), and between generally more intensive indoor group housing (GS, B, ESF) versus more extensive group housing (O, Fam). The first group (T, IS) scored worst; the latter (O, Fam) scored best for overall welfare (Figure 1; Table 1). No differences were found between the mean scores for T and IS, between GS, B and ESF, or between O and Fam. The T system was ranked significantly worse for welfare than the IS system (one-sided Sign test, $P = 0.008$). No further significant differences in ranks between the indoor group housing systems or between the O and Fam system were found with the Sign test.

Table 2 shows various measures indicating the relative importance of the attributes which experts mentioned when explaining their personal welfare scores. The main criteria to add variants of housing systems by specifying additional overall scores were straw versus no straw (8 experts did so); stable versus dynamic groups (4 ex-

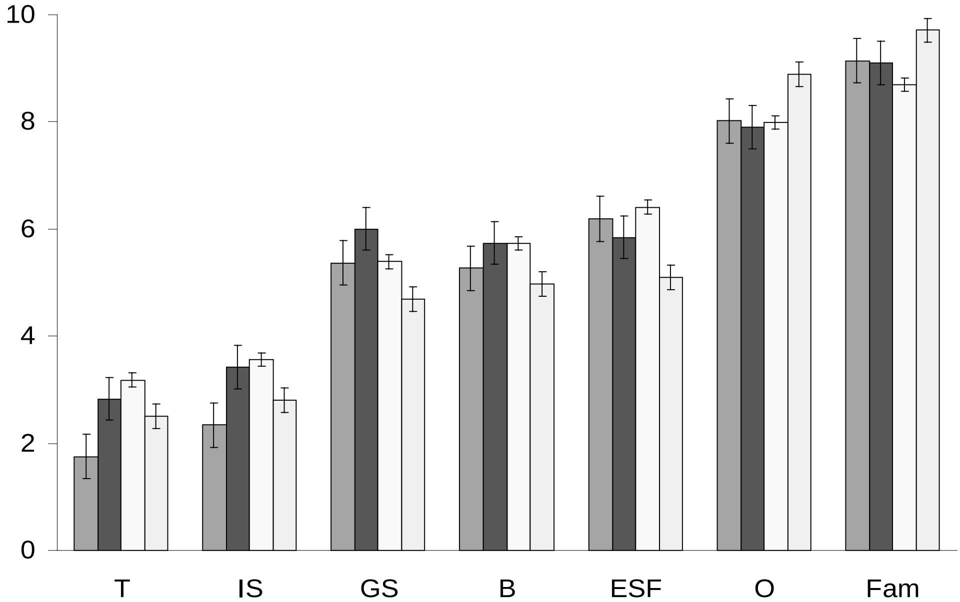

Figure 1. Mean welfare scores and their standard errors for the 7 main housing systems for pregnant sows (n = 11 experts). For every housing system the first column represents the mean expert score. The next three columns represent model A, B and C (see text) with which welfare scores are calculated. T: tether; IS: individual housing in stalls; GS: group housing with stalls; B: biofix; ESF: electronic sow feeder; O: outdoor; Fam: family pen.

perts); roughage versus no roughage (4 experts); and housing systems with different space allowances (4 experts). Other variants include GS systems with and without a social resting area; O systems with and without nose rings; and O systems on proper and improper soil types. On average 12.2 attributes were mentioned per expert to justify the scores (range 7–20). All experts mentioned space, straw, and feeding sequence. Other often mentioned attributes include social stability; the ability to separate functional areas (resting and dunging area; resting and feeding area); the provi-

Table 1. Mean scores (from 0 worst to 10 best) obtained from 11 pig experts for the 7 main housing systems for pregnant sows, and welfare scores as calculated by 3 models (see text).

	T	IS	GS	B	ESF	O	Fam	l.s.d.	F probability
Expert scores*	1.75[a]	2.34[a]	5.37[b]	5.27[b]	6.19[b]	8.02[c]	9.14[c]	1.182	<0.001
Model A	2.83[a]	3.42[a]	6.00[b]	5.74[b]	5.84[b]	7.90[c]	9.10[d]	1.127	<0.001
Model B*	3.18[a]	3.56[a]	5.39[b]	5.73[b]	6.41[c]	7.99[d]	8.69[e]	0.36	<0.001
Model C*	2.50[a]	2.81[a]	4.69[b]	4.97[b]	5.09[b]	8.89[c]	9.71[d]	0.645	<0.001

T: tether; IS: individual housing in stalls; GS: group housing with stalls; B: biofix; ESF: electronic sow feeder; O: outdoor; Fam: family pen. Means within a row lacking the same superscript character differ (P<0.05). * A significant expert effect was found by ANOVA ($P < 0.05$).

Table 2. Number of experts (n = 11) that, while explaining their personal welfare scores for the main housing systems for pregnant sows, used levels falling into any of the listed attribute classes (a); (b) as (a) but including only non-dichotomous level-ranges (see text); (c) as (a) but including only attributes which experts used to specify scores in addition to the scores for the main housing systems; HS: housing system.

Attribute class	Number of experts		
	(a) all properties	(b) ranges	(c) HS criterium
Space	11	9	4
Straw, substrate	11	2	8
Simultaneous feeding	11	1	0
Social stability and mixing	8	2	4
Separate functional areas	7	2	1
Roughage, bulk	7	1	4
Social resting	7	1	3
Hygiene & cleanliness	7	0	2
Agonism while feeding & individual feeding level	6	1	2
Group size	6	1	1
Space structure and agonism	5	1	0
Climate	5	0	0
Tether injuries	5	0	0

sion of roughage; hygiene and cleanliness; and social resting possibilities. With social resting is meant the ability for sows to rest with two or more together unrestricted by, for example, lying stalls. Several attributes were used in a non-dichotomous way, which means they took more than two (yes/no) levels. The most important one was space (9 experts). Other quantified attributes include straw quantity, separate functional areas and social resting.

Modelling resulted in three series of calculated welfare scores. These models were based on data from each individual expert, applied only within experts (model A); based on the integrated list of all attribute levels as mentioned by all experts (model B); and on a short version of this integrated list (model C). Both model B and model C generated different scores for different experts, because the experts had specified the exact descriptive properties of the housing systems somewhat differently. Model C includes the attributes pen space, substrate (straw), feeding sequence (simultaneous feeding), social stability, separation of functional areas and food quantity (roughage). In addition, the integrated list of model B also includes the proportion of time the animals can move around freely, social resting, group size, degree of individual feeding, food-intake protection, climate (separated into two: 'when hot' and 'when cold'), air quality, flooring aspects (separated into 'rest surface' and 'walking surface'), grooming facilities, the ability to avoid conspecifics and the degree of social variation (i.e. having different age groups in one pen as in Fam). Hygiene and cleanliness were incorporated in both climate and rest-surface attributes.

Whereas there was no significant difference between the mean expert scores for the O and Fam systems, in all three models the calculated scores for these housing sys-

Table 3. Comparison of expert scores with three welfare assessment models (ANOVA). Means within a row lacking a common superscript character differ (P<0.05).

	Expert scores	Model A	Model B	Model C	l.s.d.	F probability
Main housing systems, scores	0.000^{a}	0.356^{b}	0.376^{b}	0.0412^{a}	0.3071	0.021
All housing systems, scores	0.000^{ab}	0.352^{c}	0.209^{ae}	-0.303^{bf}	0.3180	0.001

tems differed significantly from each other. For model B a significant difference was also found between the calculated score for B (lower) and ESF (higher) (Table 1).

Figure 1 also shows that the three models tend to follow the expert scores for the 7 main housing systems. With statistical analysis the failure to find a significant difference between a model and the expert scores indicates statistical similarity between the model and the expert scores. When the expert scores are taken as a standard, failure to find a significant difference between a model and the expert scores, indicate that it is better than those models which do differ. When only the scores for the seven main housing systems were tested, model A and model B failed, and only model C, the shortened version of the integrated model, was not significantly different from the expert scores. However, when all scores were tested, i.e. including those added by the experts to specify variations of the main housing systems, then only model A failed. Model B and C did not fail, i.e. the scores calculated by these models did not differ significantly from the expert scores (Table 3).

Discussion

The list of main housing systems we obtained by asking experts is similar to the one mentioned by Broom (1989). In Table 4 we have given a tentative distribution of welfare scores over the main housing systems as the general picture emerging from our interviews. We chose a numerical scale (0–10) based on welfare ranks within attributes. For example, increasing the amount of space per pen was generally considered to be increasingly better for welfare. When pen space is similar for T and IS, for GS and B, and for ESF and Fam we get a welfare value distribution as shown in Table 4 in the first row. The overall scores are calculated as the average score (without giving any attribute level additional weight).

Svendsen & Svendsen (1997) produced a similar table, but with plusses and minuses and they did not include the O and Fam system. They also did not compute overall welfare scores. However, when their plusses and minuses are added overall scores are found which are similar to our findings (Table 1 and Figure 1). The scores calculated from Svendsen & Svendsen (1997) would be (in number of plusses): T = IS = 1.5; GS as cubicles = 4; B = 6.5; ESF = 5 plusses. They include other variations of the GS system (with a social resting area, deep straw or straw-bedded kennel) which get higher scores (up to 10.5 plusses). We also found GS, defined as group housing with stalls, to be the system with the widest range (expert scores 2–8). Included are systems with feeding stalls and a communal resting area, but also systems

Table 4. Tentative welfare distribution table. Within attributes (rows) a welfare value is assigned between 10 (best) and 0 (worst) on the basis of ranking the levels of the corresponding housing systems. Intermediate ranks get intermediate welfare values in a proportional way. Attribute levels may vary considerably within a given housing system, especially those marked with *. Attributes included in the table and rankings of attribute levels are tentatively based on the interviews with pig experts. Overall welfare is calculated as the average of the attribute scores. Mean subjective welfare scores of experts are also given. T: tether; IS: individual housing in stalls; GS: group housing with stalls; B: biofix; ESF: electronic sow feeder; O: outdoor; Fam: family pen.

Attribute	T	IS	GS	B	ESF	O	Fam
Space per pen	0	0	3.3	3.3	6.7*	10	6.7
Space per sow	0	0	3.3*	3.3	3.3	10	6.7
Straw, substrate	0	0	5*	0*	5*	10	10
Simultaneous feeding	10	10	10	10	0*	10	10
Social stability and mixing	5	5	0*	0	0	0	10
Separate functional areas	0	0	3.3*	3.3	6.7	10	10
Roughage, bulk	0	0	5*	0*	5*	10	10
Social resting	0	0	0*	10	10	10	10
Hygiene & cleanliness	0	0	5	10	10	5	10
Agonism while feeding & individual feeding level	10	10	10	0	10	0	10
Group size	0	0	10*	10	5*	10*	10
Climate (when cold)	10	10	5*	10	10	0	5
Climate (when hot)	0	0	5	5	5	10	5
Tether injuries	0	10	10	10	10	10	10
Overall score	2.5	3.2	5.4	5.4	6.2	7.5	8.8
Expert score	1.75[a]	2.34[a]	5.37[b]	5.27[b]	6.19[b]	8.02[c]	9.14[c]

with cubicles, i.e. feeding-lying stalls. GS systems with a substantial amount of space, substrate and a separate resting area are among the best rated housing systems. The Fam system in its revised form (e.g. Wechsler, 1996) also can be classified as such a system. However, this system has additional features including strictly separated functional areas for resting, activity/rooting and dunging, and a stable family structure of 4 sows, which are housed together with a boar during the breeding period and together with their offspring until these are about 5 months of age. Such features mimic the natural environment and are generally considered to be beneficial for welfare. Fam was included as a reference system. It received an average score of 9.14. The O system received the next highest score (8.02). The mean scores for the Fam and O system did not differ significantly. The one-sided Sign test also failed to identify the Fam system as ranking significantly better than the O system (P = 0.145). However, the O system was generally (but not by all experts) described as a system without nose rings, on pasture, with a proper soil type, sufficient and well-insulated arcs and a mud pool for wallowing. We suspect there may have been a tendency to take the better O farms as typical examples rather than the typical or 'average' ones.

The single most important finding of this paper may be that the experts, who were all pig-welfare scientists that have been involved in the design of welfare friendly housing systems, give significantly higher scores to group housing systems as compared to the individual, confined systems (T and IS). On our scale from 0 to 10 T and IS are clearly 'poor' welfare systems, whereas GS and B, and to a lesser extent also ESF, get scores close to the midline of the scale (5.5). This indicates that these systems are assessed as neither very poor nor very good welfare systems. However, we must warn against drawing overhasty conclusions based on the present findings. The scores obtained for the different housing systems are relative to each other and the scores contain no explicit reference to which scores would be acceptable. Furthermore, variation between farms within a certain type of housing systems and between individual animals may be much larger than between housing systems (e.g. Rushen & De Passillé, 1992; Signoret & Vieuille, 1996). In particular, the best individual housing systems are almost certainly better than the worst group housing systems. Our integrated model B also suggests that individual housing may reach similar scores as the group housing systems when non-social parameters are optimised. However, even though our models allow predicting the effects of changing such attribute levels on the overall welfare status, they have not been properly tested to make such predictions.

In this paper we have identified the most important attributes to assess the welfare status of pregnant sows in relation to housing, firstly by limiting the amount of time an expert had to justify his/her welfare scores, and secondly by counting the number of experts that use similar attributes. However, several caveats can be indentified.

Firstly, experts may have attached different weighting factors to different attributes.

Secondly, overlap exists between various items in the list of attribute classes (Table 2). For example, between straw and roughage and between space, group size and group stability. It is, therefore, not correct to conclude that straw (mentioned by all experts) is more important than roughage (mentioned by 7 experts, Table 2), because straw is also a form of roughage. However, it is not possible to construct a mutually exclusive list of attributes. For example, straw is also used for resting and rooting, and therefore cannot be lumped together with the roughage category.

Thirdly, the experts only evaluated typical examples of the main housing systems. This may have obstructed identification of certain attributes such as disease levels or stockmanship, which may vary especially within housing systems rather than between housing systems.

Fourthly, certain attributes seem to be missing. In accordance with Rushen and De Passillé (1992) our list contains both design criteria (environmental factors) and performance criteria (animal factors). Agonism was the most important performance criterion. Other animal factors such as stereotyped-behaviour levels were rarely mentioned, but do seem to be important. The limited time available for each interview (up to 2 hours) forced each expert to focus on only the most important attributes. Furthermore, 'missing' animal factors may be predictable from environmental factors that were specified by the experts. Examples are stereotyped-behaviour levels, which depend on feeding regime and a restricted space allowance (Terlouw *et al.*, 1991), and simultaneous feeding, which is an indicator of the ability to synchronise activities.

Fifthly, our results only identify opinions of a group of pig-welfare scientists. It remains to be shown that other scientifically trained professionals, such as veterinarians, have similar opinions and that the opinions we found are valid. This may be done by identifying their scientific basis and the consequences for the animals involved with respect to the whole range of needs the animals have (cf. Baxter & Baxter, 1984; De Koning, 1984; Bracke *et al*., 1997).

In principle the present findings support the development of a decision support system for welfare assessment in several ways:

Firstly, they create anchor points to evaluate system performance by indicating major input-output relationships, i.e. between the descriptions of the main housing systems for pregnant sows and their respective overall welfare score.

Secondly, our findings indicate that there may be a substantial degree of consensus in overall welfare assessment, at least among welfare scientists; more so than would be expected from the diversity of literature concerning welfare definitions and animal welfare assessment (e.g. Rushen & De Passillé, 1992; Mason & Mendl, 1993). Our experts all seem to choose similar attributes to assess welfare, which indicates that they were talking about the same thing. No fundamentally different perspectives on welfare seemed to be involved as, for example, in constructing the integrated model (B) no cases of logical inconsistencies were found between experts (with one expert ranking a certain attribute-level better than another attribute-level and another expert ranking in reversed order). Differences were found between experts: not all experts mentioned exactly the same properties. One example concerned the amount of space. One expert specified that the amount of space in the family pen was more than strictly necessary for optimal welfare. This expert did not rate the extra space as allocated in the O system as contributing to welfare. Several other experts, however, used the extra space in the O system as compared to the Fam system to explain having given a higher welfare score to the O system. However, we believe it would be a mistake to focus on such differences at the expense of the broad consensus that is evident from our findings (Table 1).

Thirdly, overall welfare assessment seems to allow itself to be modelled. Therefore, the development of a formalised procedure to assess welfare may be feasible. Quite simple modelling ignoring weighting factors and interactions already seems to capture major aspects of reasoning about welfare. More sophisticated models may be constructed, where these factors or factors concerning the hierarchy of needs are being taken into account.

In this paper we have presented a novel approach to the problem of overall welfare assessment. This method involves interviewing experts and employing an indirect way of collecting information about the importance of housing attributes by asking experts to rank and give scores for housing systems. Such an approach may help scientists to progress from pure subjectivity in overall welfare assessment through intersubjectivity toward the ideal of 'objectivity'. Respectively, these stages are characterised by personal welfare scores, subjectively agreed-upon group welfare-scores and welfare scores based on rational decision making based on all known facts including what is known scientifically. Inherent in the notion of overall welfare is the postulate that positive aspects of animal environments may compensate negative aspects to a certain degree. Such a quantified approach to welfare may therefore also help to get beyond cut-off point thinking (cf. Mendl (1991) for a critique of cut-off points) and its associated legislative objective in terms of means-prescriptions. Overall welfare assessment may eventually allow legislation to be formulated in terms of goal prescriptions in which minimum overall welfare standards can be defined generically and where individual farmers are allowed to reach such standards in farm specific ways.

Acknowledgements

This work was funded by the Technology Foundation (STW) and the Dutch Society for the Protection of Animals (NVBD).

References

Anonymous, 1993. Genstat 5 Release 3 Reference Manual, 1993. Oxford University Press, Oxford, 796 pp.

Anonymous, 1997a. The welfare of intensively kept pigs. Scientific Veterinary Committee, Brussels, 199 pp.

Anonymous, 1997b. Booklets of the Pig Welfare Advisory Group: no 1 Introduction of Sows into Groups. 5 pp.; no 2 Non-straw or Low Straw Systems for Housing Dry Sows. 4pp.; no 4 Cubicles and Free-Access Stalls. 8 pp.; no 5 Yards and Individual Feeders. 8 pp.; no 6 Yard or Kennels with Short Stall Feeders (Trickle Feeding or Wet Feed Systems). 8 pp.; no 7 Yards or Kennels with Floor Feeding. 10 pp.; no 8 Outdoor Sows. 8 pp.; no 9 Electronic Sow Feeders (ESF). 16 pp. MAFF Publications, London.

Backus, G.B.C. (Ed.), 1997. Comparison of four housing systems for non-lactating sows (In Dutch). Report P 1.171. Research Institute for Pig Husbandry, Rosmalen, 176 pp.

Bartussek, H., 1988. Haltung. In: Haiger, A., R. Storhas & H. Bartussek (Eds.), Natur-gemässe Viehwirtschaft. Verlag Ulmer, Stuttgart, pp. 147–248.

Bartussek, H., 1997. Animals Needs Index ANI (Tiergerechtheitsindex TGI-35L) for assessment of animals well being in housing systems for Austrian proprietary products and legislation. In: Arendonk, J.A.M. (Ed.), Book of Abstracts of the 48th Annual Meeting of the European Association for Animal Production. Wageningen Pers, Wageningen, pp. 190.

Baxter, M.R. & S.H. Baxter, 1984. How to delineate the sow's welfare needs. *Annales de Recherches Véterinaires* 15: 281–286.

Bracke, M.B.M., J.H.M. Metz & A.J. Udink Ten Cate, 1997. Assessment of animal welfare in husbandry systems. In: J.T. Sørensen (Ed.), Livestock Farming Systems – more than food production – Proceedings of the fourth International Symposium on Livestock Farming Systems, Wageningen Pers, Wageningen, pp. 231–237.

Bracke, M.B.M., J.H.M. Metz, A.A. Dijkhuizen & B.M. Spruijt. Decision support system for assessing farm animal welfare: strategy and prototype. Submitted to *Livestock Production Science*.

Broom, D.M., 1989. The assessment of sow welfare. *Pig Veterinary Journal* 22: 100–111.

Dawkins, M.S., 1997. D.M.G. Wood-Gush Memorial Lecture: Why has there not been more progress in animal welfare research? *Applied Animal Behaviour Science* 53: 59–73.

De Koning, R. 1984. An environmental model for sows: materializing physical and behavioural needs. *Annales de Recherches Véterinaires* 15: 303–307.

Fraser, D., 1995. Science, values and animal welfare: exploring the 'inextricable connection'. *Animal Welfare* 4: 103–117.

Mason, G. & M. Mendl, 1993. Why is there no simple way of measuring animal welfare? *Animal Welfare* 2: 301–319.

Mendl, M., 1991. Commentary. Some problems with the concept of a cut-off point for determining when an animal's welfare is at risk. *Applied Animal Behaviour Science* 31: 139–146.
Olsson, A.-Ch., M. Andersson, D. Rantzer, J. Svendsen & T. Hellström, 1986. Group housing of sows in gestation: comparison of computer-controlled individual feeding system with a group feeding system based on biological fixation (In Swedish with tables and summary in English). Report 51. Swedish Univiversity of Agricultural Sciences, Department of Farm Buildings, Lund, 221 pp.
Rushen, J. & A.M.B. De Passillé, 1992. The scientific assessment of the impact of housing on animal welfare: a critical review. *Canadian Journal of Animal Science* 72: 721–743.
Signoret, J. & C. Vieuille, 1996. Effectiveness and limitations of physiological versus ethological criteria to assess the welfare of pigs in relation to the housing system. *Pig News and Information* 17: 115–121.
Stolba, A., 1981. A family pen system in enriching pens as a novel method of pig housing. Alternatives to intensive husbandry systems. In: Anonymous (Eds.), Proceedings, Universities Federation for Animal Welfare, Potters Bar, pp. 52–67.
Svendsen, J. & L.S. Svendsen, 1997. Intensive (commercial) systems for breeding sows and piglets to weaning, *Livestock Production Science* 49: 165–179.
Sundrum, A., 1997. Assessing livestock housing conditions in terms of animal welfare – possibilities and limitations. In: Sørensen, J.T. (Ed.), Livestock Farming Systems – more than Food Production. Proceedings of the fourth International Symposium on Livestock Farming Systems, Wageningen Pers, Wageningen, pp. 238–246.
Sundrum, A., R. Andersson & G. Postler, 1994. Der Tiergerechtheitsindex-200/1994 – ein Leitfaden zur Beurteilung von Haltungssystemen für Rinder, Kälber, Legehennen und Schweine. Verlag Köllen, Bonn, 211 pp.
Terlouw, E.C.M., A.B. Lawrence & A.W. Illius, 1991. Influences of feeding level and physical restriction on development of stereotypies in sows. *Animal Behaviour* 42: 981–991.
Wechsler, B., 1996. Rearing pigs in species-specific family groups. *Animal Welfare* 5: 25–35.