

Full Length Research Paper

The application of animal welfare standards in intensive production systems using the assessment protocols of Welfare Quality®: Fattening pig husbandry in Northwest Germany

D. Otten*, E. Annas and H. F. A. Van den Weghe

Department of Animal Sciences, Division: Process Engineering, Georg-August University of Goettingen, Universitaetsstr, 7, D-49377 Vechta, Germany.

Accepted 26 March, 2013

The increased requirements for animal welfare have raised the need for a comprehensive on-farm assessment system. This paper is a first step to analysing the reliability and feasibility of the on-farm welfare assessment with regard to the animal-related measures of Welfare Quality® in intensive fattening pig husbandries. Based on the 2009 Welfare Quality® assessment protocols for pigs, six analyses were undertaken by one observer on three farms. It became apparent that the system in essence, fulfils the requirements of a sound assessment of animal welfare under intensive production with low on-farm variability. The behaviour-based measurements had a higher degree of within-farm variability than clinical- and resource-based measurements as the assessment involves a greater degree of subjectivity. Some measurements seemed to be of low sensitivities as there were no or very low variations in many of the indicators being assessed. Despite this, this preliminary study indicates that the assessment system is a reliable and feasible tool for the evaluation of animal welfare status in intensive pig production for fattening pig.

Key words: Welfare Quality®, animal welfare, intensive livestock production.

INTRODUCTION

The welfare of animals used for food production has increasingly become an area of interest at all levels of the value-added chain (Blokhuis et al., 2008). The reasons are due to the changes in domestic animal husbandry, which has become more and more specialised and intensified (Temple et al., 2011a; Aparicio Tovar and Vargas Giraldo, 2006; Hughes and Duncan, 1988). Additionally, consumers increasingly demand animal-welfare-friendly products (Tuytens et al., 2010; Ellis et al., 2009; Carlsson et al., 2007; Harper and Makatouni, 2002; Velde et al., 2002; McGlone, 2001). The food

industry has reacted to this situation and is at present discussing the introduction of numerous different labels to guarantee high standards of animal welfare.

Although, a large body of approaches has been developed, none of them could be established in practice so far. Against this background, the Welfare Quality® system was developed within the European Union's Sixth Framework Programme on Food Quality and Safety (2006 to 2010). The project involved a total of 44 institutions based in 13 European and 4 Latin American countries. Its aim was to develop reliable standardised

*Corresponding author. Email: dennis.otten@agr.uni-goettingen.de. Tel: 0172 411 66 27. Fax: +49 (0) 4441 - 15 448.

methods for the assessment of animal welfare at farm level (Welfare Quality®, 2009).

A first evaluation with a prototype of the Welfare Quality® system for sows and piglets was undertaken by Scott et al. (2009) on 82 pig farms in the UK and the Netherlands, encompassing a wide variety of farming systems. The analysis showed that the incidence of clinical welfare problems as indicated by the system was generally low. The main criticisms involved stereotypical behaviour patterns. Knierim and Winckler (2009) provide a review about the validity, reliability and feasibility of the scoring system. These authors also discussed future perspectives of using Welfare Quality® evaluations by looking at the welfare of cattle.

The first step towards the evaluation of Welfare Quality® assessment protocols on intensive pig farming systems was done by Temple et al. (2011a, b) on 30 conventional growing pig farms in Spain. Their results showed that the measures presented very little variation to differentiate farms using intensive production and could just be used to identify poor welfare levels under such conditions. In addition, the state of welfare at the slaughterhouse was analysed by an overview of the sensitivity and feasibility of the Welfare Quality® protocol for finishing pigs in ten Spanish slaughterhouses (Dalmau et al., 2009).

But so far, there exist no studies about the reliability of the Welfare Quality® assessment in intensive fattening pig farming, an important requirement to establish an assessment system in practice. The present study takes this into account to evaluate reliability within farms (test-retest reliability). Consistency of welfare assessment requires further attention in the future, particularly if farms are to be certified on the welfare status, based on infrequent recorded measurements (Knierim and Winkler, 2009).

Generally, reliability issues of animal welfare assessments have often been neglected so far and require more thorough investigation and discussion. A great absence of studies about reliability measures for animal welfare indicators is remarkable (Knierim and Winkler, 2009). In literature, consistent statistical tests for reliability are rare. While rare unequivocal scientific methods or criteria for range „good“ or „acceptable“ reliability have been established so far, some opinions are given in the literature (Knierim and Winkler, 2009). Often, the correlation coefficient is used. This was not possible in this preliminary study because of the sample size.

Therefore, another widely used analysis method (especially at preliminary studies) was chosen, the coefficient of variation (CV). It expresses the experimental error as percentage of the mean and is a very good index of the reliability (Gomez and Gomez, 1984). However, the rate of acceptable CV varies greatly with type of experience (Patel et al., 2001). Gomez and Gomez (1984) opined an acceptable CV from 8 to 15%.

In the present study, the authors chose a CV of 10% an acceptable range for a reliable assessment.

Beside the objectivities of reliability, the study discusses the feasibility of the assessment on intensive fattening production for future potential of implementation in practice. To be feasible, an assessment system should be relatively easy to perform and require little input from the farmers (Temple et al., 2011a). Also, time constraints are a main concern of an assessment system considering feasibility (Knierim and Winkler, 2009). The actual time needed for an adequate assessment of a farm is difficult to gauge (Temple et al., 2011a). Therefore, Knierim and Winkler (2009) discuss that welfare status on a farm should be carried out during a one-day visit.

MATERIALS AND METHODS

Experimental design

Three intensive fattening pig farms (described in the following as Farms A, B and C) were each analysed by six repeats (08.12.2011 to 24.12.2011 at weekly intervals (9:00 a.m. to 14:00 a.m.)) using the Welfare Quality® assessment protocols for pigs (Welfare Quality®, 2009). All analyses were done by the same assessor, who had absolved an official training course at the Welfare Quality® consortium to ensure a correct application of the Welfare Quality® protocols. The farms were situated in Northwest Germany and represented the typical production system for fattening pigs in this intensive livestock region with respect to herd size, equipment and state of technology. The farm size was 1700 to 2500 fattening (genetic hybrid) pigs. The pig houses on all three farms were insulated, had mechanical ventilation and fully-slatted concrete floors. Neither sows nor piglets were kept on the farms.

Three different feeding regimes were implemented. The pigs were fed by manual liquid feeders in troughs (limit fed, 4 x per day; Farm A), by automatic liquid feeders (*ad libitum*; Farm B) or by automatic sensor-controlled liquid feeders in troughs (*ad libitum*; Farm C) with a feeding place / pig ratio of 1:1 (Farms A and C) or 1:6 (Farm B) (Table 1).

Assessment of growing pigs by Welfare Quality®

The core element of the Welfare Quality® scoring system is an animal-based assessment followed by an evaluation of certain resource- and management-based measures. The final evaluation is comprised of four principles: „Good feeding“, „Good housing“, „Good health“ and „Appropriate behaviour“. These four principles are based on twelve criteria which are calculated from various indicators. Table 2 shows the three-stage structure of the Welfare Quality® system (Welfare Quality®, 2009; Botreau et al., 2007) (Table 2).

The evaluation was done exactly according to the 2009 Welfare Quality® assessment protocol for pigs. At the start of each assessment, the farmer was interviewed about general information concerning the management of feeding and hygiene, the records of production and mortality, the regulation of ambient temperature, castration and tail docking routines, and the use of anaesthetics and the prevention of disease.

At the beginning of each investigation, a sketch of the husbandry was made involving each pen individual. Then, the number of required pens was selected randomly and arbitrarily from the sketch for ensuring a representative random sample. Therefore at every

Table 1. Description of the three pig fattening farms.

Characteristics	Farm		
	A	B	C
No. of animals	1700	1700	2500
Breed	PIC sows x Pietrain	BHZP ¹⁾ sows x Pietrain	TOPIGS 20
Animals per pen	12	30	12
Dimensions of pen	4.00 x 2.60 m	4.00 x 4.58 m	4.00 x 2.35 m
Floor space / pig	0.87 m ²	0.61 m ²	0.78 m ²
Pen design	Insulated stable, fully-slatted floor, mechanical ventilation, exhaust system; (inlet air through slots)	Insulated stable, fully-slatted floor, mechanical ventilation, exhaust system, (inlet air through slots)	Insulated stable, fully-slatted floor, mechanical ventilation, exhaust system; (inlet air through pores in the ceiling)
Feeding system	Automatic liquid feeder; (through across)	Wet feeder; (feeding pan)	Sensor-controlled liquid feeding; (through across)
Pig, feeding place ratio	1:1	1:6	1:1
Waterer per pen (without feeder)	1	1	1
Type of waterer	Conventional nipple drinker (without drip cup)		
Body weight	Starting weight 30 kg – finishing weight 120 kg		

1) Bundes Hybrid Zucht Programm [German federal hybrid breeding program].

Table 2. Structure of the three steps of the Welfare Quality® assessment system.

Step 3	Step 2	Step 1
Welfare principles	Welfare criteria	Indicators
Good feeding	1. Absence of prolonged hunger	Body condition score
	2. Absence of prolonged thirst	Water supply
Good housing	3. Comfort around resting	Bursitis, absence of manure on the body
	4. Thermal comfort	Shivering, panting, huddling
	5. Ease of movement	Space allowance
Good health	6. Absence of injuries	Lameness, wounds on body, tail biting
	7. Absence of disease	Mortality, coughing, sneezing, pumping, twisted snouts, rectal prolapse, scouring, skin condition, ruptures and hernias
	8. No painful managerial procedures	Castration, tail docking
Appropriate behaviour	9. Expression of social behaviour	Social behaviour
	10. Expression of other behaviours	Exploratory behaviour
	11. Good human-animal relationship	Fear of humans
	12. Positive emotional state	Qualitative behaviour assessment (QBA)

Source: (Welfare Quality® 2009).

farm visit, the average sample ages was of the animals in husbandry were analysed. The sample size of the observed measures was: QBA = 6 observation points x 25 animals (20 min.); coughing and sneezing 6 observation points x 25 animals; social behaviour and exploratory behaviour = 3 observation point x 50 animals; huddling, shivering, panting, fear of humans, bursitis, absence of manure on body, tail bitten, lameness, pumping, twisted

snouts, rectal prolapse, scouring, skin condition, ruptures and hernias = 15 observation points x 10 animals.

Thereafter, the observations started with an assessment of the principle “Appropriate behaviour” (Table 2). “Appropriate behaviour” was evaluated by the animals’ social behaviour, exploratory behaviour, the fear of humans and a qualitative behaviour assessment (QBA). In the QBA, the emotional status of an animal

was assessed by discerning the intensity of the occurrence of ten positive and ten negative behaviour patterns within a 20 min period: positive - active, calm, content, enjoying, happy, lively, playful, positively occupied, relaxed, sociable; negative - agitated, aimless, bored, distressed, fearful, frustrated, indifferent, irritable, listless, tense.

To undertake the QBA, the assessor entered the room and ensured that all the pigs get up. After waiting five minutes, the observer started the assessment from outside the pen in the run passage; they have to be in a partly active state to show behavior for assessment. The pigs were scaled (0 to 120) on the basis of the number of pigs showing the behaviour pattern and the intensity of the behaviour. To evaluate their social and exploratory behaviour, the pigs were scored as to whether they were active or inactive by scan samplings (five scan samples made at two-minute intervals). The active ones were scaled in positive social behaviour (sniffing, nosing, licking, moving gently away from another animal), negative behaviour (aggressive behaviour or social behaviour as a response from a disturbed animal), exploratory behaviour or other behaviours (not classified). The pigs' exploratory behaviour was also divided into pen behaviour (sniffing, nosing, licking part of pen) or other (behaviour patterns not included above) (Temple et al., 2011a; Welfare Quality®, 2009).

The assessment of the social and exploratory behaviour should be applied at three different ages of the fattening period if possible: at the beginning but at least one week after being mixed (before the establishment of a social hierarchy), in the middle of the fattening period (around 70 kg live weight), and at the end of the fattening period (Welfare Quality®, 2009).

The indicator for human-animal relationship (HAR) based on the fear of humans test, whereby the reaction of the pigs to the farmer entering the pen is analysed (the farmer goes very slowly along the passage and waits there for 30 s). For the test, 10 pens were analysed on every visit to each farm. Each pen was analysed as a whole. In the HAR, two reactions are possible: 0 = no panic present; 2 = more than 60% showing panicking behaviour.

The animal-based measures to evaluate "Good feeding", "Good housing" and "Good health" followed the "Appropriate behaviour" assessment. Table 3 shows the indicator assessments and their scoring (Welfare Quality®, 2009) (Table 3).

Most measurements were scored according to a three-point scale (0 to 2): 0 = welfare is good; 1 = welfare is acceptable (compromises made); 2 = poor and unacceptable welfare. In some cases, just the presence or absence of a particular behaviour was scored: 0 (present) or 2 (absent).

The data evaluation of the analysed indicators and the calculation of the algorithm were done by the Welfare Quality® consortium (National Institute for Agricultural Research INRA, France). Here, the overall evaluation (range of scores = 0 - 100) of a farm is given one of four values excellent, enhanced, acceptable or not classified. In the overall evaluation, the individual criteria within a particular principle do not compensate for each other (that is, a high score in one will not compensate for a low score in another). A farm is considered to be excellent if it scores more than 55 on all principles and more than 80 on two of them. It is considered to be enhanced if it scores more than 20 on all principles and more than 55 on two of them. An acceptable level of animal welfare score will be obtained by scores more than 10 on all principles and more than 20 on three of them. If a farm does not reach this minimum standard it will not be classified (Welfare Quality®, 2009).

Statistical analysis

The statistical analysis of the welfare data by the Welfare Quality® consortium was carried out with the software program SPSS,

Version 19 (PASW Statistics - SPSS 19. for Windows). The coefficients of variation and the upper and lower confidence intervals (95%) of each respective farm's observations were calculated to evaluate the reproducibility and variability of the welfare assessment. Additionally, an analysis of variance (Tukey's RM) was performed to analyse the differences between the farms. A value of $P \leq 0.05$ was considered as statistically significant. A log transformation was done before calculation.

RESULTS

In the final evaluation according to the Welfare Quality® system Farms A and B achieved an overall evaluation of "enhanced" at every evaluation. Farm C achieved this only on the first one and thereafter would be classed an "acceptable" range. Table 6 shows the rated values of the criteria, principles and the overall evaluations and the reliability of the assessments within the three farms. The comparison of means demonstrate significant differences between the farms (Table 4).

In the overall assessment, there were no differences in the evaluations of Farm A and B; Farm C was found to be significantly worse. The differences between farms occurred in the criteria absence of prolonged hunger, absence of prolonged thirst, ease of movement and expression of social behaviour.

The principle of "Good feeding" was assessed as having the worst principle of all. Especially deficits of the criterion absence of prolonged thirst had a significant impact. Criticism was not only the number of waterers (Farms A and C: 12 pigs / waterer; Farm B: 30 pigs/waterer) but also their functionality and cleanliness. Especially Farm B was significantly worse in this aspect. Farm A had a worse evaluated absence of prolonged hunger because there were worse body conditions scored. Farm B had a high variation because of the measured absence of prolonged thirst. Farm A and C did not have any variation in this aspect.

The middle range of the principle Good housing was scored due to evaluation of the criteria comfort of resting (presence of indicators manure on body and bursitis) and ease of movement. Especially Farm C had the significant worse criterion easy of movement due to the average minimum space available per animal at the start of the fattening period $2.50 \text{ m}^2 / 100 \text{ kg}$ falling to only $0.62 \text{ m}^2 / 100 \text{ kg}$ at the end. Therefore, a high variation became apparent because of changing number of pigs per pen.

The best value in the Welfare Quality® assessment was achieved on the principle of "Good health" by the farms. The criterion thermal comfort and the criterion absence of pain by management procedures did not have any variation at all. The assessment of absence of pain induced by management procedures gave a uniform value due to castration and tail docking being undertaken on all three farms. The use of local anaesthetic during these procedures was, however, considered as being a positive factor.

Table 3. Measurements and evaluation scores for animal welfare in the Welfare Quality® assessment protocols for pigs.

Measures	Method	Evaluation
Good feeding		
Body condition score	The spine, hip and pin bones are visually inspected	0 Good body condition
		2 Visible spine, hip and pin bones
Water supply	Number of drinking places, their functionality and cleanliness are investigated	0 Function correctly / clean
		2 Do not function properly / dirty
Good housing		
Bursitis	Individually scored: small 1.5-2.0 cm; large 3.0-5.0 cm; extremely large 5.0-7.0 cm	0 No evidence of bursa / swelling
		1 One or several small bursa on the same leg or one large bursa
		2 Several large bursa on the same leg, or extremely large bursa or any bursas that are eroded
Absence of manure on the body	Assessed on one side of the pig if the body surface is soiled	0 Less than 20% of body is soiled
		1 More than 20% but less than 50%
		2 More than 50%
Shivering	Observed in resting animals; defined as the slow and irregular vibration of body part, or of the body in whole	0 No pigs shivering
		1 Up to 20% shivering
		2 More than 20% shivering
Panting	Observed in resting animals; defined as breathing rapidly in short gasps and carrying out with the mouth	0 No pigs Panting
		1 Up to 20% Panting
		2 More than 20% Panting
Huddling	Observed in resting animals; defined: pig is lying with more than half of its body in contact with another pig, except when just lying side by side and alongside another animal.	0 No pigs huddling
		1 Up to 20% huddling
		2 More than 20% huddling
Space allowance	Information from farmer about average weight of pigs. Assess size and no. of animals per pen	Space allowance expresses in m ² / 100 kg animal

Table 3. Contd.

Good health		
Lameness	Observation of pigs' locomotion. Pigs should walk a certain distance before assessment.	0 Normal gait or slight difficulty but using all 4 legs 1 Severely lame, minimum weight-bearing on the affected limb 2 No weight-bearing on the affected limb, or not able to walk
Wounds on body	Assessed visually. The body is considered in five separate regions: Ears, Front (head to back shoulder), Middle (back of shoulders), Hind-quarters, Legs. Distance of approximately 0.5 m one side of animals should be assessed	1 Ears 2 Front (head to back of shoulder) 3 Middle (back of shoulder to hindquarters) 4 Hindquarters 5 Legs (from the accessory digit upwards)
Tail biting	All pigs must be standing up. The assessor must assess according to the following scale:	0 No evidence of tail biting or superficial biting along the length of the tail, but no fresh blood or any swelling 2 Bleeding tail and / or swollen infected tail lesion and / or part of tail tissue missing and presence of scabs
Mortality	Information from farmer about mortality	Percentage mortality during the previous 12 months
Coughing	Assessed during a 5-min period after pigs have stood up	Average frequency of coughing per animal per 5 minutes
Sneezing	Assessed during a 5-min period after pigs have stood up	Average frequency of sneezing per animal per 5 minutes
Pumping	Score number of pigs showing heavy or laboured breathing (when it is heavy and labored and it is easy to see chest rising and falling with each breath)	0 Percentage with no evidence of pumping 2 Percentage with evidence of pumping
Twisted snouts	Score number of pigs showing heavy or laboured breathing	0 Percentage with no evidence twisted 2 Percentage with evidence twisted
Rectal prolapse	Score number of pigs showing heavy or laboured rectal prolapsed. It is heavy and labored and it is easy to see chest rising and falling with each breath.	0 Percentage with no evidence prolapse 2 Percentage with evidence prolapse
Scouring	Identify areas in pen where dung is visible and record number of animals in the pen:	0 No liquid manure visible 1 Some liquid manure visible 2 All visible faeces are liquid

Table 3. Contd.

Skin condition	Assess one side of body and consider the total area affected in relation to the rest of body. Certain diseases can cause characteristic inflammation or discoloration of the skin.	0	No evidence of skin inflammation or discoloration
		1	0-10% of skin is inflamed or discoloured or spotted
		2	More than 10% of the skin has an abnormal colour or texture
Ruptures and hernias	Animals are observed from the front, back and sides	0	No hernias / ruptures
		1	Hernias or ruptures present, but the affected area not bleeding, not touching the floor and not affecting locomotion
		2	Bleeding lesions, hernias / ruptures and they are touching the floor
Castration	Ask farmer about castration technique	0	No castration done
		1	Castration with use of anesthetics
		2	Castration without use of anesthetics
Tail docking	Ask farmer about technique of tail docking	0	No tail docking done
		1	Tail docking with use of anesthetics
		2	Tail docking without use of anesthetics

With the principle “Appropriate behaviour” only the criterion expression of social behaviour had a positive result. The criterion expression of other behaviours was negative due to the low investigative behaviour shown by the pigs in their pens. The evaluation of the “Appropriate behaviour” also had the highest coefficient of variation. Almost all measurement had a high variation. The results of the QBA are shown in Table 5. On average, the means of positive behaviours were found significantly more commonly than the expression of negative behaviours on all three farms (Farm A: P-value 0.009; Farm B = P-value 0.003; Farm C: P-value = 0.039) (N = 10). There was a strikingly low number of pigs showing the behaviour patterns positively occupied, playful and enjoying. With respect to the negative behaviour patterns, the

raised values for *tense* and *agitated* indicated that there was a degree of unrest present in the pigs. The variation of the QBA is significantly higher than with the other criteria at Farm A (P-value = 0.0454) and Farm C (P-value = 0.0029); (Farm B; P-value = 0.0735) (Table 5).

Feasibility and practicability

The observation takes about 250 min time (QBA = 25.0 min, social behaviour and respiratory behaviour = 75.0 min, assessments in pen = 150.0 min). In addition, variations in time were caused by the necessary conversations with the farmer for the analysis and the discussion of the results. Some assessments have to be carried out at certain times. It was very important that the

assessment was not carried out during feeding time, the changing of pigs between pens or when any treatments are given because these can influence the results. The Welfare Quality’s® guidelines also point this out.

Sometimes, the natural curiosity of the pigs complicated the observations making the evaluation difficult. Poor visibility due to the lighting conditions on one Farm (A) in addition to the large number of pigs in the pens also made observation problematic. Another problem became apparent: Sick or injured animals are often taken out of pens and placed in a hospital-pen so that they can no longer be matched back to their original pens and therefore can no longer be used for the assessment. Otherwise, the application of the assessment protocols was easy to perform.

Table 4. The variation in the criteria, principles and the overall assessment of the six analyses on each farm (coefficient of variation and confidence interval 95%) and the differences between the farms (comparison of means). Different letters^{a,b,c} following the means define significant differences between farms.

Overall assessment	Farm A			Farm B			Farm C			P-Values
	CV	Confidence interval (95%)	Mean	CV	Confidence interval (95%)	Mean	CV	Confidence interval (95%)	Mean	
	0.000	2.0-2.0	2.0 ^a	0.000	2.0-2.0	2.0 ^a	0.144	2.4-3.3	2.8 ^b	0.000
Good feeding	0.004	42.7-43.0	42.3 ^a	0.311	22.6-44.4	33.5 ^b	0.000	43.0-43.0	43.0 ^a	0.023
Absence of prolonged hunger	0.034	93.5-100.5	97.0 ^a	0.000	100.0-100.0	100.0 ^b	0.000	100.0-100.0	100.0 ^b	0.022
Absence of prolonged thirst	0.000	40.0-40.0	40.0 ^a	0.365	18.5-41.5	30.0 ^b	0.000	40.0-40.0	40.0 ^a	0.022
Good housing	0.057	54.8-61.7	58.2 ^a	0.071	53.2-61.7	57.4 ^a	0.261	32.1-56.4	44.2 ^b	0.007
Comfort around resting	0.078	52.3-61.6	57.0 ^a	0.052	57.5-64.0	60.8 ^a	0.089	51.1-61.6	56.4 ^a	0.188
Thermal comfort	0.000	100.0-100.0	100.0 ^a	0.000	100.0-100.0	100.0 ^a	0.000	100.0-100.0	100.0 ^a	---
Ease of movement	0.191	52.6-1.6	65.8 ^a	0.097	50.9-62.4	56.7 ^a	0.353	25.2-54.8	40.0 ^b	---
Good health	0.031	63.1-67.3	65.2 ^a	0.056	58.2-65.4	61.8 ^a	0.037	59.8-64.6	62.2 ^a	0.086
Absence of injures	0.061	82.8-94.2	88.5 ^a	0.075	80.8-94.7	87.8 ^a	0.035	89.1-95.8	92.5 ^a	0.279
Absence of disease	0.087	86.0-103.3	88.5 ^a	0.191	65.6-98.6	87.8 ^a	0.125	71.0-92.4	82.5 ^a	0.132
Absence of pain induced by management practice	0.000	53.0-53.0	53.0 ^a	0.000	53.0-53.0	53.0 ^a	0.000	53.0-53.0	53.0 ^a	---
Appropriate behaviour	0.089	42.7-58.9	50.8 ^a	0.110	37.6-47.4	42.5 ^a	0.143	39.1-52.7	45.9 ^a	0.110
Expression of social behaviour	0.188	58.0-86.5	72.3 ^a	0.189	59.8-89.5	74.7 ^a	0.082	70.9-84.3	77.6 ^a	0.748
Expression of other behaviours	0.045	49.1-54.0	51.6 ^a	0.181	27.9-41.0	34.5 ^b	0.248	27.4-46.7	37.0 ^{bc}	0.001
Good human-animal relationship	0.312	22.7-44.8	33.7 ^a	0.461	22.9-65.9	44.4 ^a	0.268	32.4-57.1	44.6 ^a	0.380
Positive emotional state	0.155	51.5-71.5	61.5 ^a	0.181	39.6-58.0	48.8 ^a	0.311	36.5-71.8	54.2 ^a	0.232

DISCUSSION

Because of the increasing demands regarding on-farm welfare assessment the Welfare Quality® protocols were developed. The study was a first step to evaluate the reliability and feasibility of the assessment protocols for fattening pigs in intensive production systems.

The Welfare Quality® protocols evaluated general criticism was an inadequate water supply. Furthermore, the indicator body condition score was found to be critical in the evaluation of the principle "Good feeding". Signs of malnutrition or

dehydration were not visible on the analysed farms. Certain husbandry mistakes also became apparent by the presence of bursitis and manure on body. These deficits in the welfare criterion ease of movement were due to the concrete flooring and stocking density used in the intensive production conditions on the farms (Mouttoutu et al., 1999). This criterion is therefore, a sensitive and important indicator of animal welfare in intensive production systems (Waiblinger et al., 2001). Scott et al. (2009) found in their study that the clinical welfare problems in sow and piglet production were rather low.

However, the main criticisms of these authors relate to the presence of stereotypical behaviour patterns. The present study confirms their results with the general deficits in the principle "Appropriate behaviour". The consideration of the animals' behaviour has gained imperative that in intensive production systems, farmers pay greater attention to the species-specific behaviour of their animals to ensure animal-friendly husbandry.

A reliability assessment is often not easy to achieve in the Welfare Quality® evaluation (Temple et al., 2011a; Knierim and Winkler, 2009). In the present investigation, the coefficient

Table 5. The variation in the qualitative behaviour assessment (QBA) of the six analyses on each farm (coefficient of variation and confidence interval 95%). Different letters^{a,b,c} following the means define significant differences between farms.

Terms	Farm A			Farm B			Farm C			P-Values
	Mean	CV	Confidence interval (95%)	Mean	CV	Confidence interval (95%)	Mean	CV	Confidence interval (95%)	
Positive										
Active	101.7 ^a	0.149	101.7-97.8	33.5 ^b	0.331	33.6-36.4	42.0 ^b	0.516	36.4-47.6	0.002
Relaxed	12.4 ^a	0.409	11.1-13.7	86.3 ^b	0.074	84.7-88.0	81.0 ^b	0.227	76.3-85.7	0.000
Calm	45.1 ^a	0.527	38.9-51.2	70.8 ^b	0.306	65.2-76.4	86.7 ^c	0.194	82.7-91.0	0.047
Content	70.4 ^a	0.151	67.6-73.1	59.0 ^a	0.278	54.8-63.2	68.8 ^a	0.557	58.9-78.7	0.666
Enjoying	68.7 ^a	0.242	64.4-73.0	27.8 ^b	0.481	24.4-31.3	19.5 ^c	0.729	15.8-23.2	0.004
Sociable	56.0 ^a	0.226	52.5-59.4	51.0 ^a	0.334	48.2-53.8	79.0 ^a	0.342	70.4-87.6	0.199
Playful	74.5 ^a	0.077	73.0-76.0	28.7 ^b	0.347	26.1-31.2	20.2 ^c	0.613	17.0-23.4	0.002
Positively occupied	55.8 ^a	0.200	52.9-58.7	29.3 ^b	0.298	27.1-31.6	30.0 ^c	0.848	23.4-36.6	0.042
Lively	84.3 ^a	0.150	81.0-87.5	52.0 ^b	0.165	49.8-54.2	59.5 ^b	0.457	52.5-66.5	0.020
Happy	64.7 ^a	0.226	60.9-68.5	44.5 ^a	0.334	40.7-48.3	54.7 ^a	0.342	49.8-59.5	0.158
Negative										
Fearful	6.1 ^a	0.507	5.3-6.4	2.5 ^b	0.308	2.3-2.6	1.7 ^c	0.451	1.5-1.9	0.004
Agitated	23.9 ^a	0.294	22.0-25.7	8.3 ^b	1.031	6.1-10.6	6.8 ^c	1.794	3.7-10.0	0.004
Tense	26.7 ^a	0.331	24.4-29.0	10.3 ^b	0.827	8.1-12.5	6.5 ^c	1.364	4.2-8.8	0.004
Frustrated	3.1 ^a	0.510	2.6-3.4	2.7 ^a	0.605	2.3-3.1	6.3 ^a	1.194	4.3-8.2	0.421
Bored	7.4 ^a	1.058	5.4-9.4	13.0 ^a	0.675	10.7-15.3	12.5 ^a	0.935	9.5-15.5	0.298
Listless	5.0 ^a	0.635	4.1-5.7	4.2 ^a	0.381	3.8-4.6	2.3 ^b	0.322	2.1-2.5	0.046
Indifferent	4.3 ^a	0.636	3.6-4.9	11.7 ^a	1.228	8.0-15.4	11.2 ^a	1.037	8.2-14.2	0.787
Irritable	3.9 ^a	0.345	3.5-4.2	6.5 ^a	0.960	4.9-8.1	7.5 ^a	1.360	4.9-10.1	0.952
Aimless	7.3 ^a	1.139	5.1-9.4	7.5 ^a	0.972	5.6-9.4	10.8 ^a	1.152	7.6-14.1	0.882
Distressed	6.6 ^a	1.150	4.6-8.5	3.3 ^b	0.285	3.1-3.6	4.5 ^b	0.885	3.5-5.5	0.450

of variation showed that the reliability of the assessment for fattening pigs over the six repeats is to be rated rather acceptable. The assessment of “Appropriate behaviour” had a higher variation than the other indicators as the assessment involves a greater degree of subjectivity (Temple

et al., 2011b). Nevertheless, even when the “Appropriate Behaviour” in the Welfare Quality® assessment is subject to a certain degree of subjectivity (Temple et al., 2011b), it still fulfils the need to include both psychological and physiological parameters in the welfare

assessment (Duncan and Petherick, 1991). Especially intensive husbandry systems are often criticised with respect to species-specific behaviour (Van de Weerd and Day, 2009). Besides reliability of the “Appropriate behaviours”, the problem of sporadic behaviours exists and

cannot reliably be assessed during a short time of observation (Knierim and Winkler, 2009).

Temple et al. (2011a) reported a low rate (<2%) in the occurrence of the indicators panting, pumping, shivering, huddling, wounds on the body, tail biting, lameness, hernia, and scouring in growing pigs kept under intensive conditions. In the present study the minimal observation of these measures also became apparent. Furthermore, the indicators skin condition, twisted snouts and rectal prolapse were not indicated at all. In addition, the indicators castration and tail docking and the criteria absence of pain by management procedures and thermal comfort showed no variation in the observations. This lack of variation could partly be due to the small scales used in the Welfare Quality® scoring system.

The majority of the measurements were scored according to a three-point scale (0 to 2) or just grouped according to the presence or absence of an indicator. The use of a narrow scale means that the reliability of the Welfare Quality® is increased and that different observers will reach the same results (Knierim and Winkler, 2009). However, with intensive husbandry, the narrow scale might lead to a low degree of sensibility of the assessment by some indicators. Further studies are necessary, if the merging score is (or is not) a problem in intensive systems for fattening pig. The present sample sizes allow no representative statement in this respect, but seem to be suggesting Temple et al. (2010a).

The data collected from slaughterhouses could also be useful for the assessment of animal welfare. However, studies have shown that the reliability and validity is unsatisfactory at present (Bahlmann, 2009). The predictive value (the validity and reliability) of the collected data from slaughtering processes for animal welfare is not without controversy and better (especially conform) procedures of assessments at slaughterhouses are necessary. The data could especially be very useful to document long-term changes of animal health.

Simple arrangements to improve the reliability are intensity training for observers and refining definitions or data recording design (Knierim and Winkler, 2009). In the Welfare Quality® protocols, both the training of the consortium and mentoring given after training were already judged as being good by the authors.

Feasibility and practicability

The time needed is certainly to be rated efficient for a correct evaluation of on-farm animal welfare situation. Like the animal based approach of Welfare Quality®, these measures take much more time (2/3) than the management- and of resource-based one (1/3). With increasing herd size, the total time needed also increase. However, the time per pig is reduced as the time needed for the herd analysis, collection of general information and discussion with the farmer is similar to that needed

for smaller units. The feeding times and other managerial practices needs to be taken into consideration and reduces flexibility when the investigations can take place but otherwise the system is flexible regarding time. The only part of the assessment that requires the farmers input is the interview at the beginning. Otherwise, participation from the farmers is not required.

However, a final meeting should be held to explain the assessment's results and to discuss any recommendations for future practice. Especially the animal-based measurements (particularly the behaviour assessments) require some explaining; enough time should be taken to ensure an adequate transfer of knowledge. Knierim and Winckler (2009) even emphasise the high degree of interest of their farmers in the animal-based parameters.

Also, it must be clarified what happens with sick and injured pigs which have been taken out of their original pens and placed in a hospital pen. Such animals can often then no longer be exactly placed with respect to their original pen, making their inclusion in the assessment difficult. Apart from this weakness, from the authors' point of view, the Welfare Quality® system fulfils the requirements of a feasible assessment with a short duration that is easy to perform under intensive production conditions. However, for a valid implementation of this method, schooling of the observers by the Welfare Quality® consortium is of paramount importance.

Conclusions

The present study must be considered as a preliminary study which analyses the reliability and feasibility of the assessment protocols for pigs in intensive production systems. It is a first step to give an overview of the reliability and feasibility of the assessments protocols in these productions systems and does not allow giving a general statement about improvement measurements. The results indicate, however, pioneering clues on which further studies have to follow. As a conclusion, it can be said that the Welfare Quality® protocols seem to be a right step in the context of the ongoing discussion regarding a reliable on-farm assessment system for these production systems.

REFERENCES

- Aparicio Tovar MA, Vargas Giraldo JD (2006). Considerations on ethics and animal welfare in extensive pig production: Breeding and fattening Iberian pigs. *Livest. Sci.* 103:237-242.
- Bahlmann J (2009). Koordination in Food Supply Chains: Status quo und Potenziale ITbasierter Informationssysteme [Coordination in Food supply-Chains: Status puo and potential of IT-based information systems, Dissertation Georg-August-Universität Göttingen, Göttingen.
- Blokhuis HJ, Keeling LJ, Gavinelli A, Serratos J (2008). Animal welfare's impact on the food chain. *Trends Food Sci. Technol.* 19(1): 79-87.
- Botreau R, Bracke MBM, Pery P, Butterworth A, Capdeville J, Van

- Reenen CG, Veissier I (2007). Aggregation of measures to produce an overall assessment of animal welfare. Part 2: Analysis of constraints. *The Animal Consortium* 2007. *Animal* 1(8):1188-1197.
- Carlsson F, Fryblam P, Lagerkvist CJ (2007). Consumer willingness to pay for farm animal welfare: mobile abattoirs versus transportation to slaughter. *Eur. Rev. Agric. Econ.* 34(3):321-344.
- Dalmau A, Temple D, Rodríguez P, Llonch P, Velarde A (2009). Application of the Welfare Quality® protocol at pig slaughterhouses. *Anim. Welfare* 18(4):497-505.
- Duncan IJ, Petherick JC (1991). The implications of cognitive processes for animal welfare. *J. Anim. Sci.* 69(12):5017-5022.
- Ellis KA, Billington K, McNeil B, McKeegan DEF (2009). Public opinion on UK milk marketing and dairy cow welfare. *Anim. Welfare* 18:267-282.
- Gomez KA, Gomez AA (1984). *Statistical Procedures for Agricultural Workers*, John Wiley and Sons, New York, 2nd edn.
- Harper GC, Makatouni A (2002). Consumer perception of organic food production and farm animal welfare. *Br. Food J.* 104(3/4/5):287-299.
- Hughes BO, Duncan IJH (1988). The notion of ethological need, models of motivation and animal welfare. *Anim. Behav.* 36(6):1696-1707.
- Knierim U, Winkler C (2009). On-farm welfare assessment in cattle: Validity, reliability and feasibility issues and future perspectives with special regard to the Welfare Quality approach. *Anim. Welfare* 19:451-458.
- McGlone J (2001). Farm animal welfare in the context of other society issues: Toward sustainable systems. *Livest. Prod. Sci.* 72(1-2):75-81.
- Mouttotou N, Hatchell FM, Green LE (1999). Prevalence and risk factors associated with adventitious bursitis in live growing and finishing pigs in south-west England. *Prev. Vet. Med.* 39(1):39-52.
- Patel JK, Patel NM, Shiyani RL (2001). Coefficient of variation in field experiments and yardstick thereof – An empirical study. *Curr. Sci.* 81: 1163-1164.
- Scott K, Binnendijk GP, Edwards SA, Guy JH, Kiezebrink MC, Vermeer HM (2009). Preliminary evaluation of a prototype welfare monitoring system for sows and piglets (Welfare Quality® project). *Anim. Welf.* 18:441-449.
- Temple D, Dalmau A, Ruiz de la Torre JL, Manteca X, Velarde A (2011a). Application of the Welfare Quality® protocol to assess growing pigs kept under intensive conditions in Spain. *J. Vet. Behav. Clin. Appl. Res.* 6(2):138-149.
- Temple D, Manteca X, Velarde A, Dalmau A (2011b). Assessment of animal welfare through behavioural parameters in Iberian pigs in intensive and extensive conditions. *Appl. Anim. Behav. Sci.* 131(1-2):29-39.
- Tuytens FAM, Vanhonacker F, Van Poucke E, Verbeke W (2010). Quantitative verification of the correspondence between the Welfare Quality® operational definition of farm animal welfare and the opinion of Flemish farmers, citizens and vegetarians. *Livest. Sci.* 131:108-114.
- Van de Weerd HA, Day JEL (2009). A review of environmental enrichment for pigs housed in intensive housing systems. *Appl. Anim. Behav. Sci.* 116(1):1-2.
- Velde HD, Aarts N, Woerkum CV (2002). Dealing with ambivalence: Farmers' and consumers' perceptions of animal welfare in livestock breeding. *J. Agric. Environ. Ethics* 15:203-219.
- Waiblinger S, Knierim U, Winkler C (2001). The development of an epidemiologically based on-farm welfare assessment system for use with dairy cows. *Acta Agriculturae Scandinavica Section A. Anim. Sci. Suppl.* 30:73-77.
- Welfare Quality® (2009). Welfare Quality® assessment protocol for pigs (sow and piglets, growing and finishing pigs). Welfare Quality® Consortium, Lelystad, The Netherlands.