Use of animal based measures for assessing farm-animal welfare.

George Stilwell
Animal Behaviour and Welfare Research Lab,
Centre for Interdisciplinary Research in Animal Health
Veterinary Medicine Faculty, Lisbon University
Av. Universidade Técnica, Alto da Ajuda, 1330-477 Lisboa, Portugal
stilwell@fmv.ulisboa.pt

Welfare indicators

Two broad categories of indicators can be used to assess animal welfare at farm level: a) animal-based indicators (e.g. behavioural measurements, productivity and health records) and b) resource- (e.g. manpower, installation details, stocking density, feedstuff quality, etc.) and management-based indicators (e.g. how and when animals are fed, moved and mixed with other animals, breeding strategies, how routine practices like tail docking or dehorning are performed).

Resource- and management-based indicators (also known as in-put measures) were frequently included in the first welfare assessment protocols (e.g. Animal Needs Index TGI 35L). These indicators were used to estimate the welfare status of animals at farm level, because they are less subjective, often easier to audit, very convenient and highly repeatable, frequently having high inter and intra-observer repeatability. The first reports published by the European Food Safety Authority (EFSA) tried to identify environment and management indicators of poor welfare, because this corresponded to a quick and easy assessment. However, most descriptions were very ambiguous. For example: "Floors (…) designed as not to cause injury or suffering to calves standing or lying on them."

Consequently most European legislation on farm animal welfare was supported on these types of indicators (e.g. Council Directive 91/629/EEC, laying down minimum standards for the protection of calves). The fact that legislation focused mostly on housing and husbandry aspects meant that monitoring did not always reflect an overall status of animal welfare and did result in very inflexible rules.[1] Considering that legislation is produced at an European level, it was very difficult to produce legislation detailing different production practices throughout European countries, as these countries are subjected to different farming, cultural and environmental conditions. So a solution to overcome these difficulties and to allow for the development of legislative documents that can be readily applied across different European countries was deemed crucial. At least this is what European farmers expect from new legislation on animal welfare.

It became evident that good management and access to an adequate environment did not necessarily result in a high standard of welfare. So a different approach was suggested by EFSA, emphasising the role of animal-based indicators as they seem more appropriate for measuring the way the animal is responding and coping with its surroundings. In risk assessment terminology these responses are the consequences of the factors acting upon the animal. If they are negative they are referred to as adverse effects and if positive they are referred to as benefits.[2]

It should be said that management and resource-based indicators continue to be of extreme value but should be perviewed more as risk factors or a way to predict and prevent welfare problems than direct measures of animal welfare. They are also frequently used when record of certain animal-based indicators is not feasible or is too

Proceedings of AVA Annual Conference, Adelaide, 2016.
Stilwell, G - Use of animal based measures for assessing farm-animal welfare

subjective. For example, the EU Strategy for the Protection and Welfare of Animals 2012–2015 highlighted the possibility of using scientifically validated enviornment-based indicators to complement animal-based requirements in EU legislation.

Why animal-based indicators?

So, although both categories of indicators – environmental and animal-based – are significant aspects of animal welfare, and so the most valid assessment of animal welfare seem to be achieved when these are used in association, at the moment animal-based measures are considered by EFSA to be "the most appropriate indicator of animal welfare" and "…a carefully selected combination of animal-based measures can be used to assess the welfare of a target population in a valid and robust way".[3] The use of animal-based measures is becoming widespread so international organizations [International Standard Organisation (ISO) and World Organisation for Animal Health (OIE)] are working together so as to develop common standards and consensual thresholds for animal welfare indicators.

Welfare Quality® (see elsewhere in these Proceedings) was the first large European project to focus primarily on animal-based measures to be monitored and used during a single short farm visit by an independent inspector to assess actual levels of welfare instead of trying to identify factors that may eventually lead to good or poor welfare.

Animal-based measure is evaluative, obtained in a precise way and usually quantitative. It gives an indication of an animal's welfare, but a set of measures taken from an *toolbox* (see below) is normally needed to provide a good assessment of welfare.

Animal-based measures can be collected on-farm either by observation or inspection of the animal or by assessing the effects of a response on the environment or by consulting production records. EFSA has divided welfare indicators in four categories: health indicators, physiological indicators, behavioural indicators and productive indicators.[3]

The most commonly used are health indicators which include injuries and disease signs, both of which can be directly assessed by looking at the animals. Behavioural indicators, either in the form of tests performed (e.g., avoidance distance) or simple observations (e.g. signs of lameness or stereotypies) are also very commonly used in welfare assessment schemes. Some of the advantages attributed to these indicators are that they are non-invasive and non-intrusive. Physiological indicators, that includes blood parameters, heart-rate and hormones, have a lot of limitations in on-farm welfare assessments as they are difficult to interpret, not always associated with poor welfare, and entail several feasibility issues. In the case of records it should be said that production can be satisfactory even if some animals are in poor conditions[2], which makes performance indicators unspecific and therefore difficult to interpret.

Sometimes new measures are generated by combining different animal-based measures from the same individual. An example of *composite animal-based measure* is the production record (milk yield, milk composition and calving interval) for an individual cow over three lactations.

Although the aim of animal-based measures is to represent what the animals is going through and what is its actual welfare status, some measures can be used to detect those animals whose welfare is at risk – *predictive welfare indicator*.[3] An example is reduced fertility – although not becoming pregnant is not in itself a welfare problem it may indicate that, for example, nutrition quality is low and a disease is more probable in the future. Or high somatic cell count may indicate that animals will more easily develop clinical mastitis.

Proceedings of AVA Annual Conference, Adelaide, 2016.
Stilwell, G - Use of animal based measures for assessing farm-animal welfare

Independently of the type, all indicators to be used in welfare assessment must obey to certain prerequisites.

- validity - Does the indicator really measures what we think it measures? Does it relate to the animals experience? Validity is determined by accuracy, specificity and scientific validity. Two things are important to mention: an overall assessment system is only as valid as the indicators used to establish it; and, validity might be situation dependent – sometimes measures that are valid during one stage of the production cycle or in a certain production system are not applicable at other times or other conditions. This is what the World Organisation for Animal Health refers as being both fit for the *intended purpose* and fit for the *intended use*.[3]

- reliability - Reliability of an indicator shows the relative similarity of results when assessing one animal on several occasions, being associated with repeatability (De Passillé & Rushen, 2005). Would indicators be recorded in the same way by more than one assessor? Would the same assessor record indicators in the same way on more than one occasion? Essential features of reliability are inter-observer and intra-observer reliability, and also test–retest reliability. EFSA group these attributes into one called **robustness.**[3]

- feasibility - Can it be measured on farm in a reasonable manner (i.e. time, manpower, tools…)? Feasibility refers to the degree to which the evaluation or collection procedure is possible, practicable and worthwhile.[4] Indicators importance should also justify the cost, effort involved and the disturbance caused to the subjects. The development of feasible indicators is one of the main challenges in welfare assessment research

Assessment tool-box
Waiblinger et al.[5] suggested that for an assessment protocol to be effectively used on different farms it must include the following characteristics: contain indicators that are reliable and valid; be easily employed by trained people; need limited time so that repeated assessments on several farms are conceivable; expose the reasons of reduced welfare and subsequently present possible improvements for the farm.
There are many animal-based welfare measures that obey to these conditions. To this assemblage the term *"assessment tool-box"* is sometimes applied. It is obvious that it is unrealistic and unnecessary to use all tools on every occasion the welfare of an animal is to be assessed. Some will have very low prevalence, some will give repeat information, some will need too much time to collect, and some will demand complex interpretation…

In the other hand some indicators may reflect more than one welfare problem or may indicate a very complex welfare status. For example: low body condition score, may result from undernutrition, mineral deficiency, chronic disease or chronic pain;[6] hair coat condition in goats, an indicator recently developed by Battini et al.,[7] was shown to reflect not only nutritional status but also health status (e.g. chronic pneumonia). These indicators have a low specificity and are designated as *iceberg measures* because they have the potential to provide an overall assessment of welfare.

There are multiple other interactions between animal-based measures. For example, a lame cow may have difficulty in accessing the food trough therefore increasing its risk of metabolic disease. At the same time, it may lie for longer periods of time, so increasing its risk of mastitis.

Animal-based indicators may also vary along the day or between seasons. For this reason it is important to specify when the measure is to be taken or repeat evaluation at different moments to ensure consistency.

Proceedings of AVA Annual Conference, Adelaide, 2016.
Stilwell, G - Use of animal based measures for assessing farm-animal welfare

Thus a selection of tools must be made so as to address the specific objectives of a specific assessment for a particular species and category of animal at each time. It is the careful selection of these measures that will lead to the development of useful welfare assessment protocols such as those presented elsewhere in these Proceedings.

Other methodological issues have to be address to guarantee reliable and useful data. For example, sampling is a very important aspect to consider in large farms as it may be impossible to collect animal-based indicators from all individuals. To ensure reliability and feasibility it is essential that a valid, unbiased and representative sampling strategy is implemented. There are several problems associated with sampling strategies. For example, the implementation of these strategies require previous knowledge and determination of parameters such as expected prevalence, level of precision, and the confidence level of the estimation that frequently are not available.

Equally important is training of the assessors. To correctly collect and interpret animal-based indicators and especially those associated with behaviour, a perfect knowledge of natural or normal behaviour is essential. This is done by training assessors to the gold standard for all measures and repeated training at regular intervals so that observers are recalibrated.

Challenges of animal-based indicators
The integration of animal-based indicators in welfare assessment schemes entails several challenges. For instance, there are very few indicators already validated, especially for small ruminants. For example, the association of injuries and diseases, mainly linked to pain, is relatively straightforward, but for other measures proving a correlation with welfare is much more difficult (e.g. human-animal relationship, flight distance or positive emotional state).Therefore, there is a strong need for further research, validation or adaptations.
It is also essential to constantly verify reliability as an observers' assessment might be biased by the personal concern for the animals. Error may be minimized by adequate and frequent training, but bias should always be a target of evaluation.

For indicators' collection and interpretation a scoring system has to be defined to allow their integration in any welfare assessment scheme. The assessment objective (e.g. legislative or advisory) usually determines the indicators to be integrated. For example, the assessment of only the extreme BCS cases (very thin and very fat animals) may be enough for legislative purposes, but perhaps not for advisory ones where a more sensitive scale is needed.[6]

There are different types of scoring systems available. In welfare assessment two types of scoring systems are more frequently used: ordinal and continuous scales.[8] Ordinal scales, also called NRSs, are explicit scoring systems in which each individual is scored accordingly to different descriptors that correspond to a whole number. In welfare assessment, ordinal scales are the scoring systems most frequently used, due to being generally easy to deliver as it merely requires a comparability assumption. However, being in most cases artificial constructs, they are associated with difficulties in their use and interpretation. This is particularly evident if we use them to assess variables that are continuous by nature, such as lameness or pain levels. There are several problems associated with ignoring the continuous nature of variables. The first one is that by itself it introduces error in the answers, as different people will interpret differently the descriptors. Another source of error will be the limited choice of descriptors usually available. The third problem is the associated loss of efficiency and sensibility of the scoring system, with potential loss of valuable information.[8] These problems are frequently prevented by the use of continuous scoring systems. The most used

continuous scoring systems in animal welfare assessment is the visual analogue scale (VAS). The VAS is a nonverbal scale allowing for the recording of clinical events, such as lameness, in a continuous way. The VASs were developed for use in pain assessment in humans but are also used to measure a variety of subjective phenomena in the behavioural and social sciences, and are considered to have potential value for the measurement of different clinical phenomena.[8] The VAS has the advantage of not imposing a choice of descriptors, being possible to score a change even if a change between categories would not occur, and therefore overcoming some drawbacks associated with ordinal scales. However, VASs are generally viewed as being too subjective, with low user-acceptance, and difficult to use in farm conditions. Current research in scale development has recently been focusing on the development of modified VASs to assess different health indicators. A modified VAS is a VAS that besides being anchored on its extreme limits, it presents the thresholds of a categorical scale specified on the VAS continuum, dividing it into segments that are complemented with descriptors. The objective is to help observers understand the scale and make consistent choices, therefore increasing the inter-observer reliability by reducing data variation.[9] The modified VAS keeps the advantages of the VAS, such as allowing for the indicator to be recorded in a continuous way, but showing a higher resolution and being more sensible to smaller changes than NRS existing associations between the three layers, i.e. the measure (the indicator or test), the outcome (welfare consequence) and the factor (hazard). The AWIN studies leading to the welfare assessment protocols, include VAS for BCS and lameness in goats.[6]

In conclusion, when establishing science based, practical and suitable protocols, using animal based indicators seem preferable. Although perhaps more subjective, these indicators seem to better reflect what the animal is feeling and the success of its coping with the environment. However, exceptional attention should be given to validity, reliability, and feasibility studies.

Potentially there are many different animal-based measures that can be used as tools to assess animal welfare. However, not all are applicable at all moments. Decisions on the selection from the tool-box and the way to employ these tools depend on the purpose of the assessment, the specie and breed, the training and skills of the assessor, the conditions under which the indicator is to be gathered, the time available to collect it, the financial constraints etc

References
1. Blokhuis H, Miele M, Veissier I et al. *Improving farm animal welfare - Science and society working Ed. H.J. Blokhui. Springer. Wageningen Acad Publ* 2013.
2. Botreau R, Bracke MBM, Perny P et al. Aggregation of measures to produce an overall assessment of animal welfare. Part 2: analysis of constraints. *Animal* 2007;1:1188–97.
3. EFSA - Panel on Animal Health and Welfare. Statement on the use of animal - based measures to assess the welfare of animals. *EFSA J* 2012;10:2767. *http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_docume nts/2767.pdf*.
4. Martin P, Bateson P. *Measuring Behaviour: An Introductory Guide*. Cambridge University Press, 2007.
5. Waiblinger S, Knierim U, Winckler C. The Development of an Epidemiologically Based On-Farm Welfare Assessment System for use with Dairy Cows. *Acta Agric Scand Sect A — Anim Sci* Taylor & Francis Group, 2010;51:73–77..
6. Vieira A, Brandão S, Monteiro A et al. Development and validation of a visual body condition scoring system for dairy goats with picture-based training. *J Dairy*

Proceedings of AVA Annual Conference, Adelaide, 2016.
Stilwell, G - Use of animal based measures for assessing farm-animal welfare

       Sci 2015;98:6597–6608..

7.   Battini M, Peric T, Ajuda I et al. Hair coat condition: A valid and reliable indicator for on-farm welfare assessment in adult dairy goats. *Small Rumin Res* 2015;123:197–203..

8.   Scott EM, Nolan AM, Fitzpatrick JL. Conceptual and Methodological Issues Related to Welfare Assessment: A Framework for Measurement. *Acta Agric Scand Sect A — Anim Sci* Taylor & Francis Group, 2010;51:5–10.

9.   Kaler J, Wassink GJ, Green LE. The inter- and intra-observer reliability of a locomotion scoring scale for sheep. *Vet J* 2009;180:189–94

Proceedings of AVA Annual Conference, Adelaide, 2016.
Stilwell, G - Use of animal based measures for assessing farm-animal welfare

812