How reliable is the multi-criteria evaluation system of the Welfare Quality® protocol for growing pigs?

I Czycholl*,†, C Kniese‡, L Schrader‡ and J Krieter†

† Institute of Animal Breeding and Husbandry, Christian-Albrechts-University, Olshausenstrasse 40, D-24098 Kiel, Germany
‡ Institute of Animal Welfare and Animal Husbandry, Friedrich Loeffler Institute, Dornbergstrasse 25/27, D-29223 Celle, Germany
* Contact for correspondence and requests for reprints: iczycholl@tierzucht.uni-kiel.de

Abstract

This paper focuses on the reliability of the multi-criteria evaluation model included in the Welfare Quality® protocol for growing pigs to aggregate the animal-based indicators, first to criteria, then to principle level and finally to an overall welfare score. This assessment was carried out in a practical application study on a sample of 24 farms in Germany. Altogether, 102 protocol assessments were carried out in repeated visits to these farms in order to evaluate the inter-observer and test-retest repeatability of the overall scores calculated by the multi-criteria evaluation system. Reliability is then assessed by the calculation of different reliability and agreement parameters: Spearman Rank Correlation Coefficients (RS), Intraclass Correlation Coefficients (ICC), Smallest Detectable Changes (SDC) and Limits of Agreement (LoA). Inter-observer repeatability was insufficient for the criteria comfort around resting, absence of injuries, expression of social behaviours, expression of other behaviours, good human-animal relationship and positive emotional state as well as for the principles good housing and appropriate behaviour. This is probably due in the main to insufficient repeatability of the underlying indicators that have been revealed in previous studies. Test-retest repeatability is predominantly insufficient. Overall, the present results highlight the importance of absolutely reliable indicators at the baseline level. Furthermore, it could be shown that the calculation procedure is partly incorrect and consequently needs correction. Therefore, this study is an important contribution to the future progression of the Welfare Quality® protocols and animal welfare assessment tools in general.

Keywords: animal welfare, multi-criteria evaluation, pigs, reliability, repeatability, Welfare Quality®

Introduction

Animal welfare has become an important topic within public and political discussion over recent decades (Hobbs et al 2002). There is, therefore, a growing need for stakeholders to comply with and increase animal welfare standards (Vapnek & Chapman 2010). In order to meet consumers’ concerns in the form of, eg animal welfare certification schemes, it is necessary to carry out an objective evaluation of the welfare status on-farm (Blokhuis et al 2013).

Animal welfare comprises different aspects, such as the absence of thirst, hunger, discomfort, disease, pain, injuries and stress as well as the possibility to express normal behaviour (FAWC 1993). Thus, it is a multi-dimensional concept. A welfare assessment system has to take into account all different aspects in order to gain general acceptance. Different approaches have been proposed in recent years (Czycholl et al 2015). As such, the Welfare Quality® protocols are very promising as a systematic welfare evaluation tool since all the different dimensions of animal welfare are addressed and focus is given to animal-based indicators. According to Blokhuis et al (2013), animal-based parameters are the only parameters which assess the true value with regard to animal welfare.

In detail, in the Welfare Quality® protocols, the implementation of the multi-dimensionality of the concept of animal welfare takes place in the form of four principles, which are good feeding, good housing, good health and appropriate behaviour. These principles are divided into twelve criteria. These are then measured, on-farm, by a set of approximately 30 predominantly animal-based indicators. After the on-farm assessment, the measures are usually expressed as percentages of affected animals. These percentages are standardised into a dimensionless number between 0 and 100 by a multi-criteria evaluation system. Depending on the scores reached on principle level, as an overall assessment the farms are labelled as excellent, enhanced, acceptable or not classified (Welfare Quality® 2009). This multi-criteria evaluation system again places particular emphasis on the multi-dimensionality of the concept of animal welfare. Moreover, the aggregation of the on-farm measures for the total evaluation of the farm is especially required for labelling purposes (Botreau et al 2007).