

Repeatability of indicators for the on-farm assessment of sheep welfare

C. J. Phythian, M. J. Clarkson, A. C. Winter, P. H. Jones,
D. Grove-White, P. J. Cripps, E. Michalopoulou, & J. S. Duncan

School of Veterinary Science, Leahurst, Neston, CH64 7TE

C.J.Phythian@liverpool.ac.uk



Introduction

A number of welfare issues including poor body condition, lameness and mastitis are associated with the on-farm period of sheep production. Ideally the welfare of sheep would be assessed using a gold standard measure that was both robust to the diversity of sheep farming and feasible for on-farm use. To date, indicators fulfilling the scientific criteria of validity, repeatability and feasibility do not exist. In the absence of a gold standard measure, a scientific literature review and expert consultation process were used to identify potential indicators for the welfare assessment of sheep in England and Wales. The aim of this study was to determine the inter-observer repeatability for three potential sheep welfare indicators: 1. body condition, 2. lameness and 3. mastitis.

Methodology

Sample population

- 38 flocks located in Northern England and North to Mid Wales were assessed during July to December 2008

Observer characteristics

- Pool of 8 observers: veterinary surgeons (n = 5) and students (n = 3)
- Observers were trained (n = 6) or untrained (n = 2) in assessments
- Observers were categorised either previously experienced (n = 4) or inexperienced (n = 4) in sheep welfare assessment
- One experienced observer; a veterinary surgeon, who developed the indicators, was classed the 'test standard' (gold standard)

On-farm testing

- Multiple observers (range 2-3), independently assessed a sample of sheep (n = 30) on a variable number of farms
- Individual sheep were assessed using the welfare indicators:
 - Body condition** (0 – 5 ordinal scoring scale: full score precision)
 - Lameness** (0 – 1: binary scoring scale)
 - Mastitis** (0 – 2: categorical scoring scale)



Statistical analysis

Inter-observer repeatability was evaluated using a proportional chance-adjusted agreement analysis known as 'kappa' (κ):

$$\kappa = \frac{\text{Pr}(o) - \text{Pr}(c)}{1 - \text{Pr}(c)}$$

Pr (o) = proportion of observed agreement
Pr (c) = proportion of agreement expected by chance

- Repeatability of multiple observer assessments were determined by Fleiss's kappa
- Repeatability with the 'test standard' was determined by Cohen's kappa
- Kappa results were interpreted according to Fleiss (1981) as per table 1:

Table 1: Interpretation of Fleiss's Kappa

Kappa value	Interpretation
> 0.74 – 1.0	Excellent
< 0.75 – 0.41	Good - Fair
< 0.41	Poor

- Kendall's coefficient of concordance (W) further evaluated the agreement for the ordinal scoring indicator; body condition. Kendall's W was interpreted on a scale of 0 (no agreement) to 1 (perfect agreement)

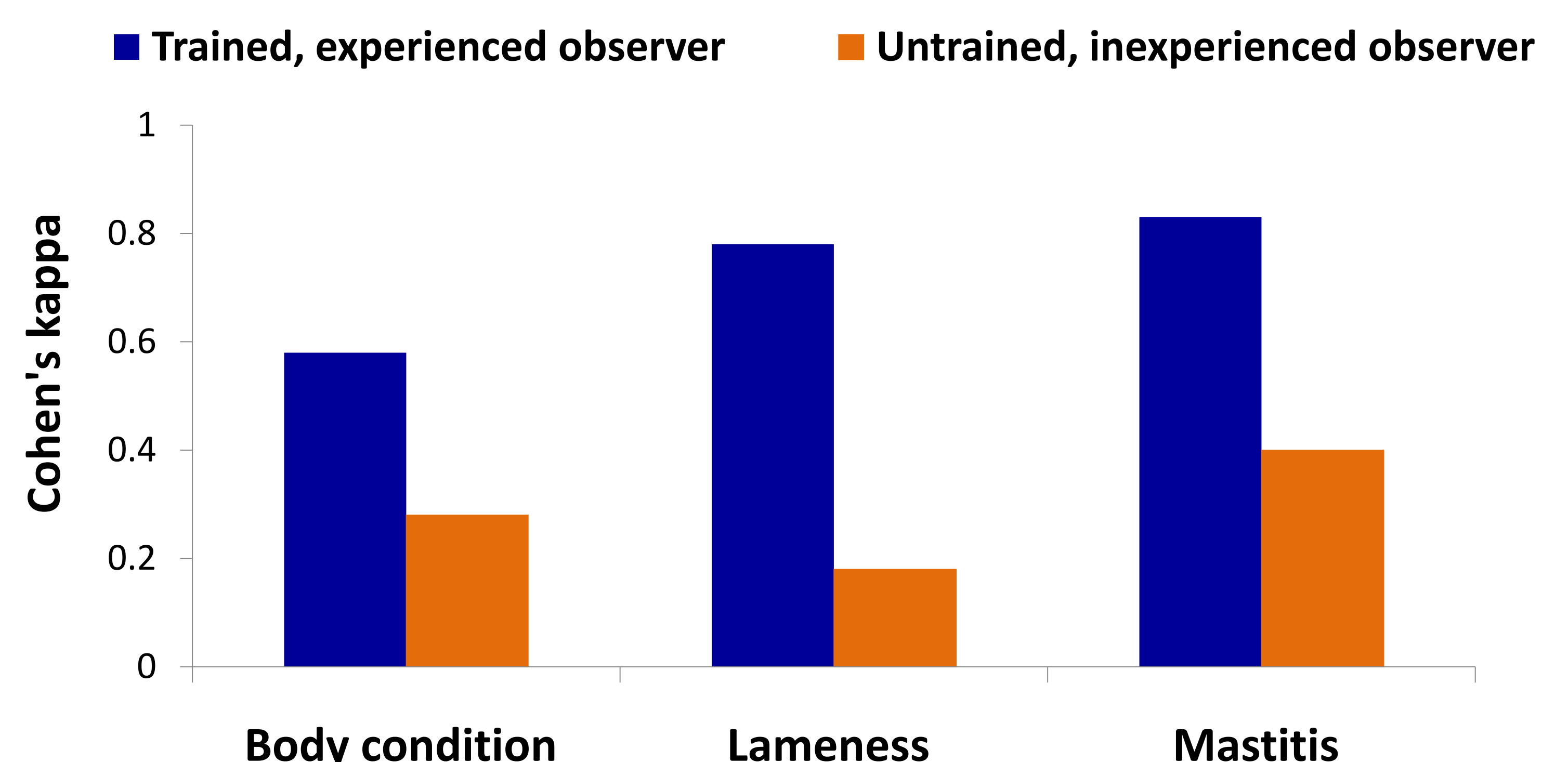
Results

Multiple observers independently assessed indicators of body condition, lameness and mastitis on a total of 1159 adult sheep. Overall observer repeatability results are shown in table 2:

Table 2: Overall inter-observer repeatability:

Indicator	Fleiss Kappa
Body condition	0.46
Lameness	0.66
Mastitis	0.44

Figure 1: Repeatability with the 'test standard' observer



Kendall's W: Trained experienced = 0.83, Untrained inexperienced = 0.79

The main findings of this study are:

- Overall good to fair levels of inter-observer repeatability for indicators of body condition, lameness and mastitis (table 2)
- An effect of observer training and experience on the level of inter-observer repeatability (figure 1)

Conclusion

Further investigation into the effect of on-farm conditions and observer training on the level of repeatability are required before these indicators can be recommended for inclusion in formal assessments into the on-farm welfare of sheep