Stoplight System for Assessing Equine Welfare

Penny Lawlis, Gayle Ecker and Vanessa Pasquale

UNIVERSITY of GUELPH
Improve Life.

EQUINE GUELPH
helping horses for life ™

Introduction

Measures of animal welfare need to be practical, economical and feasible. Measures selected need to show good interobserver reliability (IOR). Visual assessment is practical and accurate for broad categories (Burkholder, 2000) and is a valuable tool used by animal welfare inspectors. Animal welfare inspectors/first responders may not have the opportunity for extensive equine training in behavior/husbandry. Equine Guelph (EG), University of Guelph developed a simplified animal welfare assessment protocol (i.e.,The Stoplight System) for inspectors.

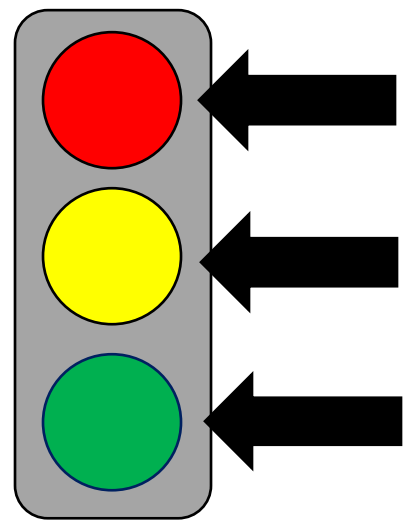

Figure 1. Stoplight scoring system with colour description

Objectives

The objectives of this preliminary study were to test the inter-observer reliability of the Stoplight System and to investigate the potential of the system for use by investigators or first responders.

Materials & Methods

The Stoplight System includes animal-based and resource-based protocols for assessing hooves, lameness, body condition, feed, and water. Parameters were chosen based on the requirements in the Equine Code of Practice published by the National Farmed Animal Care Council (NFACC). Categories of the measures used colour cues versus numbers as per most assessment scoring systems. An expert in horse welfare assessment trained five behaviour and welfare graduate students with no prior horse experience during three hours of classroom training which included photographs, written descriptions, verbal explanations and videos. Subsequently the expert and the students performed a live assessment of these welfare measures on twelve horses at a therapeutic riding stable. Data from the in-class and on-farm assessments were assembled into agreement matrices to compute kappa statistics for each welfare indicator using SAS. Use of the kappa statistic was chosen based on the small number of observations.

Results

- Highest percentage agreement on farm was lameness
- feed, body condition score, and water were all tied at 50% agreement in-class
- Students tended to score under but not over score the expert gave.
- On average, inter-observer reliability was higher for animal-based measures (BCS and Lameness) than for resource-based (feed and water).

Welfare Measure	In-Class Weighted Kappa	On-Farm Weighted Kappa
Body Condition Score	0.7991	0.1014
Lameness	0.4118 *(simple kappa)*	–
Feed	0.6513	0.6467
Water	0.4318	–

Discussion and Conclusion

The results suggested that both the system and the training needed refinement, specifically expanding the training to include examples of acceptable resources such as water and feed sources. Although students appreciated the simplicity of the system, the data suggested that inexperienced observers tend to under-score which can lead to under-reporting and less prompt intervention in the field. The small sample size was a limitation when analyzing the scoring data, as categories like lameness are disproportionately high because all the horses scored did not show lameness. In a subsequent study using the same stoplight system but with dairy cattle, the student observers were able to make more accurate results compared to the results with horses. In conclusion, this preliminary study was useful in highlighting gaps in training required for novice first responders as well as providing information to improve protocols.