Original Article

A systematic review of animal based indicators of sheep welfare on farm, at market and during transport, and qualitative appraisal of their validity and feasibility for use in UK abattoirs

P. Llonch [a, c], E.M. King [a], K.A. Clarke [b], J.M. Downes [b], L.E. Green [a]

[a] *School of Life Sciences, University of Warwick, Coventry, CV4 7AL, UK*
[b] *One Health Veterinary Services, Newcastle upon Tyne, NE20 9JU, UK*
[c] *Current address: SRUC, Roslin Institute Building, Easter Bush, Edinburgh, EH25 9RG, UK*

* Corresponding author. Tel.: +44 131 6519362.
E-mail address: pol.llonch@sruc.ac.uk (P. Llonch).

Abstract

In the UK, it has been suggested that abattoirs are ideal locations to assess the welfare of sheep because most sheep are slaughtered at abattoirs, either as finished lambs or cull ewes. Data from abattoirs could therefore provide benchmarks for welfare indicators at a national level, as well as demonstrating how these change over time. Additionally, feedback could be provided to farmers and regulatory authorities to help improve welfare and identify high or low standards for quality assurance or risk-based inspections. A systematic review of the scientific literature was conducted, which identified 48 animal-based indicators of sheep welfare that were categorised by the Five Freedoms. Their validity as measures of welfare and feasibility for use in abattoirs were evaluated as potential measures of prior sheep welfare on the farm of origin, at market, or during transportation to the abattoir.

A total of 19 indicators were considered valid, of which nine were considered theoretically feasible to assess sheep welfare at abattoirs. These were body cleanliness, carcass bruising, diarrhoea, skin lesions, skin irritation, castration, ear notching, tail docking and 'obviously sick'. Further investigation of these indicators is required to test their reliability and repeatability in abattoirs. Novel welfare indicators are needed to assess short-term hunger and thirst, prior normal behaviour and long-term fear and distress.

Keywords: Abattoir; Animal-based welfare indicators; Sheep; Systematic review; Validity

Introduction

In the UK, most sheep are slaughtered in abattoirs. Observation of sheep in abattoirs, using indicators that are transparent and fair, might provide an assessment of prior health and welfare on farm, at market and during transportation. While such inspections would not replace all inspections elsewhere, data could be used to benchmark the prevalence of welfare indicators (Farm Animal Welfare Council, 1993), to inform risk-based selection for inspections measuring compliance with animal welfare legislation, evaluate whether an assurance scheme's welfare conditions are met (Kilbride et al., 2012), or provide farmers with information on the health and welfare of their livestock to assist in health planning.

Assessment of sheep in abattoirs requires valid welfare indicators. Animal-based (outcome-based) indicators of welfare use direct assessment of an animal's mental and physical welfare. They are considered the most valid method of assessing animal welfare because the animals themselves are assessed, not their resources, and comparisons can be made across all systems of husbandry (Main et al., 2007). For industry to use such indicators in abattoirs, they must be valid (measure what they intend, i.e. animal welfare), repeatable (the same result for repeated observations of the same animal by the same and different observers), reliable (consistent results across observation of different animals) and feasible (in terms of speed, cost and not compromising normal operating procedures; Knierim and Winckler, 2009; Napolitano et al., 2009).

With these parameters in mind, we conducted a systematic review of sheep welfare indicators. We categorised these into the Five Freedoms and within each freedom grouped indicators that measured a similar aspect of welfare. Finally, we used the published literature to inform on the validity of each indicator and qualitatively assessed their feasibility for use

62 in abattoirs, to assess the prior welfare of sheep on farm, at market and during transportation
63 to the abattoir.
64
65 **Materials and methods**
66 *Search criteria and strategy*
67 We conducted a systematic review of peer-reviewed scientific literature published
68 from 1 January 1995 to 15 December 2012[1]. All experimental and observational studies of
69 sheep welfare (including research papers, conference proceedings and literature reviews)
70 referring to welfare assessment of adult sheep or lambs (*Ovis aries)* were included. Searches
71 were performed using the same search terms in four search engines: (1) Pubmed[2]; (2)
72 ScienceDirect[3]; (3) Scopus[4]; and (4) Web of Knowledge[5].
73
74 The search terms used (including all titles, abstracts and keywords) were:
75 '(assess* OR indicator* OR monitor* OR audit OR evaluation OR "animal based" OR
76 clinical AND "animal welfare" OR "sheep welfare" OR welfare AND slaughter* OR abattoir
77 OR mortem* OR farm OR on-farm AND ovine OR "ovis aries" OR sheep OR ram OR "dairy
78 sheep" OR "sheep farm" OR "sheep flock" OR ewe OR lamb) AND PUBYEAR > 1994'
79

[1] See: Green, S., Higgins, J.P.T., Alderson, P., Clarke, M., Mulrow, C.D., et al., 2008. Chapter 1: What is a systematic review? In: Higgins, J.P.T., Green, S. editors. Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available: http://www.cochrane-handbook.org/. (accessed 22 September 2015)

[2] See: PubMed http://www.ncbi.nlm.nih.gov/pubmed (accessed 22 September 2015)

[3] See: ScienceDirect http://www.sciencedirect.com (accessed 22 September 2015)

[4] See: Scopus www.scopus.com (accessed 22 September 2015)

[5] See: Web of Knowledge http://wok.mimas.ac.uk (accessed 22 September 2015)

Only documents written in English were included. Duplicates and documents not directly related to sheep welfare were removed. A second filter was applied to remove references containing no information about the methodology of assessment of sheep, which resulted in 349 articles remaining.

Criteria for selection of animal-based welfare indicators

A total of 349 articles were retrieved and read to identify all animal-based indicators of sheep welfare. Physiological measurements (e.g. serum cortisol concentration, heart rate) and resource-based observations (e.g. water availability, bedding quality) were rejected. One hundred and twenty-one papers on animal-based indicators were then reduced to papers with indicators made by visual inspection. A total of 218 animal-based indicators in 53 papers were obtained. Related indicators assessing the same welfare problem e.g. gait score and lameness score, were combined to give 48 separate indicators (Table 1). Each indicator was then allocated, using the Five Freedoms (Farm Animal Welfare Council, 1993), to the aspect within a freedom that it measured; the Five Freedoms have been acknowledged as an appropriate framework to assess all aspects of animal welfare (McCulloch, 2013).

Since none of the indicators addressed freedom from thirst, additional literature was reviewed to identify potential indicators reported in other species. This was performed by including the specific term thirst and removing the terms searching species (e.g. sheep, ovine etc.) from the search criteria.

The validity and feasibility of measuring each indicator in an abattoir was categorised as high, medium or low. High validity indicators were those validated in previous research, medium validity indicators were those where the current method of assessment did not

necessarily indicate poor welfare e.g. body condition score (BCS) is a measure of welfare in adults, but might not be a valid measure of welfare in growing lambs. Low validity indicators were those suggested as indicators in the scientific literature, but lacking evidence that they actually assess welfare. Indicators with high feasibility were those that could be recorded in abattoirs, regardless of the number of animals, the space available for the animals, and the speed of the processing line. Medium feasibility indicators were those that needed special requirements (e.g. extra space or time) for appropriate assessment and low feasibility indicators were those that could not be routinely assessed in commercial abattoirs.

Results

The 48 indicators by category and feasibility are presented in Table 1. They were assigned to freedom from hunger and thirst (n = 5), freedom from discomfort (n = 5), freedom from pain, injury and disease (n = 17), freedom to express normal behaviour (n = 8) and freedom from fear and distress (n = 13).

Discussion

This systematic review of animal-based measures of sheep welfare is the first step in the identification of valid and feasible indicators that could be used in abattoirs to monitor the prior welfare of sheep on farm, at market, or during transportation. There were 19 valid indicators were identified which provided information on long-term hunger, discomfort, injury and disease and short-term distress, but only nine were considered feasible for measurement in abattoirs (body cleanliness, carcass bruising, diarrhoea, skin lesions, skin irritation, castration, ear notching, tail docking and obviously sick). In addition, conformation and fat carcass classification (two medium validity indicators), were considered feasible to measure and useful to take forward. No valid, feasible indicators were identified that

130 measured short-term hunger or thirst, long-term normal behaviour or long-term fear and
131 distress, which could be used in abattoirs to assess welfare on the farm of origin, in markets
132 or in transit. Indicators are discussed below.
133
134 In adult sheep, chronic under nutrition can be measured by low bodyweight or low
135 BCS (Jefferies, 1961; Phythian et al., 2012a). Bodyweight varies by age, sex and breed, and
136 since the mature weight of adult sheep varies widely depending on breed, only within-animal
137 comparisons are likely to be valid. Consequently, BCS is generally used as a measure of
138 nutritional status. Although BCS does not indicate current hunger, it does provide
139 information on long-term nutritional status. It is assessed by manual palpation of the lumbar
140 region (Phythian et al., 2012a) or the ribs (Shands et al., 2009) and provides an estimate of
141 body fat and muscle. BCS is valid and reliable; variability between observers can occur but
142 this is reduced by training (Phythian et al., 2012a). Very low BCS (<1.5) indicates emaciation
143 that arises from inadequate feed, chronic disease, or parasitism (Sargison and Scott, 2010),
144 implying severe consequences for sheep welfare.
145
146 Post-mortem, the EUROP carcass classification[6, 7] indicates the shape and volume of
147 muscle in relation to bone structure; the 1 - 5 fat classification assesses the amount of visible

[6] See: Council Regulation (EEC) No 2137/92 of 23 July 1992 concerning the Community scale for the classification of carcases of ovine animals and determining the Community standard quality of fresh or chilled sheep carcases and extending Regulation (EEC) No 338/91. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31992R2137 (accessed 22 September 2015)

[7] See: Commission Regulation (EEC) No 461/93, 1993 of 26 February laying down detailed rules for the Community scale for the classification of carcasses of ovine animals. http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1442955191643&uri=CELEX:31993R0461 (accessed 22 September 2015)

148 fat (Stubsjøen et al., 2011). This classification system is designed for growing lambs, which
149 have a lower fat:muscle ratio than mature animals and consequently adult animals appear
150 'fatter' using this system. Currently, the relationship between BCS and carcass classification
151 has not been assessed; this is an important area for future research.
152
153 An expert panel identified rumen fill as an animal based measure of access to feed
154 (Phythian et al., 2011). Rumen fill can be used to assess nutritional welfare in the short period
155 before slaughter but it is not useful to assess the long-term nutritional state of the animal.
156
157 There is currently no indicator for thirst that fulfils the criteria of direct animal-based
158 assessment in sheep. In veal calves, dehydration is measured by testing duration of skin
159 tenting when skin is pinched between thumb and forefinger (Mellor and Stafford 2004);
160 dehydration is detected by a delay in the skin returning to its normal position. In horses, skin-
161 tenting time is not a valid measure of thirst (Pritchard et al., 2008). This indicator has not
162 been evaluated in sheep.
163
164 Dirt irritates the skin and attracts bacteria, ectoparasites and other pathogens and
165 demonstrates the level of hygiene in which an animal has been kept or transported (Stubsjøen
166 et al., 2011). Assessment can be based on a numerical scale from absolute cleanliness to
167 complete coverage of the body with dirt or faeces (Napolitano et al., 2009; Stubsjøen et al.,
168 2011). Phythian et al. (2012b) focused on certain areas of the body, such as the ventral
169 abdomen and the breech, to give a global score based on visual assessment. For good
170 repeatability between observers, training with clear instructions of assessment must be
171 provided (Stubsjøen et al., 2011).
172

Lying down is frequently related to resting; however, it also relates to other welfare states in sheep. For instance, increased lying time was related to heat stress in Awassi sheep (Dikmen et al., 2011). In lambs, a reduction in lying time is an indicator of pain after castration (Thornton and Waterman-Pearson, 2002). Conversely, lying, due to an inability to stand, has been suggested to be negatively correlated with fitness in newborn lambs (Phythian et al., 2011). The reasons for lying behaviour are diverse and depend on age, management procedures and other factors. Since its relationship with animal welfare is situation-specific, it is not useful in abattoirs.

Thermal stress can arise from extremely low temperatures, provoking hypothermia, or extreme high temperatures, causing hyperthermia. Sheep have behavioural and physiological coping strategies for these conditions. Shivering is an increase in muscular activity to increase body temperature in hypothermic lambs (Mellor and Stafford, 2004), it could, therefore, be a valid indicator of hypothermia, especially in young lambs. An increase in respiratory rate above 40 breaths per min, together with open-mouthed breathing (panting) indicates manageable heat stress in otherwise healthy sheep (Silanikove, 2000). however, 300 breaths per minute indicates severe heat stress (Hales and Brown, 1974; Silanikove, 2000). According to Phythian et al. (2011), fleece cover could also provide information about thermal welfare. Fleece cover can increase resistance to cold temperatures but also increase body temperature during hot weather. Hence it is likely that unshorn animals experience heat stress in high environmental temperatures.

Disease can have a major adverse impact on animal welfare, and some diseases (e.g. clostridial diseases or maedi-visna) are linked to welfare in the published literature (Fitzpatrick et al., 2006). Injury typically leads to inflammation, which is painful. Chronic

198 pain can lead to hyperalgesia or allodynia (Dolan and Nolan, 2000), which contribute to poor
199 welfare. Pain and sub-clinical disease can be difficult to assess in animals. Animal-based
200 indicators of injury and disease need to include a visible physical abnormality that can be
201 detected ante- or post-mortem. Bruises can be used as a measure of trauma during handling
202 (Jarvis and Cockram, 1994; Miranda-de la Lama et al., 2009), and thus constitute a sign of
203 poor welfare. Assessment can be performed once the hide has been removed. The location of
204 each bruise can be used to assess prior welfare; lesions on highly sensitive areas (e.g. face,
205 abdomen) could have highly deleterious effects on animal welfare. According to Cockram
206 and Lee (1991), the age of a bruise can be estimated by colour and consistency: a reddish and
207 haemorrhagic bruise is recent (0 – 10 h), dark-coloured and watery consistency is older (24-
208 48 h), yellowing >48 h. Hence, bruising might be a valid measure for assessing short-term
209 trauma that occurred in the few days or hours before death.
210
211 Diarrhoea is a clinical sign of gastrointestinal disease, especially in lambs, and can be
212 caused by diet or pathogens. Sweeny et al. (2012) define diarrhoea in sheep as the presence of
213 loose or liquid faeces, or faecal soiling on the breech fleece. Faecal consistency or staining of
214 the perineum can be used to assess diarrhoea. Cabaret et al. (2006) suggested the following
215 scoring system: 1, normal sheep faeces in pellets; 2, ‘soft’ faeces (similar to cow pat); and 3,
216 diarrhoea (semi-liquid faeces). Parasitic infections lead to reduced growth rate and wool
217 production, increased mortality (especially in young sheep), reduced reproductive success
218 and increased susceptibility to other parasites (Coop, 1979; Festa-Bianchet, 1988; Dwyer and
219 Bornett, 2004). Some parasitic infections, including *Echinococcus granulosus, Fasciola*
220 *hepatica, Dicrocoelium dendriticum*, *Cysticercus ovis,* and *Sarcocystis* spp., can affect the
221 muscles and viscera of sheep and can be detected by examination of the carcass and viscera

post-mortem (Borji et al., 2012). Round worms, however, are within the gastrointestinal lumen and so they are not readily detectable without opening the stomach or intestines.

The integument comprises the skin and modified skin structures, including head and hoof horn. Most integumentary structures are innervated. Damage is frequently caused by agonistic interactions with conspecifics and abrasions or collisions with physical structures (e.g. barbed wire fences, slatted doors), or by ingrowing head horn. New and old lesions will present differently e.g. dried blood and an open sore in recent injuries, to hyperkeratosis or hairless patches in older lesions. Stubsjøen et al. (2011) suggested two scoring systems for integumentary alterations, one for skin lesions and one for skin irritation. The skin lesion score is based on the following grading scale: 1, no skin lesions; 2, more than one lesion of >1 cm; and 3, ulceration present.

Ectoparasites can cause intense irritation to the skin and heavily infested sheep rub and bite affected areas (Plant, 2006). Dwyer and Bornett (2004) and Plant (2006) suggested that intense and regular rubbing and biting in localized areas could provoke skin lesions that help identify infected animals. For instance, *Psoroptes ovis* (sheep scab) produces intensely pruritic lesions and wool loss (Wells et al., 2013), while the lesions of the myiasis fly larvae (flystrike) induce inflammation, ulceration and wool loss (Hall and Wall, 1995). The scoring system for skin irritation validated by Stubsjøen et al. (2011) can be used to assess parasite-induced lesions and is as follows: 1, normal skin; 2, loss of wool regions; 3, redness/swelling of regions; and 4, presence of parasites or flies.

Lameness is one of the major welfare concerns of sheep according to farmers (Goddard et al., 2006). Eight papers (Table 1) describe gait assessment systems to categorise

locomotion in sheep. Kaler et al. (2009) proposed a valid and highly repeatable scale from 0 (normal) to 6 (unable to stand or walk) to assess lameness in sheep. Seven papers (Table 1) state that lameness can be assessed post-mortem by examining feet for lesions, once the lower limbs have been removed from the carcass. Hodgkinson (2010) described a systematic foot examination and compared all feet for subtle abnormalities and deformities. Scoring of foot lesions is repeatable between and within observers (Foddai et al., 2012); however, scoring mild abnormalities could overestimate the prevalence of lameness, because not all sheep with foot lesions are lame (Kaler et al., 2011).

It is self-evident that animals experience acute pain and distress at the time of mutilation and medium-term pain arising from tissue damage, with chronic pain also possible (Lomax et al., 2009; Edwards and Bennett, 2014). Mutilations can be assessed by visual inspection and in most cases provide reliable data (EFSA, 2012).

Eye condition was mentioned in two studies (Table 1). Assessment can be based on either inspection of the eyes, or ocular discharge. Blind sheep can only be detected ante-mortem. After death the eye becomes glazed and the eyelid droops, this restricts observation of some abnormalities including trauma, tumours or phthisis, which could be assessed post-mortem. Eye condition has not been validated as a welfare measure in sheep and so needs further investigation.

Coughing, dyspnoea and nasal discharge are signs of respiratory disease detected during clinical examination (Table 1). Post-mortem signs of respiratory disease include pulmonary inflammation or necrosis. These parameters have not been validated as welfare indicators, although lung lesions have been associated with increased age at finishing (Green

et al., 1995). Protocols scoring lung lesion type and severity post-mortem have been developed for cattle and pigs (Dalmau et al., 2009; Leruste et al., 2012) and such a system is required for sheep.

Napolitano et al. (2011) considered vulvar discharge as a measure of compromised health in female sheep. According to Aitken and Longbottom (2007), purulent vulvar discharge indicates pathology in reproductive organs and so it might be an indicator of reduced reproductive health. Although vulvar abnormalities have not been identified as measures of poor welfare (and so were not included as animal-based indicators of welfare; Table 1), the vulva is a highly sensitive area. Lovatt (2010) suggested that vulvar swelling, prolapse and injury were abnormalities to be considered during clinical examination of the reproductive organs in sheep. More evidence is needed to evaluate these as indicators of poor welfare.

Changes in general demeanour including lethargy and apathy are clinical signs of pain and systemic disease (Gougoulis et al., 2010). When these signs become severe, authors refer to animals that are 'obviously sick' (Mellor and Stafford, 2003; Stubsjøen et al., 2011), and this can be identified in a visual inspection. According to Gregory (1998), sickness is associated with listlessness, fatigue, reduced social interaction, inappetance, discomfort and mental confusion.

Experts suggest that any change from normal behaviour of an individual sheep can indicate a health or welfare problem (Phythian et al., 2011). The location and duration of assessments of behaviour affect the observations recorded and must be considered to avoid misinterpretation. The frequency of abnormal behaviour in farm animals can provide

297 information about their emotional state and welfare (Mason, 1991; Miranda-de la Lama et al.,
298 2012). Abnormal behaviours occur more frequently in animals living in confinement that
299 does not allow the expression of natural behaviours. Consequently, abnormal behaviours are
300 less frequent in sheep because they are less frequently confined (Dwyer and Bornett, 2004).
301
302 Forms of abnormal behaviour include stereotypies and redirected behaviour (Dwyer
303 and Bornett, 2004; Gougoulis et al., 2010). Wool biting or pulling are redirected behaviours
304 (Dwyer and Bornett, 2004), which could indicate a lack of environmental enrichment.
305 Stereotypic behaviour (i.e. mouthing bars, biting and chewing pen fixtures) is more frequent
306 in animals undergoing stress caused by maladaptation to their environment (Rushen and de
307 Passillé, 1992; Dwyer and Bornett, 2004).
308
309 Behaviours such as aggression or threats to pen mates have a negative impact on
310 welfare and misdirected behaviour and attacks on conspecifics are indicators of poor welfare
311 (Broom, 1988). Increased aggression (e.g. butting or chasing episodes), can be observed
312 during sudden environmental or social change in food, feed space or living space restrictions
313 and in large social group size (Dwyer and Bornett, 2004). Actions that harm other animals
314 can be recorded at the abattoir and might be an indicator of prior poor welfare. A sheep
315 showing normal behaviour for its sex, maturity (lamb vs. adult), or season, provides valuable
316 information about its current welfare. For example, lambs would be expected to show play
317 behaviour when not feeding or resting, while adult sheep would routinely ruminate for one
318 third of the day (Moquin et al., 2010). An interruption in expected behavioural pattern could
319 suggest welfare problems (Gougoulis et al., 2010). Qualitative Behaviour Assessment (QBA)
320 is a subjective list of behaviours that have been validated in sheep (Wemelsfelder and Farish,
321 2004), giving a final holistic estimate that takes into account all behaviours expressed. In

sheep, QBA has high repeatability between assessors (Phythian et al., 2013) and correlates significantly with physiological variables, including heart rate and stress leukogram results (Wickham et al., 2012).

A review of the relationship between social behaviour and welfare in goats suggests that in contrast to negative behaviours, affiliative behaviours, defined as positive, reciprocal behaviours between two or more individuals without reproductive interest, can improve the welfare state of a flock by helping to reduce aggression (Miranda-de la Lama and Mattiello, 2010). Two papers (Table 1) have proposed the identification of positive behaviours, such as nibbling or licking conspecifics, as a valid measure of good welfare in sheep.

Fear is one of the emotions that can severely influence the state of welfare of an animal. Stress is the biological response elicited when an individual perceives a threat to its homeostasis. The consequences of stress can be non-harmful, often referred as 'good stress' (e.g. caloric restriction in chronic hunger can promote longevity and better health), or distress, or negative stress, which weakens the immune system (Moberg, 2000). Direct observation of animal behaviour can provide a practical approach to the measurement of fear. In episodes of fear and anxiety, sheep might increase their vigilance behaviour, defined as head in an upright position and ears perpendicular to the head (Wemelsfelder and Farish, 2004; Deiss, et al., 2009). Fear has also been associated with frequency and duration of episodes of immobility, often referred as 'freezing' (Bouissou and Vandenheede, 1995; Cockram, 2004), and with ear-posture changes (Reefmann et al., 2009). Hemsworth et al. (2011) reported a correlation between head position and serum cortisol concentration when sheep were approached by a stockperson. The most frequently cited measure to assess fear in sheep is increased vocalisation (Table 1). In livestock species other than sheep (pigs, poultry

and cattle), high-pitched vocalisations are thought to signal appeasement in fear-associated contexts, whereas low-pitched vocalisations are attributed to more aggressive emotions (Manteuffel et al., 2004). In sheep, vocalisations can occur in response to numerous situations including social isolation, social attraction, and the presence of humans (Cockram, 2004; Boissy et al., 2005 and Deiss et al., 2009). Since vocalisation is performed in numerous situations other than fear, its suitability as a measure of fear requires further analysis. Wemelsfelder and Farish (2004) reviewed a set of qualitative categories, including fearfulness, for the assessment of sheep behaviour by direct observation of the whole flock. This was developed into a QBA protocol, which has been validated for the assessment of sheep welfare (Wickham et al., 2012). Although good reliability and repeatability have been obtained in overall QBA scores between assessors (Wemelsfelder and Farish, 2004), there is no information on whether this tool is reliable when only one category is assessed, because this was not its purpose.

The human-animal relationship (HAR) is a major determinant of sheep welfare because it is an important source of fear in farmed sheep (Waiblinger et al., 2006). This is particularly pertinent to extensive systems with limited interactions with people (Turner and Dwyer, 2007). The degree of aversion to human handling can also be influenced by the quality and sensitivity of the animals' interaction with the farm stockperson (Dwyer, 2009) and therefore, presumably, abattoir staff. Thus, inferences can also be drawn about social attachment to humans, the nature (positive, neutral or negative) of past experiences with people and the quality of stockmanship (Waiblinger et al., 2006). In sheep, HAR has been measured using alertness to human approach in the field, escape attempts and an avoidance distance test (Table 1). In a comprehensive review of fear tests in farm animals, Forkman et al. (2007) demonstrated that fear of humans correlated with increased heart rate. While these

372 tests had good repeatability, the review also concluded that more evidence was needed to
373 confirm their validity. The HAR is also affected by other variables, including the type of
374 farming (intensive vs. extensive; Turner and Dwyer, 2007). This impairs its validity unless it
375 is carried out in a controlled environment, which is not compatible with commercial abattoir
376 conditions.
377
378 Published reviews have reported the deleterious effects of pre-slaughter stress on meat
379 quality in ruminants (Ferguson and Warner, 2008), including sheep (Sañudo et al., 1998). The
380 organic changes occurring during pre-slaughter stress can lead to rapid decline in pH in
381 muscle due to increased ATPase activity and lactate accumulation (Monin, 1988; Liste et al.,
382 2011). This increases the rate of protein denaturation post-mortem and reduces water holding
383 capacity (WHC) of muscle, leading to dark, dry meat (Bond et al., 2004; Ferguson and
384 Warner, 2008). The most cited meat quality indicators of pre-slaughter stress are pH and meat
385 colour (Table 1), although the assessment of tenderness and WHC have also been used (Liste
386 et al., 2011). Further research is needed to validate meat tenderness and WHC as indicators of
387 prior distress in sheep.
388
389 The animal-based indicators described above and listed in Table 1 were developed for
390 the assessment of sheep welfare at farm, during transport, or at the abattoir. From our
391 assessment, 19 indicators can be regarded as high validity indicators from previous research
392 work (Table 1); 13 are of medium validity and need further research, but four of these have
393 high feasibility for measurement in abattoirs (conformation carcass classification, fat carcass
394 classification, meat colour, meat pH); 15 are of low validity. Welfare indicators validated in
395 environments other than the abattoir might be invalid in an abattoir. A summary of the factors
396 that might affect validity or feasibility of the measurements is discussed below.

Live sheep in abattoirs are highly likely to be experiencing some degree of stress. Stress probably occurs during transport and market and might be increased further by mixing or close proximity of unfamiliar sheep. On arrival at the abattoir, during unloading and penning, sheep move rapidly en masse because of their flocking instinct and this increases stress, thereby reducing the validity of indicators that require observation of animals in a consistent environment (e.g. stereotypies, QBA). Additionally, there is often a restricted view of the whole animal (i.e. the torso, belly and legs are difficult to observe in tightly packed sheep), both when sheep are being moved and when penned, reducing the feasibility of measuring indicators that require observation of the whole animal (e.g. gait, vulvar discharge). Abattoir policy is to avoid handling sheep ante-mortem to minimise bruising, reducing the feasibility of taking measurements that require touching sheep (e.g. body condition scoring, close inspection of fleece, feet, eyes etc.).

Post-mortem, there is the potential to inspect the external surface of the carcass, with and without the hide, and the internal organs and carcass. However, the dressing line often moves rapidly and there is separation of carcass from hide and internal organs early in processing that can reduce the traceability between parts of the sheep. Abattoirs minimise handling of animals post-mortem, especially hide, lower limb, head and gastrointestinal organs, to maintain high standards of hygiene. Therefore, although one possible advantage of post-mortem inspection is that carcasses, organs, hide, head and feet could be examined thoroughly after a batch of animals has been processed, this increases the risk of meat contamination and would require a separate space from the line and different personnel to make the inspections. This reduces the feasibility of observing these at abattoirs.

422 Finally, to assess prior welfare, an indicator needs to be present for a period of time,
423 so indicators that are highly variable temporarily are unlikely to be valid indicators of prior
424 welfare (e.g. panting, shivering and indicators of fear and distress and normal behaviour). In
425 this study, high validity indicators were initially considered by category to minimise
426 repetition and maximise critical comparison; where an indicator was defined as having high
427 validity in its original setting, but did not appear likely to be valid or feasible when measured
428 in an abattoir, we considered the possible use of medium validity indicators with high
429 feasibility, or alternative novel technologies.
430
431 Low BCS indicates prior long term poor welfare. Scoring BC is not possible in
432 abattoirs because handling of lambs and ewes ante-mortem is not permitted. Carcass
433 classification is an alternative to BCS and is likely to be more reliable than BCS in lambs.
434 Video image analysis (VIA) technologies have been tested to provide an automated carcass
435 classification (Rius-Vilarrasa et al., 2009; Einarsson et al., 2014), which might improve
436 feasibility and validity of the assessment compared with more subjective evaluation using the
437 EUROP classification system. Carcass classification was rated as medium validity in our
438 study, because while it can be used to assess body condition of adult sheep (typically cull
439 ewes), it was originally developed for slaughter lambs[8]. Therefore, the method needs more
440 validation, as would VIA, before being used to assess adults in abattoirs. Such validation is
441 important because emaciation (defined as BCS<1.5) should be correlated with carcass
442 conformation and fat grades or VIA. In cattle, there is significant correlation between BCS
443 and carcass grade (Emenheiser et al., 2014), but this correlation is lower at low and high BCS

[8] See: Anderson, J., 2003. Planned carcase production. Sheep management matters – A series on sheep management topics from the meat and livestock commission. http://www.mlcsl.co.uk/pdf/Planned%20Carcase%20Production.pdf (accessed 23 September 2015).

444 values (Apple, 1999; Apple et al., 1999). Carcass classification has the advantage of not
445 causing bruising or distress to a live animal. However, assessment must take place on the
446 processing line because carcasses are trimmed before being passed fit for human
447 consumption and chilled.
448
449 Rumen fill indicates whether sheep were hungry prior to slaughter. It can be assessed
450 by palpation of the left abdominal wall in live sheep or by inspection of the rumen or
451 weighing rumen contents post-mortem. Palpation of live animals in abattoirs is not
452 acceptable, as explained previously. Inspecting the rumen requires handling intestinal
453 material, which is also avoided in abattoirs to avoid contamination. Therefore, to assess short-
454 term hunger in the abattoir, either new indicators would need to be developed or a dedicated
455 individual would have to be present to handle intestinal material in a separate area from the
456 line. As mentioned previously, there is currently no validated measure to assess thirst in
457 sheep and skin-tenting time would not be acceptable in abattoirs because it requires handling
458 of sheep.
459
460 Body cleanliness and ectoparasite infestations could be observed both ante- and post-
461 mortem. Body cleanliness could indicate poor welfare at any point prior to arrival and at the
462 abattoir, while ectoparasite infestations are most likely to have occurred on farm.
463
464 Indicators such as shivering or panting provide information on the thermal comfort at
465 the moment they are assessed, but they are not valid measures of prior thermal comfort .
466 Therefore, this highly valid indicator is not useful to assess prior welfare. The possible
467 exception is fleece cover, specifically excess or insufficient fleece observed in sheep in very
468 hot or very cold weather, respectively. This low validity indicator might indicate prior

thermal discomfort on farm, at market or during transit to the abattoir. Further research is required to evaluate fleece cover as an indicator of prior thermal comfort.

Disease states can be acute or chronic. Some signs of disease will indicate prior welfare (e.g. dried faecal staining, skin irritation), but many signs of disease might be acute in onset (e.g. lameness, dyspnoea) and so of limited use in the investigation of prior welfare. Some clinical signs of disease must be assessed ante-mortem (e.g. general demeanour, obviously sick, coughing, nasal discharge, dyspnoea, diarrhoea, lameness [gait score], ocular health, scratching or rubbing, vulvar discharge). These could be assessed without touching the animal and so could be feasible to assess at abattoirs, but sufficient time and space are needed to make appropriate observations of the whole animal. Obviously sick animals would also be detected post-mortem because the carcass would not set (undergo normal post-mortem changes, including rigor mortis).

Disease indicators that require inspection of the skin or fleece (skin lesions or irritation, faecal staining) could be performed ante-mortem without handling, immediately post-mortem, or after the fleece has been removed in a separate area. Large skin lesions could be observed, but small lesions (<2 cm diameter) and lesion depth are more difficult to identify; however, they have important implications for animal welfare. Such indicators are valid to assess prior sheep welfare.

Old injuries and bruising would indicate poor prior welfare. Some injuries would be detectable ante-mortem by examining animals, subject to good visibility. The extent of injuries might only become apparent once the hide is removed. Bruising could be recorded

493 post-mortem. Since it is included in assessment protocols for welfare in cattle [9] (WQ, 2009),
494 it is likely that it could be assessed in routine carcass assessments in sheep at abattoirs.
495 Recording of bruising would require dedicated staff on the line because carcasses are
496 trimmed before grading and many bruises would be removed.
497
498 Mutilations would typically be incurred on farm and so are indicators of prior poor
499 welfare. They can be observed ante-mortem (tail docking, short tail docking, ear notching), or
500 early post-mortem (castration, tail docking, ear notching).
501
502 While some behavioural assessments are considered more valid than others, all
503 assessments of behaviour, including QBA, are unlikely to reflect past welfare. This is because
504 animals arriving from dealers, collection centres and markets might have been mixed and so
505 their behaviour will alter. Moreover, an abattoir is a novel situation and it is likely that in
506 such conditions sheep behaviour would not correspond to their normal behaviour on farm.
507 Thus, all the indicators reviewed that assess behaviour can be considered valid indicators for
508 welfare at the time of inspection at the abattoir, but it is unlikely they would provide valid
509 information about prior welfare. Similarly, indicators that assess the animal's current mental
510 state (typically those described under freedom from fear and distress) are influenced by the
511 animal's response to a new environment, novel humans and sheep. Consequently, when
512 assessed at the abattoir they might assess current welfare, but they are unlikely to provide an
513 accurate reflection of past freedom from fear and distress. In contrast, indicators of meat

[9] See: Welfare Quality, 2009. Welfare Quality assessment protocol for cattle. Welfare Quality Consortium, Lelystad, Netherlands. http://www.welfarequality.net/network/45848/7/0/40 (Accessed 23 September 2015).

514 quality as a measure of distress in carcasses might be more indicative of recent past distress,
515 although further research is required to validate these.
516
517 This study is the first phase of a project that aims to identify animal-based welfare
518 indicators of prior welfare of sheep that could be measured in commercial abattoirs in the
519 UK. The Five Freedoms was chosen as a framework to explore existing animal-based
520 indicators because it covers all aspects of welfare. We hypothesised that if at least one
521 indicator within a freedom could be identified for use in abattoirs, then prior welfare on that
522 indicator could be assessed at abattoirs (e.g. freedom from hunger and thirst would require
523 indicators for hunger and thirst). While some indicators for freedom from pain, injury and
524 disease could be measured, others were not considered feasible, and some of those might be
525 important. Gait score is one example; lameness is common in sheep but the current layout
526 and movement of sheep in some abattoirs makes observation of sheep walking impossible.
527 The next stage in the development of a complete list of valid, feasible indicators that could be
528 used to assess prior welfare in abattoirs is to test the indicators listed above to investigate
529 whether they can be recorded with high reliability and repeatability at abattoirs as at their site
530 of development and to investigate whether there are novel indicators that can be developed to
531 assess prior animal welfare for all Five Freedoms. The ultimate aim is to have a complete set
532 of indicators to benchmark the prevalence of welfare indicators to inform risk-based selection
533 for inspections measuring compliance with animal welfare legislation, to evaluate whether an
534 assurance scheme's welfare conditions are met and to provide farmers with information on
535 the health and welfare of their livestock to assist in health planning.
536
537 **Conclusions**

This systematic review identified some existing high validity indicators that might be useful to assess prior welfare of sheep at abattoirs. These were body cleanliness, carcass bruising, diarrhoea, skin lesions, skin irritation, castration, ear notching, tail docking and obviously sick. In addition, four medium validity, high feasibility, indicators were considered and two of those (carcass and fat classification) were considered useful to take forward. Other high validity indicators could not be used to measure prior welfare because either they were situation-specific (freedom to express normal behaviour or freedom from fear and distress) and measure current or recent welfare, or they were not feasible to measure in abattoirs. No indicators were identified that measured prior freedom to express normal behaviour or freedom from fear and distress. Freedom from discomfort could only be identified through scoring body cleanliness. For a set of indicators for all Five Freedoms to be completely robust, the potentially useful welfare indicators identified need further investigation to test their validity at abattoirs. Some novel welfare indicators are also necessary, assuming that measurement of historic fear and distress or historic ability to express normal behaviour could be measured.

Acknowledgements

This research was funded by the Department for Environment, Food and Rural Affairs (Grant code AW1028). We thank the two anonymous reviewers for the useful comments and suggestions and Dr Corinna Clark for proof reading the paper.

Conflict of interest statement

None of the authors of this paper has a financial or personal relationship with other people or organisations that could inappropriately influence or bias the content of the paper.

References

Aitken, I.D., Longbottom, D., 2007. Chlamydial abortion. In: Aitken, I.D. (Ed.), Diseases of Sheep. Blackwell Publishing, Oxford, UK, pp. 105-111.

Apple, J.K., 1999. Influence of body condition score on live and carcass value of cull beef cows. Journal of Animal Science 77, 2610-2620.

Apple, J.K., Davis, J.C., Stephenson, J., Hankins, J.E., Davis, J.R., Beaty, S.L., 1999. Influence of body condition score on carcass characteristics and subprimal yield from cull beef cows. Journal of Animal Science 77, 2660-2669.

Boissy, A., Bouix, J., Orgeur, P., Poindron, P., Bibé , B., Le Neindre, P., 2005. Genetic analysis of emotional reactivity in sheep: effects of the genotypes of the lambs and of their dams. Genetics Selection Evolution 37, 381-401.

Boissy, A., Manteuffel, G., Jensen, M.B., Moe, R.O., Spruijt, B., Keeling, L.J., Winckler, C., Forkman, B., Dimitrov, I., Langbein, J., et al., 2007. Assessment of positive emotions in animals to improve their welfare. Physiology and Behaviour 92, 375-397.

Bond, J.J., Can, L.A., Warner, R.D., 2004. The effect of exercise stress, adrenaline injection and electrical stimulation on changes in quality attributes and proteins in Semimembranosus muscle of lamb. Meat Science 68, 469-477.

Bouissou, M.F., Vandenheede, M., 1995. Fear reactions of domestic sheep confronted with either a human or a human-like model. Behavioural Processes 34, 81-91.

Borji, H., Azizzadeh, M., Kamelli, M., 2012. A retrospective study of abattoir condemnation due to parasitic infections: economic importance in Ahwaz, southwestern Iran. Journal of Parasitology 98, 954-957.

Broom, D.M., 1988. The scientific assessment of animal welfare. Applied Animal Behaviour Science 20, 5-19.

Cabaret, J., Gonnord, V., Cortet, J., Sauvé, C., 2006. Indicators for internal parasitic infections in organic flocks: The diarrhoea score (Disco) proposal for lambs. Proceedings of Joint Organic Congress, Odense, Denmark, May 30-31, 2006.

Caroprese, M., Casamassima, D., Rassu, S.P.G., Napolitano, F., Sevi, A., 2009. Monitoring the on-farm welfare of sheep and goats. Italian Journal of Animal Science 8, 343-354.

Cockram, M.S., Lee, R.A., 1991. Some preslaughter factors affecting the occurrence of bruising in sheep. British Veterinary Journa1 147, 120-125.

Cockram, M.S., 2004. A review of behavioural and physiological responses of sheep to stressors to identify potential behavioural signs of distress in sheep. Animal Welfare 13, 283-291.

612 Cockram, M.S., Murphy, E., Ringrose, S., Wemelsfelder, F., Miedema, H.M., Sandercock,
613 D.A., 2012. Behavioural and physiological measures following treadmill exercise as
614 potential indicators to evaluate fatigue in sheep. Animal 6, 1491-1502.
615
616
617 Coop, R.L., 1979. Production loss in sub-clinical helminth infections. Veterinary Record 105,
618 189.
619
620 Dalmau, A., Temple, D., Rodriguez, P., Llonch, P., Velarde, A., 2009. Application of the
621 Welfare Quality protocol at pig slaughterhouses. Animal Welfare 18, 497-505.
622
623 Deiss, V., Temple, D., Ligout, S., Racine, C., Bouix, J., Terlouw, C., Boissy, A., 2009. Can
624 emotional reactivity predict stress responses at slaughter in sheep? Applied Animal
625 Behaviour Science 119, 193-202.
626
627 Dikmen, S., Orman, A., Ustuner, H., 2011. The effect of shearing in a hot environment on
628 some welfare indicators in Awassi lambs. Tropical Animal Health Production 43,
629 1327-1335.
630
631 Dolan, S., Nolan, A.M., 2000. Behavioural evidence supporting a differential role for group I
632 and II metabotropic glutamate receptors in spinal nociceptive transmission.
633 Neuropharmacology 39, 1132-1138.
634
635 Dwyer, C.M., Bornett, H.L.L., 2004. Chronic stress in sheep: Assessment tools and their use
636 in different management conditions. Animal Welfare 13, 293-304.
637
638 Dwyer, C.M., 2008. The welfare of the neonatal lamb. Small Ruminant Research 76, 31-41.
639
640 Dwyer, C.M., 2009. Welfare of sheep: providing for welfare in an extensive environment.
641 Small Ruminant Research 86, 14-21.
642
643 Edwards, S., Bennett, P., 2014. Tales about tails: Is the mutilation of animals justifiable in
644 their best interests or in ours? In: Dilemmas in animal welfare, First Ed. CABI,
645 Boston, USA, pp 6-27.
646
647 European Food Safety Authority (EFSA), 2012. Scientific Opinion on the use of animal-
648 based measures to assess welfare in swine. EFSA Journal 10, 2512-2597.
649
650 Einarsson, E., Eythórsdóttir, E., Smith, C.R., Jónmundsson, J.V., 2014. The ability of video
651 image analysis to predict lean meat yield and EUROP score of lamb carcasses.
652 Animal 8, 1170-1177.
653
654 Emenheiser, J.C., Tait, R.G., Shackelford, S.D., Kuehn, L.A., Wheeler, T.L., Notter, D.R.,
655 Lewis, R.M., 2014. Use of ultrasound scanning and body condition score to evaluate
656 composition traits in mature beef cows. Journal of Animal Science 92, 3868-3877.
657
658 Farm Animal Welfare Council, 1993. Second report on priorities for research and
659 development in farm animal welfare. Ministry of Agriculture, Fisheries and Food.
660

661 Ferguson, D.M., Warner, R.D., 2008. Have we underestimated the impact of pre-slaughter
662 stress on meat quality in ruminants?. Meat Science 80, 12-19.
663
664 Festa-Bianchet, M., 1988. Nursing behaviour of bighorn sheep: correlates of ewe age,
665 parasitism, lamb age, birthdate and sex. Animal Behaviour 36, 1445-1454.
666
667 Fitzpatrick, J., Scott, M., Nolan, A., 2006. Assessment of pain and welfare in sheep. Small
668 Ruminant Research 62, 55-61.
669
670 Foddai, A., Green, L.E., Mason, S.A., Kaler., J., 2012. Evaluating observer agreement of
671 scoring systems for foot integrity and footrot lesions in sheep. BMC Veterinary
672 Research 8, 189-94.
673
674 Forkman, B., Boissy, A., Meunier-Salaün, M.C., Canali, E., Jones, R.B., 2007. A critical
675 review of fear tests used on cattle, pigs, sheep, poultry and horses. Physiology and
676 Behavior 92, 340-374.
677
678 Gougoulis, D.A., Kyriazakis, I., Fthenakis, G.C., 2010. Diagnostic significance of behaviour
679 changes of sheep: A selected review. Small Ruminant Research 92, 52-56.
680
681 Goddard, P., Waterhouse, Y., Dwyer, C., Stott, A., 2006. The perception of welfare sheep in
682 extensive systems. Small Ruminant Research 62, 215-225.
683
684 Green, L.E., Berriatua, E., Cripps, P.J., Morgan, K.L., 1995. Lesions in finished early born
685 lambs in southwest England and their relationship with age at slaughter. Preventive
686 Veterinary Medicine 22, 115-126.
687
688 Greiveldinger, L., Veissier, I., Boissy, A., 2007. Emotional experience in sheep:
689 Predictability of a sudden event lowers subsequent emotional responses. Physiology
690 and Behavior 92, 675-683.
691
692 Gregory, N.G., 1998. Physiological mechanisms causing sickness behaviour and suffering in
693 diseased animals. Animal Welfare 7, 293-305.
694
695 Hales, J.R.S., Brown, G.D., 1974. Net energetic and thermoregulatory efficiency during
696 panting in the sheep. Comparative Biochemistry and Physiology 49, 413-422.
697
698 Hall, M., Wall, R., 1995. Myiasis of humans and domestic animals. Advances in Parasitology
699 35, 257-334.
700
701 Hemsworth, P.H., Rice, M., Karlen, M.G., Calleja, L., Barnett, J.L., Nash, J., Coleman, G.J.,
702 2011. Human–animal interactions at abattoirs: Relationships between handling and
703 animal stress in sheep and cattle. Applied Animal Behaviour Science 135, 24-33.
704
705 Hodgkinson, O., 2010. The importance of feet examination in sheep health management.
706 Small Ruminant Research 92, 67-71.
707
708 Jarvis, A.M. Cockram, M.S., 1994. Effects of handling and transport on bruising of sheep
709 sent directly from farms to slaughter. Veterinary Record 135, 523-527.
710

711 Jarvis, A.M., Cockram, M.S., McGilp, I.M., 1996. Bruising and biochemical measures of
712 stress, dehydration and injury determined at slaughter in sheep transported from
713 farms or markets. British Veterinary Journal 152, 719-722.
714
715 Jefferies, B.C., 1961. Body condition scoring and its use in management. Tasmanian Journal
716 of Agriculture 32, 19-21.
717
718 Kaler, J., Wassink, G.J., Green, L.E., 2009. The inter- and intra-observer reliability of a
719 locomotion scoring scale for sheep. The Veterinary Journal 180, 189-194.
720
721 Kaler, J., George, T.R.N., Green, L.E., 2011. Why are sheep lame? Temporal associations
722 between severity of foot lesions and severity of lameness in 60 sheep. Animal
723 Welfare 20, 433-438.
724
725 Kilbride, A.L., Mason, S.A., Honeyman, P.C., Pritchard, D.G., Hepple, S., Green, L.E., 2012.
726 Associations between membership of farm assurance and organic certification
727 schemes and compliance with animal welfare legislation. The Veterinary record 170,
728 152-157.
729
730 King, E.M., Green, L.E., 2011. Assessment of farmer recognition and reporting of lameness
731 in adults in 35 lowland sheep flocks in England. Animal Welfare 20, 321-328.
732
733 Knierim, U., Winckler, C., 2009. On-farm welfare assessment in cattle: Validity, reliability
734 and feasibility issues and future perspectives with special regard to the Welfare
735 Quality approach. Animal Welfare 18, 451-458.
736
737 Leruste, H., Brscic, M., Heutinck, L.F.M., Visser, E.K., Wolthuis-Fillerup, M., Bokkers,
738 E.A.M., Stockhofe-Zurwieden, N., Cozzi, G., Gottardo, F., Lensink, B.J., et al.,
739 2012. The relationship between clinical signs of respiratory system disorders and
740 lung lesions at slaughter in veal calves. Preventive Veterinary Medicine 105, 93-100.
741
742 Liste, G., Miranda-de la Lama, G.C., Campo, M.M., Villarroel, M., Muela, E., Maria, G.A.,
743 2011. Effect of lairage on lamb welfare and meat quality. Animal Production
744 Science 51, 952-958.
745
746 Lomax, S., Sheil, M., Windsor, P.A., 2009. Use of local anaesthesia for pain management
747 during husbandry procedures in Australian sheep flocks. Small Ruminant Research
748 86, 56-58.
749
750 Lovatt, F.M., 2010. Clinical examination of sheep. Small Ruminant Research 92, 72-77.
751
752 Main, D.C.J., Whay, H.R., Lee, C., Webster, A.J.F., 2007. Formal animal-based welfare
753 assessment in UK certification schemes. Animal Welfare 16, 233-236.
754
755 Manteuffel, G., Puppe, B., Schön, P., 2004. Vocalization of farm animals as a measure of
756 welfare. Applied Animal Behaviour Science 88, 163-182.
757
758 Mason, G.J., 1991. Stereotypies: A critical review. Animal Behaviour 41, 1015-1037.
759

760 Mattiello, S., Battini, M., Andreoli, E., Minero, M., Barbieri, S., Canali, E., 2010. Avoidance
761 distance test in goats: A comparison with its application in cows. Small Ruminant
762 Research 91, 215-218.
763
764 McCulloch, S.P., 2013. A critique of FAWC's five freedoms as a framework for the analysis
765 of animal welfare. Journal of Agricultural and Environmental Ethics 26, 959-975.
766
767 Mellor, D.J., Stafford, K.J., 2003. Assessing the welfare status of newborn farm animals.
768 Animal Welfare 12, 695-698.
769
770 Mellor, D.J., Stafford, K.J., 2004. Animal welfare implications of neonatal mortality and
771 morbidity in farm animals. The Veterinary Journal 168, 118-133.
772
773 Miranda-de la Lama, G.C., Villarroel, M., Olleta, J.L., Alierta, S., Sañudo, C., Maria, G.A.,
774 2009. Effect of the pre-slaughter logistic chain on meat quality of lambs. Meat
775 Science 83, 604-609.
776
777 Miranda-de la Lama, G.C., Mattiello, S., 2010. The importance of social behaviour for goat
778 welfare in livestock farming. Small Ruminant Research 90, 1-10.
779
780 Miranda-de la Lama, G.C., Villarroel, M., Maria, G.A., 2012. Behavioural and physiological
781 profiles following exposure to novel environment and social mixing in lambs. Small
782 Ruminant Research 103, 158-163.
783
784 Moberg, G.P., 2000. Biological response to stress: Implications for animal welfare. In: The
785 Biology of Animal Stress: Basic Principles and Implications for Animal Welfare,
786 First Ed. CABI, New York, USA, p. 1-22.
787
788 Monin, G., 1988. Stress d'abattage et qualités de la viande. Recueil de Médecine Vétérinaire
789 164, 835-842.
790
791 Moquin, P., Curry, B., Pelletier, F., Ruckstuhl, K.E., 2010. Plasticity in the rumination
792 behaviour of bighorn sheep: Contrasting strategies between the sexes? Animal
793 Behaviour 79, 1047-1053.
794
795 Morgan-Davies, C., Waterhouse, A., Pollock, M.L., Milner, J.M., 2008. Body condition score
796 as an indicator of ewe survival under extensive conditions. Animal Welfare 17, 71-
797 77.
798
799 Napolitano, F., Caroprese, M., Girolami, A., Marino, R., Muscio, A., Sevi, A., 2006. Effects
800 of early maternal separation of lambs and rearing with minimal and maximal human
801 contact on meat quality. Meat Science 72, 635-640.
802
803 Napolitano, F., De Rosa, G., Ferrante, V., Grasso, F., Braghieri, A., 2009. Monitoring the
804 welfare of sheep in organic and conventional farms using an ANI 35 L derived
805 method. Small Ruminant Research 83, 49-57.
806
807 Napolitano, F., De Rosa, G., Girolami, A., Scavone, M., Braghieri, A., 2011. Avoidance
808 distance in sheep: Test–retest reliability and relationship with stockmen attitude.
809 Small Ruminant Research 99, 81– 86.

Nowak, R., Porter, R.H., Blache, D., Dwyer, C.M., 2008. Behaviour and the welfare of sheep. In: The Welfare of Sheep. First Ed. Springer, The Netherlands.

Phythian, C.J., Michalopoulou, E., Jones, P.H., Winter, A.C., Clarkson, M.J., Stubbings, L.A., Grove-White, D., Cripps, P.J., Duncan, J.S., 2011. Validating indicators of sheep welfare through a consensus of expert opinion. Animal 5, 943-952.

Phythian, C.J., Hughes, D., Michalopoulou, E., Cripps, P.J., Duncan, J.S., 2012a. Reliability of body condition scoring of sheep for cross-farm assessments. Small Ruminant Research 104, 156-162.

Phythian, C.J., Cripps, P.J., Michalopoulou, E., Jones, P.H., Grove-White, D., Clarkson, M.J., Winter, A.C., Stubbings, L.A., Duncan, J.S., 2012b. Reliability of indicators of sheep welfare assessed by a group observation method. The Veterinary Journal 193, 257-263.

Phythian, C., Michalopoulou, E., Duncan, J., Wemelsfelder, F., 2013. Inter-observer reliability of qualitative behavioural assessments of sheep. Applied Animal Behaviour Science 144, 73-79.

Plant, J.W., 2006. Sheep ectoparasite control and animal welfare. Small Ruminant Research 62, 109-112.

Pritchard, J.C., Burn, C.C., Barr, A.R.S., Whay, H.R., 2008. Validity of indicators of dehydration in working horses: A longitudinal study of changes in skin tent duration, mucous membrane dryness and drinking behaviour. Equine Veterinary Journal 40, 558-564.

Reefmann, N., Bütikofer, K.F., Wechsler, B., Gygax, L., 2009. Ear and tail postures as indicators of emotional valence in sheep. Applied Animal Behaviour Science 118, 199-207.

Rius-Vilarrasa, E., Bünger, L., Maltin, C., Matthews, K.R., Roehe, R., 2009. Evaluation of Video Image Analysis (VIA) technology to predict meat yield of sheep carcasses on-line under UK abattoir conditions. Meat Science 82, 94-100.

Rushen, J., de Passillé, A.M.B., 1992. The scientific assessment of the impact of housing on animal-welfare - a critical review. Canadian Journal of Animal Science 72, 721-743.

Sargison, N.D., Scott, P.R., 2010. The implementation and value of diagnostic procedures in sheep health management. Small Ruminant Research 92, 2-9.

Sañudo, C., Sanchez, A., Alfonso, M., 1998. Small ruminant production systems and factors affecting lamb meat quality. Meat Science 49, S29-S64.

Scott, P.R., 2011. Treatment and control of respiratory disease in sheep. Veterinary Clinics of North America: Food Animal Practice 27, 175-186.

Scott, E.M., Fitzpatrick, J.L., Nolan, A.M., Reid, J., Wiseman, M.L., 2003. Evaluation of welfare state based on interpretation of multiple indices. Animal Welfare 12, 457-468.

Shands, C.G., Mcleod, B., Lollback, M.L., Duddy, G., Hatcher, S., O'Halloran, W.J., 2009. Comparison of manual assessments of ewe fat reserves for on-farm use. Animal Production Science 49, 630-636.

Silanikove, N., 2000. Effects of heat stress on the welfare of extensively managed domestic ruminants. Livestock Production Science 67, 1-18.

Stubsjøen, S.M., Hektoen, L.A., Valle, P.S., Janczak, A.M., Zanella, A.J., 2011. Assessment of sheep welfare using on-farm registrations and performance data. Animal Welfare 20, 239-251.

Sweeny, J.P.A., Ryan, U.M., Robertson, I.D., Jacobson, C., 2012. Prevalence and on-farm risk factors for diarrhoea in meat lamb flocks in Western Australia. The Veterinary Journal 192, 503-510.

Teixeira, D.L., Miranda-de la Lama, G.C., Villarroel, M., Garcia-Belenguer, S., Sañudo, C., Maria, G.A., 2012. Effect of straw on lamb welfare, production performance and meat quality during the finishing phase of fattening. Meat Science 92, 829-836.

Thornton, P.D., Waterman-Pearson, A.E., 2002. Behavioural responses to castration in lambs. Animal Welfare 11, 203-212.

Turner, S.P., Dwyer, C.M., 2007. Welfare assessment in extensive animal production systems: Challenges and opportunities. Animal Welfare 16, 189-192.

van Burgel, A.J., Oldham, C.M., Behrendt, R., Curnow, M., Gordon, D.J., Thompson, A.N., 2011. The merit of condition score and fat score as alternatives to liveweight for managing the nutrition of ewes. Animal Production Science 51, 834-841.

Waiblinger, S., Boivin, X., Pedersen, V., Tosi, M.V., Janczak, A.M., Visser, E.K., Jones, R.B., 2006. Assessing the human-animal relationship in farmed species: A critical review. Applied Animal Behaviour Science 101, 185-242.

Wells, B., Innocent, G.T., Eckersall, P.D., McCulloch, E., Nisbet, A.J., Burgess, S.T., 2013. Two major ruminant acute phase proteins, haptoglobin and serum amyloid A, as serum biomarkers during active sheep scab infestation. Veterinary Research 44, 103.

Wemelsfelder, F., Farish, M., 2004. Qualitative categories for the interpretation of sheep welfare: A review. Animal Welfare 13, 261-268.

Wickham, S.L., Collins, T., Barnes, A.L., Miller, D.W., Beatty, D.T., Stockman, C., Blache, D., Wemelsfelder, F., Fleming, P.A., 2012. Qualitative behavioral assessment of transport-naïve and transport-habituated sheep. Journal of Animal Science 90, 4523-4535.

Table 1

Animal-based indicators of sheep welfare derived from a systematic review, classified by the Five Freedoms and their validity and likely feasibility for use in abattoirs to assess prior sheep welfare

Freedom	Category	Indicator	Observation (ante-mortem, A; post-mortem, P)	Validity [a]	Feasibility in abattoir [b]	References (first author and year of publication) [c]
From hunger and thirst [d]	Body condition	Bodyweight	A or P	Low	High	Phythian, 2011
		Body condition score (BCS)	A or P	High	Low	Morgan-Davies, 2008; Caroprese, 2009; Dwyer, 2009; Napolitano, 2009; Phythian, 2011; Stubsjoen, 2011; van Burgel, 2011; Phythian, 2012a
		Conformation carcass classification	P	Medium in finished lambs	High	Stubsjoen, 2011
		Fat carcass classification	P	Medium	High	Stubsjoen, 2011
	Access to feed	Rumen fill	P	Medium	Low	Phythian, 2011
From discomfort	Cleanliness	Body cleanliness [e]	A or P	High	High	Caroprese, 2009; Napolitano, 2009; Phythian, 2011; Stubsjoen, 2011; Phythian, 2012a; Phythian, 2012b;
	Resting comfort	Lying behaviour	A	Low	Low	Dikmen, 2011; Phythian, 2011; Stubsjoen, 2011; Cockram, 2012
	Thermal comfort	Fleece cover	A or P	Low	High	Phythian, 2011
		Shivering	A	High in lambs	Low	Mellor, 2003
		Panting	A	High	Low	Thornton, 2002; Cockram, 2004; Gougoulis, 2010; Phythian, 2011; Phythian, 2012b
From pain, injury and disease	Bruises	Carcass bruising [e]	P	High	High	Jarvis, 1996; Miranda-de la Lama, 2010; Liste, 2011; Teixeira, 2012
	Gastrointestinal health	Diarrhoea [e]	A	High in lambs	High	Napolitano, 2011; Stubsjoen, 2011
		Endoparasitism	P	High	Medium	Dwyer, 2004
	Integument alterations	Skin lesions [e]	A or P	High	High	Scott, 2003; Fitzpatrick, 2006; Dwyer, 2008; Caroprese, 2009; Napolitano, 2009, 2011; Phythian, 2011; Stubsjoen, 2011
		Skin irritation [e]	A or P	High	High	Fitzpatrick, 2006; Plant, 2006; Stubsjoen, 2011; Phythian, 2012b

	Lameness	Foot lesions	P	Medium	Medium	Fitzpatrick, 2006; Caroprese, 2009; Hodgkinson, 2010; Scott, 2003; Napolitano, 2009; Stubsjoen, 2011; Foddai 2012
		Gait assessment	A	High	Medium	Fitzpatrick, 2006; Caroprese, 2009; Kaler, 2009; King, 2011; Napolitano, 2009; Gougoulis, 2010; Phythian, 2011; Phythian, 2012b
	Mutilations	Castration [e]	A or P	High	High	Scott, 2003; Fitzpatrick, 2006; Stubsjoen, 2011
		Ear notching [e]	A or P	High	High	Phythian, 2011
		Tail docking [e]	A or P	High	High	Scott, 2003; Fitzpatrick, 2006; Stubsjoen, 2011
	Ocular health	Eye condition	A	Low	Medium	Phythian, 2011; Stubsjoen, 2011
	Respiratory health	Cough	A	High	Medium	Napolitano, 2011; Scott, 2011; Stubsjoen, 2011; Phythian, 2012b
		Dyspnoea	A	High	Medium	Mellor, 2003; Scott, 2011
		Nasal discharge	A	Medium	Medium	Napolitano, 2009; Phythian, 2011; Scott, 2011; Stubsjoen, 2011
	Reproductive health	Vulvar discharge	A or P	Medium	Medium	Napolitano, 2011
	Systemic disease	General demeanour	A	Medium	Low	Gougoulis, 2010; Phythian, 2011; Phythian, 2012b
		Obviously sick [e]	A	High	High	Mellor, 2003; Stubsjoen, 2011
To express normal behaviour	Abnormal behaviour	Redirected behaviour	A	High	Low	Dwyer, 2004; Dwyer, 2008; Nowak, 2008; Teixeira, 2012
		Stereotypies	A	High	Low	Dwyer, 2004; Dwyer, 2008; Nowak, 2008; Teixeira, 2012
	Agonistic behaviour	Aggression	A or P	Low	Low	Dwyer, 2004; Wemelsfelder, 2004; Gougoulis, 2010; Teixeira, 2012
		Threats	A	Low	Low	Teixeira, 2012
	Normal behaviour	Play behaviour	A	Low	Low	Boissy, 2007; Dwyer, 2008; Phythian, 2011
		Ruminatory behaviour	A	Low	Low	Phythian, 2011
		Qualitative Behaviour Assessment (QBA)	A	High	Low	Wemelsfelder, 2004; Dwyer, 2008; Phythian, 2011
	Positive behaviour	Affiliative behaviour	A	Low	Low	Boissy, 2007; Teixeira, 2012
From fear and distress	Animal behaviour	Ear posture	A	Medium	Low	Reefmann , 2009
		Freezing	A	Medium	Low	Bouissou, 1995; Cockram , 2004
		Head position	A	Low	Low	Cockram , 2004; Hemsworth, 2011
		QBA	A	Low	Low	Wemelsfelder, 2004

		Vigilance behaviour	A	Low	Low	Dwyer, 2008; Deiss, 2009
		Vocalisations	A	Low	Low	Bouissou, 1995; Cockram , 2004; Boissy, 2007; Greiveldinger, 2007; Nowak, 2008; Deiss, 2009; Gougoulis, 2010; Hemsworth, 2011;
	Human-animal relationship	Alertness to approach in the field	A	Low	Low	Phythian , 2011
		Avoidance distance test (ADT)	A	Low	Low	Napolitano, 2006; Waiblinger, 2006; Dwyer, 2008; Caroprese, 2009; Mattiello, 2010; Napolitano, 2011; Stubsjoen, 2011
		Escape attempts	A	Low	Low	Bouissou, 1995; Cockram , 2004
	Meat quality	Meat colour	P	Medium	High	Dwyer, 2008; Liste, 2011; Napolitano, 2011; Teixeira, 2012
		pH meat	P	Medium	High	Dwyer, 2004; Napolitano, 2006; Deiss, 2009; Liste, 2011; Teixeira, 2012
		Tenderness	P	Medium	Low	Liste, 2011
		Water holding capacity	P	Medium	Low	Liste, 2011

[a] Validity: High validity indicators were those validated in previous research, medium validity indicators were those without a reliable method of assessment and low validity indicators were those that have been suggested as indicators in scientific literature but without evidence that they actually assess welfare.

[b] Feasibility is based on the likely ability to make the observation of the indicator in an abattoir.

[c] Reference column shows publications that identify the proposed indicator as a measure to assess sheep welfare. Only first author and year are presented (full references are in the reference list).

[d] No indicators of thirst were identified in the systematic literature review.

[e] Indicators graded as both highly valid and feasible and which might be used in abattoirs to assess prior sheep welfare (on farm, at market and during transportation).