



The assessment of dairy goats' welfare in alpine summer ranges

Manuela Renna, Benedetta Torsiello, Silvana Mattiello, Luca Battaglini & Monica Battini

To cite this article: Manuela Renna, Benedetta Torsiello, Silvana Mattiello, Luca Battaglini & Monica Battini (2025) The assessment of dairy goats' welfare in alpine summer ranges, Italian Journal of Animal Science, 24:1, 292-308, DOI: [10.1080/1828051X.2025.2452251](https://doi.org/10.1080/1828051X.2025.2452251)

To link to this article: <https://doi.org/10.1080/1828051X.2025.2452251>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 22 Jan 2025.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

RESEARCH ARTICLE



The assessment of dairy goats' welfare in alpine summer ranges

Manuela Renna^a , Benedetta Torsiello^b , Silvana Mattiello^c , Luca Battaglini^b  and Monica Battini^c 

^aDipartimento di Scienze Veterinarie, Università degli Studi di Torino, Grugliasco, TO, Italy; ^bDipartimento di Scienze Agrarie, Forestali e Alimentari, Università degli Studi di Torino, Grugliasco, TO, Italy; ^cDipartimento di Scienze Agrarie e Ambientali - Produzione, Territorio, Agroenergia, Università degli Studi di Milano, Italy Milano

ABSTRACT

Aim of this study was to evaluate the results of the application of a modified version of the AWIN welfare assessment protocol for goats in alpine farming systems during summer grazing, focusing on inter-observer reliability (IOR) of the indicators and feasibility of the protocol. Three assessors applied the protocol in nine farms exploiting six alpine pastures (APs) in NW Italy. IOR of individual-level indicators was excellent or good, but unacceptable for hair coat condition, probably due to insufficient light conditions that sometimes the assessors faced inside the milking pens. IOR of group-level indicators was unacceptable in most cases. These indicators were collected from a distance, and with limited visibility due to adverse weather conditions, thick vegetation cover or rocky areas, thus explaining the unacceptable reliability. On average, 4 h were needed for the protocol application in each AP. The level of welfare was good. Exceptions regarded the hair coat condition, possibly due to external parasitism, the use of shelter while resting and the level of synchronisation during resting. For the latter, we hypothesise that the time selected for data collection was not ideal, but further studies are required to confirm such hypothesis. The application of the protocol provided useful information about goat welfare in alpine summer ranges but posed issues in terms of feasibility and reliability, especially when considering group-level indicators. Future studies are required to set up a more feasible and reliable protocol for these farming systems, possibly with the aid of Precision Livestock Farming techniques.

HIGHLIGHTS

- A modified AWIN protocol was applied for the welfare assessment of dairy goats during alpine summer grazing
- Reliability of individual-level indicators was generally excellent or good, while that of group-level indicators was unacceptable in most cases
- Poor visibility due to adverse weather conditions, thick vegetation and use of binoculars impaired the overall feasibility of group-level indicators

ARTICLE HISTORY

Received 21 September 2024
Revised 19 December 2024
Accepted 7 January 2025

KEYWORDS

Animal welfare indicators;
Capra hircus; extensive
systems; feasibility;
inter-observer reliability

Introduction

The importance of animal welfare in food production has been constantly increasing. This has led to the development of welfare assessment protocols for many species in various farming systems (Webster 2005). Most of the welfare assessment protocols have been developed for intensive husbandry systems, with animals usually kept indoors (Welfare Quality[®] 2009a, 2009b; AWIN 2015a). The reason is that, in intensive farming, welfare issues are more likely to occur due to, for example, high stocking density, limited space available for the animals and certain restrictions on the expression of their normal behaviour (Temple and

Manteca 2020). It is undeniable that freedom of movement and choice (e.g. feed selection, preferred companions, resting areas) are greater in extensive rather than in intensive farming systems. However, Windsor (Windsor 2021) highlighted several hazards to the welfare of livestock kept in extensive husbandry systems, among which prolonged periods of feed and/or water scarcity, exposure to adverse climatic conditions, and risk of predation. In the mountain, the climate can be rather harsh, with low temperatures subjected to wide variations, strong winds, high solar radiation, and reduced partial oxygen pressure (Zemp et al. 1989). These conditions can be quite challenging and, if the

CONTACT Silvana Mattiello  silvana.mattiello@unimi.it

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

animals are not provided with adequate shelters, their welfare could be compromised (Redbo et al. 2001). Thunderstorms are a further threat for animal welfare in mountain areas, as thunders may lead to accidents, or even death from electrocution (Lucchesi and Mattiello 2008). The welfare assessment of goats released to pasture in the Alps could be impacted by ongoing climate changes observed in the region during the twentieth century (Barry 2002), including fluctuating temperatures. These climate variations may also affect the vegetation, as changes in temperature and precipitation alter the biotic factors in the area. In particular, climate warming has led to increased shrubland coverage, which could reduce visibility and pose challenges for monitoring goats that tend to browse in shrubs (Cannone et al. 2007). Villalba et al. (2016) reported that livestock are not always raised in the habitat in which they have evolved, so the natural environment might also present some critical issues. This last assumption is often ignored by stakeholders (e.g. consumers, farmers), who assume that animals are automatically well adapted (Doughty et al. 2017). Furthermore, some surveys reported that, for consumers, extensive systems equal positive welfare, while farmers, although aware of potential risks, feel that issues faced by livestock in extensive systems, being “natural”, cannot be avoided or mitigated (Doughty et al. 2017). Sweeney et al. (2022) found that consumers spontaneously distinguish between welfare status of farm animals in relation to the perceived intensity level of the farming system. Small and extensive husbandry systems are perceived as the ideal conditions to ensure (Sweeney et al. 2022).

Furthermore, according to the results of many surveys, consumers' expectation of welfare conditions in extensive farming is high (Sweeney et al. 2022), because extensive systems are considered the most suitable for improving the social and physical conditions of farm animals, and consequently their welfare conditions (Riggio et al. 2023). It is fundamental not to betray the trust placed in such husbandry systems, and this means that they should be regularly monitored to check animal welfare. However, most welfare protocols and indicators have been developed for intensive husbandry systems, and they may not be suitable to assess the welfare of livestock in extensive farming systems (Kaurivi et al. 2019). This is why welfare assessment in extensively managed farm animals is still frequently neglected (Temple and Manteca 2020). Some obstacles for the development of practical protocols for beef (Kaurivi et al. 2019; Kaurivi,

Hickson, et al. 2020; Kaurivi, Laven, et al. 2020) and dairy cattle (Sapkota et al. 2020) in extensive systems in New Zealand have been well described: the authors highlighted that even existing protocols already developed for very extensive rangeland welfare assessment cannot be easily applied to all pasture-based systems, since specific operations largely differ among similar husbandry systems.

Within the European Animal Welfare Indicators (AWIN) project, a welfare assessment protocol for sheep in both intensive and extensive farming systems was developed (AWIN 2015b; Richmond et al. 2017) and applied in many studies (e.g. Munoz et al. 2017, 2019; Nenadovic et al. 2020). The AWIN welfare assessment protocol for goats was instead developed for intensive and semi-intensive farming systems, only (AWIN 2015a). Some authors applied the AWIN welfare assessment protocol for goats in extensive and semi-intensive conditions, partly modifying the original protocol (e.g. Leite et al. 2020; Battini et al. 2021; Nenadović et al. 2021). Leite et al. (2020) reported no specific feasibility problems in Brazilian semi-arid areas when applying a modified version of the AWIN welfare assessment protocol for goats. However, extensive husbandry systems largely differ one from each other, and, for example, the welfare assessment of goats in mountain ranges (e.g. alpine grazing areas) might be challenging for assessors due to the lack of barns or racks to restrain animals for individual inspection, or to adverse (e.g. fog, rain) and fast-changing weather conditions, or to dense vegetation (trees and shrubs) that reduce the possibility of observing the animals. It could also be difficult for assessors to walk and follow the animals due to uneven and slippery terrains, as reported by Battini et al. (2021) in Northern Italy.

Inter-observer reliability (IOR; defined as the agreement between different observers during a simultaneous observation; (Bateson and Martin 2021) is a fundamental pre-requisite for the inclusion of an indicator in a welfare assessment protocol (Giammarino et al. 2021). IOR has already been adopted in several studies in different species (e.g. goats: (Vieira et al. 2018); sheep: (Phythian et al. 2012); cows: (Katzenberger et al. 2020); horses: (Czycholl et al. 2019)). However, most of the research has been carried out in intensive or semi-intensive/semi-extensive farming conditions, where observations are easier to carry out due to the possibility of close inspection and of individual restraining of the animals. On the contrary, scarce information is currently available for extensive farming conditions (Spigarelli et al. 2020). IOR is strictly linked to the feasibility of an indicator (defined as the practical chance of using it

during the welfare assessment) (Battini et al. 2014): low feasibility negatively influences the reliability of the indicator itself.

The aim of the present study is to evaluate the results of the application of a modified version of the AWIN welfare assessment protocol for dairy goats in extensive farming conditions during summer mountain pasture in the Italian Alps, paying particular attention to the feasibility of the protocol and to the reliability of the indicators in alpine conditions. We chose to use the AWIN protocol because, so far, it is the only validated protocol to assess goat welfare on farm. Although this protocol was developed for intensive goat farms, most of the indicators have already proved to be feasible in semi-extensive systems (Battini et al. 2021), while others had to be removed because they were not applicable in pasture conditions, and others need further investigation to confirm their feasibility and reliability (Battini et al. 2021). Therefore, we used a modified version of AWIN protocol, integrated with possible alternative indicators, to test its suitability in alpine environment. The outcomes of the welfare assessment in alpine summer ranges are presented and discussed in comparison with the outcomes of goats' welfare evaluation in other farming systems.

Materials and methods

Farms and study area

As detailed in Battini et al. (2021), goat farm selection was achieved by extracting the information from a database of 163 farms, provided by the ASL (Local Health Center) TO3 territory of Pinerolo-Collegno (Province of Turin, Piedmont, NW Italy). From the database, we initially selected the farms based on the following inclusion criteria: (i) raising goats for dairy or dual purposes; (ii) making use of outdoor grazing in spring and autumn in proximity to the winter housing buildings; (iii) breeding prevalently Alpine and Valdostana breeds (and their crossbreeds, Alpine $\times$ Valdostana); (iv) voluntary acceptance of the farmer. Only 13 farms satisfied all the inclusion criteria and were included in the previous study on the application of the modified AWIN welfare assessment protocol in dairy goat farms under semi-extensive farming conditions (outdoor grazing in spring and autumn in proximity to the winter barns) (Battini et al. 2021). For the purpose of the present study, an additional selection criterion was the use of alpine pastures during the summer grazing season. This selection resulted in the inclusion for the current study of nine out of the original 13 farms. The selected farms were located in

the North-Western Italian Alps (Province of Turin, Piedmont) and reared Alpine and Valdostana (and their crossbreeds Alpine $\times$ Valdostana) goats for dairy purpose. During summer 2021, the goats of the nine selected farms were brought to six alpine pastures (APs), with an altitude ranging from 1177 m to 2141 m a.s.l. (mean $\pm$ SD: 1782 ± 356.2 m a.s.l.). The number of the farms does not correspond to the number of the APs included in the present study (Table 1) since merging flocks from different farms during the summer grazing period is a common practice in the Italian Alps, with the aim to reduce human labour costs (Battaglini 2007). Farms were either merged with other recruited farms or with farms not involved in this study (Table 1).

Goats

The total number of lactating goats in the nine recruited farms ranged from 5 to 37 goats (mean $\pm$ SD: 18 ± 10.8). In the six APs, the total number of goats, including those from the other flocks that did not specifically belong to our nine recruited farms, ranged from 40 to 220 goats (107 ± 65.7), while the total number of lactating goats ranged from 7 to 100 goats (45 ± 34.9). In the period when the APs were visited, the individual milk yield ranged, on average, from 0.4 to 1.0 L/day. Bucks were present in the flock in four APs, while in the other two APs they were kept indoor. Detailed information on the characteristics of the grazing animals is reported in Table 1.

Goat management in the alpine pasture areas

Detailed information on the management of the six APs in the summer months is reported in Table 2. The summer grazing period started in June (July for one flock, only) and ended in September or October (depending on the characteristics of the AP). All the farmers responsible for keeping the goats in the APs were males, with an average age of 46 ± 11.0 years (min 32; max 62 years). In two flocks, the stockmen were not the same who managed the goats during the rest of the year, as they were specifically recruited *ad interim* for the summer grazing period, only.

Even if some farms raised other animal species (Table 2), the latter never grazed in the same areas with goats. In all the visited APs, from two to four working dogs were also present. Farmers used the dogs to gather the animals and/or to protect them from wolves.

The lactating goats were manually milked twice a day, except for two flocks where the goats were milked

Table 1. Characteristics of the goats in the six alpine pastures (AP) and in the nine selected farms (F).

	AP1			AP2			AP3			AP4		AP5		AP6		
	F1	Other farm(s)		F2			F3	F9	Other farm(s)	F4	Other farm (s)	F5	F6	F7	F8	Other farm(s)
Number of lactating goats	F	17	3	7			5	15	50	20	25	13	13	37	33	30
	AP	20		7				70			45		26		100	
Total number of goats ^a	AP	54		40				112			135		80		220	
Goat breeds ^b	AP	Alpine		Alpine				Alpine			Alpine		Alpine		Alpine	
		Valdostana		Valdostana				Valdostana			Valdostana		Valdostana		Valdostana	
		A × V						A × V			A × V		A × V		A × V	
Presence of bucks	AP	Yes		Yes				Yes			No		No		Yes	
Average daily milk yield (L/head)	AP	0.8		0.5				0.4			1.0		0.6		0.9	
Average age of goats (years)	AP	6		3				4.5			7		5		5.5	
Presence/absence of horns ^c	AP	Horned only		Horned only			Both horned and acorns			Horned only		Horned only		Both horned and acorns		

^aIncluding lactating goats, dry goats, yearlings, goat kids and bucks.^bA, Alpine; V, Valdostana.^cAll our sample farms had horned goats only (disbudding was not performed), but in some alpine pastures they were occasionally mixed with dehorned goats from other farms.**Table 2.** Characteristics and management of the six alpine pastures (AP).

	Alpine pasture					
	AP1	AP2	AP3	AP4	AP5	AP6
Altitude (m a.s.l.)	1991	1117	2141	1861	1970	1549
Surface area (hectares)	58	20	300	200	250	146
Summer grazing period	June–Sept (90 days)	June–Sept (100 days)	Jul–Sept (80 days)	June–Oct (100 days)	June–Sept (80 days)	June–Oct (130 days)
Average temperature (°C) of the summer grazing period ^a	16.9	17.6	16.8	16.7	18.9	16.7
Minimum temperature (°C) of the summer grazing period ^a	8.6	9.5	14.3	8.6	9.5	6.5
Maximum temperature (°C) of the summer grazing period ^a	26.0	27.5	20.0	26.0	31.3	27.5
Average precipitation (mm) during the summer grazing season ^a	1.8	2.4	1.2	1.7	1.9	3.7
Average annual temperature (°C) ^a	8.6	9.5	8.6	8.6	9.9	9.5
Minimum annual temperature (°C) ^a	−10.5	−9.4	−10.5	−10.5	−9.5	−9.4
Maximum annual temperature (°C) ^a	26.0	27.5	26.0	26.0	31.3	27.5
Average annual precipitation (mm) ^a	2.5	3.0	2.5	2.5	2.2	3.0
Number of flocks from different farms	2	1	3	6	2	4
Raising of other farm animals	Cows Sheep	No	Cows	Cows Sheep Laying hens Donkeys	Cows	Cows Horses
Herding dogs	2	3	2	2	3	3
Guardian dogs	0	1	1	2	0	0
Presence of shelters	Yes (Rocks/Caves)	No	Yes (Rocks)	Yes (Rocks)	No	No
Presence of water points in night shelter	No	Yes	No	No	No	Yes
Water points at pasture	1	0	1	5	3	1
Supplementary feeding	No	Flaked cereals	No	No	No	No
Management during the night	Barn	Barn	Barn	Outdoor fence	Barn	Barn
Milking time, morning (h)	06:30	09:30	08:00	09:00	08:00	05:30
Milking time, evening (h)	19:00	/	/	21:00	19:30	18:30
Exit to pasture (h)	08:00	10:30	09:30	10:00	09:00	07:00
Return from pasture (h)	18:30	13:00	17:00	20:00	19:00	18:00
Stock person (gender)	Male	Male	Male	Male	Male	Male
Stock person (age, years)	62	44	32	40	43	56
Same stock person as in valley bottom	No	Yes	Yes	Yes	No	Yes

^aData provided by Banca Dati Meteorologica – ARPA Piemonte. Available at: www.arpa.piemonte.it (last accessed 02 December 2024).

only in the morning (Table 2). Milking occurred inside a barn for three flocks and in an open fenced area for the other three flocks. The visited milking barns were in two cases (AP2 and AP6) dark, filthy and small when compared to the number of the raised goats (thus leading to overcrowding problems), while in one case (AP3) the barn was newly built, luminous, and clean.

The goats were released to the grazing areas after the morning milking and were gathered before the evening milking, except for the flocks adopting the

morning milking only (Table 2). Different ways of gathering the goats were used in the visited APs. The goats were gathered by farmers helped by herding dogs (AP1; AP3; AP4), by herding dogs only (AP2; AP5), or the animals used to spontaneously go back close to the milking area at milking time (AP6). The goats fed fresh grass and bushes available in the grazing areas. Specific details on goats' nutrition and additional details on management are specified in Table 2.

During the night, the goats were kept inside a barn, except for one flock where the animals were kept in an outdoor fenced paddock; water was available only for two of the flocks that were sheltered inside the barn.

Assessors

The welfare assessment protocol was applied during the summer grazing period in each AP. With the aim of testing the reliability of the welfare protocol, the assessment was independently carried out by three assessors at the same time. The assessors had different backgrounds: two assessors were students of the first (assessor A) and second (assessor B) year of the MSc in Animal Science at the University of Turin (Italy), while the third assessor (assessor C) was a student of the first year of the MSc in Sciences and Technologies of Forest Systems and Territories at the same University. The three assessors had no specific previous experience with dairy goats. They received a common training before the beginning of data collection, including both theoretical and practical sessions. They also received, as training material, both the original AWIN welfare assessment protocol for goats kept in intensive or semi-intensive production systems (AWIN 2015a) and our previous publication on the application of the AWIN welfare assessment protocol for goats under semi-extensive conditions (Battini et al. 2021). The training was given by one author of the original AWIN welfare assessment protocol for goats.

Data collection

The farmers were contacted in advance to plan the visit and to collect some preliminary information (e.g. milking time, pasture management). The three assessors then visited each of the six APs (on six different days) and simultaneously and individually applied the protocol in each AP. All the assessors were unknown to the flocks.

Before scheduling the visits, the assessors checked the weather forecast with the attempt to perform the welfare assessment on sunny days. The assessments took place under partly cloudy (three flocks), sunny (one flock), cloudy (one flock) or foggy (one flock) weather conditions, but never when raining. Wind was absent (on four occasions) or light (on two occasions). Environmental temperatures and relative humidity were recorded by using a portable thermo-hygrometer (HOBO MX2301A, Onset, Bourne, MA, USA). The environmental temperatures on the days of the assessment ranged from 12 °C to 22 °C ($17^{\circ}\text{C} \pm 3.5^{\circ}\text{C}$). The relative

humidity ranged between 50% and 80% ($63\% \pm 12.8\%$) (Table 3).

The applied welfare assessment protocol, already described in detail in Battini et al. (2021), is a modified version of the AWIN welfare assessment protocol for goats (AWIN 2015a), integrated with some indicators derived from the AWIN welfare assessment protocol for sheep (AWIN 2015b). The protocol applied in the current study consisted of ten individual-level (body condition score (BCS), hair coat condition, abscesses, overgrown claws, udder asymmetry, faecal soiling, nasal discharge, ocular discharge, improper disbudding, and severe lameness) and seven group-level (Qualitative Behaviour Assessment – QBA, synchrony at grazing, oblivion, thermal stress, Familiar Human Approach Test (FHAT), synchrony at resting, and use of shelter while resting) indicators. Specifically, most of the original AWIN individual animal-based welfare indicators were included, as they could be evaluated in a way very similar to the one adopted in intensive (AWIN 2015a) and semi-extensive (Battini et al. 2021) conditions, when the animals are restrained for milking. In the present study the evaluation of severe lameness had also to be evaluated at individual level, when the goats left the milking area, because it was unfeasible to evaluate it in the pen at group level as in the original AWIN protocol (AWIN 2015a). Some of the group-level (i.e. Queuing at feeding, Queuing at drinking, Kneeling at the feeding rack, Latency to first contact test) and resource-based (i.e. bedding quantity and cleanness) welfare indicators included in the original protocol had to be removed, because they were not feasible in extensive conditions, and they had to be replaced with other more feasible indicators, which were included in the modified protocol for semi-extensive systems (i.e. synchrony at grazing, FHAT, synchrony at resting, and use of shelter while resting), and were fitted to the alpine grazing conditions. Finally, for the evaluation of the Human-Animal Relationship (HAR) it was impossible to apply the Latency to First Contact test, which was therefore replaced by a modified version of the FHAT, which had previously demonstrated to be suitable for assessing the quality of HAR in dairy goats kept outdoor (Battini et al. 2021). The flow-chart of data collection was also published previously (Battini et al. 2021). The individual-level indicators were recorded during the morning milking, except for severe lameness which was recorded while the goats were moved from the milking area to the grazing area. The group-level indicators were recorded while the goats were at pasture. The individual-level welfare assessment regarded

Table 3. Weather conditions during welfare assessment in the alpine pastures (AP).

	Alpine pasture					
	AP1	AP2	AP3	AP4	AP5	AP6
Visibility	Partly cloudy	Partly cloudy	Sunny	Foggy	Cloudy and foggy	Partly cloudy
Presence of wind	Light wind	/	/	Light wind	/	/
Average air temperature (°C)	12	15	20	17	17	22
Average relative humidity (%)	73	50	50	70	80	56

exclusively the adult lactating goats of the nine recruited farms ($n = 17.8 \pm 10.84$; min = 5; max = 37). In flocks merged with others, the farmers helped with the identification of their goats. However, when animals were at pasture, it was impossible to distinguish from a distance the goats belonging to one farm from those of another farm, therefore the group-level welfare assessment considered all the goats which were present in the AP, and which belonged not only to the recruited farms, but also to all the other farms merged in the same AP ($n = 106.8 \pm 65.73$; min = 40; max = 220). QBA observations were performed after the release of goats to the grazing area. Synchrony at grazing, oblivion and thermal stress were recorded immediately after QBA, followed by FHAT and finally by synchrony at resting and use of shelter. To record synchrony at grazing, oblivion, thermal stress, synchrony at resting and use of shelter, a scan sampling method was applied (three scans at 30-min intervals), as detailed in Battini et al. (2021). The quality of the human-animal relationship (HAR) was assessed by using a simplified version of the FHAT initially developed for sheep (AWIN 2015b), considering three possible classes of reactions towards the farmer: Avoidance (withdrawal of the whole flock from the farmer), Approach (at least one goat voluntarily approaches the farmer, but no physical interaction occurs), Contact (at least one goat voluntarily approaches the farmer and physical interaction occurs) (Battini et al. 2021). This test was later validated for dairy goats bred at pasture (Battini et al. 2023).

At the end of the welfare assessment, Assessor B asked the farmers to fill in a questionnaire, which contributed to complement the information directly recorded by the assessors and to better understand environmental and management factors (e.g. stockpersons' characteristics, animals, characteristics of the alpine pasture area, and management of goats at pasture) that could affect animal welfare. The questionnaire was structured into six sections: (i) weather conditions during the day of data collection; (ii) personal information about the farmer (e.g. gender, age); (iii) raised animals (e.g. breed, number of animals per category); (iv) characteristics of the alpine pastures (e.g. total surface of the pasture area, water points);

(v) animal management during summer grazing (e.g. feed supplementation, night shelter).

Finally, at the end of each on-farm assessment, assessors were asked to report the major constraints experienced during the application of the protocol.

Data analysis

For individual-level indicators, percentage frequencies of animals without welfare problems were calculated. For group-level indicators, data were calculated depending on the indicator: for synchrony at grazing, synchrony at resting and use of shelter, the proportion of goats grazing, resting or using shelter while resting in each scan was calculated out of the number of goats at pasture; for thermal comfort and oblivion, the proportion of goats without welfare problems was calculated. FHAT results were reported as the number of flocks out of the total number of flocks where goats avoided the farmer, spontaneously approached the farmers or accepted the contact with the farmer, out of the total number of flocks. Principal Component Analysis (PCA, correlation matrix, no rotation) was used to explore the results of descriptors' scores obtained from the QBA.

In the results section, the welfare outcomes are presented referring to data collected by all three assessors (A; B; C). However, only the results for assessor B were discussed concerning the welfare problems, because we presumed that this assessor was the most qualified, being at the end of a master's degree course specifically devoted to animal farming, while the other assessors were either in first year of their master degree or were engaged in a course that was not specifically devoted to farming, and therefore had a less specific background on animal science. For categorical data (individual-level indicators), the IOR was calculated over all goats, regardless of the flock. The percentage of concordance agreement (i.e. concordance rate, P_0) was calculated for each pair of assessors (A vs B, B vs C and A vs C) and for the three assessors together (A vs B vs C). The $\gamma(AC_1)$ index (Gwet 2008), which has been shown to be among the best agreement indexes to assess IOR in the case of dichotomous and trichotomous categorical animal-based

Table 4. Absence of welfare problems (mean $\pm$ SE; min–max) observed for nine dairy goat farms reared in six alpine pastures during summer, recorded during individual-level assessment ($N = 160$)^a.

Indicator	A	B	C
	Mean% $\pm$ SE% (min%–max%)	Mean% $\pm$ SE% (min%–max%)	Mean% $\pm$ SE% (min%–max%)
Normal body condition	88.4 $\pm$ 8.46 (71.4–100.0)	90.8 $\pm$ 2.45 (80.0–100.0)	83.5 $\pm$ 8.79 (71.4–100.0)
Good hair coat	83.1 $\pm$ 11.30 (60.0–100.0)	45.8 $\pm$ 2.43 (32.0–54.5)	43.7 $\pm$ 19.32 (7.7–100.0)
Absence of abscesses	89.8 $\pm$ 10.66 (69.2–100.0)	91.7 $\pm$ 2.69 (78.8–100.0)	85.3 $\pm$ 9.01 (70.6–100.0)
Regular claws	93.3 $\pm$ 8.42 (76.5–100.0)	93.8 $\pm$ 2.87 (76.5–100.0)	83.0 $\pm$ 17.32 (57.1–100.0)
Symmetric udder	92.8 $\pm$ 7.74 (76.9–100.0)	95.0 $\pm$ 2.60 (77.0–100.0)	86.7 $\pm$ 13.87 (66.7–100.0)
Absence of faecal soiling	94.0 $\pm$ 10.96 (70.6–100.0)	85.6 $\pm$ 5.72 (60.0–100.0)	80.8 $\pm$ 23.24 (40.0–100.0)
Absence of nasal discharge	99.4 $\pm$ 1.67 (95.0–100.0)	95.2 $\pm$ 2.90 (80.0–100.0)	94.6 $\pm$ 7.62 (80.0–100.0)
Absence of ocular discharge	97.7 $\pm$ 4.85 (86.7–100.0)	93.8 $\pm$ 2.45 (80.0–100.0)	93.3 $\pm$ 9.08 (73.3–100.0)
Absence of severe lameness	100.0 $\pm$ 0.00 (100.0–100.0)	99.0 $\pm$ 0.68 (94.1–100.0)	99.1 $\pm$ 2.05 (94.1–100.0)

^aResults for proper disbudding are not reported, as this procedure was not in use in any of the nine sample farms.

welfare indicators in the presence of two (Giammarino et al. 2021; Torsiello et al. 2024) and three (Torsiello et al. 2023a, 2023b) observers, was then calculated for evaluating the IOR between pairs of assessors and among the three assessors together. In addition, confidence intervals for $\gamma(AC_1)$ index were calculated to give information about the variability of the values assumed by the index itself. For continuous data (group-level indicators), intra-class correlation (ICCs) coefficients were calculated (95% confidence intervals, based on absolute agreement, random effects type, mean-rating). According to Bateson and Martin (Bateson and Martin 2021), the following thresholds were used to evaluate the quality of the reliability: <0.50 = poor; $0.50 - 0.75$ = moderate; $0.76 - 0.90$ = good; >0.90 = excellent. For QBA, the IOR of Principal Component (PC) scores attributed by the three assessors to each flock on the first two PCs was analysed by using the Kendall Correlation Coefficient W . The results were interpreted according to Bateson and Martin (Bateson and Martin 2021), where W : $0.0 - 0.2$ = slight correlation; $0.2 - 0.4$ = low correlation; $0.4 - 0.7$ = moderate correlation; $0.7 - 0.9$ = high correlation; $0.9 - 1.0$ = very high correlation. Several authors (Knierim and Winckler 2009) proposed to use a correlation coefficient threshold ≥ 0.7 as a guideline when assessing IOR in categories that are challenging to measure. This threshold provides stronger and more reliable results; hence we will follow this guideline for the discussion of our results.

The software IBM SPSS v. 26.0 for Windows (IBM Corp. Released 2019) was used to perform a PCA using QBA outcomes. The software R (R Commander, version $R \times 64$ 4.2.2) was instead used to calculate Gwet's $\gamma(AC_1)$ index and related confidence intervals for individual-level

indicators (R package: library(irrCAC); R function: gwe.tac1), and ICCs and related confidence intervals for group-level indicators (R function: icc).

Results

The results of the application by the three assessors (A, B, C) of the proposed modified version of the AWIN welfare assessment protocol for dairy goats during the summer grazing period in alpine pastures are shown in Table 4 (individual-level assessment) and Table 5 (group-level assessment), except for improper disbudding (this procedure was not performed in our sample farms), QBA and FHAT. The main identified welfare problems regarded the hair coat condition (more than half of the assessed goats were scored as having poor hair coat condition), synchrony at resting (on average only a maximum of 11.1% of the goats rested simultaneously) and use of shelter while resting (more than 88% of the goats rested without using a shelter).

Regarding QBA, the analysis identified four main PCs with eigenvalues >1 (4.860, 3.792, 2.299, 1.513 for PC1, PC2, PC3, and PC4, respectively). The first two PCs together explained 66.56% of the total variance among the APs (PC1: 37.39%; PC2: 29.17%). Figure 1 shows the distribution of the descriptors' loadings and of the scores of the AP along the first two PCs. Descriptors on PC1 (that describes the valence of emotions) ranged from Relaxed to Lively, whereas descriptors on PC2 (that describes the arousal) ranged from Frustrated to Alert.

As to the assessment of HAR quality, the results of FHAT showed that in four APs (i.e. Contact, in AP2, AP3, AP5 and AP6) the goats actively moved towards the

Table 5. Absence of welfare problems (mean $\pm$ SE; min–max) observed for six flocks reared in six alpine pastures during summer, recorded during group-level assessment ($N=641$)^a by the three assessors.

Indicator	Assessor A SCAN 1	SCAN 2	SCAN 3
	Mean% $\pm$ SE% (min%–max%)	Mean% $\pm$ SE% (min%–max%)	Mean% $\pm$ SE% (min%–max%)
Synchrony at grazing	55.0 $\pm$ 34.85 (4.7–93.0)	56.3 $\pm$ 44.00 (0.0–96.3)	76.2 $\pm$ 21.83 (44.4–100.0)
Absence of oblivious goats	3.9 $\pm$ 4.65 (0.0–12.5)	7.9 $\pm$ 9.74 (0.0–25.0)	3.8 $\pm$ 4.96 (0.0–13.0)
Synchrony at resting	14.7 $\pm$ 22.18 (0.0–48.8)	7.0 $\pm$ 11.47 (0.0–27.8)	1.0 $\pm$ 2.24 (0.0–5.56)
Use of shelter while resting	1.8 $\pm$ 2.95 (0.0–7.0)	1.5 $\pm$ 2.20 (0.0–5.6)	6.8 $\pm$ 14.85 (0.0–37.0)

Indicator	Assessor B SCAN 1	SCAN 2	SCAN 3
	Mean% $\pm$ SE% (min%–max%)	Mean% $\pm$ SE% (min%–max%)	Mean% $\pm$ SE% (min%–max%)
Synchrony at grazing	71.9 $\pm$ 31.65 (11.6–100.0)	66.5 $\pm$ 18.81 (34.9–92.6)	68.7 $\pm$ 17.42 (44.4–87.5)
Absence of oblivious goats	91.8 $\pm$ 4.70 (85.0–96.4)	89.12 $\pm$ 7.44 (75.0–97.0)	90.6 $\pm$ 3.27 (87.0–96.0)
Synchrony at resting	11.1 $\pm$ 21.41 (0.0–53.5)	6.3 $\pm$ 7.65 (0.0–11.6)	3.8 $\pm$ 6.55 (0.0–16.7)
Use of shelter while resting	2.5 $\pm$ 3.65 (0.0–9.3)	2.4 $\pm$ 4.66 (0.0–11.6)	0.4 $\pm$ 0.73 (0.0–1.79)

Indicator	Assessor C SCAN 1	SCAN 2	SCAN 3
	Mean% $\pm$ SE% (min%–max%)	Mean% $\pm$ SE% (min%–max%)	Mean% $\pm$ SE% (min%–max%)
Synchrony at grazing	66.8 $\pm$ 23.75 (40.0–105.6)	70.3 $\pm$ 24.00 (45.1–96.3)	61.3 $\pm$ 38.19 (3.0–101.9)
Absence of oblivious goats	12.4 $\pm$ 4.90 (5.6–17.5)	9.7 $\pm$ 3.28 (5.9–14.0)	11.5 $\pm$ 7.81 (2.2–21.0)
Synchrony at resting	4.7 $\pm$ 6.59 (0.0–14.9)	5.3 $\pm$ 6.39 (0.0–14.8)	1.2 $\pm$ 2.24 (0.0–5.6)
Use of shelter while resting	0.9 $\pm$ 2.28 (0.0–5.6)	1.0 $\pm$ 2.05 (0.0–5.1)	0.1 $\pm$ 0.36 (0.0–0.9)

^aResults for thermal comfort are not reported, as no sign of thermal discomfort was ever recorded.

farmers and interacted with them (sniffing, nosing), while in two APs (i.e. Approach, in AP1 and AP4) the goats remained motionless in response to human approach. In none of the APs the animals withdrew from the farmer.

IOR calculated for individual-level indicators is reported in Table 6. For nasal discharge, ocular discharge and severe lameness, the reliability between pairs of assessors and among the three assessors together was excellent. For abscesses, overgrown claws, udder asymmetry and faecal soiling, excellent reliability was obtained between the assessors A and B, while good reliability was obtained between the other pairs of assessors and among the three assessors together. For BCS, the reliability was always good, except between the assessors A and C, where moderate reliability was obtained. Poor reliability was observed both between pairs of assessors and among the three assessors together for hair coat condition. In summary, an acceptable (≥ 0.7) reliability threshold was reached for all individual-level observations, excluding hair coat condition.

As to the group-level indicators, IOR results are reported in Table 7, except for thermal stress which

was never detected during the welfare assessment. For all group-level indicators, IOR highly varied among both scans and assessors involved, showing unacceptable (<0.7) levels of agreement in several cases.

The agreement among the three assessors for QBA was considered unacceptable for both the PC1 and the PC2, due to the low Kendall's W (0.443 and 0.251, respectively). The loadings of the descriptors on the first two PCs resulting from the three assessors are presented in Table 8.

Regarding the FHAT, perfect agreement was reached among the three assessors in four (AP3; AP4; AP5; AP6) out of six APs.

The collection of individual-level indicators took about 1 min/goat. The time required to apply the whole protocol was of approximately 4h (minimum 3h, maximum 4:30h per AP). The assessors reported the presence of some constraints that hindered the overall welfare assessment and that reduced the feasibility of both individual- and group-level indicators (Table 9). Due to the position and orography of the

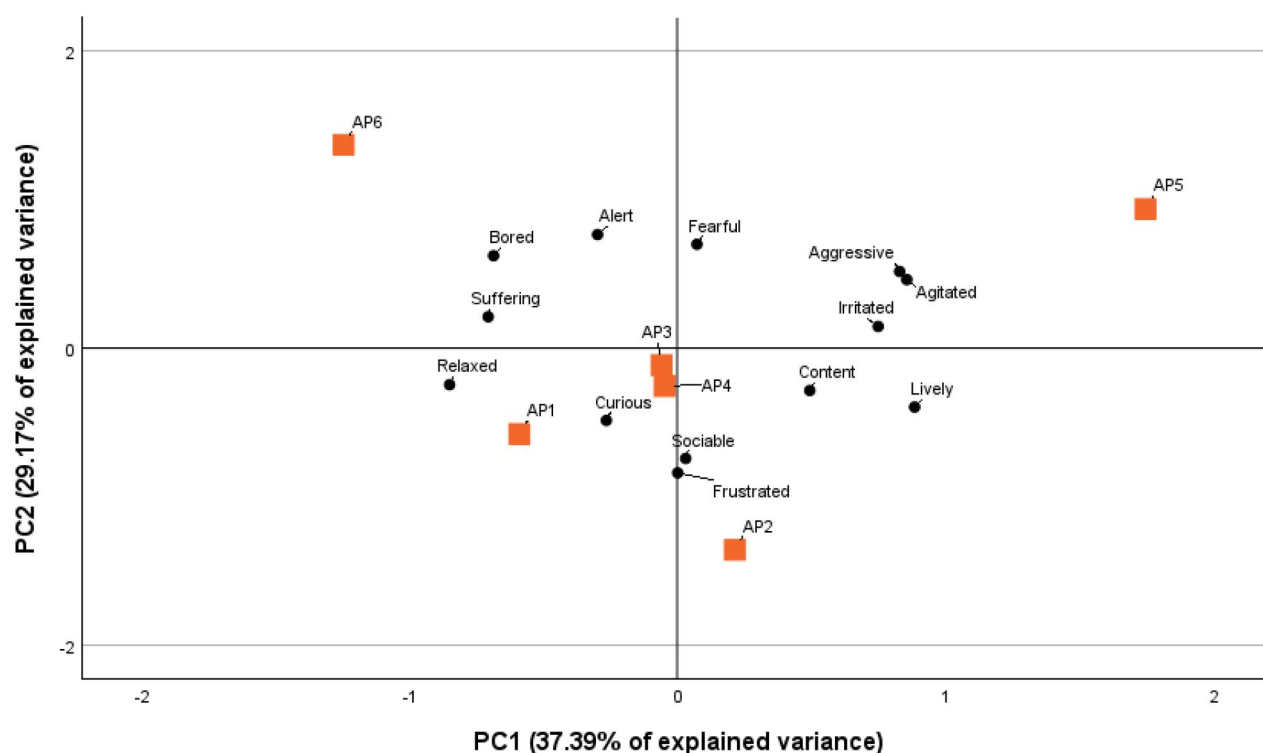


Figure 1. Biplot showing the loadings of the descriptors (·) and the scores of the alpine pastures (■) on PC1 and PC2 generated by assessor B.

Table 6. Inter-observer reliability among assessors A, B, and C for individual-level indicators ($N = 160$)^a.

Indicator	P_0 (A–B)	$\gamma(AC_1)$	P_0 (A–C)	$\gamma(AC_1)$	P_0 (B–C)	$\gamma(AC_1)$	P_0 (A–B–C)	$\gamma(AC_1)$
Body Condition Score	86%	0.84 (0.78–0.91)	76%	0.72 (0.64–0.81)	83%	0.80 (0.73–0.87)	81%	0.79 (0.73–0.85)
Hair coat condition	61%	0.29 (0.13–0.45)	62%	0.30 (0.14–0.46)	71%	0.41 (0.27–0.56)	65%	0.32 (0.21–0.43)
Abscesses	94%	0.93 (0.89–0.98)	90%	0.87 (0.81–0.94)	89%	0.86 (0.80–0.93)	91%	0.89 (0.84–0.94)
Overgrown claws	98%	0.98 (0.95–1.00)	84%	0.80 (0.71–0.88)	84%	0.79 (0.70–0.87)	89%	0.86 (0.81–0.91)
Udder asymmetry	94%	0.93 (0.88–0.97)	88%	0.86 (0.79–0.93)	92%	0.90 (0.85–0.96)	91%	0.90 (0.85–0.94)
Faecal soiling	92%	0.91 (0.86–0.96)	89%	0.86 (0.79–0.93)	93%	0.90 (0.85–0.96)	91%	0.89 (0.84–0.94)
Nasal discharge	96%	0.96 (0.93–1.00)	95%	0.95 (0.91–0.99)	94%	0.93 (0.89–0.98)	95%	0.95 (0.91–0.98)
Ocular discharge	96%	0.95 (0.92–0.99)	96%	0.96 (0.93–0.99)	97%	0.97 (0.94–1.00)	96%	0.96 (0.93–0.99)
Severe lameness	99%	0.99 (0.97–1.00)	99%	0.99 (0.97–1.00)	99%	0.99 (0.97–1.00)	99%	0.99 (0.97–1.00)

P_0 = concordance rate. Thresholds: <0.50 = poor; 0.50–0.75 = moderate; 0.76–0.90 = good; >0.90 = excellent (Bateson and Martin 2021). In bold, indicators where correlation coefficient is ≥ 0.7 .

^aResults for proper disbudding are not reported, as this procedure was not in use in any of the nine sample farms.

APs, the use of an off-road vehicle was required to reach five out of the six APs. For individual-level indicators, which were collected inside the milking pens, low-light conditions were reported in two out of six visited APs, where goats were milked inside the barn, and this affected the quality of the assessment; in these cases, the use of a torch was required, but in one AP the assessors reported difficulties due to scarce visibility in the barn even when using a torch.

Poor visibility was also reported as a major constraint during the assessment of group-level indicators, due to bad weather conditions (cloudy and/or foggy conditions in five out of six APs) and/or presence of thick vegetation cover or rocky areas (in two out of six APs). The presence of aggressive dogs, that prevented the assessors from approaching the goats, was reported as a major constraint in half of the visited APs. The assessors significant challenge in

Table 7. Inter-observer reliability among assessors A, B, and C for group-level indicators ($N = 641$).

	Reliability ICC (95% CI)			
	A-B	A-C	B-C	A-B-C
SCAN 1				
Synchrony at grazing	0.80 (-0.10 < ICC < 0.97)	0.66 (-1.08 < ICC < 0.95)	0.64 (-2.72 < ICC < 0.95)	0.79 (0.21 < ICC < 0.97)
Oblivion	0.67 (-0.44 < ICC < 0.95)	0.20 (-0.36 < ICC < 0.80)	0.26 (-1.37 < ICC < 0.87)	0.46 (-0.305 < ICC < 0.90)
Synchrony at resting	0.94 (0.66 < ICC < 0.99)	0.66 (-0.62 < ICC < 0.95)	0.67 (-1.02 < ICC < 0.95)	0.86 (0.48 < ICC < 0.98)
Use of shelter while resting	0.96 (0.73 < ICC < 0.99)	0.90 (0.40 < ICC < 0.99)	0.85 (-0.01 < ICC < 0.98)	0.94 (0.73 < ICC < 0.99)
SCAN 2				
Synchrony at grazing	0.65 (-1.68 < ICC < 0.95)	0.77 (-0.31 < ICC < 0.97)	0.84 (-0.14 < ICC < 0.98)	0.81 (0.23 < ICC < 0.97)
Oblivion	0.92 (0.44 < ICC < 0.99)	0.48 (-5.15 < ICC < 0.93)	0.34 (-11.33 < ICC < 0.92)	0.79 (0.14 < ICC < 0.97)
Synchrony at resting	0.71 (-2.06 < ICC < 0.96)	0.88 (0.22 < ICC < 0.98)	0.90 (0.28 < ICC < 0.99)	0.88 (0.46 < ICC < 0.98)
Use of shelter while resting	0.18 (-19.37 < ICC < 0.90)	0.26 (-18.16 < ICC < 0.91)	0.83 (0.05 < ICC < 0.98)	0.65 (-0.53 < ICC < 0.95)
SCAN 3				
Synchrony at grazing	0.94 (0.26 < ICC < 0.99)	0.23 (-5.02 < ICC < 0.89)	0.33 (-9.88 < ICC < 0.914)	0.56 (-0.892 < ICC < 0.935)
Oblivion	0.32 (-0.57 < ICC < 0.86)	0.21 (-0.85 < ICC < 0.85)	-0.45 (-0.88 < ICC < 0.46)	-0.11 (-2.00 < ICC < 0.80)
Synchrony at resting	0.70 (-0.50 < ICC < 0.96)	0.97 (0.81 < ICC < 0.97)	0.73 (-0.41 < ICC < 0.96)	0.81 (0.29 < ICC < 0.97)
Use of shelter while resting	-0.05 (-6.16 < ICC < 0.85)	-0.02 (-5.15 < ICC < 0.85)	0.81 (-0.05 < ICC < 0.97)	-0.05 (-3.25 < ICC < 0.84)

ICC, intraclass correlation coefficient; CI, confidence intervals.

Thresholds: <0.50 = poor; 0.50–0.75 = moderate; 0.76–0.90 = good; >0.90 = excellent (Bateson and Martin 2021). In bold, indicators where correlation coefficient is ≥ 0.7 .

tracking the animals immediately after their release at the end of the morning milking. In half of the cases, the animals moved so quickly and over such steep terrains that the assessors were unable to follow them. As a result, they had to position themselves on the opposite side of the valley and use binoculars to observe the animals.

Discussion

Reliability of animal-based welfare indicators

Although many of the welfare indicators used in this study were originally developed for intensive and semi-intensive systems, their reliability needed to be tested due to the different context (i.e. alpine pasture, extensive conditions).

IOR of individual-level indicators was good (with the only exception of hair coat conditions) and this could be explained by the decision to assess the goats during the milking procedure that guaranteed a practical access to the animals. According to Battini et al. (2016a), the restraining method could affect the feasibility of the assessment and in turn the reliability: in fact, when applying the AWIN protocol, the reliability of individual indicators was better when goats could be restrained in the feed rack, than when they had to be

Table 8. Loadings of the QBA descriptors on the first two PCs resulting from the three assessors. Loadings with positive or negative values higher than 0.6 are shown in bold.

Descriptors	Assessor A		Assessor B		Assessor C	
	PC1	PC2	PC1	PC2	PC1	PC2
Aggressive	-0.811	0.395	-0.952	0.18	0.826	0.516
Agitated	0.797	0.581	-0.961	0.228	0.853	0.461
Alert	0.779	-0.55	0.482	0.504	-0.301	0.764
Bored	-0.559	-0.379	0.126	-0.983	-0.689	0.622
Content	0.423	-0.589	-0.862	0.288	0.491	-0.285
Curious	0.817	0.132	-0.05	0.925	-0.268	-0.486
Fearful	0.755	-0.279	0.787	0.334	0.007	0.699
Frustrated	0.435	-0.124	0.739	0.371	-0.002	-0.839
Irritated	0.9	0.076	0.942	0.114	0.746	0.146
Lively	0.4	0.877	0.279	-0.031	0.882	-0.396
Relaxed	-0.749	-0.588	-0.194	-0.815	-0.853	-0.246
Sociable	-0.559	0.694	0.696	-0.37	0.028	-0.742
Suffering	-0.046	-0.197	0.66	0.045	-0.709	0.211

QBA, qualitative behaviour assessment.

restrained manually in the pen. This explains why in the present study, even if the barn was dark, the inter-observer reliability was higher compared to that obtained in well-illuminated pens (Vieira et al. 2018).

The only individual-level indicator that did not reach an acceptable level of agreement (<0.7) was the hair coat condition and wide confidence intervals suggested the high dispersion of the values. The same indicator performed well in intensive farming conditions (Vieira et al. 2018). This discrepancy can probably

Table 9. Major constraints reported by the assessors before and during the application of the welfare assessment protocol.

	Alpine pasture					
	AP1	AP2	AP3	AP4	AP5	AP6
Before the application of the protocol						
Off-road vehicle necessary to reach the pasture area	×		×	×		
During the application of the protocol						
Scarce light inside the pen for the individual-level assessment (use of torches)	×					×
Poor visibility at pasture due to adverse weather conditions	×	×		×	×	×
Presence of thick vegetation cover and/or rocky areas	×				×	
Presence of aggressive dogs		×		×	×	
Use of inaccessible pasture areas (use of binoculars)	×				×	

be explained by the different assessing method: in the original AWIN welfare assessment protocol (AWIN 2015a), the hair coat condition was assessed at group level, counting the animals showing poor quality of the coat from outside the pen. In the current adapted protocol, this indicator was instead assessed on individual animals during the milking procedures. Even if the assessment of individual animals usually helps the assessor to better evaluate the indicators, it is likely to assume that insufficient light conditions may have affected the ability of the assessors to properly evaluate the hair coat during the indoor morning milking (Table 9). This issue was already reported in semi-extensive conditions (Battini et al. 2021).

Last, the slightly lower reliability values obtained for BCS (trichotomous variable) when compared to other individual-level indicators (dichotomous variables) could also be imputable to the higher number of categories in this indicator. In fact, the higher is the number of categories for a given variable, the higher is the possibility of choice for the assessors to assign the variable to a category, and this implies a higher heterogeneity during the assignment. As a consequence, the agreement among the assessors decreases, resulting in lower P_0 and agreement values (Torsiello et al. 2024).

Vieira et al. (2018) reported high reliability of group-level indicators in intensive farming systems, but this was not confirmed in our study (Table 7). This could be explained by the fact that data collection in extensive systems is usually performed from a distance (sometimes with the aid of a binocular; Table 9), with potentially limited visibility, due to adverse weather conditions (Table 3) and/or to the presence of thick vegetation cover or rocky areas (Table 9). Even if the assessors always checked the forecasts before scheduling the visits, weather can be unpredictable on the mountains and change quickly (Bender et al. 2011). Unfortunately, the unexpected presence of fog and low clouds during five out of six days of assessment (Table 3) drastically reduced the visibility of the animals, thus compromising the feasibility during the

evaluations and, consequently, reducing the reliability of the indicators. The visibility was further compromised by the large size of the APs and by the impossibility to reach some sites due to very steep terrains (Table 2). Therefore, in two APs the use of binoculars was required to perform the observations (Table 9). Large open spaces offer to goats the possibility to express their natural selective behaviour, searching for the most palatable types of forage (Dias-Silva and Filho 2020). This implies that in some cases the exploratory behaviour of goats made it difficult to observe the animals, as they reached pasture areas sometimes impracticable for the observers. Moreover, in two APs the presence of grazable forestland made the assessment difficult due to the dense vegetation cover.

The limitations experienced in the present study had been already reported in semi-extensive conditions (Battini et al. 2021). However, IOR results in alpine pastures were worse than the ones obtained in semi-extensive conditions. This clearly shows that the assessment in extensive conditions might be much more complicated than in semi-extensive farming systems.

An additional factor that may potentially affect the reliability of animal-based welfare indicators is the scarce training of the assessors, as reported by Katzenberger et al. (2020), who recorded very low ICC when assessing the reliability of group level-indicators in South Tyrolean (North-Eastern Italy) dairy farms. However, such concern does not seem to apply to our study, because the assessors had received a proper training prior the start of data collection, similar to the training provided in the study conducted by Vieira et al. (2018), which was considered sufficient to overcome differences in experience and background of the assessors. Such consideration is also confirmed by the fact that, in the current study, good or even excellent reliability was obtained for individual-level indicators in most of the cases (Table 6). Unexpectedly, the agreement among the assessors for the use of shelters by animals while resting was scarce (Table 5). For this group-level indicator, the wide range of confidence

intervals indicates cases in which the agreement was very poor and cases in which it was excellent. This is in contrast with the results obtained in semi-extensive conditions, where almost a perfect agreement was reached among the assessors for the use of shelters while resting. Such discrepancy may suggest a difficulty for the assessors to identify biotic and abiotic elements of the mountain environment as actual shelters. Therefore, for improving future assessments, it would be advisable to clearly specify which elements characterising the mountain environment should be considered as shelters.

QBA reliability was low compared to the reliability previously obtained by applying this methodology in goats both indoors and outdoors (Grosso et al. 2016). This could be explained by the fact that, as mentioned above, alpine pastures present a challenging environment in terms of visibility. The presence of trees, shrubs, rocks, fog and/or low clouds negatively affect the visibility and sometimes even forces to use binoculars to improve the quality of the observations. Additionally, it is worth mentioning that the assessors had limited experience with QBA, and the training, which was sufficient for other indicators, could have been insufficient for allowing them to properly use this specific indicator (Battini et al. 2016b; Grosso et al. 2016; Minero et al. 2018). A possible insufficient training may also help to understand the unclear distribution of terms (Figure 1), where PC1 and PC2, that conventionally represent valence and arousal, respectively, are swapped. However, descriptors of high arousal (e.g. aggressive, lively) are all on the same side of PC1 (right side), whereas descriptors of low arousal (e.g. relaxed, bored) are on the opposite one (left side). This suggests that for the assessors the evaluation of the level of activity may be easier than the judgement of the valence of goats' emotions. We hypothesise that the assessors applied QBA similarly to quantitative behavioural observations rather than qualitative observations, due to the impossibility to capture subtle clues of behavioural expressions from a distance.

Regarding the assessment of HAR quality, no Avoidance reaction was ever detected, and the three assessors showed perfect agreement in four out of six APs when distinguishing between the Approach and Contact reactions. The FHAT applied in the current study to assess the HAR quality is a modified version of the one described in the AWIN welfare assessment protocol for sheep (AWIN 2015b) and had already been adopted by Battini et al. (2021) in semi-extensive conditions. More recently, a validated FHAT was proposed by Battini et al. (2023), who better described

the classes of reaction and, hopefully, this improvement will contribute to increase the agreement among assessors, while providing sufficiently reliable information on HAR quality.

Welfare problems in extensive alpine farming systems

In the current study, some important issues related to the prevalence of welfare problems in extensive alpine farming systems were highlighted. Despite the low detected reliability, for group-level indicators the main problem was recorded for synchrony at resting and for the use of shelter while resting, which were low in all the three scans. For activities like resting synchronisation, suggested thresholds for a good welfare range from 70 to 90% (Mattiello et al. 2024), which is much higher than what observed in the present study, and also in the similar study we conducted in semi-extensive farming conditions (Battini et al. 2021). Resting patterns of goats are usually well synchronised (Rook and Penning 1991). A possible explanation for the low level of resting synchronisation observed in the current study is the reduced availability of shelters in some pasture areas. In fact, the use of shelter while resting was also quite low, probably because sheltered areas were not enough, and goats do not like to rest in completely open areas where they are more exposed to predation risks, which is an important issue when animals are released at pasture (Aquilani et al. 2022). This is confirmed also in intensive farms, where goats prefer to rest by a wall than in an open area (Battini et al. 2014). It is also possible that we chose the wrong timing to perform the observations. Our observations were performed about two hours after the release of goats on pasture following the morning milking. This decision was based on the hypothesis that, thanks to their browsing feeding habits, goats exhibit short and frequent feeding cycles, alternated by rumination cycles, and therefore after two hours dedicated to feeding activity they would start resting for ruminating. However, (Giannetto et al. 2010) showed that activity rhythms of goats can be affected both by photic and non-photic (i.e. management factors, such as feed distribution or milking), and that the maximum amount of activity occurs during the photo phase; this would suggest that the peak of resting behaviour and synchronisation can be observed during the dark hours, as recently observed by Kok et al. (2023) in cows, and would make it unfeasible to perform such observations. An additional reason for re-thinking the time schedule of the

protocol is that resting observations were performed after the FHAT, and this implies some disturbance and moving of the animals, which may have affected the validity of synchrony at resting. This decision was due to feasibility reasons with the aim to avoid time-consuming welfare assessments that last the whole day. However, this needs to be considered for future developments.

As to the individual-level indicators, no welfare problems were detected, except for hair coat condition as, on average, only the 45.8% of goats showed a good quality of hair coat (Table 4). This percentage is almost half of the percentage reported for goats in semi-extensive (Battini et al. 2021) or intensive/semi-intensive farming conditions (Battini et al. 2016b; Can et al. 2016). A high prevalence of goats with poor hair coat condition was also observed by Leite et al. (2020) while applying the AWIN protocol in semi-intensive and extensive husbandry systems on meat goat does in Brazil. These authors hypothesised a relationship between poor hair coat and pathologies or diseases. Also, (Paskaš et al. 2024) applied the AWIN protocol to assess the welfare of dairy goats bred in large (no access to pasture areas), medium and small (both with access to pasture areas) farms in North Serbia. These authors found poor hair coat condition in the small and medium farms. In our study, poor hair coat conditions may be due to the presence of external parasites (e.g. mange mite, lice) on the skin of goats (Kaufman et al. 2006). In fact, external parasitism is a major concern in grazing goats (Fthenakis and Papadopoulos 2018), and parasites can spread from one flock to another when animals are gathered at pasture. Specifically, the seasonality of rainfall, combined with the warmth of the dry season, can encourage the growth of these ectoparasites, boosting their reproductive rates and increasing the likelihood of infestations in the host (Shilereyo et al. 2022). Furthermore, the temperature also plays an important key-role for the spread of ectoparasites (Amat-Valero et al. 2013) state that the effect of the temperatures on the survival of the above-mentioned pathogens is correlated to the type of ectoparasites involved, and to their life cycle. For example, arthropods parasites live their whole life on the body of the hosting animal, factor that protects them from the external temperature. On the contrary, other species of ectoparasites are free-living for some phases of their life, and for this reason they are exposed to a higher risk of death if the external temperature is not favourable. Considering the prevalence of goats characterised by a normal body condition (on average 90.8%) (Table 4), it seems

reasonable to exclude the presence of haematophagous parasites, which often cause a reduction of feed intake and a loss of weight in the affected animals (Fthenakis and Papadopoulos 2018). However, it is important to specify that, due to their natural browsing behaviour, goats are less sensitive to the internal infections caused by nematodes than other farm species. Indeed, goats feed on the most palatable parts of the plants, often rich in polyphenolic compounds (e.g. tannins), which have anthelmintic effects (Kahiya et al. 2003). Despite this, our results are in line with those reported by Nenadović et al. (2021) in goats reared under extensive farming systems, where a poor hair coat condition was recorded in 62.79% of the animals, caused by parasite infestation.

Body condition was normal in most of the goats. This may be explained by the fact that, at pasture, goats can get access to feed resources at different heights and levels, and this reduces feed competition and limits the presence of very thin and fat goats (Neave and Zobel 2020). This is supported by the higher prevalence of goats in poor body condition in Serbian large goat farms without access to external areas when compared to small and medium farms in which the goats had access to pastures (Paskaš et al. 2024). The latter authors imputed such welfare issue to the strong hierarchy present inside the flock when the animals are gathered, so that dominant goats control the space in front of the feeding rack. Finally, while no cases of faecal soiling were detected in semi-extensive conditions (Battini et al. 2021), in the current study about the 15% of goats suffered from this welfare problem, which may again be imputable to high parasite load due to the very rainy season recorded in 2021 (Sharma et al. 2023).

Given the variety of situations and the inconsistent application of QBA, it is difficult to explain the emotions expressed by goats in relation with the environment or the management. To give a few examples, it could be expected that the mixing of different flocks or the presence of guardian dogs could influence the emotional state of goats. Indeed, goats in different APs showed different emotions, but the reason of it is unclear.

Regarding the quality of HAR, this was good in all the six APs, as no flight responses by the goats were observed while the farmers gathered them. Similar results were recorded in semi-extensive farming conditions (Battini et al. 2021). Unexpectedly, the HAR quality does not seem to be related with the presence or absence of the same stockperson in the valley bottom and in the alpine pastures.

Overall feasibility of the welfare assessment protocol

The evaluation of both individual and group-level welfare indicators in alpine pastures was very time-consuming, due to all the above-mentioned challenges. In two out of six APs, the total assessment time reached 4h 30 min. This is 30 min longer than the application of the same protocol in semi-extensive farming conditions (Battini et al. 2021) and it doubled the time required for the assessment in intensive farming systems (Battini et al. 2016b). These findings highlight the difficulties in assessing animal welfare in alpine pastures and, more in general, in extensive conditions. Regarding feasibility from farmers' perspective, their presence is only required during the application of FHAT, which has to be performed by the farmer when goats are released at pasture (Battini et al. 2023) and during the individual-level assessment in the case of merged flocks to support the identification of goats belonging a specific farm. There are no costs related to the application of the protocol, however additional costs for the farmer could be encountered in the future if Precision Livestock Farming Systems (PLF) are implemented. These costs are supposed to reduce the time required for the assessment and to improve its reliability and feasibility. Our results suggest that the feasibility of the animal-based welfare indicators may affect their reliability (Mattiello et al. 2019). As suggested by Plesch et al. (2010), indicators need to be simple and feasible to be assessed, to increase and improve the agreement between the assessors. For these reasons, it is necessary to adapt already developed indicators or to identify new ones to improve their feasibility when applied to mountain pastures, where assessing the welfare could be particularly challenging. In the next future, the development of PLF technologies for a continuous and non-invasive monitoring of behavioural (e.g. synchronisation level), physiological (e.g. body temperature) and nutritional (e.g. BCS) parameters, as well as of indicators of emotional states (e.g. tail or ear postures) of goats will be of great help for the evaluation of their welfare. Furthermore, the collection of individual indicators that requires herding and restraining the animals is time-consuming and potentially stressful for extensively managed animals (Spigarelli et al. 2020), and PLF could overcome this further constraint. In agreement with (Vieira et al. 2018), our results underline that reliability is not an absolute attribute of an indicator, as it is strongly affected by the situation in which the indicator is collected. This implies that all the indicators need to be tested in each context where the assessment takes place.

Further considerations are relevant for the group assessment of animals kept on pasture that, even if limiting the accuracy of the evaluation, cannot be ignored as they reflect the reality of the situation. It is common for animals of different categories and ages (males, females, young, adults, subadults) to be grazed together, and it is often not possible to distinguish or separate them (this is an issue for example in the application of QBA). Goats, being gregarious animals, both influence and are influenced by the behaviour of others, and these potential differences in the management of groups complicate comparisons among farms. The presence of young animals or males can distinctly alter the overall behaviour of the group and affect the results for example of tests assessing the quality of HAR (Battini et al. 2016c). In particular, males can have a dominant behaviour over the females (Haenlein et al. 1992), affecting the stability of the group, and so the reliability of the welfare assessment. Furthermore, the presence of young animals can also affect both the feasibility and the reliability of the assessment as the presence of kids may change the activities of the mothers (Miranda-de la Lama and Mattiello 2010). Additionally, merging flocks from different farms, as commonly practiced in alpine summer ranges, may alter the hierarchy and the cohesion within the group. Goats compete for resources (Miranda-de la Lama and Mattiello 2010) and the introduction of new members into a group can affect the social stability, leading to an increase of agonistic behaviours (Addison and Baker 1982; Barroso et al. 2000) and possibly altering the synchronisation of the activities, thus affecting the reliability of welfare evaluations. For future developments, it is also important to consider that the practice of mixing multiple species for grazing together on the same pasture (mixed-livestock stocking) is growing (Anderson et al. 2012). Goats can obviously be distinguished from other species, so that the feasibility is not affected, but changes in their behaviour should be carefully considered (Bojkovski et al. 2014; Mahieu et al. 2020).

Conclusions

The application of the proposed modified version of the AWIN protocol in alpine summer pasture ranges provided some potentially useful information about the welfare of goats in these farming systems but posed some issues in terms of feasibility and reliability, especially when considering group-level indicators. While most of the individual-level indicators can be included as they are for welfare evaluation in alpine summer ranges, some methodological issues still need

to be solved for group level indicators. For example, further studies are required to identify the best observation schedule to assess behavioural synchronisation, and new feasible and reliable indicators of goats' emotional state need to be validated. Recent developments of PLF may hopefully contribute to provide valid, feasible, and reliable welfare indicators to be included in assessment protocols for these peculiar and welfare-challenging farming systems.

Last comment regards the involvement of key stakeholders, including farmers, researchers, and policymakers, that is essential in refining welfare protocols. Farmers provide practical insights based on day-to-day animal management, while researchers offer scientific expertise on animal behaviour and welfare. Then policymakers play a relevant role in establishing and enforcing regulatory frameworks that support welfare standards. Collaboration among these stakeholders ensures the development of evidence-based, effective welfare protocols that benefit both animal welfare and the farming system.

Acknowledgements

The authors would like to thank the farmers who allowed us to visit their farms. We also acknowledge the help of Dr. Andrea Bonzanino, Dr. Mauro Masino and Dr. Gabriele My for the collection of data.

Ethical approval

This study was approved by the Bioethics Committee of the University of Turin (Italy) (protocol n° 0587791).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research received no external funding.

ORCID

Manuela Renna  <http://orcid.org/0000-0003-4296-7589>
 Benedetta Torsiello  <http://orcid.org/0000-0002-3859-2609>
 Silvana Mattiello  <http://orcid.org/0000-0003-1885-949X>
 Luca Battaglini  <http://orcid.org/0000-0002-2136-3826>
 Monica Battini  <http://orcid.org/0000-0002-6134-7759>

Data availability statement

The data that support the findings of this study are available from the corresponding author [S.M.] upon reasonable request.

References

- Addison WE, Baker E. 1982. Agonistic behavior and social organization in a herd of goats as affected by the introduction of non-members. *Appl Anim Ethol.* 8(6):527–535. doi:10.1016/0304-3762(82)90216-4.
- Amat-Valero M, Calero-Torralbo MA, Valera F. 2013. Temperature during the free-living phase of an ectoparasite influences the emergence pattern of the infective phase. *Parasitology.* 140(11):1357–1367. doi:10.1017/S0031182013000929.
- Anderson DM, Fredrickson EL, Estell RE. 2012. Managing livestock using animal behavior: mixed-species stocking and flocks. *Animal.* 6(8):1339–1349. doi:10.1017/S175173111200016X.
- Aquilani C, Confessore A, Bozzi R, Sirtori F, Pugliese C. 2022. Review: precision livestock farming technologies in pasture-based livestock systems. *Animal.* 16(1):100429. doi:10.1016/j.animal.2021.100429.
- AWIN. 2015a. AWIN welfare assessment protocol for goats. University of Milan. doi:10.13130/AWIN_GOATS_2015.
- AWIN. 2015b. AWIN welfare assessment protocol for sheep. University of Milan. doi:10.13130/AWIN_SHEEP_2015.
- Barroso FG, Alados CL, Boza J. 2000. Social hierarchy in the domestic goat: effect on food habits and production. *Appl Anim Behav Sci.* 69(1):35–53. doi:10.1016/S0168-1591(00)00113-1.
- Barry RG. 2002. Past and potential future changes in mountain environments. In: Beniston M, editor. *Mountain environments in changing climates.* Routledge; p. 3–33.
- Bateson M, Martin P. 2021. *Measuring behaviour - An introductory guide.* 4th ed. Cambridge University Press.
- Battaglini LM, Quaderno SoZooAlp. 2007. Sistemi ovicapriini nelle Alpi Occidentali: realtà e prospettive. In: Battaglini LM, editor. *L'allevamento Ovino e Caprino Nelle Alpi: Tra Valenze Eco-Culturali e Sostenibilità Economica.* 4th ed. Trento: Nuove Arti Grafiche; p. 9–23.
- Battini M, Barbieri S, Vieira A, Can E, Stilwell G, Wemelsfelder F, Mattiello S. 2016a. Inter-observer reliability in applying Qualitative Behaviour Assessment to dairy goats: the importance of assessor training. In: Dwyer C, Haskell M, Sandilands V, editors. *Proceedings of the 50th Congress of the International Society for Applied Ethology.* Edinburgh, UK; p. 456.
- Battini M, Barbieri S, Vieira A, Stilwell G, Mattiello S. 2016b. Results of testing the prototype of the AWIN welfare assessment protocol for dairy goats in 30 intensive farms in Northern Italy. *Ital J Anim Sci.* 15(2):283–293. doi:10.1080/1828051X.2016.1150795.
- Battini M, Barbieri S, Waiblinger S, Mattiello S. 2016c. Validity and feasibility of Human-Animal Relationship tests for on-farm welfare assessment in dairy goats. *Appl Anim Behav Sci.* 178:32–39. doi:10.1016/j.applanim.2016.03.012.
- Battini M, Renna M, Giammarino M, Battaglini L, Mattiello S. 2021. Feasibility and reliability of the AWIN welfare assessment protocol for dairy goats in semi-extensive farming conditions. *Front Vet Sci.* 8(October):731927. doi:10.3389/fvets.2021.731927.
- Battini M, Renna M, Torsiello B, Battaglini L, Mattiello S. 2023. Is the familiar human approach test a valid method for evaluating the quality of human-goat relationship in pasture-based systems? *Appl Anim Behav Sci.* 269:106088. doi:10.1016/j.applanim.2023.106088.

- Battini M, Vieira A, Barbieri S, Ajuda I, Stilwell G, Mattiello S. 2014. Invited review: animal-based indicators for on-farm welfare assessment for dairy goats. *J Dairy Sci.* 97(11):6625–6648. <http://linkinghub.elsevier.com/retrieve/pii/S0022030214006444>. doi:10.3168/jds.2013-7493.
- Bender O, Borsdorf A, Fischer A, Stotter J. 2011. Mountains under climate and global change conditions – Research results in the alps. In: *Climate change - Geophysical foundations and ecological effects*. InTech.
- Bojkovski D, Štuhec I, Kompan D, Zupan M. 2014. The behavior of sheep and goats co-grazing on pasture with different types of vegetation in the karst region 1. *J Anim Sci.* 92(6):2752–2758. doi:10.2527/jas.2013-7199.
- Can E, Vieira A, Battini M, Mattiello S, Stilwell G. 2016. On-farm welfare assessment of dairy goat farms using animal-based indicators: the example of 30 commercial farms in Portugal. *Acta Agric Scand A: Anim Sci.* 66(1):43–55. doi:10.1080/09064702.2016.1208267.
- Cannone N, Sgorbati S, Guglielmin M. 2007. Unexpected impacts of climate change on alpine vegetation. *Front Ecol Environ.* 5(7):360–364. doi:10.1890/1540-9295(2007)5[360:UOCCO]2.0.CO;2.
- Czycholl I, Klingbeil P, Krieter J. 2019. Interobserver reliability of the animal welfare indicators welfare assessment protocol for horses. *J Equine Vet Sci.* 75:112–121. doi:10.1016/j.jevs.2019.02.005.
- Dias-Silva TP, Filho AL. 2020. Sheep and goat feeding behavior profile in grazing systems. *Acta Sci.* 43:e51265. doi:10.4025/actascianimsci.v43i1.51265.
- Doughty AK, Coleman GJ, Hinch GN, Doyle RE. 2017. Stakeholder perceptions of welfare issues and in Australia. *Animals.* 7(4):28. doi:10.3390/ani7040028.
- Fthenakis GC, Papadopoulos E. 2018. Impact of parasitism in goat production. *Small Rumin Res.* 163:21–23. doi:10.1016/j.smallrumres.2017.04.001.
- Giammarino M, Mattiello S, Battini M, Quatto P, Battaglini LM, Vieira ACL, Stilwell G, Renna M. 2021. Evaluation of inter-observer reliability of animal welfare indicators: which is the best index to use? *Animals (Basel).* 11(5):1445. doi:10.3390/ani11051445.
- Giannetto C, Casella S, Caola G, Piccione G. 2010. Photoc and non-photoc entrainment on daily rhythm of locomotor activity in goats. *Anim Sci J.* 81(1):122–128. doi:10.1111/j.1740-0929.2009.00707.x.
- Grosso L, Battini M, Wemelsfelder F, Barbieri S, Minero M, Dalla Costa E, Mattiello S. 2016. On-farm qualitative behaviour assessment of dairy goats in different housing conditions. *Appl Anim Behav Sci.* 180:51–57. doi:10.1016/j.applanim.2016.04.013.
- Gwet KL. 2008. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol.* 61(Pt 1):29–48. doi:10.1348/000711006X126600.
- Haenlein GFW, Caccese R, Sammelwitz PH. 1992. Behaviour. In: *Goat extension handbook*. Newark, NJ: University of Delaware.
- IBM Corp. Released. 2019. IBM SPSS statistics for windows, Version 26.0. Armonk, NY: IBM Corp.
- Kahiya C, Mukaratirwa S, Thamsborg SM. 2003. Effects of *Acacia nilotica* and *Acacia karoo* diets on *Haemonchus contortus* infection in goats. *Vet Parasitol.* 115(3):265–274. doi:10.1016/S0304-4017(03)00213-9.
- Katzenberger K, Rauch E, Erhard M, Reese S, Gauly M. 2020. Evaluating the need for an animal welfare assurance programme in South Tyrolean dairy farming. *Ital J Anim Sci.* 19(1):1146–1156. doi:10.1080/1828051X.2020.1823897.
- Kaufman PE, Koehler PG, Butler JF. 2006. External parasites of sheep and goats. 2006(6). Gainesville, FL: EDIS. doi:10.32473/edis-ig129-2006.
- Kaurivi YB, Hickson R, Laven R, Parkinson T, Stafford K. 2020. Developing an animal welfare assessment protocol for cows in extensive beef cow-calf systems in New Zealand. Part 2: Categorisation and Scoring of Welfare Assessment Measures. *Animals (Basel).* 10(9):1592. doi:10.3390/ani10091592.
- Kaurivi YB, Laven R, Hickson R, Parkinson T, Stafford K. 2020. Developing an animal welfare assessment protocol for cows in extensive beef cow – Calf systems in New Zealand. Part 1: Assessing the Feasibility of Identified Animal Welfare Assessment Measures. *Animals.* 10(9):1597. doi:10.3390/ani10091597.
- Kaurivi YB, Laven R, Hickson R, Stafford K, Parkinson T. 2019. Identification of suitable animal welfare assessment measures for extensive beef systems in New Zealand. *Agriculture.* 9(3):66. doi:10.3390/agriculture9030066.
- Knierim U, Winckler C. 2009. On-farm welfare assessment in cattle: validity, reliability and feasibility issues and future perspectives with special regard to the Welfare Quality® approach. *Anim Welf.* 18(4):451–458. doi:10.1017/S0962728600000865.
- Kok A, van Kneegsel A, Bokkers EAM, Kemp B, Thorup V. 2023. Exploring synchrony of lying on commercial dairy farms in relation to management. *Appl Anim Behav Sci.* 262:105906. doi:10.1016/j.applanim.2023.105906.
- Leite LO, Stamm FO, Corrêa Maceno MA, Camarinha Filho AJ, Garcia RCM. 2020. On-farm welfare assessment in meat goat does raised in semi-intensive and extensive systems in semi-arid regions of Ceará, Northeast, Brazil. *Cienc Rural.* 50(10):1–13. doi:10.1590/0103-8478cr20190745.
- Lucchesi V, Mattiello S. 2008. Indagine sul benessere dei bovini di razza “Pezzata Rossa d’Oropa” nel Biellese. In: *Benessere animale e sistemi zootecnici alpini*. Vol. 5. Quaderno SoZooAlp; p. 143–153.
- Mahieu M, Arquet R, Fleury J, Bonneau M, Mandonnet N. 2020. Mixed grazing of adult goats and cattle: lessons from long-term monitoring. *Vet Parasitol.* 280:109087. doi:10.1016/j.vetpar.2020.109087.
- Mattiello S, Battini M, De Rosa G, Napolitano F, Dwyer C. 2019. How can we assess positive welfare in ruminants? *Animals (Basel).* 9(10):758. doi:10.3390/ani9100758.
- Mattiello S, Celozzi S, Soli FM, Battini M. 2024. Exploring positive welfare measures: preliminary findings from a prototype protocol in loose housing dairy cattle farms. *Front Vet Sci.* 11:1368363. doi:10.3389/fvets.2024.1368363.
- Minero M, Dalla Costa E, Dai F, Canali E, Barbieri S, Zanella A, Pascuzzo R, Wemelsfelder F. 2018. Using qualitative behaviour assessment (QBA) to explore the emotional state of horses and its association with human-animal relationship. *Appl Anim Behav Sci.* 204:53–59. doi:10.1016/j.applanim.2018.04.008.
- Miranda-de la Lama GC, Mattiello S. 2010. The importance of social behaviour for goat welfare in livestock farming. *Small Rumin Res.* 90(1–3):1–10. doi:10.1016/j.smallrumres.2010.01.006.

- Munoz C, Campbell AJD, Hemsworth PH, Doyle RE. 2017. Animal-based measures to assess the welfare of extensively managed ewes. *Animals* (Basel). 8(1):2. doi:10.3390/ani8010002.
- Munoz C, Campbell AJD, Hemsworth PH, Doyle RE. 2019. Evaluating the welfare of extensively managed sheep. *PLoS One*. 14(6):e0218603. doi:10.1371/journal.pone.0218603.
- Neave HW, Zobel G. 2020. Personality of dairy goats affects competitive feeding behaviour at different feeder heights. *Small Rumin Res*. 192:106222. doi:10.1016/j.smallrumres.2020.106222.
- Nenadović K, Ilić T, Jovanović N, Bugarski D, Vučinić M. 2021. Welfare of native goat breeds in Serbia—Emphasis on parasitological infections. *Front Vet Sci*. 8(August):678880. doi:10.3389/fvets.2021.678880.
- Nenadovic K, Karac P, Vucinic M, Teodorovic R, Živanov D, Trailovic R, Beckei Z, Jankovic L. 2020. Assessment of the welfare of extensively managed autochthonous sheep breed Vlasicka Zackel using animal based measurements. *Acta Veterinaria-Beograd*. 70(2):207–218. doi:10.2478/acve-2020-0015.
- Paskaš S, Miočinović J, Savić M, Djukić-Stojčić M, Pihler I, Becskei Z. 2024. Welfare assessment on different-sized dairy goat farms in the northern Serbian Province of Vojvodina. *J Appl Anim Welf Sci*. 27(2):210–222. doi:10.1080/10887705.2022.2025537.
- Phythian CJ, Cripps PJ, Michalopoulou E, Jones PH, Grove-White D, Clarkson MJ, Winter AC, Stubbings LA, Duncan JS. 2012. Reliability of indicators of sheep welfare assessed by a group observation method. *Vet J*. 193(1):257–263. doi:10.1016/j.tvjl.2011.12.006.
- Plesch G, Broerkens N, Laister S, Winckler C, Knierim U. 2010. Reliability and feasibility of selected measures concerning resting behaviour for the on-farm welfare assessment in dairy cows. *Appl Anim Behav Sci*. 126(1–2):19–26. doi:10.1016/j.applanim.2010.05.003.
- Redbo I, Ehrlemark A, Redbo-Torstensson P. 2001. Behavioural responses to climatic demands of dairy heifers housed outdoors. *Can J Anim Sci*. 81(1):9–15. doi:10.4141/A00-071.
- Richmond SE, Wemelsfelder F, de Heredia IB, Ruiz R, Canali E, Dwyer CM. 2017. Evaluation of animal-based indicators to be used in a welfare assessment protocol for sheep. *Front Vet Sci*. 4(December):210. doi:10.3389/fvets.2017.00210.
- Riggio G, Angori E, Menchetti L, Diverio S. 2023. The link between the perception of animal welfare and the emotional response to pictures of farm animals kept in intensive and extensive husbandry systems: An Italian survey. *Vet Sci*. 10(11):652. doi:10.3390/vetsci10110652.
- Rook AJ, Penning PD. 1991. Synchronisation of eating, ruminating and idling activity by grazing sheep. *Appl Anim Behav Sci*. 32(2–3):157–166. doi:10.1016/S0168-1591(05)80039-5.
- Sapkota S, Laven R, Müller K, Kells N. 2020. Animal welfare assessment: can we develop a practical, time-limited assessment protocol for pasture-based dairy cows in New Zealand? *Animals* (Basel). 10(10):1918. doi:10.3390/ani10101918.
- Sharma A, Sharma S, Kour S, Avatsingh AU, Perveen K, Alsulami JA, Singh N. 2023. Gastrointestinal nematodes and protozoa in small and large ruminants from rural agro-climatic regions of northern India. *Diversity* (Basel). 15(11):1131. doi:10.3390/d15111131.
- Shilereyo M, Magige F, Ranke PS, Ogutu JO, Røskoft E. 2022. Ectoparasite load of small mammals in the Serengeti Ecosystem: effects of land use, season, host species, age, sex and breeding status. *Parasitol Res*. 121(3):823–838. doi:10.1007/s00436-022-07439-1.
- Spigarelli C, Zuliani A, Battini M, Mattiello S, Bovolenta S. 2020. Welfare assessment on pasture: a review on animal-based measures for ruminants. *Animals* (Basel). 10(4):609. doi:10.3390/ani10040609.
- Sweeney S, Regan Á, McKernan C, Benson T, Hanlon A, Dean M. 2022. Current consumer perceptions of animal welfare across different farming sectors on the Island of Ireland. *Animals* (Basel). 12(2):185. doi:10.3390/ani12020185.
- Temple D, Manteca X. 2020. Animal welfare in extensive production systems is still an area of concern. *Front Sustain Food Syst*. 4:545902. doi:10.3389/fsufs.2020.545902.
- Torsiello B, Giammarino M, Battaglini L, Battini M, Mattiello S, Quatto P, Renna M. 2023a. Evaluation of inter-observer reliability of dichotomous and four-level animal-based indicators. 74th Annual Meeting of the European Federation of Animal Science; Lyon, France. p. 889.
- Torsiello B, Giammarino M, Battaglini L, Battini M, Mattiello S, Quatto P, Renna M. 2023b. Evaluation of inter-observer reliability of animal-based welfare indicators in the case of trichotomous variables with two or more observers. *Ital J Anim Sci*. 22:148–149.
- Torsiello B, Giammarino M, Quatto P, Battini M, Mattiello S, Battaglini L, Renna M. 2024. Evaluation of inter-observer reliability in the case of trichotomous and four-level animal-based welfare indicators with two observers. *Ital J Anim Sci*. 23(1):938–960. doi:10.1080/1828051X.2024.2367681.
- Vieira A, Battini M, Can E, Mattiello S, Stilwell G. 2018. Inter-observer reliability of animal-based welfare indicators included in the Animal Welfare Indicators welfare assessment protocol for dairy goats. *Animal*. 12(9):1942–1949. doi:10.1017/S1751731117003597.
- Villalba JJ, Manteca X, Vercoe PE, Maloney SK, Blache D. 2016. Integrating nutrition and animal welfare in extensive systems. In: Phillips CJC, editor. *Nutrition and the Welfare of Farm Animals*. Berlin Heidelberg: Springer; p. 135–163.
- Webster J. 2005. The assessment and implementation of animal welfare: theory into practice. *Rev Sci Tech Off Int Epiz*. 24:723–734.
- Welfare Quality®. 2009a. Welfare Quality® assessment protocol for cattle. Lelystad, The Netherlands.
- Welfare Quality®. 2009b. Welfare Quality® assessment protocol for pigs (sow and piglets, growing and finishing pigs). Lelystad, Netherlands.
- Windsor PA. 2021. Progress with livestock welfare in extensive production systems: lessons from Australia. *Front Vet Sci*. 8:827685. doi:10.3389/fvets.2021.674482.
- Zemp M, Leuenberger H, Künzi N, Blum JW. 1989. Influence of high altitude grazing on productive and physiological traits of dairy cows. *J Anim Breed Genet*. 106(1–6):278–288. doi:10.1111/j.1439-0388.1989.tb00242.x.