



Original Research

A Novel Chart to Score Rumen Fill Following Simple Sequential Instructions

Marie Schneider^{1,2,*}, Leonie Hart¹, Eva Gallmann², Christina Umstätter³¹ Department of Competitiveness and System Evaluation, Agroscope, CH-8356 Ettenhausen, Switzerland² Institute of Livestock Systems Engineering, University of Hohenheim, DE-70599 Stuttgart, Germany³ Thünen Institute of Agricultural Technology, Thünen Institute, DE-38116 Braunschweig, Germany

ARTICLE INFO

Article history:

Received 7 October 2021

Revised 9 February 2022

Accepted 16 February 2022

Keywords:

Dairy cow
Feed intake
Flowchart
Pasture
Rumen fill

ABSTRACT

Knowledge about feed intake and adequate nutritional status of dairy cows is important to achieve high milk yield in grazing systems. A possible, but subjective, method to estimate changes in grazing intake could be rumen fill scoring. Thus, a novel flow chart, based on existing criteria for rumen fill scoring was developed to simplify this approach.

The first objective of this study was to develop a flowchart to reduce training time for untrained observers and to support the observers in their decision-making process. Therefore, the quality of scoring of four trained observers that used the criteria, originally published by Zaaijer and Noordhuizen (2003), was assessed first. After that, the interobserver reliability of trained observers (control group) and untrained observers that used the novel flowchart was determined. The novel flowchart was tested twice. The second version of the flowchart included the feedback of the first round. To assess if rumen fill scoring is a suitable method for detecting changes in grazing intake was the second objective of this study. Therefore, it was investigated if rumen fill scores could be used to detect decreasing feed intake during a 6-d grazing period with limited feed allocation.

The experiments demonstrated that overall agreement between trained observers who used the criteria originally published was moderate ($\kappa_F = 0.53$). The overall agreement of untrained observers who used the novel flowchart increased from fair ($\kappa_F = 0.40$) to substantial ($\kappa_F = 0.66$) between the two versions of the flowchart. Finally, it was found that rumen fill scores decreased, according to decreasing feed allocation for this small sample. In conclusion, the novel flowchart simplified rumen fill scoring, thus facilitating the assessment of rumen fill as a potential method for estimating changes in pasture intake in research and practice.

© 2022 The Authors. Published by Elsevier Inc. on behalf of The Society for Range Management. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Introduction

Grazing systems in dairy farming will be more important in the future as housing systems with integrated pasture allowance are often associated with higher animal welfare (Dillon et al. 2008; Hofstetter et al. 2014). Therefore, grazing systems are more and more supported by society and policy makers (Die Borchert Kommission 2020). Nevertheless, despite the aspiration of high standards in animal welfare, farming systems have to be profitable. Therefore, a high milk yield is one of the necessary factors (Shalloo et al. 2004), which is mainly influenced by the consistent feed intake of high-quality feed. Thus, assessing changes in feed intake on pasture could be a helpful approach for managing herbage utilization

and detecting inadequate feed allocation on an individual cow level.

Numerous methods are available for estimating feed intake on pasture, such as the n-alkane method (Mayes et al. 1986; Dillon and Stakelum 1988), measurement of grass height through rising plate meters (RPM) following a conversion into herbage mass, herbage dry matter intake estimation models (Rombach et al. 2019), and GrazeIn model by Delagarde et al. (2011).

Although there are numerous methods of estimating feed intake of cows on pasture individually or groupwise, none of them is suitable for practical use. Particularly for long-term use in practice, measurements should be noninvasive, repeatable, user-friendly, and cheap (Garnick et al. 2018). Therefore, changes in body condition scores (BCSs) are often used in practice (Zaaijer and Noordhuizen 2003). However, the response rate of this parameter is too slow to use it for adjusting the ration composition (Zaaijer and Noordhuizen 2003) or for detecting a short-term decrease in feed in-

* Correspondence and current address: Marie Schneider, Lindenfelser Weg 11, 64385 Reichelsheim, Germany, 00491606704744.

E-mail address: marieschnellbaecher@oulook.de (M. Schneider).

take. Therefore, Zaaier and Noordhuizen (2003) developed a scoring system that can be used to detect changes in feed intake. One part of that system was a subjective scoring of rumen fill, by visually observing the left paralumbar fossa. The scoring system was based on five categories. According to the authors, rumen fill is influenced by dry matter intake (DMI), ration composition, digestion, and passage rate.

These developed rumen fill scores of Zaaier and Noordhuizen (2003) were later validated by Burfeind et al. (2010). The authors calculated a Spearman's rank correlation (r_s) to find a relation between rumen fill scores and DMI at the barn, which correlated moderately ($r_s = 0.68$). In accordance with their findings, Hart et al. (2022) showed that rumen fill scores decreased by 0.73 scores, if feed allocation on pasture was reduced to 80% of the cow's demand. Furthermore, Burfeind et al. (2010) measured a high variability in depth of paralumbar fossa within a short time of 70 minutes. Götze et al. (2019) found that the lactation day influences the sensitivity of the rumen fill score. During the first 9 d post partum, changes in rumen fill score are much more influenced by changes in DMI than in the period 24 d ante partum (Götze et al. 2019). Furthermore, Burfeind et al. (2010) stated that the interobserver reliability of trained observers, measured by Cohen's Kappa (κ), was substantial ($\kappa = 0.68$).

The present study was conducted as part of a larger study on feed intake of grazing cows, using the rumen fill scores as a potential indicator for evaluating short-term changes in feed intake on pasture. To the best of our knowledge, this is one of the first studies on rumen fill scoring that focused on the detection of changes in grazing intake. The method was not used to estimate dry matter feed intake in general but rather to detect inadequate feed intake on pasture. During the preparation of this study, the idea of a flowchart based on the criteria of Zaaier and Noordhuizen (2003) was born to make the scoring decision easier and to improve the agreement between the observers. In general, subjective scoring systems are known to require intensive training in order to be reliable and comparable. Previous studies have shown that reliability increased according to the training time, while simplification reduced the training time (Brenninkmeyer et al. 2007; D'Eath 2012).

Therefore, the objectives of the present study were to evaluate if the developed novel flowchart reduces training time and simplifies the scoring of rumen fill, as well as if the rumen fill scores are suitable as indicators for changes in feed intake on pasture. Consequently, the experimental periods were designed to focus on 1) the quality of scoring of trained observers that did not use the chart, 2) the quality of scoring of trained and untrained observers that used the novel flowchart, and 3) the suitability of rumen fill scoring to detect decreasing feed intake on pasture.

Methods

The study was conducted at the research farm of Agroscope in Tänikon (Ettenhausen, Thurgau, Switzerland). All animal experiments were approved by the local authorities (TG01/19), according to animal protection law. The study was split into three different parts, as described previously. The experiments were undertaken over a consecutive period from July 7 until July 19 and on 2 single days, on August 12 and September 17, 2019.

Animals and Housing

In this study, 20 Brown Swiss cows were used for rumen fill scoring. Thirteen cows were multiparous and seven were primiparous. Their mean BCS, scored once before the experimental period in July, was 3.05 ± 0.50 (standard deviation [SD]). The mean

bodyweight, measured with a tape measure every second day during the experimental period in July, was 643.31 ± 67.94 kg (SD). The cows were divided into 2 groups of 10 cows each and were kept in 2 separate areas of the barn. The free stall barn was equipped with full concrete floors and low bed cubicles filled with straw. In the barn, the cows were fed with hay and corn pellets. In addition, they received concentrate individually at a feeding station. During the summer months, the animals were turned out daily for 11 h during the night. In autumn, they grazed 9 h during the day. The groups grazed in separate plots.

Recorded Parameters and Experimental Design

Scoring

The criteria for rumen fill scoring were defined by Zaaier and Noordhuizen (2003) and ranged from score 1 to 5. To increase the accuracy of the scoring system, additional steps of 0.5 were added. Each observer group consisted of four observers, following the approach of Burfeind et al. (2010), who used three trained observers in their study. When the flowchart was tested, only three trained observers (Observers 1, 2, and 3) acted as a control group (Fig. 1). The fourth trained observer (Observer 4) was excluded as he or she was not available anymore.

The process of scoring was similar in each period of the study. The cows only received a small amount of corn pellets before scoring, when they were fixated at the feeding gate, to prevent the impact of hay intake on rumen fill. Due to the fixation, the cows were standing on a level, full-concrete floor. To reduce error, it was important that all observers carried out the scoring at the same time, because of the variability of the paralumbar fossa. Therefore, all observers were standing on the left side of the cow, next to her pelvic bone, assessing the left paralumbar fossa. The given scores were noted down by hand on paper lists clipped to clipboards and kept hidden from other observers.

Quality assessment of scoring

Interobserver reliability of trained observers To quantify the interobserver reliability of trained observers, Observers 1–4 scored rumen fill of each cow on the basis of the criteria of Zaaier and Noordhuizen (2003) with the additional steps of 0.5 included. The novel flowchart was not available to them at that point. Furthermore, they scored rumen fill over 6 d with different levels of feed allocation.

Two of the four observers (Observers 1 and 2) had knowledge about the feed allocation. However, the observers did not know previous scores during scoring. Additionally, the scoring dates were not consistent with the allocation days on pasture.

Control group and untrained observers After the previous experimental period, a flowchart for scoring rumen fill based on the scoring system, defined by Zaaier and Noordhuizen (2003) (Fig. 2), was developed and evaluated. The novel flowchart guides the observer through his or her decision-making process for a specific rumen fill score in simple sequential instructions.

During the development, the chart was tested twice. On both test days, four untrained observers and the control group (Observers 1–3) scored rumen fill of the 20 cows. The untrained observers of the first test day, using the first chart version (see Fig. 2) were Observers 5–8 (see Fig. 1). Observers 9–12 were the untrained observers on the second test day (Supplementary 1). The second chart version (see Supplementary 1) included modifications suggested by a feedback round after the first test day.

The untrained observers were scientists of different ages (23–58 yr) and with different knowledge on livestock. Some of them mainly worked in crop research and others in labor science or process engineering in animal husbandry. None of them was working with cows on a regular basis.

Details of the study	Quality assessment of scoring			Quality assessment of estimating feed intake
	Interobserver reliability of trained observers	Interobserver reliability of untrained observers and the control group		
Scoring method	Criteria of Zaaijer and Noordhuizen (2003)*	Flow Chart Version 1	Flow Chart Version 2	Criteria of Zaaijer and Noordhuizen (2003)*
Introduced modification	Steps of 0.5	Pictures and the hand width as measurement tool	Additional pictures included	Steps of 0.5
Observer groups	Observers 1 - 4	Observers 5 - 8 & Control Group (Observers 1 - 3)	Observers 9 - 12 & Control Group (Observers 1 - 3)	Observers 1 - 4
Database	20 cows, observed on 6 days = 120 observations per observer	20 cows, one times observed by 4 naive untrained observers & 3 control group members	20 cows, one times observed by 4 naive untrained observers & 3 control group members	20 cows, observed by 4 observers on each of the 6 days = 80 observations per day

*Zaaijer, D., Noordhuizen, J.P.T.M., 2003. A novel scoring system for monitoring the relationship between nutritional efficiency and fertility in dairy cows. *Irish Veterinary Journal*, 145–151.

Figure 1. Schematic overview over the components of the study, including scoring methods, introduced modifications, observer groups, and the database.

This part of the study should show if the flowchart simplified rumen fill scoring to reduce training time and to make it more objective by using a defined hand width as a measuring aid (see [Supplementary 1](#)).

Quality assessment of detecting feed intake changes

To assess the potential of detecting changes in feed intake, four trained observers scored rumen fill over 6 d while herbage availability on pasture decreased over time. This resulted in six different levels of feed supply. The treatment was part of a larger study, looking at cow behavioral responses to decreasing grazing allocation. Therefore, during an experimental period of 12 d, every sixth day, fresh plots were allocated to the two groups of cows. The experimental design of the larger study demanded a timely shift of 2 d between the two groups, for allocating a new plot ([Fig. 3](#)). Therefore, during the experimental period of our study the two cow groups grazed on five different plots, in total.

The size of each grazing plot was calculated 1 d before the new plot was allocated. It was targeted to supply 80% of the estimated dry matter feed demand for 10 cows over 6 grazing days. The amount of feed allocation was calculated on the basis of three factors: pregrazing herbage mass, measured by an RPM (Grasshopper, TrueNorth Technologies, Shannon, Ireland); expected grass growth over 6 d, determined on the basis of guidelines of [Mosimann and Stettler \(2004\)](#); and expected feed intake on pasture in a part-time grazing system with 11 h of grazing per day. The compressed grass height was measured every day post grazing using the RPM in order to monitor the herbage availability. Over the 6-d periods, the amount of available herbage on pasture was targeted to decrease until below the required daily feed demand of the cows.

Additionally, the amount of feed allocated in the barn was reduced to 80% as well except for Day 1, when the cows were fed ad libitum to ensure that they were not hungry when turned out onto a new pasture plot. That means on Days 2–6, the amount of feed dropped from 6.5 kg hay and 3.5 kg corn pellets to 5.2 kg hay and

2.8 kg corn pellets per day and cow, whereas on Day 1, the cows received hay ad libitum and 4.0 kg corn pellets each.

To avoid bias from the observers while scoring, the scoring days were distributed over the complete experimental period (see [Fig. 3](#)). In total, Observers 1–4 scored rumen fill on 6 d. It was ensured that each of the scoring days corresponded with one of the grazing Days 1–6 to account for the different levels of feed supply. The scoring had to be conducted soon after the daily grazing period because the method detects the feed intake of the last 2–6 h ([Hulsen 2010](#)). As cows grazed during the night, the observers scored rumen fill after the morning milking. In this part of the study, the observers did not use the flowchart but rather the criteria of [Zaaijer and Noordhuizen \(2003\)](#) (see [Fig. 1](#)), as the idea of a scoring chart was born at this point.

To check the influence of the gradually decreasing feed availability on the cows' feed intake, the daily milk yield of each cow was recorded. The milk yields of the morning and evening milking were recorded by the milking parlor and saved as daily sum.

Data analysis and statistics

The scores, noted by hand, were checked and transferred to *Excel 2016* (Microsoft Corporation, Redmond, WA). The statistical analysis was performed using R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

Quality assessment of scoring

To analyze the interobserver reliability pairwise between each possible pair of observers, Spearman's rank correlation (r_s) and Cohen's weighted Kappa (κ) were calculated. The correlation r_s of all observer pairs was calculated to identify any linear relationship between them. The criteria for the interpretation of r_s were in agreement with the definition of [Hinkle et al. \(2003\)](#): 0.00–0.30 = negligible correlation, 0.31–0.50 = low corre-

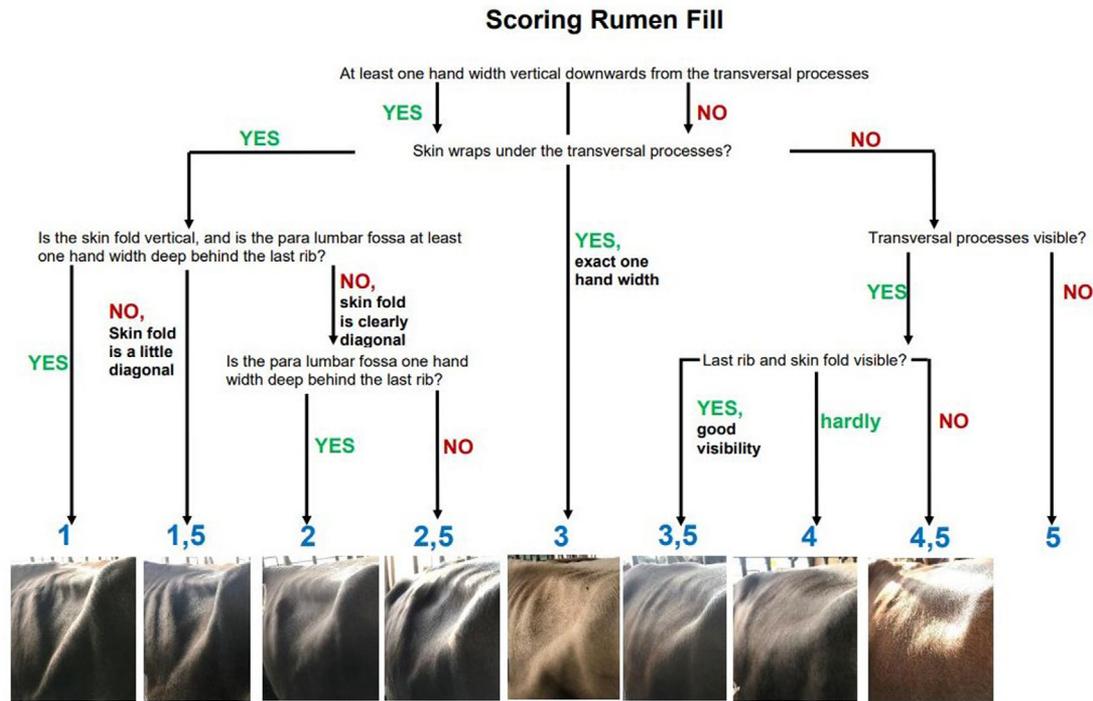


Figure 2. First version of the novel flowchart for scoring rumen fill.

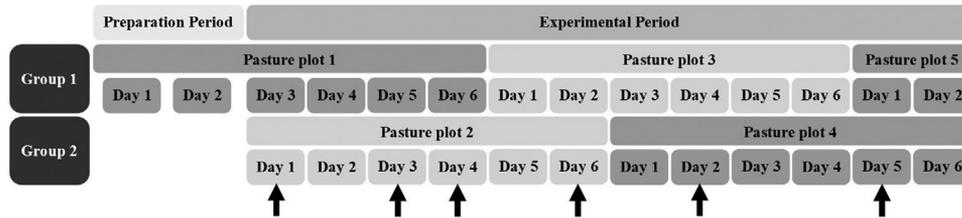


Figure 3. Experimental design of pasture allocation. The preparation period includes Day 1 and Day 2 of the first pasture plot of Group 1. Each group consisted of 10 cows. Additionally, days of rumen fill scoring are shown as black arrows.

lation, 0.51–0.70 = moderate correlation, 0.71–0.90 = high correlation, and 0.91–1.00 = very high correlation.

Additionally, κ was used to identify the degree of pairwise agreement of each pair of observers. It was weighted quadratically to penalize large deviations harder than small deviations.

Moreover, weighted Fleiss' kappa (κ_F) was used to determine the degree of the overall agreement among all observers of a group (e.g., trained observers). The weighted Fleiss' kappa was also weighted quadratically. The interpretation of Kappa coefficients was based on the criteria defined by Landis and Koch (1977): < 0.00 = poor agreement, 0.00–0.20 = slight agreement, 0.21–0.40 = fair agreement, 0.41–0.60 = moderate agreement, 0.61–0.80 = substantial agreement and 0.81–1.00 = almost perfect agreement.

To test if there was a significant difference among the observers, an analysis of variance (ANOVA) and the test of least significant difference (LSD) were conducted with the scores of the reduction rate of the observers. The significance level of the LSD test was 5% for the whole study.

Quality assessment of detecting feed intake changes

The size of change between the scores on consecutive allocation days is here referred to as reduction rate of scores. At first, the reduction rate of scores between consecutive days (e.g., 1 and 2, 2 and 3 ...) for each cow and each observer was calculated. After

that, the geometric mean of the reduction rate of all cows at each day and observer was calculated.

To assess the detection of changes in feed intake, the calculated reduction rates for each cow and each observer between each day were used and summarized using the geometric mean. This resulted in five geometric means for the reduction rates, one for every pair of consecutive days. These five calculated means of reduction were continuously subtracted from the baseline of 0%, which corresponds to the ad libitum feed intake of the cows.

The reduction rate, as described earlier, was also calculated for the daily milk yield. These reduction rates were continuously subtracted from the baseline of 0%, which corresponds with the milk yield of the cows under conditions of ad libitum feed intake. The daily milk yield of each cow was linked to her feed allocation.

The influence of BCS and body weight on rumen fill was considered in the analysis. However, due to the small sample size, the validity of the information was questioned and therefore not included in the study.

Results

Quality assessment of scoring

Interobserver reliability of trained observers The interobserver reliability was analyzed by using κ and r_s to explain agreement and

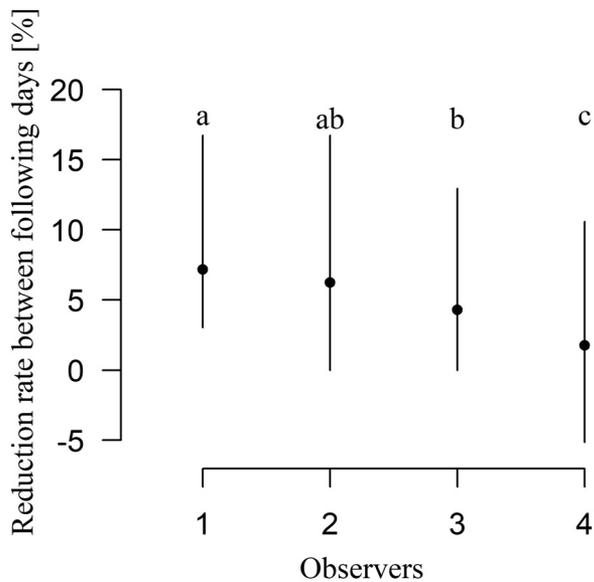


Figure 4. Results of the analysis of variance and the test of least significant difference are indicated by letters. The size of change between the scores on consecutive days with increasing allocation is here referred to as reduction rate of scores ($n = 120$ observations per observer). The lines represent the range of reduction rate.

correlation, respectively. Three of the observers (Observers 1–3) correlated highly, whereas Observer 4 had low to moderate correlations with every other observer. In conclusion, the pairwise agreement among the four observers ranged from fair ($\kappa = 0.30$) to almost perfect ($\kappa = 0.85$) and the correlation ranged from low ($\kappa = 0.49$) to high ($\kappa = 0.84$). Additionally, the overall agreement was moderate ($\kappa_F = 0.53$).

These findings were supported by the ANOVA and LSD analysis shown in Fig. 4. No significant difference was found between the scoring of Observers 1 and 2 and Observers 2 and 3. The scoring of Observer 4 was significantly different from all other observers with a lower mean reduction rate.

Following discrepancies among the observers, a flowchart was developed (see Fig. 2) and modified after a feedback round. This version additionally includes a higher number of photographs. The complete flowchart is displayed in Supplementary 1.

Quality of the control group

Using the first version of the flowchart, all observers of the control group demonstrated a substantial agreement ($\kappa = 0.71$ to $\kappa = 0.82$) and the correlation between each pair of the three observers was high ($r_s = 0.76$ to $r_s = 0.88$). The overall agreement of the observers was $\kappa_F = 0.74$.

The range of pairwise agreement and correlation among the three trained observers of the control group, using the second version, is similar but slightly lower. The pairwise agreement ranged from moderate to substantial ($\kappa = 0.52$ to $\kappa = 0.79$). In addition, the scores of the three trained observers were moderately to highly correlated ($r_s = 0.65$ to $r_s = 0.79$). The overall agreement was $\kappa_F = 0.71$.

The results of the two conducted ANOVAs with LSD can be found in Fig. 5 and support these findings. The ANOVA of the first scoring shows that Observers 1 and 3 had the same range (2.5 to 4.5) and Observer 2 showed a slightly higher scattering of scores (1.5–4.5). Moreover, the ANOVA of the second scoring demonstrated that Observers 2 and 3 had the same range of assigned scores (2–4.5), whereas Observer 1 scored slightly different. However, the observers did not score significantly different from each other.

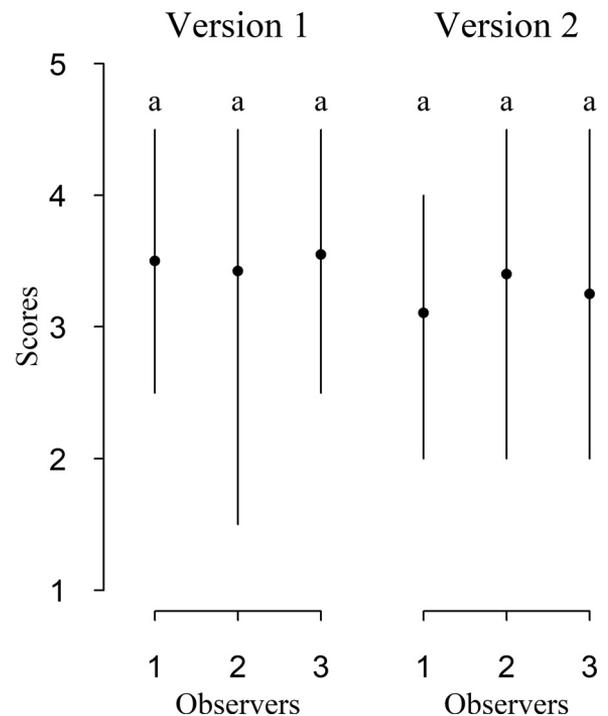


Figure 5. Rumen fill scores of the observers of the control group who used the flowchart versions 1 and 2 ($n = 20$ observations per observer in each version). Results of the analysis of variance and the test of least significant difference are indicated by letters. Lines represent the range of scores.

Interobserver reliability of untrained observers

The pairwise agreement between the untrained observers, using Version 1, ranged from slight ($\kappa = 0.10$) to substantial ($\kappa = 0.71$). Moreover, the correlation coefficients between the untrained observer pairs were varying, ranging from a negligible ($r_s = 0.21$) to a high correlation ($r_s = 0.83$). There was one untrained observer (Observer 5) whose scores differed substantially from the others due to less variation in his scoring. The overall agreement among the four observers can be interpreted as fair ($\kappa_F = 0.40$). These results are supported by the ANOVA and LSD (Fig. 6).

The pairwise agreement among the observers, using Version 2, ranged from moderate ($\kappa = 0.50$) to substantial ($\kappa = 0.73$) and the correlation from low ($r_s = 0.50$) to high ($r_s = 0.72$). The overall agreement of the observers was substantial ($\kappa_F = 0.66$) and thus higher than in the first version of the flowchart. Again, the results of the ANOVA and LSD shown in Fig. 6 supported the previously mentioned findings. There was no significant difference among the four observers who used the second version. Additionally, the variation in the given scores was remarkably similar for all observers. There was no single observer whose variation of scores differed as much as Observer 5, using the first version (see Fig. 6).

Quality of the novel flowchart

The mean of κ between the control group and the untrained observers who used the first version was $\kappa = 0.64$. This indicated a substantial agreement between the untrained observers and control group. The mean of κ between the control group and untrained observers, using the second version, was $\kappa = 0.72$. Therefore, the agreement between untrained observers and the control group was also substantial. This indicates that there was a positive development, and Version 2 seems to support untrained observers better in their decision-making process.

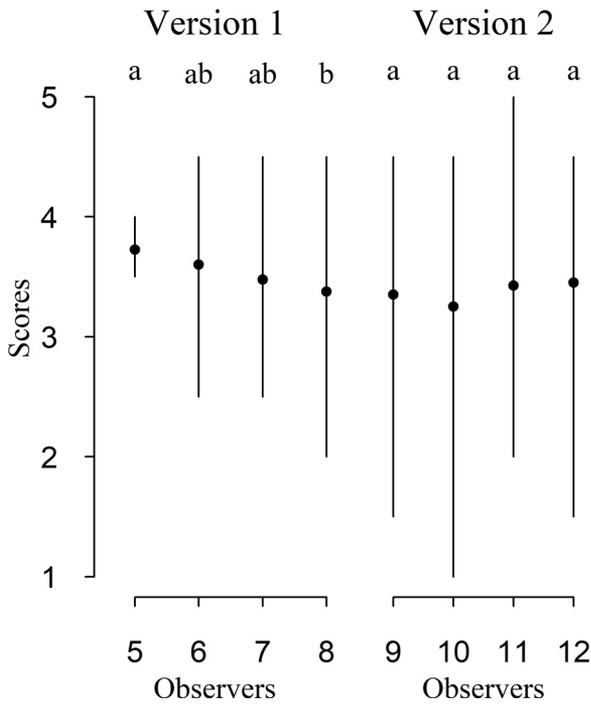


Figure 6. Rumen fill scores of untrained observers who used the flowchart versions 1 and 2 ($n=20$ observations per observer in each version). Results of the analysis of variance and the test of least significant difference are indicated by letters. The lines represent the range of scores.

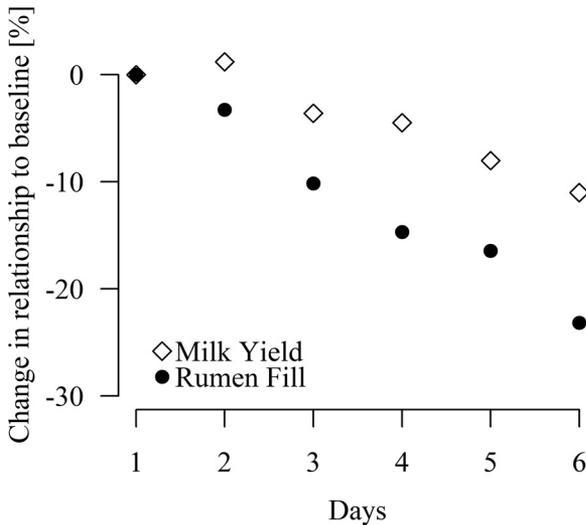


Figure 7. Mean percentage of changes in feed intake calculated on the basis of the reduction rate of rumen fill scores of the trained Observers 1–4 over a 6-d period with increasing level of feed supply ($n=80$ observations; 20 observations per observer on each day). The experiment set out to reduce grazing allocation continuously from the baseline on Day 1 to -20% on Day 6. Additionally, the average milk yield across all cows and the two allocation periods for each of the 6 d is shown ($n=40$ measurements; 20 measurements per day in both periods).

Quality assessment of detecting feed intake changes

The potential to use the rumen fill scores to detect changes in feed intake on pasture was assessed by looking at the scores for cows with six different grazing allocation levels. The changes to the baseline, identified via scoring of the trained Observers 1 – 4, is shown in Fig. 7. It reveals a continuous daily reduction of 24% over the 6-d period, from the baseline on Day 1 to Day 6. This finding is in accordance with the experimental aim of the larger study

to reduce grazing allocation from sufficient (baseline) to scarce, defined as 20% reduction, over the 6-d period. The reduction in milk yield of the cows added up to 11% at Day 6 compared with the baseline.

Discussion

Quality assessment of scoring

One of the aims of the study was to develop a flowchart to reduce the training time for observers when they start using the rumen fill scoring system and to support the observers in their decision-making process during scoring. An indication was indeed found that the results of the interobserver reliability improved between using the first and second versions of the flowchart, as the pairwise agreement range changed from “slight to substantial” to “moderate to substantial.” However, the sample size is small and the results need to be considered carefully. Yet it should be pointed out that on the last experimental date, no single untrained observer varied widely from the others. This result is also supported by the fact that the overall agreement among the untrained observers was higher than the overall agreement of the four trained observers while they were not using the flowchart. Although most of the four untrained observers who used the second version of the flowchart had little to no experience with cows, the overall agreement was just slightly lower than the agreement of trained observers measured by Burfeind et al. (2010). Therefore, an indication was found—that the developed novel flowchart, based on the criteria of Zaaijer and Noordhuizen (2003), simplifies rumen fill scoring and reduces subjectivity and training time.

Quality assessment of detecting feed intake changes

The second objective of this study was to assess if rumen fill scoring is a suitable method for detecting changes in feed intake on pasture. To do so, the reduction rate of the average of the entire group of cows was calculated and subtracted daily from the baseline. In accordance with the feed allocation treatment, a reduction in grazing allocation from the baseline on Day 1 down to -20% on Day 6 was targeted. However, the practical implementation of the targeted reduction of grass availability could have been affected by external and internal factors, such as unexpected regrowth patterns of grass and varying feed intake of cows due to weather impact. Nevertheless, the targeted aim coincides with the finding that rumen fill scores reduced from the baseline to Day 6 by 24%. Moreover, the reduction of feed intake is supported by the reduction in milk yield. The size of the effect is in accordance with the findings of Herve et al. (2019) and Vanbergue et al. (2018), which were summarized by Leduc et al. (2021), and with the findings of Hart et al. (2022). In their studies, milk yield was reduced by about 9% (Herve et al. 2019), 10.4% (Hart et al. 2022), and 12% (Vanbergue et al. 2018), when feed supply was restricted to 80% of the cow’s ad libitum DMI (Vanbergue et al. 2018; Herve et al. 2019; Hart et al. 2022). This high agreement among the targeted feed intake, reduction in milk yield, and reduction of rumen fill scores indicates that rumen fill scoring could be a suitable method for detecting changes in DMI during grazing.

Therefore, we conclude that rumen fill scores can support grazing studies by complementing the RPM method because they have the advantage of providing information on an individual animal level. In comparison with other intake estimation models, rumen fill scoring is less labor intensive, cheap, and user-friendly. Rumen fill scoring can therefore provide rangers and researchers with valuable management information.

Nevertheless, rumen fill scoring has some obvious limits. Although the novel flowchart reduces subjectivity, rumen fill scoring

stays a subjective observation. Additionally, there are some aspects to consider, such as the time of scoring, influence of rumen activity, and fixation of the cows during scoring. Moreover, it is important to focus on the group and not on individual animals because decreasing feed intake and thus a reduction in rumen fill score of a single cow could also be caused by other reasons, such as various diseases (Bareille et al. 2003; Norring et al. 2014). Moreover, the amount of change of rumen fill is individually different among cows. We hypothesize that this could be caused by rank, individual feed intake, or individual body type. Another point is that it is important to focus on relative changes and not absolute scores. Therefore, rumen fill scores do not indicate absolute DMI of the cows. The result is a relative change and needs a baseline.

However, to implement rumen fill scoring as a commonly used method for evaluating changes in feed intake on pasture, it needs to be investigated further. Important topics for future studies are the possibility of technical scoring of rumen fill, through computed image analysis; the impact of knowledge of the observers about feed allocation on their scoring; the impact of different sward types, including the impact of rumen gas; the influence of different passage rates; and the effect of different breeds and body types of cattle on the scoring results.

Implications

The novel chart simplified rumen fill scoring in a way that the overall agreement of untrained observers was nearly similar to the overall agreement of the trained observers. Therefore, the method proved to be practical and easy to use for grazing allocation studies with Brown Swiss cattle. This study demonstrated that rumen fill scoring might be useful to detect short-term changes of grazing intake. However, to implement rumen fill scoring as a commonly used method for evaluating changes in feed intake on pasture, it needs to be investigated further.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank all observers for making this project possible. In addition, thank you to the farm staff for helping to conduct this study. Moreover, thanks to Dr. Abu Zar Shafullah for statistical support.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.rama.2022.02.007](https://doi.org/10.1016/j.rama.2022.02.007).

References

Bareille, N., Beaudeau, F., Billon, S., Robert, A., Faverdin, P., 2003. Effects of health disorders on feed intake and milk production in dairy cows. *Livestock Production Science* 83, 53–62.

- Brenninkmeyer, C., Dippel, S., March, S., Brinkmann, J., Winckler, C., Knierim, U., 2007. Reliability of a subjective lameness scoring system for dairy cows. *Animal Welfare* 16, 127–129.
- Burfeind, O., Sepúlveda, P., Keyserlingk, M.A.G.von, Weary, D.M., Veira, D.M., Heuwieser, W., 2010. Technical note: evaluation of a scoring system for rumen fill in dairy cows. *Journal of Dairy Science* 93, 3635–3640.
- D'Eath, R.B., 2012. Repeated locomotion scoring of a sow herd to measure lameness: consistency over time, the effect of sow characteristics and interobserver reliability. *Animal Welfare* 21, 219–231.
- Delagarde, R., Faverdin, P., Baratte, C., Peyraud, J.L., 2011. Grazeln: a model of herbage intake and milk production for grazing dairy cows. 2. Prediction of intake under rotational and continuously stocked grazing management. *Grass and Forage Science* 66, 45–60.
- Die Borchert Kommission, 2020. Recommendations of the competence network livestock husbandry. The Borchert Commission.
- Dillon, P., Hennessy, T., Shalloo, L., Thorne, F., Horan, B., 2008. Future outlook for the Irish dairy industry: a study of international competitiveness, influence of international trade reform and requirement for change. *International Journal of Dairy Technology* 61, 16–29.
- Dillon, P., Stakelum, G., 1988. The use of n-alkanes and chromic oxide as markers for determining feed intake, faecal output and digestibility in dairy cows. In: *Proceedings of the 12th General Meeting of the European Grassland Federation*, Dublin, pp. 154–158.
- Garnick, S., Barboza, P.S., Walker, J.W., 2018. Assessment of animal-based methods used for estimating and monitoring rangeland herbivore diet composition. *Rangeland Ecology & Management* 71, 449–457.
- Götze, K., Crivellaro, P., Pieper, L., Engelhard, T., Staufienbiel, R., 2019. Evaluation of the rumen filling of dairy cows to assess the individual feed intake in the herd management. *Veterinary Practice, Issue G, Large Animals/Farm Animals* 47, 5–13.
- Hart, L., Dickhoefer, U., Paulenz, E., Umstaetter, C., 2022. Evaluation of a binary classification approach to detect herbage scarcity based on behavioral responses of grazing dairy cows. *Sensors* 22, 968.
- Herve, L., Quesnel, H., Veron, M., Portanguen, J., Gross, J.J., Bruckmaier, R.M., Boutinaud, M., 2019. Milk yield loss in response to feed restriction is associated with mammary epithelial cell exfoliation in dairy cows. *Journal of Dairy Science* 102, 2670–2685.
- Hinkle, D.E., Wiersma, W., Jurs, S.G., 2003. *Applied statistics for the behavioral sciences*, 5th ed.. Houghton Mifflin, Boston, MA, USA, p. 756.
- Hofstetter, P., Frey, H.-J., Gazzarin, C., Wyss, U., Kunz, P., 2014. Dairy farming: indoor v. pasture-based feeding. *Journal of Agricultural Science* 152, 994–1011.
- Hulsen, J., 2010. Beurteilungskarte pansenfüllung. In: Hulsen, J. (Ed.), *Kuhsignale Kontrollbuch*. Roodbont, p. 18.
- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.
- Leduc, A., Souchet, S., Gelé, M., Le Provost, F., Boutinaud, M., 2021. Effect of feed restriction on dairy cow milk production: a review. *Journal of Animal Science* 99.
- Mayer, R.W., Lamb, C.S., Colgrove, P.M., 1986. The use of dosed and herbage n-alkanes as markers for the determination of herbage intake. *Journal of Agricultural Science* 107, 161–170.
- Mosimann, E., Stettler, M., 2004. Pasture of dairy cows. Calculation of the adjusted stocking strength. *Information AGFF W10*.
- Norring, M., Häggman, J., Simojoki, H., Tamminen, P., Winckler, C., Pastell, M., 2014. Short communication: lameness impairs feeding behavior of dairy cows. *Journal of Dairy Science* 97, 4317–4321.
- Rombach, M., Südekum, K.-H., Münger, A., Schori, F., 2019. Herbage dry matter intake estimation of grazing dairy cows based on animal, behavioral, environmental, and feed variables. *Journal of Dairy Science* 102, 2985–2999.
- Shalloo, L., Dillon, P., O'Loughlin, J., Rath, M., Wallace, M., 2004. Comparison of a pasture-based system of milk production on a high rainfall, heavy-clay soil with that on a lower rainfall, free-draining soil. *Grass and Forage Science* 59, 157–168.
- Vanbergue, E., Peyraud, J.L., Ferlay, A., Miranda, G., Martin, P., Hurtaud, C., 2018. Effects of feeding level, type of forage and milking time on milk lipolytic system in dairy cows. *Livestock Science* 217, 116–126.
- Zaaijer, D., Noordhuizen, J.P.T.M., 2003. A novel scoring system for monitoring the relationship between nutritional efficiency and fertility in dairy cows. *Irish Veterinary Journal* 145–151.