Journal of the American Association for Laboratory Animal Science
Copyright 2023
by the American Association for Laboratory Animal Science

Vol 62, No 1
January 2023
Pages 81–86

Effects of Human Observer Presence on Pain Assessment Using Facial Expressions in Rabbits

Renata H Pinho,[1,*] André A Justo,[2] Daniela S Cima,[3] Mariana W Fonseca,[3] Bruno W Minto,[4] Fabiana D L Rocha,[4] Matthew C Leach,[5] Stelio P L Luna[6]

The goal of this study was to evaluate the effect of a human observer on Rabbit Grimace Scale (RbtGS) scores. The study scored video footage taken of 28 rabbits before and after orthopedic surgery, as follows: 24 h before surgery (*baseline*), 1 h after surgery (*pain*), 3 h after analgesia administration (*analgesia*), and 24 h after surgery (*24h*) in the presence and absence of an observer. Videos were assessed twice in random order by 3 evaluators who were blind to the collection time and the presence or absence of an observer. Responses to pain and analgesia were evaluated by comparing the 4 time points using the Friedman test, followed by the Dunn test. The influence of the presence or absence of the observer at each time point was evaluated using the Wilcoxon test. Intra- and interrater reliabilities were estimated using the intraclass correlation coefficient. The scale was responsive to pain, as the scores increased after surgery and had decreased by 24 h after surgery. The presence of the observer reduced significantly the RbtGS scores (median and range) at *pain* (present, 0.75, 0 to 1.75; absent, 1, 0 to 2) and increased the scores at *baseline* (present, 0.2, 0 to 2; absent, 0, 0 to 2) and *24h* after surgery (present, 0.33, 0 to 1.75; absent, 0.2, 0 to 1.5). The intrarater reliability was good (0.69) to very good (0.82) and interrater reliability was moderate (0.49) to good (0.67). Thus, the RbtGS appeared to detect pain when scored from video footage of rabbits before and after orthopedic surgery. In the presence of the observer, the pain scores were underestimated at the time considered to be associated with the greatest pain and overestimated at the times of little or no pain.

Abbreviations: FAU (facial action unit); RbtGS: Rabbit Grimace Scale

DOI: 10.30802/AALAS-JAALAS-22-000056

Introduction

Humans and other animals express emotions such as fear, joy, sadness, and pain through their faces.[3] Facial expressions can help identify pain in nonverbal patients, such as neonates and people with dementia.[13,28] Subjective scoring systems have been developed to assess pain based on facial expressions for several species,[2,9-11,16,32] including animals used in research.[18,19,34] The *Rabbit Grimace Scale* (RbtGS)[18] comprises 5 facial action units (FAU); orbital tightening, nostril shape, cheek flattening, whisker position, and ear position, with each action scored as 0 (absent), 1 (moderately present), or 2 (obviously present). This method has been used to evaluate pain in both research[4,17,30] and pet[1] rabbits. In a recent review of facial scales, the RbtGS was considered as having a moderate level of evidence for pain assessment due to the limited number of studies that have evaluated the validity and reliability of the scale.[8]

A potential limitation of facial expression analysis is that an animal may not be positioned such that manipulation of the animal may be necessary to obtain an appropriate image. The need to manipulate the animal is not ideal, as it impaired the assessment of facial expressions in lambs.[14] In addition, the presence of male human observers suppressed facial expressions of pain in rodents,[33] and the presence of a female observer inhibited the duration and frequency of pain behaviors in rabbits.[29] Therefore, these findings strongly suggest that pain assessment should be performed without human interference at least in these species. The presence of humans could lead to a false negative (a conclusion that the animal is either free of pain or has less pain than it is actually experiencing). Conversely, an observer can also inhibit the expression of normal behavior in pain-free rabbits, for instance by causing a reduction in activity levels and the duration of exploration behavior. This could lead to a false positive conclusion (that is, the mistaken belief that the animal is suffering pain).[29] Therefore, the influence of an observer on the expression in the RbtGS requires further investigation.

To address this issue, the current study used video footage to evaluate how the presence of a human observer affected RbtGS scores. Our hypothesis was that the presence of an observer alters the facial expression of rabbits that are both free of and experiencing pain.

Materials and Methods

The study was conducted in accordance with the Brazilian legislation of the National Council for the Control of Animal Experimentation and was approved by the Ethics Committee in the Use of Animals under protocol number 0156/2018 of the School of Veterinary Medicine and Animal Science of São Paulo State University, Botucatu Campus.

Submitted: 4 Jun 2022. Revision requested: 9 Jul 2022. Accepted: 17 Oct 2022
[1]Department of Surgical Specialties and Anesthesiology, Botucatu Medical School, São Paulo State University, Botucatu, São Paulo, Brazil; Faculty of Veterinary Medicine, University of Calgary, Calgary, AB, Canada, [2]Department of Surgery, School of Veterinary Medicine and Animal Science, University of São Paulo, São Paulo, SP, Brazil, [3]Department of Surgical Specialties and Anesthesiology, Botucatu Medical School, São Paulo State University, Botucatu, São Paulo, Brazil, [4]Department of Veterinary Clinics and Surgery, School of Agricultural and Veterinary Sciences, São Paulo State University, Jaboticabal, São Paulo, Brazil, [5]School of Natural and Environmental Science, Newcastle University, Newcastle upon Tyne, United Kingdom, [6]Department of Veterinary Surgery and Animal Reproduction, School of Veterinary Medicine and Animal Science, São Paulo State University, Botucatu, São Paulo, Brazil
**Corresponding author. Email: renata.haddad@unesp.br*

Animals, surgery, and data collection. This study used video recordings of rabbits that were undergoing partial ostectomy of the radius as part of another study.[29] Twenty-eight (11 females and 17 males) intact New Zealand white rabbits (*Oryctolagus cuniculus*) from the central vivarium of the São Paulo State University in Botucatu were used. Rabbits were 159 ± 5 d old and weighed 3.7 ± 0.4 kg, The rabbits were considered healthy based on a hemogram and physical exams, including inspection, cardiac and respiratory auscultation, and rectal temperature measurement.

Rabbits were housed in individual cages (60 × 60 × 60 cm) in a shelter adapted to house experimental animals. The shelter had natural ventilation controlled by curtains, with a local natural photoperiod of approximately 12 h of light per day and a mean temperature of 21 °C (70 °F). Animals were fed a rabbit commercial dry feed (Fri-Coelho, Trouw Nutrition, São Paulo, SP, Brazil), and water was supplied in open dishes, both ad libitum. Pieces of carrot were offered daily, and Tifton hay (*Cynodon spp*.) was provided twice a week. Pinecones were provided to supply interaction and chewing needs.[29] The videos were recorded at the following time points: *baseline* (24 h before surgery); *pain* (1 h after recovery from anesthesia); *analgesia* (3 h after analgesia and 4 h after anesthetic recovery); and *24 h* (24 h after anesthetic recovery). A GoPro Hero 5 camera (GoPro, San Mateo, CA) was used for the recording. Before each assessment, a new piece of carrot and a new pinecone were offered. After 5 min of acclimation, the footage of the rabbits was recorded for 5 min in the absence of the observer. Subsequently, a female observer who was familiar with the rabbits (RHP) entered the room, and the rabbits were then recorded for another 5 min (that is, with the observer present).[29]

For surgery, rabbits were premedicated with 5mg/kg of pethidine intramuscularly to promote mild sedation and to reduce stress during induction. Anesthesia was induced using isoflurane administered in oxygen via a face mask and was maintained with isoflurane after tracheal intubation or, when intubation could not be achieved, with a face mask. This protocol was used rather than injectable anesthetics because it is relatively painless, anesthetic recovery is quick, and residual anesthetic effects that could affect pain assessment are minimized. Fentanyl (2 µg/kg) was administered intravenously immediately before surgery. One hour after anesthetic recovery, after filming the *pain* time point, morphine 2 mg/kg and meloxicam 1 mg/kg were administered intramuscularly.

Video editing. To reduce the time taken to evaluate the videos and thus avoid rater fatigue, the original 5-min recordings were reduced to 2- to 3-min video clips by using a Movavi editor (Movavi Software Inc., St. Louis, MO). The reduced videos included the proportional duration and frequency of the behaviors that occurred during the original 5 min (that is, if the rabbit lay down for 30 s [10%] in the 5-min (300 s) original video, the 2-min (120 s) reduced video clips showed the rabbit lying down for 12 s [10%]).

RbtGS assessment. This was an opportunistic study in which the RbtGS assessments were carried out at the same time as a recently published study assessing a behavioral scale in rabbits.[15] Both the current study and the published study that validated a behavioral scale[15] used videos collected in a previous study on rabbit behavior.[29] Three raters (one man and two women) who were unaware of the time points scored the RbtGS in each of the 224 videos (112 in the presence of the observer and 112 in the absence of the observer, using the same rabbits and time points in both evaluations).

All the raters were residency-trained veterinarians, with similar experience working with veterinary anesthesia (4 y) and no previous experience using the RbtGS. No special training was provided, but before starting assessments the raters received the images and descriptors of each Facial Action Unit (FAU) according to the original study in which the scale was developed[18] as a guide to scoring the RbtGS. One of the raters (RHP) was responsible for handling the animals during the recordings and editing the videos.

The videos were scored randomly by raters who were blind to the time points and the presence or absence of the observer. The evaluations were repeated one month later for calculation of intraobserver reliability between the 2 phases. The raters were instructed to assign a single score for each video by considering the highest score of each FAU throughout the entire video period and not to score the RbtGS when the rabbits were eating or grooming. The RbtGS was scored on a 3-point scale of intensity: absence of a FAU corresponded to a score of "0," a moderate intensity corresponded to a score of "1," and an obvious presence corresponded to a score of "2." "Missing data" were assigned when the rater was unable to identify a FAU due to the position of the rabbit, poor image quality, or inability to see the FAU. These cases were considered missing data in the analysis.

Statistical analysis. If a FAU was not observed in the presence of the observer, then that FAU from that time point of that rabbit was excluded from analysis in both the presence and absence of the observer and vice-versa. FAU identified by other raters or by the same rater during another evaluation phase were used for statistical analysis. The RbtGS score was calculated as the adjusted mean of each individual FAU. The scores of all FAU that were observed were summed, and the value was divided by the number of FAU assessed. For example, when all FAU were identified, the RbtGS score was calculated as follows: orbital tightening + nostril shape + cheek flattening + whisker position + ear position/5. If the observer was not able to identify the whisker position, the RbtGS score was calculated as follows: orbital tightening + nostril shape + cheek flattening + ear position/4. Thus, the adjusted mean score of the RbtGS ranged from 0 to 2.[27] Data are presented as medians and range.

The frequency of occurrence (percentage) of each possible score (0, 1, or 2) and missing data were described for each FAU at each time point and for all time points grouped (GM).

Responsiveness of RbtGS scoring and influence of the observer's presence. The responsiveness of the RbtGS was evaluated in terms of the increase or decrease in scores in response to a pain stimulus, the administration of analgesics, and after 24 h. We hypothesized that scores would increase after surgery as compared with baseline and decrease after the administration of analgesics as compared with *pain* (after surgery). Intermediate scores were expected at 24 h after surgery because at a longer time period after surgery, the rabbit's pain is less than that occurring immediately after surgery (*pain*).

Data were analyzed using Graph Pad Prism 5.0 software (GraphPad Software, San Diego, California, USA) and R software (RStudio, Rstudio Team, 2020). Data were nonparametric based on to the Shapiro–Wilk test. To assess responsiveness of the scale, the time points (*baseline, pain, analgesia, and 24h*) were compared using the Friedman test followed by the Dunn post-hoc test. To compare RbtGS scores in the presence and absence of the observer at each time point, the Wilcoxon test for paired samples was used. Differences were considered statistically significant at $P < 0.05$.

Intra- and interrater reliability. Intra- and interrater reliability were estimated using the intraclass correlation coefficient (ICC), with a confidence interval (CI) of 95%. Intrarater reliability was defined by comparing the data from each rater during Phase 1 and Phase 2. Interrater reliability was determined by comparing the 2 phases of evaluations among the 3 raters.

Results

Frequency of occurrence and missing data. A score of 0 was the most common (that is, over 55% occurrence) for all FAU at *baseline*. At the *pain* and *analgesia* time points, scores 1 or 2 were predominant for the FAU orbital tightening and ear position; score 0 (36%) and missing data (27%) prevailed for cheek flattening. Data for nostril shape and whisker position were missing in 63% and over 75% of the assessments, respectively. At *24h*, the score 0 was most common of the 5 FAU, except for ear position in the observer´s presence, which scored as 1 in 50% of the assessments (Figure 1). Of the 1,344 total evaluations [28 animals × 2 (presence and absence of the observer) × 4-time points × 3 evaluators × 2 evaluation phases], all 5 FAU could be identified in only 683 (50.8%).

Responsiveness, influence of the observer, and reliability. RbtGS scores increased after surgery ($P < 0.0001$) in both the absence and presence of the observer. Scores (median, range) were significantly higher at *pain* (presence: 0.75, 0 to 1.75; absence: 1, 0 to 2) as compared with *baseline* (presence: 0.2, 0 to 2; absence: 0, 0 to 2). Analgesic administration was not associated with significant differences between *pain* and *analgesia* time points. The *24h* scores were higher than at *baseline* only in the presence of the observer ($P < 0.001$) (Table 1).

The RbtGS scores were higher at *Baseline* ($P < 0.02$) and *24h* ($P < 0.0001$) time points and lower at *pain* ($P < 0.002$) when the observer was present compared with when she was absent. Therefore, based in our descriptive data, the presence of an observer at *baseline* was associated with a 20% overestimate of median scores, and with an underestimate of 25% at the *pain* time point (Table 1).

The RbtGS intrarater reliability varied from good (0.69) to very good (0.82). The interrater reliability varied from moderate (0.49) to good (0.67)[6] (Table 2).

Discussion

The RbtGS was able to detect pain and was influenced by the presence of an observer. This result is in line with a previous study[29] in which the observer´s presence reduced both pain behaviors in rabbits experiencing pain and the normal behavior in pain-free rabbits, which could lead to false-negative and false-positive results, respectively.

Both in the presence and absence of the observer, a score of 0 occurred in more than 55% of all FAU at *baseline*. Scores 1 or 2, corresponding to a moderate or obvious presence of the FAU of orbital tightening and ear position were observed at the time points likely to be associated with the most postoperative pain (*pain*). The other FAU were associated with missing data in most of the *pain* and *analgesia* time point assessments. The significant amount of missing data may have contributed to underestimation of the presence of scores 1 and 2 at the *pain* time point (in other words, if the missing data FAU had been detected, higher scores might have been observed at these time points). At *24h*, the score 0 predominated for all FAU except ear position, which showed a large number of scores of 1. This finding could be due to the venous access performed in the marginal ear vein during anesthesia. Therefore, when postoperative pain is assessed by the RbtGS, catheterization should be performed in veins other than the auricular.

Previous research has reported difficulty, similar to ours, in scoring RbtGS FAU when using still images[17,18] or videos.[24] As in our study, the whisker position was excluded from the analysis during development of the RbtGS because this FAU could not be not identified in most images.[18]A possible explanation for the considerable missing data observed in the present study at *pain* and *analgesia* time points as compared with *baseline* and *24h* time points, was the position of the rabbits during the recordings, in which their faces were looking toward the back of the cage (that is, away from the camera). Further, in the previous study of RbtGS,[17] rabbits were undergoing maxillary surgery. The nostril shape and cheek flattening FAU were not detected at any time point and assessment of the whisker position was difficult; therefore these items were excluded from the statistical analysis. Nevertheless, when scoring was performed in person, this difficulty was not reported,[1] suggesting that real-time observation of the RbtGS may reduce the amount of missing data. However, automated home-cage monitoring systems are still a suitable alternative for pain assessment. Their use reduces the effects of human presence and minimizes observer effects.[12]

Studies in other species have also reported difficulty in assessing particular FAU. In mice,[20,23] the whisker item was excluded because it was not easy to identify, and it was not statistically representative in ferrets.[32] In piglets,[10] lip contraction, nostril dilation, and mandible profile were excluded because they were difficult to identify.

The RbtGS was responsive to pain, as indicated by higher scores at the time point most likely to be associated with pain and lower scores at 24 h after surgery. However, regardless of the presence or absence of the observer, the RbtGS did not detect effect of analgesia. Pain is expected to be relieved by analgesia, shown by a reduction in pain scores, as reported for the Rabbit Pain Behavior Scale (RPBS).[15] This failure to detect expected analgesia could be due to poor efficacy of the analgesics used, the influence of opioids on facial expressions, and/or lower sensitivity of the RbtGS to identify the effect of analgesia as compared with a behavior-based analysis.[15] As for the influence of the opioids, the Sedation Scale for Use in Rabbits[31] includes orbital tightening, which may support an influence of opioids on this FAU. In horses, eye-related FAU can be distinguished between those caused by sedation (that is, orbital closure) and those related to facial muscle contraction caused by pain (that is, orbital tightening).[2,26] However, this differentiation does not appear to be feasible in rabbits, because of their smaller size and more dense coat; the FAU orbital tightening is described identically on the sedation scale[31] and in the RbtGS.[18] In cats, the orbital tightening was more scored after administration of buprenorphine and acepromazine when compared with baseline.[7] The residual effect of inhalant anesthetic in mice[23] and rats[25] is another example of the effect of drugs influencing facial pain assessment.

As for the behavioral pain assessment in rabbits undergoing orthopedic surgery,[29] the presence of a female human observer, even if familiar to the rabbits, reduced the pain scores after surgery and overestimated pain in rabbits expected to be devoid of pain (*baseline*) or in reduced pain (*24h*). These results suggest the possibility that both behavioral and facial alterations represent emotions other than pain, such as fear,[5] leading to false-positive results (that is, identifying pain when it is not present). During pain evaluation, other emotions in animals (for example, stress and anxiety) should be avoided because they can be similar

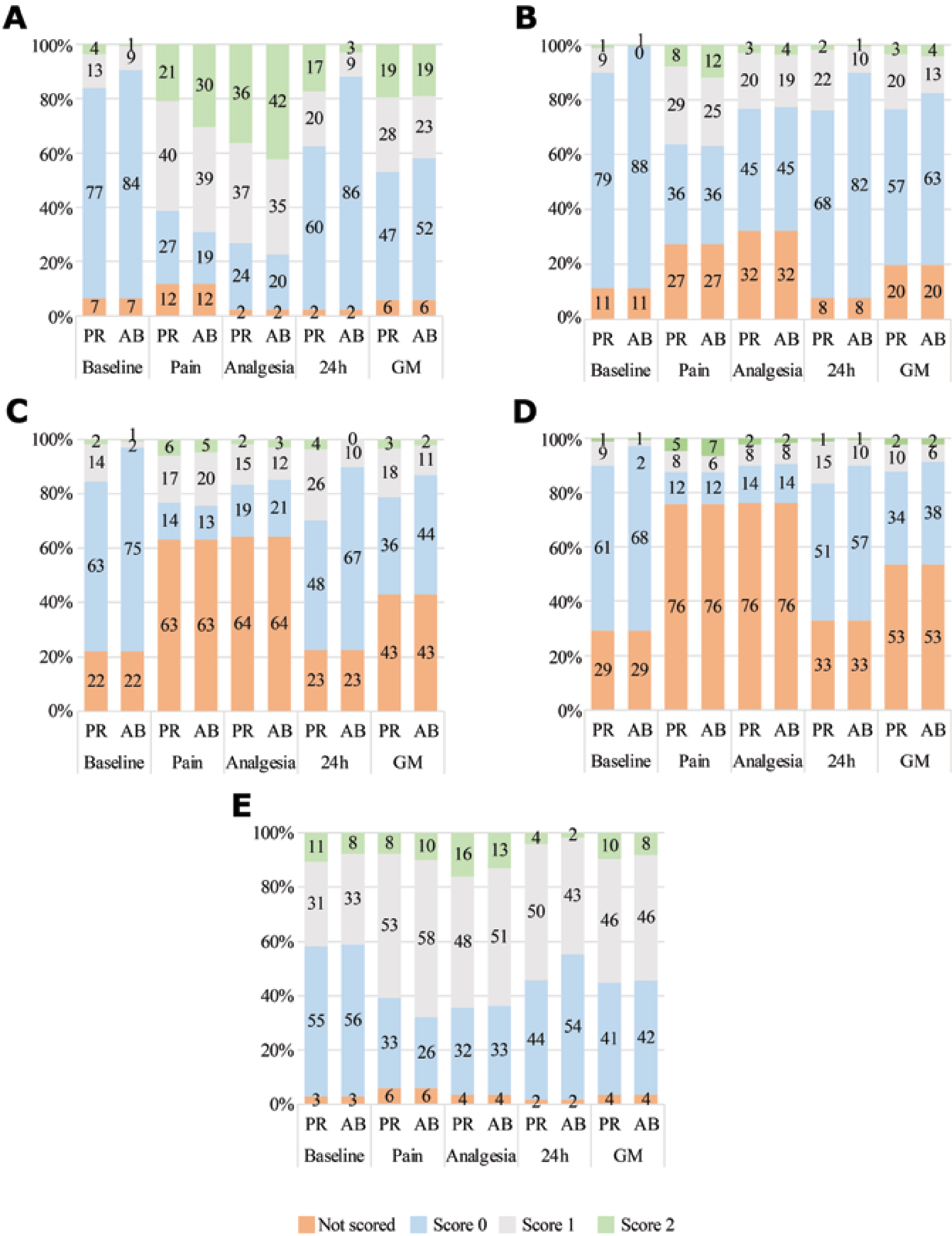

Figure 1. Percentage of frequency of occurrence of the scores of each FAU and missing data from the RbtGS. A) orbital tightening; B) cheek flattening; C) nose shape; D) whisker position; E) ear position. *Baseline*: 24 h before surgery; *Pain*: 1 h after anesthetic recovery; *Analgesia*: 3 h after rescue analgesia and 4 h after anesthetic recovery; *24h*: 24 h after anesthetic recovery; GM: grouped moments (*baseline* + *pain* + *analgesia* +*24h*); PR: presence of the observer; AB: Absence of the observer.

Table 1. Median (range) of RbtGS Scores in the presence and absence of the observer

Time point	Baseline	Pain	Analgesia	24h
Observer present	0.2 (0–2)[c*]	0.75 (0–1.75)[a]	1 (0–2)[a]	0.33 (0–1.75)[b*]
Observer absent	0 (0–2)[b]	1 (0–2)[a*]	1 (0–2)[a]	0.2 (0–1.5)[b]

Small letters indicate significant differences between time points (a > b > c) and * indicates significant differences between the presence and absence of the observer.

to pain. However, despite the differences seen in the presence and absence of the observer, the RbtGS did recognize pain after surgery in the present study, both in the presence and absence of the observer.

Although rabbits in our study had no signs of pododermatitis, rabbits housed in cages can develop this problem, which may cause pain. The low baseline pain scores probably suggest that the rabbits in the sample were not suffering from pododermatitis. Factors other than surgery that may cause discomfort should be avoided to preserve wellbeing and avoid interference in pain assessment.

The intrarater reliability of the RbtGS ranged from good to very good. However, the interrater reliability was moderate to good and lower than in previous studies of the RbtGS that either evaluated using photographs[18] or in person.[1] Providing training for evaluators who were not experienced in using the RbtGS could have increased the reliability in our study, as demonstrated in the evaluation of facial expressions of pain in rats.[35]

This current study had the following limitations. First, the amount of missing data may have been affected by the video assessment, which may be harder to evaluate as compared with static images. Also, insufficient quality of image capture from videos directly affects the assessment of facial scales,[2,18,22] with higher accuracy being associated with higher image quality.[19] The standard method for developing and validating facial scales is through photographs or screenshots, which provide high-quality and static images that facilitate the identification of FAU.[9,10,18,19,21,32,34] However, images may not be consistent with real-time assessment because 1) facial expressions may be difficult to perceive in a moving animal; 2) expression may vary within a period of time; and 3) images can capture a transitory expression that may not be visible in real-time evaluation. Reducing the original videos to shorter videos may have also been a limitation, as the FAU observed in an abbreviated time frame may not reflect what would be observed in the original video. Finally, the presence of cage grids in the videos overlaid the face and may also have hindered the assessment of FAU. Therefore, the use of transparent caging would be preferable. Given the difficulty in seeing the rabbit's face depending on its position in the cage, the ideal housing would provide a 360° view. Pen housing might improve the visualization of FAU, allow group housing, and improve wellbeing, and should therefore be considered in future studies.

Table 2. RbtGS intra- and interrater reliability

Reliability	ICC (CI)
Intrarater	
Evaluator 1	0.76 (0.71–0.82)
Evaluator 2	0.77 (0.71–0.82)
Evaluator 3	0.75 (0.69–0.80)
Interrater	
Evaluator 1 × 2	0.61 (0.55–0.66)
Evaluator 1 × 3	0.55 (0.49–0.61)
Evaluator 2 × 3	0.61 (0.55–0.67)

ICC: intraclass correlation coefficient; CI: confidence interval of 95%, Interpretation: very good 0.81–1.0; good: 0.61–0.80; moderate: 0.41–0.60; reasonable: 0.21–0.4; poor < 0.2.[6]

Conclusion

The RbtGS detected postoperative pain in rabbits undergoing orthopedic surgery. The presence of a human observer who was familiar with the rabbits seemed to reduce scores at the time of greatest pain and increase scores at times of little or no pain, suggesting that remote assessments are preferable if possible. The significant amount of missing data was an important limitation of the methodology used in this study; therefore, high-quality images are recommended when using this tool to evaluate pain in rabbits. New studies should be carried out to determine whether FAU that are difficult to assess (nostril shape and whisker position) are really necessary or could be excluded without compromising the validity of the RbtGS. Home cage monitoring, especially automated systems, can also be used to assess pain in rabbits because they do not require the presence of an observer and allow monitoring of animals in both active and inactive phases.

Acknowledgments

This study was funded by the thematic project (2017/12815-0) and grant (2018/17839-7), São Paulo Research Foundation (FAPESP), and by the Coordination for the Improvement of Higher Education Personnel (CAPES) - grant.

References

1. **Banchi P, Quaranta G, Ricci A, von Degerfeld MM, Von Degerfeld MM**. 2020. Reliability and construct validity of a composite pain scale for rabbit (CANCRS) in a clinical environment. PLoS One **15:**e0221377. https://doi.org/10.1371/journal.pone.0221377.
2. **Dalla Costa E, Minero M, Lebelt D, Stucke D, Canali E, Leach MC**. 2014. Development of the horse grimace scale (HGS) as a pain assessment tool in horses undergoing routine castration. PLoS ONE **9**:e92281. https://doi.org/10.1371/journal.pone.0092281
3. **Darwin C**. 1872. The expression of the emotions in man and animals. London: Albermale.
4. **DiVincenti L, Meirelles LAD, Westcott RA**. 2016. Safety and clinical effectiveness of a compounded sustained-release formulation of buprenorphine for postoperative analgesia in New Zealand white rabbits. J Am Vet Med Assoc **248:**795–801. https://doi.org/10.2460/javma.248.7.795.
5. **Dolensek N, Gehrlach DA, Klein AS, Gogolla N**. 2020. Facial expressions of emotion states and their neuronal correlates in mice. Science **368:**89–94. https://doi.org/10.1126/science.aaz9468.
6. **Altman DG.** 1991. Practical statistics for medical research. London: Chapman and Hall/CRC.
7. **Evangelista MC, Benito J, Monteiro BP, Watanabe R, Doodnaught GM, Pang DSJ, Steagall PV**. 2020. Clinical applicability of the Feline Grimace Scale: Real-time versus image scoring and the influence of sedation and surgery. PeerJ **8:**e8967. https://doi.org/10.7717/peerj.8967.
8. **Evangelista MC, Monteiro BP, Steagall PV**. 2022. Measurement properties of grimace scales for pain assessment in nonhuman mammals: a systematic review. Pain **163:**e697–e714. https://doi.org/10.1097/j.pain.0000000000002474.

9. **Evangelista MC, Watanabe R, Leung VSY, Monteiro BP, O'Toole E, Pang DSJ, Steagall PV**. 2019. Facial expressions of pain in cats: the development and validation of a Feline Grimace Scale. Sci Rep **9:**19128. https://doi.org/10.1038/s41598-019-55693-8.
10. **Di Giminiani P, Brierley VLMH, Scollo A, Gottardo F, Malcolm EM, Edwards SA, Leach MC**. 2016. The assessment of facial expressions in piglets undergoing tail docking and castration: Toward the development of the Piglet Grimace Scale. Front Vet Sci **3:**100. https://doi.org/10.3389/fvets.2016.00100.
11. **Gleerup KB, Forkman B, Lindegaard C, Andersen PH**. 2015. An equine pain face. Vet Anaesth Analg **42:**103–114. https://doi.org/10.1111/vaa.12212.
12. **Grieco F, Bernstein BJ, Biemans B, Bikovski L, Burnett CJ, Cushman JD, van Dam EA, Fry SA, Richmond-Hacham B, Homberg JR, Kas MJH, Kessels HW, Koopmans B, Krashes MJ, Krishnan V, Logan S, Loos M, McCann KE, Parduzi Q, Pick CG, Prevot TD, Riedel G, Robinson L, Sadighi M, Smit AB, Sonntag W, Roelofs RF, Tegelenbosch RAJ, Noldus LPJJ**. 2021. Measuring behavior in the home cage: Study design, applications, challenges, and perspectives. Front Behav Neurosci **15:**735387. https://doi.org/10.3389/fnbeh.2021.735387.
13. **Grunau RVE, Craig KD**. 1987. Pain expression in neonates: Facial action and cry. Pain **28:**395–410. https://doi.org/10.1016/0304-3959(87)90073-X.
14. **Guesgen MJ, Beausoleil NJ, Leach M, Minot EO, Stewart M, Stafford KJ**. 2016. Coding and quantification of a facial expression for pain in lambs. Behav Processes **132:**49–56. https://doi.org/10.1016/j.beproc.2016.09.010.
15. **Haddad Pinho R, Luna SPL, Esteves Trindade PH, Augusto Justo A, Santilli Cima D, Werneck Fonseca M, Watanabe Minto B, Del Lama Rocha F, Miller A, Flecknell P, Leach MC**. 2022. Validation of the rabbit pain behaviour scale (RPBS) to assess acute postoperative pain in rabbits (*Oryctolagus cuniculus*). PLoS ONE **17:**e0268973. https://doi.org/10.1371/journal.pone.0268973
16. **Häger C, Biernot S, Buettner M, Glage S, Keubler LM, Held N, Bleich EM, Otto K, Müller CW, Decker S, Talbot SR, Bleich A**. 2017. The Sheep Grimace Scale as an indicator of post-operative distress and pain in laboratory sheep. PLoS One **12:**e0175839. https://doi.org/10.1371/journal.pone.0175839.
17. **Hedenqvist P, Trbakovic A, Thor A, Ley C, Ekman S, Jensen-Waern M**. 2016. Carprofen neither reduces postoperative facial expression scores in rabbits treated with buprenorphine nor alters long term bone formation after maxillary sinus grafting. Res Vet Sci **107:**123–131. https://doi.org/10.1016/j.rvsc.2016.05.010.
18. **Keating SCJ, Thomas AA, Flecknell PA, Leach MC**. 2012. Evaluation of EMLA cream for preventing pain during tattooing of rabbits: Changes in physiological, behavioural and facial expression responses. PLoS One **7:**e44437. https://doi.org/10.1371/journal.pone.0044437.
19. **Langford DJ, Bailey AL, Chanda ML, Clarke SE, Drummond TE, Echols S, Glick S, Ingrao J, Klassen-Ross T, LaCroix-Fralish ML, Matsumiya L, Sorge RE, Sotocinal SG, Tabaka JM, Wong D, van den Maagdenberg AM, Ferrari MD, Craig KD, Mogil JS**. 2010. Coding of facial expressions of pain in the laboratory mouse. Nat Methods **7:**447–449. https://doi.org/10.1038/nmeth.1455.
20. **Leach MC, Klaus K, Miller AL, Scotto di Perrotolo M, Sotocinal SG, Flecknell PA**. 2012. The assessment of post-vasectomy pain in mice using behaviour and the Mouse Grimace Scale. PLoS ONE **7:**e35656. https://doi.org/10.1371/journal.pone.0035656.
21. **McLennan KM, Miller AL, Dalla Costa E, Stucke D, Corke MJ, Broom DM, Leach MC**. 2019. Conceptual and methodological issues relating to pain assessment in mammals: The development and utilisation of pain facial expression scales. Appl Anim Behav Sci **217:**1–15. https://doi.org/10.1016/j.applanim.2019.06.001.
22. **McLennan KM, Rebelo CJB, Corke MJ, Holmes MA, Leach MC, Constantino-Casas F**. 2016. Development of a facial expression scale using footrot and mastitis as models of pain in sheep. Appl Anim Behav Sci **176:**19–26. https://doi.org/10.1016/j.applanim.2016.01.007.
23. **Miller A, Kitson G, Skalkoyannis B, Leach M**. 2015. The effect of isoflurane anaesthesia and buprenorphine on the mouse grimace scale and behaviour in CBA and DBA/2 mice. Appl Anim Behav Sci **172:**58–62. https://doi.org/10.1016/j.applanim.2015.08.038.
24. **Miller AL, Clarkson JM, Quigley C, Neville V, Krall C, Geijer-Simpson A, Flecknell PA, Leach MC**. 2022. Evaluating pain and analgesia effectiveness following routine castration in rabbits using behavior and facial expressions. Front Vet Sci **9:**782486. https://doi.org/10.3389/fvets.2022.782486.
25. **Miller AL, Golledge HDR, Leach MC**. 2016. The influence of isoflurane anaesthesia on The Rat Grimace Scale. PLoS One **11:**e0166652. https://doi.org/10.1371/journal.pone.0166652.
26. **de Oliveira AR, Gozalo-Marcilla M, Ringer SK, Schauvliege S, Fonseca MW, Trindade PHE, Filho JNPP, Luna SPL**. 2021. Development and validation of the facial scale (FaceSed) to evaluate sedation in horses. PLoS One **16:**e0251909. https://doi.org/10.1371/journal.pone.0251909.
27. **Oliver V, De Rantere D, Ritchie R, Chisholm J, Hecker KG, Pang DSJ**. 2014. Psychometric assessment of the rat grimace scale and development of an analgesic intervention score. PLoS One **9:**e97882. https://doi.org/10.1371/journal.pone.0097882.
28. **Oosterman JM, Zwakhalen S, Sampson EL, Kunz M**. 2016. The use of facial expressions for pain assessment purposes in dementia: A narrative review. Neurodegener Dis Manag **6:**119–131. https://doi.org/10.2217/nmt-2015-0006.
29. **Pinho RH, Leach MC, Minto BW, Rocha FDL, Luna SPL**. 2020. Postoperative pain behaviours in rabbits following orthopaedic surgery and effect of observer presence. PLoS ONE **15:**e0240605. https://doi.org/10.1371/journal.pone.0240605
30. **Raillard M, Detotto C, Grepper S, Beslac O, Fujioka-Kobayashi M, Schaller B, Saulacic N**. 2019. Anaesthetic and perioperative management of 14 male New Zealand white rabbits for calvarial bone surgery. Animals (Basel) **9:**896. https://doi.org/10.3390/ani9110896.
31. **Raulic J, Leung VS, Doss GA, Graham JE, Keller KA, Mans C, Sadar MJ, Vergneau-Grosset C, Pang DS**. 2021. Development and Testing of a Sedation Scale for Use in Rabbits (Oryctolagus cuniculus). J Am Assoc Lab Anim Sci **60:**549–555. https://doi.org/10.30802/AALAS-JAALAS-21-000002.
32. **Reijgwart ML, Schoemaker NJ, Pascuzzo R, Leach MC, Stodel M, De Nies L, Hendriksen CFM, Van Der Meer M, Vinke CM, Van Zeeland YRA**. 2017. The composition and initial evaluation of a grimace scale in ferrets after surgical implantation of a telemetry probe. PLoS One **12:**e0187986. https://doi.org/10.1371/journal.pone.0187986.
33. **Sorge RE, Martin LJ, Isbester KA, Sotocinal SG, Rosen S, Tuttle AH, Wieskopf JS, Acland EL, Dokova A, Kadoura B, Leger P, Mapplebeck JC, McPhail M, Delaney A, Wigerblad G, Schumann AP, Quinn T, Frasnelli J, Svensson CI, Sternberg WF, Mogil JS**. 2014. Olfactory exposure to males, including men, causes stress and related analgesia in rodents. Nat Methods **11:**629–632. https://doi.org/10.1038/nmeth.2935.
34. **Sotocinal SG, Sorge RE, Zaloum A, Tuttle AH, Martin LJ, Wieskopf JS, Mapplebeck JCS, Wei P, Zhan S, Zhang S, McDougall JJ, King OD, Mogil JS**. 2011. The Rat Grimace Scale: A partially automated method for quantifying pain in the laboratory rat via facial expressions. Mol Pain **7:**55. https://doi.org/10.1186/1744-8069-7-55.
35. **Zhang EQ, Leung VS, Pang DS.** 2019. Influence of rater training on inter- and intrarater reliability when using the Rat Grimace Scale. J Am Assoc Lab Anim Sci **58**:178–183. https://doi.org/10.30802/AALAS-JAALAS-18-000044