

Access this article online
Quick Response Code:

Website: www.joas.in
DOI: 10.4103/joas.joas_8_18

# Rating visualization in shoulder arthroscopy: A comparison of the visual analog scale versus a novel shoulder arthroscopy grading scale

Vince W. Lands, Daniel M. Avery III, Ajith Malige<sup>1</sup>, Jill Stoltzfus<sup>1</sup>, Brett W. Gibson, Gregory F. Carolan

## Abstract:

**PURPOSE:** To assess the interobserver reliability and intraobserver variability of the visual analog scale (VAS) for visualization in shoulder arthroscopy and compare it to a less variable, more objective novel grading scale, the shoulder arthroscopy grading scale (SAGS).

**METHODS:** Twenty separate 30-s length video clips were created from a library of shoulder arthroscopies. Video clips were randomized and distributed to six sports medicine fellowship-trained surgeons at two time points with a 1-month interval. Each rated visualization according to an adapted VAS and a novel grading scale, the SAGS.

**RESULTS:** The VAS and SAGS both showed an excellent degree of consistency with interobserver reliability among raters with intraclass correlation coefficients (ICCs) of 0.96 and 0.97, respectively. Five of six raters demonstrated strong intraobserver variability with the VAS and SAGS with ICC ranging from 0.87 to 0.97 and 0.61 to 0.93, respectively.

**CONCLUSION:** Given the strong-to-excellent degree of consistency in using the VAS and the SAGS, either can be reliably used as a measurement of visualization in shoulder arthroscopy.

## Keywords:

Arthroscopy, shoulder, visual analog scale

## Introduction

A satisfactory visual field is essential for surgeons to perform arthroscopic surgery effectively. During arthroscopy, the quality of the visual field is negatively affected by blood mixing with the irrigation fluid. In knee arthroscopy, the visual field can be improved by controlling bleeding with the use of a tourniquet. Since this is not an option in shoulder arthroscopy, a variety of techniques to control bleeding have been utilized to improve visualization. These techniques include the use of flow-controlled or pressure-controlled pumps, the use of hypotensive anesthesia,

and the use of epinephrine in the irrigation fluid.<sup>[1-4]</sup>

The visual analog scale (VAS) is a subjective tool originally developed by clinicians to assess pain in patients.<sup>[5]</sup> Its use has grown recently and it has been used to describe multiple subjective measurements, such as chondropathy of the knee, fatigue, functional capacity, tension, and psychiatric state.<sup>[6,7]</sup> Two recent studies applied the VAS to describe the quality of visualization in shoulder arthroscopy. They both utilized the VAS rating to compare the hematocrit level in irrigation fluid with and without the use of epinephrine.<sup>[1,2]</sup> While the authors have been able to show utility in the use of the VAS in

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

**How to cite this article:** Lands VW, Avery DM, Malige A, Stoltzfus J, Gibson BW, Carolan GF. Rating visualization in shoulder arthroscopy: A comparison of the visual analog scale versus a novel shoulder arthroscopy grading scale. J Orthop Allied Sci 2019;7:8-11.

Departments of  
Orthopaedic Surgery and  
'Data Management and  
Outcomes Assessment,  
St. Luke's University  
Health Network,  
Bethlehem,  
Pennsylvania, USA

## Address for correspondence:

Dr. Vince W. Lands,  
St. Luke's Orthopaedic  
Specialists, St. Luke's  
University Hospital  
and Health Network,  
Bethlehem, PA 18015,  
USA.  
E-mail: vince.lands@  
sluhn.org

the setting of visualization in shoulder arthroscopy, it has not been formally validated. Due to its wide scale (0–10) and subjectivity, the VAS may not be an optimal scoring system for rating arthroscopic visualization and a novel rating system with a smaller range of scores and more objective criteria may be needed.

The purpose of this study was to assess the interobserver reliability and intraobserver variability of the VAS for visualization in shoulder arthroscopy and compare it to a less variable, more objective grading scale, the shoulder arthroscopy grading scale (SAGS) validating their use. Our hypothesis was that the less variable, more objective SAGS would show significantly improved interobserver reliability and intraobserver variability among raters.


## Methods

After obtaining approval from our Institutional Review Board, 20 separate 30-s length video clips were created from a larger library of shoulder arthroscopies. The library was derived from cumulative procedures performed by all sports medicine-trained physicians. All arthroscopic cases were performed using the same arthroscopic equipment (Arthrex AR-3200 Synergy HD3 Tower with dual wave pump and standard 30° scope). Thirty seconds was deemed appropriate to allow each rater adequate time to determine if full visualization of anatomic structures was achieved and if surgical task was completed. Each video clip was de-identified and created to capture different examples of visualization. The 20 videos consisted of 15 video clips depicting visualization in the subacromial space and five within the glenohumeral joint. The clips were then randomized utilizing a random number generator (random.org) to an order of 1–20 and distributed to six different raters. After their initial evaluation, there was a 1-month interval in data collection following which the 20 video clips were re-randomized and distributed to the six raters for a second round of evaluation. One month was agreed on collectively by raters. All raters were board-certified and fellowship-trained sports medicine orthopedic surgeons with widespread shoulder arthroscopy experience. Each rater was asked to complete an evaluation form during each round [Figure 1]. The evaluation forms asked each surgeon to rate the visualization according to the VAS<sup>[2]</sup> and to the SAGS (as detailed below).

### Visual analog score

The VAS is a method of measuring subjective experience on a longitudinal scale commonly depicted with “smiley faces” to “sad faces,”<sup>[1,7]</sup> similar to Figure 1. As in Avery *et al.*<sup>[2]</sup> for VAS grading, each surgeon was asked to rate their ability to visualize anatomic structures during the procedure. A score of 10 signifies perfect visualization

**VAS**  
Please watch each video clip of an arthroscopic shoulder procedure. At the end, we would ask you to rate your visualization according to a VAS scale for your overall impression. This should be based on your ability to identify anatomic structures within the shoulder.



10   9   8   7   6   5   4   3   2   1   0

**Grade**  
The second part for each video clip is a visualization grade. Please use the following description for what grade you think the visualization would most likely fall into.

**Grade 1- Excellent-** adequate visualization of anatomic structures with minimal need for additional intervention such as electrocautery or temporarily altering pump pressure

**Grade 2- Good-** adequate visualization of anatomic structures with occasional need for electrocautery or temporarily altering pump pressure

**Grade 3- Average-** visualization of anatomic structures requires frequent use of both electrocautery and altering pump pressure

**Grade 4- Poor-** inadequate visualization of anatomic structures despite using both electrocautery and altering pump pressure

Video 1  
VAS  
10   9   8   7   6   5   4   3   2   1   0  
Grade  
1   2   3   4

**Figure 1:** Example of evaluation form showing instructions to each rater with descriptions for visual analog scale and shoulder arthroscopy grading scale

of all structures, while a score of 0 signifies the inability to visualize structures.

### Shoulder arthroscopy grading scale

The SAGS was developed to provide a more objective means of grading visualization and create a simpler form of communication among surgeons. Grades are classified according to the ability to adequately visualize anatomic structures and accomplish surgical task. Grade 1 is described as being able to visualize all anatomic structures and accomplish surgical tasks (biceps tendon, humeral head, glenoid, subacromial bursa, labrum, rotator cuff muscles, glenohumeral ligaments), Grade 2 is described as identifying most anatomic structures and accomplishing surgical tasks (50% of anatomic structures), Grade 3 is described as visualizing some anatomic structures but not accomplishing surgical tasks (25% anatomic structures), and finally, Grade 4 is listed as not being able to visualize anatomic structures or accomplish surgical tasks.

### Statistical analysis

Separate intraclass correlation coefficients (ICCs) with a two-way random effects model were calculated to assess interobserver reliability (average measures) and intraobserver reliability (single measures) in the VAS and SAGS (IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp). The objective of this analysis was to determine consistency of responses rather than absolute agreement.<sup>[8]</sup> The ICC quantitative values are categorized as: excellent for values 0.75–1.0; strong, 0.60–0.74; moderate, 0.40–0.59; poor, <0.40. We calculated ICCs in lieu of weighted kappa coefficients, which are commonly applied to ordinal data, because

weighted kappa coefficients have some notable limitations and are therefore not universally endorsed.<sup>[9]</sup> Using NCSS software (Hintze, J.[2011]. PASS 11. NCSS, LLC. Kaysville, Utah, USA), a sample size of six raters with 40 observations per subject achieves 100% power to detect an ICC of at least 0.50 under the alternative hypothesis (null hypothesis ICC = 0.00), at alpha = 0.05.

## Results

### Visual analog scale

The interobserver reliability showed an excellent degree of consistency (ICC = 0.96, 95% confidence interval [CI] = 0.93–0.98). Likewise, the intraobserver variability exhibited a strong degree of consistency, with five of the six raters' average ICCs ranging from 0.87 to 0.97.

### Shoulder arthroscopy grading scale

The interobserver reliability showed an excellent degree of consistency (ICC = 0.97, 95% CI = 0.94–0.98). The intraobserver variability also demonstrated strong consistency, with five of the six raters average ICCs ranging from 0.61 to 0.93.

It is important to note that one rater demonstrated poor consistency regarding intraobserver reliability when using the VAS and SAGS, with ICCs of 0.06 and 0.27, respectively. Tables 1 and 2 demonstrate recorded results of the SAGS and VAS.

## Discussion

Our study found a strong-to-excellent degree of consistency in rating intraoperative visualization using

**Table 1: Interobserver reliability averages for visual analog scale and shoulder arthroscopy grading scale**

	ICC, 95% CI
VAS	0.96 (0.93-0.98)
SAGS	0.97 (0.94-0.98)

CI=Confidence interval, ICC=Intraclass correlation coefficient, VAS=Visual analog scale, SAGS=Shoulder arthroscopy grading scale

**Table 2: Intraobserver reliability for each rater for visual analog scale and shoulder arthroscopy grading scale**

	VAS	SAGS
	ICC for single measures, 95% CI	ICC for single measures, 95% CI
Reviewer #1	0.89 (0.76-0.96)	0.89 (0.75-0.96)
Reviewer #2	0.05 (-0.39-0.47)	0.27 (-0.19-0.63)
Reviewer #3	0.90 (0.76-0.96)	0.61 (0.24-0.82)
Reviewer #4	0.88 (0.72-0.95)	0.90 (0.76-0.96)
Reviewer #5	0.92 (0.80-0.97)	0.85 (0.66-0.94)
Reviewer #6	0.87 (0.69-0.95)	0.83 (0.62-0.93)

CI=Confidence interval, ICC=Intraclass correlation coefficient, VAS=Visual analog scale, SAGS=Shoulder arthroscopy grading scale

both the VAS and the SAGS. The use of VAS for rating visualization in shoulder arthroscopy was initially supported by Jensen *et al.*,<sup>[1]</sup> who demonstrated that the addition of epinephrine to irrigation fluid seems to reduce intra-articular bleeding during routine arthroscopic shoulder surgery and improved visualization. The VAS was then validated by Avery *et al.*, who also found that visualization improved when epinephrine was added to irrigation fluid. However, Chierichini *et al.*<sup>[10]</sup> evaluated the effects of different additives into irrigation fluid on visualization and found that the VAS was unreliable for this application. The need for a formally validated scoring system is evident and lacking in literature despite its use and support in previous studies.<sup>[1,2,10]</sup>

Proper communication is the primary goal of any universal scoring system. A successful system should demonstrate simplicity, accuracy, and reliability.<sup>[6]</sup> Due to the wide scoring range and subjectivity of the VAS, we expected lower interobserver and intraobserver variability for rating arthroscopic visualization. However, our results show that the VAS was comparable in consistency to our more objective four-point grading scale. Five of our six raters demonstrated a high degree of consistency (83.3%), providing preliminary evidence of the reliability of the two scoring systems.

Our study has several strengths. First, a large number of evaluators scored visualization in a large number of video clips from a variety of shoulder arthroscopic procedures. This should reduce the bias that any individual evaluator or video clip would have upon the results. Second, all evaluators were sports medicine fellowship-trained physicians who have widespread experience with arthroscopic images; therefore, their scrutiny of scoring visualization should be higher than that of other subspecialties, which comports with previous studies.<sup>[1,2,10]</sup> Finally, similar to previous work by Bellamy *et al.*,<sup>[11,12]</sup> our use of ICCs to measure consistency of responses is important for validating each scoring system since small variations (i.e., one-point change in VAS) are unlikely to affect technique.

### Limitations

The main limitation to our study is the lack of generalizability of these results to the orthopedic community as a whole. It would not be surprising if inclusion of general orthopedic surgeons lowered the degree of consistency in either system, especially in the VAS system where more variability in number of response and subjectivity exists. Another would be the length of video clips used and its representation of assessment for an entire procedure. While assessing visualization for a procedure in its entirety would be optimal, doing so would decrease the number of scoring

opportunities available for assessment. Selection bias for time point used for each video is also important. Videos were chosen to fall in a variety of grading categories, with each being succinctly different from the previous. Providing a larger number of videos would help cover the entire scale of visualizations and provide a more accurate analysis of both grading systems based on visualization ability. Combining visualization scores for the glenohumeral joint and subacromial space may not be appropriate since these represent two anatomically distinct regions of the shoulder with potentially poorer visualization on average for the subacromial space. However, given the strong degree of consistency in our results, with the majority depicting the subacromial space, this probably is inconsequential. Finally, we did have one outlier in our ratings, which may have been due to multiple reasons, including inaccuracy of recording responses. Our results were reported as a whole as to not violate the integrity of our study. Excluding such outliers and including additional raters could show more of a discrepancy between the two systems.

## Conclusion

Given the strong-to-excellent degree of consistency in using the VAS and the SAGS, either can be reliably used as a measurement of visualization in shoulder arthroscopy. The objective criteria and the smaller discrete options of the SAGS may lead to more reliable communication among surgeons. However, the relative ease of understanding and use of the VAS can also allow it to be the easier grading scale to use among different specialties within and outside of orthopedic surgery.

## Financial support and sponsorship

Nil.

## Conflicts of interest

There are no conflicts of interest.

## References

1. Jensen KH, Werther K, Stryger V, Schultz K, Falkenberg B. Arthroscopic shoulder surgery with epinephrine saline irrigation. *Arthroscopy* 2001;17:578-81.
2. Avery DM 3<sup>rd</sup>, Gibson BW, Carolan GF. Surgeon-rated visualization in shoulder arthroscopy: A randomized blinded controlled trial comparing irrigation fluid with and without epinephrine. *Arthroscopy* 2015;31:12-8.
3. Ogilvie-Harris DJ, Weisleder L. Fluid pump systems for arthroscopy: A comparison of pressure control versus pressure and flow control. *Arthroscopy* 1995;11:591-5.
4. Morrison DS, Schaefer RK, Friedman RL. The relationship between subacromial space pressure, blood pressure, and visual clarity during arthroscopic subacromial decompression. *Arthroscopy* 1995;11:557-60.
5. Huskisson EC. Measurement of pain. *Lancet* 1974;2:1127-31.
6. Ayril X, Gueguen A, Ike RW, Bonvarlet JP, Frizziero L, Kalunian K, et al. Inter-observer reliability of the arthroscopic quantification of chondropathy of the knee. *Osteoarthritis Cartilage* 1998;6:160-6.
7. Hasson D, Arnetz B. Validation and findings comparing the VAS vs. likert scales for psychosocial measurements. *Int Electron J Health Educ* 2005;8:178-92.
8. Hallgren KA. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutor Quant Methods Psychol* 2012;8:23-34.
9. Jakobsson U, Westergren A. Statistical methods for assessing agreement for ordinal data. *Scand J Caring Sci* 2005;19:427-31.
10. Chierichini A, Frassanito L, Vergari A, Santoprete S, Chiarotti F, Saccomanno MF, et al. The effect of norepinephrine versus epinephrine in irrigation fluid on the incidence of hypotensive/bradycardic events during arthroscopic rotator cuff repair with interscalene block in the sitting position. *Arthroscopy* 2015;31:800-6.
11. Bellamy N, editor. Reliability. In: *Musculoskeletal Clinical Metrology*. Dordrecht: Kluwer Academic Publishers; 1993. p. 11-24.
12. Bellamy N, Campbell J, Haraoui B, Gerez-Simon E, Buchbinder R, Hobby K, et al. Clinimetric properties of the AUSCAN osteoarthritis hand index: An evaluation of reliability, validity and responsiveness. *Osteoarthritis Cartilage* 2002;10:863-9.