Taylor & Francis
Taylor & Francis Group

Journal of Manual & Manipulative Therapy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/yjmt20

Inter-rater reliability of scapular dyskinesis classification in overhead athletes by entry-level physical therapy students

Juan Guerrero-Henriquez, Mauricio Oyarce-Mella, Jimmy Reyes Rocabado, David Olivares-Ponce, Omar Olivares-Lee, Stefano Pérez-Pasten & Martin Vargas-Matamala

To cite this article: Juan Guerrero-Henriquez, Mauricio Oyarce-Mella, Jimmy Reyes Rocabado, David Olivares-Ponce, Omar Olivares-Lee, Stefano Pérez-Pasten & Martin Vargas-Matamala (2021): Inter-rater reliability of scapular dyskinesis classification in overhead athletes by entry-level physical therapy students, Journal of Manual & Manipulative Therapy, DOI: 10.1080/10669817.2021.1972653

To link to this article: https://doi.org/10.1080/10669817.2021.1972653

Published online: 30 Aug 2021.

JOURNAL OF MANUAL & MANIPULATIVE THERAPY
https://doi.org/10.1080/10669817.2021.1972653

Inter-rater reliability of scapular dyskinesis classification in overhead athletes by entry-level physical therapy students

Juan Guerrero-Henriquez[a], Mauricio Oyarce-Mella[b], Jimmy Reyes Rocabado[c], David Olivares-Ponce[d], Omar Olivares-Lee[d], Stefano Pérez-Pasten[d] and Martin Vargas-Matamala[a]

[a]Rehabilitation and Human Movement Sciences Department, Faculty of Health Sciences, University of Antofagasta, Antofagasta, Chile; [b]Biomedical Department, Faculty of Health Sciences, University of Antofagasta, Antofagasta, Chile; [c]Mathematics Department, Faculty of Basics Sciences, University of Antofagasta, Antofagasta, Chile; [d]Kinesiology, Faculty of Health Sciences, University of Antofagasta, Antofagasta, Chile

ABSTRACT

Background: Scapular dyskinesis (SD) is a dysfunction associated with shoulder pathologies in overhead athletes and its classification is highly complex, especially in entry-level physical therapy students.

Objective: This research aimed to determine the effect of a video training method on inter-rater reliability of physical therapy entry-level students with no clinical experience while performing SD ratings under 4-type, severity, and 'yes or no' criteria in overhead athletes.

Methods: A total of 49 overhead athletes were evaluated by video recordings in posterior view of lifting tasks. Two entry-level students without clinical experience performed SD classifications according to previously mentioned criteria on both scapula in two separate sessions, with 1 week of difference. Data analysis was conducted based on the kappa coefficient (k) and its confidence intervals (CI).

Results: The 'yes or no' criteria present better results than the 4-type and severity criteria. In addition, a better inter-rater agreement is evidenced in the second measurement compared to the first one. Furthermore, the k coefficients reach values like those achieved by experienced evaluators.

Conclusion: Video training settings for scapular dyskinesis rating in entry-level physical therapist students with no clinical experience allows better inter-rater reliability values.

KEYWORDS
Reliability; scapular dyskinesis; overhead athletes; physical therapy specialty

Introduction

Entry-level physical therapy students training in sports injuries management is oriented to develop skills in intervention and clinical reasoning based on diagnostic procedures. Physical examination is a central part of shoulder pain conditions and alteration in scapulothoracic joint movements, also called scapular dyskinesis (SD), suggests dysfunction of the stabilizing muscles of the scapula during a functional task [1]. Whereas evidence is inconclusive that SD is a pathological entity [2], this dysfunction has been associated with different shoulder pathologies [3,4]. On the other hand, visual evaluation of the scapular motion is largely used in shoulder dysfunctions clinic assessment [5–7], especially in the sports population in which there are overhead gestures [8,9].

Overhead gestures in sports imply a highly coordinated musculoskeletal sequence that involves multidirectional and high muscle torques in the shoulder complex [10]. Identification of scapular movement impairments, the inherent workload of the sport, and psychosocial variables had been proposed for a broader approach to shoulder assessment and effective therapeutic interventions [6,11,12].

SD has been assessed preferably by visual methods [8,13] and classified according to three classification criteria: 4-type according to visible prominence of the scapula [14]; according to its severity [15,16] and 'yes or no' criteria, according to its presence or absence [17] with varying levels of reliability.

Reliability of each of the SD classification methods, inter-rater reliability values (between two or more evaluators) expressed as kappa values and agreement percentages have been used [18–20]. Assessment of SD is often carried out by physical therapists with postgraduate studies in the musculoskeletal area and clinical experience, whose kappa values range from 0.42 to 0.61 [20]. Few studies use novice evaluators, such as students or physical therapists with short clinical experience within their SD classification methodologies [1,21].

While reliability values in SD classification methods are generally low, little is known regarding the influence of the training of the evaluators involved on the agreement they can reach, therefore, teaching assessment methods to evaluate scapular motion has to be carefully interpreted [22,23]. In this context [15], suggest that standardized training using videotaped

CONTACT Juan Guerrero-Henriquez juan.guerrero@uantof.cl Universidad de Antofagasta, Avenue #02800, Antofagasta 1240000, Chile

examples of normal and abnormal scapular kinematics could improve the reliability of SD classification [15,16]. Furthermore, as previously stated, little is known about inter-rater reliability among evaluators with short or no clinical experience, therefore the purpose of this study was to determine the effect of a video training method on inter-rater reliability of physical therapy entry-level students with no clinical experience while performing SD ratings under 4-type, severity, and 'yes or no' criteria in overhead athletes.

Materials and methods

Participants

Through a descriptive non-experimental cross-sectional design study and using a non-probabilistic convenience sampling method, 49 athletes who perform sports involving overhead gestures from the basketball, volleyball, and handball disciplines of the Sports Service of the University of Antofagasta (Deportes UA) were recruited. All athletes that reported surgeries or any shoulder pathology in the last 9 months and movement limitations or pain during protocol application were excluded. Demographic data are presented in Table 1. Considering the methodological design of the present study (observational, conducted on a healthy population and without intervention), approval by an ethics committee was not required. Notwithstanding the above, this study follows the principles of the Declaration of Helsinki and all participants signed informed consent before their inclusion in the study.

Raters

Two entry-level physical therapy students (rater A and B) volunteered to participate to conduct SD assessments. Students' curriculum only includes clinical assessment and intervention on patients in the last year of their entry-level education, so they had no clinical experience at the time of SD assessment.

Table 1. Participants demographic data.

Variable	All (n = 49)	Male (n = 30)	Female (n = 19)
Age, years	20.81 ± 3.28	21 [17; 27]	20 [18; 26]
Weight, Kg	73.46 ± 1.87	74.8 [57; 119.40]	68 [50; 89]
Height, cm	169.28 ± 9.36	174.63 ± 7.22	160.85 ± 5.34
BMI	21.78 ± 0.44	21.07 [17.17; 31.26]	20.86 [16.51; 28.71]
Dominance			
Right-handed, n (%)	45 (91.8)	27 (55.1)	18 (36.7)
Left-handed, n (%)	4 (8.2)	3 (6.1)	1 (2.0)
Discipline			
Basketball, n (%)	18 (36.7)	12 (24.5)	6 (12.2)
Volleyball, n (%)	14 (28.6)	11 (22.4)	3 (6.1)
Handball, n (%)	17 (34.7)	7 (14.3)	10 (20.4)

Data expressed as mean ± standard deviation and as median [minimum value; maximum value]. For dominance and discipline, data are expressed as the number of cases (n) and percentage (%).

Video recording

Male participants were asked to remove the clothing that covered their trunk and female participants were asked to wear appropriate clothing that allowed visualization of the scapula. Participants performed five continuous elevations in the scapular plane bilaterally for 6 seconds [3 seconds of ascent and 3 seconds of descent) using a metronome. Participants were instructed to hold an external load proportional to their body weight following 15,recommendations: if the subject's body weight was less than 68.1 kg, the external load used was 1.4 kg; instead, if the body weight was 68.1 kg or more, the external load used was 2.3 kg [15].

Before recording, participants were instructed on the execution of the elevation test. A posterior view video was recorded by an independent investigator, not involved in SD rating. Videos were recorded against a single-color tone background (white or green) (Figure 1). All videos were labeled and stored in an encrypted virtual folder.

Rater's training and assessment

Both evaluators were trained with slides that included clinical aspects, operational definitions, and examples of SD [15]. Operational definitions of each SD classification are provided in Table 2. During this training session, an evaluation sheet was given for subjects' SD classification.

Access to the encrypted virtual folder was given to raters for SD. They were instructed to perform ratings on the participant's right and left scapula. Once their classifications were registered on the evaluation guideline, videos were randomized, and a new classification was requested one week after the first rating. Raters were not allowed to communicate with each other at the time, during or after the evaluation, nor were they allowed to know the results of each other's evaluations.

Statistical analysis

All statistical analyses were performed in IBM® SPSS® Statistics Version 25 (IBM Corp., Armonk, NY, USA). The Shapiro-Wilk test was used to evaluate the normality of continuous demographic variables. Data were expressed as mean and standard deviation and as median and minimum and maximum values for normal and non-normal distributions, respectively. Categorical variables were presented as the number of cases and percentage.

For inter-rater reliability analysis, Cohen's kappa (k) statistic was estimated. In addition, for each k coefficient, a 95% confidence interval (CI) was

Figure 1. Examples of recording, in posterior view, of female (A) and male (B) participants.

Table 2. Operational definitions for classification methods for scapular dyskinesis.

Author	Classification	Definition
Kibler [14]	Type 1	Prominent inferior medial scapular during arm elevation.
	Type 2	Prominent medial scapular border during motion.
	Type 3	Prominent superomedial angle during motion.
	Type 4	Symmetrical movement (no SD).
Uhl [17]	Yes	Types 1, 2, and 3: Patterns of scapular asymmetry.
	No	Type 4: Symmetric scapular motion (no SD).
McClure [15,16]	Obvious	Excessive or premature movement of the scapula during elevation or lowering of the arm.
	Subtle	Questionable evidence of abnormality, inconsistently present.
	Normal	Symmetric scapular movement (no SD).

Abbreviations: Scapular dyskinesis (SD).

estimated [CI = $k \pm 1.96$ x standard measurement error]. The range of k values were from 0 to 1, using the following scale for interpretation: <0 as poor agreement; 0.01–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1, almost perfect agreement [20].

Results

The results of the reliability analysis are summarized in Table 3. A better inter-rater agreement is evidenced in the second measurement compared to the first one, specifically, SD classification according to 'yes or no' criteria presents better results than 4-type and severity criteria.

Discussion

The purpose of this research was to determine the effect of a video training method on inter-rater reliability of physical therapy entry-level students with no clinical experience while performing SD ratings under the 4-type, severity, and 'yes or no' criteria in overhead athletes. Our results indicate that inter-rater agreement is slight to fair for 4-type, slight to moderate for severity criteria, and fair to substantial for 'yes or no' (Table 3). To the best of our knowledge, only two articles compare similar results using a similar study population.

21,evaluated shoulder muscle strength and kinematics, which included scapular control in an elevation task in an athletic population. They classified SD according to McClure's criteria [15]. Their evaluators were final year physical therapy students who achieved k values in the range of 0.67 to 0.84 (moderate agreement), which is consistent with our research [21]. In another inter-rater reliability study, 1,used a dichotomous classification of SD, similar to 'yes or no' criteria [17], with reliability results ranging from 0.31 to 0.71 (moderate to substantial agreement), also consistent with our investigation [1]. It should be noted that the evaluators in this research were physical therapists with a minimum of 6 months of clinical experience.

Findings of our research show that, in the absence of clinical experience, the reliability results between SD classification methods are similar to those studies

Table 3. Inter-rater statistics.

	k [CI]		
	4-type	Severity	'Yes or no'
Right scapula			
First rating	0.476 [0.288; 0.664]	0.289 [0.095; 0.483]	0.714 [0.504; 0.924]
Second rating	0.500 [0.296; 0.704]	0.599 [0.399; 0.800]	0.781 [0.581; 0.981]
Left scapula			
First rating	0.374 [0.184; 0.564]	0.548 [0.356; 0.740]	0.528 [0.281; 0.775]
Second rating	0.378 [0.166; 0.590]	0.616 [0.418; 0.814]	0.733 [0.515; 0.951]

Kappa coefficient *(k)* and confidence intervals (CI): [lower bound; upper bound].

whose evaluators had extensive training or clinical experience in the area [18,20]. Although the results are similar, care should be taken in interpreting them given the wide heterogeneity, as evidenced by the confidence intervals.

Similar *k* may be due in part to the applied methodology, in which, before video classification, evaluators are trained with images and videos. An advantage of this type of training is that uses video recordings of subjects with SD, which allows the evaluator to become familiarized with this methodology [15,16]. The methodology applied (i.e. video classification) both in the training and in the first evaluation could also explain the improvement in the *k* coefficients in the second evaluation.

Regarding the criteria used, SD classification is particularly complex in the clinical context. The use of 4-type and severity criteria tends to be more complex to define since it requires more specific observational criteria regarding the most prominent bony eminences and severity, respectively [5,15,16], which would explain the decreased reliability and the width observed in confidence intervals. On the other hand, the 'yes or no', criteria allows a dichotomous classification which facilitates the classification of the dysfunction [17]. The latter operational definition may improve inter-rater agreement; therefore, the authors recommend that in reliability studies, rigorous training methodologies should be followed.

Although SD reliability is low and its clinical relevance has also been questioned, it continues to be part of shoulder girdle health conditions evaluation. Furthermore, SD classification systems continue to be part of clinical training in entry-level, post-professional, and continuing education settings in physical therapy and manual therapy schools. This research aims to contribute to the improvement of methodologies to achieve better reliability in SD assessment and our results suggest that video-based training could accomplish this goal even in students without clinical experience.

Limitations

We should mention some limitations of our study. For this research, overhead university athletes were selected, which could limit the extrapolation of our inter-rater agreement to the general population. Furthermore, only asymptomatic participants were selected, limiting the extrapolation of these findings to a symptomatic population or to a population with shoulder health conditions. The sampling methodology used in this research could bias the inclusion of more complex SD cases which, added to the complexity described in the evaluation, could modify the inter-rater reliability results. Regarding the methodology, the use of video ratings considers the same subject in the two instances of classification and does not consider the variability in the same subject between different evaluations.

Conclusion

We can conclude that the use of video training settings for scapular dyskinesis rating in entry-level physical therapist students with no clinical experience allows similar inter-rater reliability values compared to those evidenced in clinicians trained in the classification of this dysfunction in overhead athletes.

Acknowledgments

This research was promoted by the Nucleus of Research in Rehabilitation and Human Movement Sciences (Decree 1076/2015). The authors would also like to thank the Sports Service of the University of Antofagasta (Deportes UA) staff, especially Daniel Orrego for providing the facilities for the application of the methodology and the athletes included in this research.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Juan Guerrero-Henriquez, PT, MSc, currently works as an academic in the Rehabilitation and Human Movement Sciences Department at University of Antofagasta, Chile. He has a master's degree in Orthopedic Manual Therapy from Andres Bello University, Chile. He is member of the Chilean Association of Movement Sciences (ACCM). His research lines are human movement analysis, musculoskeletal science, and data science.

Mauricio Oyarce-Mella, PT, MSc, currently works as an academic in the Biomedical Department at University of Antofagasta, Chile and as physical therapist at MEDS Sports Medicine Clinic. He has a master's degree in Orthopedic Manual Therapy from Andres Bello University, Chile. His research interests are human anatomy and musculoskeletal physical therapy.

Jimmy Reyes Rocabado is a State Professor of Mathematics from the Universidad de Chile and holds a master's degree in Mathematical Statistics from the Inter-American Center for Statistics Education. He currently works as a full-time academic at the Universidad de Antofagasta, teaching at undergraduate and graduate level. His research line is distribution theory and applications, in addition to research in areas of teaching and health sciences. He is an institutional member of the Chilean Statistical Society (SOCHE).

David Olivares-Ponce, PT, currently works in the attention of elderly people. His professional development interests are physical therapy rehabilitation for the elderly.

Omar Olivares-Lee has a degree in physical therapy from the Universidad de Antofagasta, Chile. Currently, he is an intern at the Valdivieso Family Health Center in Antofagasta, Chile. His development interest is physical therapy intervention in elderly people.

Stefano Pérez-Pasten, PT, currently works as sanitary customs personnel in Aeródromo El Loa, due to the COVID-19 pandemic. His development interest are musculoskeletal and sports physical therapy.

Martín Vargas-Matamala, PT, DO, MSc, currently works as an academic in the Rehabilitation and Human Movement Sciences Department at University of Antofagasta, Chile. He is member of the Chilean Association of Movement Sciences (ACCM). His research lines are human movement analysis and musculoskeletal rehabilitation.

ORCID

Juan Guerrero-Henriquez http://orcid.org/0000-0002-9071-3586

References

[1] Larsen CM, Søgaard K, Eshoj H, et al. Clinical assessment methods for scapular positioning and function. An inter-rater reliability study. Physiother Theory Pract. 2019. DOI:10.1080/09593985.2019.1579284

[2] Plummer H, Sum J, Pozzi F, et al. Observational scapular dyskinesis: known-groups validity in patients with and without shoulder pain. J Orthop Sports Phys Ther. 2017;14(8):1–25.

[3] Burn MB, McCulloch PC, Lintner DM, et al. Prevalence of scapular dyskinesis in overhead and nonoverhead athletes: a systematic review. Orthop J Sports Med. 2016;4(2):1–8.

[4] Hickey D, Solvig V, Cavalheri V, et al. Scapular dyskinesis increases the risk of future shoulder pain by 43% in asymptomatic athletes: a systematic review and meta-analysis. Br J Sports Med. 2018;52(2):102–110.

[5] Kibler B, Ludewig P, McClure P, et al. Clinical implications of scapular dyskinesis in shoulder injury: the 2013 consensus statement from the "scapular summit". Br J Sports Med. 2013;47(14):877–885.

[6] Klintberg IH, Cools AMJ, Holmgren TM, et al. Consensus for physiotherapy for shoulder pain. Int Orthop. 2015;39(4):715–720.

[7] Struyf F, Nijs J, Meeus M, et al. Does scapular positioning predict shoulder pain in recreational overhead athletes? Int J Sports Med. 2014;35:75–82.

[8] Huang T, Lin J, Ou H, et al. Movement pattern of scapular dyskinesis in symptomatic overhead athletes. Sci Rep. 2017;7(1):1–7.

[9] Kibler B, Sciascia A. The shoulder at risk: scapular dyskinesis and altered glenohumeral rotation. Operative Tech Sports Med. 2016;24(3):162–169.

[10] Bakshi N, Freehill M. The overhead athletes shoulder. Sports Med Arthrosc Rev. 2018;26 (3):88–94.

[11] Cools AMJ, Struyf F, De Mey K, et al. Rehabilitation of scapular dyskinesis: from the office worker to the elite overhead athlete. Br J Sports Med. 2014;48(8):692–697.

[12] Hogan C, Corbett JA, Ashton S, et al. Scapular dyskinesis is not an isolated risk factor for shoulder injury in athletes: a systematic review and meta-analysis. Am J Sports Med. 2020;1–11. DOI:10.1177/0363546520968508

[13] Lefèvre-Colau M, Nguyen C, Palazzo C, et al. Recent advances in kinematics of the shoulder complex in healthy people. Ann Phys Rehabil Med. 2018;61 (1):56–59.

[14] Kibler WB, McMullen J. Scapular dyskinesis and its relation to shoulder injury. Journal of the American Academy of Orthopaedic Surgeons. 2003;11 (2):142–151.

[15] McClure P, Tate A, Kareha S, et al. A clinical method for identifying scapular dyskinesis, Part 1: reliability. J Athl Train. 2009;44(2):160–164.

[16] Tate A, McClure P, Kareha S, et al. A clinical method for identifying scapular dyskinesis, part 2: validity. J Athl Train. 2009;44(2):165–173.

[17] Uhl T, Kibler B, Gecewich B, et al. Evaluation of clinical assessment methods for scapular dyskinesis. Arthroscopy. 2009;25(11):1240–1248.

[18] De Groef A, Van Kampen M, Vervloesem N, et al. Inter-rater reliability of shoulder measurements in middle-aged women. Physiotherapy (United Kingdom). 2017;103(2):222–230.

[19] Johnston J. Weighted kappa for multiple raters. Percept Mot Skills. 2016;107(6):837–848.

[20] Rossi DM, Pedroni CR, Martins J, et al. Intrarater and interrater reliability of three classifications for scapular dyskinesis in athletes. PLoS ONE. 2017;12 (7):1–10.

[21] Møller M, Attermann J, Myklebust G, et al. The inter- and intrarater reliability and agreement for field-based assessment of scapular control, shoulder range of motion, and shoulder isometric strength in elite adolescent athletes. Phys Ther Sport. 2018;32:212–220.

[22] Lange T, Struyf F, Schmitt J, et al. The reliability of physical examination tests for the clinical assessment of scapular dyskinesis in subjects with shoulder complaints: a systematic review. Phys Ther Sport. 2017;26:64–89.

[23] Nordqvist JM, Öberg BE, Adolfsson LE, et al. Heterogeneity among patients with subacromial pain – variabilities within clinical presentation and its impact on daily life. Physiotherapy. 2021;112:113–120.