Rate of Agreement between an Experienced and Novice Clinician in the Application of a Clinical Diagnostic Rule for Subacromial Impingement Syndrome: A Pilot Study

By:
Andrew C. Thomas, SPT, CSCS
Laura E. Sherrill, SPT, CSCS
Jan Jurado, SPT, CSCS
James F. Villers, DPT, FAAOMPT

Evidence Based Practice III Assignment
Submitted in partial fulfillment of the
Requirements of the course: PT 7344
Physical Therapy Program
Angelo State University

ABSTRACT

Background and Purpose: To determine the rate of agreement between a novice clinician and an expert clinician based on their ability to apply special tests from a clinical decision rule (CDR) for subacromial impingement syndrome (SAIS) proposed by Michener et al (2009).

Design: One group, repeated-measures reliability study.

Methods: Examination of 12 subjects with shoulder pain by a novice and expert rater included five SAIS tests: painful arc, Neer, Hawkins-Kennedy, empty can, and the external rotation test. A kappa analysis was performed to assess rate of agreement between examiners.

Results: 100% agreement (K = 1.000) existed between novice and expert findings for the Hawkins-Kennedy test and the diagnosis of SAIS through the CDR. Moderate agreement was established for the painful arc test (K = 0.567), and fair agreement was indicated by the Neer (K = 0.316), Empty Can (K = 0.297), and External Rotation (K = 0.217) tests.

Conclusions: The results of this study suggest usage of CDRs may normalize clinical knowledge between novice and expert clinicians in the physical diagnosis of SAIS.

Keywords: physical therapy, shoulder impingement, diagnosis, manual tests, inter-rater reliability

INTRODUCTION

Shoulder pathology is a common etiology of pain and disability presented to physicians.[5] Shoulder pain accounts for 33.2% of physician office visits for musculoskeletal patients in the United States from 1995 to 2005.[35] On average, 50% of the general population experience shoulder pain every year.[32] Between 44-65% of shoulder pain reports are related to subacromial impingement syndrome (SAIS).[14,28]

The subacromial space is defined inferiorly by the humeral head and superiorly by the under the surface of the anterior third of the acromion, coracoacromial ligament and the acromioclavicular joint.[28] The space between the acromion and humeral head ranges from 1.0 to 1.5 centimeters as seen on radiographs.[28] In this space lies the rotator cuff (RC) tendons, the long head of the biceps tendon, the subacromial bursa, and the coracoacromial ligament.[4] Any abnormality that disturbs the relationship of these subacromial structures may lead to SAIS.[4]

SAIS is defined as an encroachment of the subacromial tissues as a result of narrowing in the subacromial space and represents a spectrum of pathology ranging from subacromial bursitis to (RC) tendinopathy and full-thickness rotator cuff tears.[26] SAIS is a common complaint for patients of all ages and different activity levels.[21] The main consequences of SAIS are functional loss and disability, most commonly from RC tendinopathy.[15,22] Mechanisms of RC tendinopathy

have been classically described as extrinsic, intrinsic, or a combination of both. It is theorized that with intrinsic impingement, partial or full thickness RC tears are the result of degeneration from overuse, tension overload, or trauma.[22] Extrinsic impingement occurs as a result of mechanical compression by external structures causing inflammation and degeneration of the tendon.[17] Narrowing of the subacromial space is not a hallmark finding in SAIS but degeneration within the tendon is.[27] An understanding of the anatomy of the RC tendons and the underlying pathogenesis aids in the diagnosis of SAIS.[5]

There exists a wide variety of physical examination tests which are used to diagnose SAIS in clinical trials. A subset of tests examine for possible symptoms with compression of the RC tendons in the subacromial space, contraction of supraspinatus muscle, and contraction of the infraspinatus muscle. The studies which examine the psychometrics of the these tests have provided varying results.[9,18,21,25] The limitations of the studies include use of a reference standard that has its own error, a retrospective design, and non-blinding of the examiner who determined the reference standard.[21] Studies examining the reliability of SAIS diagnoses have also lead to wide range of results, with poor to almost perfect agreement (Kappa 0.18 to 1.0).[13,19,21,24]

Recent research supports the use of positive test clusters, termed clinical diagnostic rules (CDRs), for improving the clinical diagnostic accuracy for SAIS.[21] Michener et al.[21] used highly trained clinicians to perform the SAIS specialty tests.[21] In the article, a cluster of 3 + of 5 tests produced a sensitivity of 0.75 and specificity of 0.74 with a +LR of 2.93.[21] The test cluster included the painful arc test, Neer impingement test, Hawkins-Kennedy test, empty can test, and the external rotation test.[21] CDRs such as the one proposed by Michener assist clinicians in making a diagnosis, establishing a prognosis, and in the implementation of interventions.[3] Several other shoulder pathologies that have similar patterns and findings with physical examination can be mistakenly diagnosed as SAIS, thus further exemplifying the need for CDRs in diagnostic testing within the physical therapy setting.[2]

Inter tester-reliability between novice and expert raters has not been reviewed in regards to the diagnosis of SAIS as assessed by the current literature base. Current research articles specify the use of expert clinicians as it is assumed that the expert clinicians are more accurate in execution of tests and measures with pattern recognition.[20] Pattern recognition is an alternative model of clinical reasoning that is based on recognition of patterns from clinical presentations.[20] In this model, if the current patient presents similar to patients who previously demonstrated successful outcomes, similar management techniques are utilized.[20] Expert clinicians possess enhanced recall and also engage in ‘forward’ reasoning.[7] This refers to the finding with routine problems in which experts tend to work forward from the given information to a solution to the problem.[7] Expert performance may be the result of vast amounts of knowledge and pattern based retrieval acquired over many years of experience in the associated domain.[7]

The question exists if clinicians without advanced musculoskeletal training and experience can replicate similar diagnostic accuracy. A gap in knowledge exists regarding the ability of a non-expert in musculoskeletal disorders in the application and accurate interpretation of the SAIS CDR. Thus, the purpose of the study is to determine the inter-rater reliability

between a novice clinician and an expert clinician in regards to the application and interpretation of a CDR for SAIS.

METHODS

Study design

The study design was a one group, repeated-measures reliability study approved by Angelo State University's Institutional Review Board. Participants with chief complaint of shoulder pain were recruited through two local physical therapy clinics. Each participant's rights were safeguarded by an informed consent, which was read and signed. Before the examination, a medical history was completed by participants to ensure inclusion and exclusion criteria. Participants signed a waiver and agreed to withhold any history of their shoulder pathology to the examiners to eliminate bias. Testing was completed in a private examination room that allowed autonomous assessments by each examiner. One examiner was a novice clinician in his third year of physical therapy school. The second examiner (expert) was a fellowship trained orthopedic manual physical therapist with 32 years of clinical experience. No materials were needed as this study utilized manual contacts from the clinicians.

Participants

A purposive sample of 12 subjects, 4 males and 8 females, a total of 13 shoulder pathologies, with one female participant having bilateral shoulder involvement. Mean age of the participants was 60.1 (± 12.06) years with a range of 43–81 years. Eligibility criteria included adults aged from 18-85 years old. Exclusion criteria included history of previous shoulder fracture or surgery.

Procedures

The CDR proposed by Michener et al.[21] was utilized, which consisted of 5 impingement shoulder tests. To be consistent, the order of the tests given by each examiner was: painful arc, Neer, Hawkins-Kennedy, empty can, and lastly external rotation resistance test utilizing the methods as described by Meichner et al.[21] Each clinician read the article and completed his exam based on the instructions illustrated in the article. No training or practices were performed prior to testing participants. The order of which the clinicians examined the participants was randomized through a randomization engine, Research Randomizer.[29]

The painful arc test was performed by asking the participant to actively abduct their shoulder, reporting any pain. If pain was reported between 60° and 120° of abduction, the test was considered positive. The Neer test was performed by stabilizing the superior aspect of the scapular spine with an inferior force while passively fully flexing the shoulder, applying overpressure at the sign of first resistance. A positive test was noted as reproduction of pain near the acromioclavicular joint. The Hawkins-Kennedy test was performed with the examiner flexing the participant's shoulder and elbow to 90° then maximally internally rotating the shoulder, applying overpressure at the sign of first resistance. A positive test was noted as report of pain in

the superior shoulder. The empty can test was performed by elevating the subject's shoulder to 90° in the scapular plane and then asking the participants to rotate the shoulder until his/her thumb pointed toward the floor. The examiner then applied an inferiorly directed force at the wrist while the participants attempted to resist. A positive test was recorded as weakness or pain in the involved shoulder. The external rotation resistance test was performed by placing the arm at the participant's side and flexing the elbow to 90°. A medially directed force was exerted on the distal forearm to resist shoulder external rotation. A positive test was classified as weakness or pain of the involved. All tests were scored on a dichotomous scale. Participants were given a 5-minute rest period between each examination to provide the shoulder time to recover to baseline.

Statistical Analysis

SPSS for Windows version 19.0 (SPSS Inc, Chicago, IL) was used for all statistical analyses. Descriptive statistics were used to examine the measure of agreement utilizing the kappa analysis coefficient. Kappa values < 0.20 have a poor agreement, 0.21 - 0.40 have a fair agreement, 0.41 - 0.60 have moderate agreement, 0.61 - 0.80 have a good agreement, and 0.81 - 1.00 have a very good agreement.[32]

RESULTS

A total of 13 subjects (60.08 ± 12.06 years old; 4M, 8F; 43-81 years old) participated in the study. **Table 1** illustrates the results recorded by the novice and expert clinicians in regards to the test performed and the resulting diagnosis of SAIS. **Table 2** illustrates the results of the SPSS kappa analysis from for each test and the resulting diagnostic assessment.

The novice and expert demonstrated 100% agreement (K = 1.000) for the Hawkins-Kennedy test and the diagnosis of SAIS through the application of the CDR (+3/5 tests). Moderate agreement was found for the painful arc test (K = 0.567), whereas the Neer (K = 0.316), empty can (K = 0.297), and external rotation (K = 0.217) tests all demonstrated fair agreement.

DISCUSSION

This study demonstrates that even though a difference in the rate of agreement exists for the majority of individual tests between novice and expert, the CDR still demonstrated 100% agreement. This result speaks to the importance of a CDR with the fact that each examiner read an article, and regardless of experience in the clinic, came to the same conclusion. Cook et al.[6] suggested that well-constructed CDRs can improve clinical decision making and practice, which is evident in this present study. [6,26] Furthermore, the diagnosis of shoulder pathology is complex and requires more than one test secondary to the limited diagnostic accuracy of each individual test.[1,8]

Of the tests that were administered by the examiners, the Hawkins-Kennedy provided the greatest agreement at 100%. There is large variability in research in regards to the inter-tester

reliability of the Hawkins-Kennedy test. Ostor et al.[24] demonstrated a K = .18-.43 using clinicians of varying levels of training.[24] Whereas, May et al.[19] performed a systematic review of physical examination tests of patients with shoulder problems, and found that the inter-tester reliability ranged from 0.18-0.91.[19] The agreement that was demonstrated in this study could be secondary to the straightforward nature of the Hawkins-Kennedy test.

The painful arc test's agreement of 0.567 is within the range found in current research in regards to inter-rater reliability. May et al.[19] found the range of reliability scores from a systematic review was between 0.47-0.93. Michener et al.[21] found reliability score of 0.45.[19,21] Painful arc is dependent on the description given by the clinician, and the patient following those directions. The experience garnered by the expert clinician in administering the test most likely has allowed for a scripted description to increase consistency that the novice has yet to develop.

The tests that had the greatest variability in performance and require the clinician to make a clinical decision, (Neer, empty can, and external rotation) showed the least amount of agreement in this study. Each of these tests required the clinicians to utilize an amount of pressure or stabilization that they choose. The expert clinician has performed these tests on numerous patients previously and may be less apprehensive in their application of force in order to accurately perform the tests. This is consistent with the current research in which the reliability variables range from 0.1-1.00 for Neer, 0.44-.94 for the empty can test, and 0.18-0.45 for the external rotation test.[19] Jensen et al.[11] suggested that expertise and consistency with tests develops when the clinician tests and refines propositions, hypothesis, and principle-based expectations in practical situations.[11] Each of the values from this study fall into the range found in the research except for the empty can test. This could be the result of inconsistency of testing, and lack of experience by the novice with shoulder pain/SAIS patients.

In regards to the reference article from Michener et al.[21] in which the CDR was based from experienced clinicians, the results of the kappa analysis differ from the current study and are illustrated in **Table 3**. Michener et al.[21] found the greatest rate of agreement from the external rotation test, and the least was the Hawkins-Kennedy test.[21] The opposite was true in regards to the agreement in this study between the novice and expert clinicians. This could be a result of clinical decision making, expert clinicians using a directive factor approach, or novice clinician having a lack of experience.[34] May et al.[19] demonstrated that there is a significant lack of clinical reasoning that exists with the novice therapists. Novice therapists utilized hypothetico-deductive reasoning only whereas the expert therapists utilize pattern recognition and hypothetico-deductive reasoning to form a diagnosis.[20] Experience in the clinic provides the novice with the ability to transition to use hypothetico-deductive reasoning with pattern recognition to provide more accurate diagnoses from both the history and physical examination.[20]

Diagnostic values of various tests for SAIS vary considerably when considered individually.[14] The research demonstrates mixed results in terms of the sensitivity, specificity and reliability of physical tests for SAIS.[19,25] In a systematic review performed by Hughes et al.[10] examined the sensitivity and specificity of shoulder impingement tests. The Hawkins-Kennedy test's sensitivity and specificity range from 71.5%-100% and 23.0%-66.3% respectively. The

Neer Impingement's range from 59.3% - 91.6% and 28.5% - 68.7% respectively. The painful arc's range from 9.5% - 75.8% and 47.0% - 88.4% respectively. The empty can's range from 32.1% - 98.5% and 40.0% - 89.9% respectively. Lastly, the external rotation's range from 19.4% - 84% and 53% - 90.1% respectively.[10]This variability demonstrates that the individual tests cannot be performed to determine rule in/out criteria with confidence and a CDR would provide the clinician with the best clinical diagnosis.

There are some limitations to this study. The first limitation is the small sample population (n = 13). Second, all subjects were older than 43, therefore neglecting those in the younger age bracket. The results from this study are more indicative of a population with joint degeneration as opposed to traumatic injury. A third limitation to this study was the noticeable age difference between the expert and novice examiners, making participant blinding to the level of experience of the examiner a non-option.

CONCLUSION

In conclusion, the results of this study suggest that CDRs may help normalize clinical reasoning and physical diagnosis between novice and expert clinicians. Specific tests to diagnose SAIS demonstrated decreased inter-rater reliability between novice and expert clinicians in comparison to the CDR attesting to the strength of the CDR. The Hawkins-Kennedy test was the only test to provide excellent agreement between the novice and expert clinicians. Therefore, the Hawkins-Kennedy test may be the most reliable single test for subacromial impingement between novice and expert clinicians. The data gathered from this study, though limited, demonstrates the need for further research in the comparison of a novice and expert clinician in the application of a CDR to compare their results to one another.

ACKNOWLEDGEMENTS

The researchers would like to acknowledge Heidi S. Moyer for her assistance with manuscript formatting, editing, and content clarity.

References

1. Alqunaee M, Galvin R, Fahey T. Diagnostic Accuracy of Clinical Tests for Subacromial Impingement Syndrome: A Systematic Review and Meta-Analysis. Archives Of Physical Medicine & Rehabilitation [serial online]. February 2012;93(2):229-236. Available from: CINAHL with Full Text, Ipswich, MA. Accessed April 6, 2015.
2. Bas de Witte et al. Study protocol subacromial impingement syndrome: the identification of pathophysiologic mechanisms (SISTIM). BMC Musculoskeletal Disorders [serial online]. January 2011;12(1):282-292. Available from: Academic Search Complete, Ipswich, MA. Accessed February 25, 2015.
3. Beneciuk J, Bishop M, George S. Clinical prediction rules for physical therapy interventions: a systematic review. Physical Therapy [serial online]. February 2009;89(2):114-124. Available from: CINAHL with Full Text, Ipswich, MA. Accessed February 26, 2015.
4. Bigliani LU, Levine WN. Subacromial impingement syndrome. J Bone Joint Surg Am. 1997;79:1854–68.
5. Boykin R, Heuer H, Vaishnav S, Millett P. Rotator cuff disease-basics of diagnosis and treatment. Italy Rheumatology Reports. 2010;2:e1:1-12.
6. Cook C. Potential pitfalls of clinical prediction rules [editorial]. J Man Manip Ther. 2008;16:69 –71
7. Doody C, McAteer. Clinical Reasonng of Expert and Novice Physiotherapists in an Outpatient Orthopaedic Setting. Physiotherapy. 2002:88: 258-268.
8. Hegedus E, Goode A, Wright A, et al. Which physical examination tests provide clinicians with the most value when examining the shoulder? Update of a systematic review with meta-analysis of individual tests. British Journal Of Sports Medicine [serial online]. November 2012;46(14):964-978. Available from: CINAHL with Full Text, Ipswich, MA. Accessed April 6, 2015.
9. Holtby, R. and Razmjou, H. Validity of the supraspinatus test as a single clinical test in diagnosing patients with rotator cuff pathology. J Orthop Sports Phys Ther. 2004; 34: 194–200
10. Hughes P, Taylor N, Green R. Most clinical tests cannot accurately diagnose rotator cuff pathology: a systematic review. Australian Journal Of Physiotherapy [serial online]. January 1, 2008;54:159-170. Available from: ScienceDirect, Ipswich, MA. Accessed May 1, 2015.
11. Jensen GM, Shepard KF, Hack LM. The novice versus the experienced clinician: insights into the work of the physical therapist. *Phys Ther.* 1990;70:314-323.
12. Jobe, F.W. and Moynes, D.R. Delineation of diagnostic criteria and a rehabilitation program for rotator cuff injuries. Am J Sports Med. 1982; 10: 336–339.
13. Johansson, K and Ivarson,S, Intra- and interexaminer reliability of four manual shoulder maneuvers used to identify subacromial pain. Manual Therapy. 2009:14, 2,:231-239. Available from: SwePub, Ipswich, MA. Accessed March 18, 2015.
14. Kelly S. The value of physical tests for subacromial impingement syndrome: a study of diagnostic accuracy. Clinical Rehabilitation [serial online]. February 2010;24(2):149-158. Available from: Psychology and Behavioral Sciences Collection, Ipswich, MA. Accessed February 26, 2015.
15. Koester MC, George MS, Kuhn JE. Shoulder Impingement Syndrome. Am J Med. 2005 May; 118(5):452-5.

16. Lewis JS. Rotator cuff tendinopathy/subacromial impingement syndrome: is it time for a new method of assessment? Br J Sports Med 2009;43:259–64.
17. Ludewig PM, Cook TM. Translations of the humerus in persons with shoulder impingement symptoms. J Orthop Sports Phys Ther. 2002;32:248–59.
18. MacDonald, P.B., Clark, P., and Sutherland, K. An analysis of the diagnostic accuracy of the Hawkins and Neer subacromial impingement signs. J Shoulder Elbow Surg. 2000; 9: 299–301
19. May S et al. Systematic Review: Reliability of physical examination tests used in the assessment of patients with shoulder problems: a systematic review. Physiotherapy [serial online]. January 1, 2010;96:179-190. Available from: ScienceDirect, Ipswich, MA. Accessed February 28, 2015.
20. May S, Withers S, Reeve S, Greasley A. Limited clinical reasoning skills used by novice physiotherapists when involved in the assessment and management of patients with shoulder problems: a qualitative study. J Man Manip Ther. 2010 Jun; 18(2): 84–88.
21. Michener et. al. Reliability and Diagnostic Accuracy of 5 Physical Examination Tests and Combination of Tests for Subacromial Impingement. Arch Phys Med Rehabil. 2009:90:1898-1903.
22. Michener LA, McClure PW, Karduna AR. Anatomical and biomechanical mechanisms of subacromial impingement syndrome. Clin Biomech (Bristol, Avon) 2003;18:369–79.
23. Moezy M, Sepehrifar S, Dodaran MS. The effects of scapular stabilization based exercise therapy on pain, posture, flexibility and shoulder mobility in patients with shoulder impingement syndrome: a controlled randomized clinical trial. Med J Islam Repub Iran.2014;28:87.
24. Ostor A, Richards C, Prevost A, Hazleman B, Speed C. Interrater reproducibility of clinical tests for rotator cuff lesions. Annals Of The Rheumatic Diseases [serial online]. October 2004;63(10):1288-1292. Available from: MEDLINE, Ipswich, MA. Accessed February 28, 2015.
25. Park H, Yokota A, Gill H, El Rassi G, McFarland E. Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndrome. Journal Of Bone & Joint Surgery, American Volume [serial online]. July 2005;87A(7):1446-1455. Available from: CINAHL with Full Text, Ipswich, MA. Accessed February 26, 2015.
26. Reilly B, Evans A. Translating Clinical Research into Clinical Practice: Impact of Using Prediction Rules To Make Decisions. Annals Of Internal Medicine [serial online]. February 7, 2006;144(3):201-W36. Available from: Academic Search Complete, Ipswich, MA. Accessed February 26, 2015.
27. Seitz AL, McClure PW, Finucane S, et al. Mechanisms of rotator cuff tendinopathy: intrinsic, extrinsic, or both? Clin Biomech (Bristol, Avon) 2011;26:1–12.
28. Umer M, Qadir I, Azam M. Subacromial impingement syndrome. Orthop Rev (Pavia) 2012;4(2):e18.
29. Urbaniak, G. C., & Plous, S. (2013). Research Randomizer (Version 4.0) [Computer software]. Retrieved on June 22, 2013, from http://www.randomizer.org/
30. van de Sande M, Stoel B, Rozing P. Subacromial space measurement. A reliable method indicating fatty infiltration in patients with rheumatoid arthritis. Clinical Orthopaedics & Related Research [serial online]. October 2006;451:73-79. Available from: CINAHL with Full Text, Ipswich, MA. Accessed February 28, 2015.

31. Van der Windt DA, KoesBW, de Jong BA, Bouter LM. Shoulder disorders in general practice: incidence, patient characteristics, and management. Ann Rheum Dis. 1995;54:959–64.
32. Viera A, Garrett J. Understanding interobserver agreement: the kappa statistic. Family Medicine [serial online]. May 2005;37(5):360-363. Available from: MEDLINE, Ipswich, MA. Accessed February 28, 2015.
33. Vind M, Bogh S, Larsen C, Knudsen H, Søgaard K, Juul-Kristensen B. Inter-examiner reproducibility of clinical tests and criteria used to identify subacromial impingement syndrome. BMJ Open [serial online]. May 18, 2011;1(1):e000042. Available from: MEDLINE, Ipswich, MA. Accessed February 26, 2015.
34. Wainwright S, Shepard K, Harman L, Stephens J. Factors that influence the clinical decision making of novice and experienced physical therapists. Physical Therapy [serial online]. 2011;91(1):87-101. Available from: CINAHL with Full Text, Ipswich, MA. Accessed February 26, 2015.
35. Wofford JL, Mansfield RJ, Watkins RS. Patient characteristics and clinical management of patients with shoulder pain in U.S. primary care settings: secondary data analysis of the National Ambulatory Medical Care Survey. BMC Musculoskelet Disord 2005;6:4

Table 1: Expert vs. Novice Test Results					
Test			Novice		Total
			Negative	Positive	
Painful Arc	Expert	Negative	2	1	3
		Positive	1	9	10
	Total		3	10	13
Neer	Expert	Negative	1	0	1
		Positive	3	9	12
	Total		4	9	13
Hawkins-Kennedy	Expert	Negative	2	0	2
		Positive	0	11	11
	Total		2	11	13
Empty Can	Expert	Negative	2	3	5
		Positive	1	7	8
	Total		3	10	13
External Rotation	Expert	Negative	3	3	6
		Positive	2	5	7
	Total		5	8	13
CDR	Expert	Negative	3	0	3
		Positive	0	10	10
	Total		3	10	13
Total	Expert	Negative	13	7	20
		Positive	7	51	58
	Total		20	58	78

Table 2: Kappa Analysis Symmetric Measures						
Test			Value (95% CI)	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Painful Arc	Measure of Agreement	Kappa	.567 (0.03,1.00)	.273	2.043	.041
	N of Valid Cases		13			
Neer	Measure of Agreement	Kappa	.316 (-0.18,0.81)	.253	1.561	.118
	N of Valid Cases		13			
Hawkins-Kennedy	Measure of Agreement	Kappa	1.000 (1.00, 1.00)	.000	3.606	.000
	N of Valid Cases		13			
Empty Can	Measure of Agreement	Kappa	.297 (-0.22,0.81)	.263	1.145	.252
	N of Valid Cases		13			
External Rotation	Measure of Agreement	Kappa	.217 (-0.31,0.74)	.269	.792	.429
	N of Valid Cases		13			
CDR	Measure of Agreement	Kappa	1.000 (1.00,1.00)	.000	3.606	.000
	N of Valid Cases		13			
Total	Measure of Agreement	Kappa	.529 (0.31,0.74)	.110	4.675	.000
	N of Valid Cases		78			

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.

Table 3: Comparison with Michener et al.		
Inter-tester Kappa Reliability Coefficients		
Test	Results	Michener et al.
Painful Arc	0.567	0.45
Neer	0.316	0.4
Hawkins-Kennedy	1	0.39
Empty Can	0.297	0.47
External Rotation	0.217	0.67