Construct Validity and Interrater Agreement of the Sport Injury Rehabilitation Adherence Scale


Context: Adherence to clinic-based rehabilitation might influence outcomes. Objective: To examine the construct validity and interrater agreement of a measure of adherence to clinic-based rehabilitation. Design: Repeated-measures in both study 1 and study 2. Participants: 43 student rehabilitation practitioners in study 1 and 12 patients undergoing rehabilitation after anterior cruciate ligament reconstruction in study 2. Interventions: Participants in study 1 rated the adherence of a simulated videotaped patient exhibiting high, moderate, and low adherence. Two certified athletic trainers rated the adherence of patients at 4 consecutive appointments in study 2. Main Outcome Measure: The Sport Injury Rehabilitation Adherence Scale. Results: In study 1, adherence ratings increased in a linear fashion across the 3 levels of adherence, and \( r = .84 \) and rater-agreement-index values ranging from .84 to .95 were obtained. In study 2, the rater-agreement index was .94. Conclusions: Strong support was found for the construct validity and interrater agreement of the Sport Injury Rehabilitation Adherence Scale. Key Words: compliance, measurement, psychology


Injury is a common occurrence in sport participation, both across sports and around the world. For athletes whose injuries are sufficiently severe to warrant participation in injury-rehabilitation programs, adherence to the prescribed treatment protocol is considered by sports-medicine professionals to be an important contributor to successful rehabilitation.
outcomes. Empirical support for this position has been found in several studies in which higher levels of adherence were related to better sport-injury rehabilitation outcomes.

Sport-injury rehabilitation programs typically involve some combination of clinic- and home-based activities. Adherence to the clinic-based aspect of sport-injury rehabilitation has traditionally been assessed with indices of attendance at rehabilitation sessions, which generally involve calculating the ratio of the number of sessions attended to the number of sessions scheduled. An important limitation of attendance indices is that they do not measure what patients actually do during their rehabilitation sessions. Some patients might show up for their rehabilitation appointments but just “go through the motions” and fail to exert themselves during appointments.

The Sport Injury Rehabilitation Adherence Scale (SIRAS) was designed to go beyond attendance at rehabilitation sessions and measure sport-injury rehabilitation practitioners’ ratings of patient adherence during clinic-based rehabilitation sessions. On the SIRAS (see Table 1), practitioners respond to 3 items (scored on 5-point Likert-type scales) regarding the degree to which patients exert themselves, follow practitioner instructions and advice, and are receptive to changes in the rehabilitation program during a given rehabilitation session. Scores for the 3 items are summed to create a composite score. In the initial validation studies, preliminary support was provided for the test–retest reliability (ICC = .77 over a 1-week period), internal consistency (Cronbach α = .82), and unidimensionality of the SIRAS. Significant positive correlations have been obtained between SIRAS scores and other measures of adherence.

### Table 1  Sport Injury Rehabilitation Adherence Scale*

<table>
<thead>
<tr>
<th>Item</th>
<th>Score Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Circle the number that best indicates the intensity with which this patient completed rehabilitation exercises during today’s appointment:</td>
<td>1  2  3  4  5</td>
</tr>
<tr>
<td>minimum effort</td>
<td>maximum effort</td>
</tr>
<tr>
<td>2. During today’s appointment, how frequently did this patient follow your instructions and advice?</td>
<td>1  2  3  4  5</td>
</tr>
<tr>
<td>never</td>
<td>always</td>
</tr>
<tr>
<td>3. How receptive was this patient to changes in the rehabilitation program during today’s appointment?</td>
<td>1  2  3  4  5</td>
</tr>
<tr>
<td>very unreceptive</td>
<td>very receptive</td>
</tr>
</tbody>
</table>

*The scale can also be used with reference to adherence tendencies in general by using the present tense (without reference to “today’s appointment”). Reprinted from Physical Therapy in Sport, 1, Brewer et al, Preliminary psychometric evaluation of a measure of adherence to clinic-based sport injury rehabilitation, 68-74, 2000, by permission of the publisher Churchill Livingstone.
and both attendance at rehabilitation sessions\textsuperscript{13,14} and adherence to home-based rehabilitation activities,\textsuperscript{9} providing partial evidence of the construct validity of the SIRAS. A relationship between SIRAS scores and adherence to clinic-based rehabilitation activities has, however, yet to be established.

For the SIRAS to be used more confidently, it is important to ensure the comparability of scores across rehabilitation practitioners. An interrater reliability coefficient of .57 for the SIRAS has been reported,\textsuperscript{13} but this value was based on the correspondence between mean SIRAS scores of different practitioners across different rehabilitation sessions. A more accurate indication of the interrater agreement of the SIRAS could be obtained having the same rehabilitation practitioners complete the SIRAS for the same patients across the same series of rehabilitation appointments. Interrater agreement was considered a more relevant psychometric property of the SIRAS than interrater reliability for 3 main reasons. The first reason relates to the SIRAS being based on subjective judgments with no "correct" responses.\textsuperscript{15} Second, because SIRAS scores in clinical settings tend to cluster at the high end of the scale,\textsuperscript{9,14,16} an index of interrater agreement takes into account the fact that the range of the scale is paramount.\textsuperscript{17,18} Third, from a practical standpoint, it is more critical for different practitioners to make similar judgments about a given patient than it is for the practitioners to make proportional judgments across a series of patients without regard to the similarity of the judgments.\textsuperscript{19}

The purpose of study 1 was to further examine the psychometric properties of the SIRAS. In particular, the construct validity and interrater agreement of the SIRAS were evaluated by having student rehabilitation practitioners complete the SIRAS for a hypothetical videotaped patient across high, moderate, and low adherence conditions. Student rehabilitation practitioners were selected as participants because, given that the clinical sport-rehabilitation environment is populated with practitioners varying extensively in training and experience (e.g., athletic trainers, occupational therapists, physical therapists, physical therapy assistants, physician assistants, student trainees), it was important to demonstrate that the SIRAS could be used reliably even by individuals with relatively little clinical experience. It was assumed that if student practitioners could use the SIRAS reliably, the scale could be used at least as reliably by more experienced professionals.

\textbf{Study 1}

\textbf{Method}

\textit{Participants.} Participants were 26 students enrolled in an undergraduate athletic training program (mean = 3.20, SD = 0.71 years in the program; mean = 656.00, SD = 272.08 hours of clinical experience) and 17 students enrolled in an undergraduate/graduate physical therapy program (mean = 5.65, SD = 1.62 years in the program; mean = 729.41, SD = 279.83 hours of
clinical experience) at a small private college. Participants in the convenience sample ranged in age from 20 to 43 years. Thirty-one (72%) were women and 12 (28%) were men. Thirty-nine (91%) were Caucasian, 2 (5%) were Asian/Pacific Islander, and 2 (5%) were Hispanic.

Procedure. This study was approved by the Institutional Review Board at Springfield College. After being read a brief description of the study, participants completed an informed-consent document and a demographic questionnaire that included items pertaining to gender, age, race, academic major, year in program, and amount of clinical experience. Participants then viewed a video program displaying 3 different rehabilitation vignettes in a counterbalanced order. Each vignette was roughly 9 minutes in length. The 3 vignettes, which depicted a highly adherent patient, a moderately adherent patient, and a minimally adherent patient, showed the same rehabilitation procedure, the same patient, the same therapist, and standardized dialogue. The attitudes and behaviors of the patient differed across the vignettes in a manner consistent with high, moderate, and low levels of adherence. At the conclusion of each vignette, the participants were asked to rate the adherence of the patient in the video using the SIRAS. Ratings were completed independently and collected after the third vignette.

Statistical Analyses. Descriptive statistics were calculated for SIRAS scores across the 3 vignettes, age, year in academic program, and amount of clinical experience. A mixed analysis of variance (ANOVA) with 1 within-subjects (repeated-measures) factor (the 3 video vignettes) and 1 between-subjects factor (presentation order of video vignettes) was computed to compare participants’ SIRAS scores across the 3 vignettes and the 6 possible video-presentation orders. Significant effects were followed up with Bonferroni pairwise comparisons. Eta-squared effect-size calculations were performed to determine the proportion of variance in SIRAS scores accounted for by experimental condition. To ensure that differences in SIRAS scores were not simply a function of repeated exposure to the vignettes, a 1-way independent-groups ANOVA was computed to examine participants’ SIRAS scores for the first vignette that they viewed, independent of the influence of viewing other vignettes. This analysis was followed up with a Fisher least-significant-difference procedure. Interrater agreement was computed for the SIRAS with the $r_{WCG0}$ statistic and the rater-agreement index (RAI) statistic for each of the 3 vignettes, as well as by using the RAI across all 3 vignettes. Descriptive statistics, eta-squared, and ANOVA statistical analyses were computed using the Statistical Package for the Social Sciences. Post hoc comparisons for the independent-groups ANOVA and interrater-agreement statistics were computed by hand.

Results

The mixed ANOVA performed on SIRAS scores across the 3 video vignettes revealed that the main effect for video-vignette presentation order was not statistically significant, $F_{5,37} = 1.73, p > .10$. Although the 2-way interaction
between video vignette and video-vignette presentation order was statistically significant, $F_{10,74} = 1.20, P = .05$, examination of the interaction revealed that there were no significant differences in SIRAS scores among the presentation orders for any of the 3 video vignettes and that there was a consistent pattern of significant differences among the 3 video vignettes for all 6 presentation orders. The video-vignette main effect was statistically significant, $F_{2,74} = 533.44, P < .001$. Bonferroni pairwise comparisons indicated that participants' SIRAS scores for the highly adherent vignette (mean = 14.00, SD = 1.27) were significantly higher ($P < .001$) than their SIRAS scores for the moderately adherent vignette (mean = 8.93, SD = 1.67) and the minimally adherent vignette (mean = 4.79, SD = 1.93). Participants' SIRAS scores for the moderately adherent vignette were significantly higher ($P < .001$) than those for the minimally adherent vignette. The eta-squared analyses revealed that 94% of the variance in SIRAS scores was accounted for by experimental condition.

The 1-way independent-groups ANOVA performed on SIRAS scores for participants' first vignette viewed (exclusive of the second and third vignettes viewed) was statistically significant, $F_{2,40} = 76.84, P < .01$. Post hoc comparisons calculated using a Fisher least-significant-difference procedure revealed that participants who viewed the highly adherent vignette first (mean = 13.63, SD = 1.63) had SIRAS scores that were significantly higher ($P < .01$) than those of participants who viewed the moderately adherent vignette first (mean = 8.00, SD = 1.71) and those who viewed the minimally adherent vignette first (mean = 5.23, SD = 2.28). Participants who viewed the moderately adherent vignette first had SIRAS scores significantly higher ($P < .01$) than those of participants who viewed the minimally adherent vignette first.

As shown in Table 2, the $r_{WG(j)}$ and RAI values for the highly adherent, moderately adherent, and minimally adherent vignettes ranged from .84 to .95. Along with the RAI value of .87 obtained for the aggregate SIRAS scores across the 3 vignettes, these values represent a high degree of interrater agreement for the SIRAS.

### Table 2 Interrater-Agreement Values for the SIRAS in Study 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>$r_{WG(j)}$</th>
<th>Rater-agreement index</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>.95</td>
<td>.90</td>
</tr>
<tr>
<td>Moderate</td>
<td>.89</td>
<td>.86</td>
</tr>
<tr>
<td>Low</td>
<td>.90</td>
<td>.84</td>
</tr>
<tr>
<td>Aggregate</td>
<td>—</td>
<td>.87</td>
</tr>
</tbody>
</table>
Discussion

The findings for the mixed and 1-way independent-groups ANOVAs provide strong support for the construct validity of the SIRAS in that SIRAS scores increased in a linear fashion with the degree of adherence to clinic-based sport-injury rehabilitation activities. These results converge with the positive correlations between SIRAS scores and attendance at rehabilitation sessions obtained in previous research\(^{13-14}\) in suggesting that the SIRAS taps the construct of adherence to clinic-based sport-injury rehabilitation.

The use of inexperienced practitioners as evaluators has been shown to compromise the reliability of some measurement instruments,\(^{21-25}\) whereas the reliability of other instruments has not been adversely affected when used by inexperienced evaluators.\(^{26-30}\) The results of the current study suggest that the SIRAS falls among the latter assessment indices, because advanced student rehabilitation practitioners were able to generate acceptable levels of interrater agreement.\(^{15-17}\) It is possible, of course, that more experienced practitioners would have even higher levels of interrater agreement.

Although the interrater-agreement values found for the SIRAS are encouraging, they are limited in that they correspond to only a single patient engaging in 3 hypothetical videotaped sessions. Study 2 was conducted to address this limitation. Specifically, to gain a better estimate of the interrater agreement of the SIRAS in a clinical setting, 2 certified athletic trainers completed the SIRAS independently after 4 consecutive appointments with actual sport-injury rehabilitation patients.

Study 2

Method

Participants. Participants were 12 individuals (9 men and 3 women) undergoing outpatient physical therapy after reconstructive surgery performed on the anterior cruciate ligament (ACL) of the knee. The mean age of participants was 29.33 (SD = 11.44) years. Nine participants identified their race as White and 3 reported their race to be Hispanic.

Procedure. Participants completed an informed-consent form as part of a larger study on psychological aspects of rehabilitation after ACL reconstruction. The study was approved by the Institutional Review Board at Springfield College. Two certified athletic trainers, 1 of whom was responsible for the care of participants and 1 of whom was purely an observer, completed the SIRAS independently after each of 4 consecutive rehabilitation appointments for all participants.

Data Analysis. An overall RAI value was calculated for the SIRAS across the 4 appointments observed. The \(r_{\text{WCI}}^{\text{PP}}\) statistic is not appropriate for analyses with fewer than 10 raters\(^{31}\) and was therefore not used in this study.
Results and Discussion

The RAI value across the 4 appointments observed was .94, indicating a high level of interrater agreement for the SIRAS. This finding suggests that SIRAS scores for different rehabilitation practitioners observing the same patient will be substantially similar in a clinical setting.

Comments

Together with previous research,13 the results of study 1 and study 2 provide evidence that the SIRAS is a brief, psychometrically sound measure of adherence during sport-injury rehabilitation sessions. Thus, it appears that SIRAS scores can be obtained from multiple practitioners over the course of a patient’s involvement in clinic-based rehabilitation activities, with the assumption that variations in SIRAS scores are more attributable to changes in the patient’s behavior than to the fact that there are different practitioners completing the SIRAS. The SIRAS can therefore be used with confidence in conjunction with rehabilitation-session attendance data to describe patient adherence to the clinic-based portion of sport-injury rehabilitation.

It is important to note that despite the tendency for patients with sport injuries to display high levels of adherence to clinic-based activities,9,14,16 participants in study 1 used the full range of possible SIRAS scores across the 3 vignettes. This suggests that the SIRAS would be able to indicate when patients depart from the norm of adherence to the clinic-based treatment protocol, which might occur for reasons such as low motivation, poor articulation of the rehabilitation regimen, and problems in the patient–practitioner relationship.

The primary practical implication of the results of study 1 and study 2 is that researchers and practitioners can have increased confidence that the SIRAS taps the adherence construct and that different raters use the SIRAS in a similar way to assess adherence in the same patient. The SIRAS can be used to examine patient adherence to clinic-based rehabilitation activities over the course of treatment, identify patients who might be having difficulty in adhering to the prescribed treatment regimen, and evaluate the correspondence between adherence to clinic-based treatment and rehabilitation outcomes.

Acknowledgments

This article was supported in part by grant number R29 AR44484 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases. Its contents are solely the responsibility of the authors and do not represent the official views of the institute. We gratefully acknowledge the assistance of Thomas Ghazil and the cooperation of Baystate Outpatient Rehabilitation in conducting the investigation.
References


