Title

Rationale for a Parsimonious Measure of Subjective Knee Function after ACL Reconstruction: A Rasch Analysis

Authors

T. Ryan Duckett, MA¹
Christine Fox, PhD¹
Joseph M. Hart, PhD, ATC, FACSM, FNATA²
Grant E. Norte, PhD, AT, ATC, CSCS³

Affiliations

¹ University of Toledo, Department of Educational Research and Measurement, Toledo, OH
² University of Virginia, Kinesiology Department, Sports Medicine Program, Charlottesville, VA
³ University of Toledo, School of Exercise Science and Rehabilitation, Athletic Training Program, Toledo, OH

Corresponding Author

Grant E. Norte, PhD, AT, ATC, CSCS
University of Toledo
School of Exercise and Rehabilitation Sciences
Athletic Training Program
2801 W. Bancroft St.
HH 2505H, Mail Stop 119
Toledo, OH 43606
Office: (419) 530-5305
Fax: (419) 530-2477

Readers should keep in mind that the in-production articles posted in this section may undergo changes in the content and presentation before they appear in forthcoming issues. We recommend regular visits to the site to ensure access to the most current version of the article. Please contact the JAT office (jat@slu.edu) with any questions.
Rationale for a Parsimonious Measure of Subjective Knee Function among Individuals with ACL Reconstruction: A Rasch Analysis

Context: The International Knee Documentation Committee (IKDC) Subjective Knee Evaluation is the most frequently used patient-reported measure of subjective knee function among individuals with ACL reconstruction (ACLR). Yet, limitations with traditional validation approaches leave it unclear whether the IKDC measures knee function as intended. Rasch analysis offers a robust validation approach, which may enhance clinical interpretation of the IKDC. Objective: 1) To assess the psychometric properties, 2) ability to classify health status, and 3) relationships between the IKDC and objective measures of strength and functional performance relative to a newly proposed reduced-item instrument. Design: Cross-sectional. Setting: Laboratory. Patients or other Participants: Seventy-seven individuals with primary, unilateral ACLR (21.9±7.8 years, 6.2±1.0 months post-surgery), and seventy-six age-matched controls. Main Outcome Measure(s): Rasch analysis was used to assess the psychometric properties of the IKDC. Receiver-operator-characteristic curves and logistic regression were used to assess the accuracy of classifying ACLR versus control participants. Correlations (Pearson and Spearman) were used to assess relationships between subjective knee function, quadriceps torque, and single-limb hop performance. Results: Rasch analysis aided the development of a reduced 8-item instrument (IKDC-8), which yielded improved psychometric properties in misfit analysis, percent of variance accounted for by one dimension (IKDC-8=71.5%; IKDC=56.7%), and item reliability. The IKDC was an outstanding diagnostic tool and the IKDC-8 was excellent, correctly classifying 87.2% and 82.7% of cases, respectively. The Hanley-McNeil formula found that there was no significant difference in the areas under the respective ROC curves. Equivalent associations between subjective and objective knee function were observed regardless of instrument used. Conclusions: We observed evidence of enhanced reliability and validity for a parsimonious measure of subjective knee function. The proposed instrument reduces the number of items, increases score interpretability as measuring a single construct, and improves the rating scale functioning, while not significantly diminishing its ability to classify ACLR versus control participants or changing
existing relationships with objective measures of recovery. We suggest the IKDC-8 may enhance clinical
use by reducing administration time, improving the interpretation of the subjective knee function score,
and clarifying functional ability.

**Keywords:** Patient-reported outcome measures, instrument validity; subjective knee function

**Abstract word count:** 337

**Body of manuscript word count:** 4,239

**Key points:**

1. Use of the Rasch measurement model aided in the development of a reduced instrument that
measures subjective knee function.

2. Compared to the IKDC, the reduced IKDC-8 demonstrated greater evidence of measuring
subjective knee function alone and no other constructs, improved function of the response
categories for each item, and placed respondents on a reliable continuum of knee function.

3. The IKDC-8 – with its fewer items and clearer response options – can help clinicians accurately
map respondents’ progress in their rehabilitation following ACLR.
Over the past several decades, the use of patient-reported outcome measures (PROMs) has continued to increase in clinical importance.\(^1\) PROMs provide information on an individual’s experience during the medical treatment and recovery processes, and serves as one of the primary pillars of quality healthcare.\(^2\) The increase in consideration of patient-oriented outcomes corresponds to more emphasis placed on patient involvement and satisfaction with healthcare decision-making.\(^3\) Alongside the prominent role of objective measures, recent meta-analyses have reported the increasing emphasis placed on PROMs in return-to-sport (RTS) decisions for individuals with anterior cruciate ligament reconstruction (ACLR).\(^4\) Yet, in spite of the increased utilization of PROMs, practitioners today echo the limitations noted nearly 25 years ago regarding insufficient reliability standards and the role of PROMs in guiding clinical diagnoses and recommendations.\(^5,6\)

The International Knee Documentation Committee (IKDC) Subjective Knee Evaluation is one of the most prominent PROMs used to assess knee function after ACLR.\(^7\) The IKDC has been included in a number of RTS test batteries in addition to commonly used objective measures of muscle strength and functional performance (i.e. hop tests).\(^8\) Moreover, the relationships between subjective and objective knee function have been widely reported.\(^9,10\) For example, higher subjective knee function has been associated with several critical functional measures among individuals with ACLR, including greater unilateral and symmetric knee extensor torque,\(^9\) better single-leg hop (SLH) performance,\(^11\) quadriceps strength symmetry,\(^12\) and overall successful performance on RTS tests with high sensitivity.\(^13\) These data suggest the IKDC may be a useful clinical indicator of functional recovery after ACLR, and may have utility as a screening tool when incorporated in a battery of tests aimed at assessing overall patient health status.

Previous studies investigating the reliability of the IKDC\(^14,15\) have relied primarily on classical test theory, an approach with long noted limitations in establishing a stable psychometric instrument.\(^1\) Classical test theory techniques used to develop instruments can typically ensure internal scale consistency (i.e. correlation among items), but do not sufficiently address other assumptions of construct validity including specific objectivity and unidimensionality. For a PROM to be useful as a meaningful
outcome in any clinical setting, it is necessary to establish how reliably the responses on the instrument can be summed to produce a valid measure of the intended construct (e.g. subjective knee function). No study has fully evaluated the IKDC with the Rasch measurement model (RMM), which is considered the most robust approach to providing defensible evidence of the scientific rigor of PROM instruments. Specifically, instrument validation using RMM uses probabilistic expectations to estimate the extent to which participant response patterns match the expected development for subjective knee function.

Therefore, our primary aim was to use RMM to assess the psychometric properties of the IKDC as a patient-reported measure of knee function. Based on the interpretation of the psychometric diagnostics and content expertise, we proposed a shortened version of the IKDC. In order to evaluate the construct validity of the resultant IKDC, we investigated its performance on several clinical metrics compared. Accordingly, our secondary aim was to evaluate the ability of the original and resultant IKDC instrument to classify with and without a history of ACLR using identified target values of subjective knee function. Our final aim was to assess the existing relationships between subjective knee function and common objective measures of patient recovery using each instrument.

METHODS

Participants

We performed a retrospective analysis of data collected using a cross-sectional design to investigate the individual IKDC item responses in both individuals after ACLR and a control group of similar age and sex distribution. Eighty individuals with a history of primary, unilateral ACLR were recruited from a university sports medicine clinic, university student body, and local community near the time of physician clearance over a three-year period. Eighty age-matched individuals were enrolled as control participants. Three participants from the ACLR group and four participants from the control group did not complete the IKDC, and were removed from analysis, leaving 77 ACLR and 76 control participants, respectively (Table 1). To be eligible, patients must have undergone surgery with either a
bone-patellar tendon-bone or hamstrings tendon autograft, and could not have had a history of failed ACLR or previous lower extremity surgery. Control participants could not have had a history of lower extremity surgery or injury within the previous 6 months. This study was approved by the Institutional Review Board at the University of (XXX), and all participants provided written and verbal consent.

Procedures

All participants completed the study procedures in the same order as described below. Outcomes were measured bilaterally beginning with the uninvolved limb (ACLR participant) or control limb in a counterbalanced order (control participant). The non-dominant control limb was matched with the involved ACLR limb.

Quadriceps Strength

Knee extension maximal voluntary isometric contraction (MVIC) torque was measured using a stationary dynamometer (Systems 3, Biodex Systems Inc., Shirley, NY) with the knee flexed to 90° as previously described. Participants were instructed to kick out as hard as possible following a brief period of familiarization. Visual feedback of the torque output was provided and verbal encouragement by the investigator was used to ensure maximal effort was given by each participant. The average of 3 MVIC trials was recorded and normalized to body mass (Nm/kg). Limb symmetry indices (LSI) were expressed as a percentage of the average involved limb performance divided by the average uninvolved limb performance.

Single-Leg Hop Performance

The SLH for distance was measured as previously described. Participants performed 3 practice trials followed by 3 test trials. The average of the 3 test trials was recorded, normalized to body height, and expressed as a unit-less ratio. The LSI was expressed as a percentage of average involved limb performance divided by the average uninvolved limb performance.
Subjective Knee Function

The IKDC$^{14}$ – consisting of 19 knee-specific questions related to symptoms, sport activities, and function – was used to quantify subjective knee function. The traditional method was used to calculate scores on a scale from 0-100. Higher scores indicated lower symptoms and higher overall function. Previous research has reported an internal consistency coefficient of 0.92 amongst a sample with a variety of knee injuries and disorders (24% reported ACL injury)$^{14}$.  

Statistical Analyses

Psychometric Properties

Rasch analysis of the IKDC data was conducted using the Rating Scale Model$^{21}$ in WINSTEPS (version 4.5.2, Beaverton, OR).$^{22}$ An initial analysis of the IKDC responses of the 153 participants established baseline values for key measurement diagnostics: 1) rating scale performance, 2) unidimensionality, 3) person and item fit indices with their corresponding reliability statistics. Then, based on interpretation of the diagnostics, changes were made iteratively to the IKDC.$^{23}$

Rating scale performance was assessed by examining the item response categories to determine if they functioned as intended. For the IKDC items, this implied that participants with higher levels of knee function would endorse higher response categories in a linear and consistent manner across the items. Additionally, each response option should have represented a distinct qualitative and quantitative meaning.

Unidimensionality tests were conducted to assess whether the measure of knee function was confounded by items that represent more than 1 construct. Rasch diagnostics indicated the extent to which the variance in participant responses to IKDC items could be explained solely by the knee function of the participant and the difficulty of the items. Greater amounts of variance explained by the principal...
component indicated better adherence to the principle of unidimensionality, with values > 60% typically required.24

Person and item fit indices were explored to examine how well individuals met the expectation of the RMM. To distinguish between levels of subjective knee function, person separation was expected to be > 2.0 and reliability > 0.8, and item separation was expected to be > 3.0 and reliability > 0.9. Residual mean square fit statistics were explored to determine the extent to which the IKDC items fit a unidimensional linear measure of subjective knee function, with an expected value of 1.0 for each item and those > 2.0 interpreted as a possible threat to the measurement system.23

Classification of ACLR versus Control Participants

Receiver-operator-characteristic (ROC) curves were used to assess the diagnostic utility of the IKDC and a reduced IKDC. The positive actual state was identified as history of ACLR with the implicit understanding that lower scores on the respective IKDC instruments indicated greater impairments in subjective knee function. The area under the ROC curve (AUC) was used to determine the ability of each instrument to correctly discriminate ACLR versus control participants, and interpreted as no discrimination (< 0.5), acceptable discrimination (0.7-0.8), excellent discrimination (0.8-0.9), or outstanding discrimination (> 0.90).25 Each ROC curve was visually inspected to identify a target value that maximized the sensitivity and specificity of the instrument using the Youden Index. The sensitivity, specificity, and positive (LR+) and negative (LR-) likelihood ratios were calculated for each identified target value. The Hanley-McNeil formula was used to determine if there was a significant difference in the AUCs for the instruments.26 Binary logistic regression analyses were used to assess the ability of the selected target values to accurately classify ACLR versus control participants. Chi squared analyses were used to compare the proportion of control individuals and those with ACLR that met or failed to meet established target values between instruments.

Relationships Between Subjective and Objective Function
Bivariate correlation coefficients were used to assess the associations between subjective knee function, knee extensor torque, and SLH performance in the ACLR group only using each IKDC instrument. Pearson’s correlations ($r$) were used for normally distributed outcomes and Spearman’s rank-order correlations ($\rho$) were used for non-normally distributed outcomes as determined by the Shapiro-Wilk test. Coefficients were classified as negligible (0-0.29), low (0.3-0.49), moderate (0.5-0.69), high (0.7-0.89), and very high (0.9-1).\textsuperscript{27} The correlations between the functional performance measures and the respective IKDC instruments were tested for any statistical differences by using Fisher’s $r$-to-$z$ transformation and Steiger’s formulae.\textsuperscript{28} The level of statistical significance for all analyses was set \textit{a priori} at $P \leq .05$ and $1 - \beta = 0.80$. All analyses with the exception of the Rasch analysis were preformed using SPSS (v. 25, IBM Corp., Armonk, NY).

RESULTS

Participant Demographics

Group demographics are presented in Table 1. The ACLR group was taller ($P = .040$) and demonstrated lesser knee extension torque, knee extension torque symmetry, SLH distance, SLH symmetry, and IKDC scores (all $P < .001$) compared to the control group. Individuals with ACLR were enrolled at an average of 6.2 months after surgery.

Psychometric Properties

The RMM was used to evaluate the IKDC (Table 2). First, rating scales were changed to address categories with insufficient selection (i.e. fewer than 10 observations of that category across all items) and categories that advanced in a non-monotonic fashion. Next, items with a mean square statistic $> 2$ were removed from the instrument (IKDC items 8, 9c, 9e). As a result of this iterative process, 10 of the 18 items used to score the IKDC were removed. This resulted in an 8-item version of the IKDC (IKDC-8, Table S1) with a raw score range of 0 to 24.
Rasch analysis of the IKDC-8 revealed several domains of psychometric superiority and several domains of non-inferiority when compared to the IKDC (Table 3). Compared to the IKDC, the IKDC-8 rating scales displayed better distribution of observations across rating scale categories for the three items (IKDC items 2, 3, and 10b), that used an 11-point scale in the original IKDC. The mean square statistics for item infit and outfit fell within the accepted range for high stakes tests of 0.8-1.2 across all 8 items. The analysis for unidimensionality in the IKDC-8 instrument accounted for 71.5% of the variance in participant responses to the IKDC-8 items, with no evidence the items were eliciting systematic information beyond that associated with subjective knee function. Figure 1 presents a visual representation of the unidimensional construct of subjective knee function as measured by the IKDC-8. The person separation index for both the IKDC and IKDC-8 indicated that each instrument could separate participants into at least three distinct levels of knee function. The person reliability was moderately high at 0.83, implying the items on the IKDC-8 would produce reliable measures of subjective knee function among another group of similar participants. The order of item difficulty was found to be reproducible, with an extremely high reliability of 0.99.

Classification of ACLR versus Control Participants

Results from each ROC curve analysis are presented in Table 4. The IKDC was an outstanding discriminator of ACLR versus control participants (P<.001). Achieving an IKDC value of 94.9 yielded a high sensitivity and specificity, and correctly classified 87.2% of all cases (P<.001). Similarly, the IKDC-8 was an excellent discriminator of ACLR versus control patients (P<.001). Achieving an IKDC-8 value of 93.8 yielded a high sensitivity with moderate specificity, and correctly classified 82.7% of all cases (P<.001). The proportion of controls who met (IKDC = 83.3%, IKDC-8 = 70.5%; \( \chi^2(1) = 3.61, P = .057 \)), or those with ACLR who failed to meet (IKDC = 91%, IKDC-8 = 94.9%; \( \chi^2(1) = 0.88, P = .349 \)) the identified target values did not statistically differ between instruments. The Hanley-McNeil formula found no significant difference in the areas under the respective ROC curves (\( z = 1.64, P = .10 \)).
Regardless of the instrument used (IKDC, IKDC-8), a higher rating of subjective knee function demonstrated a negligible association with greater knee extension torque \( (r = 0.251, 0.265; P = 0.898) \), and low associations with greater SLH distance \( (r = 0.403, 0.442; P = 0.693) \), and SLH symmetry \( (ρ = 0.332, 0.355; P = 0.826) \) in the ACLR group. Neither instrument was associated with knee extension torque symmetry \( (r = -0.120, -0.072; P = 0.675) \). The magnitude of association did not statistically differ between instruments.

**DISCUSSION**

Our primary aim was to evaluate the IKDC using the Rasch measurement model to assess the suitability of the instrument to measure the construct of subjective knee function. An iterative approach led to the development of the IKDC-8, a reduced version of the IKDC, that demonstrated superiority in rating scale performance, unidimensionality, and person and item fit indices. The observed improvements provide evidence of increased construct validity by means of enhanced instrument performance. Using a Rasch analysis of the IKDC-8 provided additional information clinicians could use to evaluate how well patient reports of knee function fit the underlying theoretical construct (as presented in Figure 1). Despite including fewer items, the IKDC-8 explained a larger amount of variance in how participants responded to individual items when compared to the IKDC. We observed ceiling effects, with 33% \( (n = 57/153; 3 \) ACLR) of the sample achieving a maximum measure (score of 24). Including the controls was necessary to establish a meaningful “healthy” threshold and did not change the ordering of the items or fundamentally alter the interpretation of the construct of knee function. Collectively, the IKDC scores reported in our sample of individuals with ACLR was consistent with published normative data including those with and without a history of knee injury \( (82.1 ± 11.7 \text{ vs. } 82 ± 22) \). Despite the referenced data including a heterogeneous sample (based on injury history), our findings align with previous reports of IKDC scores among individuals 6 months following primary, unilateral ACLR, suggesting our sample was representative of this population.
The Rasch analysis of the IKDC-8 developed greater clarity in meaning of response categories and flagged items in the IKDC that did not fit within the theoretical model of subjective knee function. In developing the IKDC-8, items from the IKDC that did not fit to a linear, progressive measure of knee function were removed. Several items did not fit the unidimensional construct of knee function. The two “symptom” items (IKDC 4 and 6) were identified as a secondary dimension, and further, did not have sufficient observations in all of the rating categories. The four “activity level” items (IKDC 1, 5, 7, and 8) also formed an additional dimension, in line with findings from a previous study. The “prior knee function” item (IKDC 10a) – which is not used in calculating the typical IKDC score – was removed here as well because the construct of interest pertains to current knee function. Three of the “daily activity items” were found to significantly misfit the model (mean square fit statistics > 2). Removing misfitting items and addressing issues in how participants used the rating scales resulted in a substantial increase (15%) in explaining the variability in participants’ responses. In addition to the examination of fit statistics, the combining of ambiguous categories reduced the measurement error associated with the confusion of adjacent, unlabeled categories (e.g. the difference in meaning between a score of 6 vs. 7).

An additional benefit of using the RMM to develop the IKDC-8 was placement of items on an equal interval scale, allowing a clear comparison of relative item difficulty. This feature enabled easier identification of individuals who provided highly unexpected answers that deviated from the established model of subjective knee function. Returning to Figure 1, although the two participants displayed on the right achieved the same IKDC-8 score (raw score 19, converted to traditional 0-100 scale=79.2), their response patterns revealed important differences. The responses from the first participant (ACLR #2) conformed almost perfectly to expectation (e.g. a person with a 79.2 IKDC-8 score would have no difficulty with uniplanar activities, minimal difficulty with multiplanar activities, and occasional, mild pain). The response from the second participant (ACLR #35) displayed a curious case where the individual expressed no difficulty in uniplanar and multiplanar activities, no limitation in daily activities, yet constant, moderate to severe pain. The ability to identify how well responses from individuals with ACLR in rehabilitation contexts fit the expected model of subjective knee function offers clinicians and
researchers greater diagnostic clarity and clinical guidance because deviation from expectation is where intervention is most important. Aberrant response patterns could indicate the participant misunderstood an item, or that their unexpected response reflects a deficiency in their functional status or rehabilitation program.

Additionally, we observed similar findings relative to the ability of the IKDC-8 to accurately classify ACLR versus control participants. For example, the IKDC-8 demonstrated an excellent ability (AUC 0.8-0.9) to discriminate between individuals with ACLR and controls, whereas the IKDC demonstrated an outstanding ability to do so (AUC > 0.90). However, the confidence intervals of the associated AUC values for each instrument overlapped considerably and were not statistically different, suggesting comparable diagnostic utility. The identified target values for each instrument were highly sensitive, yet the IKDC-8 demonstrated lower specificity (0.71 vs. 0.83). Accordingly, the IKDC-8 yielded a stronger negative likelihood ratio, but lesser positive likelihood ratio compared to the IKDC.

From a practical standpoint, these data indicate that scoring $\geq 93.8$ on the IKDC-8 would result in a 3-fold increase in the odds of being classified as a control participant, whereas scoring $< 93.8$ yielded a 14-fold increase in the odds of being classified as having an ACLR. In contrast, a target value of 94.9 on the IKDC yielded a 5-fold increase and 9-fold decrease in the same odds. As expected, the majority of control individuals met the identified target value of each instrument (IKDC=83.3, IKDC-8=70.5), whereas the majority of individuals with ACLR failed to meet it (IKDC=91, IKDC-8=94.9). Although the proportions of pass/fail within each group did not statistically differ, the IKDC-8 instrument appeared to be more difficult for control and ACLR individuals to pass on average, which could contribute to a more conservative assessment of perceived knee function if used in clinical practice. The identified target values appear similar to a previous investigation that reported achieving an IKDC score $\geq 94.8$ was a sensitive indicator of quadriceps strength symmetry at the time of RTS after ACLR. A separate study reported that achieving an IKDC score $\geq 93$ indicated a higher likelihood of readiness to RTS after ACLR. These values are similar to the mean IKDC scores reported in our control group (IKDC=97.7, IKDC-8=94.6), which suggests our identified target values are appropriate indicators of control status.
Better subjective knee function associated with greater quadriceps strength, SLH distance, and SLH symmetry among individuals with ACLR, which agrees with previous reports.\textsuperscript{11,34} However, we did not observe evidence of association between subjective knee function and quadriceps strength symmetry, which has been described.\textsuperscript{12,35} Although relationships between subjective and objective knee function can be influenced by the magnitude of impairment, individuals with ACLR in our sample appeared to demonstrate impairments in quadriceps strength, quadriceps strength symmetry, SLH distance, and SLH symmetry consistent with those in the literature.\textsuperscript{36} The lower self-reported current activity level among the ACLR group could partially explain the observation of objectively measured impairments, although these relationships are unclear in the literature.\textsuperscript{10} However, the observed relationships were negligible to low in magnitude, suggesting that each instrument of knee function may explain unique information about individuals recovering from ACLR that is distinct from objective measures of quadriceps and lower extremity function. This appears to indicate that perceived function differs from measured function, and supports the utility of including each in the RTS decision-making process. According to the observed relationships, the IKDC-8 appears to provide similar information to the IKDC, which offers early support for its clinical utility as an indicator of perceived functional recovery after ACLR.

Our initial findings pose important implications for measuring subjective knee function during the early rehabilitation process as an indicator of recovery. Stakeholders from clinicians and researchers, to patients and insurance providers, continue to invest in and seek improvements in the quality of care when it comes to rehabilitation outcomes.\textsuperscript{1} The role of PROMs will continue to play an important part in evaluating the success of rehabilitation treatments. Implementing an instrument, such as the IKDC-8, that was developed using the RMM may lead to greater precision and clarity in measuring the subjective knee function of individuals with ACLR. Item level analysis can provide stakeholders with more clarity as to what patients are experiencing and guide clinical decision-making. Continual reference to how patient responses fit the established continuum of knee function will facilitate better informed clinical care.

Limitations
Our findings have the potential for clearer clinical assessments, but two limitations need to be taken into consideration. First, the IKDC-8 responses were derived from the original IKDC instrument. Responses from an existing IKDC data set were analyzed and modified under the guidelines of RMM. These results should be interpreted cautiously until the IKDC-8 can be administered and analyzed from an independent sample. As with any instrument, greater precision will come with using the proposed IKDC-8 with a broader range of participants and in a variety of populations extending beyond individuals with primary, unilateral ACLR. Allowing participants to use the newly proposed rating categories could lead to future refinement or change. While all the steps undertaken in the development of the IKDC-8 followed well-established guidelines,\textsuperscript{23,37} future studies that use the IKDC-8 in the proposed format will provide greater information on the reliability and clinical utility of the IKDC-8. The second limitation pertains to the cross-sectional nature of using the same cohort for the exploratory ROC and correlational analyses. The statistical methods employed in this study to compare the results produced by the respective instruments do take into account the commonalities in the standard errors introduced by the same sample. Even with that, future studies using separate cohorts can work toward a more confirmatory analysis.

Application of the RMM led to the proposal of a reduced instrument of subjective knee function, the IKDC-8. The IKDC-8 exhibited superior person fit, item fit, and dimensionality, as well as noninferior reliability statistics relative to the IKDC. In addition, the IKDC-8 yielded a reduction in the number of response items, which may serve to minimize administration time, increase the interpretation of the resulting score by analyzing response patterns, and improve the rating scale functioning by clarifying response categories. These changes reflected in the IKDC-8 do not appear to diminish the instrument’s ability to classify ACLR versus control participants and did not change existing relationships with objective measures of recovery. Implementation of the IKDC-8 may be useful to further clarify patient recovery along a continuum of lower-to-higher levels of knee function following ACLR.
1. Cano SJ, Pendrill LR, Melin J, Fisher WP. Towards consensus measurement standards for patient-centered outcomes. *Measurement (02632241).* 2019;141:62-69.

2. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open.* 2013;3(1):1-18.

3. Wright RW. Knee injury outcomes measures. *J Am Acad Orthop Surg.* 2009;17(1):31-39.

4. Agarwalla A, Puzzitiello RN, Liu JN, et al. Timeline for Maximal Subjective Outcome Improvement After Anterior Cruciate Ligament Reconstruction. *Am J Sports Med.* 2019;47(10):2501-2509.

5. Jacob J, Edbrooke-Childs J, Law D, Wolpert M. Measuring what matters to patients: Using goal content to inform measure choice and development. *Clin Child Psychol Psychiatry.* 2017;22(2):170-186.

6. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: Are available health status surveys adequate? *Quality of Life Research: Int J Qual of Life Aspects of Treat, Care & Rehabil.* 1995;4(4):293-307.

7. Glogovac G, Schumaier AP, Grawe BM. Return to Sport Following Revision Anterior Cruciate Ligament Reconstruction in Athletes: A Systematic Review. *Arthroscopy.* 2019;35(7):2222-2230.

8. Werner JL, Burland JP, Mattacola CG, Toonstra J, English RA, Howard JS. Decision to return to sport participation after anterior cruciate ligament reconstruction, part II: self-reported and functional performance outcomes. *J Athl Train.* 2018;53(5):464-474.

9. Pietrosimone B, Lepley AS, Harkey MS, et al. Quadriceps Strength Predicts Self-reported Function Post-ACL Reconstruction. *Med Sci Sports Exerc.* 2016;48(9):1671-1677.

10. Norte GE, Solaas H, Saliba SA, Goetschius J, Slater LV, Hart JM. The relationships between kinesiophobia and clinical outcomes after ACL reconstruction differ by self-reported physical activity engagement. *Phys Ther Sport.* 2019;40:1-9.

11. Menzer H, Slater LV, Diduch D, et al. The Utility of Objective Strength and Functional Performance to Predict Subjective Outcomes After Anterior Cruciate Ligament Reconstruction. *Orthop J Sports Med.* 2017;5(12):232597117744758.

12. Zwolski C, Schmitt LC, Quatman-Yates C, Thomas S, Hewett TE, Paterno MV. The influence of quadriceps strength asymmetry on patient-reported function at time of return to sport after anterior cruciate ligament reconstruction. *Am J Sports Med.* 2015;43(9):2242-2249.

13. Logerstedt D, Dj Stasi S, Grindem H, et al. Self-reported knee function can identify athletes who fail return-to-activity criteria up to 1 year after anterior cruciate ligament reconstruction: a delaware-oslo ACL cohort study. *J Orthop Sports Phys Ther.* 2014;44(12):914-923.

14. Irrgang JJ, Anderson AF, Boland AL, et al. Development and Validation of the International Knee Documentation Committee Subjective Knee Form. *Am J Sports Med.* 2001;29(5):600-613.

15. Grevnerts H, Terwee C, Kvist J, Grevnerts HT, Terwee CB. The measurement properties of the IKDC-subjective knee form. *Knee Surg Sports Traumatol Arthrosc.* 2015;23(12):3698-3706.

16. Cano SJ, Hobart JC. The problem with health measurement. *Patient Preference And Adherence.* 2011;5:279-290.

17. Petrillo J, Cano SJ, McLeod LD, Coon CD. Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value Health.* 2015;18(1):25-34.

18. McClimans L, Browne J, Cano S. Clinical outcome measurement: Models, theory, psychometrics and practice. *Stud Hist Philos Sci.* 2017;65:67-73.

19. Norte GE, Hertel JN, Saliba SA, Diduch DR, Hart JM. Quadriceps Neuromuscular Function in ACL Reconstructed Patients with and without Knee Osteoarthritis: A Cross-Sectional Study. *J Athl Train.* 2018.
20. DiFabio M, Slater LV, Norte G, Goetschius J, Hart JM, Hertel J. Relationships of Functional Tests Following ACL Reconstruction: Exploratory Factor Analyses of the Lower Extremity Assessment Protocol. J Sport Rehabil. 2017:1-24.

21. Andrich D. A rating formulation for ordered response categories. Psychometrika. 1978;43(4):561-573.

22. Linacre JM. WINSTEPS Rasch measurement computer program. 2006.

23. Wright BD, Stone MH. Best test design. Chicago: Mesa Press; 1979.

24. Smith E. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. J Appl Meas. 2002;3:205-231.

25. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied logistic regression. Vol 398: John Wiley & Sons; 2013.

26. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. J Radiology. 1983;148(3):839-843.

27. Mukaka M. Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. Malawi Med J. 2012.

28. Steiger JH. Tests for comparing elements of a correlation matrix. J Psychological bulletin. 1980;87(2):245.

29. Anderson AF, Irrgang JJ, Kocher MS, Mann BJ, Harrast JJ. The International Knee Documentation Committee subjective knee evaluation form: normative data. Am J Sports Med. 2006;34(1):128-135.

30. Norte GE, Goetschius JW, Slater LV, Hart JM. Influence of Patient Demographics and Surgical Characteristics on Pass Rates of Return-to-Activity Tests in Anterior Cruciate Ligament-Reconstructed Patients Before Physician Clearance. Clinical J Sport Medicine: Official J of the Canadian Academy of Sport Med. 2020.

31. Bodkin S, Hertel J, Bruce A, et al. Patient Function in Serial Assessments Throughout the Post-ACL Reconstruction Progression. J Phys Ther in Sport. 2020.

32. Higgins LD, Taylor MK, Park D, et al. Reliability and validity of the International Knee Documentation Committee Subjective Knee Form. Joint Bone Spine. 2007;74(6):594-599.

33. Lentz TA, Zeppieri G, Jr., Tillman SM, et al. Return to preinjury sports participation following anterior cruciate ligament reconstruction: contributions of demographic, knee impairment, and self-report measures. J Orthop Sports Phys Ther. 2012;42(11):893-901.

34. Lepley AS, Pietrosimone B, Cormier ML. Quadriceps Function, Knee Pain, and Self-Reported Outcomes in Patients With Anterior Cruciate Ligament Reconstruction. J Athl Train. 2018;53(4):337-346.

35. Hunnicutt JL, McLeod MM, Slone HS, Gregory CM. Quadriceps Neuromuscular and Physical Function After Anterior Cruciate Ligament Reconstruction. J Athl Train. 2020;55(3):238-245.

36. Lepley LK. Deficits in Quadriceps Strength and Patient-Oriented Outcomes at Return to Activity After ACL Reconstruction: A Review of the Current Literature. Sports Health. 2015;7(3):231-238.

37. Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen, Denmark: Danish Institute for Educational Research; 1960.
**Legend to Figure**

**Figure 1.** Map of subjective knee function as measured by IKDC-8

**Left Column:** Items are ordered from most difficult at the top to easiest at the bottom, on an equal interval scale. Item difficulties for each are in parentheses. Global description of item types are bracketed to the left of the column

**Central Axis:** Normalized range of subjective function, lower scores indicating lower subjective knee function

**Right Column:** Responses from two average ACLR participants with the same total IKDC-8 score (79.2; raw = 19), achieved by different response patterns to the items
|                | ACLR (n = 77) | Healthy (n = 76) | p value |
|----------------|---------------|------------------|---------|
| Sex *a         | 37 F, 40 M    | 42 F, 34 M       | .372    |
| Age (years)    | 21.9 ± 7.8    | 22.0 ± 4.2       | .947    |
|                | Range: 13-47  | Range: 15-46     |         |
| Height (cm)    | 166.1 ± 38.5  | 155.2 ± 24.7     | .040*   |
| Mass (kg)      | 68.1 ± 3.7    | 67.9 ± 4.4       | .797    |
| IKDC score (0-100%) | 82.1 ± 11.7 | 97.7 ± 3.8       | < .001* |
| IKDC-8 score (0-100%) | 69.9 ± 18.4 | 95.0 ± 9.2       | < .001* |
| Knee extension torque (Nm/kg) | 1.57 ± 0.59 | 2.10 ± 0.62     | < .001* |
| Knee extension torque LSI | 69.0% ± 22.0% | 97.0% ± 20.0% | < .001* |
| Single hop for distance | 0.68 ± 0.20 | 0.78 ± 0.16     | < .001* |
| Single hop for distance LSI | 88.0% ± 12.0% | 99.0% ± 6.0% | < .001* |
| Tegner score (pre-injury) | 8.7 ± 1.5 | N/A             | < .001*b |
| Tegner score (current) | 6.1 ± 2.0 | N/A             | < .001* |
| Time from surgery (months) | 6.2 ± 1.0 | N/A             | N/A     |
| Graft type, n (%) | 50 BPTB (64.9) | N/A            | N/A     |
|                | 27 HT (35.1)  | N/A              | N/A     |

Abbreviations: IKDC, international knee documentation committee subjective knee evaluation; ACLR, anterior cruciate ligament reconstruction; LSI, limb symmetry index; N/A, not applicable; BPTB, bone-patellar tendon-bone; HT, hamstrings tendon

*a Chi-squared test

*b Comparison of pre-injury and current ACLR Tegner scores

* Statistically significant at p ≤ .05
Table 2. Summary of changes to IKDC items and rating scales

| Content                                    | Original IKDC | IKDC-8 |
|--------------------------------------------|---------------|--------|
| Activity level without knee pain           | Item 1        | Categories 5-point | DELETED 4-point |
| Frequency of knee pain                     | Item 2        | Categories 11-point | 1 4-point |
| Severity of knee pain                      | Item 3        | Categories 11-point | 2 4-point |
| Stiffness of knee                          | Item 4        | Categories 5-point | DELETED 4-point |
| Activity level without knee swelling       | Item 5        | Categories 5-point | DELETED 4-point |
| Locking or catching of knee                | Item 6        | Categories 2-point | DELETED 4-point |
| Activity level without knee giving way     | Item 7        | Categories 5-point | DELETED 4-point |
| Activity level able to participate in      | Item 8        | Categories 5-point | DELETED 4-point |
| Ability to go up stairs                    | Item 9a       | Categories 5-point | DELETED 4-point |
| Ability to go down stairs                  | Item 9b       | Categories 5-point | 3 4-point |
| Ability to kneel on front of knee          | Item 9c       | Categories 5-point | DELETED 4-point |
| Ability to squat                           | Item 9d       | Categories 5-point | DELETED 4-point |
| Ability to sit with knee bent              | Item 9e       | Categories 5-point | DELETED 4-point |
| Ability to rise from chair                 | Item 9f       | Categories 5-point | 4 4-point |
| Ability to run straight ahead              | Item 9g       | Categories 5-point | 5 4-point |
| Ability to jump and land on involved leg   | Item 9h       | Categories 5-point | 6 4-point |
| Ability to stop and start quickly          | Item 9i       | Categories 5-point | 7 4-point |
| Knee function prior to knee injury         | Item 10a      | Categories 11-point | DELETED 4-point |
| Current knee function                      | Item 10b      | Categories 11-point | 8 4-point |

Abbreviations: IKDC, international knee documentation committee subjective knee evaluation
Table 3. Psychometric properties of the IKDC and IKDC-8 from Rasch analyses

| Measurement category          | Description of diagnostic                                      | IKDC | IKDC-8 | Impact of change to IKDC-8                                                                 |
|-------------------------------|-----------------------------------------------------------------|------|--------|------------------------------------------------------------------------------------------|
| Rating scale performance      | Items with disordered category thresholds                      | 3    | 0      | Greater clarity in meaning of response categories                                         |
|                               | Items with OUTFIT > 2.0                                         | 3    | 0      | Elimination of items with confounding meaning                                             |
| Unidimensionality             | Variance accounted for in the measures                          | 56.7%| 71.5%  | Better understanding of knee function with fewer items                                   |
|                               | Contrasts with eigenvalues > 2.0                                | 2    | 0      | Reassurance that measuring only subjective knee function                                  |
| Person fit indices            | Person separation statistic                                     | 2.47 | 2.23   | Similar ability to differentiate participants                                             |
|                               | Reliability in estimation of person measures                   | 0.86 | 0.83   | Similar reliability                                                                       |
| Item fit indices              | Item separation statistic                                       | 5.02 | 9.79   | Increased precision in item difficulty                                                    |
|                               | Reliability in estimation of item measures                     | 0.96 | 0.99   | Similar item placement on the measure                                                    |
Table 4. Established target values and diagnostic utility for IKDC and IKDC-8 to maximize discrimination between ACLR and control participants

|                          | IKDC \(^a\) | IKDC-8 \(^a\) |
|--------------------------|-------------|---------------|
| Target value             | 94.9        | 93.8          |
| Area under the curve (95% CI) | 0.94 (0.90-0.97) \(^*\) | 0.90 (0.85-0.95) \(^*\) |
| Sensitivity              | 0.91        | 0.95          |
| Specificity              | 0.83        | 0.71          |
| Positive likelihood ratio | 5.46        | 3.22          |
| Negative likelihood ratio | 0.11        | 0.07          |

Abbreviations: IKDC, international knee documentation committee subjective knee evaluation; ACLR: anterior cruciate ligament reconstruction; CI, confidence interval

\(^a\) Data obtained or calculated from receiver operator characteristic curve analyses. Smaller values indicate stronger evidence for the positive actual state (ACLR).

\(^*\) Statistically significant at \(p \leq .05\)
1a) Rise from chair (31.6)
1b) Run straight ahead (37.1)
1c) Go down stairs (39.2)
1d) Stop and start quickly (48.9)
1e) Jump on involved leg (52.5)
2) Function of daily activities (65.5)
3) Frequency of pain (70.6)
4) Severity of pain (75.4)

Uniplanar movement activities

Multiplanar movement activities

Functional Limitations and Symptoms

Easier Item

More difficult item

ACLR #2

ACLR #35
Appendix

Table S1. Revised IKDC-8 Instrument

1. How difficult do you find the following tasks given your knee function:

|   | Not difficult at all (3) | Minimally difficult (2) | Moderately difficult (1) | Extremely difficult (0) |
|---|--------------------------|-------------------------|--------------------------|-------------------------|
| a. | Rise from a chair        |                         |                          |                         |
| b. | Run straight ahead       |                         |                          |                         |
| c. | Go down stairs           |                         |                          |                         |
| d. | Stop and start quickly   |                         |                          |                         |
| e. | Jump and land on involved leg |                   |                          |                         |

2. How would you rate the function of your knee on a scale of 0 to 3 with 3 being normal, excellent function and 0 being the inability to perform any of your usual daily activities which may include sports?
   - No limitation in daily activities (3)
   - Could perform most of my daily activities (2)
   - Could perform a limited number of my daily activities (1)
   - Couldn’t perform daily activities (0)

3. During the past 4 weeks, or since your injury, how often have you had pain?
   - Never (3)
   - Occasionally (2)
   - Frequently (1)
   - Constant (0)

4. If you have pain, how severe is it?
   - No pain (3)
   - Slight to mild (2)
   - Moderate to severe (1)
   - Worst pain imaginable (0)
Scoring Instructions for the IKDC-8

While the Rasch-Andrich rating scale model was used to compute the measures presented in the current study, a very strong correlation was found between the scores computed via the Rasch analysis and using the raw summed scores: \( r(153) = .97, p < .001 \). Although conducting a Rasch analysis on the scores provides additional diagnostic information, we have taken efforts to reproduce some of those benefits to the extent possible while using raw summed scores.

Similar to the original IKDC, the IKDC-8 uses an ordinal rating scale where a score of 0 represents the lowest level of function on a given item. All items are used to calculate the IKDC-8 measure using the following method:

1. Sum up scores provided by the participant to each item.
2. Transform the raw score to a traditional 0 to 100 scale as follows:

\[
IKDC\ measure = \left[ \frac{\text{Raw Score}}{24} \right] \times 100
\]

One important aspect discussed above is going beyond the summed score and understanding how responses to specific items inform that score. The items in the IKDC-8 are ordered from easiest to most difficult. This implies that participants with higher levels of knee function would endorse higher response categories in a linear and consistent manner across the items. See the explanation of Figure 1 in the Discussion for more details.