Journal of Athletic Training 2021;56(9):1037–1041
doi: 10.4085/1062-6050-502-19

www.natajournals.org

Invariance Testing of the Disablement in the Physically Active Scale Short Form-10

Madeline P. Casanova, PhD; Russell T. Baker, PhD, DAT, AT, CMP; Michael A. Pickering, PhD; Jayme G. Baker, DPT, PT, AT

Department of Movement Sciences, University of Idaho, Moscow

Context: Psychometrically sound instruments are needed to accurately track the effectiveness of treatment and assess the quality of patient care. The Disablement in the Physically Active (DPA) scale Short Form-10 (SF-10) was developed as a more parsimonious version of the Disablement in the Physically Active scale to assess disablement in the physically active. Psychometric assessment of the DPA SF-10 has not been completed; specifically, the scale properties must be assessed among a sample of individuals who respond only to the 10-item scale at multiple time points.

Objective: To assess the psychometric properties of the DPA SF-10 using confirmatory factor analysis and invariance procedures across multiple time points.

Main Outcome Measure(s): Confirmatory factor analyses and longitudinal invariance tests were conducted.

Results: The DPA SF-10 met contemporary fit index recommendations and demonstrated longitudinal invariance; however, localized fit concerns suggest further modification is needed.

Conclusions: Adoption of the DPA SF-10 into widespread clinical practice and research is not recommended until further psychometric testing and scale modification are performed.

Key Words: confirmatory factor analysis, psychometric analysis, longitudinal measurement

Key Points

- Confirmatory factor analysis indicated the Disablement in the Physically Active (DPA) scale Short Form-10 (SF-10) met some goodness-of-fit indices and invariant criteria. However, localized concerns previously identified for the scale were confirmed among a sample of individuals who answered only the items included in the DPA SF-10.
- Although the DPA SF-10 has improved measurement properties and a reduced response burden compared with those of the original DPA scale, adoption into clinical practice is not recommended until further psychometric testing and model alteration confirm the most psychometrically sound instrument.

The Disablement in the Physically Active (DPA) scale is a patient-reported outcome measure used to assess and track patient perceptions of the injury process using the disablement framework. Although a number of scales have been created to assess disablement, the 16-item DPA scale is unique because it was developed specifically for physically active individuals who have musculoskeletal injuries.[1] Clinicians who work with physically active populations may find this instrument particularly valuable because other instruments do not adequately assess the disablement constructs in these populations.[2] However, researchers[2] have since identified concerns with the psychometric properties of the DPA scale.

The model fit concerns led to scale modifications that produced a shortened version of the scale: the DPA scale Short Form-10 (DPA SF-10). The DPA SF-10 uses 10 of the original items from and is a more parsimonious version of the 16-item DPA scale. The structure of the DPA SF-10 closely resembles that of the DPA scale, with a second-order disablement construct (DIS) that includes the first-order factors impairment (IMP) and functional limitations (FL); the DIS covaries with a first-order construct quality of life (QOL). Initial findings indicated the DPA SF-10 addressed some of the model misspecification and multicollinearity concerns found in the original 16-item DPA scale and also resulted in improved model fit, reduced response burden, and increased administration efficiency.[2] Yet the potential multicollinearity among the constructs was not fully resolved with the DPA SF-10 despite the positive initial psychometric findings.[2] In addition, the sample tested in the development of the DPA SF-10 comprised only individuals who had responded to all 16-items of the original DPA scale. Therefore, the 6 items removed from the scale may have influenced participant responses to the remaining 10 items.[2]

Further analysis is needed to address the remaining concerns with the psychometric properties of the scale.[2] Thus, the primary purpose of our study was to evaluate the psychometric properties of the DPA SF-10 via confirmatory factor analysis (CFA) among a sample of individuals who responded only to the 10 items contained in the short form. In addition, because the DPA SF-10 is intended as a measure for tracking and evaluating patients, it is important to ensure that the underlying constructs can be adequately measured and compared across repeated measures (ie, if the instrument is invariant across time, clinicians will be able to interpret score changes across treatment sessions).[3] Hence,

the secondary purpose of our study was to assess invariance (ie, equal factor variances, equal factor covariance, and equal means) of the DPA SF-10 across repeated measures.

METHODS

Participants

Participants were recruited from 22 athletic training clinics and 2 outpatient rehabilitation clinics across the United States. The individuals were physically active, and those with chronic pain were excluded (Table 1).[1,2,4] All participants provided written informed consent, and the study was approved by our university's institutional review board.

Instrumentation

Participants completed paper versions of the DPA SF-10 and a demographic questionnaire. The DPA SF-10 was completed at 3 visits; the time of completion depended on the injury type (Table 1) and was based on the methods used to create the DPA scale.[1] Recruits who were healthy or had sustained an acute or subacute injury completed the DPA SF-10 at the initial visit (time point 1), 3 to 5 days after the initial visit (time point 2), and 6 to 10 days after the initial visit (time point 3). Those who had a persistent injury completed the DPA SF-10 at the initial visit (time point 1), 7 to 10 days after the initial visit (time point 2), and 3 weeks after the initial visit or at their discharge visit if it occurred before this time (time point 3).

The deidentified participant demographic data collected included injury category, athletic status, age, sex, sport, general injury location, specific injury location, and type of injury. The collected DPA SF-10 and demographic data were input into Qualtrics (Qualtrics XM) by the collecting athletic trainer.

Data Analysis

Data Cleaning. Data were downloaded from Qualtrics for analyses using SPSS (version 25.0; IBM Corp) and Analysis of Moment Structure (version 25.0; IBM Corp). Individuals who did not respond to the DPA SF-10 items at all 3 time points and those who did not respond to at least 90% (ie, 9 of 10) of the DPA SF-10 items were removed from the data set. Normality was assessed using histograms, skewness, and kurtosis values, and multivariate outliers were evaluated using a Mahalanobis distance set at $P < .01$.[3]

Scale Structure. Data from the full sample were included in a confirmatory factor analysis (CFA) on the DPA SF-10 by time point using the Analysis of Moment Structure software to assess the scale structure and verify the underlying constructs of DIS and QOL. A hierarchical CFA was performed, grouping the disablement variables IMP and FL to create a second-order variable, DIS, that would then covary with the first-order variable QOL. To assess correlations among the 3 latent variables (ie, IMP, FL, QOL), we conducted an additional first-order CFA. Model fit indices were evaluated on the basis of a priori values; the comparative fit index (CFI; $\geq$0.95), Tucker-Lewis Index (TLI; $\geq$0.95), root mean square error of approximation (RMSEA; $\leq$0.06), and Bollen Incremental Fit Index (IFI; $\geq$0.95) were computed to assess overall goodness of fit.[3,5,6] The likelihood ratio statistic (χ^2) was also assessed though it was not used as a primary assessment of model fit.

Table 1. Inclusion and Exclusion Criteria for Participant Activity Level, Injury, and Pain Type and Participant Athletic Status Stratification Definitions[a]

Criterion or Status	Definition
Inclusion criteria[1,2]	
Physically active and	An individual who engages in athletic, recreational, or occupational activities that require physical skills and who uses strength, power, endurance, speed, flexibility, range of motion, or agility at least 3 d/wk
Healthy or	Free from musculoskeletal injury and fully able to participate in sport or activity
Acute injury or	A musculoskeletal injury that precludes full participation in sport or activity for at least 2 consecutive d (study participation occurred within 72 h postinjury)
Subacute injury or	A musculoskeletal injury that precludes full participation in sport or activity for at least 2 consecutive d (study participation occurred within 3 d–1 mo postinjury)
Persistent injury	A musculoskeletal injury that has been symptomatic for at least 1 mo (study participation occurred at least 1 mo postinjury)
Exclusion criteria[1,2]	
Chronic pain	Pain that consistently does not get any better with routine treatment or nonnarcotic medication
Participant athletic status stratification[2,4]	
Competitive athlete	A participant who engages in a sport activity that requires at least 1 preparticipation examination, regular attendance at scheduled practices or conditioning sessions, and a coach who leads practices or competitions
Recreational athlete	A participant who meets the criteria for physical activity and participates in sport but does not meet the criteria for competitive status
Occupational athlete	A participant who meets the criteria for physical activity for occupation or recreation but does not meet the criteria for competitive or recreational athlete

[a] Adapted with permission.[1,2]

Longitudinal Invariance Testing. To ensure that individuals were interpreting the items and meanings of the items similarly across time (ie, across time points 1, 2, and 3), we performed invariance testing to assess measurement and structural invariance of the DPA SF-10 across the 3 time points. Longitudinal testing involved assessing the measurement parameters (ie, equal form, loadings, and intercepts), and if the model passed the invariant criteria, substantive parameters were also assessed (ie, variances, covariances, and means). The same criteria used for the CFAs were used to assess model fit.[3,7] Invariance was evaluated using the CFI difference test (CFI_{DIFF}), with a cutoff of $<$0.01 and the χ^2 difference test (χ^2_{DIFF}), with a cutoff of $P = .01$.[3,7] Given the sensitivity of the χ^2_{DIFF} test to sample size, the CFI_{DIFF} test held greater weight in decisions regarding invariance testing of model fit.[7]

Table 2. Demographic Data (N = 279)

Characteristic	Mean ± SD
Age, y	23.6 ± 8.9
	No. (%)[a]
Sex	
Male	129 (46.2)
Female	149 (53.4)
Other	1 (0.4)
Activity level	
Low	49 (17.6)
Medium	71 (25.4)
High	157 (56.3)
Unknown	2 (0.7)
Occupation	
Competitive athlete	169 (60.6)
Recreational athlete	27 (9.7)
Occupational athlete	37 (13.3)
Activities of daily living	44 (15.8)
Unknown	2 (0.7)
Injury category	
Persistent injury	95 (34.1)
Acute injury	66 (23.7)
Subacute injury	61 (21.9)
Healthy	57 (20.4)
Ethnicity	
Caucasian/White	201 (72.0)
African American	42 (15.1)
Hispanic	23 (8.2)
Asian/Pacific Islander	8 (2.9)
Other	5 (1.8)

[a] Percentages in each category were rounded and may not total 100%.

RESULTS

A total of 315 individuals participated in the study; 13 people did not complete the DPA SF-10 items at all 3 time points and were removed from the data set. A total of 23 participants were identified as univariate and multivariate outliers and were removed from the analysis. When examining distributional properties (ie, skewness and kurtosis values) of the sample, we found that only 1 item had a nonnormal distribution (ie, kurtosis $\geq$ 3.4 but $\leq$ 4.00). Transformation of the data was considered; however, the data transformation was not completed because it was unlikely to lead to substantial differences in results or interpretations.[8] The sample comprised 279 individuals (age = 23.6 ± 8.9 years; age range, 18–78 years; median age = 21.0 years), with females accounting for 53.4% (n = 149; Table 2). A total of 57 (20.4%) individuals were healthy, and 222 (79.6%) were injured; injuries were classified as persistent (n = 95, 34.1%), acute (n = 66, 23.7%), or subacute (n = 61, 21.9%). Most individuals were competitive athletes (n = 169, 60.6%) and had a high level of activity (n = 157, 56.3%).

Scale Structure

Time Point 1. The CFA of the DPA SF-10 model at time 1 indicated acceptable fit to the sample data (Figure). All factor loadings were different (P values $\leq$ .001), and goodness-of-fit indices met recommended values for CFI (0.978), TLI (0.969), and IFI (0.978) but slightly exceeded the RMSEA (0.073) cutoff. Five path coefficients exceeded 0.91. Inspection of the first-order model indicated moderate to high correlations between first-order latent variables: IMP and FL ($r = 0.82$, $P < .001$), IMP and QOL ($r = 0.46$, $P < .001$), and FL and QOL ($r = 0.38$, $P < .001$). Modification indices demonstrated meaningful cross-loadings.

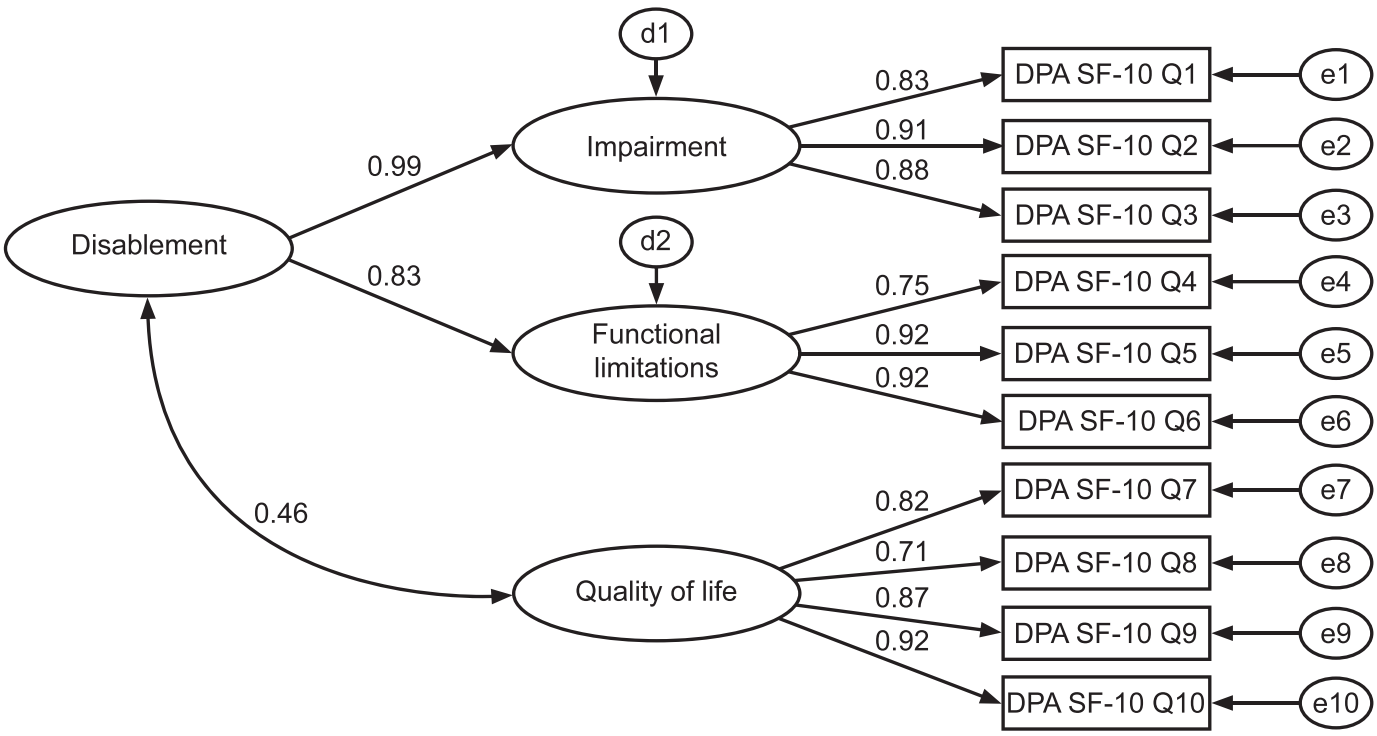

Figure. Confirmatory factor analysis for the Disablement in the Physically Active Scale Short Form-10 (DPA SF-10) for time point 1. Abbreviations: Q, question; e, error term. Adapted with permission.[1,2]

Table 3. Goodness-of-Fit Indices for Measurement Invariance Analyses Across Repeated Measures

	χ^2	Degrees of Freedom	χ^2 Difference Test (Degrees of Freedom Difference)	Comparative Fit Index	Comparative Fit Index Difference Test	Tucker-Lewis Index	Root Mean Square Error of Approximation
Measure 1 (N = 279)	79.06	32	[a]	0.978	[a]	0.969	0.073
Measure 2 (N = 279)	92.17	32	[a]	0.970	[a]	0.958	0.082
Measure 3 (N = 279)	86.12	32	[a]	0.973	[a]	0.962	0.078
Model A (configural or equal form)	585.637	339	[a]	0.970	[a]	0.970	0.051
Model B (metric or equal loadings)	597.347	352	11.71 (13)	0.970	<0.001	0.970	0.050
Model C (equal factor variances)	766.771	358	181.13 (19)[b]	0.950	0.020[b]	0.950	0.064
Model D (scalar or equal indicator intercepts)	616.932	366	31.30 (27)	0.969	0.001	0.969	0.050
Model E (equal latent means)	828.263	372	242.63 (33)[b]	0.944	0.026[b]	0.934	0.066

[a] Indicates the value is not calculated in this step.
[b] Indicates model did not pass invariance criteria.

Time Point 2. The CFA of the DPA SF-10 model at time 2 indicated acceptable fit to the sample data. All factor loadings were different (P values $\leq .001$), and goodness-of-fit indices met recommended values for CFI (0.970), TLI (0.958), and IFI (0.971) but exceeded the recommended RMSEA value (0.082). Five path coefficients exceeded 0.90. Inspection of the first-order model revealed moderate to high correlations between first-order latent variables: IMP and FL ($r = 0.89$, $P < .001$), IMP and QOL ($r = 0.48$, $P < .001$), and FL and QOL ($r = 0.44$, $P < .001$). Modification indices demonstrated meaningful cross-loadings, as well as specification among a number of error-term covariances.

Time Point 3. The CFA of the DPA SF-10 model at time 3 indicated acceptable fit to the sample data. All factor loadings were different (P values $\leq .001$), and goodness-of-fit indices met recommended values for CFI (0.973), TLI (0.962), and IFI (0.973) but exceeded the recommended RMSEA value (0.078). Four path coefficients were ≥ 0.90, with 1 exceeding 1.0. Inspection of the first-order model demonstrated moderate to high correlations between first-order latent variables: IMP and FL ($r = 0.94$, $P < .001$), IMP and QOL ($r = 0.44$, $P < .001$), and FL and QOL ($r = 0.36$, $P < .001$). Similar to time points 1 and 2, modification indices demonstrated meaningful cross-loadings, as well as specification between error-term covariances.

Longitudinal Invariance Testing

The full sample (N = 279) was included in longitudinal invariance testing. The initial model (configural) demonstrated acceptable model fit (CFI = 0.970, $\chi^2_{339} = 585.637$, RMSEA = 0.051; Table 3), indicating equal form across repeated measures. The metric model (ie, equal loadings) passed both the CFI_{DIFF} and χ^2_{DIFF} tests. Because the model satisfied the invariance criteria, an equal latent variance model was conducted. Both CFI_{DIFF} and χ^2_{DIFF} noninvariant criteria were exceeded, indicating the variances were not equal across repeated measures. Variance ranged from 0.63 to 1.04 for the first-order latent variable IMP, from 0.54 to 0.95 for FL, and from 0.15 to 0.66 for QOL. At each time point, variance for all latent variables decreased, indicating that the DPA SF-10 scale appeared to capture less variability across the sample with repeated administrations.

The scalar model (ie, equal loadings and intercepts) passed both the CFI_{DIFF} and the χ^2_{DIFF} tests. The invariant scalar model results allowed comparison of reported levels of the latent variables across repeated measures. All means for first-order latent variables IMP, FL, and QOL were different across repeated measures. The means of the latent variables decreased, indicating that across repeated measures, participants reported less IMP and FL and better QOL.

DISCUSSION

The psychometric properties of the DPA SF-10 were assessed at 3 time points. Overall, the DPA SF-10 met many model fit recommendations, which indicated it may be a suitable tool for measuring patient disablement. Although overall goodness of fit (ie, model fit indices) was met, a number of localized areas of strain in the solution indicated potential ill fit (eg, high path coefficients).[7] For example, 1 local area of concern involved the high (0.99, 0.99, >1.00, respectively) path coefficients from DIS to IMP found across repeated measures (ie, time points 1–3). Standardized path loadings of this magnitude typically indicate the presence of multicollinearity in the data and model misspecification.[7] In addition, the high correlations (≥ 0.82) between IMP and FL suggested that multicollinearity was present and that IMP and FL either were not measuring unique constructs or were being interpreted similarly by respondents. Last, model misspecification may be evidenced by the large modification indices; the indices revealed meaningful cross-loadings and suggested specification between error-term covariances.[7] The presence of multicollinearity within factors, as well as the potential model misspecification, has also been noted in previous research.[2] Our findings, when combined with those reported earlier, support the need to further remove items or reduce factors to produce a parsimonious and psychometrically sound scale.[3,7]

Longitudinal invariance was conducted to determine measurement and structural invariance of the DPA SF-10 across multiple measures. An invariant solution implies that participants, across repeated measures, interpret the questions and underlying latent factors (IMP, FL, QOL) in the same way.[3,7] Therefore, when a model is invariant, comparisons among repeated measures, levels, and groups of individuals are possible. Clinically, providers would be able to assess change over repeated measures and conclude that the measured change was a true change in the patient's perceived disablement instead of measurement error. Our invariant solution allowed us to compare sample variances

and means for IMP, FL, and QOL across repeated measures. Overall, the variance in scores decreased, and participants reported improved scores for IMP, FL, and QOL, implying that treatment improved an individual's overall disablement and QOL across repeated measures.

Although the invariant model offers a promising tool for tracking a patient's disablement, overall model fit and potential multicollinearity raise concerns. Although it can be argued that scales used to assess disablement should be multidimensional,[1,9,10] scale subdimensions must uniquely contribute without substantial overlap to provide a psychometrically sound and precise measure.[3,7] Although the DPA SF-10 meets many fit requirements and offers a lower response burden than that of the original DPA scale, our data suggested the multicollinearity and model misspecification found in a previous study[2] were still present among a sample of individuals who responded only to the 10 included items. Furthermore, our results confirm earlier findings[1,2] that demonstrated the QOL and DIS latent variables were unique constructs. The uniqueness of these constructs suggests that the same phenomenon is not being measured and that these responses should not be summed as a score of DIS if 1 construct that is different from QOL is already labeled *disablement*. Our results supported previous results[2] that suggested clinicians should score and assess the individual construct scores to evaluate patient health status across physical (ie, DIS) and mental (ie, QOL) health constructs.[11] When considered in the context of previous findings,[2,11] our overall findings suggest further alteration of the DPA SF-10 is necessary before it can be used in clinical practice and research.

LIMITATIONS AND FUTURE RESEARCH

We recruited a diverse population from sites across the United States; however, participants were primarily from a small group of clinics and consisted of younger individuals. Therefore, future investigators should assess the measurement properties of the scale in samples that include active pediatric, older adult, and geriatric individuals to ensure that the instrument is psychometrically sound. Furthermore, we did not establish the responsiveness of the DPA SF-10; future authors should include this step if the scale is to be used to assess the effectiveness of clinical practice. Previous researchers have also indicated the model fit concerns in the DPA SF-10 may be adequately addressed using an 8-item scale solution (ie, DPA scale Short Form-8 [SF-8]) that removes items 4 (ie, Stability) and 6 (ie, Skill Performance) from the DPA SF-10.[2] The DPA SF-8 solution fully addressed the multicollinearity present in both the DPA scale and the DPA SF-10; therefore, the DPA SF-8 may be an appropriate instrument to use if its psychometric properties can be validated with a new sample of participants who respond only to the 8 items included in the DPA SF-8. Future investigators should use similar methods as in our study to confirm the psychometric properties of the DPA SF-8 while also conducting invariance testing to ensure that the scale can be used to assess change with repeated measures or differences across groups among a sample of individuals who answer only those 8 items.

CONCLUSIONS

Psychometric properties of the DPA SF-10 were assessed at and across repeated measures. The CFA procedures indicated the DPA SF-10 met some goodness-of-fit indices and invariant criteria; however, a number of localized concerns previously identified in the scale were confirmed among a sample of individuals who answered only the items included in the DPA SF-10. Whereas the DPA SF-10 has improved the measurement properties and reduced the response burden compared with that of the original DPA scale, adoption into clinical practice is not recommended until further psychometric testing and model alteration confirm the most psychometrically sound instrument.

REFERENCES

1. Vela LI, Denegar CR. The Disablement in the Physically Active Scale, part II: the psychometric properties of an outcomes scale for musculoskeletal injuries. *J Athl Train*. 2010;45(6):630–641. doi:10.4085/1062-6050-45.6.630
2. Baker RT, Burton D, Pickering MA, Start A. Confirmatory factor analysis of the Disablement in the Physically Active Scale and preliminary testing of short-form versions: a calibration and validation study. *J Athl Train*. 2019;54(3):302–318. doi:10.4085/1062-6050-355-17
3. Kline RB. *Principles and Practice of Structural Equation Modeling*. 4th ed. Guilford Press; 2016.
4. US Department of Health and Human Services. *Physical Activity Guidelines for Americans*. 2nd ed. US Department of Health and Human Services; 2018.
5. Bryant FB, Yarnold PR. Principal-components analysis and exploratory and confirmatory factor analysis. In: Grimm LG, Yarnold PR, eds. *Reading and Understanding Multivariate Statistics*. American Psychological Association; 1995:99–136.
6. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6(1):1–55. doi:10.1080/10705519909540118
7. Brown TA. *Confirmatory Factor Analysis for Applied Research*. 2nd ed. Guilford Press; 2015.
8. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 4th ed. Allyn & Bacon; 2001.
9. Verbrugge LM, Jette AM. The disablement process. *Soc Sci Med*. 1994;38(1):1–14. doi:10.1016/0277-9536(94)90294-1
10. Jette AM. Outcomes research: shifting the dominant research paradigm in physical therapy. *Phys Ther*. 1995;75(11):965–970. doi:10.1093/ptj/75.11.965
11. Houston MN, Hoch JM, Van Lunen BL, Hoch MC. The development of summary components for the Disablement in the Physically Active Scale in collegiate athletes. *Qual Life Res*. 2015;24(11):2657–2662. doi:10.1007/s11136-015-1007-6

Address correspondence to Russell T. Baker, PhD, DAT, ATC, CMP, PRT-c, Department of Movement Sciences, University of Idaho, 875 Perimeter Drive, Moscow, ID 83844. Address email to russellb@uidaho.edu.