



Published in final edited form as:

*J Pain*. 2020 ; 21(5-6): 557–569. doi:10.1016/j.jpain.2019.09.003.

## Psychometric evaluation of the Optimal Screening for Prediction of Referral and Outcome Yellow Flag (OSPRO-YF) tool: factor structure, reliability, and validity

Katie A. Butera, PT, DPT<sup>1</sup>, Steven Z. George, PT, PhD, FAPTA<sup>2</sup>, Trevor A. Lentz, PT, PhD, MPH<sup>2</sup>

<sup>1</sup>Department of Physical Therapy, College of Public Health and Health Professions, University of Florida, Box 100154 UFHSC, Gainesville, FL, 32610 (USA);

<sup>2</sup>Department of Orthopaedic Surgery and Duke Clinical Research Institute, Duke University, Durham NC 27707 (USA);

### Abstract

Psychological characteristics consistently predict clinical outcomes for musculoskeletal pain conditions. The Optimal Screening for Prediction of Referral and Outcome Yellow Flag (OSPRO-YF) tool assesses negative mood, fear-avoidance, and positive affect/coping. Psychometric testing of the tool is needed to guide clinical use. The purpose of this secondary analysis was to determine the OSPRO-YF factor structure (17- and 10-item versions) and test factor reliability and concurrent validity. Data from two musculoskeletal pain cohorts ( $n_1=428$ ;  $n_2=440$ ) were used. An exploratory factor analysis (EFA) identified the factor structure in the first cohort; a confirmatory factor analysis (CFA) validated the factor structure in the second cohort. EFA yielded 4 factors for the 17-item version (catastrophizing, positive coping, negative mood, fear-avoidance) and 3 factors for the 10-item version (negative coping, negative mood positive affect/coping). CFA indicated good fit for the 10-item and 17-item models. Factors from both versions demonstrated moderate to good test-retest reliability. Regression results revealed that factors from both versions contributed significantly to variance in pain intensity, disability, and quality of life. Results indicate the OSPRO-YF is a reliable and valid multidimensional psychological assessment tool for individuals with musculoskeletal pain. Future studies are needed to establish the OSPRO-YF as a prognostic and treatment monitoring tool.

### Keywords

pain-related psychological characteristics; pain; musculoskeletal pain; multidimensional screening; psychological assessment; psychometric evaluation; yellow flags

---

Please address all correspondence to Trevor A. Lentz: Duke Clinical Research Institute, 200 Morris Street, Durham, NC 27701, Office Phone: 919-668-7495; Trevor.Lentz@duke.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

*Disclosures:* The authors have no conflicts of interest to disclose.

## Introduction

Psychological characteristics consistently predict clinical outcomes for musculoskeletal pain conditions [1,3,4,24,25,36,47,55]. National pain management initiatives [15,57] and clinical practice guidelines [11,14,37,43] recommend clinicians routinely assess psychological characteristics for pain conditions. Multiple tools exist to aid clinicians in evaluating pain-related psychological characteristics. Unidimensional tools that measure discrete psychological characteristics (e.g. pain catastrophizing or depressive symptoms) are common. However, emerging evidence suggests that evaluation of multiple psychological characteristics allows for better interpretation of patients' risk for poor outcomes [9,61]. As a result, there is interest in using multidimensional tools that concurrently estimate multiple pain-related psychological characteristics. These tools provide an efficient option for multifactorial risk assessment and can guide more informed treatment decision-making [46].

The Optimal Screening for Prediction of Referral and Outcome Yellow Flag (OSPRO-YF) Tool is a recently developed, 17-item multidimensional screening tool used for assessment of pain-related psychological characteristics among individuals with musculoskeletal pain conditions (Appendix) [39]. The tool evaluates a variety of discrete pain-related psychological characteristics across three theoretical constructs: 1) negative mood (i.e. depression, anxiety, and anger); 2) fear-avoidance (i.e. fear-avoidance beliefs, catastrophizing, kinesiophobia, and pain-anxiety); and 3) positive affect/coping (i.e. pain self-efficacy, rehabilitation self-efficacy, and chronic pain acceptance). The OSPRO-YF tool distinguishes itself from other multidimensional tools by estimating positive psychological characteristics [26,41,42]. Another unique characteristic of the OSPRO-YF tool is that it accurately estimates scores for 10 full-length, unidimensional psychological questionnaires [39]. The tool also identifies the presence of "yellow flags", defined as highest or lowest quartile (negative and positive constructs respectively) score estimates across each of the 10 unidimensional questionnaires [39].

Recent validation studies for the OSPRO-YF tool used a composite (i.e. simple summary of items) score calculated by a simple sum of the items. Using this method, OSPRO-YF scores predicted 12-month pain severity, disability, and health care utilization outcomes following physical therapy [5,23,40]. However, to date additional psychometric properties of the OSPRO-YF have not been further investigated. A better understanding of the underlying constructs measured by the tool is necessary even though the tool was developed with an established theoretical structure. This additional psychometric information will aid in its implementation and clinical interpretation [39].

The purpose of this secondary analysis was to evaluate the psychometric properties of the 17-item OSPRO-YF tool, and a shortened 10-item version, by: 1) empirically deriving the overall factor structure of the tool; 2) reporting test-retest reliability of the identified OSPRO-YF factors; and 3) determining concurrent validity of the identified OSPRO-YF factor structure by comparing associations with a) pain intensity, b) disability, and c) quality of life. We hypothesized that the 17- and 10-item versions would consist of three latent factors consistent with the selected theoretical constructs: negative mood, fear-avoidance, and positive affect/coping. Further, we hypothesized that the factors would demonstrate

adequate test-retest reliability and concurrent validity with measures commonly used for clinical outcomes (i.e. pain, self-reported disability, and quality of life).

## Methods

### Overview

This paper reports on a secondary analysis of data from the OSPRO cohort study. The OSPRO cohort study included patients seeking outpatient physical therapy for low back, neck, shoulder, and knee pain. It was conducted to develop the OSPRO-YF and a separate review of systems (OSPRO-ROS) screening tool for use in outpatient orthopedic physical therapy settings. The OSPRO cohort study included a cross-sectional phase in which the screening tools were developed (i.e. development cohort), and a separate longitudinal phase used to validate the tools for one-year clinical outcomes (i.e. validation cohort). Methodological details for the development [21,39] and validation cohorts [22] have been previously reported; key details of the cohort design are included in the following sections. Of note, this secondary analysis focuses on the OSPRO-YF only.

Participants in both OSPRO cohort studies were recruited from the Orthopaedic Physical Therapy-Investigative Network (OPT-IN). OPT-IN is a research network supported by the Academy of Orthopaedic Physical Therapy, with the purpose of performing multicenter clinical projects that examine diagnosis/classification, prognosis, and/or patient-centered treatment outcomes in patients with musculoskeletal conditions commonly managed by orthopedic physical therapists. The University of Florida Institutional Review Board approved this study and informed consent was obtained from all participants.

### Development Cohort

A convenience sample of participants (n=427) was recruited from participating OPT-IN clinical sites (n=11) during the development phase (March 2013 to May 2014). Clinical sites included outpatient physical therapy clinics in Gainesville, Florida (n=3) and Jacksonville, FL (n=8). The sample included individuals with musculoskeletal pain in four different anatomical regions: low back (n=118; 28%), neck (n=92; 22%), shoulder (n=106; 25%), and knee (n=111; 26%).

All data in this cross-sectional phase of the study were collected via self-report using standard intake forms provided to participants during their initial physical therapy evaluation. Participants were asked to answer questions regarding the musculoskeletal condition for which they were seeking treatment. All data were entered into a web-based electronic records database (REDCap; Vanderbilt University, Nashville, TN). A separate sub-sample (n=38) of the development cohort was recruited to examine test-retest reliability of the newly developed tools. Participants in the sub-sample completed the questionnaires at their initial physical therapy evaluation, and then again 48 hours later.

### Validation Cohort

A separate convenience sample of participants (n=440) was recruited from participating OPT-IN clinical sites (n=9) during the validation phase (December 2014 to December 2015).

The OPT-IN network was expanded for this phase, with clinics representing five of eight geographic regions of the United States, including the Mideast, Southeast, Great Lakes, Rocky Mountain States and Far West regions. The sample included individuals with musculoskeletal pain in four different anatomical regions: low back (n=118; 27%), neck (n=98; 22%), shoulder (n=107; 24%), and knee (n=117; 27%).

All data were collected via self-report and entered electronically by the participant in a deidentified manner. Baseline and follow-up data collection occurred online using a secure website. Follow-up time points were at 4 weeks, 6 months and 12 months; participants were notified of pending assessment by an email that directed them back to the study website to complete their follow-up assessment. Only baseline data from the validation cohort were considered for this secondary analysis.

## Participants

The following inclusion/exclusion criteria were utilized for both cohort studies.

**Inclusion criteria**—Participants were eligible to participate if they were: 1) between the ages of 18 and 75; 2) seeking outpatient physical therapy treatment for musculoskeletal pain in the low back, neck, knee, or shoulder; and 3) able to read and understanding English.

**Exclusion criteria**—Participants were excluded from participation if they: 1) were diagnosed with a widespread chronic pain syndrome (e.g. fibromyalgia, irritable bowel syndrome) or neuropathic pain syndrome (e.g. diabetic neuropathy, complex regional pain syndrome); 3) had a history of a psychiatric disorder (currently under the care of a mental health provider or taking multiple psychiatric medications); 4) were receiving treatment for active cancer; 5) had a history of a neurological disorder (e.g. stroke, spinal cord injury).

## Measures

Similar demographic/historical information and clinical measures were collected during both cohort phases and have been previously reported [22,39]. We only detail the information and measures utilized in this secondary analysis here.

**Demographic and Historical Information**—Age (years) and gender were collected from each participant. Participants also reported their primary anatomical region of pain (low back, neck, shoulder, or knee) and surgical history for the primary anatomical of pain (yes or no). Pain chronicity was determined by asking participants if they had “experienced any pain and activity limitations every day for the past 3 months?”. Participants answering yes were coded as having chronic pain. Similar operational definitions of chronic pain have been used previously [5,10,18] and are supported by national guidelines [16].

**Optimal Screening for Prediction of Referral and Outcome Yellow Flag Tool (OSPRO-YF)**—The previously described OSPRO-YF is a multidimensional screening tool that includes items from negative mood, fear-avoidance, and positive affect/coping factors [39]. The purpose of this tool is to aid with identification of pain-related psychological characteristics in outpatient orthopedic settings. Most questions inquire about general

disposition or how people feel currently/at the time of filling out the questionnaire. The OSPRO-YF has 17- and 10-item versions (Appendix). Information on the statistical process for identifying the OSPRO-YF items is reported in the original development paper [39]. Briefly, the 17- and 10-item tools were created through an iterative process that considered different item sets based on their ability to accurately predict full length questionnaire scores. Accuracy for the 17- and 10-item versions of the tool were at least 85% and 81% for full length questionnaire scores, respectively. The 10-item version was derived independently but used the same iterative process as the 17-item version. The current paper used the original 17- and 10-item OSPRO-YF tools reported in the development paper (Appendix) [39].

### **Clinical Measures**

**Average Daily Pain Intensity:** Pain intensity was assessed with a numerical pain rating scale ranging from 0 (“no pain”) to 10 (“worst pain imaginable”) [7,12,34]. Participants rated their current, best, and worst pain intensity over the preceding 24 hours. Prior evidence from other researchers indicates that single pain ratings lack validity, especially when compared to serial ratings that occur over a given period of time (e.g. daily ratings for a 7-day period) [32–35]. It was not feasible to obtain serial ratings during the routine care episodes for this study, so we incorporated a composite pain rating by averaging three pain intensity ratings—current, best, and worst—for the preceding 24-hour period. This value is reported as the “average daily pain intensity” in the current analysis. This approach was selected because a composite pain rating provides better reliability and validity than single pain ratings [32–35] and offers a pragmatic way to improve validity when collecting data in routine care settings. Indeed, we have used this approach previously in several clinical pain research studies [5,6,20,23].

**Disability:** Disability was assessed with region-specific self-report measures and z-score transformation. The Oswestry Disability Questionnaire (ODI) [19,30], Neck Disability Index (NDI) [58,59], short-version of the Disabilities of the Arm, Shoulder and Hand Questionnaire (QuickDASH) [2], and International Knee Documentation Committee Subjective Knee Form (IKDC) [31] were used for low back, neck, shoulder, and knee pain, respectively, depending on the participant’s self-reported primary pain location. Then, region-specific disability scores—ODI, NDI, QuickDASH, and IKDC—were transformed to z-scores so they could be used in the same regression models. Scores from the IKDC were reversed before standardizing to represent a measure of disability that was consistent with other measures (i.e. higher scores indicating greater disability). This method of standardization when measuring disability across multiple anatomical regions has been used previously [8,22,23].

**Mental and Physical Quality of Life:** The Medical Outcomes Study 8-item Short-Form Health Survey (SF-8) was used to measure general health quality of life [60]. Specifically, the SF-8 Mental Component Score (SF-8 MCS) and SF-8 Physical Component Score (SF-8 PCS) were used separately in the current secondary analysis [60].

## Sample Size Estimate

Sample size recommendations for studies involving factor analysis vary. Costello and Osborne (2005) indicate that factor analysis studies commonly include a case to item ratio of 5:1–10:1 and that a case to item ratio within this range is often considered an acceptable [13]. Pooled analyses from multiple studies indicated that a case to item ratio of 20:1 produced the most accurate factor analysis results [13]. Additionally, Tabachnick and Fidell (6<sup>th</sup> edition, 2013) also recommend sample sizes of at least 300 in studies involving factor analysis, especially when there are concerns regarding the strength of the data (e.g. low communalities or a small number of factors) [56]. Furthermore, a minimum of 100 participants is recommended for psychometric studies, with larger sample sizes acknowledged as being better [45,52,56]. Based on these recommendations, the goal for both cohorts was to recruit a minimum of 400 total participants (equivalent to a 23:1 case to item ratio for the 17-item OSPRO-YF version and 40:1 case to item ratio for the 10-item OSPRO-YF version).

## Data Analysis

Data analyses were conducted using IBM SPSS Statistics for Windows Version 24.0 and IBM SPSS Amos Graphics Version 25.0 (IBM Corp, Armonk, New York).

### Aim 1a: OSPRO-YF Factors

**Identifying OSPRO-YF Factors**—Data from the OPT-IN development cohort (n=427) were used to identify factors for both the 17-item and 10-item OSPRO-YF versions. There were 40 participants (less than 10%) with at least one piece of missing data from responses to OSPRO-YF items; listwise deletion was used to exclude participants with missing data from this part of the analysis. An exploratory factor analysis (EFA) was conducted on OSPRO-YF items using a polychoric correlation matrix, principal axis factoring, and oblique (i.e. PROMAX) rotation. Eigenvalues > 1 criterion determined the number of factors; factor composition was defined by items with loadings > | 0.4 |. Items that were cross-loaded (loading > | 0.4 | on more than one factor) or did not meet the loading threshold were not assigned to a factor.

**Confirming OSPRO-YF Factors**—Data from the OPT-IN validation cohort (n=440) were used to confirm factors identified in the development cohort for both the 17-item and 10-item versions. There were no participants with missing data for this part of the analysis. A confirmatory factor analysis (CFA) was conducted via structural equation modeling and maximum likelihood estimates to validate the adequacy of the EFA results. Items that were cross-loaded (loading > | 0.4 | on more than one factor) or did not meet the loading threshold (as determined in the EFA) were dropped for this part of the analysis. Overall model fit was assessed through examination of the commonly used goodness-of-fit statistical tests and their associated thresholds: Goodness of Fit (GFI>0.9, good); Comparative Fit Index (CFI>0.95, good; >0.9 acceptable); Root Mean Square Error Approximation (RMSEA <0.06, good; <0.08, acceptable); and p-close, a value that tests the RMSEA confidence interval (p-close>0.05, reject null hypothesis of poor fit) [28,29,44].

**Factor Sub-Scoring**—Based on EFA results, the items and method used for factor sub-scoring was determined. We planned to calculate factor sub-scores by summing the individual items that loaded on each factor. Scores for items #2, #13, #14, #15, and #17 (positively worded items) would be reversed before summing to maintain consistent directionality across items; thus, higher scores would indicate higher psychological distress. However, if positively worded items were the only items loaded on a factor, then scores would not be reversed.

### **Aim 1b: Reliability of OSPRO-YF Factors**

Data from the OPT-IN development cohort sub-sample (n=38) were used to determine test-retest reliability of the 17-item and 10-item simple summary scores as well as the identified factor sub-scores. The subsample of patients selected for test-retest reliability was an independent sample collected specifically to test reliability of the tool outside of the development and validation cohorts. These individuals were recruited from one of the OPT-IN sites, and a limited dataset (e.g. OSPRO-YF scores only) was collected at intake and 48 hours later during a follow up visit. Test-retest reliability between baseline and 48-hour follow-up factor sub-scores were assessed using intraclass correlation coefficients (ICC). ICC estimates were calculated based on a single measurement, absolute agreement, two-way mixed-effects model. Reliability was rated as excellent (ICC values >0.9), good (ICC values=0.75 to 0.9), or moderate (ICC values=0.5 to 0.75 [38,48]).

### **Aim 2: Concurrent Validity of OSPRO-YF Factors**

Data from the OPT-IN validation cohort was used to determine concurrent validity of the factor sub-scores for both the 17-item and 10-item versions with average daily pain intensity, disability (z-score transformed), mental quality of life, and physical quality of life as dependent variables. Separate hierarchical linear regression models were developed for each OSPRO-YF version. The first block included factor sub-scores of the 17-item or 10-item OSPRO-YF version, and second block included the interaction between factor sub-scores and a key patient-level characteristic. Characteristics were gender, anatomical region, surgical history, or chronicity. These patient level characteristics were selected to determine whether relationships between factor sub-scores and dependent variables varied as a function of the characteristics. Interactions were entered in separate models, meaning there were 4 different concurrent validity models for each combination of dependent variable and OSPRO-YF version. (i.e., factor subscore x gender, anatomical region, surgical history, and chronicity for both the 17-item and 10-item version). For models with a gender interaction, males were the reference group. For models with an anatomical region interaction, knee pain was the reference group. For models with a surgical history interaction, no surgical history was the reference group. For models with a chronicity interaction, no chronic pain was the control group. An alpha level of 0.01 was used in this portion of the analysis to account for the multiple comparisons.

Regression model assumptions for multivariate models were checked by visual inspection and statistical estimates when appropriate. Concerns for multicollinearity were determined by examining variance inflation factor estimates and tolerance estimates for each model;

multicollinearity was not a concern if variance inflation factor estimates were  $<10$  and tolerance estimates were  $>0.2$  [54].

## Results

### Aim 1a: Establishing OSPRO-YF Factors

#### Identifying OSPRO-YF Factors

**17-item Version Factors:** The EFA for the 17-item version yielded a four-factor solution with eigenvalues of 5.78 (34% variance), 1.96 (12% variance), 1.36 (8% variance), and 1.01 (6% variance). Factor loadings ranged from 0.471 to 0.884 (Table 1). Item #16 did not load above 0.4 on any factor. Item #1 was cross-loaded on two factors and thus not assigned to a factor. Factors were labeled as “Catastrophizing” (Factor 1), “Negative Mood” (Factor 2), “Fear-Avoidance” (Factor 3), and “Positive Affect/Coping” (Factor 4) based on inspection and qualitative interpretation of item content.

**10-item Version Factors:** The EFA for the 10-item version yielded a three-factor solution with eigenvalues of 3.65 (36% variance), 1.26 (13% variance), and 1.05 (11% variance). There was no cross-loading and all item loadings ranged from 0.560 to 0.915 (Table 1). Factors were labeled as “Positive Affect/Coping” (Factor 1), “Negative Coping” (Factor 2), and “Negative Mood” (Factor 3) based on inspection and qualitative interpretation of item content. We chose a “Negative Coping” label rather than the “Fear Avoidance” theoretical factor label to better reflect the items in that factor (which originally came from questionnaires on fear avoidance, catastrophizing and kinesiophobia).

#### Confirming OSPRO-YF Factors

**Model Structure and Estimates:** Based on the EFA, a four-factor model structure was developed for the 17-item OSPRO-YF CFA. Item #1 (from the Patient Health Questionnaire (PHQ-9)) and item #16 (from the Chronic Pain Acceptance Questionnaire (CPAQ)) were dropped from this CFA due to cross-loading and not meeting the threshold criteria, respectively. The four-factor model structure and results are depicted in Figure 1. Similarly, a two-factor model structure was developed for the 10-item OSPRO-YF CFA. No items were dropped from this CFA. The two-factor model structure and results are depicted in Figure 2.

**Model Fit:** The four-factor solution for the 17-item version was fully supported by the fit statistic results (Figure 1). Good model fit was demonstrated by the GFI (0.941). Additionally, acceptable model fit was demonstrated by the CFI (0.914) and RMSEA (0.056), as well as a non-significant p-close value (0.174) indicating the null hypothesis of poor fit should be rejected. The three-factor solution for the 10-item version was also fully supported (Figure 2). Good model fit was demonstrated by the GFI (0.970) and the RMSEA (0.051), as well as a non-significant p-close value (0.433) indicating the null hypothesis of poor fit should be rejected. Further, acceptable model fit was demonstrated by the CFI (0.903).

**Factor Sub-Scoring**—Based on the EFA results, factor sub-scores were generated using the calculation methods described in Table 2. Four factor sub-scores were generated for the



17-item version and two factor sub-scores were generated for the 10-item version; these factor sub-scores were then used in all subsequent analyses (i.e. to test reliability and concurrent validity of the factors).

### **Aim 1b: Reliability of OSPRO-YF Factors**

For the 17-item version, all four factors demonstrated good to moderate reliability. ICC values indicated good reliability for the catastrophizing factor, negative mood factor and fear-avoidance factor (ICCs=0.83, 0.89, and 0.83, respectively; all  $p<0.01$ ) and moderate reliability for the positive affect/coping factor (ICC=0.63;  $p<0.01$ ). Additionally, the 17-item simple summary score demonstrated good reliability (ICC=0.88;  $p<0.001$ ). For the 10-item version, ICC values indicated good reliability for both the negative coping and negative mood factors (ICCs=0.80 and 0.86, respectively; both  $p<0.01$ ) and moderate reliability for the positive affect/coping factor (ICC=0.65;  $p<0.01$ ). Additionally, the 10-item simple summary score demonstrated good reliability (ICC=0.85;  $p<0.001$ ).

### **Aim 2: Concurrent Validity of OSPRO-YF Factors**

Visual inspection of relationships between each independent variable and dependent measure scores indicated linear relationships for the multivariate models. Multicollinearity was not a concern as variance inflation factor estimates were less than 2.3 and tolerance estimates were greater than 0.4 for all independent variables across all models.

**Four-factor, 17-item OSPRO-YF Tool**—Base model and interaction model regression results for the 17-item factors and clinical measures are shown in Table 3. The four factor sub-scores combined accounted for 22%, 37%, 37%, and 24% of the variance in average daily pain intensity, disability, mental quality of life, and physical quality of life, respectively. The catastrophizing and positive affect/coping factors significantly contributed to the average daily pain measure ( $p<0.01$ ). Catastrophizing, fear avoidance, and positive affect/coping factors significantly contributed to disability ( $p<0.01$ ). Catastrophizing, negative mood, and positive affect/coping factors significantly contributed to the mental quality of life and physical quality of life measures ( $p<0.01$ ). There were no significant interactions by gender, by anatomical region, by chronicity, nor by surgical history.

**Two-factor, 10-item OSPRO-YF Tool**—Base model and interaction model regression results for the 10-item factors and clinical measures are shown in Table 4. The three factor sub-scores combined accounted for 15%, 33%, 34%, and 19% of the variance in average daily pain intensity, disability, mental quality of life, and physical quality of life, respectively. All three factors—positive affect/coping, negative coping, and negative mood—significantly contributed to the mental quality of life measure ( $p<0.01$ ), while only the positive affect/coping and negative coping factors significantly contributed to average daily pain, disability, and physical quality of life measures ( $p<0.01$ ). There were no significant interactions by gender, by anatomical region, by chronicity, nor by surgical history.

## Discussion

These psychometric analyses provide further evidence supporting the OSPRO-YF as a reliable and valid multidimensional psychological assessment tool for individuals experiencing common musculoskeletal pain conditions (i.e. low back, neck, shoulder, or knee pain). Our results indicated that the empirical factor structure of the 17-item version of the OSPRO-YF differed slightly from the underlying theoretical structure. Specifically, the theoretical structure included fear-avoidance and pain catastrophizing together as negative coping; however, in the current analyses this factor was differentiated into two separate factors for fear-avoidance and pain catastrophizing. The 17-item version was psychometrically-sound based on good test-retest reliability and concurrent validity with common clinical outcome measures for pain severity, perceived disability, and mental and physical quality of life. Factor analysis results indicated that the 10-item version of the OSPRO-YF included positive affect/coping, negative coping, and negative mood; this factor structure was similar to the original theoretical structure. The 10-item version was also psychometrically-sound based on good to moderate test-retest reliability and concurrent validity with common clinical outcome measures for pain severity, perceived disability, and mental and physical quality of life.

### Clinical Implementation of the OSPRO-YF

Consistency of psychometric testing across two separate cohorts strengthens the case that the OSPRO-YF is suitable for clinical implementation as an assessment tool for pain-related psychological factors. Our findings support the use of a simple summary score in previous OSPRO-YF validation studies [5,22,38] and suggest that summing OSPRO-YF items to create a simple summary score or factor sub-scores (i.e. simple summing of factor-specific items) offers reliable and valid scoring options beyond those originally proposed during development of the tool (i.e., estimation of parent questionnaire scores and identification of yellow flags). We recommend clinicians utilize the methods summarized in Table 2 to calculate factor sub-scores for each OSPRO-YF version.

**Use of the 17-item Version—Better Discrimination of Sub-Scores—**The 17-item OSPRO-YF demonstrated a stable factor structure stability for identifying catastrophizing, negative mood, fear avoidance, and positive affect/coping. Therefore, this version of the tool is recommended to those interested in differentiating distinct underlying constructs of negative coping (i.e., fear avoidance and catastrophizing), or evaluating positive affect/coping. The latter is a notable strength of the OSPRO-YF because other available multidimensional tools identify only negative factors. Emerging evidence demonstrates positive psychological characteristics (i.e. optimism, self-efficacy) are strong contributors to improved clinical outcomes [17,49–51,53]. Thus, use of the 17-item version may be especially valuable for simultaneously evaluating positive coping and multiple underlying constructs of negative coping, with the added benefits of reasonable accuracy in estimating total parent questionnaire scores.

**Use of the 10-item Version—Pragmatic and Predictive—**The 10-item OSPRO-YF demonstrated a three-factor structure closely aligned with the proposed theoretical factor

structure. This version is capable of separately evaluating positive and negative coping characteristics. However, it does not differentiate between fear-avoidance and pain catastrophizing like the 17-item version does. Nevertheless, it has many practical benefits. First, it is a concise version of the OSPRO tool, and it is an efficient option for accurately estimating full length questionnaire scores [39]. Use of the 10-item version minimizes patient and provider burden by limiting the total number of items. Second, simple summary scores for the 10-item version have demonstrated accuracy in predicting 12-month disability, persistence of chronic pain, and risk for subsequent surgery following an episode of physical therapy [5,23,40]. Therefore, the 10-item version may be the preferred option if the intent of assessment is to quickly determine a patient's risk for poor outcomes, or if differentiating between fear avoidance and pain catastrophizing is not needed.

### Strengths and Limitations

This study has several important strengths. First, the sample size for both the development and validation cohorts exceeded recommendations for factor analyses [13]. Moreover, follow up testing with CFA in a separate cohort following EFA is also a recommended best practice for determining the consistency of factor structure [13]. Second, we conducted several psychometric assessments of the tool, including factor analysis, reliability, and concurrent validity, all of which are necessary for informed clinical implementation of the tool. Third, the concurrent validation was conducted using multiple, common clinical patient-reported measures that are often used as primary endpoints. Additionally, the inclusion criteria for both study cohorts were intentionally broad [22,39], resulting in a pragmatic, generalizable sample of individuals for which this questionnaire would be used in orthopedic clinical practice.

This study also includes some limitations. First, the OSPRO development and validation cohorts only included patients with the four most prevalent musculoskeletal conditions: low back, neck, shoulder, or knee pain. Thus, we are unable to assess whether the factor structure would differ in other orthopedic or pain-related populations (e.g. fibromyalgia), although we expect it to be similar. Second, study cohorts were obtained via convenience sampling and only included individuals seeking physical therapy treatment. Moreover, we excluded individuals undergoing active treatment for a mental health disorder. The OSPRO-YF was designed to be a quick assessment tool for capturing pain-related distress, and it was not meant to be a tool used for those with known psychopathology. Nevertheless, chronic pain may often be associated with other comorbid psychiatric conditions that would require care by a mental health provider and use of this tool in those populations requires further study. Overall, these limitations impact the generalizability of the tool, but also provide important opportunities for future psychometric analyses.

### Future Directions

While both versions of the OSPRO-YF tool demonstrate utility, additional longitudinal studies are warranted. While the predictive validity of OSPRO-YF simple summary scores are established [5,23,40], future studies are required to establish whether the OSPRO-YF factors identified in this study predict clinical outcomes and which factors are the most accurate predictors. This will provide clinicians with better direction on best uses of the

OSPRO-YF as a prognostic and/or treatment monitoring tool. Additionally, the OSPRO-YF complements existing risk-stratification approaches and allows for concise assessment of multiple negative and positive psychological factors. Therefore, studies aimed at using the OSPRO-YF alongside already validated risk-stratification measures (i.e. the STarT Back Screening Tool or Örebro Musculoskeletal Pain Screening Questionnaire) [26,27,41,42] would be of value. Risk screening *and* accurate identification of specific psychological factors is likely necessary to optimize psychologically informed practice and develop tailored treatment-matched pathways [39].

Finally, studies examining the psychometric properties of the OSPRO-YF in other pain conditions (e.g. musculoskeletal pain in other regions, widespread pain conditions, neuropathic pain syndromes, and known psychopathology) are critical for broadening this tool's clinical application.

## Acknowledgements

The authors wish to acknowledge Sam Wu, PhD, and Yunfeng Dai for their contributions to the development of the OSPRO-YF. Giorgio Zeppieri Jr is acknowledged for his contributions in collection of the reliability sample. OPT-IN Network Participants included: University of Florida: Joel Bialosky; UF Health: Giorgio Zeppieri, Jr., Daniel Broome, Marty Huegel, Debi Jones, Steve Emery, Mike Hodges, Derek Miles, Jodi Davis, Charlene Stubbington, Mike Darcy; ATI Physical Therapy: Ellen Shanley, Thomas Denninger, Jenna Bartsokas, Elise Harris, Jordan Floyd, Wade Harrell; University of Southern California: Lori Michener, Amy Pomrantz, Brooks Rehabilitation: Jason Beneciuk, Raine Osborne, Nata Salvatori, John Leschitz, Brian Hagist, Laura Langer, Tim Shreve, Nando Malaman, Michael Bourassa, Justin Zych, Tasha Mouton Shanklin; University of Illinois at Chicago: Aaron Keil, Brad Myers, Deb Davey, Justin Payette, Adam Wielechowski, Richard Severin, Erik Martinez; Indiana State University: Ryan Hanigan, Carolina Valencia, Danielle Jena, Nicole Woodard; Arcadia University: Angela Tate; Life's Work Physical Therapy: Sandra Stryker, Aaron Leonard, Erin Courtney, Brandon Little, Kathryn Jankord, Brad Simpson, Charleen Hall, Paige Nixon, Julia Neufeld; University of Colorado, Denver: Paul Mintken, Virginia Amette, Andrea Barsch.

The original project was supported by the 2013 Clinical Research Network grant from the Academy of Orthopaedic Physical Therapy (formerly known as the Orthopaedic Section of the American Physical Therapy Association). The funding body had no role in the design of the study or collection, analysis, and interpretation of the data or in writing the manuscript.

KAB received support from the NIH T-32 Neuromuscular Plasticity Pre-Doctoral Fellowship (NIH T32 HD 043730), the Foundation for Physical Therapy Research Promotion of Doctoral Studies Level I and II Scholarships, the International Chapter of the P.E.O. Sisterhood Scholar Award, and the Brooks-PHHP Research Collaboration while working on this project. SZG received support from NIH while working on this manuscript (NIAMS AR055899). TL received support for this project from the Foundation for Physical Therapy with Promotion of Doctoral Studies (PODS I & II) Awards.

## References

- [1]. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. *Arch Intern Med* 163:2433–45, 2003. [PubMed: 14609780]
- [2]. Beaton DE, Wright JG, Katz JN, Upper Extremity Collaborative G. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am* 87:1038–46, 2005. [PubMed: 15866967]
- [3]. Beneciuk JM, Bishop MD, Fritz JM, Robinson ME, Asal NR, Nisenzon AN, George SZ. The STarT back screening tool and individual psychological measures: evaluation of prognostic capabilities for low back pain clinical outcomes in outpatient physical therapy settings. *Phys Ther* 93:321–33, 2013. [PubMed: 23125279]
- [4]. Beneciuk JM, Fritz JM, George SZ. The STarT Back Screening Tool for prediction of 6-month clinical outcomes: relevance of change patterns in outpatient physical therapy settings. *J Orthop Sports Phys Ther* 44:656–64, 2014. [PubMed: 25098194]

- [5]. Beneciuk JM, Lentz TA, He Y, Wu SS, George SZ. Prediction of persistent musculoskeletal pain at 12 months: a secondary analysis of the Optimal Screening for Prediction of Referral and Outcome (OSPRO) validation cohort study. *Phys Ther* 98:290–301, 2018. [PubMed: 29669081]
- [6]. Beneciuk JM, Robinson ME, George SZ. Low back pain subgroups using fear-avoidance model measures: results of a cluster analysis. *Clin J Pain* 28:658, 2012. [PubMed: 22510537]
- [7]. Bolton JE. Accuracy of recall of usual pain intensity in back pain patients. *Pain* 83:533–39, 1999. [PubMed: 10568862]
- [8]. Butera KA, Lentz TA, Beneciuk JM, George SZ. Preliminary Evaluation of a Modified STarT Back Screening Tool Across Different Musculoskeletal Pain Conditions. *Phys Ther* 96:1251–61, 2016. [PubMed: 26847006]
- [9]. Campbell P, Bishop A, Dunn KM, Main CJ, Thomas E, Foster NE. Conceptual overlap of psychological constructs in low back pain. *Pain* 154:1783–91, 2013. [PubMed: 23726935]
- [10]. Carey TS, Freburger JK, Holmes GM, Jackman A, Knauer S, Wallace A, Darter J. Race, care seeking, and utilization for chronic back and neck pain: population perspectives. *J Pain* 11:343–350, 2010. [PubMed: 19853527]
- [11]. Childs JD, Cleland JA, Elliott JM, Teyhen DS, Wainner RS, Whitman JM, Sopky BJ, Godges JJ, Flynn TW, Delitto A. Neck pain: clinical practice guidelines linked to the International Classification of Functioning, Disability, and Health from the Orthopaedic Section of the American Physical Therapy Association. *J Orthop Sports Phys Ther* 38:A1–A34, 2008.
- [12]. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine* 30:1331–4, 2005. [PubMed: 15928561]
- [13]. Costello AB, Osborne JW. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Pract Assess Res Eval* 10:1–9, 2005.
- [14]. Delitto A, George SZ, Van Dillen L, Whitman JM, Sowa GA, Shekelle P, Denninger TR, Godges JJ. Low back pain: clinical practice guidelines linked to the international classification of functioning, disability, and health from the Orthopaedic Section of the American Physical Therapy Association. *J Orthop Sports Phys Ther* 42:A1–57, 2012.
- [15]. Department of Health and Human Services and National Institutes of Health. National Pain Strategy: A Comprehensive Population Health Strategy for Pain. Accessed Mar 2019 Available: [https://iprcc.nih.gov/sites/default/files/HHSNational\\_Pain\\_Strategy\\_508C.pdf](https://iprcc.nih.gov/sites/default/files/HHSNational_Pain_Strategy_508C.pdf)
- [16]. Deyo RA, Dworkin SF, Amtmann D, Andersson G, Borenstein D, Carragee E, Carrino J, Chou R, Cook K, DeLitto A. Report of the NIH Task Force on research standards for chronic low back pain. *Phys Ther* 95:e1–18, 2015. [PubMed: 25639530]
- [17]. Foster NE, Thomas E, Bishop A, Dunn KM, Main CJ. Distinctiveness of psychological obstacles to recovery in low back pain patients in primary care. *Pain* 148:398–406, 2010. [PubMed: 20022697]
- [18]. Freburger JK, Holmes GM, Agans RP, Jackman AM, Darter JD, Wallace AS, Castel LD, Kalsbeek WD, Carey TS. The rising prevalence of chronic low back pain. *Arch Intern Med* 169:251–8, 2009. [PubMed: 19204216]
- [19]. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther* 81:776–88, 2001. [PubMed: 11175676]
- [20]. George SZ, Beneciuk JM. Psychological predictors of recovery from low back pain: a prospective study. *BMC Musculoskelet Disord* 16:49, 2015. [PubMed: 25849159]
- [21]. George SZ, Beneciuk JM, Bialosky JE, Lentz TA, Zeppieri G Jr, Pei Q, Wu SS. Development of a review-of-systems screening tool for orthopaedic physical therapists: results from the Optimal Screening for Prediction of Referral and Outcome (OSPRO) cohort. *J Orthop Sports Phys Ther* 45:512–26, 2015. [PubMed: 25996361]
- [22]. George SZ, Beneciuk JM, Lentz TA, Wu SS. The Optimal Screening for Prediction of Referral and Outcome (OSPRO) in patients with musculoskeletal pain conditions: a longitudinal validation cohort from the USA. *BMJ Open* 7:e015188, 2017.
- [23]. George SZ, Beneciuk JM, Lentz TA, Wu SS, Dai Y, Bialosky JE, Zeppieri G Jr. Optimal screening for prediction of referral and outcome (OSPRO) for musculoskeletal pain conditions:

- results from the validation cohort. *J Orthop Sports Phys Ther* 48:460–75, 2018. [PubMed: 29629615]
- [24]. George SZ, Coronado RA, Beneciuk JM, Valencia C, Werneke MW, Hart DL. Depressive symptoms, anatomical region, and clinical outcomes for patients seeking outpatient physical therapy for musculoskeletal pain. *Phys Ther* 91:358–72, 2011. [PubMed: 21233305]
- [25]. George SZ, Stryker SE. Fear-avoidance beliefs and clinical outcomes for patients seeking outpatient physical therapy for musculoskeletal pain conditions. *J Orthop Sports Phys Ther* 41:249–59, 2011. [PubMed: 21335927]
- [26]. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, Hay EM. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum* 59:632–41, 2008. [PubMed: 18438893]
- [27]. Hill JC, Dunn KM, Main CJ, Hay EM. Subgrouping low back pain: a comparison of the STarT Back Tool with the Örebro Musculoskeletal Pain Screening Questionnaire. *Eur J Pain* 14:83–9, 2010. [PubMed: 19223271]
- [28]. Hooper D, Coughlan J, Mullen M. Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods* 6:53–60, 2008.
- [29]. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 6:1–55, 1999.
- [30]. Hudson-Cook N. T-NK. *Back Pain: New Approaches to Rehabilitation and Education*. New York, NY: Manchester University Press, 1989.
- [31]. Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, Richmond JC, Shelborne KD. Development and validation of the international knee documentation committee subjective knee form. *Am j Sports Med* 29:600–613, 2001. [PubMed: 11573919]
- [32]. Jensen MP, Castarlenas E, Tomé-Pires C, de la Vega R, Sánchez-Rodríguez E, Miró J. The Number of Ratings Needed for Valid Pain Assessment in Clinical Trials: Replication and Extension. *Pain Med* 16:1764–72, 2015. [PubMed: 26178637]
- [33]. Jensen MP, McFarland CA. Increasing the reliability and validity of pain intensity measurement in chronic pain patients. *Pain* 55:195–203, 1993. [PubMed: 8309709]
- [34]. Jensen MP, Turner JA, Romano JM, Fisher LD. Comparative reliability and validity of chronic pain intensity measures. *Pain* 83:157–62, 1999. [PubMed: 10534586]
- [35]. Jensen MP, Turner LR, Turner JA, Romano JM. The use of multiple-item scales for pain intensity measurement in chronic pain patients. *Pain* 67:35–40, 1996. [PubMed: 8895229]
- [36]. Keeley P, Creed F, Tomenson B, Todd C, Borglin G, Dickens C. Psychosocial predictors of health-related quality of life and health service utilisation in people with chronic low back pain. *Pain* 135:142–50, 2008. [PubMed: 17611036]
- [37]. Kelley MJ, Shaffer MA, Kuhn JE, Michener LA, Seitz AL, Uhl TL, Godges JJ, McClure PW, Altman RD, Davenport T. Shoulder pain and mobility deficits: adhesive capsulitis: clinical practice guidelines linked to the international classification of functioning, disability, and health from the Orthopaedic Section of the American Physical Therapy Association. *J Orthop Sports Phys Ther* 43:A1–31, 2013.
- [38]. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 15:155–63, 2016. [PubMed: 27330520]
- [39]. Lentz TA, Beneciuk JM, Bialosky JE, Zeppieri G, Dai Y, Wu SS, George SZ. Development of a Yellow Flag Assessment Tool for Orthopaedic Physical Therapists: Results From the Optimal Screening for Prediction of Referral and Outcome (OSPRO) Cohort. *J Orthop Sports Phys Ther* 46:327–43, 2016. [PubMed: 26999408]
- [40]. Lentz TA, Beneciuk JM, George SZ. Prediction of healthcare utilization following an episode of physical therapy for musculoskeletal pain. *BMC Health Serv Res* 18:648, 2018. [PubMed: 30126409]
- [41]. Linton SJ. *Manual for the Örebro Musculoskeletal Pain Screening Questionnaire: the early identification of patients at risk for chronic pain*. Örebro, Sweden: 1999.

- [42]. Linton SJ, Boersma K. Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Örebro Musculoskeletal Pain Questionnaire. *Clin J Pain* 19:80–86, 2003. [PubMed: 12616177]
- [43]. Logerstedt DS, Snyder-Mackler L, Ritter RC, Axe MJ, Godges J, Altman RD, Briggs M, Chu C, Delitto A, Ferland A. Knee pain and mobility impairments: meniscal and articular cartilage lesions: clinical practice guidelines linked to the international classification of functioning, disability, and health from the orthopaedic section of the American Physical Therapy Association. *J Orthop Sports Phys Ther* 40:A1–597, 2010.
- [44]. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods* 1:130, 1996.
- [45]. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. *Psychological Methods* 4:84, 1999.
- [46]. Nicholas MK, Linton SJ, Watson PJ, Main CJ, Decade of the Flags” Working G. Early identification and management of psychological risk factors (“yellow flags”) in patients with low back pain: a reappraisal. *Phys Ther* 91:737–53, 2011. [PubMed: 21451099]
- [47]. Picavet HS, Vlaeyen JW, Schouten JS. Pain catastrophizing and kinesiophobia: predictors of chronic low back pain. *Am J Epidemiol* 156:1028–34, 2002. [PubMed: 12446259]
- [48]. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. New Jersey: Prentice Hall, 2000.
- [49]. Powell R, Johnston M, Smith WC, King PM, Chambers WA, Krukowski Z, McKee L, Bruce J. Psychological risk factors for chronic post-surgical pain after inguinal hernia repair surgery: A prospective cohort study. *Euro J Pain* 16:600–610, 2012.
- [50]. Ronaldson A, Poole L, Kidd T, Leigh E, Jahangiri M, Steptoe A. Optimism measured pre-operatively is associated with reduced pain intensity and physical symptom reporting after coronary artery bypass graft surgery. *J Psychosom Res* 77:278–82, 2014. [PubMed: 25129850]
- [51]. Rosenberger PH, Kerns R, Jokl P, Ickovics JR. Mood and attitude predict pain outcomes following arthroscopic knee surgery. *Ann Behav Med* 37:70–76, 2009. [PubMed: 19169766]
- [52]. Sapnas KG, Zeller RA. Minimizing sample size when using exploratory factor analysis for measurement. *J Nurs Meas* 10:135–54, 2002. [PubMed: 12619534]
- [53]. Sipilä K, Ylöstalo PV, Ek E, Zitting P, Knuutila ML. Association between optimism and self-reported facial pain. *Acta Odontol Scand* 64:177–82, 2006. [PubMed: 16809196]
- [54]. Stevens JP. *Applied multivariate statistics for the social sciences*. 4th ed. Mahwah, New Jersey: Lawrence Erlbaum, 2002.
- [55]. Sullivan MJ, Stanish W, Sullivan ME, Tripp D. Differential predictors of pain and disability in patients with whiplash injuries. *Pain Res Manag* 7:68–74, 2002. [PubMed: 12185370]
- [56]. Tabachnick BG, Fidell LS. *Using multivariate statistics*. 6th ed. Boston, Massachusetts: Pearson, 2013.
- [57]. The Interagency Pain Research Coordinating Committee and the National Institutes of Health Office of Pain Policy. *Federal Pain Research Strategy*. Accessed Mar 2019 Available: [https://iprcc.nih.gov/sites/default/files/iprcc/FPRS\\_Research\\_Recommendations\\_Final\\_508C.pdf](https://iprcc.nih.gov/sites/default/files/iprcc/FPRS_Research_Recommendations_Final_508C.pdf)
- [58]. Vernon H. The Neck Disability Index: state-of-the-art, 1991–2008. *J Manipulative Physiol Ther* 31:491–502, 2008. [PubMed: 18803999]
- [59]. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther* 14:409–15, 1991. [PubMed: 1834753]
- [60]. Ware JE. *How to score and interpret single-item health status measures: a manual for users of the SF-8 health survey:(with a supplement on the SF-6 health survey)*. Lincoln, RI: QualityMetric Incorporated, 2001.
- [61]. Wideman TH, Hill JC, Main CJ, Lewis M, Sullivan MJ, Hay EM. Comparing the responsiveness of a brief, multidimensional risk screening tool for back pain to its unidimensional reference standards: the whole is greater than the sum of its parts. *Pain* 153:2182–91, 2012. [PubMed: 22800410]

### Highlights

- Psychometric evaluation indicates the OSPRO-YF is reliable and valid.
- The OSPRO-YF can be used as a multidimensional psychological assessment tool.
- Implementation of the OSPRO-YF may better guide clinical decision making.
- Implementation of the OSPRO-YF may also improve musculoskeletal pain management.



**Perspective:**

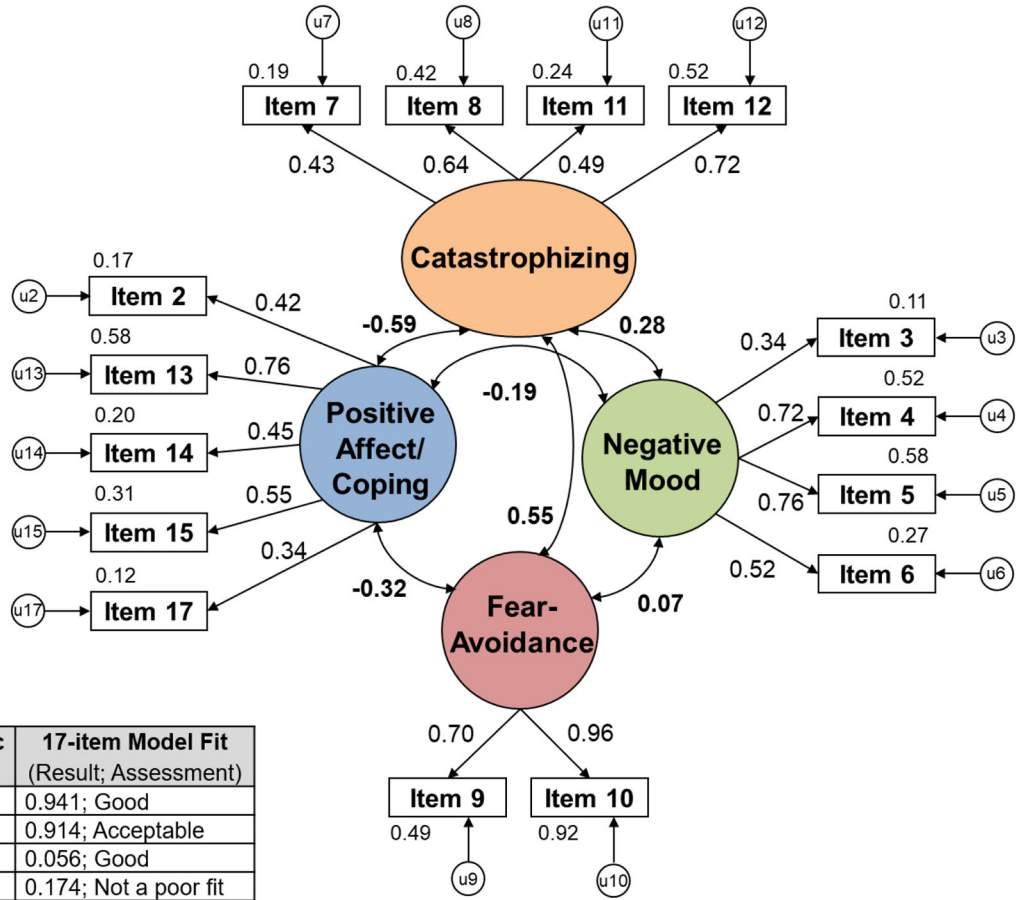
This article presents the psychometric properties of the OSPRO-YF tool. Findings indicate the OSPRO-YF is a reliable and valid multidimensional psychological assessment tool for individuals with musculoskeletal pain. Implementation of the OSPRO-YF tool may better guide clinical decision making and may lead to improved musculoskeletal pain management strategies.

Author Manuscript

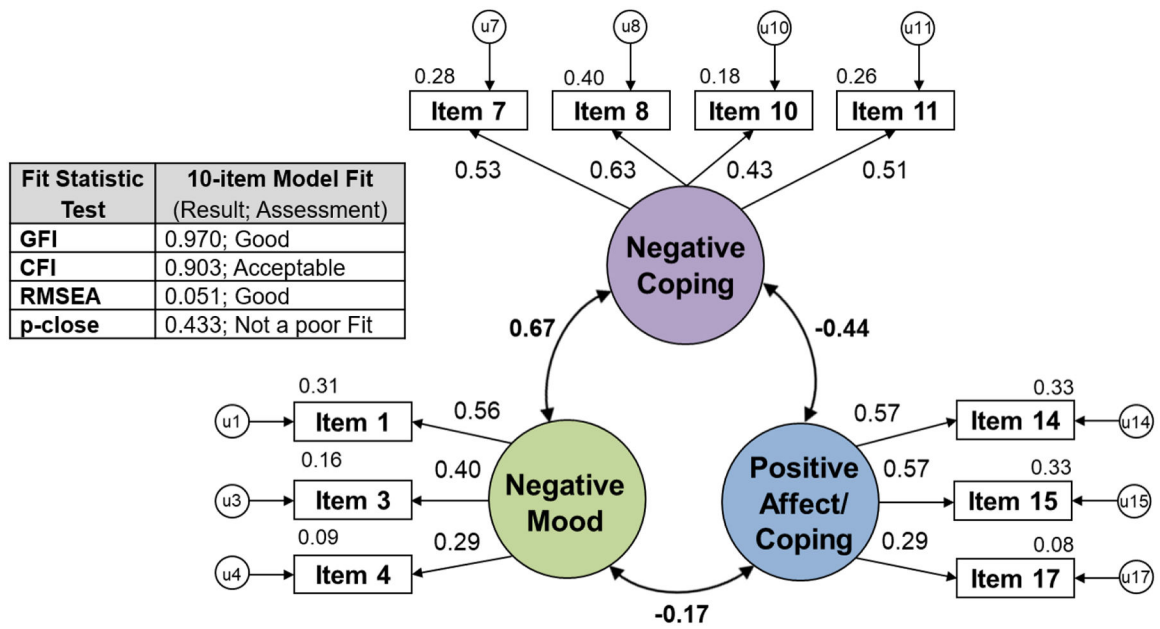
Author Manuscript

Author Manuscript

Author Manuscript



**Figure 1. 17-item OSPRO-YF Confirmatory Factor Analysis Results**  
 17-item OSPRO-YF four-factor model structure and resulting standardized parameter estimates; a summary of the fit statistic results is also shown; GFI=Goodness of Fit (>0.9, good); CFI=Comparative Fit Index (>0.95, good; >0.9 acceptable); RMSEA=Root Mean Square Error Approximation (<0.06, good; <0.08, acceptable); p-close refers to a value that tests the RMSEA confidence interval (p-close>0.05, reject null hypothesis of poor fit).



**Figure 2. 10-item OSPRO-YF Confirmatory Factor Analysis Results**

10-item OSPRO-YF three-factor model structure and resulting standardized parameter estimates; a summary of the fit statistic results is also shown; GFI=Goodness of Fit (>0.9, good); CFI=Comparative Fit Index (>0.95, good; >0.9 acceptable); RMSEA=Root Mean Square Error Approximation (<0.06, good; <0.08, acceptable); p-close refers to a value that tests the RMSEA confidence interval (p-close>0.05, reject null hypothesis of poor fit).

**Table 1.**

OSPRO-YF Item Theoretical Factors Compared to Exploratory Factor Analysis Results

OSPRO-YF Items* (Parent Questionnaire)	17-item Theoretical Factor	17-item Exploratory Factor Analysis Factor (Loading)	10-item Theoretical Factor	10-item Exploratory Factor Analysis Factor (Loading)
Item 1 (PHQ-9)	Negative Mood	(-0.457)**	Negative Mood	Negative Mood (0.568)
		(0.471)**		
Item 2 (STAI)	Negative Mood	Positive Affect/Coping (0.571)		
Item 3 (STAI)	Negative Mood	Negative Mood (0.594)	Negative Mood	Negative Mood (0.730)
Item 4 (STAXI)	Negative Mood	Negative Mood (0.783)	Negative Mood	Negative Mood (0.881)
Item 5 (STAXI)	Negative Mood	Negative Mood (0.817)		
Item 6 (STAXI)	Negative Mood	Negative Mood (0.655)		
Item 7 (TSK-11)	Fear-Avoidance	Catastrophizing (0.858)	Fear-Avoidance	Negative Coping (0.915)
Item 8 (PCS)	Fear-Avoidance	Catastrophizing (0.637)	Fear-Avoidance	Negative Coping (0.595)
Item 9 (FABQ-PA)	Fear-Avoidance	Fear-Avoidance (0.884)		
Item 10 (FABQ-PA)	Fear-Avoidance	Fear-Avoidance (0.802)	Fear-Avoidance	Negative Coping (0.560)
Item 11 (FABQ-W)	Fear-Avoidance	Catastrophizing (0.634)	Fear-Avoidance	Negative Coping (0.606)
Item 12 (PASS-20)	Fear-Avoidance	Catastrophizing (0.728)		
Item 13 (PSEQ)	Positive Affect/Coping	Positive Affect/Coping (0.782)		
Item 14 (CPAQ)	Positive Affect/Coping	Positive Affect/Coping (0.696)	Positive Affect/Coping	Positive Affect/Coping (0.824)
Item 15 (CPAQ)	Positive Affect/Coping	Positive Affect/Coping (0.861)	Positive Affect/Coping	Positive Affect/Coping (0.805)
Item 16 (CPAQ)	Positive Affect/Coping	---		
Item 17 (SER)	Positive Affect/Coping	Positive Affect/Coping (0.756)	Positive Affect/Coping	Positive Affect/Coping (0.800)

\* All item numbers correspond to OSPRO-YF 17-item version question numbers; grayed areas indicate items that are not part of the 10-item version

\*\* Cross-loaded item; not assigned to a factor

--- highest loading did not meet 0.4 threshold; not assigned to a factor

PHQ-9=Patient Health Questionnaire; STAI=State-Trait Anxiety Inventory; STAXI=State-Trait Anger Expression Inventory; TSK-11=Tampa Scale for Kinesiophobia; PCS=Pain Catastrophizing Scale; FABQ=Fear Avoidance Beliefs Questionnaire (PA=Physical Activity, W=Work); PASS-20=Pain Anxiety Symptom Scale; PSEQ=Pain Self Efficacy Questionnaire; CPAQ=Chronic Pain Acceptance Questionnaire; SER= Self-Efficacy for Rehabilitation.

**Table 2.**

Methods for Calculating OSPRO-YF Factor Sub-scores

<b>17-item OSPRO-YF</b>		
	<b>Calculation</b>	<b>Included Items</b>
<b>Catastrophizing Factor Sub-Score</b>	Sum of items loaded on Catastrophizing factor	7, 8, 11, 12
<b>Negative Mood Factor Sub-Score</b>	Sum of items loaded on Negative Mood factor	3, 4, 5, 6
<b>Fear-Avoidance Factor Sub-Score</b>	Sum of items loaded on Fear-Avoidance factor	9, 10
<b>Positive Affect/Coping Factor Sub-Score</b>	Sum of items loaded on Positive Affect/Coping factor	2, 13, 14, 15, 17
<b>10-item OSPRO-YF</b>		
	<b>Calculation</b>	<b>Included Items</b>
<b>Positive Affect/Coping Factor Sub-Score</b>	Sum of items loaded on the Positive Affect/Coping factor*	14, 15, 17
<b>Negative Coping Factor Sub-Score</b>	Sum of items loaded on the Negative Coping factor	7, 8, 10, 11
<b>Negative Mood Factor Sub-Score</b>	Sum of items loaded on the Negative Mood factor	1, 3, 4

\* Indicated items for the Positive Affect/Coping Factor have the same directionality and *do not* need to be reversed for measurement and interpretation of this sub-score.

**Table 3.**

Contribution of 17-item OSPRO-YF Factor Sub-Scores to Pain, Disability, and Quality of Life Measure Scores

	Model		Independent Variables	df1, df2	B (Unstandardized Coefficient)	$\beta$ (Standardized Coefficient)	p-value
Average Daily Pain	Base Model R <sup>2</sup> =0.22	Block 1	Catastrophizing Factor	4, 435	0.593	0.373	<0.001 *
			Negative Mood Factor		-0.192	-0.060	0.168
			Fear Avoidance Factor		-0.101	-0.063	0.176
			Positive Affect/ Coping Factor		-0.253	-0.220	<0.001 *
	<i>Interaction Model: No significant interactions</i>						
Disability	Base Model R <sup>2</sup> =0.37	Block 1	Catastrophizing Factor	4, 435	0.108	0.403	<0.001 *
			Negative Mood Factor		-0.035	-0.065	0.097
			Fear Avoidance Factor		0.046	0.172	<0.001 *
			Positive Affect/ Coping Factor		-0.044	-0.227	<0.001 *
	<i>Interaction Model: No significant interactions</i>						
Mental Quality of Life	Base Model R <sup>2</sup> =0.37	Block 1	Catastrophizing Factor	4, 435	-0.789	-0.323	<0.001 *
			Negative Mood Factor		-1.547	-0.318	<0.001 *
			Fear Avoidance Factor		0.189	0.077	0.067
			Positive Affect/ Coping Factor		0.411	0.233	<0.001 *
	<i>Interaction Model: No significant interactions</i>						
Physical Quality of Life	Base Model R <sup>2</sup> =0.24	Block 1	Catastrophizing Factor	4, 435	-0.724	-0.319	<0.001 *
			Negative Mood Factor		0.626	0.138	0.001 *
			Fear Avoidance Factor		-0.169	-0.073	0.109
			Positive Affect/ Coping Factor		0.410	0.249	<0.001 *
	<i>Interaction Model: No significant interactions</i>						

\* significant (alpha=0.01)

**Table 4.**

Contribution of 10-item OSPRO-YF Factor Sub-Scores to Pain, Disability, and Quality of Life Measure Scores

	Model		Independent Variables	df1, df2	B (Unstandardized Coefficient)	$\beta$ (Standardized Coefficient)	p-value
	<b>Average Daily Pain</b>	<i>Base Model</i> R <sup>2</sup> =0.15	Block 1	Positive Affect/ Coping Factor	3, 436	-0.296	-0.190
Negative Coping Factor				0.385		0.257	<0.001 *
Negative Mood Factor				0.350		0.084	0.068
<i>Interaction Model: No significant interactions</i>							
<b>Disability</b>	Model		Independent Variables	df1, df2	B	$\beta$	p-value
	<i>Base Model</i> R <sup>2</sup> =0.33	Block 1	Positive Affect/ Coping Factor	3, 436	-0.041	-0.159	<0.001 *
			Negative Coping Factor		0.126	0.501	<0.001 *
			Negative Mood Factor		0.033	0.047	0.247
<i>Interaction Model: No significant interactions</i>							
<b>Mental Quality of Life</b>	Model		Independent Variables	df1, df2	B	$\beta$	p-value
	<i>Base Model</i> R <sup>2</sup> =0.34	Block 1	Positive Affect/ Coping Factor	3, 436	0.466	0.196	<0.001 *
			Negative Coping Factor		-0.459	-0.200	<0.001 *
			Negative Mood Factor		-2.659	-0.419	<0.001 *
<i>Interaction Model: No significant interactions</i>							
<b>Physical Quality of Life</b>	Model		Independent Variables	df1, df2	B	$\beta$	p-value
	<i>Base Model</i> R <sup>2</sup> =0.19	Block 1	Positive Affect/ Coping Factor	3, 436	0.421	0.190	<0.001 *
			Negative Coping Factor		-0.756	-0.354	<0.001 *
			Negative Mood Factor		0.243	0.041	0.363
<i>Interaction Model: No significant interactions</i>							

\* significant (alpha=0.01)