Psychometric Assessment of Psychosocial Health Constructs: Measuring Psychological Pain and Disablement

A Dissertation
Presented in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
with a
Major in Education, Health and Human Sciences
in the
College of Graduate Studies
University of Idaho
by

Madeline Patrice Casanova

Major Professor: Russell T. Baker, Ph.D., DAT
Committee Members: Lindsay W. Larkins, DAT; Emma J. Grindley, Ed.D.;
Karen M. Appleby, Ph.D.
Department Administrator: Philip W. Scruggs, Ph.D.

August 2020

Authorization to Submit Dissertation

This dissertation of Madeline Patrice Casanova, submitted for the degree of Doctor of Philosophy with a Major in Education, Health and Human Sciences and titled "Psychometric Assessment of Psychosocial Health Constructs: Measuring Psychological Pain and Disablement," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: _____ Date: _____
Russell T. Baker, Ph.D., DAT

Committee Members: _____ Date: _____
Lindsay W. Larkins, DAT

_____ Date: _____
Emma J. Grindley, Ed.D.

_____ Date: _____
Karen M. Appleby, Ph.D.

Department Administrator: _____ Date: _____
Philip W. Scruggs, Ph.D.

Abstract

Health care professionals are tasked with incorporating evidence-based practice (EBP) and assessing practice effectiveness, which can be accomplished by using health-related patient reported outcome (PRO) instruments. Clinicians use PRO instruments to assess a wide variety of health-related constructs (e.g., quality of life, pain, disablement) and must decide which instrument is relevant and important to use for patients. Many PRO instruments, however, have not been tested using the recommended contemporary psychometric analysis techniques. Therefore, the purpose of this dissertation was to assess the psychometrics of two psychosocial health instruments used in clinical practice and research: 1) Orbach and Mikulincer Mental Pain (OMMP) scale and 2) Disablement in Physically Active Short Form-8 (DPA-SF) scale.

The OMMP is an instrument designed to assess psychological pain; however, previous research has not yielded a consistent scale structure, and the internal consistency of the subscales have not met recommended values. Therefore, confirmatory factor analysis (CFA) was conducted on the 9-factor, 44-item OMMP. Model fit indices were not met and an exploratory factor analysis (EFA) was conducted to identify a more parsimonious (i.e., OMMP-9) structure. The OMMP-9 was then tested in a covariance model and refined further to create the OMMP-8. The OMMP-8 was then subjected to invariance testing between age groups, sex, activity classification, activity level, and injury status. The 9-factor, 44-item OMMP did not meet recommended measurement criteria and should not be recommended for use in research and clinical practice in its current form. The refined OMMP-8 met recommended measurement invariance criteria and may be a more viable option to use; however, more research should be completed prior to adoption.

The DPA SF-8 has been proposed as a tool to be used in the physically active population to assess a physical summary component (PHY) and a quality of life component (QOL) of disablement; however, analysis of scale structure has not been confirmed with a sample of individuals who have only answered the eight items included in the scale. Additionally, further scale development analyses (e.g., reliability, responsiveness, longitudinal invariance) to ensure psychometrics are sound have not been completed. Therefore, confirmatory factor analyses (CFAs) were conducted on the 2-factor, 8-item scale on each time point (i.e., visit) to ensure factor structure. Additionally, the reliability of the scale and internal consistency of the subscales were assessed; a minimal detectable change (MDC) value was calculated; minimal clinically important differences (MCID) were established; and, invariance testing across three visits and groups was conducted. The CFAs at all three visits exceeded recommended model fit indices, the intraclass correlation coefficient value (.924) calculated indicated excellent scale reliability, and Cronbach's alpha for subscales PHY and QOL were within recommend values. The MDC value calculated for summative scale scoring was 5.83 points, while the MCID values for persistent injuries was 2 points and 3 points for acute injuries. The DPA SF-8 was invariant across time and across subgroups. The DPA SF-8 met CFA recommendations and criteria for multi-group and longitudinal invariance testing, which indicates the scale may be used to assess for differences between the groups or across time. Our overall analysis indicates the DPA SF-8 is a valid, reliable, and responsive instrument to assess patient improvement in the physically active population.

Acknowledgements

Thank you to Dr. Russell Baker, I am incredibly grateful to have been your student. Thank you for fueling my intellectual curiosity, for encouraging me to step outside my comfort zone, and for providing a safe environment to test out my wings. Thank you for giving me the strength and courage I needed to fulfill this dream.

Thank you to my committee, Dr. Lindsay Larkins, Dr. Karen Appleby, and Dr. Emma Grindley. You brought enthusiasm, kindness, and intellectual prowess and were a constant source of inspiration and compassion. Thank you for the stimulating conversations and for providing shoulders to lean on.

Thank you to Dr. Tony Pickering. Your courses were instrumental to my development as a scholar, thank you for providing a challenging and enjoyable environment to learn. Thank you for your thoughtful feedback and your willingness to always let me pick your brain.

Thank you to Dr. Caroline Faure. Without you and Dr. Appleby, I would not have pursued this degree. You both saw something in me that I did not see in myself and I am so grateful that our paths crossed. You are an amazing teacher and an even better friend. I cannot thank you enough for the sticky notes with words of encouragement, the phone calls and text messages reminding me that I am not alone, and for being my advocate in life.

To Dr. Megan Nelson. This has been such a long journey and I am so grateful to have had you running beside me; I would not have crossed the finish line without you. I am so lucky to have you as a colleague and as a friend. This is only the beginning; I look forward to collaborating and working together in the future.

To all of my fellow students and colleagues, in particular, Dr. Emily Miley, Dr. Ashley Reeves, and Aspen Ankney. It has been such a blessing to have each of you in my life. Thank you for providing the best cohort of support.

Lastly to the Casanova, Shea, and JKE families. In my time of need, without hesitation, you graciously offered your support. Without it, I would not have fulfilled this dream, thank you.

Dedication

To my husband, Zachary.

You are my rock, my inspiration, my happy place, my always; this journey was only possible because of your love, support, kindness, and compassion.

To my Sanderson Sisters, the Winnie and Mary to my Sarah.

You are my best friends, my confidants, my own personal calming circle; I am so blessed to have you in my corner, thank you for being my number one fans.

To the Harding Four.

During the worst parts of the storm, it was your fort that provided safety and comfort.

Thank you for being a constant light in the darkness.

To the Casanova Clan.

All of my best qualities were shaped from your unique influences, being a part of this incredible family is my greatest joy.

Thank you for teaching me how to survive in the lava.

To my parents, Rick and Peggy.

You taught me to dream big, work hard, and smile along the way.

Thank you for your love and support, I am so proud to be your daughter.

Table of Contents

List of Tables

List of Figures

Introduction

The use of patient reported outcome (PRO) instruments are increasingly important to assess practice effectiveness. Many instruments have been developed to assess psychosocial health outcomes; however, few have undergone contemporary psychometric analysis to ensure the instrument is sound and can be used in practice and research (Boateng et al., 2018). Clinicians and researchers should ensure proper psychometric analysis has been conducted regarding scale dimensionality, reliability, and validity (Boateng et al., 2018; Brown, 2014; Kline, 2015).

When developing or modifying a survey, items must be written and evaluated by experts for content, applicability, and readability and the scale must be preliminarily assessed in the target population (Boateng et al., 2018; Dillman, 2014). After item selection is completed, a scale must then be assessed for dimensionality (Boateng et al., 2018; Brown, 2014). To determine the number of dimensions in a scale, an exploratory factor analysis (EFA) can be conducted. Throughout the EFA process, inter-item correlations, correlations between constructs, and Cronbach's alpha on subscales of the solution are assessed and modifications are made if necessary (Boateng et al., 2018; Leech, 2014). Once a parsimonious solution has been identified, the scale is then tested in a new sample using confirmatory factor analysis (CFA). The CFA would ensure that the operationalization of the construct and factors identified are consistent with the scale structure (Boateng et al., 2018; Brown, 2014; Kline, 2015).

Following CFA, the scale should then undergo invariance testing between groups of interest and across time (Brown, 2014; Byrne, 2016; Van De Schoot et al., 2015). Invariance testing is necessary to determine if the association between the underlying latent constructs

and their respective items are stable and approximately equal across groups and across time (Brown, 2014; Byrne, 2016; Kline, 2015; Van De Schoot et al., 2015). An invariant scale would indicate that individuals of different groups or at different time points are interpreting the survey items and meanings of the items similarly, regardless of group membership or time, which confirms scores from the instrument truly correspond with the underlying constructs and are not due to group-specific or time attributions. Testing invariance is necessary to ensure the instrument can be used to compare hypothesized group differences (Kline, 2015). In addition to invariance testing, throughout scale development, reliability, validity (e.g., concurrent, convergent, discriminant), and responsiveness should be assessed.

Two psychosocial health instruments currently utilized are the Orbach and Mikulincer Mental Pain Scale (OMMP; Orbach et al., 2003) and the Disablement in Physically Active Short Form-8 Scale (DPA SF-8; Baker et al., 2019). Although preliminary analysis has been conducted on the scales, complete psychometric analysis has not been performed. For the OMMP, researchers have not identified a consistent scale structure (Guimarães et al., 2014; Heo, 2008; Tassani et al., 2019) and the internal consistency of the subscales has not met recommended values (Gvion et al., 2014; Levi et al., 2008; Levi-Belz et al., 2017; Soumani et al., 2011). For the DPA SF-8, preliminary analysis has supported the scale structure (Baker et al., 2019, in press); however, a scale structure must be assessed in a sample of individuals who only responded to the 8-items and invariance testing conducted between groups and across time. Additionally, scale reliability and responsive must be examined, and a minimal clinically important difference (MCID) value should be established. Therefore, the purpose of this research was to assess the psychometric properties of the OMMP and the DPA SF-8 to determine applicability in research and clinical practice.

References

Baker, R. T., Burton, D., Pickering, M. A., & Start, A. (2019). Confirmatory factor analysis of the Disablement in the Physically Active Scale and preliminary testing of short-form versions: A calibration and validation study. *Journal of Athletic Training*, *54*(3), 302–318. https://doi.org/10.4085/1062-6050-355-17

Baker, R. T., Casanova, M. P., Pickering, M. A., & Baker, J. (in press). Invariance testing of the Disablement in the Physically Active Scale. *Journal of Athletic Training*.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6,149*, 1-18.

Brown, T. A. (2014). *Confirmatory Factor Analysis for Applied Research* (Second Ed.). Guilford Publications.

Byrne, B. (2016). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming* (3rd ed.). Routledge.

Dillman, D. A. (2014). *Mail and Internet Surveys: The Tailored Design Method* (Fourth Ed.). Hoboken, New Jersey: John Wiley & Sons.

Guimarães, R., Fleming, M., & Cardoso, M. F. (2014). Validation of the Orbach & Mikulincer Mental Pain Scale (OMMP) on a drug addicted population. *Social Psychiatry and Psychiatric Epidemiology*, *49*(3), 405–415.

Heo, J. (2008). *Mental pain and suicide: Exploring a causal path among trauma, object relations, mental pain and suicidal ideation* [Doctoral Dissertation]. The California School of Professional Psychology.

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling* (Fourth ed.). Guilford Publications.

Leech, N. L., Barrett, K. C., Morgan, G. A., Barrett, K. C., & Morgan, G. A. (2014). *IBM SPSS for Intermediate Statistics* (Fifth Ed.). Routledge.

Levi, Y., Horesh, N., Fischel, T., Treves, I., Or, E., & Apter, A. (2008). Mental pain and its communication in medically serious suicide attempts: An "impossible situation." *Journal of Affective Disorders*, *111*(2–3), 244–250.

Levi-Belz, Y., Gvion, Y., Grisaru, S., & Apter, A. (2017). When the pain becomes unbearable: Case-control study of mental pain characteristics among medically serious suicide attempters. *Archives of Suicide Research*, *22*(3), 380–393. https://doi.org/10.1080/13811118.2017.1355288

Orbach, I., Mikulincer, M., Sirota, P., & Gilboa-Schechtman, E. (2003). Mental pain: A multidimensional operationalization and definition. *Suicide and Life-Threatening Behavior*, *33*(3), 219–230. https://doi.org/10.1521/suli.33.3.219.23219

Soumani, A., Damigos, D., Oulis, P., Masdrakis, V., Ploumpidis, D., Mavreas, V., & Konstantakopoulos, G. (2011). Mental pain and suicide risk: Application of the Greek version of the Mental Pain and the Tolerance of Mental Pain scale. *Psychiatriki*, *22*(4), 330–340.

Tossani, E., Ricci Garotti, M. G., Mikulincer, M., Giovagnoli, S., Calzolari, G., Landi, G., & Grandi, S. (2019). Psychometric evaluation of the Italian version of Orbach & Mikulincer mental pain scale in a non-clinical sample. *Current Psychology*. https://doi.org/10.1007/s12144-019-0128-4

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01064

Manuscript One: Measuring Psychological Pain: Psychometric Analysis of the Orbach and Mikulincer Mental Pain Scale

Abstract

Background: Suicide is a public health concern, with an estimated 1 million individuals dying by suicide each year worldwide. Several theories behind the meaning and motivation of suicide have been proposed; however, individual psychological pain is believed to be a contributing factor and has continued to be assessed. Therefore, establishing a psychometrically sound tool to adequately measure psychological pain may be valuable. The Orbach and Mikulincer Mental Pain scale (OMMP) has been proposed; however, previous psychometric analysis on the OMMP has not yielded a consistent scale structure, and the internal consistency of the subscales has not met recommended values. Therefore, the primary purpose of this study was to assess the psychometric properties of the OMMP in a diverse group of individuals.

Methods: A confirmatory factor analysis (CFA) was conducted on the 9-factor, 44-item OMMP. Because model fit indices were not met, an exploratory factor analysis (EFA) was conducted that yielded a more parsimonious structure. The EFA structure was then tested in a covariance model and multigroup invariance testing was subsequently performed.

Results: The CFA of the original 9-factor, 44-item OMMP did not meet recommended model fit indices. The EFA conducted yielded a more parsimonious scale (i.e., OMMP-9) structure. The OMMP-9 was then tested in a covariance model and refined further to create the OMMP-8. The OMMP-8 was then subjected to invariance testing between age groups, sex, activity classification, activity level, and injury status.

Conclusions: The 9-factor, 44-item OMMP does not meet recommended measurement criteria and should not be recommended for use in research and clinical practice in its current form. The refined OMMP-8 may be a more viable option to use; however, more research should be completed prior to adoption.

Introduction

Worldwide, an estimated 1 million individuals die by suicide each year (World Health Organization, 2019). In the United States, suicide ranks as the tenth leading cause of death (Heron, 2019). The rankings are more concerning when assessing causes of death by age group (Heron, 2019): suicide is the second, fourth, and eighth leading cause of death for individuals 10-34, 34-44, and 55-64 years of age, respectively. Additionally, rates of suicide have been dramatically increasing in the United States since 1999 (Stone et al., 2018). Therefore, a better understanding of suicide risk and subsequent prevention efforts continue to be of the utmost importance.

Although many meanings and motivations behind suicide have been documented (e.g., suffering pain from sickness or old age, military disasters or distress, political or social peril, stressful life events), the theory of personal agony has continued to receive attention from clinicians and researchers alike (Seidel, 1995). Leenaars (1996) wrote, "The enemy of life is [psychological] pain… it is the pain of feeling pain… the fear is that the trauma, the crisis is bottomless – an eternal suffering" (p. 224). The eternal suffering described is frequently heard by clinicians and captured in suicide notes with statements like "I can't stand the pain any longer" (Goldsmith et al., 2002). Although psychological pain (PsyPn) is extremely important to understand, the complexity and multifactorial nature of PsyPn has

resulted in challenges to conceptualize and measure the construct, thus creating significant gaps in the literature (Meerwijk & Shattell, 2012).

Nonetheless, over the last 100 years, several attempts to conceptualize PsyPn have been made. One of the first accounts can be traced back to Freud (1917), who associated PsyPn with an individual's feelings of mourning or melancholy following loss. Other researchers later elaborated and described PsyPn as feelings of suffering, emptiness, and a belief that the future was lost and no hope remained (Frankl, 1992). In the 1990s, the term *psychache* was coined to describe a model of intolerable PsyPn (Shneidman, 1998). Shneidman (1998) believed PsyPn was experienced due to frustrated or thwarted essential needs (e.g., to be loved, to protect one's image, avoid shame). The lack of essential needs caused individuals to experience a number of negative emotions (e.g., guilt, shame, defeat, hopelessness) and eventually led to a generalized experience of unbearable PsyPn. Subsequently, another model of PsyPn described by Bolger (1999), who labeled PsyPn as emotional pain, proposed that a traumatic event shattered an individual's personal identity and connection with others. The shattering left intense feelings of emotional pain, which was depicted as brokenness, woundedness, loss of self, feelings of disconnection, and the awareness of one's own negative attributes (Bolger, 1999).

Other terms, in addition to *psychache* and emotional pain, have also been used to describe PsyPn: suffering (Morse, 2011; Rehnsfeldt & Eriksson, 2004), mental pain (Orbach et al., 2003), and psychic pain (Yager, 2015). Literature reviews have been conducted on these terms and researchers have argued these terms refer to the same concept (Conejero et al., 2018; Meerwijk & Weiss, 2011); therefore, there was a call to unify the terms under the umbrella of 'psychological pain' (Meerwijk & Weiss, 2011). The recent unification efforts

led to the development of an accepted definition after careful examination of various concepts and models of PsyPn: "a lasting, unsustainable and unpleasant feeling resulting from negative appraisal of an inability or deficiency of the self" (Meerwijk & Weiss, 2011).

With a consensus definition established, there was a need to develop a psychometrically sound instrument to adequately measure PsyPn. A number of instruments to measure PsyPn have been proposed; however, each one has limitations and relatively few have undergone necessary psychometric analysis. The Psychological Pain Scale (Shneidman, 1999) requires participants to rate their PsyPn, rate perceived PsyPn of five pictures, identify three feelings prominent in their pain, and write an essay describing their PsyPn. Due to the complexity of the scale, a trained individual is needed to administer and interpret the results, and only modest scale reliability has been found (Leenaars & Lester, 2005). The Psychache Scale (Holden et al., 2001) was developed using constructs from the Psychological Pain Scale but it eliminated the need for a trained individual to administer the scale. The scale, condensed to 13-items, addressed frequency of PsyPn, but did not capture intensity of pain or the unpleasant or negative feelings associated with PsyPn. The Mee-Bunney Psychological Pain Assessment (Mee et al., 2011) was developed as a brief (i.e., 10-item scale) instrument to measure PsyPn, but the questions did not capture the unpleasant or negative feelings associated with PsyPn. Further, descriptions about scale development or testing of the scale structure were not identified in the literature.

The Orbach and Mikulincer Mental Pain Scale (OMMP) may be a more effective option because it was developed using more contemporary approaches (e.g., grounded theory and content analysis to develop questions, factor analysis to assess factor structure) and addressed some of the constraints associated with the other instruments (Orbach et al., 2003).

For example, the OMMP does not require a trained administrator and includes questions that assess both the intensity and dimensions of PsyPn (Orbach et al., 2003). The scale also includes more detailed questions regarding various cognitive and affective components of PsyPn (Pompili et al., 2008). The scale includes nine factors: experience of irreversibility, loss of control, narcissistic wounds, emotional flooding, freezing, estrangement, confusion, social distancing, and emptiness (Appendix B; Orbach et al., 2003). The OMMP has been administered in clinical populations (Conrad et al., 2009; Guimarães et al., 2014; Levi et al., 2008; Reisch et al., 2010; van Heeringen et al., 2010), college student samples (Heo, 2008; Orbach et al., 2003), and non-clinical community members (Soumani et al., 2011; Tossani et al., 2019). Researchers have primarily used the OMMP to evaluate relationships between PsyPn and depression, suicidal behavior, and anxiety.

Although assessing PsyPn, particularly between groups, is important for clinicians and researchers alike, instruments that have not undergone psychometric evaluation may not provide adequate, accurate, or reliable results. Thus, attempts to draw meaningful conclusions about scores from the instrument may not be recommended. The steps recommended to establish a psychometrically sound instrument include, but are not limited to: 1) assessing the proposed items and scale structure using exploratory factor analysis, 2) verification of the underlying dimensions and scale structure of the instrument using confirmatory factor analysis (CFA), and 3) assessing measurement invariance and population heterogeneity (Boateng et al., 2018; Brown, 2014; Kline, 2015). An established instrument will be generalizable and allow clinicians and researchers to adequately measure the constructs intended and reliably compare differences between groups and across time (Brown, 2014; Byrne, 2016; Kline, 2015).

A limited number of studies conducted on the OMMP have examined the psychometrics of the scale. A consistent scale structure using either CFA or exploratory factor analysis methods, however, has not been reported in the available literature (Table 1.1). For example, Guimarães et al., (2014) found a 5-factor, 24-item solution in a drug addicted sample of respondents. In contrast, Tossani et al. (2019) found a 5-factor, 31-item scale solution in a non-clinical sample (Table 1.1). Heo (2008) investigated the psychometrics in a Korean and US student population; in the Korean population, a 5-factor, 21-item solution was found, while a 5-factor, 20-item scale was found in the US student population (Table 1.1). Although a 5-factor solution was consistent across studies, the factors and items included in the final solutions were not identical (Table 1.1). The inconsistency between samples indicates the theoretical framework of the scale is not well-supported (Brown, 2014; Byrne, 2016; Kline, 2015).

Further, the reported internal consistency of the subscales (i.e., experience of irreversibility, loss of control, narcissistic wounds, and emotional flooding) in the initial scale development work (Orbach et al., 2003) exceed the recommended Cronbach's alpha value ≥ .90 (Leech et al., 2014; Streiner, 2003; Table 1.2). The high Cronbach's alpha values for the subscales may indicate multicollinearity, or redundancy among the items used within the subscales (Brown, 2014; Kline, 2015; Leech et al., 2014; McCrae et al., 2011; Streiner, 2003). Similarly, the social distancing subscale was initially reported to have a Cronbach's alpha of .80; however, the items have not consistently met the recommended ≥ .70 level (Leech et al., 2014; Pesudovs et al., 2007) and the items have been removed from the final scale solution in subsequent research (Guimarães et al., 2014; Heo, 2008; Levinger et al., 2015; Tossani et al., 2019). Researchers who have used the items have reported alphas that

range from .34 to .42 (Gvion et al., 2014; Levi et al., 2008; Levi-Belz et al., 2017; Soumani et al., 2011). Thus, a reduction of items and/or subscales may be necessary to create a more parsimonious and psychometrically sound scale (Brown, 2014; Kline, 2015).

Despite the use of the OMMP in practice and research, complete and robust psychometric analysis of the scale has yet to be completed. There is a need to conduct a CFA to test the hypothesized factor structure of the OMMP, ensuring that the items are indirect measures of the hypothesized latent variables (Brown, 2014; Bryant & Yarnold, 1995). Additionally, the inconsistent psychometrics reported for the scale among different populations indicate the need for invariance testing in a diverse sample of individuals to ensure the scale is generalizable and unbiased towards different groups. Therefore, the primary purpose of this study was to assess the psychometric properties of the OMMP in a diverse group of individuals using CFA. Because model fit did not meet recommended levels, an EFA was conducted to establish a more parsimonious scale structure that was then tested in a rigorous covariance model. The secondary purpose was to conduct invariance testing between age groups, sex, activity classification, activity level, and injury status on the parsimonious scale structure identified.

Methods

The University Institutional Review Board approved the study and participants provided informed consent prior to beginning the survey. Emerging adults and adult participants (Sigelman & Rider, 2017) were recruited using a combination of convenience and snowball sampling methods. Members of the research team utilized personal contacts and social media pages to contact and advertise the study to participants. Additionally, participants were recruited using ResearchMatch (Harris et al., 2012), an online volunteer platform designed to match volunteers with researchers. Participants were able to complete

an electronic or paper version of the survey. The electronic survey was developed using Qualtrics Survey Software (Qualtrics Inc., Provo, UT) and the identical paper version of the survey was developed using Microsoft Word. Individuals who completed the electronic version were sent a link to the Qualtrics survey; paper copies were printed and distributed to those who opted to complete it by hand. The survey included the OMMP, a pain questionnaire, psychosocial questionnaires, and a participant demographic questionnaire.

Orbach and Mikulincer Mental Pain Scale

The Orbach and Mikulincer Mental Pain Scale (OMMP) consists of 44 items measuring nine unique factors. Factors include experience of irreversibility (nine items; e.g., the pain will never go away), loss of control (ten items; e.g., I have no control over the situation), narcissistic wounds (five items; e.g. I am rejected by everybody), emotional flooding (four items; e.g., There are strong ups and downs in my feelings), freezing (three items; e.g., I feel paralyzed), estrangement (three items; e.g., I am a stranger to myself), confusion (three items; e.g., I have difficulties in thinking), social distancing (four items; e.g., I don't feel like talking to other people), and emptiness (three items; e.g., I can't find meaning in my life). Participants rated each statement using a 5-point Likert scale (1 = Strongly disagree, 2 = Disagree, 3 = Agree to some extent, 4 = Agree, 5 = Strongly agree).

Pain Questionnaire

To assess physical pain, individuals completed the Numerical Pain Rating Scale (NPRS; Hartrick, Kovan & Shapiro, 2003). The NPRS (Appendix C) is used to assess intensity of physical pain; participants rated their best, worst, and current pain on a 0-10 scale (0 = no pain, 10 = worst pain possible). The pain scores reported for best, current, and worst were averaged to create a score representative of the patient's level of pain over 24 hours.

The NPRS has demonstrated good test-retest reliability (intraclass correlation coefficients ranging from .80 to .99) and high correlations were found between the NPRS and two other pain measures (Visual analog scale correlations range from .86 - .99; Verbal rating scale = .93), indicating good validity (Alghadir et al., 2018; Bijur et al., 2003; DeLoach et al., 1998; Hawker et al., 2011; Phan et al., 2012; von Baeyer et al., 2009).

Psychosocial Questionnaires

The Patient Health Questionnaire (PHQ-9) was utilized to assess depression (Spitzer et al., 1999). The PHQ-9 (Appendix D) includes 10-items, nine of which correspond with the diagnostic criteria for major depressive disorder. Participants rated each question on a 4-point Likert scale (0 = not at all; 3 = nearly every day), indicating how often each item had bothered them in the past two weeks. The PHQ-9 has reported high reliability and validity to measure presence and severity of depression in both clinical and general populations (Kocalevent et al., 2013; Kroenke et al., 2001; Manea et al., 2012; Martin et al., 2006). Items were then summed to create a composite score. The PHQ-9 has demonstrated good internal reliability with α ranging from .77 – 87 (Kocalevent et al., 2013; Löwe et al., 2004; Ślusarska et al., 2019; Urtasun et al., 2019; Villarreal-Zegarra et al., 2019). Construct validity has also been demonstrated by comparing the PHQ-9 to scales of quality of life, life satisfaction, emotional well-being, psychological well-being, and mental health (Keum et al., 2018; Kocalevent et al., 2013); convergent validity has been established by comparing the scale to other measures of depression (Löwe et al., 2004; Maroufizadeh et al., 2019). Additionally, responsiveness, (i.e., the validity of the PHQ-9 across time) has also been established (Löwe et al., 2004). Psychometric properties of the scale were assessed using CFA and multi-group invariance (e.g., sex, age, education level, ethnicity socioeconomic status) techniques;

researchers found the model met fit indices and passed invariance criteria, allowing for meaningful group comparisons (Galenkamp et al., 2017; Keum et al., 2018; Villarreal-Zegarra et al., 2019).

The Self-Compassion Scale (SCS) was utilized to assess self-compassion (Neff, 2003). The SCS (Appendix E) includes 26 items to measure six factors: self-kindness (e.g., I'm kind to myself when I'm experiencing suffering), self-judgement (e.g., When times are really difficult, I tend to be tough on myself), common humanity (e.g., I try to see my failings as part of the human condition), isolation (e.g., When I fail at something that's important to me I tend to feel alone in my failure), mindfulness (e.g., When something upsets me I try to keep my emotions in balance), and over-identification (e.g., When something upsets me I get carried away with my feelings). Participants indicated how often they acted in the manner stated in each of the items using a 5-point Likert scale (1 = almost never; 5 = almost always). Items in each factor were summed to create six subscale scores; all items were also summed to create a total score (Neff et al., 2019). The SCS has demonstrated good internal reliability with α ranging from .75 to .81 and test-retest reliability with α ranging from .80 to .88 (Neff, 2003). Psychometric properties of the scale were assessed using CFA and ESEM techniques across 20 samples; excellent fit was found for the six-factor solution (Neff et al., 2019). Additionally, predictive validity has also been demonstrated by comparing the SCS to scales of neuroticism, happiness, optimism, depression, stress, anxiety, and healthier physiological responses to stress (Breines et al., 2014; Finlay-Jones et al., 2015; Friis et al., 2016; Neff, 2003; Neff et al., 2007).

The Depression Anxiety Stress Scales-21 (DASS-21) was used to assess perceived psychological distress (Lovibond & Lovibond, 1995). The DASS-21 (Appendix F) includes

21 items assessing depression (e.g., I couldn't seem to experience any positive feeling at all), anxiety (e.g., I experienced breathing difficulty), and stress (e.g., I found it hard to wind down). Participants were asked to rate each statement, indicating how much the statement applied to them over the past week using a 4-point Likert scale (0 = did not apply to me at all; 1 = Applied to me to some degree, or some of the time; 2 = Applied to me a considerable degree, or a good part of the time; 3 = applied to me very much, or most of the time). Items from each subscale were summed to create composite scores, with the cumulative score representing psychological distress. The DASS-21 has demonstrated good internal reliability with α ranging from .73 to .87 (Lovibond & Lovibond, 1995; Osman et al., 2012) and good test re-test reliability with α ranging from .77 to .89 (Asghari et al., 2008). Convergent validity has been established by comparing the scale to anxiety, depression, and stress scales (Asghari et al., 2008; Bottesi et al., 2015; Lovibond & Lovibond, 1995; Osman et al., 2012; Tonsing, 2014), and construct validity of a 3-factor model using EFA and CFA techniques was also established (Osman et al., 2012; Tonsing, 2014).

Participant Questionnaire

A participant questionnaire (Appendix G) was created to collect demographic data including sex, ethnicity, age, highest level of education, physical activity level, diagnosis of a mental illness, and injury status.

Data Analysis

A member of the research team input paper survey responses into Qualtrics. All data was then exported from Qualtrics for analysis into the Statistical Package for Social Sciences Version 26 (SPSS, Inc., Chicago, IL). Missing responses were calculated for the OMMP and individuals missing 4 or more items (i.e., 10%) were removed from the dataset. Individuals

missing less than 10% of the items (i.e., 3 items or less) were retained, and missing data was replaced with the rounded mean score of the respective item (Kline, 2015). Because the primary purpose was to assess the OMMP, individuals were not excluded if they were missing demographic information or responses to other instruments included in the survey packet. Continuous variables were reported as (mean±SD) and categorical variables were reported as (n; percentage).

Histograms and skewness and kurtosis values were used to assess for normality of the data. Univariate outliers were assessed and removed if the z-scores exceeded the cut-off value of |3.3|. Multivariate outliers were also assessed and individuals were removed if the Malahanobis distance, identified using a chi-square table with degrees of freedom and p-value of .01 (Kline, 2015), was exceeded. After assessment of normality and outliers, the full sample was used to conduct a CFA using maximum likelihood estimation. Because model fit did not meet recommended guidelines (Bryant & Yarnold, 1995; Kline, 2015), the full sample was randomly split into two datasets (n1, n2). To identify a more parsimonious solution, an EFA was conducted on sample n1. The solution found during the EFA process was then tested in a more rigorous covariance model approach (Kline, 2015) using sample n2 and further refinement led to the creation of a refined model. A latent variable model was then assessed between the refined OMMP and the original OMMP, to assess the amount of variance accounted for in the new solution. The refined OMMP covariance model then underwent multigroup invariance testing. Invariance testing was conducted across sex, age groups, activity classification, activity level, and injury status. Finally, latent variable correlations were performed to assess the relationships between the refined OMMP, the pain questionnaire, and the psychosocial measures.

Confirmatory Factor Analysis

To test the factorial validity of the original 9-factor, 44-item scale, a CFA using maximum likelihood estimation was conducted on the full sample using the Analysis of Moment Structures (AMOS) Version 26 software (IBM Corp., Armonk, NY). Overall goodness of fit was evaluated by assessing the likelihood ratio statistic (Chi-square or CMIN), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Bollen's Incremental Fit Index (IFI; Bryant & Yarnold, 1995; Hu & Bentler, 1999; Kline, 2015). Because the Chi-square statistic is heavily influenced by sample size, it was not used as a primary assessment of model fit; instead model fit was deemed acceptable if contemporary criteria were met CFI ≥ .95, TLI ≥ .95, RMSEA ≤ .06, and IFI ≥ .95. In addition to assessing overall goodness of fit, localized areas of strain in the solution were assessed, and the interpretability, size, and statistical significance of the model's parameter estimates (i.e., factor variances, covariances, and indicator errors) were reviewed (Brown, 2014).

Exploratory Factor Analysis

The dataset was randomly split into two samples (n1, n2) and a maximum likelihood extraction EFA with direct oblimin rotation was conducted on sample n1. Three criteria were utilized to determine the number of factors retained: 1) factors with an eigenvalue > 1.0, 2) scree plot inflexion point examination, and 3) factors that accounted for more than 5% of the variance (Brown, 2014; Hayton et al., 2004; Leech et al., 2014; Schönrock-Adema et al., 2009). Assessment of Bartlett's test for sphericity (<.001) and Kaiser-Meyer Olkin Measure of Sampling Adequacy (≥.70) for sampling adequacy were checked for violations (Leech et al., 2014). Following extraction, items were assessed individually and removed one at a time

until a parsimonious solution was found. Items were removed using commonly accepted recommendations: loading < .40, cross-loading ≥ .30, high bivariate correlations with another item in the scale, poor theoretical or conceptual fit, and/or the item contributed to low internal consistency (Brown, 2014; Leech et al., 2014; Pesudovs et al., 2007; Streiner, 2003). Lastly, Cronbach's alpha was assessed on each factor and set *a priori* at ≥ .70 and ≤ .89 (Leech et al., 2014; Morgado et al., 2018; Pesudovs et al., 2007).

Covariance Model

The parsimonious solution found during EFA was then tested using covariance modeling in sample n2. The same goodness-of-fit criteria that was utilized for the initial CFA was also used to assess acceptability of model fit for the covariance model (Brown, 2014; Kline, 2015). In addition, modification indices, factor loadings, and correlations between variables were observed. To determine if the refined version of the scale explained an acceptable amount of variance ($r \geq 0.90$; $R^2 = 0.81$; Raes et al., 2011) a correlational analysis was conducted on the scores of the OMMP and the refined OMMP.

Invariance Testing

Using the full sample, the refined model identified with sample n1 in EFA and tested in a covariance model with sample n2, was then subjected to multigroup invariance testing using the complete data set. AMOS (IBM Corp., Armonk, NY) software was utilized to perform the analysis across sex (i.e., male, female), age (i.e., emerging adults, adults), injury status (i.e., injured, healthy), activity level (inactive/low, moderate/high), and activity classification (i.e., individuals who participated in athletic activity, individuals who did not participate in athletic activity). Invariance testing is necessary to determine if the association between the underlying latent constructs (e.g., psychological pain, confusion, loss of control,

narcissistic wounds) and their respective items are stable and approximately equal across groups (Brown, 2014; Byrne, 2016; Kline, 2015; Van De Schoot et al., 2015). An invariant model ensures individuals of different groups are interpreting the survey items and meanings of the items similarly, regardless of group membership (e.g., male or female), which confirms scores from the instrument truly correspond with the underlying constructs and are not due to group-specific attributions. Instrument invariance is necessary to ensure the instrument can be used to compare hypothesized group differences (e.g., if females report higher mean scores on psychological pain than males).

Invariance testing involves a set of hierarchical steps with increasing levels of constraint (Brown, 2014; Byrne, 2016; Gregorich, 2006; Kline, 2015). First, individual CFAs by subgroup category (e.g., male and female, injured and healthy) were conducted, ensuring the operationalization of the construct and factors (e.g., confusion, irreversibility, social distancing) were present. Following individual CFAs, the model then underwent configural, metric, and scalar invariance. Configural invariance places both groups in the same model and ensures the same factors have identical items across groups (e.g., Emptiness has three items with substantial loadings in both males and females). Metric invariance tests if the factor loadings are equal across groups; thus, invariance at this step would ensure the meanings of the common factors are similar across groups. Finally, scalar invariance ensures that item intercepts are equal across groups, which implies the means are not driven or contaminated by outside factors (e.g., cultural norms, group specific attributes). Therefore, scalar invariance allows for means of the latent variables to be meaningfully compared across groups. If the model met metric invariance requirements, equal variances were assessed; if the model met scalar invariance requirements, equal mean models were tested. Model fit was

compared using the CFI difference test (CFI_{DIFF}) and the chi-square difference test (χ^2_{DIFF}), with a *p*-value cut-off of 0.01 (Brown, 2014; Byrne, 2016). The CFI_{DIFF} test held greater weight in decisions regarding model fit because the χ^2_{DIFF} test is sensitive to sample size (Brown, 2014; Kline, 2015). Therefore, if a model exceeded the χ^2_{DIFF} test but met the CFI_{DIFF} test, invariance testing proceeded.

Correlation Models

AMOS (IBM Corp., Armonk, NY) Version 26 was used to assess latent variable correlations between the second order refined OMMP and psychosocial questionnaires (i.e., PHQ-9, SCS, DASS-21). Additionally, correlations were assessed between the refined OMMP and subscales of the DASS-21 and the average NPRS pain score.

Results

A total of 1,535 individuals completed the survey. Seventy individuals were missing responses to more than 10% of the OMMP items and were removed from the data set. Three individuals were missing less than 10% of the OMMP; therefore, the missing values for those participants were replaced with the rounded mean for each item. A total of 97 individuals reported scores that indicated univariate (z scores ≥ 3.4) outliers, while an additional 217 reported scores that indicated multivariate outliers (Mahalanobis distance ≥ 68.71); these 314 participants were removed from the data set prior to analysis. A total of 1,151 participants, ages 18-95 (mean age = 41.01 ± 16.67), were retained for data analysis. Females accounted for 72.4% (n = 833) of the sample, while males accounted for 17.9% (n = 206). Participants were also grouped by injury classification, mental health diagnosis, education level, activity level, and by activity classification (Table 1.3).

Confirmatory Factor Analysis Orbach and Mikulincer Mental Pain Scale

The CFA of the 9-factor, 44-item OMMP goodness-of-fit indices did not meet recommended values (CFI = .856, TLI = .842, RMSEA = .072, IFI = .856; Figure 1.1). Factor loadings were significant and ranged from -.24 to .86; however, correlations between first-order latent variables (e.g., 'Irreversibility, 'Emptiness) were high, ranging from r = .52 to r = .94 (Table 1.4) and modification indices suggested a number of meaningful cross-loadings were present. The dataset was randomly split into two equal samples (n1 = 576, n2 = 575) for further analysis because of possible multicollinearity between first-order latent variables and overall model fit failing to meet recommended values. Sample n1 was used for EFA procedures, while sample n2 was used to assess fit of the refined solution in a covariance model.

Exploratory Factor Analysis

Initial EFA of the OMMP in sample n1 extracted four factors that accounted for 60.35% of the variance (Table 1.5). However, 14 items had low loadings or substantial cross-loadings. A total of 35 items were removed due to low loadings, high cross-loadings, inflated Cronbach's alpha levels, high inter-item correlation values, or lack of conceptual fit. Item removal resulted in a 3-factor, 9-item refined OMMP (i.e., OMMP-9) that accounted for 75.38% of the variance, contained items with loadings ≥ .43, and had Cronbach's alphas ranging from .767 - .856 (Table 1.6).

Factor 1 contained items 44, 29, and 32 that tapped into the belief that the experience is perpetual and retained the original label "Experience of Irreversibility." Factor 2 contained items 8, 35, and 14 and tapped into experiencing extreme emotions and feelings; it retained the original label "Emotional Flooding." Factor 3 contained items 7, 1, and 16 and tapped

into an individual's negative self-belief regarding social relationships and retained the original label "Narcissistic Wounds."

Covariance Model Refined OMMP-9

The covariance model of the OMMP-9 in sample n2 had improved model fit (Figure 1.2) with almost all goodness of-fit indices meeting recommended values (CFI = .968, TLI = .952, RMSEA = .076, IFI = .968; Hu & Bentler, 1999; Kline, 2005). Factor loadings were significant and ranged from .68 to .89, while the correlations between first-order latent variables (e.g., 'Irreversibility, 'Emptiness) were improved, ranging from r = .55 to r = .59. Modification indices indicated there was one item with meaningful cross-loadings, therefore further refinement of the model was performed. Item 32 was removed, which resulted in a 3-factor, 8-item scale (i.e., OMMP-8) with all model fit indices exceeding recommended values (CFI = .997, TLI = .995, RMSEA = .026, IFI = .997; Figure 1.3). Factor loadings were significant, ranging from .71 - .94, and moderate correlations between first-order latent variables (range = .52 - .59) were now found.

Participant scores for the original 44-item OMMP were highly correlated (r = .925, R^2 = .856) with participant scores from the OMMP-8. The high correlation value indicated participant responses on the OMMP-8 explained an acceptable amount of variance in responses on the original OMMP.

Invariance Testing of Refined OMMP-8

Invariance Analysis for Mental Health Diagnosis

Of the 1,151 individuals in the full sample, 1,029 (89.4%) reported history of mental health diagnosis (yes = 396, no = 633) and were used for analysis. The initial model (i.e., equal form) met all model fit indices (CFI = .988; χ^2 = 78.56; RMSEA = .036; Table 1.7).

The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test (CFI = .988) and the χ^2_{DIFF} test (χ^2 = 83.30). Because the metric model was invariant between groups, examination of the equal latent variable factors was warranted. The equal factor variance model passed the CFI_{DIFF} test (CFI = .978) and only slightly exceeded the χ^2_{DIFF} test (χ^2 = 122.23), indicating variances of the latent variables were equal between groups. The scalar model (i.e., equal intercepts) passed both the CFI_{DIFF} test (CFI = .985) and the χ^2_{DIFF} test (χ^2 = 100.20). Because the scalar model was invariant between groups, examination of the latent mean model was warranted. The equal latent means model did not pass the CFI_{DIFF} test (CFI = .956) or the χ^2_{DIFF} test (χ^2 = 129.60), indicating there were differences in means between groups. When means were not constrained to be equal, the group that reported a current or past mental health diagnosis exhibited substantially higher levels of psychological pain across all three constructs (i.e., experience of irreversibility, emotional flooding, and narcissistic wounds) than the group who reported no mental health diagnosis.

Invariance Analysis for Sex

Of the 1,151 individuals in the sample, 1,039 (90.3%) reported sex (male = 206, female = 833) and were used for analysis. The initial model (i.e., equal form) met all model fit indices (CFI = .987; χ^2 = 84.15; RMSEA = .038; Table 1.8). The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test (CFI = .988) and the χ^2_{DIFF} test (χ^2 = 86.61). Because the metric model was invariant between groups, examination of the equal latent variable factors was warranted. The equal factor variance model passed both the CFI_{DIFF} test (CFI = .988) and the χ^2_{DIFF} test (χ^2 = 89.75), indicating variances were equal between groups. The scalar model (i.e., equal intercepts) passed both CFI_{DIFF} test (CFI = .985) and the χ^2_{DIFF} test (χ^2 = 101.13). Because the scalar model was invariant between groups, examination of the

latent mean model was warranted. The equal latent means model passed the CFI_{DIFF} test (CFI = .978) and slightly exceeded the χ^2_{DIFF} test (χ^2 = 48.53), indicating there were no differences in means between groups.

Invariance Analysis for Injury Status

Of the 1,151 individuals in the sample, 1,050 (91.2%) reported injury status (healthy = 662, injured = 388) and were used for analysis. The initial model (i.e., equal form) met all model fit indices (CFI = .993; χ^2 = 59.49; RMSEA = .027; Table 1.9). The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test (CFI = .994) and the χ^2_{DIFF} test (χ^2 = 63.28). Because the metric model was invariant between groups, examination of the equal latent variable factors was warranted. The equal factor variance model did not pass the CFI_{DIFF} test (CFI = .961) or the χ^2_{DIFF} test (χ^2 = 190.45), indicating variances were not equal between groups. Examination of the variances when not constrained to be equal indicated that the injured group exhibited substantially more variance on the latent variable "Experience of Irreversibility" than the healthy group.

The scalar model (i.e., equal intercepts) passed both the CFI_{DIFF} test (CFI = .993) and the χ^2_{DIFF} test (χ^2 = 72.40). Because the scalar model was invariant between groups, examination of the latent mean model was warranted. The equal latent means model did not pass the CFI_{DIFF} test (CFI = .954) or the χ^2_{DIFF} test (χ^2 = 222.23), indicating there were differences in means between groups. When means were not constrained to be equal, the injured group reported higher levels of psychological pain in all three constructs (i.e., experience of irreversibility, emotional flooding, and narcissistic wounds) than the healthy group.

Invariance Analysis for Age

Of the 1,151 individuals in the sample, 1,047 (91.0%) reported age and were used for analysis. Individuals were grouped according to human developmental literature (Sigelman & Rider, 2009): emerging adulthood (ages 18-25; n = 211), early adulthood (ages 26-40; n = 388), middle adulthood (ages 41-65; n = 334), late adulthood (ages 66+; n = 114). The configural model (i.e., equal form) met all model fit indices (CFI = .993; χ^2 = 96.16; RMSEA = .020; Table 1.10). The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test (CFI = .993) and the χ^2_{DIFF} test (χ^2 = 244.59). Because the metric model was invariant between groups, examination of the equal latent variable factors was warranted. The equal factor variance model did not pass the CFI_{DIFF} test (CFI = .964) or the χ^2_{DIFF} test (χ^2 = 134.47), indicating variances were not equal between groups. Examination of the variances when not constrained to be equal indicated that the group variances differed across the three latent variables. The middle adulthood group exhibited substantially more variance on the latent variable "Experience of Irreversibility," and the late adulthood group exhibited substantially less variance on the latent variables "Emotional Flooding" and "Narcissistic Wounds."

The scalar model (i.e., equal intercepts) only slightly exceeded the CFI_{DIFF} test (CFI = .982) however, it passed the χ^2_{DIFF} test (χ^2 = 72.40) and met an additional recommendation of $RMSEA_{DIFF}$ test < .015 (RMSEA = .026; Chen, 2007), thus indicating the model was invariant between groups. Therefore, examination of the latent mean model was warranted. The equal latent means model did not pass the CFI_{DIFF} test (CFI = .940) or the χ^2_{DIFF} test (χ^2 = 341.65), indicating there were differences in means between age groups. When means were not constrained to be equal, the late adulthood group reported lower levels of psychological pain in latent constructs "Emotional Flooding" and "Narcissistic Wounds", while the middle

adulthood group exhibited higher levels of psychological pain in latent construct "Experience of Irreversibility" than the emerging and early adulthood groups.

Invariance Analysis for Activity Level

A total of 1,050 (91.2%) individuals in the sample reported activity level (inactive/low = 589, moderate/high = 461) and were used for analysis. The initial model (i.e., equal form) met all model fit indices (CFI = .995; χ^2 = 50.94; RMSEA = .022; Table 1.11). The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test (CFI = .996) and the χ^2_{DIFF} test (χ^2 = 55.33). Because the metric model was invariant between groups, examination of the equal latent variable factors was warranted. The equal factor variance model did not pass the CFI_{DIFF} test (CFI = .980) or the χ^2_{DIFF} test (χ^2 = 117.11), indicating variances were not equal between groups. Examination of the variances when not constrained to be equal indicated the inactive/low group exhibited substantially more variance on the latent variable "Experience of Irreversibility" than the healthy group.

The scalar model (i.e., equal intercepts) passed both the CFI_{DIFF} test (CFI = .995) and the χ^2_{DIFF} test (χ^2 = 62.75). Because the scalar model was invariant between groups, examination of the latent mean model was warranted. The equal latent means model did not pass the CFI_{DIFF} test (CFI = .974) or the χ^2_{DIFF} test (χ^2 = 145.27), indicating there were differences in means between groups. When means were not constrained to be equal, the inactive/low group reported higher levels of psychological pain in all three constructs (i.e., experience of irreversibility, emotional flooding, and narcissistic wounds) than the moderate/high activity group.

Invariance Analysis for Activity Classification

A total of 1,050 (91.2%) individuals in the sample reported activity classification (i.e., if they engaged in athletic, recreational, or occupational activities that require physical skills and use strength, power, endurance, speed, flexibility, range of motion, or agility at least 3 days per week) and were used for analysis (athletic activity = 455, no athletic activity = 595). The initial model (i.e., equal form) met all model fit indices (CFI = .991; $\chi^2 = 68.13$; RMSEA = .031; Table 1.12). The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test (CFI = .991) and the χ^2_{DIFF} test ($\chi^2 = 72.16$). Because the metric model was invariant between groups, examination of the equal latent variable factors was warranted. The equal factor variance model slightly exceeded the CFI_{DIFF} test (CFI = .980) and the χ^2_{DIFF} test ($\chi^2 = 116.38$), however passed the $RMSEA_{DIFF} < .015$, indicating variances were equal between groups.

The scalar model (i.e., equal intercepts) passed both the CFI_{DIFF} test (CFI = .990) and the χ^2_{DIFF} test ($\chi^2 = 82.58$). Because the scalar model was invariant between groups, examination of the latent mean model was warranted. The equal latent means model did not pass the CFI_{DIFF} test (CFI = .972) or the χ^2_{DIFF} test ($\chi^2 = 154.13$), indicating there were differences in means between groups. When means were not constrained to be equal, the group who did not participate in athletic activity reported higher levels of psychological pain in all three constructs (i.e., experience of irreversibility, emotional flooding, and narcissistic wounds) than the group who did participate in athletic activity.

Correlational Analysis

There was a significant correlation between the OMMP-8 and the latent variable models of the PHQ-9 ($R = .90$, $R^2 = .81$, $p < .001$), SCS ($R = -.85$, $R^2 = .72$, $p < .001$), and

DASS-21 ($R = .86$, $R^2 = .74$, $p < .001$). Correlations were also significant between the OMMP-8 higher-order model and the subscales of the DASS-21 (depression $R = .84$, $R^2 = .71$, $p < .001$; stress $R = .74$, $R^2 = .54$, $p < .001$; anxiety $R = .67$, $R^2 = .45$, $p < .001$), and the average NPRS pain score ($R = .56$, $R^2 = .32$, $p < .001$).

Discussion

Suicide is a public health concern, with an estimated 1 million individuals dying by suicide each year worldwide (World Health Organization, 2019). Several theories behind the meaning and motivation of suicide have been proposed; however, individual psychological pain is believed to be a contributing factor and has continued to be assessed (Seidel, 1995). Therefore, establishing a psychometrically sound tool to adequately measure psychological pain may be valuable. Previous psychometric analysis on the OMMP has not yielded a consistent scale structure (Guimarães et al., 2014; Heo, 2008; Tassani et al., 2019), and the internal consistency of the subscales has not met recommended values (Guimarães et al., 2014; Gvion et al., 2014; Heo, 2008; Levi et al., 2008; Levi-Belz et al., 2017; Soumani et al., 2011). Therefore, the primary purpose of this study was to assess the psychometric properties of the OMMP in a diverse group of individuals.

The CFA of the original 9-factor, 44-item OMMP did not meet recommended model fit indices. Therefore, an EFA was conducted to establish a more parsimonious scale (i.e., OMMP-9) structure. The OMMP-9 was then tested in a covariance model and refined further to create the OMMP-8. The OMMP-8 was then subjected to invariance testing between age groups, sex, activity classification, activity level, and injury status. The findings of our study suggest that the 9-factor, 44-item OMMP does not meet recommended measurement criteria and should not be recommended for use in research and clinical practice in its current form.

The refined OMMP-8 may be a more viable option to use; however, more research should be completed prior to adoption.

Confirmatory Factor Analysis

The original 9-factor scale structure was not supported in our study due to poor model fit indices and high latent variable correlations indicating many sub-dimensions were not measuring unique constructs. Our findings are consistent with previous research which failed to identify a consistent scale structure (Guimarães et al., 2014; Heo, 2008; Tassani et al., 2019). Correlations between first-order latent variables were moderate to very high (ranged from .52 to .94), indicating multicollinearity between factors and poor discriminant validity. Modification indices also suggested there were items with meaningful cross-loadings (i.e., items measured several factors), further suggesting multicollinearity and lack of distinction between factors. The inconsistent factor structure, poor model fit, validity concerns (i.e., factorial and discriminant), and possible multicollinearity provide evidence that the scale should not be used in its current format. Thus, scale refinement using alternate model generation was warranted to determine if a psychometrically sound version could be identified using the current items.

Refined OMMP Psychometric Analysis

An EFA was conducted in a calibration sample (i.e., n1) and a 9-item, 3-factor solution (i.e., OMMP-9) emerged. The nine items represented three of the original nine factors: three items from the "Experience of Irreversibility" factor, three items from "Emotional Flooding," and three items from "Narcissistic Wounds." The OMMP-9 was then subjected to covariance modeling procedures using the validation sample (i.e., n2). Although the model had improved fit, modification indices suggested further refinement could improve

model fit: item 32 (i.e., something in my life was damaged forever) was therefore removed from the model due to meaningful cross-loadings. The final model (i.e., OMMP-8) retained eight of the original items and represented three distinct factors (i.e., experience of irreversibility, emotional flooding, and narcissistic wounds). The retained factors capture the essence of the definition (i.e., "a lasting, unsustainable and unpleasant feeling resulting from negative appraisal of an inability or deficiency of the self"; Meerwijk & Weiss, 2011).

Although the OMMP-8 only retained 18% of the questions from the original scale, participant responses were highly correlated ($r = .925$) with the original OMMP. Participant scores on the OMMP-8 accounted for a substantial amount of the variance ($r^2 = .856$) in the responses to the original 44-item OMMP (Raes et al., 2011). On average, participant scores for the OMMP-8 (mean = 1.99) were similar to those found in previous non-clinical samples (Gvion et al., 2014; Nahaliel, 2014; Tossani, 2019) and were lower than those found in clinical populations (Guimarães et al., 2014; Gvion et al., 2014; Levi et. al., 2008; Nahaliel, 2014).

The 3-factor structure identified in our sample, however, was not consistent with previous research that identified 5-factor structures in their samples (Guimarães et al., 2014; Heo, 2008; Orbach et al., 2003). The items included in the scale were also not consistent except for items 7, 14, 35, and 8 (Guimarães et al., 2014; Heo, 2008; Orbach et al., 2003; Tassani et al., 2019). Additionally, the only factor that has emerged across the five studies was "Emotional Flooding" (Guimarães et al., 2014; Heo, 2008; Orbach et al., 2003; Tassani et al., 2019). Although our study found a parsimonious model, more research should be done to ensure the scale structure identified is replicated in subsequent samples.

Refined OMMP-8 Invariance Testing

The OMMP-8 was then subjected to multigroup invariance testing. The OMMP-8 passed multigroup measurement invariance criteria for all group classifications: sex, injury status, activity level, mental health diagnosis, age, and activity classification. Thus, researchers can use the OMMP-8 to examine differences in psychological pain among these groups through a comparison of group mean scores. We did not identify group mean differences in psychological pain between males and females or between individuals who were healthy and injured on the OMMP-8. Our results are similar to previous research that did not identify differences between males and females in the subscales of "Irreversibility" and "Narcissistic Wounds;" however, our results also differ with previous research that identified group mean differences in "Emotional Flooding" between males and females (Tossani, 2019). Although no differences were found in our sample, subsequent research should continue to assess for differences as previous literature has indicated females exhibit higher levels of ruminate which contribute to higher rates of depression (Broderick & Korteland, 2003; Johnson & Whisman, 2013).

Group mean differences in psychological pain were identified between individuals with and without a current or past mental health diagnosis. Our results indicate individuals with a past or current mental health diagnosis exhibited substantially more psychological pain than those who did not have a past or current mental health diagnosis. This finding is consistent with previous research (Gvion et al., 2014; Levi et. al., 2008; Nahaliel, 2014) and provides further evidence of content validity for the OMMP-8 (Kline, 2015). Clinical populations have reported higher levels of psychological pain and previous researchers have found that scores on subscales of the OMMP can distinguish individuals based on the

likelihood they will engage in a high-risk suicide attempt (Levi-Beltz, 2017) or if they have suicidal tendencies (Nahaliel, 2014). Although these measures were not assessed in the present study, future research should assess the ability of the OMMP-8 to distinguish individuals with and without high suicide risk.

Group differences in variances and means for psychological pain were also found between activity level groups. Individuals who were classified as being inactive or engaging in low physical activity had substantially more variance (i.e., dispersion) in their responses and exhibited substantially more psychological pain than those who were active. Similarly, those who did not engage in athletic activity (i.e., athletic, recreational, or occupational activities requiring physical skills and use strength, power, endurance, speed, flexibility, range of motion, or agility at least 3 days per week) had substantially higher scores on psychological pain than those who did participate in athletic activity. Our results differ from previous research that found athletes respond different to psychosocial health (e.g., disablement, quality of life) constructs (Huffman et al., 2008; McAllister et al., 2001) however, they are consistent with previously reported findings, which indicate individuals who are physically active have higher satisfaction with life (Bendíková & Nemček, 2016; Melin et al., 2003), higher levels of quality of life (Anokye et al., 2012), and better psychosocial health outcomes (Dunton et al., 2007; Strine et al., 2008). Therefore, the more active an individual is, the lower the risk for poor psychosocial health outcomes, including psychological pain.

Lastly, differences in variances and means for psychological pain were also found between age groups. In our sample, when comparing total scores for the OMMP-8, the 65+ group had substantially less psychological pain (total score = 13.60) than all other groups

(emerging adults = 16.20, early adulthood = 15.99, middle adulthood = 16.57). Our finding is consistent with previous researchers who found that younger individuals exhibit higher levels of psychological pain than older individuals (Orbach et. al., 2003; Tossani, 2019) and that with age, there is a decrease in psychological distress (Wood et al., 2010). Further, older individuals are more effective and motivated at regulating emotions, particularly disengaging with negative material, which also decreases psychological distress (Rösler et al., 2004; Scheibe & Blanchard-Fields, 2009). Thus, as individuals age, they may report lower scores in psychological pain because there is a decline in frequency and duration of negative emotions and a more positive view on life is developed (Carstensen et al., 2003; Charles et al., 2003).

Latent Variable Correlational Analyses to support Construct Validity

The OMMP-8 was positively correlated with the PHQ-9 (R = .90), the DASS-21 (R = .86), the subscales of the DASS (R = .67 – R = 84), and negatively correlated with the SCS (R = -.85); the findings support the construct validity of the scale (Kline, 2015). The OMMP-8 was also positively correlated with the average NPRS score (R = .56). The correlations found in our study align with the multi-factorial definition of psychological pain as measured in the OMMP-8. Additionally, the positive correlations found between the OMMP-8, the DASS-21, and the DASS-21 subscales, are consistent with previous research (Guimarães et al., 2014; Orbach, 2003). Although the correlations between the OMMP-8 and the DASS-21 were slightly higher (r = .67 to .84) than those reported previously for the OMMP and DASS-21 (Guimarães et al., 2014), our model only included three factors, whereas the previous study included five factors of the OMMP. Thus, the reduction in factors and items may have led to the higher correlation value between the scales. More research on the psychometric properties of the OMMP-8, as well as the DASS-21, should be completed to

ensure the soundness of the psychometric properties of each scale and to ensure each is measuring a distinguishable experience.

Clinical Implications

Our research identified the OMMP-8 scale (Appendix H), which meets strict contemporary measurement criteria, to be recommended for use in research and clinical practice. The OMMP-8 scale met invariance testing recommendations which allow it to be administered in different groups (e.g., males and females, athletes and non-athletes) and allows for group differences to be interpreted as true differences instead of measurement error within the scale (Kline, 2015). Additionally, our findings indicated that respondents with a history of a current or past mental health diagnosis will score higher on the scale. Our results do not support using scores for diagnostic criteria at this time, however, they do provide insight into psychological pain and individual well-being of individuals, thus positively informing patient care. Lastly, although group comparisons are supported by the invariance testing findings, clinicians and researchers should be cautious using the OMMP-8 to assess change over time until the appropriate analyses (e.g., longitudinal invariance, scale responsiveness) have been completed.

Limitations and Future Research

Although our study included a diverse group of individuals, this study is not without limitations. The OMMP-8 was assessed using a cross-validation sample with our decision to split the samples; however, the sample used participants who responded to the original 44-item scale. Thus, the responses to the OMMP-8 items, could have been influenced by the other 35 items on the scale. Therefore, future research should be done on a sample of individuals who only respond to the eight items. Additionally, we found the OMMP-8 was

highly correlated with the PHQ-9 and DASS-21. Our findings could indicate refinement of the OMMP led to a more parsimonious scale which had greater overlap with the PHQ-9 and DASS-21. However, conducting similar measurement examination of the DASS-21 and PHQ-9 may also be warranted to ensure those scales meet similar contemporary recommendations and that scale refinement would not alter the resulting correlation values between scales. The psychometric properties of these scales were not assessed in our study and future research should conduct those analyses and re-assess the correlations between scales. Additionally, our findings could have been influenced due to the timing of the scale administration. Data collection occurred at the beginning stages of the COVID-19 pandemic. It is possible that individuals experienced elevated levels of psychological pain, depression, and psychological distress compared to normal, which may have subsequently impacted participant responses and the correlation values found between scales.

Although the OMMP-8 scale is a more parsimonious scale to assess psychological pain, more work should be done to validate the scale structure in new samples. More research should be performed with adolescents, as the rates in suicide have increased in that demographic dramatically (World Health Organization, 2019). Additionally, because it may be important for clinicians and researchers to assess change over time, reliability, responsiveness, minimal clinically important differences, and longitudinal invariance analyses should be conducted to ensure that the measurement properties of the scale are invariant over time (Kline, 2015). Lastly, we have to consider the purpose of this scale and its utilization. The OMMP was designed as a comprehensive instrument to assess the unique constructs of psychological pain. While participant scores on the OMMP-8 were highly correlated ($r = .925$) with the original OMMP, the elimination of so many items and factors

should be reviewed to ensure the refined tool captures the desired multi-factorial nature of psychological pain. Researchers may want to consider further analyses that correlate OMMP-8 responses (sub-dimensions and higher order latent variables) with other scales designed to measure relevant variables of psychological pain. Researchers may also want to consider adding novel items to tap into sub-constructs of psychological pain that are not captured in the OMMP-8. In particular, rewriting items to capture the respondent experience of "Emptiness" and "Loss of Control" should be examined because researchers have found individuals who attempt suicide score significantly higher in those dimensions (Levi-Belz, et al., 2017).

Conclusions

The original scale structure of the OMMP was not supported in our study. We subsequently identified a refined 3-factor, 8-item OMMP (i.e., OMMP-8) that met contemporary recommendations for model fit and multi-group invariance testing. Our findings support the OMMP-8 as a more viable option to assess psychological pain in research and clinical practice, but caution is warranted until more research is completed to further assess the measurement properties of the refined scale.

References

Alghadir, A., Anwer, S., Iqbal, A., & Iqbal, Z. (2018). Test–retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain. *Journal of Pain Research*, *Volume 11*, 851–856. https://doi.org/10.2147/JPR.S158847

Anokye, N. K., Trueman, P., Green, C., Pavey, T. G., & Taylor, R. S. (2012). Physical activity and health related quality of life. *BMC Public Health*, *12*(1), 624

Asghari, A., Saed, F., & Dibajnia, P. (2008). Psychometric properties of the Depression Anxiety Stress Scales-21 (DASS-21) in a non-clinical Iranian sample. *Iranian Psychological Association*, *2*(2), 82–102.

Bendíková, E., & Nemček, D. (2016). Life satisfaction in healthy people and people with noncommunicable diseases: Differences between active and inactive individuals. *Sport Science*, *9*(2), 19-23.

Bijur, P. E., Latimer, C. T., & Gallagher, E. J. (2003). Validation of a verbally administered Numerical Rating Scale of acute pain for use in the emergency department. *Academic Emergency Medicine*, *10*(4), 390–392. https://doi.org/10.1111/j.1553-2712.2003.tb01355.x

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6*. https://doi.org/10.3389/fpubh.2018.00149

Bolger, E. (1999). Grounded theory analysis of emotional pain. *Psychotherapy Research*, *9*(3), 342–362.

Bottesi, G., Ghisi, M., Altoè, G., Conforti, E., Melli, G., & Sica, C. (2015). The Italian version of the Depression Anxiety Stress Scales-21: Factor structure and psychometric properties on community and clinical samples. *Comprehensive Psychiatry*, *60*, 170–181. https://doi.org/10.1016/j.comppsych.2015.04.005

Breines, J. G., Thoma, M. V., Gianferante, D., Hanlin, L., Chen, X., & Rohleder, N. (2014). Self-compassion as a predictor of interleukin-6 response to acute psychosocial stress. *Brain, Behavior, and Immunity*, *37*, 109–114. https://doi.org/10.1016/j.bbi.2013.11.006

Broderick, P. C., & Korteland, C. (2002). Coping style and depression in early adolescence: Relationships to gender, gender role, and implicit beliefs. *Sex Roles*, *46*(7-8), 201-213.

Brown, T. A. (2014). *Confirmatory Factor Analysis for Applied Research* (Second Ed.). Guilford Publications.

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In *Reading and Understanding Multivariate Statistics* (pp. 99–136). American Psychological Association.

Byrne, B. (2016). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming* (3rd ed.). Routledge.

Carstensen, L. L., Fung, H. H., & Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and Emotion*, *27*(2), 103-123.

Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General*, *132*(2), 310-323.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464-504.

Conejero, I., Olié, E., Calati, R., Ducasse, D., & Courtet, P. (2018). Psychological pain, depression, and suicide: Recent evidences and future directions. *Current Psychiatry Reports*, *20*(5). https://doi.org/10.1007/s11920-018-0893-z

Conrad, A. K., Jacoby, A. M., Jobes, D. A., Lineberry, T. W., Shea, C. E., Arnold Ewing, T. D., Schmid, P. J., Ellenbecker, S. M., Lee, J. L., Fritsche, K., Grenell, J. A., Gehin, J. M., & Kung, S. (2009). A psychometric investigation of the suicide status form II with a psychiatric inpatient sample. *Suicide and Life-Threatening Behavior*, *39*(3), 307–320. https://doi.org/10.1521/suli.2009.39.3.307

DeLoach, L. J., Higgins, S., Caplan, A. B., & Stiff, J. L. (1998). The Visual Analog Scale in the immediate postoperative period: Intrasubject variability and correlation with a numeric scale. *Aneth Analg*, *86*, 102–106.

Dunton, G. F., Schneider, M., & Cooper, D. M. (2007). An investigation of psychosocial factors related to changes in physical activity and fitness among female adolescents. *Psychology and Health*, *22*(8), 929-944.

Finlay-Jones, A. L., Rees, C. S., & Kane, R. T. (2015). Self-compassion, emotion regulation and stress among australian psychologists: Testing an emotion regulation model of

self-compassion using structural equation modeling. *PLOS ONE*, *10*(7), 1–19. https://doi.org/10.1371/journal.pone.0133481

Frankl, V. (1992). *Man's Search for Meaning: An Introduction to Logotheraphy* (Fourth Edition). Beacon Press.

Freud, S. (1917). Mourning and Melancholia. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIV (1914-1916): On the History of the Psycho-Analytic Movement, Papers on Metapsychology and Other Works*.

Friis, A. M., Johnson, M. H., Cutfield, R. G., & Consedine, N. S. (2016). Kindness matters: A randomized controlled trial of a mindful self-compassion intervention improves depression, distress, and HbA1c among patients with diabetes. *Diabetes Care*, *39*(11), 1963–1971. https://doi.org/10.2337/dc16-0416

Galenkamp, H., Stronks, K., Snijder, M. B., & Derks, E. M. (2017). Measurement invariance testing of the PHQ-9 in a multi-ethnic population in Europe: The HELIUS study. *BMC Psychiatry*, *17*(1). https://doi.org/10.1186/s12888-017-1506-9

Goldsmith, S., Pellmar, T., Kleinman, A., & Bunney, W. (2002). *Reducing suicide: A national imperative*. National Academies Press.

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, *44*(11 (Suppl 3)), S78–S94.

Guimarães, R., Fleming, M., & Cardoso, M. F. (2014). Validation of the Orbach & Mikulincer Mental Pain Scale (OMMP) on a drug addicted population. *Social Psychiatry and Psychiatric Epidemiology*, *49*(3), 405–415. https://doi.org/10.1007/s00127-013-0751-6

Gvion, Y., Horresh, N., Levi-Belz, Y., Fischel, T., Treves, I., Weiser, M., David, H. S., Stein-Reizer, O., & Apter, A. (2014). Aggression–impulsivity, mental pain, and communication difficulties in medically serious and medically non-serious suicide attempters. *Comprehensive Psychiatry*, *55*(1), 40–50. https://doi.org/10.1016/j.comppsych.2013.09.003

Harris, P. A., Scott, K. W., Lebo, L., Hassan, N., Lightner, C., & Pulley, J. (2012). ResearchMatch: A National Registry to Recruit Volunteers for Clinical Research: *Academic Medicine*, *87*(1), 66–73. https://doi.org/10.1097/ACM.0b013e31823ab7d2

Hartrick, C. T., Kovan, J. P., & Shapiro, S. (2003). The Numeric Rating Scale for clinical pain measurement: A ratio measure? *Pain Practice*, *3*(4), 310–316. https://doi.org/10.1111/j.1530-7085.2003.03034.x

Hawker, G. A., Mian, S., Kendzerska, T., & French, M. (2011). Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF. *Arthritis Care & Research*, *63*(S11), S240–S252. https://doi.org/10.1002/acr.20543

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*(2), 191–205. https://doi.org/10.1177/1094428104263675

Heo, J. (2008). *Mental pain and suicide: Exploring a causal path among trauma, object relations, mental pain and suicidal ideation* [Doctoral Dissertation]. The California School of Professional Psychology.

Heron, M. (2019). Deaths: Leading Causes for 2017. *National Vital Statistics Reports*, *68*(6), 1–76.

Holden, R. R., Mehta, K., Cunningham, E. J., & McLeod, L. D. (2001). Development and preliminary validation of a scale of psychache. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *33*(4), 224–232. https://doi.org/10.1037/h0087144

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huffman, G. R., Park, J., Roser-Jones, C., Sennett, B. J., Yagnik, G., & Webner, D. (2008). Normative SF-36 values in competing NCAA intercollegiate athletes differ from values in the general population. *JBJS*, *90*(3), 471-476.

Johnson, D. P., & Whisman, M. A. (2013). Gender differences in rumination: A meta-analysis. *Personality and individual differences*, *55*(4), 367-374.

Keum, B. T., Miller, M. J., & Inkelas, K. K. (2018). Testing the factor structure and measurement invariance of the PHQ-9 across racially diverse U.S. college students. *Psychological Assessment*, *30*(8), 1096–1106. https://doi.org/10.1037/pas0000550

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling* (Fourth ed.). Guilford Publications.

Kocalevent, R.-D., Hinz, A., & Brähler, E. (2013). Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry*, *35*(5), 551–555. https://doi.org/10.1016/j.genhosppsych.2013.04.006

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Leech, N. L., Barrett, K. C., Morgan, G. A., Barrett, K. C., & Morgan, G. A. (2014). *IBM SPSS for Intermediate Statistics* (Fifth Ed.). Routledge.

Leenaars, A. A. (1996). Suicide: A multidimensional malaise. *Suicide and Life-Threatening Behavior*, *26*(3), 221–236.

Leenaars, A. A., & Lester, D. (2005). A note on Shneidman's Psychological Pain Assessment Scale. *OMEGA - Journal of Death and Dying*, *50*(4), 301–307. https://doi.org/10.2190/WH9X-80M3-NJ54-5GCU

Levi, Y., Horesh, N., Fischel, T., Treves, I., Or, E., & Apter, A. (2008). Mental pain and its communication in medically serious suicide attempts: An "impossible situation." *Journal of Affective Disorders*, *111*(2–3), 244–250. https://doi.org/10.1016/j.jad.2008.02.022

Levi-Belz, Y., Gvion, Y., Grisaru, S., & Apter, A. (2017). When the pain becomes unbearable: Case-control study of mental pain characteristics among medically serious suicide attempters. *Archives of Suicide Research*, *22*(3), 380–393. https://doi.org/10.1080/13811118.2017.1355288

Levinger, S., Somer, E., & Holden, R. R. (2015). The importance of mental pain and physical dissociation in youth suicidality. *Journal of Trauma & Dissociation*, *16*(3), 322–339. https://doi.org/10.1080/15299732.2014.989644

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck

Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

Löwe, B., Kroenke, K., Herzog, W., & Gräfe, K. (2004). Measuring depression outcome with a brief self-report instrument: Sensitivity to change of the Patient Health Questionnaire (PHQ-9). *Journal of Affective Disorders*, *81*(1), 61–66. https://doi.org/10.1016/S0165-0327(03)00198-8

Manea, L., Gilbody, S., & McMillan, D. (2012). Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis. *Canadian Medical Association Journal*, *184*(3), E191–E196. https://doi.org/10.1503/cmaj.110829

Maroufizadeh, S., Omani-Samani, R., Almasi-Hashiani, A., Amini, P., & Sepidarkish, M. (2019). The reliability and validity of the Patient Health Questionnaire-9 (PHQ-9) and PHQ-2 in patients with infertility. *Reproductive Health*, *16*(1). https://doi.org/10.1186/s12978-019-0802-x

Martin, A., Rief, W., Klaiberg, A., & Braehler, E. (2006). Validity of the brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *General Hospital Psychiatry*, *28*(1), 71–77. https://doi.org/10.1016/j.genhosppsych.2005.07.003

McAllister, D. R., Motamedi, A. R., Hame, S. L., Shapiro, M. S., & Dorey, F. J. (2001). Quality of life assessment in elite collegiate athletes. *The American Journal of Sports Medicine*, *29*(6), 806-810.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, *15*(1), 28–50. https://doi.org/10.1177/1088868310366253

Mee, S., Bunney, B. G., Bunney, W. E., Hetrick, W., Potkin, S. G., & Reist, C. (2011). Assessment of psychological pain in major depressive episodes. *Journal of Psychiatric Research*, *45*(11), 1504–1510. https://doi.org/10.1016/j.jpsychires.2011.06.011

Meerwijk, E. L., & Shattell, M. M. (2012). We need to talk about psychological pain. *Issues in Mental Health Nursing*, *33*(4), 263–265. https://doi.org/10.3109/01612840.2011.653046

Meerwijk, E. L., & Weiss, S. J. (2011). Toward a Unifying Definition of Psychological Pain. *Journal of Loss and Trauma*, *16*(5), 402–412. https://doi.org/10.1080/15325024.2011.572044

Melin, R., Fugl-Meyer, K. S., & Fugl-Meyer, A. R. (2003). Life satisfaction in 18-to 64-year-old Swedes: In relation to education, employment situation, health and physical activity. *Journal of Rehabilitation Medicine*, *35*(2), 84-90.

Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S., & Ferreira, M. E. C. (2018). Scale development: Ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica*, *30*(1). https://doi.org/10.1186/s41155-016-0057-1

Morse, J. (2011). The praxis theory of suffering. In *Philosophies and Theories for Advances Nursing Practices*. Jones & Bartlett Learning.

Nahaliel, S., Sommerfeld, E., Orbach, I., Weller, A., Apter, A., & Zalsman, G. (2014). Mental pain as a mediator of suicidal tendency: A path analysis. *Comprehensive Psychiatry*, *55*(4), 944-951.

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, *2*(3), 223–250. https://doi.org/10.1080/15298860309027

Neff, K. D., Rude, S. S., & Kirkpatrick, K. L. (2007). An examination of self-compassion in relation to positive psychological functioning and personality traits. *Journal of Research in Personality*, *41*(4), 908–916. https://doi.org/10.1016/j.jrp.2006.08.002

Neff, K. D., Tóth-Király, I., Yarnell, L. M., Arimitsu, K., Castilho, P., Ghorbani, N., Guo, H. X., Hirsch, J. K., Hupfeld, J., Hutz, C. S., Kotsou, I., Lee, W. K., Montero-Marin, J., Sirois, F. M., de Souza, L. K., Svendsen, J. L., Wilkinson, R. B., & Mantzios, M. (2019). Examining the factor structure of the Self-Compassion Scale in 20 diverse samples: Support for use of a total score and six subscale scores. *Psychological Assessment*, *31*(1), 27–45. https://doi.org/10.1037/pas0000629

Orbach, I., Mikulincer, M., Sirota, P., & Gilboa-Schechtman, E. (2003). Mental pain: A multidimensional operationalization and definition. *Suicide and Life-Threatening Behavior*, *33*(3), 219–230. https://doi.org/10.1521/suli.33.3.219.23219

Osman, A., Wong, J. L., Bagge, C. L., Freedenthal, S., Gutierrez, P. M., & Lozano, G. (2012). The Depression Anxiety Stress Scales-21 (DASS-21): Further Examination of Dimensions, Scale Reliability, and Correlates: Depression Anxiety Stress. *Journal of Clinical Psychology*, *68*(12), 1322–1338. https://doi.org/10.1002/jclp.21908

Pesudovs, K., Burr, J. M., Harley, C., & Elliott, D. B. (2007). The development, assessment, and selection of questionnaires. *Optometry and Vision Science*, *84*(8), 663–674. https://doi.org/10.1097/OPX.0b013e318141fe75

Phan, N., Blome, C., Fritz, F., Gerss, J., Reich, A., Ebata, T., Augustin, M., Szepietowski, J., & Ständer, S. (2012). Assessment of pruritus intensity: Prospective study on validity

and reliability of the Visual Analogue Scale, Numerical Rating Scale and Verbal Rating Scale in 471 patients with chronic pruritus. *Acta Dermato Venereologica*, *92*(5), 502–507. https://doi.org/10.2340/00015555-1246

Pompili, M., Lester, D., Leenaars, A. A., Tatarelli, R., & Girardi, P. (2008). Psychache and suicide: A preliminary investigation. *Suicide and Life-Threatening Behavior*, *38*(1), 116–121. https://doi.org/10.1521/suli.2008.38.1.116

Prysak, M. (2018). *Suicide vulnerability and risk: Fragmented sense of self and psychache* [Doctoral Dissertation]. The University of Adelaide.

Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical Psychology & Psychotherapy*, *18*(3), 250–255. https://doi.org/10.1002/cpp.702

Rehnsfeldt, A., & Eriksson, K. (2004). The progression of suffering implies alleviated suffering. *Scandinavian Journal of Caring Sciences*, *18*(3), 264–272.

Reisch, T., Seifritz, E., Esposito, F., Wiest, R., Valach, L., & Michel, K. (2010). An fMRI study on mental pain and suicidal behavior. *Journal of Affective Disorders*, *126*(1–2), 321–325. https://doi.org/10.1016/j.jad.2010.03.005

Rösler, A., Ulrich, C., Billino, J., Sterzer, P., Weidauer, S., Bernhardt, T., ... & Kleinschmidt, A. (2005). Effects of arousing emotional scenes on the distribution of visuospatial attention: Changes with aging and early subcortical vascular dementia. *Journal of the Neurological Sciences*, *229*, 109-116.

Scheibe, S., & Blanchard-Fields, F. (2009). Effects of regulating emotions on cognitive performance: What is costly for young adults is not so costly for older adults. *Psychology and Aging*, *24*(1), 217-223.

Schönrock-Adema, J., Heijne-Penninga, M., Van Hell, E. A., & Cohen-Schotanus, J. (2009). Necessary steps in factor analysis: Enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. *Medical Teacher*, *31*(6), e226–e232. https://doi.org/10.1080/01421590802516756

Seidel, G. (1995). Suicide in the elderly in antiquity. *International Journal of Geriatric Psychiatry*, *10*(12), 1077–1084. https://doi.org/10.1002/gps.930101216

Shneidman, E. S. (1998). Further reflections on suicide and psychache. *Suicide and Life-Threatening Behavior*, *28*(3), 245–250.

Shneidman, E. S. (1999). The Psychological Pain Assessment Scale. *Suicide and Life-Threatening Behavior*, *29*(4), 287–294.

Sigelman, C. K., & Rider, E. A. (2017). *Life-Span Human Development* (9th ed.). Cengage Learning.

Ślusarska, B., Nowicki, G., Piasecka, H., Zarzycka, D., Mazur, A., Saran, T., & Bednarek, A. (2019). Validation of the Polish language version of the Patient Health Questionnaire-9 in a population of adults aged 35–64. *Annals of Agricultural and Environmental Medicine*, *26*(3), 420–424. https://doi.org/10.26444/aaem/99246

Soumani, A., Damigos, D., Oulis, P., Masdrakis, V., Ploumpidis, D., Mavreas, V., & Konstantakopoulos, G. (2011). Mental pain and suicide risk: Application of the Greek version of the Mental Pain and the Tolerance of Mental Pain scale. *Psychiatriki*, *22*(4), 330–340.

Spitzer, R. L., Kroenke, K., & Williams, J. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *JAMA*, *282*(18), 1737. https://doi.org/10.1001/jama.282.18.1737

Stone, D. M., Simon, T. R., Fowler, K. A., Kegler, S. R., Yuan, K., Holland, K. M., Ivey-Stephenson, A. Z., & Crosby, A. E. (2018). Vital Signs: Trends in state suicide rates—United States, 1999–2016 and circumstances contributing to suicide—27 States, 2015. *Center for Disease Control and Prevention Morbidity and Mortality Weekly Report*, *67*(22), 617–624. https://doi.org/10.15585/mmwr.mm6722a1

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, *80*(1), 99–103. https://doi.org/10.1207/S15327752JPA8001_18

Strine, T. W., Chapman, D. P., Balluz, L. S., Moriarty, D. G., & Mokdad, A. H. (2008). The associations between life satisfaction and health-related quality of life, chronic illness, and health behaviors among US community-dwelling adults. *Journal of Community Health*, *33*(1), 40-50.

Tonsing, K. N. (2014). Psychometric properties and validation of Nepali version of the Depression Anxiety Stress Scales (DASS-21). *Asian Journal of Psychiatry*, *8*, 63–66. https://doi.org/10.1016/j.ajp.2013.11.001

Tossani, E., Ricci Garotti, M. G., Mikulincer, M., Giovagnoli, S., Calzolari, G., Landi, G., & Grandi, S. (2019). Psychometric evaluation of the Italian version of Orbach & Mikulincer mental pain scale in a non-clinical sample. *Current Psychology*. https://doi.org/10.1007/s12144-019-0128-4

Urtasun, M., Daray, F. M., Teti, G. L., Coppolillo, F., Herlax, G., Saba, G., Rubinstein, A., Araya, R., & Irazola, V. (2019). Validation and calibration of the patient health questionnaire (PHQ-9) in Argentina. *BMC Psychiatry*, *19*(1). https://doi.org/10.1186/s12888-019-2262-9

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01064

van Heeringen, K., Van den Abbeele, D., Vervaet, M., Soenen, L., & Audenaert, K. (2010). The functional neuroanatomy of mental pain in depression. *Psychiatry Research: Neuroimaging*, *181*(2), 141–144. https://doi.org/10.1016/j.pscychresns.2009.07.011

Villarreal-Zegarra, D., Copez-Lonzoy, A., Bernabé-Ortiz, A., Melendez-Torres, G. J., & Bazo-Alvarez, J. C. (2019). Valid group comparisons can be made with the Patient Health Questionnaire (PHQ-9): A measurement invariance study across groups by demographic characteristics. *PLOS ONE*, *14*(9), e0221717.

von Baeyer, C. L., Spagrud, L. J., McCormick, J. C., Choo, E., Neville, K., & Connelly, M. A. (2009). Three new datasets supporting use of the Numerical Rating Scale (NRS-11) for children's self-reports of pain intensity: *Pain*, *143*(3), 223–227. https://doi.org/10.1016/j.pain.2009.03.002

World Health Organization. (2019). Suicide in the world: Global health estimate. *World Health Organization*. https://apps.who.int/iris/bitstream/handle/10665/326948/WHO-MSD-MER-19.3-eng.pdf

Yager, J. (2015). Addressing patients' psychic pain. *American Journal of Psychiatry*, *172*(10), 939–943. https://doi.org/10.1176/appi.ajp.2015.14081012

Tables

Table 1.1

Exploratory Factor Analysis Solutions Validating the OMMP

Subjects	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 4**	**Sample 5**
Item	*Factor*	*Factor*	*Factor*	*Factor*	*Factor*
10. I have lost something that I will never find again.	IR	IR	IR	IR	IR
44. The pain will never go away.	IR	IR	IR	IR	*DNF*
22. The difficult situation will never change.	IR	IR	*DNF*	*DNF*	*DNF*
26. The world has changed forever.	IR	IR	IR	*DNF*	IR
30. My life has stopped.	IR	IR	*DNF*	*DNF*	*DNF*
32. Something in my life was damaged forever.	IR	IR	IR	IR	IR
34. I will never be the same person.	IR	IR	IR	*DNF*	IR
43. I can't change what is happening to me.	IR	IR	EMP	*DNF*	*DNF*
29. I will never be able to reduce my pain.	IR	IR	EMP	IR	*DNF*
6. I am afraid of the future.	LC	LC/FRZ	*DNF*	*DNF*	*DNF*
36. I have no control over the situation.	LC	*DNF*	*DNF*	*DNF*	IR
33. There is uncertainty about my life and myself.	LC	LC/FRZ	IR	*DNF*	IR
31. I have no idea what to expect of the future.	LC	LC/FRZ	*DNF*	*DNF*	*DNF*
13. I have no control over my life.	LC	*DNF*	*DNF*	*DNF*	*DNF*
2. I am completely helpless.	LC	*DNF*	HP	LC	HP
28. I have no control over what is happening inside me.	LC	*DNF*	EF	*DNF*	POW
9. I am completely defeated.	LC	*DNF*	*DNF*	*DNF*	*DNF*
5. I will fall apart.	LC	LC/FRZ	HP	LC	HP
21. I cannot trust myself.	LC	LC/FRZ	*DNF*	LC	POW
7. I am rejected by everybody.	NW	NW	*DNF*	NW	HP
12. I feel abandoned and lonely.	NW	NW	HP	*DNF*	*DNF*
1. Nobody is interested in me.	NW	NW	HP	NW	*DNF*
16. Others hate me.	NW	NW	*DNF*	NW	*DNF*
18. I am worthless.	NW	NW	*DNF*	*DNF*	*DNF*
14. My feelings change all the time.	EF	EF	EF	EF	EF
35. There are strong ups and downs in my feelings.	EF	EF	EF	EF	EF
3. I feel an emotional turmoil inside me.	EF	LC/FRZ	EF	EF	*DNF*
8. I am flooded by many feelings.	FRZ	EF	EF	EF	EF

11. I feel numb and not alive.	FRZ	LC/FRZ	*DNF*	*DNF*	*DNF*
19. I feel paralyzed	FRZ	LC/FRZ	*DNF*	LC	*DNF*
4. I cannot do anything at all	EST	LC/FRZ	HP	LC	HP
17. I feel that I am not my old self anymore.	EST	*DNF*	*DNF*	*DNF*	*DNF*
23. I feel as if I am not real.	EST	*DNF*	CON	*DNF*	*DNF*
15. I am a stranger to myself.	EST	NW	*DNF*	*DNF*	POW
20. I cannot concentrate.	CON	LC/FRZ	CON	LC	*DNF*
24. I have difficulties in thinking.	CON	*DNF*	CON	LC	POW
27. I feel confused.	CON	*DNF*	*DNF*	*DNF*	*DNF*
37. I want to be left alone.	SD	*NI*	*NI*	*NI*	*NI*
25. I need the support of other people. ®	SD	*NI*	*NI*	*NI*	*NI*
40. I don't feel like talking to other people.	SD	*NI*	*NI*	*NI*	*NI*
42. I can't stay alone. ®	SD	*NI*	*NI*	*NI*	*NI*
41. I can't find meaning in my life.	EMP	EMP	EMP	EMP	EMP
39. I have no desires	EMP	EMP	EMP	EMP	EMP
38. I have no future goals.	EMP	EMP	EMP	EMP	EMP

Factor Names: IR = Experience of irreversibility; LC = Loss of control; NW = Narcissistic wounds; EF = Emotional flooding; FRZ = Freezing; EST = Estrangement; CON = Confusion; SD = Social distancing; EMP = Emptiness; LC/FRZ = Lack of Control and Freezing; HP = Helplessness; POW = Powerlessness. Other Abbreviations and Symbols: NI = Item Not Included in Analysis; DNF = Item Did Not Factor. Sample 1 = 513 Israeli Jewish adults (Orbach et al., 2003); Sample 2 = 544 Italian adults (Tossani et al., 2019); Sample 3 = 403 drug addicted adults (Guimarães et al., 2014); Sample 4 = 427 Korean students (Heo, 2008); Sample 5 = 229 US students (Heo, 2008).

Table 1.2

Cronbach's Alpha Across Samples

Factors	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 4**	**Sample 5**
Experience of irreversibility*	.95	.90	.82	.81	.89
Loss of control*	.95	-	-	.90	-
Narcissistic wounds*	.93	.86	-	.81	-
Emotional flooding*	.93	.85	.80	.88	.80
Freezing*	.85	-	-	-	-
Estrangement*	.79	-	-	-	-
Confusion*	.80	-	.75	-	-
Social distancing*	.80	-	-	-	-
Emptiness*	.75	.81	.83	.82	.87
Lack of Control and Freezing		.92			
Helplessness			.78		.91
Powerlessness					.88

* indicates one of the original nine factors; Sample 1 = 513 Israeli Jewish adults (Orbach et al., 2003); Sample 2 = 544 Italian adults (Tossani et al., 2019); Sample 3 = 403 drug addicted adults (Guimarães et al., 2014); Sample 4 = 427 Korean students (Heo, 2008); Sample 5 = 229 US students (Heo, 2008).

Table 1.3

Demographic Data for the OMMP

Characteristics	N	%
Sex		
Male	206	17.9
Female	833	72.4
Prefer not to answer	8	0.7
Unknown	104	9.0
Education		
Some high school, no diploma	2	0.2
High school or GED	38	3.3
Some college, no degree	126	10.9
Associate degree	60	5.2
Bachelor's degree	281	24.4
Master's degree	385	33.4
Doctoral degree	133	11.6
Other	21	1.8
Unknown	105	9.1
Mental Health Diagnosis		
Yes	396	34.4
No	633	55.0
Prefer not to answer	18	1.6
Unknown	104	9.0
Ethnicity		
Caucasian	891	77.4
African American	54	4.7
Hispanic	62	5.4
Asian/Pacific Islander	58	5.0
Other	23	2.0
Unknown	63	5.5
Activity Level		
Inactive	179	15.6
Low	410	35.6
Medium	336	29.2
High	125	10.9
Unknown	101	8.8
Athletic Classification		
Competitive athlete	32	2.8
Recreational athlete	175	15.2
Occupational athlete	128	11.1
Activities of daily living	118	10.3
No athletic participation	595	61.7
Unknown	101	8.8
Injury Status		
Healthy	662	57.5
Acute injury	22	1.9
Sub-acute injury	27	2.3
Persistent injury	110	9.6
Chronic injury	229	19.9
Unknown	101	8.8

Table 1.4

Correlations Between First-Order Latent Variables OMMP

Factors	IRR	LOSS	NW	EF	FRZ	EST	CONF	SD	EMP
IRR	1.0								
LOSS	.89	1.0							
NW	.80	.89	1.0						
EF	.71	.82	.66	1.0					
FRZ	.83	.94	.92	.67	1.0				
EST	.84	.91	.89	.77	.90	1.0			
CONF	.76	.88	.69	.84	.78	.83	1.0		
SD	.67	.71	.72	.66	.69	.76	.71	1.0	
EMP	.77	.75	.76	.52	.79	.79	.61	.71	1.0

IRR = Experience of irreversibility; LOSS = Loss of control; NW = Narcissistic wounds; EF = Emotional flooding; FRZ = Freezing; EST = Estrangement; CONF = Confusion; SD = Social distancing; EMP = Emptiness

Table 1.5

Initial Exploratory Factor Analysis OMMP

Item	Factor 1	Factor 2	Factor 3	Factor 4
4. I cannot do anything at all.	**.812**	.050	.068	-.075
2. I am completely helpless.	**.773**	-.058	-.099	-.033
9. I am completely defeated.	**.768**	.044	-.080	-.037
18. I am worthless.	**.756**	-.095	-.166	.111
19. I feel paralyzed.	**.651**	.132	-.036	.072
21. I cannot trust myself.	**.616**	.233	.061	.083
11. I feel numb and not alive.	**.599**	.052	-.122	.197
7. I am rejected by everybody.	**.561**	.040	.000	.250
5. I will fall apart.	**.559**	.314	-.012	-.005
30. My life has stopped.	**.554**	.071	-.273	-.072
13. I have no control over my life.	**.536**	.240	-.093	-.019
22. The difficult situation will never change.	**.508**	.124	-.317	-.048
41. I can't find meaning in my life.	**.481**	.049	-.178	.314
23. I feel as if I am not real.	**.474**	.077	-.107	.175
15. I am a stranger to myself.	**.468**	.119	-.056	.297
39. I have no desires.	**.464**	-.152	-.239	.326
12. I feel abandoned and lonely.	**.460**	.107	-.173	.222
38. I have no future goals.	**.420**	-.177	-.321	.295
1. Nobody is interested in me.	**.419**	.007	.006	.365
16. Others hate me.	**.398**	.097	-.077	.229
28. I have no control over what is happening inside me.	**.392**	.313	-.227	.023
8. I am flooded by many feelings.	.032	**.747**	.044	.097
35. There are strong ups and downs in my feelings.	-.079	**.708**	-.156	.159
27. I feel confused.	.280	**.637**	.069	.011
3. I feel an emotional turmoil inside me.	.138	**.636**	-.018	.135
14. My feelings change all the time.	.075	**.621**	.031	.123
25. I need the support of other people. ®	.072	**-.592**	.054	.192
33. There is uncertainty about my life and myself.	.084	**.568**	-.161	.128
20. I cannot concentrate.	.244	**.559**	.099	.147
26. The world has changed forever.	-.056	**.548**	-.086	.026
31. I have no idea what to expect of the future.	.192	**.528**	-.053	.080
34. I will never be the same person.	-.212	**.502**	-.376	.239
6. I am afraid of the future.	.358	**.490**	.133	.064
24. I have difficulties in thinking.	.247	**.477**	-.021	.151
42. I can't stay alone. ®	-.234	**-.430**	.040	.104
17. I feel that I am not my old self anymore.	.012	**.375**	-.182	.295
36. I have no control over the situation.	.314	**.329**	-.269	-.139

44. The pain will never go away.	.089	-.023	**-.847**	.001
29. I will never be able to reduce my pain.	.227	-.018	**-.771**	-.056
43. I can't change what is happening to me.	.291	.240	**-.464**	-.104
32. Something in my life was damaged forever.	.004	.312	**-.453**	.200
10. I have lost something that I will never find again.	.005	.243	**-.375**	.222
37. I want to be left alone.	.061	.155	-.017	**.631**
40. I don't feel like talking to other people.	.153	.126	-.055	**.598**

Table 1.6

Refined OMMP-9 Exploratory Factor Analysis

Item	1	2	3
44. The pain will never go away.	.957		
29. I will never be able to reduce my pain.	.855		
32. Something in my life was damaged forever.	.425		
8. I am flooded by many feelings.		.847	
35. There are strong ups and downs in my feelings.		.826	
14. My feelings change all the time.		.763	
7. I am rejected by everybody.			.857
1. Nobody is interested in me.			.715
16. Others hate me.			.600
Eigenvalues	**4.54**	**1.22**	**1.02**
% of variance	**50.48**	**13.56**	**11.34**
Cronbach's alpha	**.835**	**.856**	**.767**

Table 1.7

Goodness-of-fit Indices for Measurement Invariance Analyses Across Mental Health Diagnoses OMMP-8

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Mental health diagnosis (n = 396)	12.69	17	----	1.00	----	1.00	.000
No mental health diagnosis (n = 633)	65.89	17	----	.979	----	.965	.067
Configural (equal form)	78.56	34	----	.988	----	.980	.036
Metric (equal loadings)	83.30	39	4.74(5)	.988	<.001	.983	.033
Equal factor variances	122.23	42	**43.67(8)**	.978	.01	.978	.043
Scalar (equal indicator intercepts)	100.20	44	21.64(10)	.985	.003	.981	.035
Equal latent means	208.16	47	**129.60(13)**	.956	**.032**	.948	.058

Table 1.8

Goodness-of-fit Indices for Measurement Invariance Analyses Across Sex OMMP-8

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Males (n = 206)	40.54	17	----	.970	----	.951	.082
Females (n = 833)	43.52	17	----	.992	----	.986	.043
Configural (equal form)	84.15	34	----	.987	----	.979	.038
Metric (equal loadings)	86.61	39	2.46(5)	.988	+.001	.983	.034
Equal factor Variances	89.75	42	5.60(8)	.988	+.001	.984	.033
Scalar (equal indicator intercepts)	101.13	44	16.99(10)	.985	.002	.981	.035
Equal latent means	132.68	47	**48.53(13)**	.978	.009	.974	.042

Table 1.9

Goodness-of-fit Indices for Measurement Invariance Analyses Across Injury Status

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Healthy (n = 662)	36.94	17	----	.992	----	.986	.042
Injured (n = 388)	22.55	17	----	.996	----	.994	.029
Configural (equal form)	59.49	34	----	.993	----	.989	.027
Metric (equal loadings)	63.28	39	3.79(5)	.994	+.001	.991	.024
Equal factor variances	190.45	42	**130.96(8)**	.961	**.021**	.948	.058
Scalar (equal indicator intercepts)	72.40	44	12.91(10)	.993	NC	.991	.025
Equal latent means	222.23	47	**162.74(13)**	.954	**.032**	.945	.060

Table 1.10

Goodness-of-fit Indices for Measurement Invariance Analyses Across Age Groups

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Emerging Adulthood (n=211)	33.61	17	----	.980	----	.967	.068
Early Adulthood (n=388)	7.54	17	----	1.00	----	1.00	.000
Middle Adulthood (n=334)	32.87	17	----	.988	----	.980	.053
Late Adulthood (n = 114)	22.04	17	----	.983	----	.973	.051
Configural (equal form)	96.16	68	----	.993	----	.988	.020
Metric (equal loadings)	123.78	83	27.62(15)	.993	.003	.986	.022
Equal factor variances)	230.63	92	**134.47(24)**	.964	**.029**	.957	.038
Scalar (equal indicator intercepts)	168.23	98	72.07(30)	.982	.011	.979	.026
Equal latent means	341.65	107	**245.49(39)**	.940	**.053**	.937	.046

Table 1.11

Goodness-of-fit Indices for Measurement Invariance Analyses Across Activity Level

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Inactive/Low (n = 589)	33.43	17	----	.992	----	.987	.041
Moderate/High (n = 461)	17.52	17	----	1.00	----	.999	.008
Configural (equal form)	50.94	34	----	.995	----	.993	.022
Metric (equal loadings)	55.33	39	4.39(5)	.996	+.001	.994	.020
Equal factor variances	117.11	42	66.17(8)	.980	**.015**	.973	.041
Scalar (equal indicator intercepts)	62.75	44	11.81(10)	.995	NC	.994	.020
Equal latent means	145.27	47	94.33(13)	.974	**.021**	.969	.045

Table 1.12

Goodness-of-fit Indices for Measurement Invariance Analyses Across Athletic Classification

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Athletic activity (n = 455)	27.96	17	----	.993	----	.989	.038
No athletic activity (n = 595)	40.14	17	----	.989	----	.982	.048
Configural (equal form)	68.13	34	----	.991	----	.985	.031
Metric (equal loadings)	72.16	39	4.03(5)	.991	NC	.987	.028
Equal factor variances	116.38	42	**48.25(8)**	.980	**.011**	.974	.041
Scalar (equal indicator intercepts)	82.58	44	14.45(10)	.990	.001	.987	.029
Equal latent means	154.13	47	**86.0(13)**	.972	**.019**	.966	.047

Figures

Figure 1.1

Confirmatory Factor Analysis Orbach and Mikulincer Mental Pain Scale

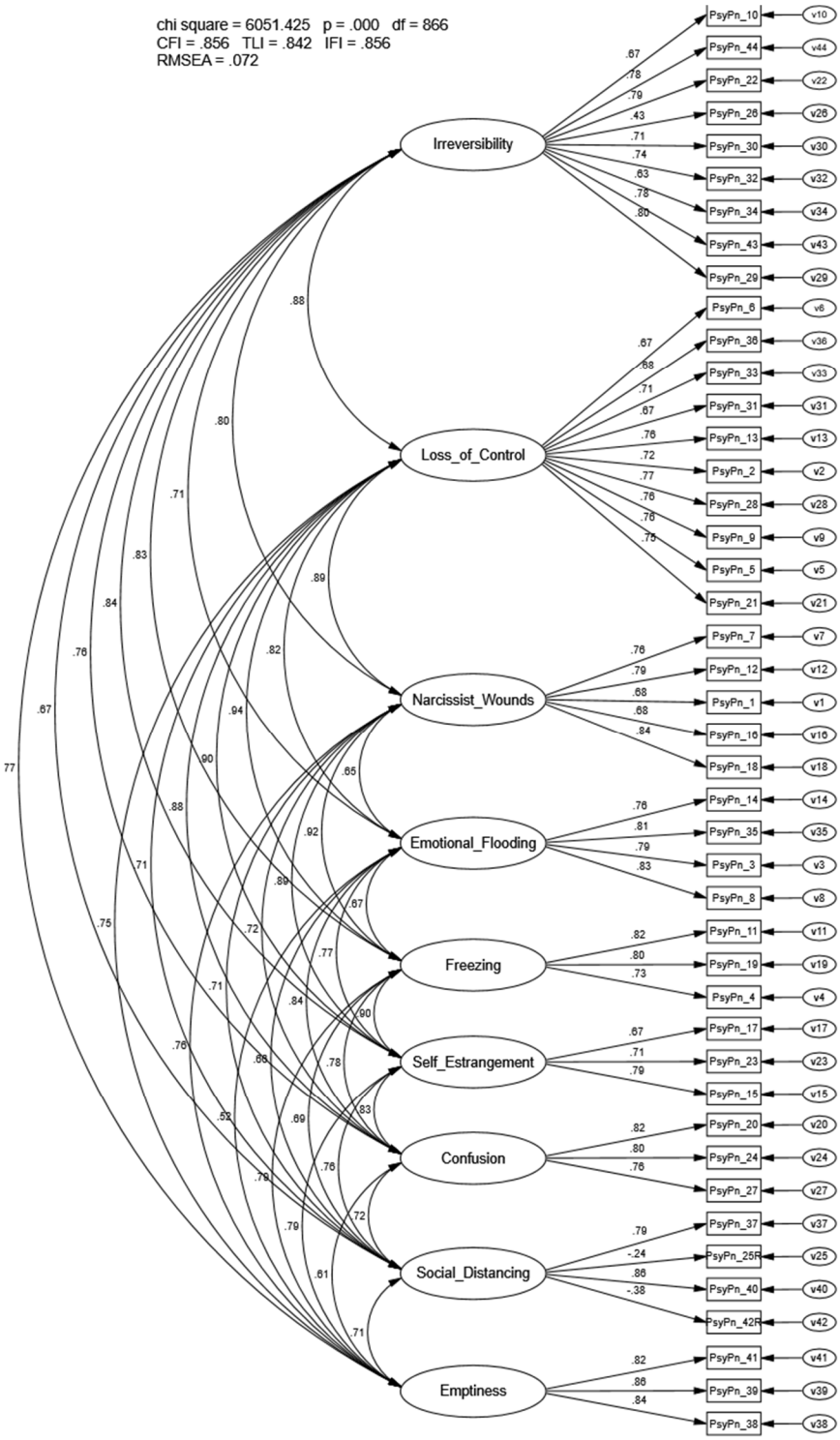


Figure 1.2

Covariance Model OMMP-9

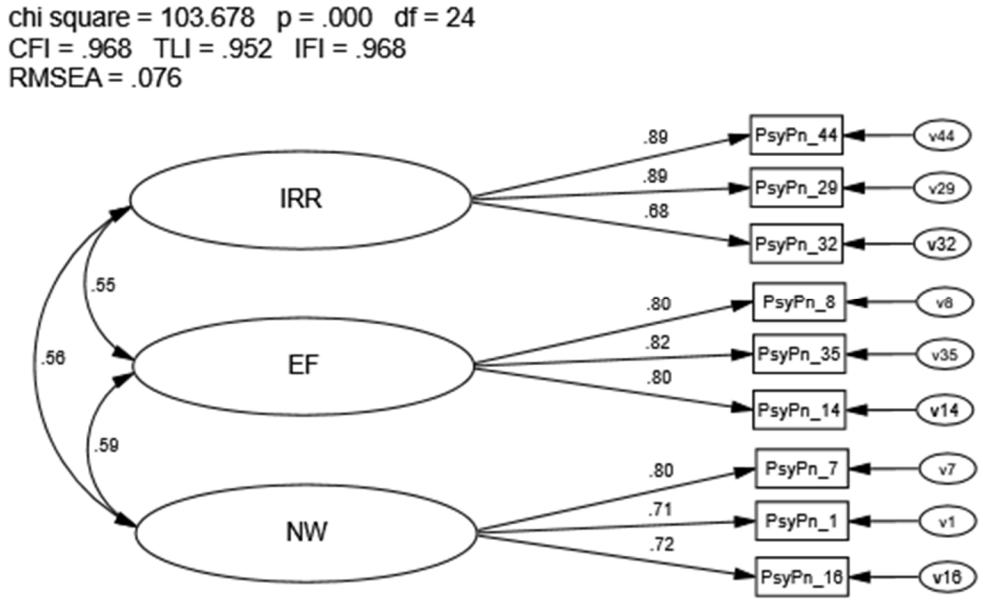
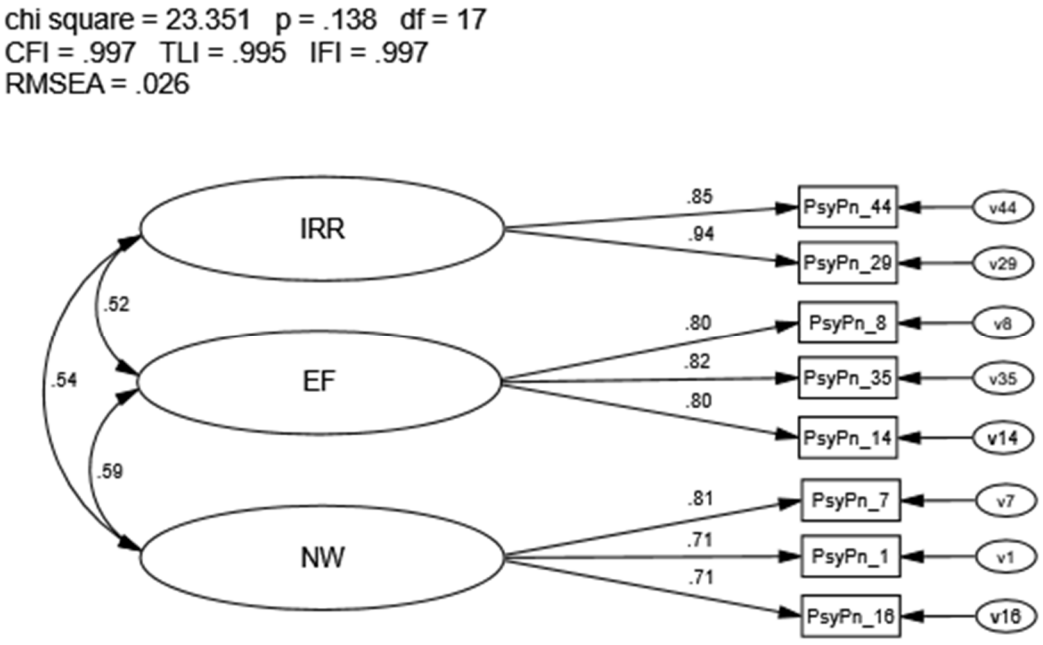


Figure 1.3

Covariance Model OMMP-8



Manuscript Two: Disablement in Physically Active Scale Short Form-8: Psychometric Evaluation

Abstract

Background: Patient-centered care and evidence-based practice (EBP) are core competencies for health care professionals. The importance of EBP has led to an increase in research involving clinical outcomes; current recommendations emphasize collecting patient focused measures (e.g., the patient's perspective and experience of their range of motion), thus increasing the need for psychometrically sound patient reported outcome measures (PROMs) of health. Disablement has been identified as a valuable multi-dimensional construct for patient care. The DPA SF-8 has been proposed as a tool to be used in the physically active population that assesses a physical summary component (PHY) of health and a quality of life component (QOL) however, further analysis is necessary to ensure the instrument is psychometrically sound.

Methods: Confirmatory factor analyses (CFAs) was conducted on the 2-factor, 8-item scale on each time point (i.e., visit) to ensure factor structure. Reliability of the scale and internal consistency of the subscales were assessed, and a minimal detectable change calculated. Additionally, minimal clinically important differences were also established, and invariance testing across three time points (i.e. visits) and groups was conducted.

Results: The CFAs at all three visits exceeded recommended model fit indices. The intraclass correlation coefficient value (.924) calculated indicated excellent scale reliability and Cronbach's alpha for subscales PHY and QOL were within recommend values. The

MDC value calculated was 5.83 and the MCID for persistent injuries were 2 points and for acute injuries, 3 points. The DPA SF-8 was invariant across time and across subgroups. **Conclusions:** The DPA SF-8 met CFA recommendations and criteria for multi-group and longitudinal invariance testing, which indicates the scale may be used to assess for differences between the groups or across time. Our overall analysis indicates the DPA SF-8 is a valid, reliable, and responsive instrument to assess patient improvement in the physically active population.

Introduction

Health care professionals have an ethical obligation to uphold core competencies, which includes providing patient-centered care and employing evidence-based practice (EBP; Institute of Medicine, 2003; National Athletic Trainers' Association, 2011). Engaging in EBP involves the integration of the best available research evidence coupled with clinical expertise and unique patient values and circumstances (Institute of Medicine, 2003; Law & MacDermid, 2014; Straus et al., 2019). The need for EBP has led to an increased emphasis on research involving clinical outcomes. Clinical outcomes may be measured using physiological or radiographic findings, patient self-report instruments, or a combination of objective clinical measures and patient-reported outcome measures (Arnold et al., 2005; Clancy & Eisenberg, 1998). The importance of EBP has led to a paradigm shift in measuring clinical outcomes; recommendations have included a reduced reliance on clinician-focused measures (e.g., range of motion scores) and have instead emphasized the need to collect patient focused measures (e.g., the patient's perspective and experience of their range of motion; Raine et al., 2016).

The emphasis on patient focused measures has increased the need for psychometrically sound patient reported outcome measures (PROMs; Law & MacDermid, 2014). The use of PROMs provides a patient-reported assessment of health (Raine et al., 2016); PROMs may measure one construct (i.e., unidimensional) or multiple constructs (i.e., multidimensional) and can be categorized as generic (e.g., general health), disease or symptom-specific (e.g., stroke), regional or body-part specific (e.g., shoulder pain), or patient-specific (e.g., occupational performance; Law & MacDermid, 2014; McDowell, 2006; Raine et al., 2016). The broad dimensions of health measured by PROMs may include physical function (e.g., mobility, range of movement), symptoms (e.g., pain, fatigue), psychological well-being (e.g., psychological illness, coping), social well-being (e.g., relationships with family, leisure activities), cognitive functions (e.g., concentration, memory), role activities (e.g., employment, financial concerns), personal constructs (life satisfaction, spirituality), satisfaction with care, or a combination of these dimensions (Raine et al., 2016).

The disablement construct has become an increasingly popular health dimension to assess in patient care. Disablement is a multidimensional construct that combines several dimensions of health status (Snyder et al., 2008); however, due to theoretical differences in disablement models, various disablement PROMs (e.g., WHO Disablement Assessment Schedule, Duke Health Profile) have been developed for clinical practice. Selecting an appropriate disablement PROM requires consideration of the underlying theoretical model, as well as reflection on the population of interest because researchers have modified or created disablement PROMs to be used in specific subgroups of patients (e.g., physically active patients; Vela & Denegar, 2010a).

The Disablement in Physically Active (DPA) Scale was developed as a multi-dimensional disablement model PROM for a physically active population (Vela & Denegar, 2010a, 2010b). The DPA Scale is a 16-item scale used to assess transient disablement dimensions of impairment, functional limitation, disability, and quality of life (Figure 2.1; Vela & Denegar, 2010b). Although the DPA Scale provided clinicians with a much-needed PROM for physically active populations, subsequent psychometric analysis of the scale indicated the instrument did not meet contemporary recommendations for scale development (Baker et al., 2019, in press; Houston et al., 2015). Specifically, researchers found the DPA Scale did not meet model fit recommendations, had potential issues of multicollinearity between factors, and did not pass testing for invariance across different populations of interests (Baker et al., 2019, in press). Alternate model generation was conducted to resolve the identified issues present in the DPA Scale; a modified, and more parsimonious version of the scale, the Disablement in Physically Active Scale Short Form-8 (DPA SF-8), was proposed (Baker et al., 2019, in press).

The DPA SF-8 uses 8-items from the original DPA Scale to assess two factors of disablement: a physical summary component and a quality of life component (Baker et al., n.d., 2019). The DPA SF-8 exceeded contemporary standards for model fit and accounted for a substantial amount of variance in participants' scores on the DPA Scale (Baker et al., n.d., 2019). The modified scale could reduce response burden for participants and provide more efficient administration for clinicians with improved scale validity and precision. Prior to adoption into clinical practice, however, further analysis is necessary to ensure the DPA SF-8 is psychometrically sound (e.g., reliable, valid, responsive) and can accurately assess disablement across subgroups and across time (Brown, 2014; Mokkink et al., 2010).

Thus, further research on the DPA SF-8 is needed to establish scale validity and should include multiple steps to ensure the scale is suitable for use in clinical practice and research: 1) a sample of individuals who only answered the 8-items contained in the DPA SF-8 version must be tested to confirm the factor structure (Baker et al., n.d., 2019), 2) scale reliability and responsive must be examined, 3) a minimal clinically important difference (MCID) value should be established to allow clinicians to evaluate if a patient has undergone a clinically significant change, and 4) invariance testing should be conducted across groups and repeated use of the scale in practice. Therefore, the purpose of the study was to evaluate the psychometric properties of the DPA SF-8 in a three-step process: 1) a confirmatory factor analysis (CFA) of factor structure, using contemporary fit recommendations, in a large heterogeneous sample to ensure model fit in respondents who only respond to the items included in the DPA SF-8, 2) psychometric analysis of scale reliability, validity, sensitivity to change, and responsiveness, and then 3) invariance testing of the scale across subgroups (e.g., sex, age, injury classification) and across visits (e.g., intake, discharge).

Methods

Approval from the university institutional review board was obtained prior to collection of participant information. All participants provided informed consent and when necessary, legal guardians provided consent prior to participation and minors provided assent. All data was deidentified prior to data analysis.

Participants

Athletic training clinics and outpatient rehabilitation clinics (n = 8) across the United States were used to recruit participants. The targeted sample for recruitment included individuals who were physically active, while those who were sedentary or inactive were

excluded. Additionally, individuals who were healthy or had an injury classified as acute, subacute, or persistent were recruited while those with chronic pain were excluded from the study (Table 2.1; Vela & Denegar, 2010; Strong et al., 2002). Participants were grouped by sex, pre-defined physical activity (i.e. competitive athlete, recreational athlete, occupational athlete) and injury categories (Table 2.1).

Instrumentation

The survey packet included the Disablement in the Physically Active Scale Short Form-8 (DPA SF-8), a numeric pain rating scale, a global functioning scale, a patient specific functional scale, the Global Rate of Change Scale (GRoC), and a demographic information questionnaire. The survey packet was completed with an athletic trainer at three different visits; completion of survey packets was determined by injury category, consistent with previous research (Vela & Denegar, 2010b). Healthy individuals or individuals with either an acute or subacute injury, completed the packet at initial intake (visit 1), 3-5 days post initial visit (visit 2), and 7-10 days post initial visit and/or at discharge (visit 3). Individuals with a persistent injury, completed the packet at initial intake (visit 1), 7-10 days post initial visit (visit 2), and 3 weeks post initial visit and/or at discharge (visit 3). All survey data and demographic information were inputted into Qualtrics (Qualtrics, LLC, Provo, UT) by the collecting athletic trainer or by a member of the research team.

Disablement in Physically Active Scale Short Form-8

The DPA SF-8 is an 8-item patient-reported outcome measure (PROM) scale designed to measure two factors with four items in each latent factor: 'Physical' (PHY) and 'Quality of Life' (QOL; Appendix I). The two latent factors, PHY and QOL are first order latent variables that covary (Figure 2.2). Participants rated each item on a 1 ("no problem")

to 5 (“severe”) Likert scale. The scores provided for each item were then added together, with 8 points being subtracted from the summed total to produce a final total score. Participant total scores could range from 0 points to 32 points. The DPA SF-8 was collected at all three visits.

Numeric Pain Rating Scale

The Numeric Pain Rating Scale (NPRS; Hartrick et al., 2003) is an instrument designed to measure intensity of pain. Participants were asked to rate their current, best, and worst pain levels over the past 24 hours (Appendix J). Participants used a 0 (“no pain”) to 10 (“worst pain imaginable”) scale. A score that represents a patient’s level of pain over 24 hours was calculated by averaging the best, current, and worst pain scores reported (Cleland et al., 2008; Mintken et al., 2009). The NPRS was collected at all three visits.

Global Functioning Scale

The Global Functioning (GF) scale is a single item question used to assess an individual’s perceived overall level of functioning (Appendix K). Participants used a 10-cm line anchored by 0 (“unable to function at a normal level”) and 100 (“able to function completely at a normal level before the injury/problem”) scale (Wilkin et al., 1992). The GF scale was collected at all three visits.

Patient Specific Functional Scale

The Patient Specific Functional Scale (PSFS) assesses participant’s perceived ability to function on specific activities or tasks (Appendix L). Participants were asked to select three important activities that they are currently not able to do or have difficulty doing as a result of the injury/problem (Livermore-Brasher et al., 2018; Stratford et al., 1995; Sterling, 2007). After selecting three activities, participants used a 0 (“unable to perform activity”) to

10 (“able to perform activity at the same level as before injury or problem”) scale to rate each individual activity (Livermore-Brasher et al., 2018; Stratford et al., 1995; Sterling, 2007). The PSFS was collected at all three visits.

Global Rate of Change Scale

The Global Rate of Change Scale (GRoC) was used to assess an individual’s rate of change throughout treatment (Appendix M). The GRoC has been proposed as the “gold standard” for change and has been previously validated in a number of studies (Fritz & Clifford, 2010; Hurst & Bolton, 2004; Jaeschke et al., 1989; Spadoni et al., 2004; Stratford et al., 1996). Participants used a 15-point scale (-7 = a very great deal worse, 0 = unchanged, 7 = a very great deal better). The GRoC was only collected at the second and third visits.

Demographic Information Questionnaire

The de-identified participant demographic information (Appendix N) collected included: injury category (i.e., persistent, acute, sub-acute, or healthy), patient athletic status (e.g., competitive athlete, recreational athlete), age, sex, sport, general injury location (e.g., lower extremity, upper extremity), specific injury location (e.g., head/neck, shoulder/arm, ankle/foot), and type of injury (e.g., arthritis, sprain, post-surgery). Demographic information was collected at the first visit.

Data Analysis

All data was input into Qualtrics (Provo, UT) by the athletic trainer or a member of the research team. Data were then downloaded and analyzed using Statistical Package for the Social Sciences (SPSS Inc., Chicago, IL, USA) Version 26 and Analysis of Moment Structure (AMOS, SPSS, Inc.) Version 26. Missing responses were calculated for the DPA SF-8 and individuals who did not respond to at least 90% (i.e., 7 of the 8) of the items were

removed from the data set. Although demographic data was assessed, individuals missing information were not removed from the data set. After missing data was calculated, assessment of univariate and multivariate outliers was conducted. Data normality was assessed by examining histograms, skewness values, kurtosis values, and examining for outliers using z-scores and Mahalanobis distance. Participants with z-scores exceeding |3.4| for an individual item were flagged and removed. Multivariate outliers for each individual were assessed, flagged, and removed from the data set if the Mahalanobis distance exceeded the cut-off value identified in the chi-square table with degrees of freedom ($p = .01$; Kline, 2015). For longitudinal invariance, individuals who did not respond to DPA SF-8 items at all three time points (i.e., visits) were not used in the analysis.

Scale Structure

The full sample was used to conduct confirmatory factor analyses (CFA) with maximum likelihood estimation in Analysis of Moment Structures (AMOS) software (IBM Corp., Armonk, NY) on the proposed 8-item, 2-factor structure of the DPA SF-8 by time point (i.e., visit; Figure 2.2). Model fit indices were evaluated based on *a priori* values. The relative goodness-of-fit indices computed were the Comparative Fit Index (CFI; $\geq .95$), Tucker-Lewis Index (TLI; $\geq .95$), Root Mean Square Error of Approximation (RMSEA $\leq .06$), and Bollen's Incremental Fit Index (IFI; $\geq .95$). Additionally, the likelihood ratio statistic (CMIN) was assessed but not used as the primary assessment measure because it is heavily influenced by sample size (Brown, 2014; Kline, 2015). Because model fit criteria were met, longitudinal and multigroup invariance testing was conducted.

Reliability

Internal consistency of the scale was assessed by calculating Cronbach's alpha for each proposed factor; Cronbach's alpha was set *a priori* at ≥ .70 and ≤ .89 (Leech et al., 2014; Morgado et al., 2018). Additionally, three intraclass correlation coefficients (ICC; 2,1) were calculated to assess test-retest reliability for the DPA SF-8 total scores, PHY subscale scores, and QOL subscale scores for healthy individuals across time points (i.e., visits). Values were set *a priori*: < .50 = poor, .50 - .75 = moderate, .75 - .90 = good, > .90 = excellent (Koo & Li, 2016). The standard error of measurement (SEM) value was then computed for total scores, PHY subscale scores, and QOL subscale scores using the formula $SEM = SD \times \sqrt{1-ICC}$; to calculate the three minimal detectable change (MDC) values, the formula $MDC = SEM \times 1.96 \times \sqrt{2}$ was used (Weir, 2005).

Validity

Correlations were assessed using a covariance modeling approach between the second-order latent variable of the DPA SF-8 and the scores of the GF, NPRS, and PSFS. Additionally, correlations were assessed between the first-order latent variables of the DPA SF-8 (i.e., PHY, QOL) and the GF, NPRS, and PSFS at each time point (i.e., visit).

Responsiveness

Responsiveness is typically understood as an aspect of validity for longitudinal research; it is the ability of an instrument to detect change over time (Hays & Hadorn, 1992; Husted et al., 2000). Clinical instruments used in an evaluative manner (e.g., is my patient getting better throughout treatment) should adequately detect changes related to the measure of interest. To detect responsiveness, a protocol from previous research establishing the responsiveness of the original DPA Scale was used (Vela & Denegar, 2010b). The protocol

included creating four Receiver Operating Characteristics (ROC) curves, two for individuals with acute or subacute injuries and two for individuals with persistent injuries. The procedure involved creating change scores for the DPA SF-8 and for classification group scales (i.e., NPRS, GF, PSFS). First, two change scores from the DPA SF-8 were calculated by subtracting the scores from visit 2 with visit 1 (V2 score) and subtracting scores from visit 3 with scores from visit 2 (V3 score). Then change scores were calculated for the NPRS, GF, and PSFS scales by subtracting the scores from visit 2 with visit 1 and subtracting scores from visit 3 with scores from visit 2.

The change scores from the DPA SF-8 (i.e., V1, V2) were then used to calculate the plots for the ROC curve based on classification groups (i.e., clinical significance, stable) that would indicate undergoing a clinically significant change. Due to the multi-dimensional nature of the DPA SF-8, as well as potential limitations of the GRoC for assessing change (Garrison & Cook, 2012), change scores from four different scales were used as criteria for determining clinically significant changes across treatment. Scores from the GRoC and the change scores from the NPRS, GF, and PSFS were used as criteria for classification groupings; individuals were placed into two classification groups: one group for visit 2 and one group for visit 3. To be placed in the clinically significant group an individual had to meet two criteria: GRoC score of 4 or greater (Fritz & Clifford, 2010; Vela & Denegar, 2010b), NPRS change score difference of 30% or more (Livermore et al., 2018), PSFS change score of 2 or greater (Nicholas et al., 2012; Livermore et al., 2018), or GF change score difference of 30% or more. If an individual did not meet at least two of the four criteria, the participant was placed in the stable group.

Sensitivity and specificity values were then calculated for V2 and V3 based on the number of individuals classified as experiencing a clinically significant change versus those who did not experience a clinically significant change (i.e., stable). A ROC curve was plotted using the sensitivity and specificity values. The area under the curve (AUC) was used to determine if the DPA SF-8 would correctly distinguish between individuals with a clinically significant change and those who did not experience a clinically significant change; an AUC value close to 1.00 indicates the test has perfect discernment between groups (Deyo & Centor, 1986; Zou et al., 2007).

Two ROC curves were calculated using participants with acute or subacute injuries only and two ROC curves were calculated using individuals with persistent injuries only. The minimal clinically important difference (MCID) value was determined by selecting the point on the ROC curve that represents the smallest overall error rate (Kelly et al., 2008; Vining & Gladish, 1992). The MCID represents the change in score on the DPA SF-8 that indicated the participant has undergone a clinically significantly change (Kelly et al., 2008; Vela & Denegar, 2010b).

Invariance testing

The same criteria utilized for the CFAs were used to assess fit for invariance models (Kline, 2015; Brown, 2015). Invariance testing with the full sample was conducted to assess measurement invariance of the DPA SF-8 across three visits (i.e., longitudinal invariance) and between subgroups of the sample (i.e., multigroup invariance.). Individuals who completed the DPA SF-8 at all three visits and had suffered an injury were used to assess invariance across time; data from visit one was used to assess multi-group invariance between injury status, sex, and activity levels. Invariance was evaluated based on a CFI

difference (CFI_{DIFF}) of less than .01, and the chi-square difference test (χ^2_{DIFF}), with a *p*-value cut-off of 0.01 (Brown, 2015; Kline, 2015). The CFI_{DIFF} test held greater weight in decisions regarding invariance testing model fit (Brown, 2015; Kline, 2015) because of the sensitivity of the χ^2_{DIFF} test regarding sample size. Therefore, if a model exceeded the χ^2_{DIFF} test but passed the CFI_{diff} test, invariance testing would continue.

Results

A total of 525 individual responses were collected. Of the 525, twenty individuals were missing more than 10% of the responses and were removed from the dataset. Five individuals reported scores that were identified as univariate outliers and 22 reported scores that were identified as multivariate outliers; the 27 individuals were subsequently removed from the dataset. A total of 478 individuals were retained for analysis. The mean age of the sample was 27.52 ± 11.55 years (range = 13 - 70; median = 22) with males accounting for 47.6% (n = 216) and females accounting for 49.4% (n = 236). Individuals with persistent injuries accounted for 36.2% (n = 177) of the sample and recreational athletes accounted for 33.3% (n = 159) of the sample. Full demographic information is presented in Table 2.2.

Scale Structure

The scale structure of the DPA SF-8 was assessed at all three time points (Visit 1, Visit 2, Visit 3). Groups means are presented in Table 2.3 by visit and injury type. A total of 478 individuals completed the DPA SF-8 at time 1 (i.e., visit 1) and were used for the analysis. The goodness of fit model indices exceeded recommended values (CFI = .997, TLI = .996, IFI = .997, RMSEA = .023; Figure 2.3). The first-order latent variable correlation (R= .40, R^2 = .16) and factor loadings were significant ($p < .001$), with loadings ranging from

.66 - .87. Modification indices did not demonstrate any significant cross-loadings or meaningful modifications were necessary.

A total of 347 individuals completed the DPA SF-8 at time 2 (i.e., visit 2) and were used for the analysis. The goodness of fit model indices exceeded recommended values (CFI = .993, TLI = .990, IFI = .993, RMSEA = .039; Figure 2.4). The first-order latent variable correlation (R = .45; R^2 = .21) and factor loadings were significant ($p < .001$), with loadings ranging from .69 - .88. Modification indices did not demonstrate any significant cross-loadings or meaningful modifications were necessary.

A total of 234 individuals completed the DPA SF-8 at time 3 (i.e., visit 3) and were used for the analysis. The goodness of fit model indices exceeded recommended values (CFI = .991, TLI = .986, IFI = .991, RMSEA = .050; Figure 2.5). The first-order latent variable correlation (R = .49; R^2 = .24) and factor loadings were significant ($p < .001$), with loadings ranging from .71 - .94. Modification indices did not demonstrate any significant cross-loadings or meaningful modifications were necessary.

Reliability

Factor 1, Physical (PHY), included DPA SF-8 items 1-4, and Factor 2, Quality of Life (QOL), included items 5-8. Cronbach's alpha was assessed by factor (PHY, QOL) across three time points (Visit 1, Visit 2, Visit 3). The PHY factor alphas were .81 (Visit 1), .86 (Visit 2), and .89 (Visit 3) while the QOL factor alphas were .87 (Visit 1, 2) and .88 (Visit 3). The ICC (2,1) for healthy individuals (n = 26) across visits was .924 with an SEM value of 2.10 and an MDC value of 5.83 points. The ICC (2, 1) for the PHY subscale was .899 with an SEM value of 1.44 and an MDC value of 4.0. The ICC (2, 1) for the QOL subscale was .841 with an SEM value of 1.69 and MDC value of 4.68.

Validity

The correlations between the second-order latent variable DPA SF-8 and the GF scores were significant at -.63 (R^2 = .40, p < .001) for visit 1, -.56 (R^2 = .32, p < .001) for visit 2, and -.65 (R^2 = .42, p < .001) for visit 3. The correlations between the second-order latent variable DPA SF-8 and the average NPRS scores were significant for visit 1 at .58 (R^2 = .34, p < .001), .80 (R^2 = .64, p < .001) for time 2, and .78 (R^2 = .61, p < .001) for visit 3. The correlations between the second-order latent variable DPA SF-8 and the average PSFS score were significant for visit 1 at -.51 (R^2 = .26, p < .001), -.69 (R^2 = .48, p < .001) for visit 2, and -.65 (R^2 = .42, p < .001) for visit 3.

For visit one, the correlations were significant between the first-order latent variable PHY and the GF score (R = -.55, R^2 = .30, p < .001), the NPRS (R = .57, R^2 = .32, p < .001), and the PSFS (R = -.51, R^2 = .26, p < .001); the correlations were also significant between the first-order latent variable QOL and the GF score (R = -.29, R^2 = .08, p < .001), the NPRS (R = .23, R^2 = .05, p < .001), and the PSFS (R = -.21, R^2 = .04, p < .001).

For visit two, the correlations were significant between the first-order latent variable PHY and the GF score (R = -.58, R^2 = .34, p < .001), the NPRS (R = .66, R^2 = .44, p < .001), and the PSFS (R = -.64, R^2 = .41, p < .001); the correlations were also significant between the first-order latent variable QOL and the GF score (R = -.24, R^2 = .06, p < .001), the NPRS (R = .44, R^2 = .19, p < .001), and the PSFS (R = -.34, R^2 = .12, p < .001).

For visit three, the correlations were significant between the first-order latent variable PHY and the GF score (R = -.64, R^2 = .41, p < .001), the NPRS (R = .72, R^2 = .52, p < .001), and the PSFS (R = -.64, R^2 = .41, p < .001); the correlations were also significant between

the first-order latent variable QOL and the GF score ($R = -.33$, $R^2 = .11$, $p < .001$), the NPRS ($R = .43$, $R^2 = .18$, $p < .001$), and the PSFS ($R = -.33$, $R^2 = .11$, $p < .001$).

Responsiveness

Persistent. One hundred individuals with a persistent injury, responded to the DPA SF-8 at visit one and two and were used for analysis. Of the 100 individuals, 26 reported experiencing a clinically significant change at the second visit. The AUC value for participants was .710 (95% confidence interval = .597, .822; P = .002; Figure 2.6). The MCID value calculated for the ROC curve for visit 2 was 2.50 points (sensitivity = .731; 1 – specificity = .392).

Ninety-seven individuals with a persistent injury responded to the DPA SF-8 at visit 2 and 3 and were used for analysis. Of the 97 individuals, 29 reported experiencing a clinically significant change at visit three. The AUC value for participants was .721 (95% confidence interval = .616, .825; P = .001; Figure 2.7). The MCID value calculated for the ROC curve for visit 3 was 1.50 points (sensitivity = .690; 1 – specificity = .397). The two values were averaged to create an MCID value of 2 points for individuals with persistent injuries.

Acute and Subacute Injuries. Seventy-seven individuals with an acute or subacute injury, responded to the DPA SF-8 at visit 1 and 2 and were used for analysis. Of the 77 individuals, 40 reported experiencing a clinically significant change at visit two. The AUC value for participants was .803 (95% confidence interval = .706, .901; P < .001; Figure 2.8). The MCID value calculated for the ROC curve on visit 2 was 3.5 point (sensitivity = 0.675; 1 – specificity = .216).

Seventy-three individuals with an acute or subacute injury responded to the DPA SF-8 at visit 2 and 3 and were used for analysis. Of the 73 individuals, 28 reported experiencing

a clinically significant change at visit 3. The AUC value for participants was .716 (95% confidence interval = .595, .837; P = .002; Figure 2.9). The MCID value calculated for the ROC curve by visit 3 was 2.50 points (sensitivity = .571; 1 – specificity = .172). The two values were averaged to create an MCID value of 3 points for individuals with acute or subacute injuries.

Invariance Testing

Longitudinal (Repeated Assessment)

A total of 206 injured individuals responded to the DPA SF-8 at three time points (i.e., visits) and were used for analysis. The configural model (i.e., equal form) goodness of fit indices met recommended values (CFI = .981; χ^2 [213] = 278.46; RMSEA = .039; Table 2.6). The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test and the χ^2_{DIFF} test, warranting examination of an equal latent variance model. The equal latent variance model passed both the CFI_{DIFF} and χ^2_{DIFF} difference test, indicating variances were equal for first-order latent variables PHY and QOL across time.

The scalar model (i.e., equal indicator intercepts) slightly exceeded the χ^2_{DIFF} test however, met the CFI_{DIFF} test, warranting examination of an equal latent means model. The equal latent means model did not pass either the CFI_{DIFF} or the χ^2_{DIFF} test, indicating means for PHY and QOL were not equal across time. Analysis of means when not constrained to be equal indicated that individuals reported lower scores for PHY and higher scores for QOL across time (i.e., repeated assessment).

Multigroup

Sex. A total of 452 individuals reported sex (male = 216; female = 236) at time one (i.e., visit one) and were used for analysis. The configural model (i.e., equal form) goodness

of fit indices met recommended values (CFI = .983; χ^2 [213] = 288.52; RMSEA = .039; Table 2.7). The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test and the χ^2_{DIFF} test, warranting examination of an equal latent variance model. The equal latent variance model passed both the CFI_{DIFF} and χ^2_{DIFF} difference test, indicating variances were equal for first-order latent variables PHY and QOL across sex.

The scalar model (i.e., equal indicator intercepts) passed both the CFI_{DIFF} test and the χ^2_{DIFF} tests, warranting examination of an equal latent means model. The equal latent means model passed both the CFI_{DIFF} and the χ^2_{DIFF} test, indicating means for PHY and QOL were equal across sex.

Activity Level. A total of 392 injured individuals reported their activity level (low = 105, medium = 179, high = 133) at time one (i.e., visit one) and were used for analysis. The configural model (i.e., equal form) goodness of fit indices met recommended values (CFI = .995; χ^2 [57] = 60.58; RMSEA = .013; Table 2.8). The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test and the χ^2_{DIFF} test, warranting examination of an equal latent variance model. The equal latent variance model passed both the CFI_{DIFF} and χ^2_{DIFF} difference test, indicating variances were equal for first-order latent variables PHY and QOL across activity level. The scalar model (i.e., equal indicator intercepts) passed both the CFI_{DIFF} test and the χ^2_{DIFF} tests, warranting examination of an equal latent means model. The equal latent means model passed both the χ^2_{DIFF} test and the CFI_{DIFF} difference test, indicating means were equal for first-order latent variables PHY and QOL across activity levels.

Injury Category. A total of 329 individuals reported having a persistent (n = 177) or a subacute or acute injury (n = 161) at time one (i.e., visit one) and were used for analysis.

The configural model (i.e., equal form) goodness of fit indices met recommended values (CFI = 1.0; χ^2 [38] = 30.89; RMSEA < .001; Table 2.9). The metric model (i.e., equal loadings) passed both the CFI_{DIFF} test and the χ^2_{DIFF} test, warranting examination of an equal latent variance model. The equal latent variance model passed both the CFI_{DIFF} and χ^2_{DIFF} difference test, indicating variances were equal for first-order latent variables PHY and QOL across injury category.

The scalar model (i.e., equal indicator intercepts) passed both the CFI_{DIFF} test and the χ^2_{DIFF} tests, warranting examination of an equal latent means model. The equal latent means model passed both the CFI_{DIFF} and χ^2_{DIFF} difference test, indicating means were equal for first-order latent variables PHY and QOL across injury category.

Discussion

Patient-centered care and evidence-based practice (EBP) are core competencies for health care professionals. The importance of EBP has led to an increase in research involving clinical outcomes (e.g., physiological findings, patient self-reported instruments); current recommendations emphasize collecting patient focused measures (e.g., the patient's perspective and experience of their range of motion; Raine et al., 2016), thus increasing the need for psychometrically sound patient reported outcome measures (PROMs) of health (Law & MacDermid, 2014). Disablement has been proposed as a valuable multi-dimensional construct for patient care; however, selecting an appropriate disablement PROM to assess disablement may depend on the specific subgroups of patients completing the scale (Vela & Denegar, 2010a).

The Disablement in Physically Active Scale Short Form-8 (DPA SF-8), (Baker et al., 2019, in press) assesses two factors of disablement: a physical summary component (PHY) and a quality of life component (QOL; Baker et al., 2019, in press). The DPA SF-8 exceeded

contemporary standards for model fit and (Baker et al., 2019, in press); however, further analysis was necessary to ensure the DPA SF-8 was psychometrically sound and could accurately assess disablement across subgroups and time (Brown, 2014; Mokkink et al., 2010). Therefore, the purposes of our study were to establish the DPA SF-8 scale reliability, validity, sensitivity to change, responsiveness, and longitudinal and multi-group invariance.

DPA SF-8 Scale Structure

The CFAs at all three visits exceeded recommended model fit indices, thus confirming the scale structure of the DPA SF-8 (Baker et al., 2019, in press). This study, however, was the first to use a large heterogeneous sample of adolescents and adults who responded to the 8-item scale. The total scores on the DPA SF-8 by injury classification (Table 2.3) were similar to scores reported in previous research (Baker et al., 2019). Individuals with a persistent or acute injury reported higher overall scores (i.e., more disablement and lower quality of life) than healthy individuals who reported lower overall scores (i.e., less disablement and higher quality of life). The correlation values between the first-order latent variables PHY and QOL ($R = .40\text{-}49$, $R^2 = .16\text{-}24$) across visits were also similar to previous research (Baker et al., 2019, in press); the findings support that the PHY and QOL constructs of disablement are unique constructs (Baker et al., 2019).

Overall, the scale structure findings indicate exceptional model fit for the DPA SF-8 in respondents who only answer the 8-items, and suggest the DPA SF-8 continues to resolve item redundancy and multicollinearity issues found in the DPA Scale or DPA SF-10 (Baker et al., 2019, in press; Vela & Denegar, 2010). Although the scale was designed for use in the physically active, our full sample included a small percentage (n = 25, 5.2%) of individuals with extremely low levels of physical activity (i.e., activities of daily living). The excellent

model fit with those individuals included, as well as the excellent model fit in studies excluding extremely low levels of physical activity individuals (Baker et al., 2019, in press), implies the scale may be psychometrically sound in both groups. However, future research should assess the scale structure of the DPA SF-8 in a larger group of individuals with extremely low levels of physical activity, as well as in inactive individuals. Additionally, multi-group invariance between physically active and inactive individuals should be performed to ensure scale structure is supported across these groups.

Internal Consistency and Reliability of the DPA SF-8

Cronbach's alpha for PHY and QOL were within recommend values, which support sound internal consistency of the constructs and reduced risk of multicollinearity in the scale. The ICC value (.924) calculated across three time points (i.e., initial visit, visit 2 = 3-5 days post initial visit, visit 3 = 7-10 days post-initial visit) indicated excellent scale reliability (Koo & Li, 2016). Our results were similar to the original ICC value (.943) found for the 16-item DPA Scale in injured individual across two time points, 24 hours apart (Vela & Denegar, 2010b); our ICC value was higher than the reliability value (.792) reported in soccer players tested on the 16-item scale during preseason, one week apart (Hoch et al., 2015). Our results indicate a true change in a patient's overall disablement when completing the DPA SF-8 multiple times is likely less than 6 points (MDC = 5.83), which was improved from a previously reported MDC value of 12.48 for the DPA Scale (Hoch et al., 2015). The improved internal consistency and MDC values of the DPA SF-8 were expected given the improved precision and model fit, as well as the reduced item redudancy, of the scale compared to the original 16-item DPA Scale (Baker et al., 2019, in press).

Validity

Criterion validity was assessed by examining the correlations between the DPA SF-8 and the scores on the GF scale, NPRS, and PSFS. The significant inverse relationship between the GF scale and DPA SF-8 is consistent with previous findings (Vela & Denegar, 2010b); however, the second-order latent variable correlation values across all participants and timepoints (i.e., visits) in our study were lower ($R = -.63$, $R^2 = .40$ [visit 1], $R = -.56$, $R^2 = .31$ [visit 2], and $R = -.65$, $R^2 = .42$ [visit 3]) than previous findings ($R = -.714$, $R^2 = .51$ for persistent and $R = -.751$, $R^2 = .56$ for acute injuries) for the original 16-item scale (Vela & Denegar, 2010b). The small difference in correlational values may be the result of study or scale differences. We utilized a larger and more diverse participant pool with a higher mean age than the previous study (Vela & Denegar, 2010b) and we included a small portion of healthy individuals; it is possible that differences in participant age between the studies or the healthy participants included in our analysis resulted in slightly different responses across items or scales. For example, people who are healthy should not be processing change from injury, while people of different ages who are injured process changes across health dimensions (e.g., physical function, quality of life) differently across the lifespan (Zullig et al., 2005); those differences may have altered the correlational values between the scales. Another potential explanation is reduced item redundancy in the DPA SF-8 due to the decreased number of PHY questions (i.e., 4 items compared to 12) present in the short form compared to the DPA Scale; the removal of highly correlated items assessing physical functioning may have also reduced the correlation between the GF scale and the DPA SF-8.

The assessment of concurrent validity should also involve correlating the DPA SF-8 to other relevant scales because the DPA SF-8 is a multidimensional scale that allows

summative (i.e., scale total) and construct (i.e., PHY and QOL) scoring. The correlational directions (e.g., inverse with GF scale and PSFS) and magnitudes met expectations and support concurrent validity. The second-order latent variable correlational analysis indicated significant positive relationships between the DPA SF-8 and the NPRS across visits (Table 2.4), with an inverse significant relationship between the DPA SF-8 and the PSFS across visits (Table 2.4). The first order latent variable correlations between the PHY construct of the DPA SF-8 and the NPRS and PSFS demonstrated a similar pattern of direction and magnitude across visits (Table 2.5). The first order correlations were also significant between QOL construct of the DPA SF-8 and the NPRS and PSFS; however, correlation values between these scales and the QOL construct were, as expected, lower than those found with the PHY construct.

The overall correlational findings support the validity of the DPA SF-8. The DPA SF-8, like the DPA Scale, was significantly and appropriately correlated with the GF Scale providing some evidence of criterion validity. Additionally, the DPA SF-8 PHY construct was highly correlated with related unidimensional scales (i.e., NPRS and PSFS) designed to measure components found in that dimension. The DPA SF-8 QOL construct was correlated with related unidimensional scales (i.e., GF Scale, NPRS, and PSFS); the correlation values fit proposed theory in that the correlations were in the same direction but of lower magnitude to those found with the PHY construct. Further, the correlation values between the DPA SF-8 PHY and QOL constructs and the GF Scale, NPRS, and PSFS increased over visits which indicated that patient improvement was being identified across both SF-8 constructs and the other scales in a more similar pattern. Finally, the primarily adult population in our study more strongly defined (i.e., more heavily weighted) improvement through physical health

changes, as opposed to QOL changes, which is consistent with the expectations developed in previous research (Zullig et al., 2005). Future research should be completed to further establish the validity of the QOL subscale by correlating the construct to another previously established quality of life instrument.

Receiver Operating Curve Responsiveness

We also evaluated the ability of the DPA SF-8 to detect change over time, or the responsiveness of the scale (Hays & Hadorn, 1992; Husted et al., 2000), using a ROC curve analysis. Previous research utilized the GRoC to classify individuals into either a clinically significant group or a stable group to develop MCIDs (Fritz & Clifford, 2010; Vela & Denegar, 2010b). We chose to utilize three additional outcome measures (i.e., NPRS, PSFS, GF) in addition to the GRoC for classification into the grouping for two reasons: 1) the GRoC has been scrutinized for poor reliability and recall bias (Garrison & Cook, 2012), and 2) the multidimensional nature of the DPA SF-8 was better represented by utilizing multiple instruments that represented the depth of the unique constructs/items of the DPA SF-8.

The four ROC curves were then constructed based on our groupings at two time points (i.e., visit two and visit three): two for individuals with persistent injuries and two for individuals with acute or subacute injuries. The four AUC values (range = .710 - .803) for the ROC curves were statistically significant and within the moderately high range, indicating the scale could detect meaningful change from the patient perspective. Overall, our range of AUC values was slightly narrower (i.e., top end was lower) than was found for the DPA Scale (range = .702-.911); however, our sample was significantly larger and more diverse, and did not have a group (i.e., acute) where all members experienced a significant change (Vela & Denegar, 2010b).

We calculated the MCID values using the ROC curve for two groups of respondents: 1) individuals with a persistent injury, 2) and individuals with acute and subacute injuries. The MCID values are beneficial for providing clinicians and researchers with insight into true clinical change as perceived by the patient. Our results indicate patients with a persistent injury will have likely experienced a clinically significant change with a 2 point or greater change in the total DPA SF-8 score. For those with a subacute or acute injury, a clinically significant change will likely have occurred if a 3 point or greater change is reported on the total DPA SF-8 score. The MCID values for the DPA SF-8 are lower than those reported for the DPA Scale for persistent (6 points) and acute (9 points) injured (Vela & Denegar, 2010b); however, the lower values are expected given the removal of items (16 to 8) resulting in improved model fit, reduced item redundancy, resolved multicollinearity, and improve scale precision.

Our findings, however, may be limited by our sample and methodology. The established MCID values may have improved accuracy if group classification included a component to more effectively assess and classify change in the QOL construct. For example, adolescents weigh responses more heavily to mental health changes than adults (Zullig et al., 2005), and those changes may not be effectively captured in any of the methodologies utilized to establish MCIDs for the DPA Scale or the DPA SF-8. Thus, future research may be needed to establish MCID for the sub-constructs of the scales, across different age groups, or using methods which classify change more effectively across both physical and mental health dimensions to better represent the multi-factorial nature of the DPA Scale and DPA SF-8.

Multi-group and Longitudinal Invariance

Our study is the first to assess invariance of the DPA SF-8 across time visits and groups of interest (i.e., sex, injury classification, activity level). Invariance testing is necessary to ensure the association between the PHY and QOL latent variables, and their items, are stable and relatively equal over visits and between groups (Brown, 2014; Byrne, 2016, Kline, 2015). An instrument that is invariant allows for comparisons across group and time (i.e., visit) by confirming individuals are interpreting the items and meaning of the items similarly, which provides evidence that score changes or group differences are true changes/differences as opposed to differences due to other group/time attributes or measurement error (Brown, 2014; Kline, 2015).

Our results indicate the DPA SF-8 was invariant across all our analyses, which allows clinicians and researchers to use the scale to compare differences in the tested groups (e.g., sex, physical activity level) or to assess individual changes in scores over the treatment period. We did not find group mean differences in the PHY or QOL constructs across sex or physical activity level. Our results differ from previous research where individuals who engaged in physical activity report higher scores on quality of life (Anokye et al., 2012); however, our results may have been confounded by the physical activity group including participants who were currently suffering an injury which likely would have reduced QOL scores compared to those who were uninjured but physically active.

Our invariance results also support the validity and responsiveness of the DPA SF-8. The DPA SF-8 was invariant across the persistent and acute/sub-acute injury groups which indicates the scale may be used to assess differences in disablement across the two groups. Further, the responses of the injured participants were invariant across visits and revealed

significant improvements in their health status (i.e., less physical impairment and disability and more quality of life) across repeated measures. The DPA SF-8 revealing significant improvement over visits for those suffering injuries would be expected when naturally healing and care from their healthcare provider is occurring across the repeated measures.

Unfortunately, our sample did not include a large enough number of healthy participants to include this group in the multi-group invariance test procedures with the persistent and acute/sub-acute injury groups. Researchers have indicated the DPA Scale did not demonstrate multi-group invariance across injured and uninjured participants (Baker et al., 2019); further research is needed to establish if the DPA SF-8 is invariant between these groups, ensuring item-level bias does not occur due to group attributes. Clinically, establishing invariance across injured and uninured participants helps ensure item interpretation and measurement are measured consistently across these two groups, which is valuable when the DPA SF-8 is used to inform return to play or discharge from care decisions (i.e., when patients transition from injured to healthy).

Clinical Implications

The results from our study indicate the DPA SF-8 is reliable, valid, and responsive instrument for the physically active population. Clinicians and researchers may use the DPA SF-8 to assess treatment efficacy across repeated measures or to compare scores between certain groups; however, caution is warranted if scores are being compared across injured and uninjured respondents at this time. The MDC (5.83 points) and MCID (acute/subacute = 3; persistent = 2) values support the responsiveness of the scale: 1) clinicians and researchers may interpret a real change outside of measurement error has occurred when a change greater or equal to 6 points has occurred; 2) a clinically significant change important to the patient

can be interpreted when a 2 point or greater or 3 point or greater change is reported by those with a persistent or acute/subacute injury, respectively. Our results also confirm previous findings (Baker et al., 2019) that the PHY and QOL constructs are unique dimensions captured within the scale to measure health status in the physically active (Smith et al., 1999).

Our results support previous findings (Baker et al., 2019, in press) for scoring the individual constructs (i.e., PHY and QOL); however, our study is the first to establish MDC values the for the PHY (MDC = 4) and QOL (MDC = 5) constructs. Clinicians may use the MDC values to determine when a patient reports a change in each construct that is greater than the expected error for repeated completion of each construct; however, further research is needed to establish MCIDs for each construct. While examining the individual construct scores is likely a more accurate portrayal of health status (Houston et al., 2015), cumulative scores can be created and assessed (e.g., MDC values, MCID values) to provide clinicians insight into the overall health status of their patient. Clinicians should consider whether the improvements in DPA SF-8 cumulative scores are primarily driven by changes in physical function assessed by the items in the PHY and QOL constructs as opposed to true changes in mental health (i.e., QOL). Clinicians who use cumulative scores should also assess subdimension scores and consider the use of additional wellness or mental health status questionnaires when appropriate for a specific patient case.

Limitations and Future Research

Although our study used a large heterogenous population, most of our respondents (mean age = 28 years) were either in the emerging and early adulthood stages of human development. Our cross-sectional sample had smaller participation from members in other

stages (e.g., adolescents, middle adulthood, late adulthood) of human development. Therefore, future research should establish model fit and multi-group and longitudinal invariance of the DPA SF-8 across these age groups as appropriate for various clinical settings. Our sample also included a small percentage of individuals who had extremely low levels of physical activity; however, the group of individuals was too small to include in multi-group invariance testing. Future research should be conducted using a larger sample of extremely low physical active individuals and inactive individuals to ensure scale structure is sound in these groups, while also performing multi-group invariance testing (i.e., active vs. inactive individuals) to ensure group differences are not due to measurement error. Similarly, the MDC and MCID values may be different across groups (e.g., adolescent, emerging adult groups, low activity, high activity) and future research should seek to determine if those values vary across relevant clinical groups.

We used a similar protocol as previous research (Vela & Denegar, 2010b) to establish group classifications for clinically significant improvement; however, this exact method has not been replicated in the literature and the methods utilized may not best capture change across a multidimensional instrument. Therefore, future research should assess the responsiveness of the scale in a diverse sample of individuals with different instruments that adequately capture the depth and uniqueness of the PHY and QOL constructs of the DPA SF-8 to improve accuracy of classifications and MCID values. Additionally, we used previously established methods (Vela & Denegar, 2010b) for assessing construct validity by correlating the GF scale to the DPA SF-8; however, we also conducted second and first order correlations between the DPA SF-8 and the GF scale, NPRS, and PSFS to assess construct validity. While the results support the validity of the DPA SF-8, further research is warranted

to establish the validity of the QOL construct, as well as validity of a cumulative DPA SF-8 score as a measure for health status. Finally, while we had a sufficient sample size for much of our analysis work, we were limited by sample size in certain cases (e.g., multi-group invariance testing including a healthy group for comparison to injured participants); we also experienced participant dropout over the course of the study (i.e., participants being unable to return for 2nd or 3rd visits due to COVID-19). Thus, future research using large samples with higher completion rates for all three time points (i.e., visits) would be valuable to confirm or refute certain study findings (e.g., MCID values).

Conclusions

The DPA SF-8 met the strictest CFA recommendations without the need for scale modification in respondents who only answered the 8-items included in the scale. The DPA SF-8 also met all criteria for applied multi-group and longitudinal invariance tests, which indicates the scale may be used to assess for differences between the groups or across time. Our overall analysis indicates the DPA SF-8 is a valid, reliable, and responsive instrument to assess patient improvement in the physically active population.

References

Anokye, N. K., Trueman, P., Green, C., Pavey, T. G., & Taylor, R. S. (2012). Physical activity and health related quality of life. *BMC Public Health*, *12*(1), 624

Arnold, B. L., Gansneder, B. M., & Perrin, D. H. (2005). *Research Methods in Athletic Training*. F.A. Davis Company.

Baker, R. T., Burton, D., Pickering, M. A., & Start, A. (2019). Confirmatory factor analysis of the Disablement in the Physically Active Scale and preliminary testing of short-form versions: A calibration and validation study. *Journal of Athletic Training*, *54*(3), 302–318. https://doi.org/10.4085/1062-6050-355-17

Baker, R. T., Casanova, M. P., Pickering, M. A., & Baker, J. (in press). Invariance testing of the Disablement in the Physically Active Scale. *Journal of Athletic Training*.

Brown, T. A. (2014). *Confirmatory Factor Analysis for Applied Research* (Second Ed.). Guilford Publications.

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In *Reading and understanding multivariate statistics* (pp. 99–136). American Psychological Association.

Clancy, C. M., & Eisenberg, J. M. (1998). Outcomes research: Measuring the end results of health care. *Science*, *282*(5387), 245–246.

Cleland, J. A., Childs, J. D., & Whitman, J. M. (2008). Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Archives of Physical Medicine and Rehabilitation*, *89*(1), 69-74.

Deyo, R. A., & Centor, R. M. (1986). Assessing the responsiveness of functional scales to clinical change: An analogy to diagnostic test performance. *Journal of Chronic Diseases*, *39*(11), 897–906. https://doi.org/10.1016/0021-9681(86)90038-X

Fritz, J. M., & Clifford, S. N. (2010). Low back pain in adolescents: A comparison of clinical outcomes in sports participants and nonparticipants. *Journal of Athletic Training*, *45*(1), 61–66. https://doi.org/10.4085/1062-6050-45.1.61

Garrison, C., & Cook, C. (2012). Clinimetrics corner: The Global Rating of Change Score (GRoC) poorly correlates with functional measures and is not temporally stable. *Journal of Manual & Manipulative Therapy*, *20*(4), 178-181.

Hartrick, C. T., Kovan, J. P., & Shapiro, S. (2003). The Numeric Rating Scale for clinical pain measurement: A ratio measure? *Pain Practice*, *3*(4), 310–316. https://doi.org/10.1111/j.1530-7085.2003.03034.x

Hays, R. D., & Hadorn, D. (1992). Responsiveness to change: An aspect of validity, not a separate dimension. *Quality of Life Research*, *1*(1), 73–75. https://doi.org/10.1007/BF00435438

Hoch, J. M., Druvenga, B., Ferguson, B. A., Houston, M. N., & Hoch, M. C. (2015). Patient-reported outcomes in male and female collegiate soccer players during an athletic season. *Journal of Athletic Training*, *50*(9), 930-936.

Houston, M. N., Hoch, J. M., Van Lunen, B. L., & Hoch, M. C. (2015). The development of summary components for the Disablement in the Physically Active scale in collegiate athletes. *Quality of Life Research*, *24*(11), 2657–2662. https://doi.org/10.1007/s11136-015-1007-6

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Hurst, H., & Bolton, J. (2004). Assessing the clinical significance of change scores recorded on subjective outcome measures. *Journal of Manipulative and Physiological Therapeutics*, *27*(1), 26–35. https://doi.org/10.1016/j.jmpt.2003.11.003

Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for assessing responsiveness: A critical review and recommendations. *Journal of Clinical Epidemiology*, 11.

Institute of Medicine. (2003). *Health Professions Education: A Bridge to Quality*. National Academies Press. https://doi.org/10.17226/10681

Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status. *Controlled Clinical Trials*, *10*(4), 407–415. https://doi.org/10.1016/0197-2456(89)90005-6

Kelly, M. J., Dunstan, F. D., Lloyd, K., & Fone, D. L. (2008). Evaluating cutpoints for the MHI-5 and MCS using the GHQ-12: A comparison of five different methods. *BMC Psychiatry*, *8*(1). https://doi.org/10.1186/1471-244X-8-10

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling* (Fourth ed.). Guilford Publications.

Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J of Chiropr Med.* 2016;15(2):155-63.

Law, M. C., & MacDermid, J. (2014). *Evidence-Based Rehabilitation* (Third edition). Slack Incorporated.

Leech, N. L., Barrett, K. C., Morgan, G. A., Barrett, K. C., & Morgan, G. A. (2014). *IBM SPSS for Intermediate Statistics* (Fifth Ed.). Routledge.

Livermore-Brasher, J. S., Baker, R., Nasypany, A., & Cheatham, S. W. (2018). Guidelines for Useful Integration of Patient-Rated Outcome Measures into Clinical Practice. *Clinical Practice in Athletic Training*, *1*(2), 14-31.

McDowell, I. (2006). *Measuring Health: A Guide to Rating Scales and Questionnaires* (3rd ed). Oxford University Press.

Mintken, P. E., Glynn, P., & Cleland, J. A. (2009). Psychometric properties of the shortened disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH) and Numeric Pain Rating Scale in patients with shoulder pain. *Journal of Shoulder and Elbow Surgery*, *18*(6), 920-926.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, *19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S., & Ferreira, M. E. C. (2018). Scale development: Ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica*, *30*(1). https://doi.org/10.1186/s41155-016-0057-1

National Athletic Trainers' Association. (2011). *Athletic Training Education Competencies* (5th edition). https://www.nata.org/sites/default/files/competencies_5th_edition.pdf

Nicholas, P., Hefford, C., & Tumilty, S. (2012). The use of the Patient-Specific Functional Scale to measure rehabilitative progress in a physiotherapy setting. *Journal of Manual & Manipulative Therapy*, *20*(3), 147-152.

Raine, R., Fitzpatrick, R., Barratt, H., Bevan, G., Black, N., Boaden, R., Bower, P., Campbell, M., Denis, J.-L., Devers, K., Dixon-Woods, M., Fallowfield, L., Forder, J., Foy, R., Freemantle, N., Fulop, N. J., Gibbons, E., Gillies, C., Goulding, L., … Zwarenstein, M. (2016). Challenges, solutions and future directions in the evaluation of service innovations in health care and public health. *Health Services and Delivery Research*, *4*(16), 1–136. https://doi.org/10.3310/hsdr04160

Smith, K. W., Avis, N. E., & Assmann, S. F. (1999). Distinguishing between quality of life and health status in quality of life research: a meta-analysis. *Quality of Life Research*, *8*(5), 447-459.

Snyder, A. R., Parsons, J. T., Valovich McLeod, T. C., Curtis Bay, R., Michener, L. A., & Sauers, E. L. (2008). Using disablement models and clinical outcomes assessment to enable evidence-based athletic training practice, part I: Disablement models. *Journal of Athletic Training*, *43*(4), 428–436. https://doi.org/10.4085/1062-6050-43.4.428

Spadoni, G. F., Stratford, P. W., Solomon, P. E., & Wishart, L. R. (2004). The evaluation of change in pain intensity: A comparison of the P4 and single-item numeric pain rating scales. *Research Report*, *34*(4), 9.

Sterling, M. (2007). Patient specific functional scale. *Australian Journal of Physiotherapy*, *53*(1), 65-66.

Stratford, P. W., Binkley, J. M., & Riddle, D. L. (1996). Health status measures: Strategies and analytic methods for assessing change scores. *Physical Therapy*, *76*(10), 1109–1123. https://doi.org/10.1093/ptj/76.10.1109

Stratford, P., Gill, C., Westaway, M., & Binkley, J. (1995). Assessing disability and change on individual patients: a report of a patient specific measure. *Physiotherapy Canada*, *47*(4), 258-263.

Straus, S. E., Glasziou, P., Richardson, W. S., & Haynes, R. B. (2019). *Evidence-Based Medicine: How to Practice and Teach EBM* (Fifth Edition). Elsevier.

U.S. Department of Health and Human Services. (2018). *Physical Activity Guidelines for Americans* (2nd edition). Washington, DC.

Vela, L. I., & Denegar, C. (2010a). Transient disablement in the physically active with musculoskeletal injuries, part I: A descriptive model. *Journal of Athletic Training*, *45*(6), 615–629. https://doi.org/10.4085/1062-6050-45.6.615

Vela, L. I., & Denegar, C. R. (2010b). The Disablement in the Physically Active Scale, part II: The psychometric properties of an outcomes scale for musculoskeletal injuries. *Journal of Athletic Training*, *45*(6), 630–641. https://doi.org/10.4085/1062-6050-45.6.630

Vining, D. J., & Gladish, G. W. (1992). Receiver operating characteristic curves: A basic understanding. *RadioGraphics*, *12*(6), 1147–1154. https://doi.org/10.1148/radiographics.12.6.1439017

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength & Conditioning Research*, *19*(1), 231-240.

Wilkin, D., Hallam, L., & Doggett, M. A. (1992). *Measures of Need and Outcome for Primary Health Care*. Oxford University Press.

Zou, K. H., O'Malley, A. J., & Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, *115*(5), 654–657. https://doi.org/10.1161/CIRCULATIONAHA.105.594929

Zullig, K. J., Valois, R. F., & Drane, J. W. (2005). Adolescent distinctions between quality of life and self-rated health in quality of life research. *Health and Quality of Life Outcomes*, *3*(1), 1-9.

Tables

Table 2.1

Study Definitions and Terminology DPA SF-8

Terminology	Definition
Physically Active	"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 days per week." (Vela & Denegar, 2010b)
Injury Classification (Vela & Denegar, 2010b)	
Healthy	"Free from musculoskeletal injury and fully able to participate in sport or activity."
Acute Injury	"A musculoskeletal injury that precludes full participation in sport or activity for at least 2 consecutive days (0-72 hours post-injury)."
Subacute Injury	"A musculoskeletal injury that precludes full participation in sport or activity for at least 2 consecutive days (3 days to 1-month post-injury)."
Persistent Injury	"A musculoskeletal injury that has been symptomatic for at least 1 month."
Chronic Injury	"Pain that consistently does not get any better with routine treatment or non-narcotic medication."
Activity Level Classification (U.S. Department of Health and Human Services, 2018)	
Extremely low	"No activity beyond baseline activity."
Low	"Activity beyond baseline, but fewer than 150 minutes of moderate intensity exercise per week."
Medium	"150-300 minutes of moderate intensity activity per week".
High	"More than 300 minutes of moderate intensity activity per week."
Athlete Level (Vela & Denegar, 2010b)	
Competitive	"A participant who engages in a sport activity that requires at least 1 pre-participation examination, regular attendance at scheduled practices and/or conditioning sessions and a coach who leads practices and/or competitions."
Recreational	"Participants who meet the criteria for physical activity and participate in sport, but do not meet the criteria for competitive status."
Occupational	"Participants who meet the criteria for physical activity for occupation or recreation, but do not meet the criteria for competitive or recreational athlete."
Physically Active in Activities of Daily Living	"Participants who do not meet the criteria for any "athlete" category, but who are physically active through their daily activities (e.g., physically active for at least 30 minutes per day, 3 days per week)."

Table 2.2

Demographic Information DPA SF-8

	Full Sample (n = 478)
Sex	N, %
Male	216 (45.2)
Female	236 (49.4)
Other	2 (0.4)
Unknown	24 (5.0)
Activity Level	
Extremely Low	25 (5.2)
Low	105 (22.0)
Medium	179 (37.4)
High	133 (27.8)
Unknown	36 (7.5)
Occupational	
Competitive Athlete	48 (10.0)
Recreational Athlete	159 (33.3)
Occupational Athlete	25 (5.2)
Activities of Daily Living	126 (26.4)
Unknown	120 (25.1)
Injury Category	
Persistent Injury	177 (37.0)
Acute Injury	69 (14.4)
Sub-Acute Injury	89 (18.6)
Healthy	30 (6.3)
Unknown	113 (23.6)
Ethnicity	
Caucasian/White	380 (79.5)
African American	7 (1.5)
Hispanic	27 (5.6)
Asian/Pacific Islander	25 (5.2)
Other	14 (2.9)
Unknown	25 (5.2)

Table 2.3

Group Mean Scores of the DPA SF-8 by Visit and Injury Classification

DPA SF-8 Collection Visit	Injury Category	N	Mean ± SD
Visit 1	Persistent Injury	177	12.69 ± 5.46
	Acute Injury	69	11.99 ± 5.39
	Sub-Acute Injury	89	10.75 ± 4.60
	Healthy	30	3.07 ± 4.11
	Total	365	11.3 ± 5.74
Visit 2	Persistent Injury	135	9.87 ± 5.71
	Acute Injury	52	7.25 ± 5.09
	Sub-Acute Injury	66	8.17 ± 4.70
	Healthy	29	1.03 ± 2.04
	Total	282	8.08 ± 5.71
Visit 3	Persistent Injury	104	8.00 ± 6.26
	Acute Injury	35	5.83 ± 4.89
	Sub-Acute Injury	47	6.98 ± 4.94
	Healthy	28	1.12 ± 2.04
	Total	214	6.52 ± 5.8

Table 2.4

Correlations between Second-Order Latent Variable DPA SF-8 and Other Instruments

DPA SF-8	Visit 1	Visit 2	Visit 3
GF	-.63	-.56	-.65
NPRS	.58	.80	.78
PSFS	-.51	-.69	-.65

Table 2.5

Correlations between First-Order Latent Variables of the DPA SF-8 and Other Instruments

DPA SF-8 PHY	Visit 1	Visit 2	Visit 3
GF	-.55	-.58	-.64
NPRS	.57	.66	.72
PSFS	-.51	-.64	-.64
DPA SF-8 QOL	**Visit 1**	**Visit 2**	**Visit 3**
GF	-.29	-.24	-.33
NPRS	.23	.44	.43
PSFS	-.21	-.34	-.33

Table 2.6

Goodness-of-fit indices for measurement invariance analyses across visit DPA SF-8

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Visit 1	20.97	19	----	.997	----	.995	.023
Visit 2	30.40	19	----	.984	----	.976	.054
Visit 3	30.66	19	----	.988	----	.982	.055
Configural (equal form)	278.46	213	----	.981	----	.976	.039
Metric (equal loadings)	294.18	225	15.72(12)	.980	.001	.976	.039
Equal factor variances	310.43	229	31.97(16)	.977	.003	.972	.042
Scalar (equal indicator intercepts)	327.38	237	**48.92(24)**	.974	.006	.970	.043
Equal latent means	452.23	241	**173.77(28)**	.940	**.04**	.931	.065

Table 2.7

Goodness-of-fit indices for measurement invariance analyses across sex DPA SF-8

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Males (n = 216)	19.66	19	----	.999	----	.999	.013
Females (n = 236)	20.60	19	----	.998	----	.997	.019
Configural (equal form)	40.26	38	----	.999	----	.998	.011
Metric (equal loadings)	51.86	44	11.60(6)	.995	.004	.994	.020
Equal factor variances	52.16	46	11.90(8)	.996	.003	.996	.017
Scalar (equal indicator intercepts)	59.98	50	19.72(12)	.994	.005	.993	.021
Equal latent means	65.27	52	25.01(14)	.992	.007	.991	.024

Table 2.8

Goodness-of-fit indices for measurement invariance analyses across activity level DPA SF-8

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Low (n = 102)	21.04	19	----	.995	----	.992	.033
Medium (n = 169)	19.01	19	----	1.00	----	1.00	.002
High (n = 121)	20.51	19	----	.996	----	.993	.026
Configural (equal form)	60.58	57	----	.997	----	.996	.013
Metric (equal loadings)	76.75	69	16.17(12)	.994	.003	.992	.017
Equal factor variances	87.24	73	26.66(16)	.988	.009	.986	.022
Scalar (equal indicator intercepts)	85.93	81	25.35(24)	.996	.001	.996	.013
Equal latent means	95.91	85	35.33(28)	.991	.006	.991	.018

Table 2.9

Goodness-of-fit indices for measurement invariance analyses across injury category DPA SF-8

	χ^2	df	χ^2_{diff} (df_{diff})	CFI	CFI_{diff}	TLI	RMSEA
Persistent (n = 177)	10.94	19	----	1.00	----	1.00	.000
Acute/subacute (n = 161)	18.79	19	----	1.00	----	1.00	.000
Configural (equal form)	30.89	38	----	1.00	----	1.00	.000
Metric (equal loadings)	40.08	44	9.19(6)	1.00	NC	1.00	.000
Equal factor variances	40.88	46	9.99(8)	1.00	NC	1.00	.000
Scalar (equal indicator intercepts)	43.47	50	12.58(12)	1.00	NC	1.00	.000
Equal latent means	55.22	52	24.33(14)	.997	.003	.997	.014

Figure 2.1

Scale Structure of the Disablement in the Physically Active Scale

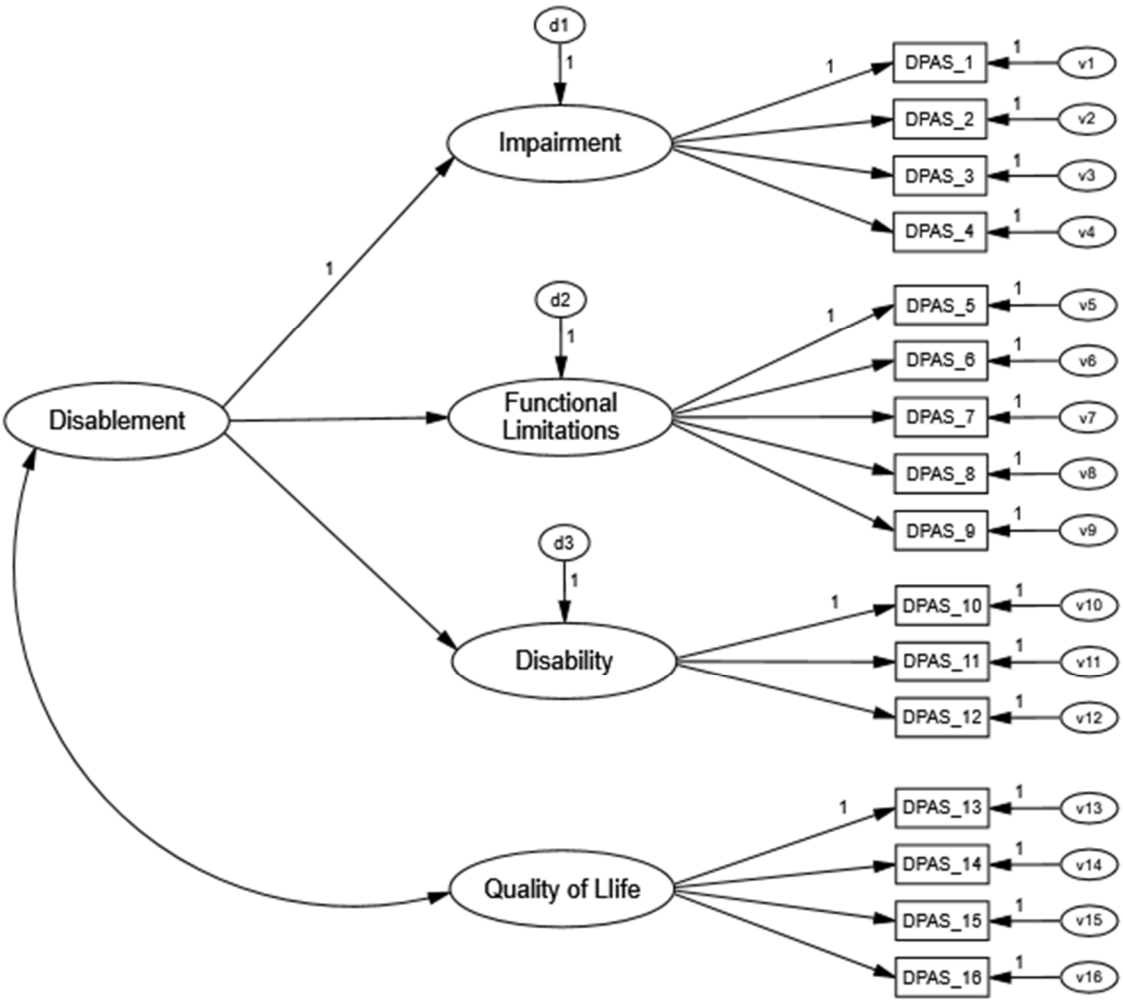


Figure 2.2

Scale Structure of the Disablement in Physically Active Scale Short Form-8

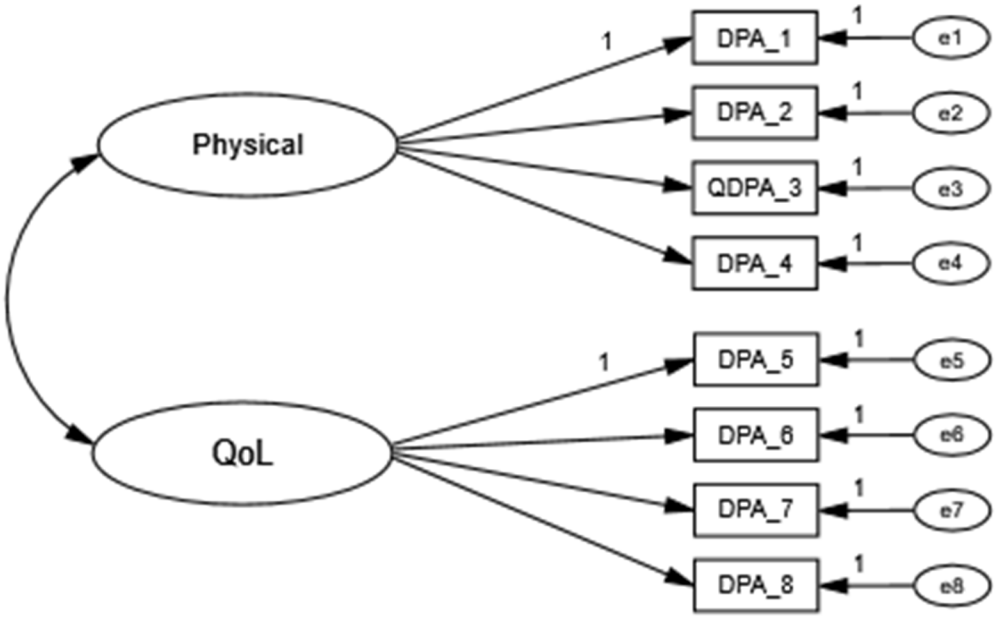


Figure 2.3

Confirmatory Factor Analysis Disablement in Physically Active Scale Short Form-8 Visit 1

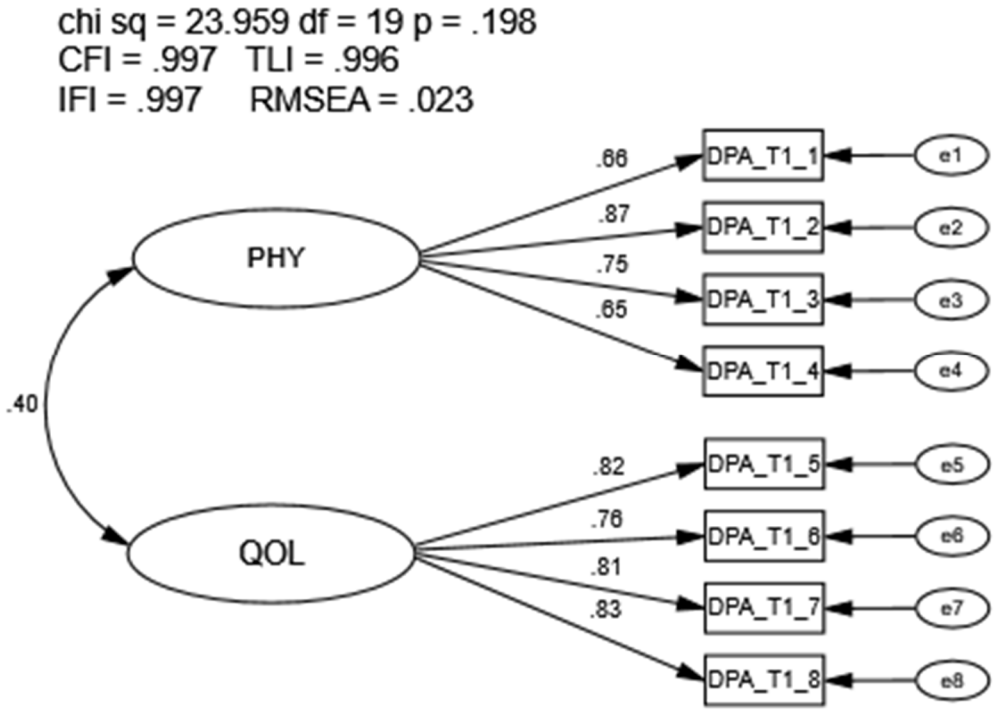


Figure 2.4

Confirmatory Factor Analysis Disablement in Physically Active Scale Short Form-8 Time 2

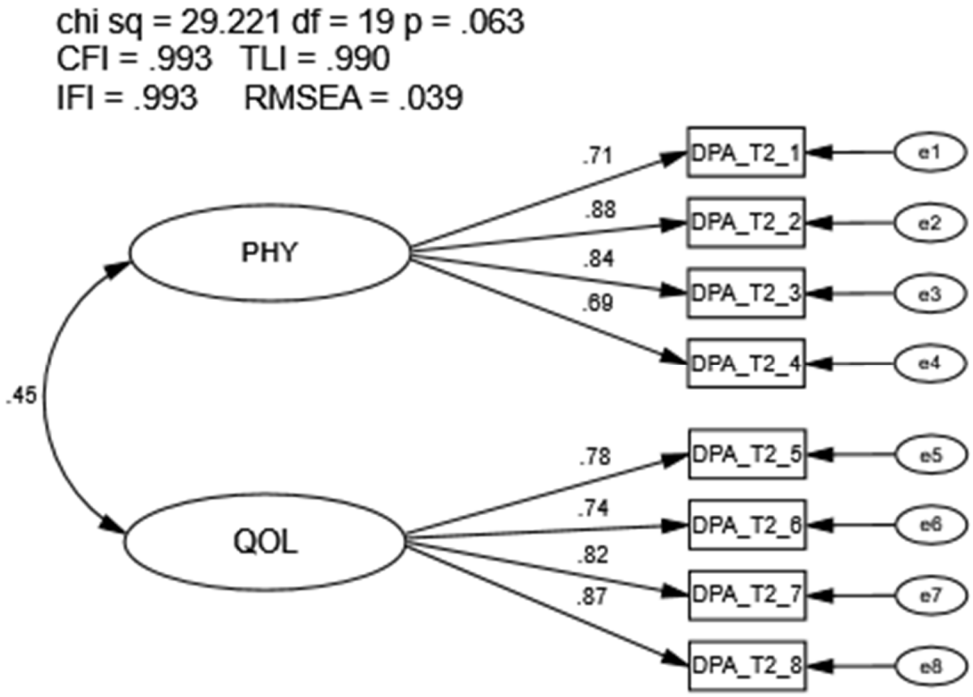


Figure 2.5

Confirmatory Factor Analysis Disablement in Physically Active Scale Short Form-8 Visit 3

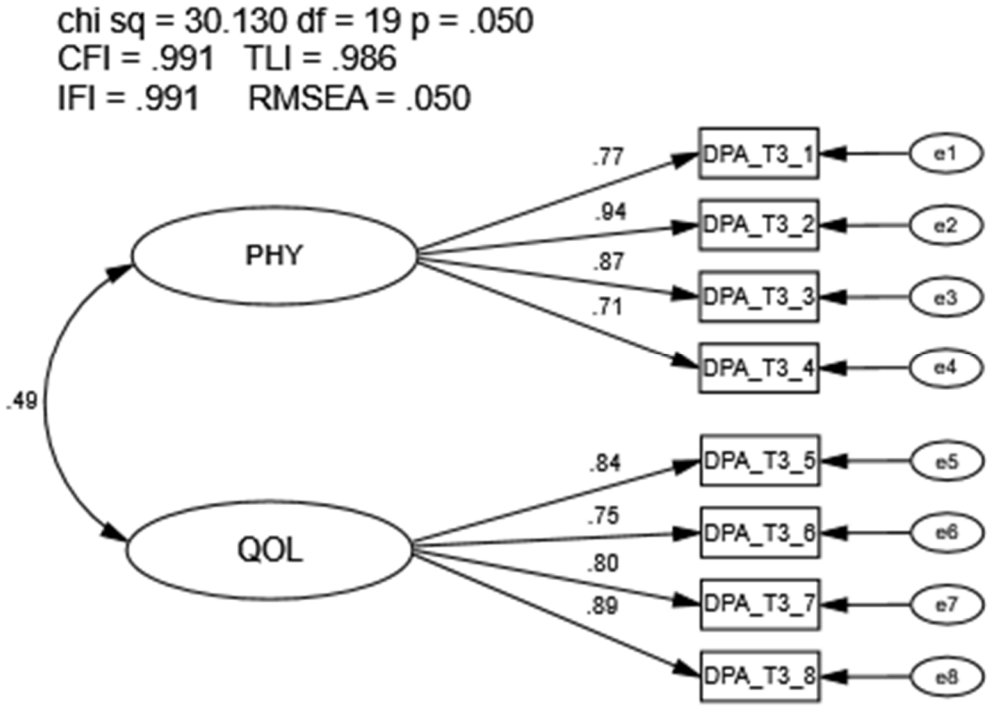
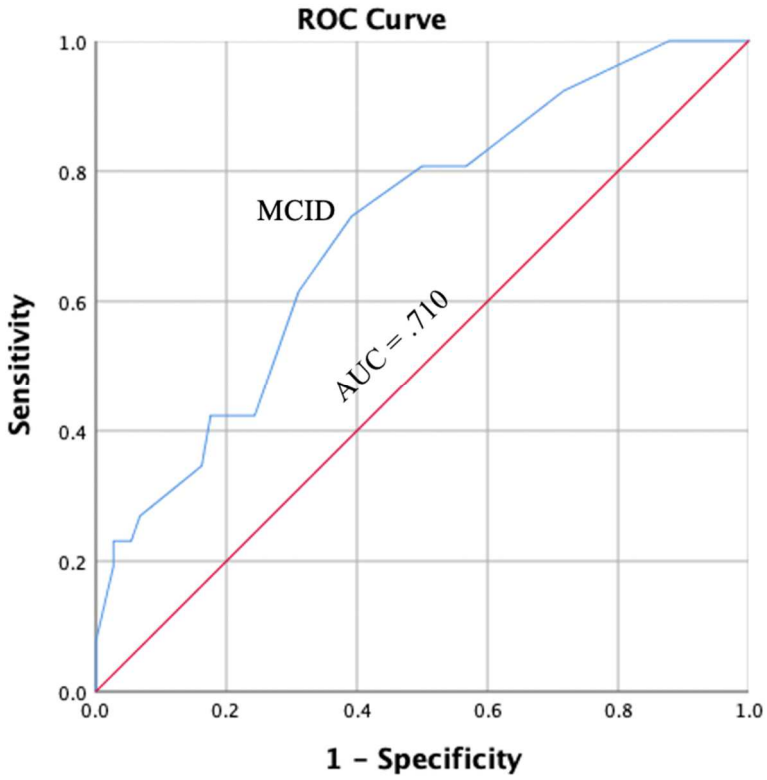


Figure 2.6

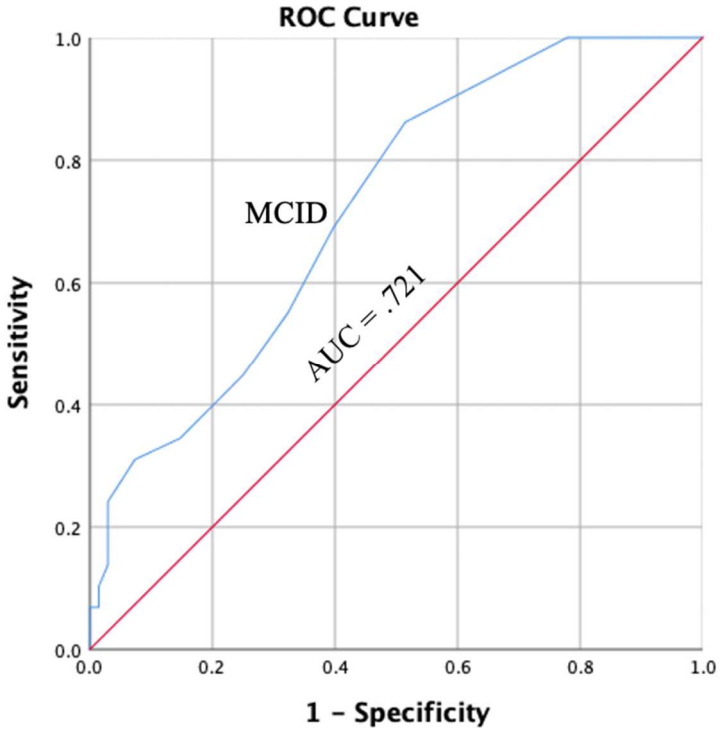
Receiver Operating Curve for Individuals with Persistent Injuries, Visit Two



Abbreviations: AUC, area under the curve; MCID, minimal clinically important difference

Figure 2.7

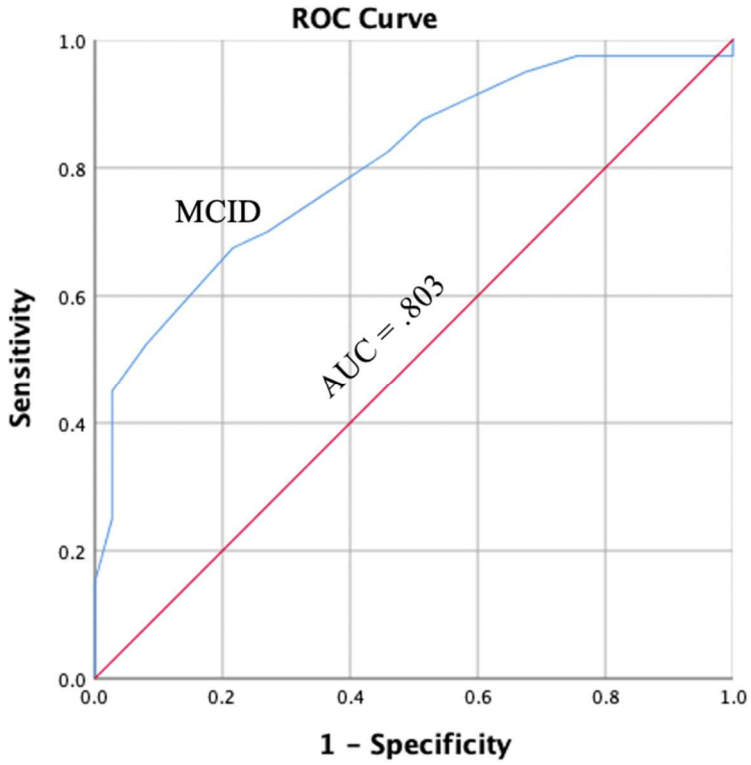
Receiver Operating Curve for Individuals with Persistent Injuries, Visit Three



Abbreviations: AUC, area under the curve; MCID, minimal clinically important difference

Figure 2.8

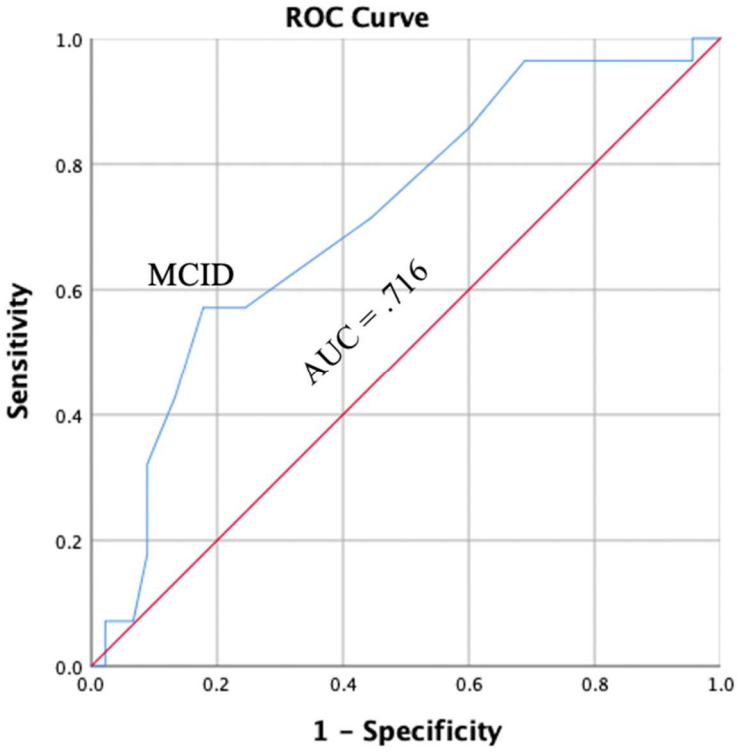
Receiver Operating Curve for Individuals with Acute and Subacute Injuries, Visit Two



Abbreviations: AUC, area under the curve; MCID, minimal clinically important difference

Figure 2.9

Receiver Operating Curve for Individuals with Acute and Subacute Injuries, Visit Three



Abbreviations: AUC, area under the curve; MCID, minimal clinically important difference

Appendix A: Institutional Review Board Letter of Approval

University of Idaho
Office of Research Assurances
Institutional Review Board
875 Perimeter Drive, MS 3010
Moscow, ID 83844-3010
Phone: 208-885-6162
Fax: 208-885-6014
Email: irb@uidaho.edu

February 20, 2020

To: Russell Thomas Baker Jr.

Cc: Madeline Casanova

From: University of Idaho Institutional Review Board

Approval Date: February 20, 2020

Title: Psychosocial Aspects of Pain and Well-Being

Protocol: 20-007, Reference: 007352

Exempt under Category 2 at 45 CFR 46.104(d)(2).

On behalf of the Institutional Review Board at the University of Idaho, I am pleased to inform you that the protocol for this research project has been certified as exempt under the category listed above.

This certification is valid only for the study protocol as it was submitted. Studies certified as Exempt are not subject to continuing review and this certification does not expire. However, if changes are made to the study protocol, you must submit the changes through VERAS for review before implementing the changes. Amendments may include but are not limited to, changes in study population, study personnel, study instruments, consent documents, recruitment materials, sites of research, etc.

As Principal Investigator, you are responsible for ensuring compliance with all applicable FERPA regulations, University of Idaho policies, state and federal regulations. Every effort should be made to ensure that the project is conducted in a manner consistent with the three fundamental principles identified in the Belmont Report: respect for persons; beneficence; and justice. The Principal Investigator is responsible for ensuring that all study personnel have completed the online human subjects training requirement. Please complete the *Continuing Review and Closure Form* in VERAS when the project is completed.

You are required to notify the IRB in a timely manner if any unanticipated or adverse events occur during the study, if you experience an increased risk to the participants, or if you have participants withdraw or register complaints about the study.

IRB Exempt Category (Categories) for this submission:
Category 2: Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met: i. The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot

University of Idaho
Office of Research Assurances
Institutional Review Board
875 Perimeter Drive, MS 3010
Moscow, ID 83844-3010
Phone: 208-885-6162
Fax: 208-885-6014
Email: irb@uidaho.edu

readily be ascertained, directly or through identifiers linked to the subjects; ii. Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation; or iii. The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by .111(a)(7).

Appendix B: Orbach and Mikulincer Pain Scale

Instructions: Using the scale below, indicate how strongly you agree or disagree with each statement.

Likert Scale: 0: Strongly disagree 1: Disagree 2: Agree to some extent 3: Agree 4: Strongly agree	Strongly disagree	Disagree	Agree to some extent	Agree	Strongly agree
1. Nobody is interested in me.	O	O	O	O	O
2. I am completely helpless.	O	O	O	O	O
3. I feel an emotional turmoil inside me.	O	O	O	O	O
4. I cannot do anything at all.	O	O	O	O	O
5. I will fall apart.	O	O	O	O	O
6. I am afraid of the future.	O	O	O	O	O
7. I am rejected by everybody.	O	O	O	O	O
8. I am flooded by many feelings.	O	O	O	O	O
9. I am completely defeated.	O	O	O	O	O
10. I have lost something that I will never find again.	O	O	O	O	O
11. I feel numb and not alive.	O	O	O	O	O
12. I feel abandoned and lonely.	O	O	O	O	O
13. I have no control over my life.	O	O	O	O	O
14. My feelings change all the time.	O	O	O	O	O
15. I am a stranger to myself.	O	O	O	O	O
16. Others hate me.	O	O	O	O	O
17. I feel that I am not my old self anymore.	O	O	O	O	O
18. I am worthless.	O	O	O	O	O
19. I feel paralyzed	O	O	O	O	O
20. I cannot concentrate.	O	O	O	O	O

21. I cannot trust myself.	O	O	O	O	O
22. The difficult situation will never change.	O	O	O	O	O
23. I feel as if I am not real.	O	O	O	O	O
24. I have difficulties in thinking.	O	O	O	O	O
25. I need the support of other people. ®	O	O	O	O	O
26. The world has changed forever.	O	O	O	O	O
27. I feel confused.	O	O	O	O	O
28. I have no control over what is happening inside me.	O	O	O	O	O
29. I will never be able to reduce my pain.	O	O	O	O	O
30. My life has stopped.	O	O	O	O	O
31. I have no idea what to expect of the future.	O	O	O	O	O
32. Something in my life was damaged forever.	O	O	O	O	O
33. There is uncertainty about my life and myself.	O	O	O	O	O
34. I will never be the same person.	O	O	O	O	O
35. There are strong ups and downs in my feelings.	O	O	O	O	O
36. I have no control over the situation.	O	O	O	O	O
37. I want to be left alone.	O	O	O	O	O
38. I have no future goals.	O	O	O	O	O
39. I have no desires	O	O	O	O	O
40. I don't feel like talking to other people.	O	O	O	O	O
41. I can't find meaning in my life.	O	O	O	O	O
42. I can't stay alone. ®	O	O	O	O	O
43. I can't change what is happening to me.	O	O	O	O	O
44. The pain will never go away.	O	O	O	O	O

®=reverse scored

Appendix C: Numeric Pain Rating Scale

Instructions: Please indicate the intensity of your current, best, and worst pain levels **over the past 24 hours** on a scale of 0 (no pain) to 10 (worst pain Imaginable).

Current Pain Rating: _____ Best Pain Rating: _____ Worst Pain Rating: _____

Appendix D: Patient Health Questionnaire-9

Instructions: **Over the past 2 weeks**, how often have you been bothered by any of the following problems?

Likert Scale: 0: Not At All 1: Several Days 2: More Than Half the Days 3: Nearly Every Day	Not at All	Several Days	More Than Half the days	Nearly Every Day
	0	**1**	**2**	**3**
1. Little interest or pleasure in doing things.	O	O	O	O
2. Feeling down, depressed or hopeless.	O	O	O	O
3. Trouble falling asleep, staying asleep, or sleeping too much.	O	O	O	O
4. Feeling tired or having little energy.	O	O	O	O
5. Poor appetite or overeating.	O	O	O	O
6. Feeling bad about yourself – or that you're a failure or have let yourself or your family down.	O	O	O	O
7. Trouble concentrating on things, such as reading the newspaper or watching television.	O	O	O	O
8. Moving or speaking so slowly that other people could have notice. Or, the opposite – being so fidgety or restless that you have been moving around a lot more than usual.	O	O	O	O
9. Thoughts that you would be better off dead or of hurting yourself in some way.	O	O	O	O

10. If you checked off any problems, how difficult have those problems made it for you to do your work, take care of things at home, or get along with other people?

O	O	O	O
Not difficult at all	Somewhat difficult	Very difficult	Extremely difficult

Appendix E: Self-Compassion Scale

Instructions: Please indicate how often you acted in the manner stated in each of the items on a scale of 1 (almost never) to 5 (almost always).

	Almost Never				Almost Always
	1	**2**	**3**	**4**	**5**
1. I try to be understanding and patient towards those aspects of my personality I don't like.	O	O	O	O	O
2. I'm kind to myself when I'm experiencing suffering.	O	O	O	O	O
3. When I'm going through a very hard time, I give myself the caring and tenderness I need.	O	O	O	O	O
4. I'm tolerant of my own flaws and inadequacies.	O	O	O	O	O
5. I try to be loving towards myself when I'm feeling emotional pain.	O	O	O	O	O
6. When I see aspects of myself that I don't like, I get down on myself.	O	O	O	O	O
7. When times are really difficult, I tend to be tough on myself.	O	O	O	O	O
8. I can be a bit cold-hearted towards myself when I'm experiencing suffering.	O	O	O	O	O
9. I'm disapproving and judgmental about my own flaws and inadequacies.	O	O	O	O	O
10. I'm intolerant and impatient towards those aspects of my personality I don't like.	O	O	O	O	O
11. When I feel inadequate in some way, I try to remind myself that feelings of inadequacy are shared by most people.	O	O	O	O	O
12. I try to see my failings as part of the human condition.	O	O	O	O	O
13. When I'm down and out, I remind myself that there are lots of other people in the world feeling like I am.	O	O	O	O	O
14. When things are going badly for me, I see the difficulties as part of life that everyone goes through.	O	O	O	O	O
15. When I fail at something that's important to me, I tend to feel alone in my failure.	O	O	O	O	O
16. When I think about my inadequacies, it tends to make me feel more separate and cut off from the rest of the world.	O	O	O	O	O

17. When I'm feeling down, I tend to feel like most other people are probably happier than I am.	O	O	O	O	O
18. When I'm really struggling, I tend to feel like other people must be having an easier time of it.	O	O	O	O	O
19. When something upsets me, I try to keep my emotions in balance.	O	O	O	O	O
20. When I'm feeling down, I try to approach my feelings with curiosity and openness.	O	O	O	O	O
21. When something painful happens, I try to take a balanced view of the situation.	O	O	O	O	O
22. When I fail at something important to me, I try to keep things in perspective.	O	O	O	O	O
23. When something upsets me, I get carried away with my feelings.	O	O	O	O	O
24. When I'm feeling down, I tend to obsess and fixate on everything that's wrong.	O	O	O	O	O
25. When something painful happens, I tend to blow the incident out of proportion.	O	O	O	O	O
26. When I fail at something important to me, I become consumed by feelings of inadequacy.	O	O	O	O	O

Appendix F: Depression Anxiety Stress Scale - 21

Instructions: Please read each statement and mark the bubble indicating how much the statement applied to you **over the past week**. There are no right or wrong answers. Do not spend too much time on any statement.

Likert Scale: 0: Did not apply to me at all 1: Applied to me to some degree, or some of the time 2: Applied to me to a considerable degree or a good part of time 3: Applied to me very much or most of the time	Did not apply to me at all	Applied to me to some degree, or some of the	Applied to me to a considerable degree or	Applied to me very much or most of the
Over the past week…	**0**	**1**	**2**	**3**
1. I found it hard to wind down.	O	O	O	O
2. I was aware of dryness of my mouth.	O	O	O	O
3. I couldn't seem to experience any positive feeling at all.	O	O	O	O
4. I experienced breathing difficulty (e.g. excessively rapid breathing, breathlessness in the absence of physical exertion).	O	O	O	O
5. I found it difficult to work up the initiative to do things.	O	O	O	O
6. I tended to over-react to situations.	O	O	O	O
7. I experienced trembling (e.g. in the hands).	O	O	O	O
8. I felt that I was using a lot of nervous energy.	O	O	O	O
9. I was worried about situations in which I might panic and make a fool of myself.	O	O	O	O
10. I felt that I had nothing to look forward to.	O	O	O	O
11. I found myself getting agitated.	O	O	O	O
12. I found it difficult to relax.	O	O	O	O
13. I felt down-hearted and blue.	O	O	O	O
14. I was intolerant of anything that kept me from getting on with what I was doing.	O	O	O	O

15. I felt I was close to panic.	O	O	O	O
16. I was unable to become enthusiastic about anything.	O	O	O	O
17. I felt I wasn't worth much as a person.	O	O	O	O
18. I felt that I was rather touchy.	O	O	O	O
19. I was aware of the action of my heart in the absence of physical exertion (e.g. sense of heart rate increase, heart missing a beat).	O	O	O	O
20. I felt scared without any good reason.	O	O	O	O
21. I felt that life was meaningless.	O	O	O	O

Appendix G: Participant Questionnaire

1. What is your self-identified sex?
   Male
   Female
   Prefer not to answer
2.
3. What is your age? _____ (years)
4.
5. What is your ethnicity?
   Caucasian/White
   African American/Black
   Hispanic
   Asian/Pacific Islander
   Other (please specify) _____
6.
7. What is the highest level of education you have completed? Please select the best choice.
   Some high school but no diploma or degree
   High school diploma or GED
   Some college but no degree
   Associate degree
   Bachelor's Degree
   Master's Degree
   Doctorate Degree (PhD, EdD, etc.)
   Other (please specify) _____
8.
9. How would you rate your overall health?
   Excellent
   Very good
   Good
   Fair
   Poor
10.
11. Do you *currently* have a physical injury?
    No, I am healthy.
    Yes, I have an ***acute*** injury (A musculoskeletal injury that occurred within the past 3 days and prevents full participation in sport/activity for at least 2 consecutive days).
    Yes, I have a ***sub-acute*** injury (A musculoskeletal injury that occurred between 4 and 30 days ago and prevented full participation in sport/activity for at least 2 consecutive days).
    Yes, I have a ***persistent*** injury (A musculoskeletal injury that has been symptomatic for at least 1 month).
    Yes, I have a ***chronic*** injury (Pain that consistently does not get any better with routine treatment or non-narcotic medication).
12.
13. How would you describe your *current* physical activity level? (Baseline activity refers to light-intensity activities of daily life [e.g., standing, walking, lifting lightweight objects]; moderate activity includes activities such as brisk walking, yoga, and lifting weights).

Inactive: No activity beyond baseline activity
Low: activity beyond baseline, but fewer than 150 minutes of moderate intensity exercise per week.
Medium: 150 to 300 minutes of moderate intensity activity per week.
High: more than 300 minutes of moderate intensity activity per week.

14.

15. Do you currently engage in athletic, recreational, or occupational activities that require physical skills and use strength, power, endurance, speed, flexibility, range of motion, or agility at least 3 days per week?
    Yes
    No

16.

17. If you answered **yes** to the last question, which definition most closely align to you?
    Competitive athlete: Someone who engages in a sport activity that requires at least 1 pre-participation examination, regular attendance at schedules practices and/or conditioning sessions, and a coach who leads practices and/or competitions.
    Recreational athlete: Someone who meets the criteria for physical activity and participation in sport but does not meet the criteria for competitive athlete.
    Occupational athlete: Someone who meets the criteria for physical activity for occupation or recreation but does not meet the criteria for competitive or recreational athlete.
    Activities of daily living: Someone who does not meet the criteria for any "athlete" category, but who is physically active on a daily basis through their daily activities.
    Not applicable

18.

19. Have you ever been diagnosed with a mental illness?
    Yes
    No
    Prefer not to answer

20. If you have been diagnosed with a mental illness (past and/or current), what diagnosis were you given? Select all that apply.
    Depression
    Anxiety
    Post-Traumatic Stress Disorder
    Eating Disorder
    Schizophrenia
    Bipolar Disorder
    Substance Use Disorder
    Other (please specify) _____
    I have never been diagnosed with a mental illness

Appendix H: Orbach and Mikulincer Mental Pain Scale – 8

Instructions: Using the scale below, indicate how strongly you agree or disagree with each statement.

Likert Scale: 0: Strongly disagree 1: Disagree 2: Agree to some extent 3: Agree 4: Strongly agree	Strongly disagree	Disagree	Agree to some extent	Agree	Strongly agree
1. The pain will never go away.	O	O	O	O	O
2. I am flooded by many feelings.	O	O	O	O	O
3. I am rejected by everybody.	O	O	O	O	O
4. I will never be able to reduce my pain.	O	O	O	O	O
5. There are strong ups and downs in my feelings.	O	O	O	O	O
6. Nobody is interested in me.	O	O	O	O	O
7. My feelings change all the time.	O	O	O	O	O
8. Others hate me.	O	O	O	O	O

Appendix I: Disablement in Physically Active Scale Short Form – 8

Instructions: Please answer **each statement** with one response by shading in the circle that most closely describes your problem(s) within the past **24 hours**. Each problem has possible descriptors under each. Not all descriptors may apply to you but are given as common examples.

KEY 1 – No Problem 2 – I have the problem(s), but it does not affect me 3 – The problem(s) slightly affects me 4 – The problem(s) moderately affects me 5 – The problem(s) severely affects me	No Problem	Does not Affect	Slight	Moderate	Severe
Physical Component	1	2	3	4	5
Pain – "Do I have **pain**?"	O	O	O	O	O
Motion – "Do I have impaired **motion**?" Ex. Decreased range/ease of motion, flexibility, and/or increased stiffness	O	O	O	O	O
Muscular Functioning – "Do I have impaired **muscle function**?" Ex. decreased strength, power, endurance, and/or increased fatigue	O	O	O	O	O
Changing Directions – "Do I have difficulty with **changing directions** in activity?" Ex. twisting, turning, starting/stopping, cutting, pivoting	O	O	O	O	O
Physical Score (Total Score – 4)	_____ / 16				
Quality of Life Component	1	2	3	4	5
Well Being – "Do I have difficulties with the following…?" 1) Increased uncertainty, stress pressure anxiety	O	O	O	O	O
Well Being – "Do I have difficulties with the following…?" 2) Altered relationships with team, friends, and/or colleagues	O	O	O	O	O
Well Being – "Do I have difficulties with the following…?" 3) Decreased overall energy	O	O	O	O	O
Well Being – "Do I have difficulties with the following…?" 4) Changes in my mood and/or increased frustration	O	O	O	O	O
Quality of Life Score (Total Score – 4)	_____ / 16				
DPA SF-8 Total Score (Sum Construct Scores)	_____ / 32				

Appendix J: Numeric Pain Rating Scale

Please indicate the intensity of your current, best, and worst pain levels over the past 24 hours on a scale of 0 (no pain) to 10 (worst pain Imaginable).

Current Pain Rating: _____

Best Pain Rating: _____

Worst Pain Rating: _____

Appendix K: Global Functioning Scale

Please circle the number that most closely represents your overall level of functioning. Please use the scale below, where "0" represents being unable to function at your normal level and "100" represents being able to function completely at your normal level of function before the injury/problem.

0-------10-------20-------30-------40-------50-------60-------70-------80-------90-------100

Appendix L: Patient Specific Functional Scale

Please indicate your ability to perform ***three*** important activities that you are unable to do or are having difficulty with as a result of your injury/problem. Please rate your ability on a scale of 0 (unable to perform activity) to 10 (able to perform activity at the same level as before injury or problem).

Functional Activity Rating 1: _____

Activity 1 Selected: _____

Functional Activity Rating 2: _____

Activity 2 Selected: _____

Functional Activity Rating 3: _____

Activity 3 Selected: _____

Appendix M: Global Rating of Change Scale

With respect to your injury/problem, how would you describe yourself now compared to immediately after your first noticed the injury/problem (please circle only one):

A very great deal worse (-7)	Unchanged (0)	A tiny bit better (1)
A great deal worse (-6)		A little bit better (2)
Quite a bit worse (-5)		Somewhat better (3)
Moderately worse (-4)		Moderately better (4)
Somewhat worse (-3)		Quite a bit better (5)
A little bit worse (-2)		A great deal better (6)
A tiny bit worse (-1)		A very great deal better (7)

Appendix N: Participant Demographic Items

1. How long have you been experiencing your health condition/pain/injury? Please select the best choice:

Less than 24 hours
24-72 hours
3 days to 1 week
1 to 4 weeks
1 to 6 months
6 months to 1 year
More than 1 year

2. How would you describe your current physical activity level (Baseline activity refers to light-intensity activities of daily life [e.g., standing, walking, lifting lightweight objects]; moderate activity includes activities such as brisk walking, yoga, and lifting weights.)?

Extremely Low: No activity beyond baseline activity.
Low: Activity beyond baseline, but fewer than 150 minutes of moderate intensity exercise per week.
Medium: 150 to 300 minutes of moderate intensity activity per week.
High: More than 300 minutes of moderate intensity activity per week.

3. What is your self-identified sex?

Male
Female
Prefer not to answer

4. What is your age (in years)? _____

5. What is your ethnicity? Please select the best choice.

Caucasian/White
African American/Black
Hispanic
Asian/Pacific Islander
Other (please specify) _____

6. What is the highest education level you have completed? Please circle the best choice.

High school diploma or GED
Some college but no degree
Associate degree
Bachelor's Degree
Master's Degree
Doctorate Degree (PhD, EdD, etc.)
Other (please specify) _____