Fear of movement in children and adolescents undergoing major surgery: A psychometric evaluation of the Tampa Scale for Kinesiophobia

Abstract

Background: The objective of this study was to evaluate the psychometric properties of the 17-item Tampa Scale for Kinesiophobia (TSK) in youth. Methods: Participants were 264 children and adolescents (58.7% female, M_{age}=14.1 years, SD_{age}=2.51) scheduled for major surgery who were assessed before surgery, while in hospital postoperatively, and at 6 and 12 months after surgery. Exploratory factor analyses (EFA) were conducted to determine the factor structure of pre-operative TSK scores. Reliability, and convergent, discriminant, and predictive validity were examined. Results: EFA on the 17-item TSK revealed a two-factor model distinguishing the 13 positively scored items from the 4 reverse scored items, but the fit was poor. A second EFA was conducted on the 13 positively scored items (TSK-13) revealing a three-factor model: Fear of injury, bodily vulnerability, and activity avoidance. The TSK-13 showed adequate internal consistency (Ω=0.82) and weak convergent validity. The TSK-13 was not correlated with postoperative, in-hospital physical activity (actigraphy; r (179) =-0.10, p=0.18) and showed adequate discriminant validity, that is correlations less than .70, with measures of depression (r(225)=0.41, p<0.001) and general anxiety (r(224)=0.35, p<0.001). Predictive validity for pain-related disability at 12 months (r(70)=0.34, p<0.001) was adequate. Conclusions: The original TSK-17 does not appear to be a meaningful measure of kinesiophobia in youth after surgery possibly because of the syntactic structure of the reverse scored items. In contrast, a modified TSK-13, comprised of only the positively scored items, revealed a 3-factor structure that is reliable and demonstrates adequate convergent, discriminant, and predictive validity.

Significance: Kinesiophobia is an important construct to evaluate in the transition from acute to chronic pain among children and adolescents. The 17 item Tampa Scale for Kinesiophobia (TSK) does not show adequate validity or reliability in youth undergoing major surgery, however, the psychometric properties of a 13-item modified scale (TSK-13) are promising.

Article type   : Original Manuscript

Corresponding author mail id: bnrosen@yorku.ca
Fear of movement in children and adolescents undergoing major surgery: A psychometric evaluation of the Tampa Scale for Kinesiophobia

Chronic post-surgical pain and disability are surgical complications reported by 11-54% of children and adolescents after major surgery (Chidambaran et al., 2017; Landman, Oswald, Sanders, Diab, & Spinal Deformity Study, 2011; Rosenbloom et al., 2019; Sieberg et al., 2013). In adults, fear of pain is one of the most prominent predictors of chronic pain and pain-related disability (Keefe, Rumble, Scipio, Giordana, & Perri, 2004; Vlaeyen & Linton, 2000), and, more specifically, fear of physical performance and perceived disability (Al-Obaidi, Nelson, Al-Awadhi, & Al-Shuwaie, 2000; Asmundson, Norton, & Allendings, 1997; Verbunt, Seelen, Vlaeyen, van der Heijden, & Knottnerus, 2003; Vlaeyen, Kole-Snijders, Boeren, & van Eek, 1995). The relationship between pain and disability may be explained more thoroughly by the cognitive-behavioural fear-avoidance model of chronic pain (Vlaeyen et al., 1995; Vlaeyen & Linton, 2000). This model proposes that after injury, one either confronts the pain and recovers, or because of a fear of pain becomes disabled through a pathway involving pain catastrophizing and avoidance.

Asmundson and colleagues (2012) proposed the Pediatric Fear-Avoidance Model of Chronic Pain and recommended use of the Fear of Pain Questionnaire (FOPQ) (Simons, Sieberg, Carpino, Logan, & Berde, 2011), which evaluates kinesiophobia, to empirically test the model (Asmundson, Noel, Petter, & Parkerson, 2012). Kinesiophobia is defined as "an excessive irrational, and debilitating fear of physical movement and activity resulting from a feeling of vulnerability to painful injury or re-injury" (Kori, Miller, & Todd, 1990) that is proposed to predict the development of pain-related disability (Vlaeyen et al., 1995). The FOPQ was developed using adult measures (i.e. Tampa Scale for Kinesiophobia (Kori et al., 1990)),

Psychological Inflexibility in Pain Scale, Fear Avoidance Beliefs Questionnaire, Pain Anxiety Symptoms Scale-20).

The TSK rarely has been used in pediatric settings. Wicksell et al (Wicksell, Melin, Lekander, & Olsson, 2009) found a significant decrease in TSK scores, measured by a Swedish version, as well as improvements in quality of life following an Acceptance and Commitment Therapy intervention for children with chronic pain. More recently, Ye et al (2020) conducted a principle components analysis and confirmatory factor analysis on the TSK in 55 adolescents undergoing spinal fusion surgery. After excluding four items, the authors report that only the Activity Avoidance subscale of the TSK revealed good psychometric properties (Ye et al., 2020). This finding is not consistent with factor analytic studies conducted in adults showing psychometric soundness for two factors (Goubert et al., 2004; Heuts et al., 2004; Swinkels-Meewisse, Swinkels, Verbeek, Vlaeyen, & Oostendorp, 2003). Moreover, the authors did not examine the convergent, discriminative, or predictive validity of the TSK.

The aim of this study was to evaluate the psychometric properties of the TSK in children and adolescents undergoing major surgery. To achieve this aim, we (1) conducted an exploratory factor analysis (EFA) to determine the factor structure of the TSK, (2) evaluated the convergent and discriminant validity of the TSK, and (3) assessed its predictive validity by examining the relationship between the TSK and chronic pain and pain-related disability 12 months after surgery.

Methods

The present article reports results from a larger study examining the development of chronic post-surgical pain in youth. Complete methods and measures have been reported in Rosenbloom et al. (2019).

Participants

Children/adolescents were eligible to participate if they were aged 8-17 years and scheduled for orthopedic surgery (i.e., osteotomy, plate insertion tibial/femur, surgery for scoliosis) or general surgery (i.e., thoracotomy, thoracoabdominal surgery, Nuss/Ravitch pectus repair, sternotomy, laparotomy, laparoscopic-assisted; colectomy, ileostomy, J-pouches). One of the child's parents/guardians was also invited to participate. Children were excluded if (1) they had a documented developmental or cognitive delay, (2) they had a diagnosis of cancer, (3) they did not speak or read English, or (4) their parent or guardian did not speak or read English.

Measures

Primary Questionnaire. *Tampa Scale for Kinesiophobia (TSK).* The TSK (Kori et al., 1990) is a 17-item scale designed for use in adults to measure fear of movement-evoked pain and injury (Table 1). Items are scored on a 4-point Likert scale as follows: 1 = strongly disagree; 2 = disagree; 3 = agree; 4 = strongly agree. It contains 13 positively scored items for which endorsing *strong agreement* is associated with the highest Likert scale score (4), and 4 reverse scored items (#4, #8, #12, #16), spread out evenly throughout the scale, for which endorsing *strong disagreement* is associated with the highest score (4) after reverse scoring. Total scores range from 17 to 68 with higher scores indicative of greater fear of movement. The TSK has been shown to have good internal consistency with Cronbach alpha's ranging from 0.68 to 0.86 (Swinkels-Meewisse et al., 2003; Vlaeyen et al., 1995). The TSK has demonstrated good test-retest reliability (Swinkels-Meewisse et al., 2003). Some studies favor a two-factor model of the TSK (Clark, Kori, & Brockel, 1996; French, France, Vigneau, French, & Evans, 2007; Goubert et al., 2004), however, others have shown one and three factor models (Roelofs, Goubert, Peters, Vlaeyen, & Crombez, 2004). Some factor analytic studies show that the reverse scored items do not load well onto any factor (Clark et al., 1996; Goubert et al., 2004; Swinkels-Meewisse et al., 2003).

Convergent Validity Measures. *The Numerical Rating Scale (NRS).* The NRS is an 11-point verbally administered scale that measures the subjective experience of pain intensity (I) or pain unpleasantness (U). The NRS-I ranged from 0 (no pain at all) to 10 (worst possible pain). The NRS-U ranged from 0 (not at all unpleasant/horrible/yucky) to 10 (most unpleasant/horrible/yucky). The NRS has excellent reliability and validity, and has been validated for acute postsurgical pain in children ages 7-17 years (Pagé et al., 2012; Von Baeyer et al., 2009).

Child Pain Anxiety Symptoms Scale (CPASS). The CPASS (Pagé, Fuss, Martin, Escobar, & Katz, 2010) is a 20-item scale that measures the fear and anxiety-related thoughts, feelings, behaviors, and physical sensations that accompany the experience and anticipation of pain. Each item is rated on a scale of 0 (never) to 5 (always) and overall scores range from 0 to 100 with higher scores indicative of greater pain-related anxiety. CPASS has excellent internal consistency (α = 0.89 to 0.903) and strong construct validity (Pagé et al., 2011). Internal

consistency for the present study was excellent at T0 ($\alpha = 0.92$), T1 ($\alpha = 0.94$), T2 ($\alpha = 0.93$), and T3 ($\alpha = 0.93$).

   Pain Catastrophizing Scale-Children (PCS-C (Crombez et al., 2003)). The PCS-C is a child version of the PCS (Sullivan, Bishop, & Pivik, 1995) that measures the thoughts and feelings children may experience when they are in pain. Each of the 13 items is rated on a 5-point scale ranging from not at all (0) to all the time (4) with the total scores ranging from 0 to 52 with higher scores indicative of greater pain catastrophizing. The PCS-C yields a total score and three subscale scores assessing (1) rumination, (2) magnification, and (3) helplessness. The PCS-C has excellent internal consistency ($\alpha = 0.90$) and strongly correlates with pain intensity ($r = 0.49$) and disability ($r = 0.50$) (Crombez et al., 2003). Internal consistency for the present study was excellent at T0 ($\alpha = 0.94$), T1 ($\alpha = 0.94$), T2 ($\alpha = 0.93$), and T3 ($\alpha = 0.93$).

   Actigraphy. The Actical movement monitor (Respironics, Inc., Bend, Oregon) is a small, non-invasive, wristwatch-sized device that contains an omnidirectional accelerometer designed to measure physical activity and caloric expenditure on a continuous basis. The movement monitor provides an objective, quantifiable measure of physical movement and has been used in past studies of children in a post-surgical hospital setting (Kudchadkar et al., 2019; Puyau, Adolph, Vohra, Zakeri, & Butte, 2004). The movement monitor was placed on the child's wrist when they arrived in the post-operative care unit after surgery and left in place until they were discharged. The physical activity monitor was placed on the wrist opposite the hand accessed for the intravenous line. The Actical has a sampling frequency of 32Hz. Data were downloaded in epochs of 30 seconds. Total daily (24 hour) movement counts were computed daily starting at 00:00 on the day after surgery until 23:59 the day before hospital discharge.

   Discriminant Validity Measures. *Childhood Anxiety Sensitivity Index (CASI).* The CASI (Silverman, Fleisig, Rabian, & Peterson, 1991) is an 18-item scale that measures the extent to which the symptoms of anxiety (e.g., increased heart rate, shortness of breath, racing thoughts) are feared due to the belief that they will have harmful somatic, psychological, and/or social consequences. Each item is rated on a scale of 1 (none) to 3 (a lot). Total scores range from 18 to 54 with higher scores indicative of greater anxiety sensitivity. The CASI has good internal consistency ($\alpha = 0.87$), satisfactory test-retest reliability ($r = 0.76$) and acceptable construct validity (Silverman et al., 1991).Internal consistency for the present study was very good at T0 ($\alpha = 0.86$), T1 ($\alpha = 0.87$), T2 ($\alpha = 0.85$), and T3 ($\alpha = 0.86$).

Children's Revised Impact of Event Scale (CRIES; (Horowitz, Wilner, & Alvarez, 1979; Smith, Perrin, Dyregrov, & Yule, 2003) The CRIES is a 13-item scale that measures posttraumatic stress disorder (PTSD) symptoms in the previous six months. Each item is rated based on its frequency of occurrence on a 4-point scale, from 0 (none), 1 (rarely), 3 (sometimes) to 5 (a lot), with total scores ranging from 0 to 65. A score of 30 or higher indicates a very likely presence of PTSD. The CRIES has good reliability (α=0.80) (Smith et al., 2003) and high validity as a screening measure for PTSD (Perrin, Meiser-Stedman, & Smith, 2005). Internal consistency for the present study was excellent at T0 ($\alpha = 0.91$), and good at T1 ($\alpha = 0.88$), T2 ($\alpha = 0.88$), and T3 ($\alpha = 0.89$).

Multidimensional Anxiety Scale for Children (MASC-10, and -39). The MASC-10 (March & Sullivan, 1999) is a 10-item, shortened version of the 39-item MASC-39. Both versions measure self-reported physiological responses, harm avoidance, and social and separation anxiety. Items are rated on a scale from 0 (never true about me) to 3 (often true about me). Total scores range from 0-30 (MASC-10) and 0-117 (MASC-39), with higher scores indicating more symptoms of anxiety. The MASC-39 has good internal consistency ($\alpha = 0.60$ to 0.85), strong test-retest reliability ($r = 0.79$ to 0.93), good convergent validity (correlates significantly with the Revised Children's Manifest Anxiety Scale) and also has good discriminant validity (March, Parker, Sullivan, Stallings, & Conners, 1997). Internal consistency of the MASC-39 for the present study was excellent ($\alpha = 0.906$). The MASC-10 has excellent internal consistency ($\alpha = 0.89$), strong test-retest reliability ($r = 0.86$), and good convergent and discriminant validity (March et al., 1997; March & Sullivan, 1999). Internal consistency of the MASC-10 for the current study was acceptable at T1 ($\alpha = 0.80$), T2 ($\alpha = 0.78$), and T3 ($\alpha = 0.79$).

The Center for Epidemiological Studies-Depression Scale for Children (CES-DC). The CES-DC consists of 20-items that examine depressed mood, worthlessness, helplessness, psychomotor retardation, eating and sleeping problems. Items are rated on a scale from 0 (not at all) to 3 (a lot) to indicate how frequently each statement was experienced 'in the past week'. Total scores range from 0 to 60 with higher scores indicating more severe depressive symptoms. The CES-DC has excellent internal consistency ($\alpha = 0.89$) and good convergent validity (Faulstich, Carey, Ruggiero, Enyart, & Gresham, 1986). Internal consistency for the present

study was excellent at T0 ($\alpha = 0.921$), T1 ($\alpha = 0.90$), and good at T2 ($\alpha = 0.89$) and T3 ($\alpha = 0.90$).

Chronic Pain Acceptance Questionnaire-Adolescents (CPAQ-A). The CPAQ-A (McCracken, Gauntlett-Gilbert, & Eccleston, 2010) is a 20-item scale that measures an adolescent's acceptance of chronic pain. Items are rated on a 5-point scale ranging from 0 (never true) to 4 (always true) with total scores ranging from 0 to 80, higher scores indicating greater acceptance of pain. The CPAQ-A has two subscales: activity engagement and pain willingness. The internal consistency for the activity engagement subscale has been shown to be good ($\alpha = 0.86$) and also adequate for pain willingness ($\alpha = 0.75$) (McCracken et al., 2010). Internal consistency for the present study was good at T0 ($\alpha = 0.88$), T2 ($\alpha = 0.87$), and T3 ($\alpha = 0.88$).

Predictive Validity Measures. *PROMIS-Pediatric Pain Interference Scale (PPIS).* The 8-item PROMIS-Pediatric Pain Interference Scale (Varni et al., 2010) assesses how the child's pain has interfered with certain aspects of their life over the past 7-days (e.g., sleep, attention, schoolwork, physical activities, emotion). Each item is rated on a 5-point scale ranging from "never" to "almost always". Scores range from 0 to 32 where higher scores indicate greater pain-related functional impairment. The PPIS consistently achieves a Cronbach's α of 0.85 (Varni et al., 2010). Internal consistency for the present study was excellent at T0 ($\alpha = 0.93$) and T3 ($\alpha = 0.92$), and good at T2 ($\alpha = 0.90$).

The Functional Disability Inventory (FDI). The FDI (Walker & Greene, 1991) is a 15-item scale that measures the extent to which children experience difficulties in completing daily tasks and activities (e.g., "Walking to the bathroom", "Eating regular meals", and "Being at school all day"). Each item is rated on a 5-point Likert Scale, which ranges from 0 (no trouble) to 4 (impossible). The total score ranges from 0 to 60 with higher scores indicative of increasing difficulty engaging in the activities. FDI has excellent internal consistency ($\alpha = 0.90$) and has good concurrent validity (Walker & Greene, 1991). The internal consistency of the FDI for the present study was excellent at T0 ($\alpha = 0.92$) and T3 ($\alpha = 0.91$), and good at T2 ($\alpha = 0.86$).

Procedure

The study was reviewed and approved by the Research Ethics Boards at The Hospital for Sick Children (SickKids) (REB file # 1000019644) and the Human Participants Review Committee at York University (Certificate # 2010 – 276). This prospective, longitudinal study involved four assessment time points over the course of a year: pre-operative, in-hospital, and 6-

and 12-months post-operative.

Pre-operative assessment (T0). The baseline assessment included administration of child and parent questionnaires. The research assistant was available for consultation if needed. The child completed questionnaires asking about previous and current pain experiences, as well as relevant psychological and emotional functioning (TSK, PCS-C, MASC-10, CRIES, CES-DC, CASI, CPAQ, FDI, PPIS). The order of questionnaire administration was randomized between subjects to minimize fatigue and order effects. The child's pre-operative medication use (analgesics and others) was obtained from the parents and confirmed by the patient's hospital medical record.

Intraoperative Anesthetic Management. Each patient received a general anesthetic. The following intraoperative factors were extracted from the surgical and anesthetic records: duration of surgery, analgesic/anesthetic regime including use of epidural/regional anesthetic techniques, systemic opioids, non-opioid adjuvants.

In-hospital Post-operative Assessment (T1). Physical movement was measured continuously while in hospital using a non-invasive Actical physical activity monitor beginning when participants arrived at the PACU after surgery. Pain intensity scores (NRS-I) and pain unpleasantness scores (NRS-U) were obtained daily by a research assistant. Postoperative analgesic use (e.g., opioid consumption, adjunct analgesics) was recorded from the child's medical record. In addition, 48-72 hours after surgery children completed self-report measures (TSK, PCS-C, MASC-10, CRIES, CES-DC, CASI, CPAQ). The research assistant was available for consultation if needed. The FDI and the CPAQ were not provided to the children to complete at T1 because they are not validated for in-hospital measurement.

Six (T2) and 12 (T3) Month Post-Operative Follow-ups. Six and twelve months after surgery, the research assistant followed up with participants by telephone to complete a set of measures to determine pain, psychological and emotional adaptation, current pain medications, incidence, intensity, quality of chronic postsurgical pain and the extent to which it interferes with daily activity (TSK, PCS-C, MASC-10, CRIES, CES-DC, CASI, CPAQ, FDI, PPIS).

Data Analysis

Factor Analysis

All factor analytic procedures were conducted in R Version 3.4.1 (R Core Team, 2014) using the packages "car" (Fox, Weisberg, & Price, 2020), "GPArotation" (Bernaards & Jennrich,

2014), and "psych" (Revelle, 2019). Polychoric correlation matrix for the 17 items were analyzed and observed for the strength and direction of the relationships between variables. An EFA was chosen over a confirmatory factor analysis because the TSK has not yet been evaluated among youth. Using ordinary least squares (OLS) to find the minimum residual (minres) solution, the number of factors within the TSK at T0 was evaluated. Specifically, the number of factors was determined by an examination of the (1) eigenvalue scree plot, (2) parallel analysis, (3) standardized root mean square residual (SRMR), (4) residual correlations, and (5) factor loadings, communality, and uniqueness and residual correlations, as well as the factor loadings with oblimin rotation.

Reliability

Reliability analyses were conducted in R using the package "psych". Cronbach's alpha (α) and McDonald's omega (Ω) were used to examine the internal consistency of the TSK prior to surgery. Cronbach's α was interpreted as: Excellent: $0.9 \leq α$; Good: $0.8 \leq α < .9$; $0.7 \leq α < .8$ Acceptable; $0.6 \leq α < .7$ Questionable; $0.5 \leq α < .6$ Poor; $α < .5$ Unacceptable (DeVellis, 2012).

Construct Validity

Construct validity was examined through convergent validity and discriminant validity. Convergent validity was determined by examining correlations between the TSK and psychological constructs linked to the fear-avoidance model (Asmundson et al., 2012); specifically the NRS pain intensity, NRS pain unpleasantness, CPASS avoidance subscale, pain catastrophizing, and physical movement levels. High correlations (i.e., $r > .70$) indicated convergent validity (Jackson, 2009). Discriminant validity was determined by examining correlations between the TSK and theoretically dissimilar (or less similar) psychological constructs, specifically general anxiety, anxiety sensitivity, symptoms of posttraumatic stress disorder, general depression, and chronic pain acceptance. A correlation below .70 between the TSK and these measures would indicate adequate discriminant validity.

Convergent and discriminant validity were also examined by comparing the magnitude of the difference in correlation coefficients between: (1) the TSK and CPASS Avoidance Subscale versus the TSK and MASC-10; (2) the TSK and CPASS Avoidance Subscale versus TSK and CES-DC; (3) the TSK and PCS-C versus the TSK and MASC-10; and (4) the TSK and PCS-C versus the TSK and CES-DC. Significantly larger correlations (i.e., minimum difference of .2) between the TSK and pain-related fear avoidance (CPASS Avoidance Subscale and PCS-C) as

compared to the TSK and non-pain-related fear avoidance (MASC-10 and CES-DC) would suggest sufficient convergent and discriminant validity. Through the use of a web-based calculator, Fisher's *r*-to-*z* transformation was performed to convert each correlation coefficient to a *z*-score (Lee & Preacher, 2013) and Steiger's equations 3 and 10 (Steiger, 1980) to compute the covariance of the estimates.

Predictive Validity

Predictive validity, or the extent to which the TSK correlated with theoretically related constructs measured at a later time, was evaluated by examining correlations between the pre-operative scores on the TSK and disability (FDI, PPIS) 12-months after surgery.

Results

Recruitment and Demographic Information

Recruitment took place between February 2011 and August 2015. Recruitment details have been reported in previous articles (Noel et al., 2019; Rosenbloom et al., 2019). Of the 349 children and parents approached for consent, 270 provided assent and informed written consent to participate, respectively. Three children withdrew consent before participating in any part of the study, one patient's surgical procedure was changed and no longer met study criteria, and 27 children were missed (i.e., the research assistant was unable to locate or reach them) for their T0 assessment. One patient was diagnosed with cancer after consent and was withdrawn from the study. A total of 264 participants completed at least one component of the in-hospital (T1) assessment (e.g., questionnaire, actigraphy, daily pain measures). Twenty-seven participants were admitted directly to the intensive care unit (ICU) from the operating room (i.e. they did not go to the PACU) and therefore the research assistant was unable to obtain daily pain measures or place the Actical physical movement monitor on the child until they were admitted to a regular surgical unit. The 6- and 12-month retention rates of participants in this study were 81.13% and 85.28%, respectively.

Significant differences at baseline were not found between participants who completed the study at 12-months and those who did not on any of the measures, except in the PPIS. Participants who completed the study had significantly lower baseline PPIS scores ($M = 14.42$, $SD = 9.59$) than those who did not complete the study ($M = 23.00$, $SD = 5.40$), $t(162) = 2.162$, $p = .031$.

The final sample consisted of 264 children [155 female (58.49%), *M* age = 14.07 years (*SD* = 2.51), range 8-18 years] and their parents or guardians [188 female (83.92%), *M* age = 45.10 years (*SD* = 5.88), range 29-70 years, 41 parents did not answer the demographic questionnaire]. The majority of children identified as Caucasian (65.74%). The majority of children underwent surgery for scoliosis (*n* = 133, 50.2%) and 35.5% (*n* = 94) underwent an osteotomy. Fourteen children (5.3%) had a Ravitch procedure, four (1.5%) had a Nuss procedure (1.5%), four (1.5%) had a thoracotomy, and fifteen (5.7%) had another type of surgery. The mean duration of surgery was 4.59 hours (*SD* = 2.07 hours, range = 0.70 − 10.70 hours) and children stayed in hospital an average of 4.94 days (*SD* = 2.91, range 1-36 days). Participants who were transferred to the ICU had significantly longer surgical times (*p* < .001) and hospital stays (*p* = .001).

Exploratory Factor Analysis

An examination of the polychoric correlation matrix among TSK items (Table 2) showed that most items were moderately positively correlated with one another. As expected, many of the reverse scored items were negatively correlated with the positively scored items. However, this was not consistent across the items (e.g., items 4 and 11 have a correlation of .11). Additionally, the reverse scored items were weakly correlated with most of the positively scored items.

The initial EFA was conducted on all 17 TSK items (TSK-17). Parallel analysis with 100 iterations suggested a five-factor model. However, Cattel's scree plot suggested that a one- or two-factor model would work best because there was a noticeable decrease between one and two factors and another decrease between two and three factors, leveling off after that point.

Given these results, the SRMRs and residual correlations were examined for one-, two-, and five-factor models. The SRMR for the one-factor. Model was .091. The SRMR for the two-factor model was .063 indicating a good fit. The SRMR for the five-factor model was .028 indicating a better fit than the two-factor model. However, this decrease in the SRMR is reflective of the fact that more factors were added resulting in smaller residual correlations. The residual correlation matrix for the one-factor model reveals mostly low residual correlations with the exception of the reverse scored items where the residual correlations range between .004 to .121, with the majority closer to .121. The residual correlation matrices for the two- and five

factor models reveal no high residual correlations between any of the items. As a result, the factor loadings, communality, and uniqueness were evaluated.

The factor loadings for the two-factor model (Table 3) were evaluated for parsimony in both columns (i.e. within factors) and rows (i.e. between factors). The positively scored items (1, 2, 3, 5, 6, 7, 9, 10, 11, 13, 14, 15, and 17) load well onto Factor 1, whereas the reverse scored items (4, 8, 12, and 16) load well onto Factor 2. There was significant cross-loading for item 6. In this model each variable can also be evaluated for the amount of total variance that is due to each factor. For example, 40% of the variance in item 1 is explained by the common factors. Only items 1, 13, and 15 have communalities greater than .40, indicating a good fit with the factors. However, the rest of the items had communalities less than .40 (i.e., $h^2 = .08 - .39$), indicating poor correlations with the factors. Inter-factor correlations were also examined and it was found that the factors in the two-factor model were very weakly correlated ($r = .01$).

In the five-factor model (Data in Appendix), the items do not load well onto any of the five factors apart from Items 1 (loaded onto factor 1) and 4 (loaded onto factor 2). The five-factor model varied in the strength of the inter-factor relationships with one another with correlations ranging from .00 to .43.

These results for the TSK-17 are consistent with what has been found in the adult literature (Goubert et al., 2004; Heuts et al., 2004; Swinkels-Meewisse et al., 2003) in that the two-factor model fits the data best. In keeping with the adult literature, we conducted an EFA on only the positively scored items, excluding the reverse scored items from analysis. The TSK-17 EFA yielded results indicative of a poor fit (e.g., very weak correlation between factors in the two-factor model, factors separated by positively and negatively scored items). Moreover, TSK-17 convergent, discriminant, and predictive validity were worse than the TSK-13. Due to space restrictions, we report the TSK-17 EFA and psychometric results in the Appendix and the results from the TSK-13 in this manuscript.

The second EFA was conducted on the 13 positively scored TSK items (henceforth referred to as the TSK-13). Parallel analysis with 100 iterations suggested a five-factor model. However, Cattel's scree plot suggested that a one-, two- or three-factor model would fit the data because there was a noticeable decrease between one, two and three factors after which the graph levelled off. The ratio between eigenvalues were further evaluated. Given that the scree plot and

parallel analysis suggest a one-, two- or three-factor model, the SRMRs, and residual correlations were examined for the two- and three-factor models.

The SRMR for the one-factor model was .07 and the SRMR for the two-factor model was .06 indicating a good fit and providing support for these models. However, the SRMR for the three-factor model was .04 indicating a slightly better fit than the one- and two-factor models. A close look at the residual correlations in the residual correlation matrix reveals that there are no high residual correlations between any of the items in the one-, two- or three-factor models. As a result, the factor loadings, communality, and uniqueness were evaluated for the one-, two- and three-factor model.

The factor loadings for the one-factor model were evaluated. Items loaded well onto the one factor; however, communality was low for each item (i.e., $h^2 = .12 - .44$).

The factor loadings for the two-factor model were evaluated for parsimony in both columns (i.e., within factors) and rows (i.e., between factors). Items 1, 2, 5, 9, 10, 14, and 15 load well onto Factor 1 (fear of injury) and items 3, 6, 7, 11, 13, and 17 load well onto Factor 2 (activity avoidance). There was significant cross-loading among items 5, 10, 13, and 17. In this model each variable can also be evaluated for the amount of total variance that is explained by each factor. For example, 54% of the variance in item 1 was accounted for by the common factors. Only items 1 and 15 have communalities greater than .40, indicating a good fit with the factors. The rest of the items had communalities less than .40 (i.e., $h^2 = .13 - .36$), indicating poor correlations with the factors. The inter-factor correlation for the two-factor model was moderate ($r = .54$).

The factor loadings for the three-factor model (Table 4) were evaluated for parsimony in both columns (i.e. within factors) and rows (i.e. between factors). Items 1, 2, 9, 14, and 15 load well onto Factor 1 (fear of injury); items 3, 5, 6, 7, and 11 load well onto Factor 2 (bodily vulnerability; and items 10, 13, 14, and 17 load well onto Factor 3 (activity avoidance). There was significant cross-loading among items 2, 5, and 14. Fifty-seven percent of the variance in item 1 was accounted for by the common factors. Only items 1, 10, and 15 have communalities greater than .40, indicating a good fit with the factors. However, the rest of the items had communalities less than .40 (i.e., $h^2 = .21 - .39$), indicating poor correlations with the factors, but better overall than the two-factor model. Inter-factor correlations in the three-factor model were moderate ($r_{f1,f2} = 0.49$, $r_{f1, f3} = .45$, $r_{f2, f3} = .38$).

In conclusion, using the TSK-13 (positively scored items only), the one-, two-, and three-factor models yield similar results with respect to overall model statistics, however, the three-factor model is more parsimonious and yields more meaningful factors. Furthermore, 10 of the 13 items load well onto either Factor 1 (fear of injury), Factor 2 (bodily vulnerability), or Factor 3 (activity avoidance), whereas three items have some cross-loading.

Reliability

Using Cronbach's α, the TSK-13 yielded adequate overall internal consistency ($\alpha = .81$). Cronbach's α per factor was adequate, specifically the fear of injury factor had an Ω of .77, the bodily vulnerability factor had an Ω of .62, and the activity avoidance factor had an Ω of .59.

Using McDonald's Ω, the TSK-13 showed good overall internal consistency ($\Omega = .82$). The internal consistency per factor was adequate, specifically the fear of injury factor had an Ω of .77, the bodily vulnerability factor had an Ω of .62, and the activity avoidance factor had an Ω of .59.

Construct Validity

The means and standard deviations for each measure used to asses construct validity are shown in Table 5.

Convergent Validity. The TSK-13 was significantly and moderately correlated with NRS pain intensity ($r (228) = .25, p <. 001$), NRS pain unpleasantness ($r (162) = .41, p <. 001$), the avoidance subscale of the CPASS ($r (227) = .53, p <. 001$) and the PCS ($r (226) = .56, p < .001$), indicating weak convergent validity with those measures given the requirement that $r > .7$. The TSK-13 was not significantly correlated with actual physical movement on post-operative day 2 ($r (179) = -.10, p = .18$) as measured by actigraphy, indicating poor convergent and predictive validity with physical movement levels. Correlations between the TSK and other measures as well as partial correlation coefficients, controlling for age and sex, are shown in Table 6.

Discriminant Validity. The TSK-13 was significantly and moderately correlated with the MASC-10 ($r (224) = .35, p <.001$), CRIES ($r (119) = .41, p <.001$), CES-DC ($r (225) = .41, p <.001$), CASI ($r (227) = .40, p <.001$), and CPAQ ($r (180)= -.52, p < .001$), indicating adequate discriminant validity given the requirement that $r <.7$. Partial correlation coefficients, controlling for age and sex, are summarized in Table 6.

Comparison of Convergent and Discriminant Validity. Convergent and discriminant validity were examined by comparing the magnitude of the correlation coefficients. The TSK-13

was significantly more highly correlated with the Avoidance Subscale of the CPASS than with the MASC-10, $Z = 2.89$, $p = .004$, however, this difference was not greater than the .20 cuff off. Further, the TSK -13 was not significantly more correlated with the Avoidance Subscale of the CPASS than with the CES-DC, $Z = 1.89$, $p = .06$, and this difference was under the .20 cut off. The TSK-13 was significantly more correlated with the PCS-C than with the MASC-10, $Z = 3.314$, $p < .001$, and this difference was greater than .20. However, while the TSK-13 was significantly more correlated with the PCS-C than the CES-DC, $Z = 2.66$, $p = .008$, the difference was less than .20.

Predictive Validity

The means and standard deviations for the FDI and PPIS are shown in Table 5. Baseline TSK-13 scores, measured prior to surgery, were weakly correlated with functional disability 12-months after surgery (r (198) = .17, $p < .05$) and moderate with pain related disability (r (70) = .36, $p < .001$), indicating adequate predictive validity for the latter measure.

Discussion

The aim of this study was to evaluate the psychometric properties of the Tampa Scale for Kinesiophobia (TSK) among children and adolescents undergoing major surgery. EFA showed a poor fit with a one-factor model (17-items) and a slightly better fit with a two-factor model that distinguished 13 positively scored items from 4 reverse scored items. This latter finding is consistent with the results of three studies in adults that found better model fit without the reverse scored items (Clark et al., 1996; Goubert et al., 2004; Swinkels-Meewisse et al., 2003). In the present study, the model fit indices were not ideal for a two-factor model with all 17 items included. With the reverse scored items removed, the EFA conducted on the 13 positively scored items revealed a stronger model fit with three factors defined as: Fear of injury (fear of moving due to risk of re-injury), bodily vulnerability (belief that the body is vulnerable to pain and injury), and activity avoidance (belief that activities should be avoided).

The 13-item TSK yielded adequate internal consistency and weak convergent validity with measures of pain intensity, pain unpleasantness, and pain avoidance while in hospital after surgery. These findings are in-line with the cognitive-behavioural fear-avoidance model of chronic pain (Vlaeyen et al., 1995; Vlaeyen & Linton, 2000), where a higher level of fear of movement/ re-injury is associated with greater pain experiences. The TSK-13 was shown to be distinct enough from a measure of general anxiety to demonstrate some discriminant validity.

However, the TSK-13 was not distinct enough from a measure of depression or the avoidance subscale of a measure of pain-related anxiety.

The pre-surgical TSK-13 measure was not correlated with actual acute postoperative physical movement on Day 2 after surgery as measured by continuous actigraphy. It is possible that fear of movement/re-injury has a negative relationship with long-term movement, not evident in the short-term. The lack of a significant correlation may be due to a floor effect in post-operative physical movement which was low across all patients. Low levels of physical movement immediately after surgery may be explained by children having strict physiotherapy protocols to follow that restrict movement in the days after surgery (e.g., sitting first, then standing, then walking). It is also possible that low levels of physical movement may be due to the residual effects of anesthetics and surgery (e.g., fatigue, nausea, indwelling catheters, sedation). Additionally, it is likely that some movements (e.g., eating, hand gestures while talking) may have inflated total activity counts while at the same time not being a source of fear of movement/injury thereby diluting the relationship between activity and the TSK-13 scores. The low correlation between fear of movement/re-injury and physical movement is also consistent with lack of a significant association found between the TSK and treadmill tests of aerobic conditioning and maximal strength tasks (French et al., 2007).

However, another possibility is that the baseline, pre-surgical measure of the TSK-13 was a poor reflection of the participants' fear of movement/re-injury on Day 2 when physical movement was assessed given the fact that participants underwent major surgery between the two assessment times. Administration of the TSK-13 on Day 2 after surgery may have produced a stronger correlation between fear of movement/re-injury and Day 2 physical movement. Exposure to intense acute postoperative pain may have altered participants' baseline level of fear of movement/re-injury rendering the pre-surgical results invalid. A similar argument has been made regarding the validity of pre-operative child pain catastrophizing as a predictor of post-operative pain and health-related quality of life two weeks after surgery (Katz, 2015). Taken together, these 'negative' findings question the usefulness of pre-operative measures in predicting acute post-operative pain and function since the intense postsurgical pain experienced in the days after surgery may, for certain vulnerable children, establish a new baseline. Future studies should examine the TSK-13 and physical movement concurrently in the acute post-operative period.

In terms of long-term outcomes, the TSK-13 showed adequate predictive validity for pain-related disability at 12 months after surgery but its relationship to general functional disability 12 months after surgery was less compelling. These results are similar to Swinkels-Meewisse et al.'s (2003) cross-sectional study in adults, showing a positive association between the TSK and disability, as measured by the Roland Disability Questionnaire, among acute low back pain patients. Together these results support the relationship between fear of movement/re-injury and subsequent disability, such that greater fear of movement/re-injury is associated with greater disability.

There are several limitations to the present study. First, participants who completed the study had significantly lower baseline pain interference scores than those who did not complete the study. Consequently, the sample is likely not representative of the level of pain interference in the population of youth who undergo major surgery. It is not known whether, or in what way, the lack of representativeness of the sample influenced the psychometric properties of the TSK. Second, the TSK-17 had poorer model fit statistics than the TSK-13 because of the four negatively scored items in the former scale. It is not known whether these reverse scored items were a poor fit because of the wording (syntax) or content (meaning). Future studies might evaluate whether changing the wording of the reverse scored items to make them positively scored would improve the psychometric properties of the 17-item TSK. For example, changing item 8 from "Just because something aggravates my pain does not mean it is dangerous" to "Things that aggravate my pain are dangerous", or other similar wording, would help to determine whether the problem we, and others, encountered with the reverse scored items is due to item wording/syntax or item content/meaning.

Overall, the results of the present study provide initial support for the use of the TSK-13 as an adequately valid and reliable measure of the fear of movement/re-injury construct in children and adolescents undergoing surgery. The TSK-13 revealed a 3-factor structure that is reliable and demonstrates adequate convergent, discriminant, and predictive validity within the context of pediatric surgery. Further validation of this measure is warranted in other populations, such as youth with referred to pediatric chronic pain clinics.

Acknowledgements

References

Al-Obaidi, S. M., Nelson, R. M., Al-Awadhi, S., & Al-Shuwaie, N. (2000). The role of anticipation and fear of pain in the persistence of avoidance behavior in patients with chronic low back pain. *Spine, 25*(9), 1126-1131. doi:10.1097/00007632-200005010-00014

Asmundson, G. J., Noel, M., Petter, M., & Parkerson, H. A. (2012). Pediatric fear-avoidance model of chronic pain: foundation, application and future directions. *Pain Research and Management, 17*(6), 397-405. doi:10.1155/2012/908061

Asmundson, G. J., Norton, G. R., & Allendings, M. D. (1997). Fear and avoidance in dysfunctional chronic back pain patients. *Pain, 69*, 231-236. doi:10.1016/s0304-3959(96)03288-5

Bernaards, C., & Jennrich, R. (2014). GPArotation: Gradient Projection Algorithm Rotation for Factor Analysis. Version 2014.11-1. Retrieved from https://cran.r-project.org/web/packages/GPArotation/GPArotation.pdf

Chidambaran, V., Ding, L., Moore, D. L., Spruance, K., Cudilo, E. M., Pilipenko, V., . . . Sadhasivam, S. (2017). Predicting the pain continuum after adolescent idiopathic scoliosis surgery: A prospective cohort study. *Eur J Pain, 21*(7), 1252-1265. doi:10.1002/ejp.1025

Clark, M. E., Kori, S. H., & Brockel, J. (1996). Kinesiophobia and chronic pain: psychometric characteristics and factor analysis of the Tampa Scale. *American Pain Society Abstract, 77*.

Crombez, G., Bijttebier, P., Eccleston, C., Mascagni, T., Mertens, G., Goubert, L., & Verstraeten, K. (2003). The child version of the pain catastrophizing scale (PCS-C): a preliminary validation. *Pain, 104*(3), 639-646. doi:10.1016/s0304-3959(03)00121-0

DeVellis, R. F. (2012). *Scale development: Theory and applications*. Los Angeles, CA.: Sage.

Faulstich, M. E., Carey, M. P., Ruggiero, L., Enyart, P., & Gresham, F. (1986). Assessment of depression in childhood and adolescence: An evaluation of the center for epidemiological studies depression scale for children (CES-DC). *Am J Psychiatry, 143*(8), 1024-1027.

Fox, J., Weisberg, S., & Price, B. (2020). car: Companion to Applied Regression. Version 3.0-7. Retrieved from https://cran.r-project.org/web/packages/car/index.html

French, D. J., France, C. R., Vigneau, F., French, J. A., & Evans, R. T. (2007). Fear of movement/(re)injury in chronic pain: a psychometric assessment of the original English version of the Tampa scale for kinesiophobia (TSK). *Pain, 127*(1-2), 42-51. doi:10.1016/j.pain.2006.07.016

Goubert, L., Crombez, G., Van Damme, S., Vlaeyen, J. W., Bijttebier, P., & Roelofs, J. (2004). Confirmatory factor analysis of the Tampa Scale for Kinesiophobia: invariant two-factor model across low back pain patients and fibromyalgia patients. *Clinical Journal of Pain, 20*(2), 103-110. doi:10.1097/00002508-200403000-00007

Heuts, P. H., Vlaeyen, J. W., Roelofs, J., de Bie, R. A., Aretz, K., van Weel, C., & van Schayck, O. C. (2004). Pain-related fear and daily functioning in patients with osteoarthritis. *Pain, 110*(1-2), 228-235. doi:10.1016/j.pain.2004.03.035

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: a measure of subjective stress. *Psychosomatic medicine, 41*(3), 209-218. doi:10.1097/00006842-197905000-00004

Jackson, S. L. (2009). Correlational methods and statistics. In S. L. Jackson (Ed.), *Research methods and statistics: a critical thinking approach.* (pp. 140–162). Belmont, CA: Wadsworth.

Katz, J. (2015). Establishment of a new pain catastrophizing baseline after pediatric major surgery? *Journal of Pain, 16*(4), 388. doi:10.1016/j.jpain.2015.01.003

Keefe, F. J., Rumble, M. E., Scipio, C. D., Giordana, L. A., & Perri, L. M. (2004). Psychological aspects of persistent pain: current state of the science. *Pain, 5*, 195-211. Retrieved from 10.1016/j.jpain.2004.02.576

Kori, S. H., Miller, R. P., & Todd, D. D. (1990). Kinesiophobia: a new view of chronic pain behavior. *Pain management, 3*(1), 35-43.

Kudchadkar, S. R., Aljohani, O., Johns, J., Leroux, A., Alsafi, E., Jastaniah, E., . . . Crainiceanu, C. (2019). Day-Night Activity in Hospitalized Children after Major Surgery: An Analysis of 2271 Hospital Days. *The Journal of pediatrics, 209*, 190-197 e191. doi:10.1016/j.jpeds.2019.01.054

Landman, Z., Oswald, T., Sanders, J., Diab, M., & Spinal Deformity Study, G. (2011). Prevalence and predictors of pain in surgical treatment of adolescent idiopathic scoliosis. *Spine (Phila Pa 1976), 36*(10), 825-829. doi:10.1097/BRS.0b013e3181de8c2b

Lee, I. A., & Preacher, K. J. (2013). Calculation for the test of the difference between two dependent correlations with one variable in common. Retrieved from http://quantpsy.org.

March, J. S., Parker, J. D., Sullivan, K., Stallings, P., & Conners, C. K. (1997). The Multidimensional Anxiety Scale for Children (MASC): factor structure, reliability, and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(4), 554-565. doi:10.1097/00004583-199704000-00019

March, J. S., & Sullivan, K. (1999). Test-retest reliability of the Multidimensional Anxiety Scale for Children. *J Anxiety Disord, 13*(4), 349-358. doi:10.1016/s0887-6185(99)00009-2

McCracken, L. M., Gauntlett-Gilbert, J., & Eccleston, C. (2010). Acceptance of pain in adolescents with chronic pain: validation of an adapted assessment instrument and preliminary correlation analyses. *European Journal of Pain, 14*(3), 316-320. doi:10.1016/j.ejpain.2009.05.002

Noel, M., Rosenbloom, B., Pavlova, M., Campbell, F., Isaac, L., Pagé, M. G., . . . Katz, J. (2019). Remembering the pain of surgery one year later: a longitudinal examination of anxiety in children's pain memory development. *Pain*. doi:10.1097/j.pain.0000000000001582

Pagé, M. G., Campbell, F., Isaac, L., Stinson, J., Martin-Pichora, A. L., & Katz, J. (2011). Reliability and validity of the Child Pain Anxiety Symptoms Scale (CPASS) in a clinical sample of children and adolescents with acute postsurgical pain. *Pain, 152*(9), 1958-1965. doi:10.1016/j.pain.2011.02.053

Pagé, M. G., Fuss, S., Martin, A. L., Escobar, E. M., & Katz, J. (2010). Development and preliminary validation of the Child Pain Anxiety Symptoms Scale in a community sample. *Journal of Pediatric Psychology 35*(10), 1071-1082. doi:10.1093/jpepsy/jsq034

Pagé, M. G., Katz, J., Stinson, J., Isaac, L., Martin-Pichora, A. L., & Campbell, F. (2012). Validation of the numerical rating scale for pain intensity and unpleasantness in pediatric acute postoperative pain: sensitivity to change over time. *Journal of Pain, 13*(4), 359-369. doi:10.1016/j.jpain.2011.12.010

Perrin, S., Meiser-Stedman, R., & Smith, P. (2005). The children's revised impact of events scale (CRIES): validity as a screening instrument for PTSD. *Behavioral and Cognitive Psychotherapy, 33*, 487-498. doi:10.1017/S1352465805002419

Puyau, M. R., Adolph, A. L., Vohra, F. A., Zakeri, I., & Butte, N. F. (2004). Prediction of activity energy expenditure using accelerometers in children. *Medicine and Science in Sports and Exercise, 36*(9), 1625-1631. doi:10.1249/01.MSS.0000139898.30804.60

R Core Team. (2014). R: A language and environment for statistical computing.

Revelle, R. (2019). psych: Procedures for Psychological, Psychometric, and Personality Research. Version 1.9.12.31. Retrieved from https://cran.r-project.org/web/packages/psych/psych.pdf

Roelofs, J., Goubert, L., Peters, M. L., Vlaeyen, J. W., & Crombez, G. (2004). The Tampa Scale for Kinesiophobia: Further examination of psychometric properties in patients with

chronic low back pain and fibromyalgia. *European Journal of Pain, 8*(5), 495-502. doi:10.1016/j.ejpain.2003.11.016

Rosenbloom, B. N., Pagé, M. G., Isaac, L., Campbell, F., Stinson, J., Wright, J. G., & Katz, J. (2019). Pediatric chronic postsurgical pain and functional disability: A prospective study of risk factors up to one-year after major surgery. *J Pain Res, 12*, 3079-3098. doi: https://doi.org/10.2147/JPR.S210594

Sieberg, C. B., Simons, L. E., Edelstein, M. R., DeAngelis, M. R., Pielech, M., Sethna, N., & Hresko, M. T. (2013). Pain prevalence and trajectories following pediatric spinal fusion surgery. *Journal of Pain, 14*(12), 1694-1702. doi:10.1016/j.jpain.2013.09.005

Silverman, W. K., Fleisig, W., Rabian, B., & Peterson, R. A. (1991). Childhood anxiety sensitivity index. *Journal of Clinical Child and Adolescent Psychology, 20*(2), 162-168. doi:10.1207/s15374424jccp2002_7

Simons, L. E., Sieberg, C. B., Carpino, E., Logan, D., & Berde, C. (2011). The Fear of Pain Questionnaire (FOPQ): assessment of pain-related fear among children and adolescents with chronic pain. *Journal of Pain, 12*(6), 677-686. doi:10.1016/j.jpain.2010.12.008

Smith, P., Perrin, S., Dyregrov, A., & Yule, W. (2003). Principal components analysis of the impact of event scale with children in war. *Personality and Individual Differences, 34*(2), 315-322. doi:10.1016/S0191-8869(02)00047-8

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological bulletin, 87*(2), 245-251.

Sullivan, M. J., Bishop, S. R., & Pivik, J. (1995). The Pain Catastrophizing Scale: Development and Validation. *Psychological Assessment, 7*(4), 524-532. doi:10.1037//1040-3590.7.4.524

Swinkels-Meewisse, E. J., Swinkels, R. A., Verbeek, A. L., Vlaeyen, J. W., & Oostendorp, R. A. (2003). Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. *Manual Therapy, 8*(1), 29-36. doi:S1356689X02904844 [pii]

Varni, J. W., Stucky, B. D., Thissen, D., Dewitt, E. M., Irwin, D. E., Lai, J. S., . . . Dewalt, D. A. (2010). PROMIS Pediatric Pain Interference Scale: an item response theory analysis of the pediatric pain item bank. *Journal of Pain, 11*(11), 1109-1119. doi:10.1016/j.jpain.2010.02.005

Verbunt, J. A., Seelen, H. A., Vlaeyen, J. W., van der Heijden, G. J., & Knottnerus, J. A. (2003). Fear of injury and physical deconditioning in patients with chronic low back pain. *Archives of Physical Medicine and Rehabilitation, 84*, 1227-1232. doi:10.1016/s0003-9993(03)00132-1

Vlaeyen, J. W., Kole-Snijders, A. M., Boeren, R. G., & van Eek, H. (1995). Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain, 62*(3), 363-372. doi:10.1016/0304-3959(94)00279-n

Vlaeyen, J. W., & Linton, S. J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: A state of the art. *Pain, 85*, 317-332. doi:10.1016/s0304-3959(99)00242-0

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. doi:10.1016/j.pain.2009.03.002

Walker, L. S., & Greene, J. W. (1991). The functional disability inventory: measuring a neglected dimension of child health status. *Journal of pediatric psychology, 16*(1), 39-58. doi:10.1093/jpepsy/16.1.39

Wicksell, R. K., Melin, L., Lekander, M., & Olsson, G. L. (2009). Evaluating the effectiveness of exposure and acceptance strategies to improve functioning and quality of life in longstanding pediatric pain--a randomized controlled trial. *Pain, 141*(3), 248-257. doi:10.1016/j.pain.2008.11.006

Ye, D. L., Plante, I., Roy, M., Ouellet, J. A., & Ferland, C. E. (2020). The Tampa Scale of Kinesiophobia: Structural Validation among Adolescents with Idiopathic Scoliosis Undergoing Spinal Fusion Surgery. *Phys Occup Ther Pediatr*, 1-11. doi:10.1080/01942638.2020.1720054

Table 1. Items from the Tampa Scale for Kinesiophobia. Items from Kori et al. (1990).

1. I'm afraid that I might injury myself if I exercise
2. If I were to try to overcome it, my pain would increase
3. My body is telling me I have something dangerously wrong
4. * My pain would probably be relieved if I were to exercise
5. People aren't taking my medical condition seriously enough
6. My accident has put my body at risk for the rest of my life
7. Pain always means I have injured my body
8. * Just because something aggravates my pain does not mean it is dangerous
9. I am afraid that I might injure myself accidentally
10. Simply being careful that I do not make any unnecessary movements is the safest thing I can do to prevent my pain from worsening
11. I wouldn't have this much pain if there weren't something potentially dangerous going on in my body
12. * Although my condition is painful, I would be better off if I were physically active
13. Pain lets me know when to stop exercising so that I don't injure myself
14. It's really not safe for a person with a condition like mine to be physically active
15. I can't do all the things normal people do because it's too easy for me to get injured
16. * Even though something is causing me a lot of pain, I don't think it's actually dangerous
17. No one should have to exercise when he/she is in pain

*negatively worded, reversed scored items

Table 2. Polychoric correlations between TSK items. For ease of viewing and to be consistent with data analyses, positively scored TSK items are listed first (i.e. items 1, 2, 3, 5, 6, 7, 9, 10, 11, 13, 14, 15, 17) and negatively scored TSK items are listed after (i.e. items 4, 8, 12, 16).

TSK Item #	1	2	3	5	6	7	9	10	11	13	14	15	17	4	8	12	16
1.	-																
2.	.55	-															
3.	.32	.44	-														
5.	.33	.19	.31	-													
6.	.26	.20	.39	.38	-												
7.	.10	.20	.26	.22	.33	-											
9.	.51	.34	.28	.35	.23	.13	-										
10.	.37	.34	.35	.12	.24	.18	.38	-									
11.	.26	.18	.42	.23	.31	.30	.28	.29	-								
13.	.29	.32	.27	.06	.18	.27	.39	.47	.25	-							
14.	.49	.37	.31	.28	.26	.35	.33	.35	.33	.21	-						
15.	.55	.43	.32	.30	.27	.21	.48	.46	.20	.22	.62	-					

17.	.10	.29	.23	.03	.06	.19	.20	.30	.18	.27	.28	.33	-				
4.	.05	.11	-.09	-.11	-.22	.00	-.11	-.06	-.14	-.02	.11	-.06	.16	-			
8.	.13	-.01	.01	.01	-.15	-.05	.01	-.24	-.15	-.10	-.02	.02	.03	.28	-		
12.	.05	.16	-.07	-.07	-.09	.08	-.05	-.08	-.10	-.08	.17	.03	.16	.51	.18	-	
16.	.02	.06	.10	.01	.09	-.01	-.02	-.08	.02	-.10	.02	-.07	-.01	.13	.34	.18	-

TSK; Tampa Scale for Kinesiophobia

Table 3. Pattern matrix including factor loadings, communality, and uniqueness for the two-factor model with quartimin rotation. For ease of viewing and to be consistent with data analyses, positively scored TSK items are listed first (i.e. items 1, 2, 3, 5, 6, 7, 9, 10, 11, 13, 14, 15, 17) and reverse scored TSK items are listed after (i.e. items 4, 8, 12, 16).

	Λ			
Item	f_1	f_2	Communality (h^2)	Uniqueness (u^2)
1	.62	.08	.40	.60
2	.58	.16	.37	.63
3	.53	.04	.28	.72
5	.35	-.08	.13	.87
6	.37	-.23	.19	.81
7	.35	-.03	.12	.88
9	.57	-.11	.34	.66
10	.56	-.16	.34	.66
11	.42	-.19	.21	.79
13	.47	-.16	.47	.75
14	.61	.12	.39	.61
15	.66	.02	.44	.56
17	.36	.14	.15	.85
4	-.01	.65	.42	.58
8	-.04	.41	.17	.83
12	.06	.60	.37	.64

| 16 | .03 | .27 | .08 | .93 |

TSK; Tampa Scale for Kinesiophobia

Table 4. Pattern matrix including factor loadings, communality, and uniqueness for the three-factor model with quartimin rotation using the TSK-13 (positively scored items only).

		Λ			
Item	f_1	f_2	f_3	Communality (h^2)	Uniqueness (u^2)
1	.80	-.05	-.04	.57	.43
2	.44	.03	.21	.34	.66
3	.11	.43	.16	.33	.67
5	.30	.35	-.24	.24	.76
6	.04	.56	-.09	.30	.70
7	-.11	.48	.14	.25	.75
9	.48	.04	.16	.35	.65
10	.20	.07	.49	.41	.59
11	-.04	.46	.17	.38	.72
13	.04	.07	.58	.39	.61
14	.46	.21	.04	.37	.63
15	.67	.02	.08	.52	.48
17	.06	.01	.42	.21	.79

TSK; Tampa Scale for Kinesiophobia

Table 5. Means and standard deviations for pre-surgical and in hospital factors used in the validation of the TSK-13.

Pre-Surgery Variables	Mean	Standard Deviation
TSK	37.74	6.94
TSK-13	27.59	6.52
NRS Pain at Rest	2.00	2.33
NRS Pain Unpleasantness	2.55	2.97
CPASS	32.88	18.55
PCS	19.64	11.93
Movement Post-Operative Day 2	56924.86	69965.88
CASI	29.65	6.74
CRIES	22.96	14.32
MASC-10	11.45	5.55
CED-DC	16.72	12.13
CPAQ	46.68	13.16
FDI	11.66	11.23
PPIS	14.73	9.59

TSK, Tampa Scale for Kinesophobia; NRS, Numeric Rating Scale; CPASS, Child Pain Anxiety Symptoms Scale; PCS, Child Pain Catastrophizing Scale; CASI, Childhood Anxiety Sensitivity Index; CRIES, Children's Revised Impact of Event Scale; MASC-10, Multidimensional Anxiety Scale for Children; CES-DC; Center for Epidemiological Studies-Depression Scale for Children; CPAQ, Chronic Pain Acceptance Questionnaire - Adolescents; PPIS, PROMIS-Pediatric Pain Interference Scale; FDI, Functional Disability Index

Table 6. Correlation coefficients (below the diagonal) and partial correlation coefficients (above the diagonal) between the TSK-13 (positively scored items only) measured at baseline before surgery (T0) and other variables measured at various times in relation to surgery.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. T0 TSK-13		.23**	.40**	.53**	.52**	-.11	.35**	.40**	.40**	.39**	-.52**	.18*	.33*
2. T0 NRS Pain Intensity	.25**		.82**	.11	.28**	.01	.18**	.22**	.35**	.15*	-.15*	.12	.17
3. T0 NRS unpleasantness	.41**	.83**		.33**	.46**	.004	.27**	.26**	.34**	.24**	-.35**	.10	.11
4. T0 CPASS Avoidance	.53**	0.12	.33**		.61**	-.11	.40**	.30**	.357**	.32**	-.64**	.15*	.27*
5. T0 PCS	.56**	.30**	.46**	.61**		-15*	.39**	.50**	.51**	.42**	-.57**	.27**	.45**
6. POD2 Activity	-.10	-.01	.01	-.12	-.13		-.14	-.12	-.16*	-.04	.01	-.09	.11
7. T0 MASC-10	.35**	.18**	.25**	.39**	.38**	-.18*		.47**	.46**	.55**	-.35**	.16*	.29*
8. T0 CRIES	.41**	.26**	.29**	.30**	.51**	-.09	.43**		.59**	.53**	-.27**	.22**	.25*
9. T0 CES-DC	.41**	.38**	.35**	.36**	.51**	-.17*	.45**	.59**		.52**	-.43**	.19*	.17
10. T0 CASI	.40**	.19**	.26**	.32**	.43**	-.05	.54**	.54**	.54**		-.28**	.20	.30*
11. T0 CPAQ	-.52**	-.16*	-.36**	-.64**	-.57**	-.01	-.31**	-.28**	-.42**	-.27**		-.24*	-.16
12. T3 FDI	.17*	.10	.11	.15*	.27**	-.11	.17*	.21**	.19**	.12	-.24**		-.84**
13. T3 PPIS	.34**	.19	.13	.30*	.44**	.043	.33**	.25*	.22	.32**	-.17	.81**	

Note: Pearson correlation coefficients are presented below the diagonal space. Partial correlation coefficients after controlling for age and gender are presented above the diagonal space.

T0 TSK, Tampa Scale for Kinesiophobia assessed prior to surgery; T0 PCS, Pain Catastrophizing Scale assessed prior to surgery; POD2, Post-Operative Day 2; T0 MASC, Multidimensional Anxiety Scale for Children assessed prior to surgery; T0 CRIES, Children's Revised Impact of Events Scale assessed prior to surgery; T0 CESDC, Centre for Epidemiological Studies – Depression Scale Children assessed prior to surgery; T0 CASI, Childhood Anxiety Sensitivity Inventory assessed prior to surgery; T0 CPAQ, Chronic Pain Acceptance Questionnaire for Adolescents assessed prior to surgery; T3 FDI, Functional Disability Inventory assessed 12-months after surgery; T3 PPIS, PROMIS Pediatric Pain Interference Scale assessed 12-months after surgery

$*p < .05$

$**p < .01$

Appendix: Results from the Exploratory Factor Analysis (EFA) and validity testing on the TSK-17.

EFA

The following table show the results from the five-factor model with quartinim rotation in terms of factor loadings, communality, and uniqueness. For ease of viewing, the items are ordered first with positively scored items and second with negatively scored items.

			Λ				
Item	f_1	f_2	f_3	f_4	f_5	Communality (h^2)	Uniqueness (u^2)
1	.63	.00	.06	.08	.17	.59	.41
2	.28	.17	.18	.29	.10	.37	.63
3	.06	.02	.61	.20	.02	.50	.50
5	.23	-.19	.28	-.16	.18	.24	.76
6	-.04	-.30	.37	-.02	.25	.32	.68
7	-.28	.00	.23	.14	.39	.29	.71
9	.44	-.12	.08	.27	.04	.41	.59
10	.11	-.05	.03	.54	.10	.42	.58
11	-.09	-.18	.31	.21	.17	.28	.72
13	.06	-.03	.05	.67	-.09	.45	.55
14	.09	.10	.03	-.01	.72	.59	.41
15	.37	-.01	-.07	.09	.51	.58	.42
17	-.07	.25	.01	.42	.14	.26	.74
4	.00	.67	-.06	.03	.05	.45	.55
8	.23	.35	.17	-.17	-.18	.25	.75
12	-.07	.59	.11	-.05	.13	.38	.62

16	.09	.21	.45	-.21	-.16	.28	.72

Construct Validity

　　Convergent Validity. The TSK-17 was significantly and moderately correlated with NRS pain intensity (r (228) = .24, $p <$. 001), NRS pain unpleasantness (r (162) = .43, $p <$. 001), the avoidance subscale of the CPASS (r (227) = .54, $p <$. 001) and the PCS (r (226) = .55, $p < $.001), indicating weak convergent validity with those measures given the requirement that $r > $.7. The TSK-13 was not significantly correlated with actual physical movement on post-operative day 2 (r (179) = -.08, $p = $.319) as measured by actigraphy, indicating poor convergent and predictive validity with physical movement levels. Correlations between the TSK and other measures are shown in the table below.

　　Discriminant Validity. The TSK-13 was significantly and moderately correlated with the MASC-10 (r (224) = .30, $p <$.001), CRIES (r (219) = .34, $p <$.001), CES-DC (r (225) = .37, $p <$.001), CASI (r (227) = .35, $p <$.001), and CPAQ (r (180)= -.53, $p < $.001), indicating adequate discriminant validity given the requirement that $r <$.7.

Predictive Validity

　　Baseline TSK-17 scores, measured prior to surgery, were not correlated with functional disability 12-months after surgery (r (198) = .12, $p = $.085) or pain related disability (r (70) = .22, $p = $.061), indicating poor predictive validity.

Correlation coefficients between the TSK-17 measured at baseline before surgery (T0) and other variables measured at various times in relation to surgery.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. T0 TSK-17													
2. T0 NRS Pain Intensity	.239**												
3. T0 NRS unpleasantness	.426**	.826**											
4. T0 CPASS Avoidance	.537**	0.124	.326**										
5. T0 PCS	.550**	.295**	.463**	.609**									
6. POD2 Activity	-0.075	-0.008	0.008	-0.119	-0.13								
7. T0 MASC-10	.301**	.176**	.253**	.392**	.375**	-.178*							
8. T0 CRIES	.344**	.256**	.289**	.300**	.507**	-0.092	.434**						
9. T0 CES-DC	.374**	.378**	.351**	.361**	.513**	-.174*	.452**	.592**					
10. T0 CASI	.351**	.188**	.259**	.322**	.429**	-0.047	.544**	.537**	.535**				
11. T0 CPAQ	-.531**	-.161*	-.360**	-.636**	-.566**	-0.005	-.312**	-.278**	-.420**	-.269**			
12. T3 FDI	0.122	0.100	0.105	.147*	.266**	-0.11	.168*	.213**	.189**	0.121	-.242**		

13. T3 PPIS	0.222	0.194	0.13	.301*	.442**	0.043	.328**	.252*	0.217	.320**	-0.172	.807**

T0 TSK, Tampa Scale for Kinesiophobia assessed prior to surgery; T0 PCS, Pain Catastrophizing Scale assessed prior to surgery; POD2, Post-Operative Day 2; T0 MASC, Multidimensional Anxiety Scale for Children assessed prior to surgery; T0 CRIES, Children's Revised Impact of Events Scale assessed prior to surgery; T0 CESDC, Centre for Epidemiological Studies – Depression Scale Children assessed prior to surgery; T0 CASI, Childhood Anxiety Sensitivity Inventory assessed prior to surgery; T0 CPAQ, Chronic Pain Acceptance Questionnaire for Adolescents assessed prior to surgery; T3 FDI, Functional Disability Inventory assessed 12-months after surgery; T3 PPIS, PROMIS Pediatric Pain Interference Scale assessed 12-months after surgery

*$p < .05$

**$p < .01$