ORIGINAL ARTICLE

Are Constructs of the Transtheoretical Model for Physical Activity Measured Equivalently Between Sexes, Age Groups, and Ethnicities?

Raheem J. Paxton, Ph.D. · Claudio R. Nigg, Ph.D. ·
Robert W. Motl, Ph.D. · Kelly McGee, B.A. · Dana McCurdy, M.P.H. ·
Caroline Horwath Matthai, Ph.D. · Rod K. Dishman, Ph.D.

Abstract

Background/Purpose Identifying mediators of physical activity change requires measurement instruments that are reliable, valid, and generalizable to multiple populations. Despite continued application of the transtheoretical model (TTM) to the study of physical activity, the structural components of the TTM measurement instruments have been understudied in diverse populations. **Methods** A multiethnic sample ($N=700$, $M_{age}=47$, 63% women, 38% Caucasian) of participants living in Hawaii completed TTM measures. The factor validity and measurement equivalence/invariance (ME/I) of *decisional balance*, *barrier self-efficacy*, *temptations*, and *processes of change* instruments were explored between men, women, age groups, and ethnicities. **Results/Conclusions** Measurement models of barrier self-efficacy and revised models of temptations and processes of change demonstrated sufficient evidence for ME/I among all subgroups. A revised model of decisional balance demonstrated sufficient evidence for ME/I between genders and among ethnicities, but not among age groups. Future research should examine the stability of these constructs across time.

Keywords Decisional balance · Measurement equivalence/invariance · Processes of change · Self-efficacy · Temptations

R. J. Paxton
Cancer Research Center of Hawaii,
Honolulu, HI, USA

C. R. Nigg · K. McGee · D. McCurdy
University of Hawaii,
Honolulu, HI, USA

R. W. Motl
University of Illinois at Urbana-Champaign,
Urbana-Champaign, IL, USA

C. H. Matthai
University of Otago,
Dunedin, New Zealand

R. K. Dishman
University of Georgia,
Athens, GA, USA

C. R. Nigg (✉)
Department of Public Health Sciences,
John A. Burns School of Medicine,
University of Hawaii—Manoa,
1960 East West Rd,
Honolulu, HI 96822, USA
e-mail: cnigg@hawaii.edu

Introduction

Despite the numerous benefits associated with moderate to vigorous physical activity (PA), more than half of US adults are not meeting participation recommendations [i.e., 30 min of moderate PA per day, five or more days of the week or vigorous PA at least 20 min per day, three or more days a week [1]]. Approximately 25% report that they engage in no PA for exercise in their leisure time [2]. Although physical inactivity is common among Americans, rates are higher among women, minorities, and those having little or no social capital [3]. Adults who continue to lead sedentary lifestyles as they age considerably increase their risks of developing cardiovascular disease, diabetes, obesity, and premature death [4]. It is therefore imperative to better understand behavior change models that explain variation in PA in efforts to design effective interventions that increase the adoption and maintenance of PA in adult populations [5].

The transtheoretical model (TTM) has emerged as a framework to understand how individuals initiate and adopt regular PA. The framework proposes that individuals move through a temporal sequence of five stages: *precontemplation* (no intention of becoming regularly physically active), *contemplation* (intending to become regularly physically active within the next 6 months), *preparation* (intending to become regularly physically active within the next 30 days), *action* (being regularly physically active 30 min per day, most days of the week, but only within the last 6 months), and *maintenance* (meeting the requirements of PA for at least 6 months) [6]. At each individual stage of readiness to become physically active, different factors are hypothesized to influence stage progression [7]. The constructs include *self-efficacy*, *temptations*, *decisional balance*, and ten *processes of change*.

Self-efficacy can be described as a persons' self-confidence to perform a specific task in challenging and tempting situations [8]. Of particular importance for stage transition, barrier self-efficacy can be described as a person's belief in capabilities to overcome personal, social, and environmental barriers to exercising [9, 10]. Generally, cross-sectional analyses have shown that the confidence to overcome barriers to PA increases linearly across stages [7]. Conceptually related to barrier self-efficacy, temptations describe urges to engage in a specific habit in the midst of difficult situations [6, 11]. Construct validity of temptations was recently established by significantly lower levels of temptations in the later stages [12]. Based on expectancy theory, decisional balance is a multidimensional set of values perceived as advantages and disadvantages of behavioral change [13]. Construct validity of the decisional balance inventory was demonstrated by an increase in pros and a decrease in cons across stages [7].

Finally, the processes of change are defined as overt and covert strategies that individuals use to alter their experiences and environment [14, 15]. The processes are divided into two higher order factors: (a) experiential and (b) behavioral processes, each consisting of five constructs [16]. Variable findings for the measurement characteristics of the processes of change instruments have led to some criticism of this construct [17]. For example, various item indicators have been known to load significantly on more than one construct, and the construct of social liberation does not appear to be unidimensional [17].

TTM constructs have previously been applied to the study of PA by a number of researchers [6, 7, 18–20]. However, few investigators have examined the structural and generalizability aspects of construct validity of the measurement scales. Structural aspects of construct validity test the fit of a theoretically based measurement model, which describes the pattern of relationships among a set of indicators and provides the basis for calculating scores that are used in all other tests of validity [21]. Of equal importance is the establishment of generalizability aspects of construct validity by testing of multi-group equivalence/invariance, the extent to which an instrument is measuring a construct or its operations (i.e., factor loadings) similarly between groups [22]. Although some evidence supports the invariance of individual scales [23], to our knowledge, the invariance of all the TTM constructs applied to PA have not been examined simultaneously in a single but diverse sample of adults.

Despite preliminary evidence for the construct validity of instruments used to assess individual TTM constructs in the PA domain, two important limitations of the evidence must be addressed. First, researchers have noted that the construct and factorial validity of the instruments has been understudied among multiethnic populations [17, 24], as well as among age and gender subgroups. Second, the processes of change are the least studied constructs in terms of factorial validity, which has been highly inconsistent (e.g., one to two constructs or ten constructs) across reports [7]. Given these limitations, the purposes of this study were to examine the factorial validity of the TTM measures and to determine if the underlying structure was equivalent/invariant between genders and among age groups and ethnicities. Generally, the establishment of multi-group factorial invariance is necessary before meaningful inferences between groups can be drawn about variables of interest [21].

Method

Participants

This cross-sectional study used a random sample of adults (18 years or older) from Hawaii (N=700; 63.3% women; mean age=47.0 years, SD=17.1; mean education= 14.6 years, SD=2.8; 51.6% married; median income= $40,000 to $50,000, SD=$28,000; 31.6% Asian, 22.2% native Hawaiian/Pacific Islander, 37.8% Caucasian, and 8.4% other). Participant characteristics are reported in Table 1. All data from this study are part of an ongoing observational study.

Procedure

The questionnaire was programmed into a computer-assisted telephone interview system by a local survey firm. Before survey administration, the questionnaire was pilot tested among seven individuals for interpretability and ease of administration. Once finalized, participants were recruited using random digit dialing procedures. A qualified individual whose birthday was closest to the date of the phone call was asked to participate. Trained interviewers informed potential participants that they would receive a

Table 1 Demographic characteristics

Characteristic	Frequency (%)
Gender	
Male	256 (36.6)
Female	442 (63.1)
Missing	1 (0.3)
Ethnicity	
Caucasian	262 (37.4)
Pacific Islanders	155 (22.1)
Asian	218 (31.1)
Other	60 (8.6)
Missing	5 (0.8)
Age group	
18–34	186 (26.1)
35–54	274 (39.1)
55+	231 (34.1)
Missing	1 (0.1)

$10 incentive if they agreed to participate in a 30-min interview regarding their PA and nutritional behaviors. Informed consent ensuring privacy and confidentially was obtained from participants. The University of Hawaii Institutional Review Board approved all procedures.

Measures

Self-Efficacy

This six-item instrument measures confidence to be physically active in the presence of barriers. Each item is rated on a five-point scale (1=not at all confident to 5= completely confident) and represents one of six-specific domains: negative affect, excuse making, being active alone, equipment access, resistance from others, and weather [16]. The self-efficacy scale was internally consistent (alpha=0.85) in this study.

Temptations

The two factor (i.e., affect and competing demands), ten-item temptations scale assesses how tempted an individual is not to be physically active [18]. The items are preceded by the sentence "Using the scale below, please indicate how TEMPTED you are NOT to exercise in the following situations". The responses were rated on a scale ranging from 0% (not at all tempted) to 100% (extremely tempted). Internal consistencies for this study were 0.87 and 0.91 for affect and competing demands subscales, respectively.

Pros and Cons (Decisional Balance)

This two factor, ten-item Likert-type scale measures the importance of the pros and cons of PA using a five-point

scale (1=not important to 5=extremely important) [25]. Internal consistencies for this study were 0.83 and 0.71 for pros and cons, respectively.

Processes

The Process of Change questionnaire includes 30 statements that participants are asked to rate in terms of frequency of occurrence over the past month. The questionnaire contains three items for each of the ten specific processes of change and provides individual scores (ranging from 1=Never to 5= Repeatedly) [14]. For this study, alpha coefficients ranged from 0.72 to 0.88 for experiential processes and from 0.76 to 0.85 for the behavioral processes.

Data Analysis

Before examining a model fit, individual subgroups were created for the demographic variables. Age was categorized into three groups (i.e., aged 18–34, 34–54, and 55 and older), representing younger, middle, and older adulthood. The study sample consisted of 17 separate ethnicities. Based on local interpretation, three ethnic groups were created: Caucasian (i.e., Caucasian and Portuguese), Pacific Islanders (i.e., Hawaiian/Part Hawaiian, Samoan/Tongan, other Pacific Islanders, and Guamanian/Chamorro), and Asians (i.e., Chinese, Japanese, Korean, Vietnamese, Indian—from India, and other Asians). A fourth group (i.e., African American, Hispanic, Native American, mixed non-Hawaiian, and Others) was excluded from the analysis because of low frequencies and heterogeneity among the ethnic groups.

Confirmatory Factor Analysis

To examine the structural aspects of construct validity of TTM components, individual measurement models were tested using confirmatory factor analysis with full-information maximum likelihood (FIML) estimation in Mplus version 3.13 [26]. FIML uses iterative simultaneous equations to estimate model parameters in the presence of missing data by computing a likelihood function for each individual based on all available data [27, 28]. In contrast to other techniques such as pairwise and listwise deletion of cases, FIML yields accurate fit indices and parameter estimates with up to 25% simulated missing data [28, 29]. The extent of missing data ranged from less than 1% for items on the self-efficacy and temptations scales to 11% for the processes of change items.

Model Fit

To establish a model fit, a series of estimates were used to determine if the structural models resembled a close, exact,

and absolute fit to the data. The chi-square statistic (χ^2) reveals an absolute fit to the data when the χ^2 is not statistically significant. The χ^2 is too sensitive to sample size; therefore, the comparative fit index (CFI) and the standardized root mean square residual (SRMR) were also used to judge model fit [30, 31]. Values of the CFI and SRMR reveal close fit to the data when values are ≥0.95 and ≤0.08, respectively [32, 33]. The SRMR was used for assessing model data fit because it results in lower probabilities of type I and type II errors when compared to the root mean square error approximation and the Tucker–Lewis index in sample sizes ≤250 [33]. Hu and Bentler [33] proposed that using cut of values of 0.96 for CFI in combination with values of SRMR<0.10 results in the least sum of type I and type II error rates [33]. In addition, estimates of factor loadings, intercepts, variances, residual variances, and z-scores (>1.96) were inspected for sign and magnitude.

Model Modifications

Modifications to the hypothesized structure were based on substantive and empirical information provided by modification indices in Mplus [26]. Modifications were made to the measurement model only when a change resulted in improved fit based on reduction in chi-square value, improved CFI or SRMR values, and if it was theoretically plausible.

Multi-group Factorial Invariance

The invariance of the selected instruments was measured using a multi-step approach [30, 34]. Initially, the hypothesized measurement model was tested individually in each group (e.g., men, Caucasians, 18–34 year olds, etc.). Secondly, we examined the extent to which parameters in the variance–covariance matrices (equal sigmas) were invariant between and among groups. The test of equal sigmas often produces inconsistent results as an initial test of invariance [34] and may not be always be an indication that item parameters are invariant between or among groups. Accordingly, additional tests are required. We then tested sequential comparisons of two nested models in which additional successive constraints were imposed on model parameters to ensure equality of the measurement structure and factor loadings. Evidence of equal factor loadings (i.e., weak factorial invariance) provides the minimal requirement that a measurement instrument is operating similarly between/among groups [34]. The two nested models were compared based on evaluating the difference in χ^2 in relation to change in (Δ) degrees of freedom (df) of the model with no or less constraints to the model with more constraints [30]. Change in CFI less than or equal to 0.01 suggests that the invariance of an instrument should not be rejected [35].

Therefore, if the χ^2 difference test is significant but the CFI change is less than 0.01, there is some evidence for the equivalence/invariance of the model structure or parameters between groups.

Results

The Measurement Model of Barrier Self-Efficacy

The hypothesized single-factor model for barrier self-efficacy displayed in Fig. 1 represented a good fit to the data in the entire sample (χ^2=68.88, df=9, CFI=0.95, SRMR=0.03). Factor loadings, intercepts, variances, factor variances, and z-scores (>1.96) were appropriate in sign and magnitude.

Factorial Validity of Barrier Self-Efficacy

The measurement model for the barrier self-efficacy scale provided an acceptable fit for five of the eight subgroups. Marginal fit of the measurement model was observed for participants between the ages of 18 to 34 years and 35 to 54 years and for Pacific Islanders. Fit statistics and alpha coefficients for each subgroup are reported in Table 2.

Multi-group Factorial Invariance of Barrier Self-Efficacy

The test of equal sigmas between men and women provided a good fit (χ^2=26.20, df=21, CFI=0.99, SRMR=0.04) and supported that the structure underlying the items was invariant between men and women. The test of equal sigmas was only marginally supported among age groups (χ^2=93.72, df=42,CFI=0.96, SRMR=0.11) and ethnicities (χ^2=72.65, df=42, CFI=0.98, SRMR=0.11). The two nested tests in the multi-group factorial invariance routine indicated that the factor structure and factor loadings were invariant between sexes ($\Delta\chi^2$=5.46, Δdf=5, p=NS; ΔCFI=0.00), among age groups ($\Delta\chi^2$=15.30, Δdf=10, p=NS; ΔCFI=0.00), and ethnicities ($\Delta\chi^2$=11.4, Δdf=10, p=NS; ΔCFI=0.00).

The Measurement Model of Temptations

The hypothesized two-factor model of temptations displayed in Fig. 1 did not provide a good fit to the data (χ^2= 385.83, df=34, CFI=0.92, SRMR=0.07). Two items (i.e., when you're alone and when you're out of shape) were removed from the affect scale, and one item (i.e., when you feel lazy) was removed from the competing demands scale. The resulting model revealed a good fit to the data for the total sample (χ^2=76.26, df=13, CFI=0.97, SRMR=0.04). There
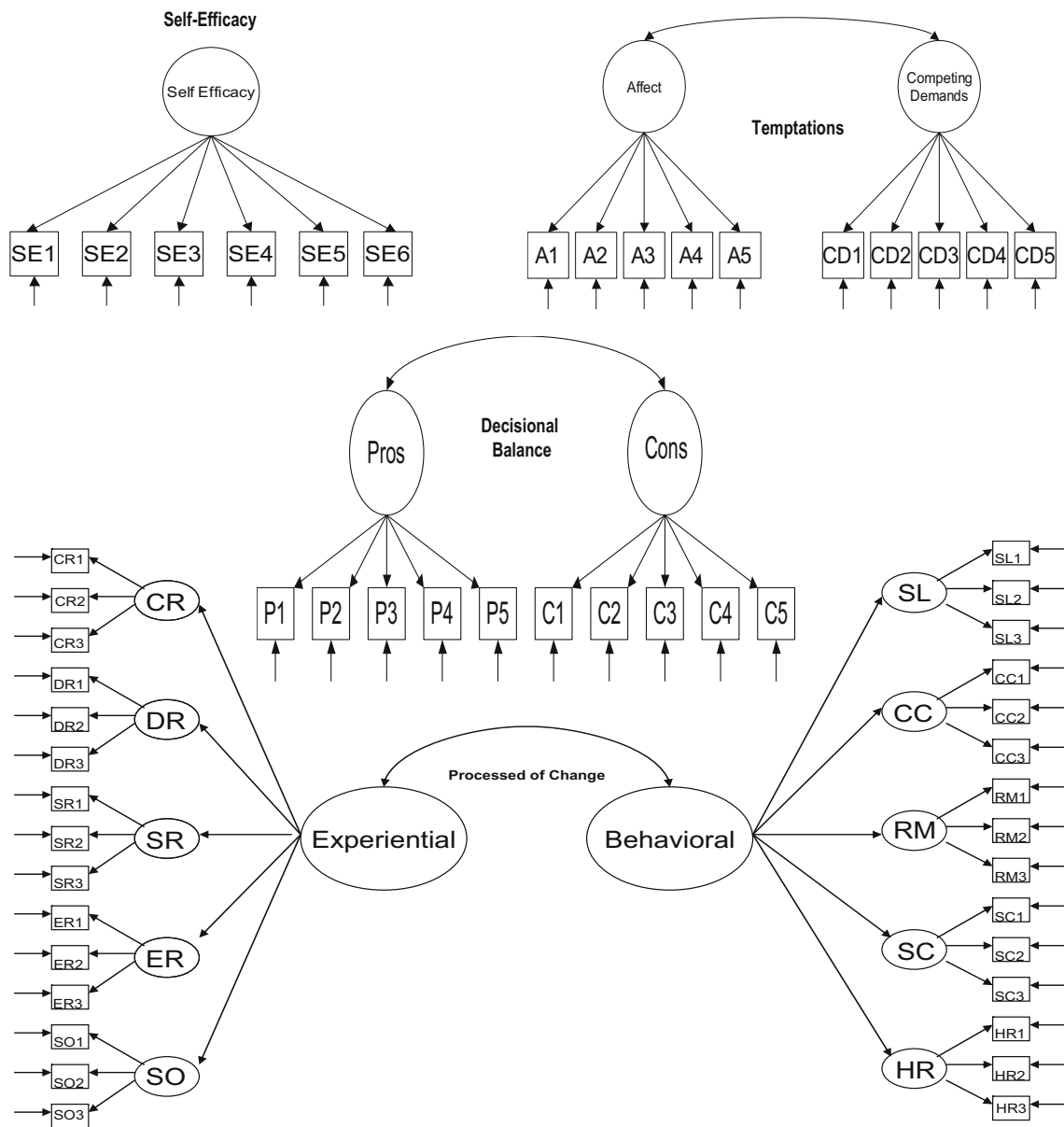
Fig. 1 Individual hypothesized measurement models for self-efficacy, temptations, decisional balance, and the processes of change. *SE1– SE6* represent item indicators for barrier self-efficacy. *A1–A5* represent the item indicators for the affect component of temptations measure. *CD1* to *CD5* represent the item indicators for the competing demands component of temptations measure. *P1* to *P5* represent the item indicators for the pro component of decisional balance measure. *C1* to *C5* represent the item indicators for the con component of the decisional balance measure. *CR*, *DR*, *SR*, *ER*, and *SO* represent consciousness raising, dramatic relief, self-reevaluation, environmental reevaluation, and social liberation of the experiential processes of change, respectively. *SL*, *CC*, *RM*, *SC*, and *HR* represent self-liberation, counter conditioning, reinforcement management, stimulus control, and helping relationships of the behavioral processes of change, respectively. There are three-item indicators for each processes construct

was a statistically significant correlation observed between the affect and competing demands components ($\varphi=0.59$, $p<0.01$). Factor loadings, intercepts, variances, factor variances, and z-scores (>1.96) were appropriate in sign and magnitude.

Factorial Validity of Temptations

The revised measurement model of temptations not to be physically active provided an appropriate fit to the data for all the subgroups analyzed. Fit statistics and alpha coefficients for each subgroup are reported in Table 2.

Multi-group Factorial Invariance of Temptations

The test of equal sigmas provided a good fit in analyses comparing men and women ($\chi^2=40.15$, $df=28$, CFI=0.99, SRMR=0.02), age groups ($\chi^2=74.69$, $df=56$, CFI=0.99, SRMR=0.09) and ethnicities ($\chi^2=99.92$, $df=56$, CFI=0.98,

SRMR=0.07). This suggests that the structure underlying item responses was invariant between and among the groups. The nested analyses in the multi-group factorial invariance routine indicated that the factor structure and factor loadings were invariant between sexes ($\Delta\chi^2$=3.53, Δdf=5, p=NS; ΔCFI=0.00), among age groups ($\Delta\chi^2$= 11.6, Δdf=10, p=NS; ΔCFI=0.00), and ethnicities ($\Delta\chi^2$= 7.44, Δdf=10, p=NS; ΔCFI=0.00).

The Measurement Model of Decisional Balance

The hypothesized model of decisional balance displayed in Fig. 1 revealed a good fit to the data in the entire sample (χ^2=111.98, df=34, CFI=0.95, SRMR=0.04). The correlation between pro and con scales was not significant (φ= 0.14, p>0.05). Factor loadings, intercepts, variances,

residual variances, and z-scores (>1.96) were appropriate in sign and magnitude.

Factorial Validity of Decisional Balance

The test of factorial validity of decisional balance suggested that there was differential model fit among the subgroups analyzed. Of particular importance was the lack of fit observed among the men. Therefore, we re-specified the measurement model for the entire sample by eliminating the final item (my exercise put an extra burden on my significant other) of the con (barrier) factor. We removed the con item because nearly 50% of the sample reported never being married; this modification significantly improved the fit of the measurement model to the data (χ^2= 69.02, df=26, CFI=0.97, SRMR=0.04). A statistically significant correlation was observed between the pro and

Table 2 Fit indices and reliabilities for barrier self-efficacy, temptations, and decisional balance by genders, age groups, and ethnicities

Measure subgroup	χ^2 (df)	CFI	SRMR	Cronbach α			
				Affect	Competing demands	Pros	Cons
Barrier self-efficacy							
Total sample	68.88 (9)	0.95	0.03	0.85			
Males	30.33 (9)	0.96	0.03	0.84			
Females	55.18 (9)	0.95	0.04	0.84			
18–34 years	31.75 (9)	0.94	0.04	0.83			
35–54 years	52.02 (9)	0.93	0.05	0.84			
55+ years	20.30 (9)	0.98	0.03	0.84			
Caucasian	36.75 (9)	0.95	0.04	0.83			
Pacific Islanders	34.86 (9)	0.94	0.04	0.87			
Asian	26.41 (9)	0.95	0.04	0.81			
Temptations (i.e., affect and competing demands)							
Total sample	76.26 (13)	0.97	0.04	0.87	0.91		
Men	31.01 (13)	0.98	0.04	0.79	0.90		
Women	58.57 (13)	0.98	0.04	0.77	0.91		
18–34 years	29.51 (13)	0.98	0.05	0.73	0.91		
35–54 years	24.83 (13)	0.99	0.04	0.78	0.91		
55+ years	49.13 (13)	0.96	0.04	0.79	0.90		
Caucasian	41.10 (13)	0.97	0.05	0.81	0.90		
Pacific Islanders	21.12 (13)	0.99	0.04	0.76	0.92		
Asian	37.54 (13)	0.97	0.05	0.74	0.91		
Decisional Balance (i.e., pros and cons)							
Total sample	69.02 (26)	0.97	0.04			0.83	0.71
Men	56.14 (26)	0.95	0.07			0.83	0.68
Women	66.74 (26)	0.96	0.04			0.81	0.70
18–34 years	39.29 (26)	0.96	0.05			0.80	0.64
35–54 years	58.95 (26)	0.95	0.06			0.79	0.74
55+ years	35.05 (26)	0.99	0.04			0.88	0.71
Caucasian	40.07 (26)	0.97	0.04			0.82	0.60
Pacific Islanders	26.87 (26)	0.99	0.04			0.81	0.72
Asian	66.75 (26)	0.94	0.08			0.86	0.74

α alpha or internal consistency, *df* degrees of freedom, χ^2 chi-square, *CFI* comparative fit index, *SRMR* standardized root mean square residual

con factor ($\varphi=0.11$, $p<0.05$). This revised measurement model provided an appropriate fit for all subgroups analyzed. See Table 2 for subgroup fit statistics and internal consistencies.

Multi-group Factorial Invariance of Decisional Balance

The test of equal sigmas among age groups was partially supported ($\chi^2=161.54$, $df=90$, CFI=0.96, SRMR=0.10), suggesting that the structure underlying item responses was partially invariant among age groups. The tests of equal sigmas comparing sexes ($\chi^2=175.86$, $df=45$, CFI=0.92, SRMR=0.16) and ethnicities ($\chi^2=227.26$, $df=90$, CFI=0.91, SRMR=0.15) were not acceptable, indicating that the structure underlying item responses differed based on groups. The nested analyses in the multi-group factorial invariance routine indicated that the factor structure and factor loadings were invariant between sexes ($\Delta\chi^2=9.48$, $\Delta df=4$, $p=$NS; ΔCFI=0.00) and among ethnicities ($\Delta\chi^2=27.6$, $\Delta df=14$, $p<0.05$; ΔCFI=0.01), but that the factor loadings were not invariant among age groups ($\Delta\chi^2=36.6$, $\Delta df=14$, $p<0.05$; ΔCFI=0.01).

The Measurement of the Processes of Change Model

The hypothesized ten-factor solution for the processes of change is presented in Fig. 1 and was not admissible owing to a negative residual variance of the self-liberation factor. The residual variance of the self-liberation items was therefore fixed to zero to enable model conversion. The resulting model did not reveal an acceptable fit to the data ($\chi^2=1,595.36$, $df=396$, CFI=0.86, SRMR=0.06). Two revised measurement models for the processes of change displayed in Fig. 2 were then tested for factorial validity. In an effort to preserve parsimony and be theoretically consistent, an iterative process was applied to revise the measurement model of the processes of change. This iterative process consisted of running sequential exploratory factor analyses extracting one item at a time until a parsimonious model was created. Figure 2a consists of a two-factor second order model represented by 18 items from 9 of the original processes of change factors. The individual processes factors were positively correlated with the original behavioral [r (range)=0.31 to 0.87] and experiential [r (range)=0.38 to 0.74] processes. This model
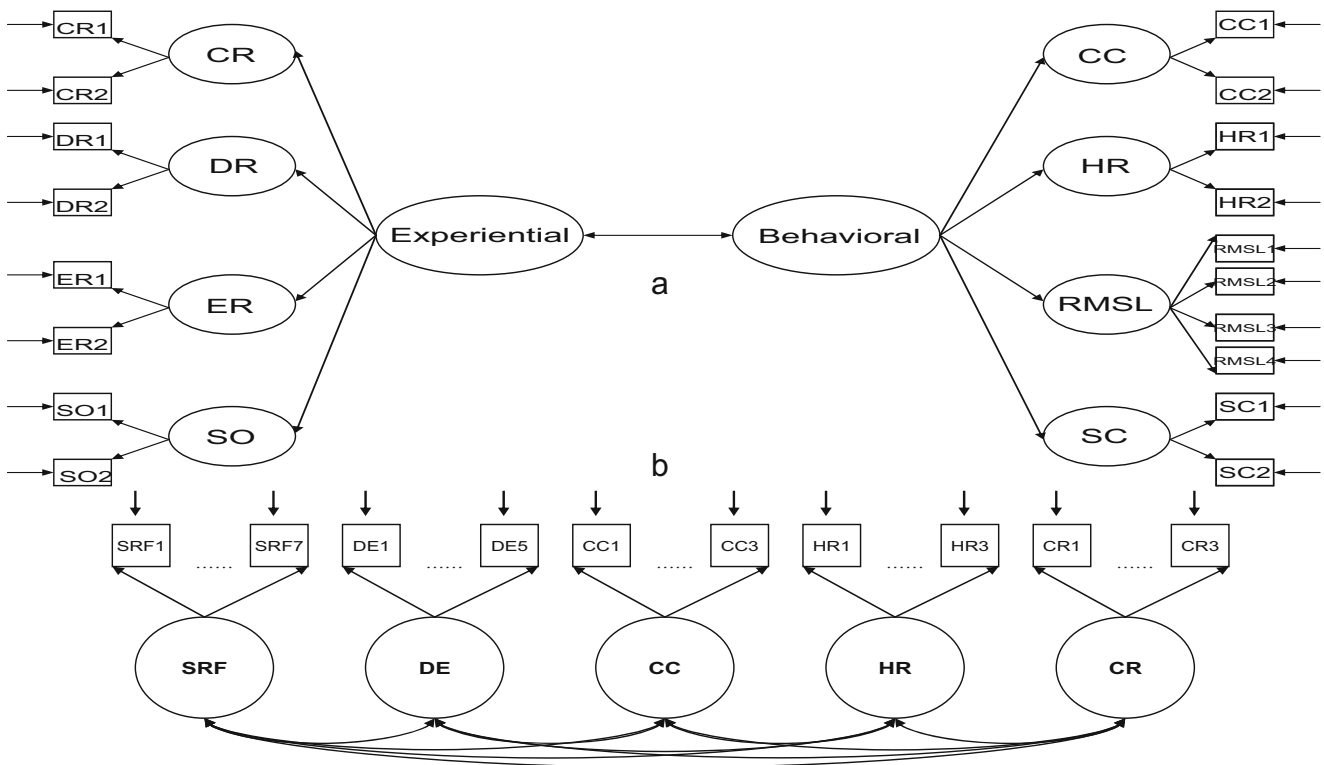

Fig. 2 Proposed measurement models for the process of change. **a** Depicts a two-factor second order process of change model. The experiential higher order factor is represented by first order factors consciousness raising (*CR*), dramatic relief (*DR*), environmental reevaluation (*ER*), social liberation (*SO*), and their indicators. The behavior factor is represented by first order factors of reinforcement management and self-liberation (*RMSL*), counter conditioning (*CC*), helping relationships (*HR*), and stimulus control (SC). **b** Depicts a five-factor measurement model of the processes of change. Factor 1 is represented seven items from the constructs of self-revaluation, reinforcement management, and self liberation (*SRF*). Five items represent factor 2 from dramatic relief and environmental reevaluation (*DE*). Factors 3 to 5 are represented by three-item indicators each for counter conditioning (*CC*), helping relationships (*HR*), and consciousness raising (*CR*)

provided marginal fit to the data (χ^2=382.91, df=126, CFI=0.94, SRMR=0.05). A significant correlation was observed between behavioral and experiential second order factors (φ=0.82, p<0.05). Based on recommendations and recognition of model fit observed in another multi-ethnic sample (R. K. Dishman, personal communication, September 15, 2006), a five-factor correlated model was also created. Two factors (i.e., stimulus control and social liberation) were removed from the analysis because of lack of simple structure. In addition, there were two factors that were represented by indicators from more than one of the hypothesized constructs. Factor 1 was created with indicators of self-reevaluation (three items), reinforcement management (two items), and self-liberation (two items). Factor 2 was created with indicators of dramatic relief (two items) and environmental reevaluation (three items). Counter conditioning, helping relationships, and consciousness raising were represented by their original item indicators. The five-factor processes scale was positively correlated with the original behavioral [r (range)=0.54–0.81] and experiential [r (range)=0.40–0.82] processes. This five-factor model also represents a marginal fit to our data (χ^2=579.95, df=179, CFI=0.93, SRMR=0.05). The five-factor model is depicted in Fig. 2b.

Factorial Validity of the Revised Eight- and Five-Factor Processes of Change Model

Results of the analyses for the two-factor second order model provided a marginal fit for most of the subgroups, and less than reasonable fit was observed for those 35 to 54 years old, Pacific Islanders, and Asians. Similar indices of model fit were observed for the five-factor correlated model, with less than reasonable fit observed for Pacific Islanders. A fit index for all subgroups for both revised measurement models is reported in Table 3.

Multi-group Factorial Invariance of the Revised Two-Factor Higher Order Model

The test of equal sigmas between sexes (χ^2=175.07, df=170, CFI=0.999, SRMR=0.050), among age groups (χ^2=487.03, df=340, CFI=0.966, SRMR=0.070), and ethnicities (χ^2=497.52, df=340, CFI=0.960, SRMR=0.075) was supported, suggesting that the structure underlying item responses was invariant between and among the groups. The nested analysis in the multi-group factorial invariance routine indicated that the factor structure and factor

Table 3 Fit indices and reliabilities for two proposed measurement models of the processes of change by genders, age groups, and ethnicities

Measure subgroup	χ^2 (df)	CFI	SRMR	Cronbach α						
				Behavioral	Experiential	SRF	DE	CC	HR	CR
Revised two-factor second order model										
Total sample	184.74 (126)	0.96	0.05	0.80	0.72					
Men	184.74 (126)	0.96	0.05	0.80	0.72					
Women	318.54 (126)	0.93	0.05	0.78	0.72					
18–34 years	199.48 (126)	0.94	0.06	0.79	0.67					
35–54 years	290.84 (126)	0.90	0.06	0.78	0.76					
55+ years	233.59 (126)	0.93	0.05	0.81	0.74					
Caucasian	195.24 (126)	0.95	0.05	0.80	0.73					
Pacific Islanders	210.09 (126)	0.91	0.06	0.78	0.74					
Asian	272.61 (126)	0.90	0.07	0.80	0.74					
Revised five factor process of change model										
Total sample	330.38 (179)	0.93	0.06			0.88	0.72	0.82	0.77	0.80
Men	330.38 (179)	0.93	0.06			0.88	0.72	0.82	0.77	0.80
Women	492.95 (179)	0.91	0.06			0.88	0.77	0.83	0.79	0.75
18–34 years	327.59 (179)	0.92	0.07			0.90	0.77	0.85	0.74	0.77
35–54 years	382.41 (179)	0.92	0.06			0.87	0.77	0.84	0.82	0.80
55+ years	299.43 (179)	0.93	0.06			0.88	0.72	0.79	0.77	0.77
Caucasian	327.54 (179)	0.93	0.06			0.88	0.73	0.79	0.78	0.77
Pacific Islanders	329.52 (179)	0.89	0.07			0.87	0.82	0.84	0.74	0.77
Asian	308.58 (179)	0.97	0.06			0.90	0.74	0.85	0.81	0.77

α alpha or internal consistency, *F1 to F5* factors 1 through factors 5, *df* degrees of freedom, χ^2 chi-square, *CFI* comparative fit index, *SRMR* standardized root mean square residual, *SRF* self-revaluation, reinforcement management, and self liberation, *DE* dramatic relief and environmental reevaluation, *CC* counter conditioning, *HR* helping relationships, *CR* consciousness raising

loadings were invariant between sexes ($\Delta\chi^2$=6.02, Δdf= 10, p=NS; ΔCFI=0.001), age groups ($\Delta\chi^2$=18.05, Δdf= 20, p=NS; ΔCFI=0.000), and ethnicities ($\Delta\chi^2$=18.21, Δdf=20, p=NS; ΔCFI=0.001).

Multi-group Factorial Invariance of the Revised Five-Factor Model

The test of equal sigmas for the five-factor processes of change model suggested that the structure underlying item responses was invariant between genders (χ^2=340.38, df= 322, CFI=0.982, SRMR=0.064), among age groups (χ^2= 615.83, df=463, CFI=0.975, SRMR=0.061), and ethnicities (χ^2=636.08, df=463, CFI=0.969, SRMR=0.067). The nested analysis in the multi-group factorial invariance routine also indicated that the factor structure and factor loadings were invariant between genders ($\Delta\chi^2$=30.7, Δdf=16, p<.05; ΔCFI=0.000), among age groups ($\Delta\chi^2$= 37.9, Δdf=32, p=NS; ΔCFI=0.001), and ethnicities ($\Delta\chi^2$= 25.3, Δdf=32, p=NS; ΔCFI=0.001).

Discussion

TTM constructs are widely used in PA studies in the USA and abroad [7]. Despite extensive use of TTM construct measures, there is no research that we know of that has investigated the factorial validity and appropriateness of these measures among men and women of different age groups and from ethnically diverse backgrounds. Therefore, the purpose of this study was to examine the factorial validity and multi-group equivalence/invariance of scales measuring barrier self-efficacy, temptations not to be physically active, decisional balance (pros and cons), and the processes of change.

The hypothesized measurement model for self-efficacy adequately fit the data for the entire sample and represented sufficient evidence of equivalence/invariance (i.e., equivalence/invariance of variance–covariance matrices, factor structure, and factor loadings) for each gender, age group, and ethnicity. Less than ideal fit of the measurement model to the data was observed for Pacific Islanders and for participants aged 18–34 and 35–54 years. However, the general pattern suggests that the model fits adequately to the data and appears to be operating similarly among the subgroups. Such results are expected because many studies have used similar items for barrier self-efficacy inventories [9, 19].

Temptations not to be physically active are one of the least studied constructs of the TTM. Only recently developed, the hypothesized measurement model proposed by Hausenblas et al. [12] did not adequately fit the data for our sample, but the model was significantly improved by the removal of three items. The revised

model of affect and competing demands closely fit the data for all subgroups and provided evidence of equivalence/ invariance.

Decisional balance measured by two correlated factors (i.e., pros and cons) closely fit the data for the total sample; however, some subgroup variation was observed. Particularly, less than ideal fit was observed for Asians, men, and participants between the ages of 18–34 and 35–54 years old. In addition, we also observed that the correlation between the pro and con factors was nonsignificant for all populations with the poorest fit. Generally, theory would suggest that the correlation would be small but significant and negative, as observed elsewhere in other decisional balance instruments [36, 37]. A large proportion of the population sampled within this study was in the maintenance phase, where some uncoupling of the relation between barriers (pros) and benefits (cons) is expected. Previous studies have reported that both the benefits (pros) and barriers (cons) of PA tend to level off with stage increase and become irrelevant factors in sustained participation in PA [20, 38].

We re-specified the measurement model because of the insufficient fit of the measurement model to the data. Improvement of fit was observed when the question "my exercise put an extra burden on my significant other" was removed. Once the modifications were made to the measurement model, the scale demonstrated the required evidence of equivalence/invariance among genders and ethnicities. Measurement equivalence/invariance for decisional balance among age groups was not supported as observed by the test of nested model, which, when compared to the test of equal sigmas, produces more consistent results [34]. When constraints were applied to the factor loadings, fit indices suggested that the factor loadings were not entirely equivalent/invariant among age groups; however, there was some evidence that the variance–covariance matrices were invariant among age groups. Previous studies have found evidence of factorial validity of a decisional balance instrument with six- and four-item inventories [37, 39], yet no test of invariance has been explored with these instruments.

Self-efficacy, temptations, and decisional balance provided sufficient evidence of factorial validity and ME/I for most subgroups; however, the hypothesized solution for the processes of change did not fit the data. Less than ideal fit was observed among all subgroups examined in the study, as observed elsewhere [16–18], suggesting that the data, measurement, or concepts are problematic. The measurement does not seem to be the culprit because studies have used a variety of measures as operationalizations of the processes constructs. Therefore, the two likely problems of data or concept need to be addressed within a longitudinal framework.

We proposed two different models to describe the processes of change. The two-factor second order model preserves seven of the original model processes of change factors. The construct of self-reevaluation was deleted because of colinearity or cross-loadings observed with two processes (i.e., reinforcement management and self-liberation) from the second order behavioral factor. Although not reported, we also explored the measurement model for an eight-factor correlated model with no second order factor. The eight-factor correlated model provided a better fit to the data when compared to two-factor second order model; however, to preserve the original structure of the process of change model, we decided to proceed with the higher order model. Of importance was the correlation observed between the experiential and behavioral factors. Previous studies have reported correlations above 0.95 between the experiential and behavioral factors [14, 16, 18, 20]. For our data, we observed a correlation of 0.82, suggesting that the items were not redundant. This is a significant contribution for the literature and suggests that the second order factor of behavioral and experiential processes may be related but independent constructs and should be measured as so. The alternative more parsimonious five-factor process model appeared to represent adequate fit to the data in this cross-sectional assessment, preserve three of the original processes of change, and represent both experiential and behavioral processes.

In regard to the revised process of change models, either case appears to hold in our sample, but more research is needed to determine how correlated these models are with other psychosocial variables and behavior over time and in different populations. Either revision of the processes provides direction for potential progress in the development of a parsimonious scale, because previous models of the processes reported lower CFIs than we reported here [14, 16, 18]. Despite some variations among subgroups, the revised measurement models provided evidence of equivalence/invariance, indicating that the scale was operating similarly among subgroups.

Establishing that a measurement model is invariant between subgroups is a prerequisite to comparing scores between subgroups, but multi-group invariance is often overlooked in health behavior research [40]. Traditional psychometric evaluations (e.g., internal consistency, predictive validity) are important, but not sufficient in determining whether a construct is being measured equivalently between groups. Tests of measurement equivalence/invariance ensure that psychosocial instruments are being measured equivalently between groups [40]. When such tests are not applied, there is increased risk of comparing subgroups on nonequivalent measures, thus biasing the interpretation from study results. Our results provide initial evidence that the TTM constructs applied to PA can be measured similarly between groups that differ according to gender, age, and ethnicity.

Overall, the study observed slight variation between subgroups in the measurement models for the scales assessing TTM constructs. However, this is a common phenomenon with small/borderline sample sizes [41]. Measurement models explored in the presence of small sizes create potentially unstable parameter estimates and fit indices [42]. Although there is a common rule that sample sizes of 200 or more are appropriate for confirmatory factor analysis, additional factors such as the number of item indicators and non-normal distributions substantially affect power in most instances [41]. In addition, the fit statistics observed here may be lower for certain populations because subgroups may be confounded by stage given most (>50%) of the subjects within the sample self-reported as being in the maintenance stage.

In sum, our study results are promising despite small sample sizes and lack of ideal fit (CFIs≥0.95). Future studies should explore the temporal stability of the measurement models over time and examine whether the structure may change over time. In addition, future studies should examine how the processes of change, decisional balance, barrier self-efficacy, and temptations moderate and mediate, respectively, self-initiated change in PA.

References

1. Centers for Disease Control and Prevention. Adult participation in recommended levels of physical activity—United States, 2001 and 2003. *MMWR*. 2005; 54: 1208–1212.
2. Centers for Disease Control and Prevention. Trends in leisure-time physical activity by age, sex, and race/ethnicity—United States, 2001 and 2003. *MMWR*. 2005; 4: 991–994.
3. Crespo CJ, Smith E, Andersen RE, Carter-Pokras O, Ainsworth BE. Race/ethnicity, social class and their relation to physical inactivity during leisure time: Results from the Third National Health and Nutrition Examination Survey, 1988–1994. *Am J Prev Med*. 2000; 18: 46–53.
4. Allison DB, Fontaine KR, Manson JE, Stevens J, VanItallie TB. Annual deaths attributable to obesity in the United States. *JAMA*. 1999; 282: 1530–1538.
5. Lewis BA, Marcus BH, Pate RR, Dunn AL. Psychosocial mediators of physical activity behavior among adults and children. *Am J Prev Med*. 2002; 23: 26–35.
6. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot*. 1997; 12: 38–48.
7. Marshall SJ, Biddle SJ. The transtheoretical model of behavior change: A meta-analysis of applications to physical activity and exercise. *Ann Behav Med*. 2001; 23: 229–246.

8. Bandura A. The anatomy of stages of change. *Am J Health Promot*. 1997; 12: 8–10.

9. McAuley E, Mihalko S. Measuring exercise related self efficacy. In: Duda J, ed. *Advancement in Measurement in Sport and Exercise Psychology*. Fitness Technology Publishers, 1998.

10. Benisovich SV, Rossi JS, Norman GJ, Nigg CR. Development of a multidimensional measure of exercise self-efficacy. *Ann Behav Med*. 1998; 20: 190.

11. Grimley D, Prochaska JO, Velicer WF, Blais WF, DiClemente CC. The transtheoretical model of change. In: Brinthaupt TM, Lipka RP, eds. *Changing the Sell: Philosophies, Techniques and Experiences*. Albany: State University of New York; 2004: 201–227.

12. Hausenblas HA, Nigg CR, Dannecker EA, et al. A missing piece of the transtheoretical model applied to exercise: Development and validation of the temptation to not exercise scale. *Psychol Health*. 2001; 16: 381–390.

13. Janis I, Mann M. *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: The Free Press; 1997.

14. Nigg CR, Norman GJ, Rossi JS, Benisovich SV. Processes of exercise behavior change: Redeveloping the scale. *Ann Behav Med*. 1998; S79.

15. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change. Applications to addictive behaviors. *Am Psychol*. 1992; 47: 1102–1114.

16. Prochaska JO, Velicer WF, DiClemente CC, Fava J. Measuring processes of change: Applications to the cessation of smoking. *J Consult Clin Psychol*. 1988; 56: 520–528.

17. Rhodes RE, Berry T, Naylor P, Higgins SJ. Three-step validation of exercise behavior processes of change in an adolescent sample. *Meas Phys Educ Exerc Sci*. 2004; 8: 1–20.

18. Marcus BH, Rossi JS, Selby VC, Niaura RS, Abrams DB. The stages and processes of exercise adoption and maintenance in a worksite sample. *Health Psychol*. 2002; 11: 386–395.

19. Marcus BH, Selby VC, Niaura RS, Rossi JS. Self-efficacy and the stages of exercise behavior change. *Res Q Exerc Sport*. 1992; 63: 60–66.

20. Marcus BH, Rakowski W, Rossi RS. Assessing motivational readiness and decision-making for exercise. *Health Psychol*. 1992; 11: 257–261.

21. Messick S. Standards of validity and the validity of standards in performance assessment. *Educ Meas*. 1995; 14: 5–8.

22. Byrne BM, Shavelson RJ, Muthen B. Testing for the equivalence of factor covariance and mean structures: The issues of partial measurement invariance. *Psychol Bull*. 1989; 105: 456–466.

23. Anatchkova MD, Redding CA, Rossi JS. Development and validation of Decisional Balance and Temptations measures for Bulgarian adolescent smokers. *Addict Behav*. 2006; 31: 155–161.

24. Whitelaw S, Baldwin S, Bunton R, Flynn D. The status of evidence and outcomes in Stages of Change research. *Health Educ Res*. 2000; 15: 707–718.

25. Nigg CR, Rossi JS, Norman GJ, Benisovich SV. Structure of decisional balance for exercise adoption. *Ann Behav Med*. 1998; 20: S211.

26. Muthen LK, Muthen BO. Mplus: *Statistical analysis with latent variables*. 3rd Ed. Los Angeles: Muthen & Muthen, 1998–2004.

27. Arbuckle J, Wothke W. *AMOS 4 user's reference guide*. Chicago: Smallwaters Corporation; 1999.

28. Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct Equ Modeling*. 2001; 8: 430.

29. Arbuckle JL. Full information estimation in the presence of incomplete data. In: Marcoulides GA, Schumacker RE, eds. *Advanced Structural Equation Modeling: Issues and Techniques*. Mahwah: Erlbaum; 1996: 243–277.

30. Bollen K. *Structural equations with latent variables*. New York: Wiley; 1989.

31. Joreskog KG. Testing structural equation models. In: Bollen LJ, ed. *Testing Structural Equation Models*. Newbury Park: Sage Publications; 1993: 294–316.

32. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990; 107.

33. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999; 6: 1–55.

34. Marsh HW. Confirmatory factor analysis models of factorial invariance: A multifaceted approach. *Struct Equ Modeling*. 1994; 1: 5–34.

35. Cheung GW, Rensvold RB. Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Struct Equ Modeling*. 2002; 9: 233–255.

36. Plummer BA, Velicer WF, Redding CA, et al. Stage of change, decisional balance, and temptations for smoking: Measurement and validation in a large, school-based population of adolescents. *Addict Behav*. 2001; 26: 551–571.

37. Marcus BH, Eaton CA, Rossi JS, Harlow LL. Self-efficacy, decision-making, and stages of change: An integrative model of physical exercise. *J Appl Soc Psychol*. 1994; 24: 489–508.

38. Taylor C, Boudreaux E, Jeffries S, Scarinci I, Brantley P. Applying exercise stage of change to a low income underserved population. *Am J Health Behav*. 2003; 27: 99–107.

39. Velicer WF, DiClemente CC, Prochaska JO, Brandenburg N. A decisional balance measure for assessing and predicting smoking status. *J Pers Soc Psychol*. 1985; 48: 1279–1289.

40. Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: Suggestions, practices and recommendations for organizational research. *Organ Res Methods*. 2000; 3: 4–70.

41. Hatcher L. *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary: SAS Institute; 1994.

42. Kaplan D, George R. A study of the power associated with testing factor mean differences under violations of factorial invariance. *Struct Equ Modeling*. 1995; 2: 101–118.

43. Nigg CR, Riebe D. The transtheoretical model: Research review of exercise behavior and older adults. In: Burbank PM, Riebe D, eds. *Promoting Exercise and Behavior Change in Older Adults: Interventions with the Transtheoretical Model*. New York: Springer; 2002: 147–180.