Iran J Public Health, Vol. 51, No.11, Nov 2022, pp.2555-2563

Original Article

Perceived Health Outcomes of Recreation Scale: Measurement Invariance over Gender

Elif Köse [1], Tennur Yerlisu-Lapa [1], Nezaket Bilge Uzun [2], *Evren Tercan Kaas [3], Emrah Serdar [4], Gökşen Aras [5]

1. *Department of Recreation, Sport Sciences Faculty, Akdeniz University, Antalya, Türkiye*
2. *Department of Educational Measurement and Evaluation, Faculty of Education, Mersin University, Mersin, Türkiye*
3. *Department of Sport Management, Sport Sciences Faculty, Akdeniz University, Antalya, Türkiye*
4. *Department of Recreation, Sport Sciences Faculty, İstanbul University-Cerrahpaşa, İstanbul, Türkiye*
5. *Department of English Language and Literature, Faculty of Arts and Sciences, Atılım University, Ankara, Türkiye*

***Corresponding Author:** Email: evrentercan@akdeniz.edu.tr

(Received 15 May 2022; accepted 23 Jul 2022)

Abstract

Background: Research handling structural differences among groups presume that the measurement tool works similarly among the groups and the results of measurements provide similar psychometric properties. Therefore, the aim of the study is to provide evidence for measurement invariance of the construct validity Perceived Health Outcomes of Recreation Scale (PHORS).

Methods: The research sample consisted of a total of 1984 adults who exercise, including 864 women and 1120 men during 2021- 2022 in Antalya City, Turkey. The MI of the PHORS was tested by multigroup confirmatory factor analyses, which test the invariance of the covariance structures within the scope of structural equation modelling. Invariance tests were gradually conducted for the implicit variables in the model, CFI (comparative fit index criteria) and AIC (Akaike information criterion) were inquired between structural invariance, where no restriction was applied on the analyses and the other invariance tests (metric invariance, scalar invariance and string invariance respectively) where more restraints are applied.

Results: The study yielded evidence showing that the measurement model defined for the factor structure of the scale provided measurement invariance by gender. ΔCFI values were ≤0.010 in all subscales for metric and scalar invariance.

Conclusion: The items of PHORS represented the same psychological structure, different groups responded to the items in the same way, the constant values in regression equations generated for the items in regression equations were equal/invariable between the groups.

Keywords: Perceived health outcomes; Measurement invariance; Gender

Introduction

Researchers developed concepts and conducted studies for gathering evidence concerning the positive evaluations of leisure experience. Among them, leisure benefit is a broad concept including

physiological, psychological and social aspects (1). Leisure benefit is defined as a subjective experience helping individuals improve their physical and mental health (2). According to the Kao's (3) three-factor model, leisure benefit is explained in three dimensions; balanced life experience, healthy lifestyle, and improved life quality. Among these dimensions, the health aspect is studied very often in leisure literature especially in studies related to leisure-time physical activity. In order to measure the perceived health benefits of leisure participation, PHORS was developed by Gómez et al. (4). The scale was based on Driver's conceptualization of leisure benefits in three dimensions; providing an improved situation, prevention of an undesired situation and realization of a satisfying psychological experience (5). A study on perceived health outcomes of park visitors was conducted in Oklahoma and recognition of psychological experience subscale achieved the highest scores, followed by improvement of health condition (6).

Some of the studies on perceived health benefits in leisure are causal-comparative studies focusing on differences among groups. These studies focus on determining whether there is a significant difference between perceived health outcomes scores based on variables such as gender, age, marital status, ethnicity, sexual orientation, activity type (active-passive/ indoor-outdoor), membership, region etc. (7-13). Some of the studies did not find a significant difference between the sub-dimensions of PHORS based on gender (7-11), marital status (8-10), sexual orientation and age (7), ethnicity (8), membership (9-13), activity type (9) and regional (13). Others reveal that there is a significant difference in PHORS scores based on income level and activity type (12), gender and age (13). However, such comparisons lose their validity unless evidence showing that the groups are equivalent on the measured feature is collected (14). In studies comparing different groups with the same measurement tool, if there is no evidence of MI, it is not possible to make scientific inferences about the results of these studies (15).

Similar to other attitude scales commonly utilized in leisure area, comparisons according to demographic variables especially gender are practiced very often. To find out the real effect of gender variable on perceived health outcomes, the differences resulting from the measurement scale itself should be neutralized by obtaining MI condition. Therefore, we aimed to provide the MI condition according to gender variable for providing more reliable results for future studies.

Methods

Participants and Procedure

The study population consisted of 52.750 individuals registered to sport centers in Antalya City, Turkey. Data collection started at the 2nd quarter of year 2021 and finished at the 1st quarter of year 2022. Overall, 1932 individuals were contacted for the research and 1801 (93.21%) individuals gave feedback. 47.19% of the participants were women (n=850, Mage=29.75, SD=10.18) and 52.80% were men (n=951, M_{age}=28.38, SD=9.21). 18-74 years old adults who participate in exercise at least daily 45 minutes and weekly 3 days were included in the study. Individuals that do not participate in exercise regularly and who filled the forms incomplete (n=131) were excluded from the study. In order to reach the target population, data were collected from two different provinces. Convenience sampling technique, one of the non-probability-sampling methods, was used in the research. Voluntary adults who exercise regularly were included in the study.

Ethical approval was taken from Akdeniz University Health Sciences Scientific Research and Publication Ethical Committee (Document number: 44863).

Data Collection Tools

In the study, the PHORS, developed by Gómez et al. (4) and adapted to Turkish by Yerlisu-Lapa et al. (16), was used. The scale is a 7-point Likert scale consisting of 3 sub-dimensions and 16 items. In the adaptation study of the scale, expert opinion was consulted in order to ensure con-

struct validity, and the entire scope of the relevant construct was discussed (17). The results of the exploratory factor analysis (EFA) of the scale demonstrate that the explained variance is 66.34%. In the confirmatory factor analysis (CFA) of the scale, the fit indices obtained after the modification between the 11th and 12th questions are as follows: χ^2/df=1.56; GFI=0.91; NFI=0.94; TLI=0.96; CFI=0.97; RMSEA=0.063; SRMR=0.059. For convergent validity, it was expected that CR>AVE; AVE>0.5 and for ensuring divergent validity, it was expected that MSV<AVE; ASV<MSV; the square root of AVE>inter-factor correlation and the CR value>0.70 for all sub-dimensions (18,19). CR>0.70 and CR>AVE conditions were met in all sub-dimensions of the scale. The AVE value varies between 0.41 and 0.47. However, AVE value being slightly below 0.50 is considered to be acceptable as the other reliability criteria were met (20). All these empirical proofs provide important evidence for the validity and reliability of the measurement tool.

Data Analysis

In order to provide evidence for the structural validity, MI according to gender was studied. Multi-group confirmatory factor analysis (MG-CFA) which is one of the structural equation modelling methods was used to test the invariance between gender and maximum likelihood estimation method was preferred. During the first phase of the data analysis, basic assumptions were tested in order to improve the quality of the study (21). The requirements of normality (For all sub-dimensions=Skewness and kurtosis varies between -1.5/+1.5. It is seen that the mode, median and arithmetic mean are close to each other) (22); multivariate normality and linearity distribution were fulfilled. In order to determine the sample size based on mean, margin of error was set at d=0.05, standard deviation=0.50 and confidence level (1-α)=0.95. According to this formula, n=381 was sufficient for the sample (23) although non-probability sampling method was utilized in this study, a large sample size (n=1894) was reached in order to increase the level of reliability and validity for the research results (24). Variance Inflation Factor (VIF) and Condition Index (CI) indexes were examined (25). VIF<10 (26); CI<10 (27) and multicollinearity problem does not exist. In order to provide methodological qualifications of PHORS scale COSMIN standards were checked (28,29). During the second phase of the study, the fit indices used as criteria for CFA are: $0\leq\chi^2/df\leq2$ (30); $0.95\leq$CFI≤1.00; $0.95\leq$NFI≤1.00 (31-33); $0.00\leq$RMSEA≤0.05; $0.00\leq$SRMR≤0.05 (34).

During the third phase of the analysis the most widely used MG-CFA was realized according to 8 guiding principles of Cross-cultural validity\MI in the COSMIN checklist (35). MG-CFA starts from the model in which no constraints are introduced, and the equivalence of parameters between groups is examined up to the most limited model, and this process (Fig.1) continues gradually (36).

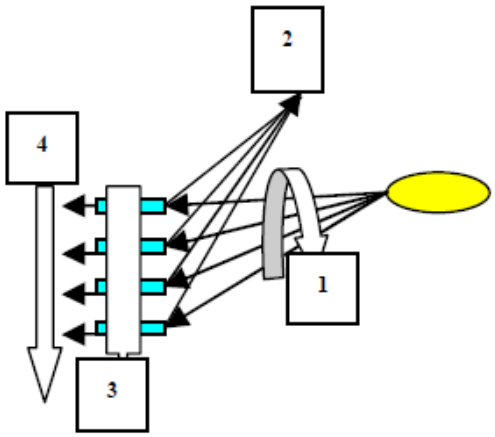

Fig. 1: MI hierarchy (37)

While Structural invariance is based on the hypothesis that the model structure is the same for subgroups (36,38); metric invariance is subgroups respond in the same way to scale items (39); scalar invariance is both item tendencies and item constants are equal between the groups (40); and the strict invariance is the scores of the error variances of the items are invariant for the groups (41).

MG-CFA is a useful method to test the equivalence of covariance structures (42). The CFI val-

ue and RMSEA values were taken into account instead of the GFI value which is affected by the sample size (43,44). There are differences of opinion regarding which fit indices to use in the reporting process of MI (44,45). To examine the results of MI, ΔCFI fit index is used as it is more convenient to explain the relationship between implicit scores and observed scores (46). ΔCFI was found to be more proper than $\Delta\chi^2$ value, which was strict and sensitive to sample size (43,44,47-49). Akaike information criterion (AIC) is another criterion which is the determinant of goodness of fit of a statistical model (50,51) and the lowest AIC value is preferred for model comparisons (52).

Results

Table 1 shows the results of the CFA analysis by gender. Women have χ^2/df=11.49 and RMSEA=0.11, 90% CI[0.11-0.12], while men have χ^2/df=12.29 and RMSEA=0.11, 90% CI[.10-0.11].

Table 1: The equality of covariance matrices of women and men

Group	*χ^2*	*df*	*χ^2/df*	*p<0.01*	*CFI*	*NFI*	*NNFI*	*RMSEA (90% CI)*	*SRMR*
Women	1160.84	101	11.49	0.00	0.98	0.97	0.97	0.11(0.11-0.12)	0.035
Men	1241.60	101	12.29	0.00	0.97	0.97	0.97	0.11(0.10-0.11)	0.047
Total	2274.44	101	22.51	0.00	0.98	0.98	0.97	0.11(0.091-0.098)	0.040

It was determined that χ^2/df and RMSEA values did not have acceptable fit indices for both groups. However, it was observed that CFI, NFI, NNFI and SRMR values have perfect fit criteria for both women and men. For women CFI=0.98; NFI=0.97 and NNFI=0.97 while SRMR=0.035. For men, these values were CFI=0.97; NFI=0.97 and NNFI=0.97 and SRMR=0.0047. CFI, NFI, NNFI values of ≥0.95 (31-33) and SRMR values of ≤0.05 indicate a perfect fit. The fact that more than one fit indices indicate a perfect fit can be interpreted as that CFA was confirmed in both groups.

Table 2 presents the distribution of standardized factor loads by gender. While the standardized factor loads of the PHORS scale for women ranged from 0.81 to 0.92, these values ranged from 0.81 to 0.90 for men. Standardized factor loads for all groups were between 0.81 and 0.91. It was observed that all t values were significant at the P<0.01 level. Table 3 shows the fit indices of the MI stages of the three sub-dimensions of the PHORS. While evaluating MI, the differences between fit indices of structural invariance, which is the first stage of MI, and other limited models (ΔCFI) were compared and AIC values were examined to provide additional empirical evidence. When the results of the MGCFA analysis were examined, the fit indices for both subgroups (gender) for PSYC, PREV and IMPV sub-dimensions revealed that structural invariance is provided. χ^2/df and RMSEA values are high in all sub-dimensions. Taken into account that the χ^2 value is affected by the sample size, other fit indices were evaluated. Except for χ^2/df and RMSEA values, other fit indices seem to indicate perfect fit. In this context, it is possible to say that structural invariance is ensured. In order to decide which variables provide metric invariance, the factor loadings between groups were also limited to be the same. To comment on metric invariance, MGCFA results, ΔCFI and AIC values were interpreted. In scalar invariance, in addition to metric invariance, similar items were assumed to be equal in men and women.

Table 2: Standardized factor loads for PHORS per gender

Items	*Women*	*t-value for women*	*Men*	*t-value for men*	*All group*	*t-value for all groups*
PSYC						
Item1	0.81	19.06	0.85	18.89	0.83	27.00
Item2	0.88	17.66	0.88	17.78	0.88	25.06
Item3	0.91	16.29	0.88	17.94	0.90	24.26
Item4	0.91	16.80	0.87	18.17	0.89	24.73
Item5	0.86	18.30	0.81	19.59	0.84	26.75
Item6	0.86	18.16	0.80	19.72	0.84	26.78
Item7	0.86	18.13	0.81	19.68	0.84	26.73
PREV						
Item8	0.87	17.35	0.82	18.59	0.84	25.43
Item9	0.85	17.98	0.77	19.43	0.81	26.53
Item10	0.92	14.97	0.89	15.41	0.91	21.43
Item11	0.88	17.19	0.85	17.72	0.86	24.66
Item12	0.89	16.52	0.84	17.93	0.87	24.47
IMPV						
Item13	0.91	15.72	0.88	16.30	0.89	22.59
Item14	0.91	15.53	0.90	16.11	0.90	21.57
Item15	0.92	15.25	0.85	17.74	0.88	23.54
Item16	0.89	16.50	0.84	17.94	0.87	24.42

All t values were significant at the $p<0.01$ level.

For this reason, factor loads and regression coefficients for women and men were limited, and error variances were released. Considering the MGCFA results, ΔCFI and AIC values, it was observed that ΔCFI≤0.010 and there was no sharp increase in AIC value. Strict invariance was provided for PSYC and PREV sub-dimensions (ΔCFI≤.0.010) but ΔCFI value for IMPV sub-dimension was ≥0.010 and because of a sharp increase in AIC value, strict invariance was not achieved.

Table 3: Fit statistics of MI stages

	Gender	*χ^2*	*df*	*NFI*	*NNFI*	*CFI*	*SRMR*	*RMSEA (90 %CI)*	*ΔCFI*	*AIC*	*Decision*
PSYC	Structural	1057.73	28	0.95	0.93	0.96	0.039	0.20(0.19-0.21)		1113.73	H_0 Accept
	Metric	1091.96	35	0.95	0.95	0.95	0.085	0.18(0.17-0.19)	-0.01	1113.96	H_0 Accept
	Scalar	1294.09	48	0.94	0.95	0.95	0.10	0.17(0.16-0.18)	-0.01	1338.09	H_0 Accept
	Strict	1348.85	55	0.94	0.96	0.95	0.11	0.16(0.15-0.17)	-0.01	1378.85	H_0 Accept
PREV	Structural	497.93	10	0.95	0.91	0.95	0.034	0.23(0.22-0.25)		573.93	H_0 Accept
	Metric	503.53	15	0.95	0.94	0.95	0.056	0.19(0.18-0.20)	0	533.53	H_0 Accept
	Scalar	584.50	24	0.94	0.95	0.94	0.076	0.16(0.15-0.17)	-0.01	616.50	H_0 Accept
	Strict	670.92	29	0.93	0.96	0.94	0.075	0.15(0.15-0.17)	-0.01	692.92	H_0 Accept
IMPV	Structural	93.74	4	0.99	0.96	0.99	0.012	0.15(0.13-0.19)		125.74	H_0 Accept
	Metric	98.92	8	0.99	0.98	0.99	0.049	0.11(0.09-0.13)	0	122.92	H_0 Accept
	Scalar	129.73	15	0.98	0.99	0.98	0.054	0.09(0.07-0.11)	-0.01	155.73	H_0 Accept
	Strict	202.27	19	0.97	0.98	0.97	0.049	0.10(0.09-0.12)	-0.02	220.27	H_0 Reject

Discussion

In this study, we aimed to test the MI in order to test the differences between the groups using PHORS stem from the measurement tool. We believe that MI provides a significant contribution as it allows discussion of differences between groups.

The results of this study, which evaluated the MI of the three sub-dimensions of the PHORS - PSYC, PREV and IMPV-according to gender, reveal that structural invariance was achieved in all three sub-dimensions. Achieving structural invariance shows that the structure is invariant according to gender, that is, the latent variables are similar in men and women. Achieving structural invariance can be interpreted that the subgroups (women and men) have the same conceptual point of view for all three sub-dimensions when responding to scale questions (53). As to metric invariance condition, it can be stated that it is ensured for PSYC, PREV, and IMPV of the PHORS, and that men and women participating in physical activity respond to the scale items in the same way. It is possible to say that it will be significant to compare the gender -based scores of the observed items by achieving metric invariance (54).

During the development of the original form of the scale, the construct validity of perceived health outcomes derived by individuals trekking in three different park ours was tested. The results presented that, individuals trekking in different park ours had similar conceptual point of view (structural invariance) and they answered the items in the same way (metric invariance). Therefore, the comparison of scores obtained from these groups is meaningful (4). In many national and international studies conducted with PHORS, differences between groups based on gender were discussed (5,7,10,13,55,56). In this way, it became possible to discuss on the significance of results.

The results regarding scalar invariance reveal that scalar invariance is also provided for all three sub-dimensions. Achieving scalar invariance means that differences on items can be compared according to gender. When the strict invariance conditions are examined, it is seen that strict invariance is achieved in the context of PSYC and PREV sub-dimensions, yet it is not achieved for the IMPV sub-dimension. This shows that the error terms related to the items of the PSYC and PREV sub-dimensions are invariant according to gender.

The main reason for the measurement invariance to be made in the adult population in this study is: Comparisons made according to the gender variable in the literature mostly deal with the adult population (10,56-59).

Limitations and Future Research

The results obtained from this study are limited to the adults participating in the study. The limitation of our research is that the study group consists of adult individuals who regularly participate in physical activity in fitness centers in Antalya and Istanbul. For this reason, we think that it is appropriate to reconsider the invariance according to gender to be made with PHORS with individuals living in different cities. In some cases, it is stated that all components (items) of the measurement tool used cannot ensure invariance (49,60). In the comparisons to be made, it may be recommended to conduct a partial invariance study for the groups in which invariance is investigated if intergroup invariance is not ensured in some components of the PHORS. It is considered to be important to strengthen the PHORS structure by conducting MI studies according to different groups in terms of the type of participation in leisure activities, age, marital status and region of individuals who exercise in their leisure time, based on the studies available in the literature, apart from the recommendations based on the research results.

Conclusion

The psychometric qualities obtained from the measurement model, which consists of items that reveal the perceived health outcomes of the

adults participating in physical activity, can be generalized according to gender. In other words, PHORS does not entail any gender bias and proves that it provides valid and reliable results in determining the characteristics of adult individuals participating in physical activity regarding this structure. The measurement model measures women and men in the same way, and there is no problem with comparing the scores obtained from this measurement tool.

Journalism Ethics considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Acknowledgements

No financial support was received for this study.

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. Xu H, Yuan M, Li J (2019). Exploring the relationship between cycling motivation, leisure benefits and well-being. *Int Rev Spat Plan Sustain Dev*, 7(2): 157-71.
2. Eskiler E, Yıldız Y, Ayhan C (2019). The effect of leisure benefits on leisure satisfaction: Extreme sports. *Turkish Journal of Sport and Exercise*, 21(1): 16-20.
3. Kao, CH (1995). A three-factor model of leisure benefits. *Journal of Outdoor Recreation Study*, 8(1): 67-78.
4. Gómez E, Hill E, Zhu X, et al (2016). Perceived health outcomes of recreation scale (PHORS): reliability, validity and invariance. *Meas Phys Educ Exerc Sci*, 20(1): 27-37.
5. Gómez E, Hill, E (2016). First landing state park: Participation patterns and perceived health outcomes of recreation at an urban-proximate park. *J Park Recreat Admi*, 34(1): 68-83.
6. Wu I-C (N), Liu, H-L(S), Cox MA (2018). An exploration of perceived health outcomes of state park visitors in Oklahoma. *Recreation, Parks, and Tourism in Public Health*, 2, 57-64.
7. Hill E, Morgan T, Hooper B (2018). Hosting triathlons on a college campus: Perceived health outcomes and satisfaction. *Journal of Outdoor Recreation, Education, and Leadership*, 10(3): 256-8.
8. McIntosh T, Hill E, Morgan T (2019). Hosting triathlons on a college campus: Perceived health outcomes and values. *Journal of Outdoor Recreation, Education, and Leadership*, 11(3): 262-4.
9. Berry C, Ahl P, Hill E, et al (2019). Examining the health outcomes of college climbers: Applying the perceived health outcomes of recreation scale. *Journal of Outdoor Recreation, Education, and Leadership*, 11(3): 258–61.
10. Elçi G, Doğan M, Gürbüz B (2019). Investigation the level of individuals' perceived health outcomes of recreation and life satisfaction. *International Journal of Sport Exercise and Training Sciences*, 5(3): 93-106.
11. Serdar E (2021). Perceived health outcomes of recreation, exercise dependence, and life satisfaction. *Balt J Health Phys Act*, 13(5): 43-50.
12. Temel AS, Tukel Y (2021). Examining the health outcomes and happiness levels that result from engaging in physical recreation: A study on university students. *International Journal of Research in Education and Science*, 7(2): 545-61.
13. Hill E, Gómez E (2020). Perceived health outcomes of mountain bikers: A national demographic inquiry. *J Park Recreat Admi*, 38(2): 31-44.
14. Korkmaz M, Somer O, Güngör D (2013). Measurement equivalence across gender with mean and covariance structure of five factor personality inventory for adolescent sample. *Education and Science*, 38(170): 121-34.
15. Mark BA, Wan TT (2005). Testing measurement equivalence in a patient satisfaction instrument. *West J Nurs Res,* 27(6): 772-787.
16. Yerlisu Lapa T, Serdar E, Tercan Kaas E, et al (2020). Psychometric properties of perceived health outcomes of recreation scale-Turkish version. *Hacettepe Journal of Sport Sciences,* 31(2): 83-95.

17. Haladyna TM, Rodriguez MC (2013). *Developing and validating test items*. Routledge: Taylor & Francis. London.
18. Hair JF, Black W, Babin BJ, et al (2006). *Multivariate data analysis*. 6th ed. Upper Saddle River: Prentice Hall. New Jersey.
19. Raykov T (1998). Coefficient Alpha and composite reliability with interrelated nonhomogeneous items. *Appl Psychol Meas*, 22(4): 375-85.
20. Fornell C, Larcker DF (1981). Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res*, 18(1): 39-50.
21. Mertler CA, Vannatta RA (2016). *Advanced and multivariate statistical methods: Practical application and interpretation*. Routledge: Taylor & Francis. London.
22. Rosenthal R, Rosnow RL (2008). *Essentials of behavioral research: Methods and data analysis*. 3rd ed. McGraw-Hill. New York.
23. Büyüköztürk Ş, Kılıç-Çakmak E, Akgün ÖE et al (2014). *Scientific Research Methods*. 17th ed. Pegem Akademi: Ankara
24. Faber J, Fonseca LM (2014). How sample size influences research outcomes. *Dental Press J Orthod*, 19(4): 27-9.
25. Salkind NJ (2006). *Encyclopedia of measurement and statistics*. SAGE Publications: Road Thousand Oaks. California.
26. Craney TA, Surles JG (2002). Model-dependent variance inflation factor cut off values. *Qual Eng*, 14(3): 391-403.
27. Paulson DS (2006). *Handbook of regression and modeling: Applications for the clinical and pharmaceutical industries*. Chapman and Hall/CRC. New York.
28. De Boer MR, Moll AC, De Vet HCW, et al (2004). Psychometric properties of vision-related quality of life questionnaires: A systematic review. *Ophthalmic Physiol Opt*, 24(4): 257-73.
29. Veenhof C, Bijlsma JW, Van den Ende CH, et al (2006). Psychometric evaluation of osteoarthritis questionnaires: A systematic review of the literature. *Arthritis Rheum*, 55(3): 480-92.
30. Kline RB (2005). *Principles and practice of structural equation modeling*. 2nd ed. Guilford Press. New York.
31. Baumgartner H, Homburg C (1996). Applications of structural equation modeling in marketing and consumer research: A review. *Int J Res Mark*, 13(2): 139-61.
32. Bentler PM, Bonett DG (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychol Bull*, 88(3): 588-606.
33. Marsh HW, Hau K-T, Artelt C (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *Int J Test*, 6(4): 311-60.
34. Browne MW, Cudeck R (1992). Alternative ways of assessing model fit. *Social Methods Res*, 21(2): 230-58.
35. Prinsen CAC, Mokkink LB, Bouter LM, et al (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*, 27(5): 1147-57.
36. Horn JL, McArdle JJ, Mason R (1983). When is invariance not invariant: A practical scientist's look at the ethereal concept of factor invariance. *The Southern Psychologist*, 1(4): 179-88.
37. Başusta B (2010). Measurement invariance. *J Meas Eval Educ Psychol*, 1(2): 58-64.
38. Chen FF, Sousa KH, West SG (2005). Testing measurement invariance of second-order factor models. *Struct Equ Modeling*, 12(3): 471-92.
39. Steenkamp J-BEM, Baumgartner H (1998). Assessing measurement invariance in cross-national consumer research. *J Consum Res*, 25(1): 78-90.
40. Tucker KL, Ozer DJ, Lyubomirsky S, et al (2006). Testing for measurement invariance in the satisfaction with life scale: A comparison of Russians and North Americans. *Soc Indic Res*, 78(2), 341-60.
41. Önen E. Examination of measurement invariance with structural equation modelling techniques [PhD thesis]. Ankara University, Türkiye; 2009.
42. Dimitrov D (2010). Testing for factorial invariance in the context of construct validation. *Meas Eval Couns Dev*, 43(2): 121-49.
43. Cheung GW, Rensvold RB (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *J Cross-Cult Psychol*, 31(2): 187-212.
44. Cheung GW, Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measure-

ment invariance. *Struct Equ Modeling*, 9(2): 233-55.
45. Chen FF (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Struct Equ Modeling*, 14(3): 464-504.
46. Wu AD, Li Z, Zumbo BD (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Pract Assess Res Evaluation*, 12(3): 1-26.
47. Quintana SM, Maxwell SE (1999). Implications of recent developments in structural equation modeling for counseling psychology. *Couns Psychol*, 27(4): 485-527.
48. Meade AW, Michels LC, Lautenschlager GJ (2007). Are internet and paper-and-pencil personality tests truly comparable? An experimental design measurement invariance study. *Organ Res Methods*, 10(2): 322-45.
49. Vandenberg RJ, Lance CE (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ Res Methods*, 3(1): 4-70.
50. Lin YJ, Tzeng WN (2009). Modelling the growth of Japanese eel Anguilla Japonica in the lower reach of the Kao-Ping River, Southern Taiwan: An information theory approach. *J Fish Biol*, 75(1): 100-12.
51. Zhu L, Li L, Liang Z (2009). Comparison of six statistical approaches in the selection of appropriate fish growth models. *Chin J Ocean Limnol*, 27: 457.
52. Burnham KP, Anderson DR (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer-Verlag. New York.
53. Vandenberg RJ, Lance CE (1998). A summary of the issues underlying measurement equivalence and their implications for interpreting group differences. *In 1998 Research Methods Forum*, (3): 1-10.
54. Bryne BM, Watkins D (2003). The issue of measurement invariance revisited. *J Cross-Cult Psychol*, 34(2): 155-75.
55. Hill E, Ahl P, Gabriele C, et al (2018). Perceived health outcomes of college climbers: Exploring why they climb. *Journal of Outdoor Recreation, Education, and Leadership*, 10(3): 259-62.
56. Beşikçi T, Emir E, Özdemir E, et al (2021). Investigation of psychological resilience levels of individuals and perceived health outcomes of recreation during the covid-19 pandemic process. *Journal of Sport Sciences Researches*, 6(2): 447-58.
57. Serdar E. Relationship among leisure attitudes, leisure satisfaction and perceived health outcomes in recreation [PhD thesis]. İstanbul University-Cerrahpasa, Türkiye; 2020.
58. Kayapınar K. Investigation of perceived health outcomes, perceived leisure boredom and life satisfaction levels of individuals participating in recreational tennis [Master Thesis]. Ankara University, Türkiye; 2020.
59. Bekar Ö. Examining the relationship between recreational exercise motivation and perceived health outcomes from individuals who exercise regularly [Master Thesis]. Akdeniz University, Türkiye; 2019.
60. Cheung GW (2007). Testing equivalence in the structure, means, and variances of higher-order constructs with structural equation modeling. *Organ Res Methods*, 11(3): 593-613.