Psychology of Sport and Exercise 14 (2013) 709–718

Contents lists available at SciVerse ScienceDirect

Psychology of Sport and Exercise

journal homepage: www.elsevier.com/locate/psychsport

Coaching competency and (exploratory) structural equation modeling: A substantive-methodological synergy

Nicholas D. Myers

Department of Educational and Psychological Studies, University of Miami, Merrick Building 311E, Coral Gables, FL 33124-2040, United States

A R T I C L E I N F O

Article history:
Received 9 January 2013
Received in revised form
24 April 2013
Accepted 25 April 2013
Available online 9 May 2013

Keywords:
athlete's evaluations
Coaching effectiveness
Geomin
Target
Categorical data

A B S T R A C T

Objectives: The purpose of this manuscript was to provide a substantive-methodological synergy of potential importance to future research in the psychology of sport and exercise.
Design: The substantive focus was the emerging role for, and particularly the measurement of, athletes' evaluations of their coach's competency within conceptual models of effective coaching. The methodological focus was exploratory structural equation modeling (ESEM), a methodology that integrates the advantages of exploratory factor analysis and confirmatory factor analysis (CFA) within the general structural equation model.
Methods: The synergy was a demonstration of when a new and flexible methodological framework, ESEM, may be preferred as compared to a more familiar and restrictive methodological framework, CFA, by reanalyzing existing data.
Results: ESEM analysis on extant datasets suggested that for responses to the Athletes' Perceptions of Coaching Competency Scale II – High School Teams (APCCS II-HST), a CFA model based on the relevant literature plus one post hoc modification, offered a viable alternative to a more complex ESEM model. For responses to the Coaching Competency Scale (CCS), a CFA model based on the relevant literature did not offer a viable alternative to a more complex ESEM model.
Conclusions: The ESEM framework should be strongly considered in subsequent studies validity studies - for new and/or existing instruments in the psychology of sport and exercise. A key consideration for deciding between ESEM and the accompanying rotation criterion and CFA in future validity studies should be level of a priori measurement theory.

Factor analysis has been closely linked with investigations of construct validity in psychology for several decades (Nunnally, 1978). Investigations of construct validity (e.g., scale development) in exercise and sport have frequently occurred in studies where only factor analytic measurement model(s), exploratory and/or confirmatory, were specified – typically guided by important, yet incomplete, a priori substantive measurement theory (Myers, Ahn, & Jin, 2011a). Incomplete substantive measurement theory often manifests as model error and offers a possible explanation as to why the majority of construct validity studies in psychology fail the test of exact fit under a strictly confirmatory approach (Jackson, Gillaspy, & Purc-Stephenson, 2009).[1] Construct validity studies that incorporate an exploratory methodological approach typically reduce model error, as compared to a confirmatory approach, but may fail to adequately incorporate a priori substantive measurement theory (Asparouhov & Muthén, 2009).

What constitutes a sufficient degree of a priori substantive measurement theory, and therefore should lead to a preference for a more restrictive confirmatory approach versus a more flexible exploratory approach, is an emerging area of methodological research of potential importance to substantive research in the psychology of sport and exercise. Simulation research suggests that a key consideration for deciding between exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) and confirmatory factor analysis (CFA; Jöreskog, 1969) in future validity studies should be level of a priori substantive measurement theory (Myers, Ahn, & Jin, 2011b). In the Myers et al. simulation study model-data fit was manipulated (as a proxy for level of a priori substantive measurement theory) with small but non-zero values in both the pattern coefficient matrix and in off-diagonals of the residual covariance matrix. Under exact (with ESEM having errors of inclusion) and close (with both ESEM and CFA having errors of

E-mail address: nmyers@miami.edu.

[1] Other possible explanations exist for failing the likelihood ratio test of exact fit (e.g., Type I error). The degree to which failing the exact fit test is regarded as a meaningful occurrence (e.g., an imperfect model may be useful) is an ongoing debate (e.g., Vernon & Eysenck, 2007).

1469-0292/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.psychsport.2013.04.008

exclusion) model-data fit, CFA generally performed at least as well as ESEM in terms of parameter bias, and was cautiously recommended when a priori substantive measurement theory can be viewed as at least close. Under only approximate model-data fit, on balance, ESEM performed at least as well as CFA and was cautiously recommended when a priori substantive measurement theory can be viewed as only approximate.

This invited manuscript demonstrates by way of an extant conceptual model (i.e., coaching competency) when a new and flexible methodological framework, ESEM, may be preferred as compared to a more familiar and restrictive methodological framework, CFA, within a substantive-methodological synergy format (Marsh & Hau, 2007). For spatial reasons this manuscript focuses primarily on the synergy by way of relevant empirical examples. Fuller reviews of both the relevant substance (Myers & Jin, 2013) and technical aspects of the broader methodology (e.g., Asparouhov & Muthén, 2009) are available elsewhere. From this point forward the acronym, (E)SEM, is used when referring to both ESEM and CFA simultaneously.

Coaching competency: the substance

The Coaching Competency Scale (CCS; Myers, Feltz, Maier, Wolfe, & Reckase, 2006) was derived via minor changes to the Coaching Efficacy Scale (CES; Feltz, Chase, Moritz, & Sullivan, 1999). The CES was developed to measure a *coach's belief* in his or her ability to influence the learning and performance of his or her athletes. The specific factors measured—instructional technique, motivation, character building, and game strategy—purposely overlap with key expected competency domains articulated in the *National Standards for Athletic Coaches* (National Association for Sport and Physical Education, 2006) and are congruent with self-efficacy theory (Bandura, 1997). According to Horn's (2002) model of coaching effectiveness, however, *why* a coach's beliefs (e.g., coaching efficacy) are related to athletes' self-perceptions and performance is due to the influence that these beliefs exert on a coach's behavior. But, the influence that a coach's behavior exerts on athletes' self-perceptions, motivation, and performances is mediated, at least in part, by athletes' evaluations of their coach's behavior. The purpose of the CCS, therefore, was to measure *athletes' evaluations* of their head coach's ability to affect the learning and performance of athletes. There is evidence that measures derived from both the CCS relate to theoretically relevant variables (e.g., Bosselut, Heuzé, Eys, Fontayne, & Sarrazin, 2012).

The measurement model for the CCS posited that four latent variables covary and influence responses to 24 items. From this point forward the terms competency and efficacy were generally not included alongside any particular dimension of coaching competency. *Motivation* was defined as athletes' evaluations of their head coach's ability to affect the psychological mood and skills of athletes. *Game strategy* was defined as athletes' evaluations of their head coach's ability to lead during competition. *Technique* was defined as athletes' evaluations of their head coach's instructional and diagnostic abilities. *Character building* was defined as athletes' evaluations of their head coach's ability to influence the personal development and positive attitude toward sport in their athletes. Myers, Feltz, et al. (2006) referred to this multidimensional construct as *coaching competency*. High school and lower division collegiate athletes of team sports comprise populations for which the CCS was intended.

Model-data fit within the CFA framework for the CCS (as well as the CES - see Myers & Jin, 2013) has generally not met heuristic values for close fit (e.g., Hu & Bentler, 1999). Myers, Feltz, et al. (2006) reported the following model-data fit for non-division I collegiate athletes ($N = 585$) clustered within men's ($g = 8$) and women's ($g = 24$) teams: $\chi^2(246) = 1266$, CFI = 0.91, Tucker-Lewis index (TLI) = 0.90, and RMSEA = 0.09. Post hoc modifications by Myers, Feltz, et al. included allowing a motivation item to also indicate game strategy and freeing the covariance between two pairs of residual variances (see Fig. 1). Interfactor correlations in the Myers, Feltz et al. study ranged from $r_{\text{technique, character building}} = 0.80$ to $r_{\text{game strategy, technique}} = 0.92$. Given the high interfactor correlations, Myers, Feltz, et al. explored models with fewer latent variables, all of which yielded statistically significant worse fit.

Responses to the CCS have a history of dependence based on athletes nested within teams. Myers, Feltz, et al. (2006) reported item-level intraclass correlation coefficients (ICC) ranging from 0.22 to 0.44, $M = 0.32$, $SD = 0.05$. Multilevel CFA is an appropriate methodology when data violate the assumption of independence (Muthén, 1994). Simulation research has indicated, however, that a relatively large number of groups (~ 100) may be necessary for optimal estimation, particularly at the between-groups level (Hox & Maas, 2001). When group-level sample size is not large, imposing a single-level model on the within-groups covariance matrix, $\mathbf{S_W}$, as opposed to the total covariance matrix, $\mathbf{S_T}$, controls for probable biases in fitting single-level models to multilevel data (Julian, 2001). Myers, Feltz, et al. modeled the $\mathbf{S_W}$.

The Coaching Efficacy Scale II-HST (CES II-HST) was derived via major changes to the CES as detailed in Myers, Feltz, Chase, Reckase, and Hancock (2008) and Myers, Feltz, and Wolfe (2008). The Athletes' Perceptions of Coaching Competency Scale II – High School Teams (APCCS II-HST) was derived via minor changes to the CES II-HST as detailed in Myers, Chase, Beauchamp, and Jackson (2010). Thus, moving from either CES to the CES II-HST, or the CCS to the APCCS II-HST, resulted in non-equivalent measurement models. There is evidence that measures derived from the APCCS II-HST relate to theoretically relevant variables (Myers, Beauchamp, & Chase, 2011) in a doubly latent multilevel model (Marsh et al., 2012).

The measurement model for the APCCS II-HST occurred at two-levels in Myers et al. (2010) because there was strong evidence for dependence due to the clustering of athletes ($N = 748$) within teams ($G = 74$). Specifically, ICCs ranged from 0.18 to 0.35, $M = 0.29$, $SD = 0.05$. There was evidence for close fit for the multilevel measurement model: $\chi^2_R(201)$ ranged from 375, $p < 0.001$, to 405, $p < 0.001$, RMSEA = 0.04, CFI = 0.99, TLI = 0.99, $\text{SRMR}_{\text{within}} = 0.03$, and $\text{SRMR}_{\text{between}} = 0.04$.[2] Evidence also was provided for both the stability of all latent variables and for the reliability of composite scores. There is evidence, at least in regard to a crude comparison of model-data fit, that the measurement model underlying responses to the APCCS II-HST may be better understood (by researchers) than the measurement model underlying responses to the CCS. While Myers et al. took a multilevel approach, the focus of this manuscript was on the athlete-level (Level 1) model only.

The measurement model for the APCCS II-HST specified that five dimensions of coaching competency covary and influence responses to 17 indicators. *Motivation* (M) was measured by four items and was defined as athletes' perceptions of their head coach's ability to affect the psychological mood and psychological skills of her/his athletes. *Game strategy* (GS) was measured by five items and was defined as athletes' perceptions of their head coach's ability to lead his/her athletes during competition. *Technique* (T) was measured by four items and was defined as athletes' perceptions of their head coach's ability to utilize her/his instructional and

[2] Many of the model-data fit indices reported in Myers et al. (2010) summarized model-data fit across levels. Future research in this area should consider providing level-specific indices put forth by Ryu and West (2009): $\chi^2_{PS\_W}$, $\chi^2_{PS\_B}$, RMSEA_{PS_W}, $\text{RMSEA}_{PS\_B}\text{CFI}_{PS\_W}$, and CFI_{PS_B}.