Running Head: EMERGENT QUANTITATIVE ANALYSES

A Review of some Emergent Quantitative Analyses in Sport and Exercise Psychology

This is an Author's Original Manuscript of an article published by Taylor & Francis in International Review of Sport & Exercise Psychology on 08 /05/2017, available online at http://www.tandfonline.com/10.1080/1750984X.2017.1317356

Abstract

The purpose of this manuscript is to provide a review of some key quantitative methods that are relevant to contemporary quantitative research in sport and exercise psychology. To achieve this purpose we provide a critical review of four quantitative methods that we believe are emergent in the sport and exercise psychology literature. The first quantitative method reviewed is sample size determination and power estimation in structural equation modelling (e.g., Satorra & Saris, 1985). The second quantitative method reviewed is exploratory structural equation modelling (Asparouhov & Muthén, 2009). The third quantitative method reviewed is mixture modelling (e.g., McLachlan & Peel, 2000). The final quantitative method reviewed is Bayesian structural equation modelling (e.g., Muthén, & Asparouhov, 2012). We begin each review with an overview of the methodology, followed by a summary of one or more related applications in sport and exercise psychology research, and conclude with some ideas for possible future applications in sport and exercise psychology.

Keywords: exploratory structural equation modelling, mixture modelling, Bayesian estimation, sample size determination, power estimation

A Review of some Emergent Quantitative Analyses in Sport and Exercise Psychology

> The domain of quantitative methods is constantly evolving and expanding. This means that there is tremendous pressure on researchers to stay current, both in terms of best practices and improvements in more traditional methods as well as increasingly complex new methods (Hancock, 2016, Description section, para 3).

One of the many and constant challenges academics face is to stay up to date with recent developments in statistical analyses that have implications for measurement and theory in their field. Researchers in sport and exercise psychology (SEP) are not spared of this challenge. Over the last 20 years or so, there has been a considerable expansion in the number of statistical techniques and software available to address questions of substantive and applied importance for the field of SEP. In our anecdotal experience, academics in this field (and most probably in other fields) tend to adopt one of the following three responses to this challenge. Some academics choose (for various reasons) not to keep up with the latest statistical developments and seek collaborators who have statistical expertise to apply these new methods. Other academics (probably the majority) try to keep up with developments in some analytical techniques due to a particular interest (e.g., in scale development). A third group of SEP researchers develop primarily a methodological expertise and reputation by being at the forefront of applying to their field numerous statistical innovations from applied statistics and psychology.

Although the merits of each profile can be debated, this discussion is not of interest for this paper. Instead, in this paper we aim to present in a succinct fashion some recent developments in quantitative analysis by targeting those academics in the first and second profile. We hope that our introduction to a selection of emerging quantitative analyses and a brief overview of their current applications in the SEP literature will trigger the curiosity and intrinsic interest of a greater pool of researchers to learn more about and apply these methods. Resources exist (see citations in the following sections) which provide detailed treatments of these topics, supplemented by software code. Notably, a recent book by Ntoumanis and Myers (2016) demonstrates the applications of these methods in sport and exercise science research.

A review of problematic and emergent quantitative and qualitative methods by Biddle, Markland, Gilbourne, Chatzisarantis, and Sparks (2001) was seminal and highly cited in the SEP field. However, there have been many advances in quantitative methodology since that paper. During the last 10-15 years, journal article contributions in the SEP field have focused on a detailed treatment of one particular statistical technique (e.g., Myers, Martin, Ntoumanis, Celimli, & Bartholomew, 2014, presented exploratory bi-factor analysis; Fitzpatrick, Gareau, Lafontaine, & Gaudreau, 2016, discussed dyadic data analysis using the Actor-Partner Interdependence Model). In this paper we aim to provide a concise update (as far as quantitative analyses are concerned) to Biddle and colleagues seminal paper by presenting four emergent analyses, namely 1) sample size determination and power estimation in structural equation modelling, 2) exploratory structural equation modelling, 3) mixture modelling, and 4) Bayesian structural equation modelling. We begin each section with an overview of the methodology, followed by a brief overview highlighting one or more key applications within SEP, and conclude with some suggestions for future applications. The section on sample size determination deviates from the other sections in that it presents a brief demonstration of the technique. This approach was undertaken in response to Schweizer and Furley (2016) who urged researchers in the sport and exercise field to do better with regard to sample size determination/power.

Papers utilising these four techniques in SEP research have emerged over the last five years or so, but their applications remain relatively sparse. We chose these four quantitative analyses for a variety of reasons, such as having a pragmatic length review as an end-product for a journal article, and because these methods can provide answers to many questions from a broad spectrum of research within SEP. This focus is not to imply that other emerging statistical techniques not covered in this review are ‘inferior’ in any way. Further, our review of each type of analyses is not meant to be exhaustive as our purpose was simply to highlight one or more key SEP examples for readers.

Sample Size Determination and Power Estimation in Structural Equation Modelling

Published applications of structural equation modelling (SEM) have been relatively common in original research within SEP for some time (e.g., Biddle, Markland, Gilbourne, Chatzisarantis, & Sparkes, 2001). Rarely, however, do published applications of SEM in SEP report a power analysis (Myers, Celimli, Martin, & Hancock, 2016). Providing results from a power analysis for an application of SEM is important because doing so ‘…may improve the methodological approach within a particular study and, perhaps more importantly, may positively influence the quality of related studies in the future…’ (Myers, Celimli, Martin, & Hancock, 2016, p. 281). The purpose of this section, therefore, is to review some key approaches to power analysis in SEM that are relevant to, but have yet to become commonly implemented in, contemporary quantitative original research in SEP. To achieve this purpose we provide a brief review of two types of power analysis (i.e., sample size determination; power estimation) for two different purposes (i.e., regarding model-data fit; regarding focal parameters) as implemented in a variety of available tools (e.g., tables; online utilities; software). Before providing this review, however, a few *key terms* are defined.

Statistical power can be defined as the probability of rejecting a false null hypothesis. While the utility of null hypothesis significance testing (NHST) has been debated in statistics (e.g., Wasserstein & Lazar, 2016), psychology (e.g., Cohen, 1994) and exercise science (e.g., Zhu, 2012), 'NHST is still the engine of statistical inference in most health and exercise sciences' (Buchanan & Lohse, 2016, p. 131). However, effect size (i.e., the magnitude of an effect) has been (e.g., Thomas, Salazar, & Landers, 1991) and is (e.g., Kelley & Preacher, 2012) at least as important a consideration as is statistical significance.

A type of power analysis that occurs prior to data collection (i.e., power is fixed and an estimate of sample size is desired) is referred to in the current manuscript as *sample size determination.* Sample size determination can perhaps be most beneficial at the planning stage of a study when resources related to data collection are being requested and/or allocated. Unsurprisingly, sample size determination has long been advocated for in both psychology (e.g., Cohen, 1994) and exercise science (e.g., Zhu, 2012). Given the substantial frequency of underpowered studies in SEP observed by Schweizer and Furley (2016), these authors cautioned, '…that researchers should take the issues of sample sizes seriously…' and suggested that '…researchers should calculate adequate sample sizes a priori based on to-be expected effects…' (p. 121).

A type of power analysis that occurs after data have been collected (i.e., sample size is fixed and an estimate of power is desired) is referred to in the current manuscript as *power estimation*. Power estimation can perhaps be most beneficial for providing an empirical context within which a statistically non-significant result was observed and/or providing updated power estimates (based on the newly collected data) that can be integrated into the planning of future research. Unsurprisingly, power estimation has long been regarded as an important consideration when interpreting related results of a statistical test of interest in both psychology (e.g., Cohen, 1994) and exercise science (e.g., Zhu, 2012). Schweizer and Furley (2016), however, analysed manuscripts published from 2009-2013 in four prominent journals in SEP and concluded that ‘A substantial proportion of published studies does not have sufficient power to detect effect sizes for psychological research’ (p. 114). Findings from Schweizer and Furley (2016) fit within a crisis of confidence in the broader psychological quantitative literature (e.g., Hoekstra, Morey, Rouder, & Wagenmakers, 2014).

From this point forward the expression *power analysis* is used when referring to both sample size determination and power estimation simultaneously. Power analysis in SEM relies on three core statistical concepts – null and alternative hypotheses, test statistics to assess null hypotheses, and central and non-central distributions – which for spatial reasons are not reviewed in this manuscript. Readers are referred to Hancock and French (2013) for a thorough treatment of each of these core topics.

Perhaps surprisingly given the findings of Schweizer and Furley (2016), a methodological literature on power analysis in SEM for two different purposes has been available for the past few decades (e.g., MacCallum, Browne, & Sugawara, 1996; Satorra & Saris, 1985). The first purpose focuses on the entire model, which we refer to as *power analysis regarding model-data fit*. The second purpose focuses on one or more specific parameters within an entire model, which we refer to as *power analysis regarding focal parameters*. Both types of power analysis (i.e., sample size determination; power estimation) can be used for both purposes of a power analysis (i.e., regarding model-data fit; regarding focal parameters) and often with a variety of available tools (e.g., tables; online utilities; software). Because there is recent evidence that the field of SEP does not, on average, report power analyses in SEM in published manuscripts (and thus there is not a large body of literature to review per se), we provide a few brief ‘how to’ demonstrations below.

A Related Application with Brief Demonstrations

In order to provide an overview of two types of power analysis for two different purposes as implemented in a variety of available tools, we first summarize a relevant application of SEM that we will refer to during our brief demonstrations. Myers, Park, et al. (2016) provided initial validity evidence for measuring multidimensional well-being in a Hispanic sample with the I COPPE Scale (Prilleltensky et al., 2015). More specifically, Myers, Park, et al. reported evidence that the measurement theory for responses to the I COPPE Scale emerged in an exploratory bi-factor analysis (under target rotation) and that the I COPPE subjective well-being factors exhibited convergent relations with scores from theoretically relevant comparison instruments. Figure 1 depicts standardized parameter estimates that are commonly of primary interest (i.e., 39 pattern coefficients and 7 correlation coefficients) from Myers, Park, et al.

The brief demonstrations provided below are intended to display a reasonable way to proceed in many applications of SEM in SEP. Some decisions are made, however, for the sake of textual parsimony and should be altered as justified within subsequent applications in practice. Type I error rate is set to $\alpha = .05$ and power is set to .80. Assumptions, too, are made about the model to be imposed (e.g., at least close model-population data fit), the data to be analysed (e.g., conditionally multivariate normal), and the estimation method that will be used (i.e., maximum likelihood). Readers are referred to Hancock and French (2013) for a thorough treatment of each of these assumptions.

Degrees of freedom are determined for the full model that is only partially depicted in Figure 1 by subtracting the number of parameters to be estimated (q) from the number of observations available for the analysis (*u*). Given that the means are assumed to be in the model, *u* can be determined by finding the value of: $p(p+3)/2$, where *p* is the number of observed variables. Therefore, the value of *u* is 434 (i.e., 28(28+3)/2). The value of *q* can be determined by summing the number of parameters to be estimated in the model. For example, specific parameters for the measurement model are as follows: 21 intercepts (i.e., one for each item), 126 pattern coefficients or 'loadings', 21 residual variances (i.e., one for each item) and 63 residual covariances; whereas specific parameters for the latent variable model are as follows: 7 means (i.e., one for each latent variable), 7 variances (i.e., one for each latent variable) and 70 covariances.[1] Therefore, the value of *q* is 315. The value of *df* is 119 (i.e., 434-315).

Brief Demonstration 1. This demonstration is intended to be applicable to future research (and particularly prior to data collection) in SEP when, in general, 'type' = sample size determination, 'purpose' = model-data fit, and 'tool' = table(s). In such cases, the necessary inputs to be provided by the user include: α-level, *df* for the entire model, desired level of power, a population model-data fit value for the null condition and a population model-data fit value for the alternative condition. To demonstrate, sample size is determined (for a given power value) regarding model-data fit for the Myers, Park, et al., (2016) example using the tables (e.g., Table 4.1 on p. 128) provided in Hancock and French (2013).[2] A value of population model-data fit (i.e., ε) in the root mean square error of approximation (i.e., RMSEA; Steiger & Lind, 1980) metric is specified as .05 for the null condition (i.e., ε_0). Two values of population model-data fit in the RMSEA metric are specified, .02 and .04, for the alternative condition (i.e., ε_1) consistent with the 90% confidence interval, [.018, .035], reported in Myers, Park, et al. Degrees of freedom for the entire model are rounded to 120 (from 119, as calculated above). Therefore, necessary sample size is equal to 191 when $\varepsilon_1 = .02$ and 702 when $\varepsilon_1 = .04$. Readers are referred to Hancock and French (2013) for more detailed step-by-step demonstrations of power analysis in SEM via tables.

Brief Demonstration 2. This demonstration is intended to be applicable to future research (and particularly after data collection) in SEP when, in general, 'type' = power estimation, 'purpose' = model-data fit, and 'tool' = online utility. In such cases, the necessary inputs to be provided by the user include: α-level, *df* for the entire model, sample size, a population model-data fit value for the null condition and a population model-data fit value for the alternative condition. To demonstrate, power is estimated (for a given sample size value) regarding model-data fit for the Myers, Park, et al., (2016) example using an online utility provided by Preacher and Coffman (2006) at http://quantpsy.org/rmsea/rmsea.htm. Population model-data fit values are identical to those specified in the previous paragraph. Degrees of freedom for the entire model are 119 because rounding is unnecessary in the online utility. A range of sample size values is specified (i.e., 250, 500, and 1000), consistent with relevant recommendations (e.g., Myers, Ahn, & Jin, 2011). When $\varepsilon_1 = .02$, power estimation is equal to .93 when sample size equals 250 and it approximates 1.00 when sample size equals 500 or 1000. When $\varepsilon_1 = .04$, power estimation is equal to .33 when sample size equals 250; .63 when sample size equals 500; and, .93 when sample size equals 1000. Readers are referred to Myers, Celimli, et al. (2016) for more detailed step-by-step demonstrations of power analysis in SEM via Preacher and Coffman's (2006) online utility.

Brief Demonstration 3. This demonstration is intended to be applicable to future research (and particularly after data collection) in SEP when, in general, 'type' = power estimation, 'purpose' = focal parameter(s), and 'tool' = software. In such cases (and under the user-friendly approach to be demonstrated), the necessary inputs to be provided by the user include: a dataset, a model, a population value for each focal parameter, α-level, and sample size. To demonstrate, power is estimated (for a given sample size value) regarding focal parameters for the Myers, Park, et al., (2016) example using Monte Carlo methods for a real data analysis via a two-step approach implemented in M*plus* 7.4 (Muthén & Muthén, 1998-2015).[3] Suppose that the 39 pattern coefficients and the 7 correlation coefficients depicted in Figure 1 are the focal parameters and that θ_i is used to symbolize a particular focal parameter. Monte Carlo methods can be used to determine the proportion of replications at which each $\mathrm{H}_0: \theta_i = 0$ is rejected for a particular sample size. A range of sample size values is specified: 250, 500, 1000. The number of replications is set to 10,000. The vast majority of parameter estimates for the entire model from Myers, Park, et al. were treated as population values.[4] The smallest power estimation value across all focal parameters is equal to .986 (i.e., covariance of interpersonal well-being with interpersonal comparison measure) when sample size equals 250 and 1.00 when sample size equals 500 or 1000.

Appendix A and Appendix B provide annotated input for Step 1 and Step 2, respectively.[5] Note that this code could also be used to determine sample size (for a given power value) regarding focal parameters. Appendix C provides truncated output identifying the power estimation value for each focal parameter when sample size equalled 250. Appendix D provides a simulated dataset (download file named dem_3.txt) so that readers can try running the syntax provided in Appendix A and Appendix B themselves.[6] Readers are referred to Muthén and Muthén (2002), Myers, Ahn, et al. (2011), and to Paxton, Curran, Bollen, Kirby and Chen (2001) for more detailed step-by-step demonstrations of power analysis in SEM via Monte Carlo methods with software. Readers are referred to Muthén and Muthén (2002) for a demonstration of how missing data and non-normal data may be accommodated in a power analysis in SEM via Monte Carlo methods with software.

Future Directions

Both the potential utility, and the relatively infrequent observation of, power analysis in SEM for original research in SEP have been known for some time (e.g., Biddle et al., 2001). Since the contribution of Biddle et al. (2001), however, a variety of progressively more accessible ‘how-to’ resources have been made available in an effort to increase the frequency of power analysis in SEM across disciplines for both types (i.e., sample size determination; power estimation) and both purposes (i.e., regarding model-data fit; regarding focal parameters) and with a variety of tools: tables (e.g., Hancock & French, 2013), online utilities (e.g., Preacher & Coffman, 2006), and software (e.g., Muthén & Muthén, 2002). The routine application of power analysis in SEM (and in other statistical modelling frameworks) for original research in SEP, however, has yet to fully emerge (e.g., Schweizer & Furley, 2016). The review, and the brief demonstrations of, power analysis in SEM provided in this manuscript (i.e., Brief Demonstration 1: sample size determination regarding model-data fit with a table; Brief Demonstration 2: power estimation regarding model-data fit with an online utility; Brief Demonstration 3: power estimation regarding focal parameters with software) should be viewed as an additional effort to expedite the full emergence of power analysis in SEM for contemporary quantitative original research in SEP. The expression ‘full emergence of power analysis in SEM’ should not be equated with the suggestion of a ‘golden rule’ that all studies in SEP that use SEM must report a power analysis as clearly there may be some cases where sufficient information is not available.

Exploratory Structural Equation Modelling

An Overview

Exploratory structural equation modelling (ESEM) was first proposed by Asparouhov and Muthén (2009). ESEM integrates exploratory factor analysis (EFA), independent clusters model confirmatory factor analysis (ICM-CFA) and structural equation modelling (SEM). ESEM can have an exploratory or confirmatory focus, depending on the research objectives of a study. Although ICM-CFA has typically been considered superior to EFA due to its greater parsimony and integration to the overarching SEM framework, recent research evidence has shown that forcing cross-loadings to be exactly zero tends to be overly restrictive for applied research. In contrast, using EFA typically accommodates such cross-loadings, particularly if they are small in size (Kline, 2000). ESEM allows the testing of such cross-loadings whilst at the same time preserving the advantages associated with ICM-CFA (e.g., path coefficients corrected for measurement error, testing of invariance of factor structure over time and/or groups). As noted by Asparouhov, Muthén, and Morin (2015), allowing cross-loadings does not undermine constructs by adding ‘noise’ but rather allows them to be estimated using all of the relevant information. Nevertheless, researchers should always aim to develop instruments that have small rather than large cross-loadings. It should also be noted that no cross-loadings should be allowed between factors which predict one another as this undermines the assumption of directionality of the associations.

By allowing cross-loadings on one or more factors, ESEM addresses important limitations associated with ICM-CFA. Specifically, by constraining cross-loadings to zero, ICM-CFA will result in inflated factor correlations; typically, the higher the magnitude of the cross-loadings, the greater the inflation in factor correlations (Marsh, Lüdtke, Nagengast, Morin, & VonDavier, 2013). As a result, positively biased and artificially inflated correlations undermine the discriminant validity of a multidimensional instrument and the predictive validity of its factors, due to multicollinearity (Marsh, Morin, Parker, & Kaur, 2014). Many instruments in SEP have correlated factors, hence, the use of ESEM is recommended to address this problem.

ESEM can be used when an instrument has two or more factors (because with a single-factor model there are no cross-loadings). ICM-CFA is nested under ESEM (Morin, Marsh, & Nagengast, 2013), hence the fit of the two models (and the plausibility of parameter estimates) can be compared as with any nested models (e.g., a chi-square difference test). Marsh et al. (2014) recommended that both ICM-CFA and ESEM models should be tested with the same data set; if the fit of both types of models is equivalent, the ICM-CFA model should be preferred as it is more parsimonious. However, Marsh et al. observed that the ICM-CFA is often too restrictive to provide acceptable fit for most psychological instruments; this is also the case in the field of SEP, as our brief review below indicates. It is also possible to include sets of ESEM and CFA factors in the same model.

One limitation of ESEM is that the pattern of cross-loadings and the size of the factor correlations will vary depending on the rotation method utilised (Morin et al., 2013). Examining model fit cannot help with this problem as fit indices are identical under different rotation methods. Marsh, Lüdtke, et al. (2013) recommended that the results of different estimation methods be compared. The online supplements accompanying the Morin, Marsh and Nagengast (2013) chapter suggested the potential for problems with geomin rotation in Mplus with a default epsilon value when using simulated data. In Table 1, we present the rotation method used in different ESEM studies in the SEP literature (and encourage the reporting of epsilon value(s) in future studies that use geomin rotation). In practice, the use of target rotation has been recently favored in the literature as providing a way to rely on a more confirmatory approach to the estimation of EFA factors (e.g., Myers, Jin, Ahn, Celimli, & Zopluoglu, 2015). With target rotation, researchers indicate the approximate size of expected cross-loadings. It should be noted, however, that this practice is appropriate when ESEM is used in a more confirmatory mode, in other words, when researchers have clear views of the factor structure expected. If neither ICM-CFA nor ESEM produces acceptable model fit, or if researchers do not have a clear view of the expected factor structure, ESEM can be used in an exploratory fashion (e.g., see Payne, Hudson, Akehurst, & Ntoumanis, 2013).

An advantage of ICM-CFA over EFA is the flexibility to examine the measurement invariance and compare latent means across groups and/or over time. Such an advantage is preserved under an ESEM framework. Readers are directed to Table 1 presented in Marsh et al. (2014) for a list of 13 tests of invariance that can be examined within ESEM (see also Schellenberg et al., 2014). However, unlike with ICM-CFA, some types of partial factor invariance cannot be tested via ESEM. Specifically, it is not possible to test partial invariance of factors loadings, variances, and covariances (or to separate tests of invariance of factor variances from those of factor covariances). However, it is possible to pursue tests of partial invariance of intercepts (or thresholds), uniqueness, and latent means. Therefore, Marsh, Nagengast and Morin (2013) proposed ESEM-within-CFA framework as a solution to address this problem. This technique also enables the testing of models not possible under ESEM, such as higher-order factor models, latent curve models, or models in which some but not all factors are related to other variables (e.g., demographics), or mediation models with bootsrapped confidence intervals. The readers are referred to Morin, Marsh, and Nagengast (2013) for more information on how ESEM-within-CFA can deal with such limitations of ESEM.

Related Applications

Applications of ESEM have grown substantially since the initial paper by Asparouhov and Muthén (2009). A search on the Scopus database in September 2016 indicated more than 200 articles utilising this method, with nearly 10% of them in the area of SEP (see Table 1 for an overview of select ESEM applications in SEP). Some researchers in the SEP field have utilised ESEM to test the factor structure of responses to a new questionnaires. For example, Appleton, Ntoumanis, Quested, Viladrich, and Duda (2016) developed and validated a new questionnaire that assesses young athletes' perceptions of the coaching environment, as proposed by achievement goal theory and self-determination theory (the Empowering and Disempowering Motivational Climate Questionnaire-Coach; EDMCQ-C). Drawing from various questionnaires, the authors pulled together an item pool which they then reduced by comparing alternative factor structures via ESEM, bi-factor ESEM, and ICM-CFA. A target rotation was utilised and hierarchical structures were compared using ESEM-Within-CFA and bi-factor ESEM approaches. Overall, ESEM solutions produced a better fit compared to ICM-CFA solutions, with bi-factor ESEM providing the best fit. However, some of the parameter estimates and obtained factor structures via ESEM did not conform to the theory underpinning the EDMCQ-C. The authors concluded that further work on the questionnaire was needed.

Other authors have used ESEM to adapt existing questionnaires. For example, Morin et al., (2016) tested a revised version of the short Physical Self-Inventory (PSI-S; Morin & Maïano, 2011) which included positively-worded reformulations of the original negatively-worded items. Morin et al. showed that scores from the revised PSI-S were invariant amongst samples of English and French-speaking adolescents. When compared to ICM-CFA, the ESEM produced better model fit and more orthogonal factors. For the ESEM, the authors used a target rotation in a confirmatory manner, specifying six correlated factors and cross-loadings as close to zero as possible. ESEM (as well as ICM-CFA) was used to show the longitudinal invariance of scale scores over a period of 7-8 months. Using a multiple-group multiple indicators multiple causes (MIMIC) approach, Morin et al. showed that the PSI-S scores showed no measurement bias in relation to gender, age, body mass index, or physical activity involvement.

Other authors have tested the measurement invariance of responses to a questionnaire by comparing an ICM-CFA model against an ESEM model. For example, Viladrich et al. (2013) examined the factor structure of responses to the Behavioural Regulation Sport in Questionnaire (BRSQ; Lonsdale, Hodge, & Rose, 2009) in youth soccer players from five European countries. The authors found that ESEM solutions (with target rotations) produced better model fit and lower inter-factor correlations compared to the ICM-CFA solutions. Further, ESEM-based invariance testing showed that BRSQ scores had metric invariance across the five samples. Viladrich et al. were not able to test for partial invariance as this is not possible in ESEM (unless an ESEM-within-CFA approach is implemented; see Marsh et al., 2013).

Future Directions

It would be interesting if researchers used ESEM to revisit the factor structure of scores from questionnaires that have been previously shown to have poor fit and/or poor factor discriminant validity when tested with ICM-CFA (for examples of such an effort see Perry, Nicholls, Clough, & Crust, 2015, and Fogarty, Perera, Furst, & Thomas, 2016). ESEM can be used in testing latent growth models, multi-trait multi-method (MTMM) models, bi-factor models, as well as latent path analysis models. Researchers in SEP are encouraged to explore such possibilities as they have certain advantages compared to ICM-CFA based approaches. For example, with regard to MTMM, Marsh et al. (2014) noted that compared to ESEM solutions, ICM-CFA solutions typically provide poorer tests of discriminant validity, which is particularly critical in MTMM studies. Further, a bi-factor ESEM approach (e.g., see Appleton et al. 2016 for an applications in SEP) is one way of testing factor structures within ESEM involving both a general and specific factors (see Morin, Arens, & Marsh, 2016, and Myers et al., 2014 where some distinctions are outlined between bi-factor ICM-CFA and bi-factor ESEM).

Bayesian structural equation modelling (BSEM; Muthén & Asparouhov, 2012) has a lot of similarities with ESEM, particularly when the target rotation is used (Marsh et al., 2014). BSEM allows cross-loadings via allowing researchers to provide estimated values based on previous research (or default software options). BSEM could be an alternative to ESEM when researchers are interested in testing higher-order factor structures or when the sample size is small relative to the complexity of the tested model. This is because Bayesian methodology does not require the normality assumption to be met, as is the case with frequentist tests such as ESEM and ICM-CFA (although with the latter it is possible to use estimation methods that take account non-normality). Researchers in SEP are encouraged to compare ESEM and BSEM approaches with the same data set (e.g., by examining the plausibility of obtained parameter estimates or whether solutions have converged with no error messages), particularly in cases where samples sizes are relatively small and instruments with numerous factors are modelled. A fuller review of Bayesian Statistics in SEP is provided in a subsequent section of this manuscript.

Mixture Modelling

An Overview

Researchers in SEP are often interested in examining group differences (e.g., sex) on a key variable of interest (e.g., intrinsic motivation). In this case, researchers will have an *a priori* hypothesis and therefore have collected information about a known grouping variable. However, there are times when researchers may not know if there are groups or subpopulations within their data. Mixture modelling can be used to uncover subpopulations that may exist in the data that were not known *a priori* (McLachlan & Peel, 2000), which is most likely to be the case when subpopulations exist on psychosocial variables. In contrast to the example above with sex, the researcher will not have collected information about the grouping variable and instead they rely on mixture modelling to identify the unobserved subpopulations (Nylund, Asparouhov, Muthén, 2007; Muthén & Muthén, 2000). These unobserved subpopulations are considered to be typological in that they provide a classification scheme and prototypical in that each participant has a given probability of membership to each subpopulation (Morin & Wang, 2016).

Mixture modelling is based on a *person centred approach.* The objective of a person centred approach is to examine relationships between people whereas the goal of a *variable centred approach* (e.g., classical SEM, regression methods) is to examine associations between variables (Morin & Wang, 2016; Muthén & Muthén, 2000). In mixture modelling, researchers identify relationships among people and classify or group them into categories called latent classes (for categorical indicators) or profiles (for continuous indicators). Given that most indicators used by SEP researchers are continuous, we will use the term ‘latent profile’ for the remainder of this section. Each latent profile contains people who are similar to each other (i.e., homogenous within groups) and different from people in other latent profiles (i.e., heterogeneous across groups; Muthén & Muthén, 2000) at one time point or over time (Nylund et al., 2007). Latent profiles can differ quantitatively (i.e., in levels or magnitude) and/or qualitatively (i.e., in shape or combinations of variables; Morin & Wang, 2016). For example, participants can have high, medium, and low levels of *both* autonomous and controlled motivation (i.e., only quantitative differences between profiles because within profiles there is a similar magnitude of autonomous and controlled motivation). Participants could also have differing levels of *each* type of motivation within one profile (e.g., profile 1 = high controlled, low autonomous motivation; profile 2 = high controlled, high autonomous motivation) and these qualitative differences within and between profiles of motivation (e.g., high/low and high/high) may lead to differential outcomes such as higher/lower physical activity participation.

Mixture modelling is considered to be an exploratory approach because researchers must fit several models specifying differing numbers of latent profiles in each model (Bauer & Curran, 2003). Typically, combinations of statistical criteria are used to determine the best model delineating the appropriate number of latent profiles. Simulation research (see Morin & Wang, 2016 for a recent review) has shown that the consistent Akaike information criterion (CAIC), Bayesian information criterion (BIC), the sample size adjusted BIC (ABIC), and the bootstrap likelihood ratio tests (bootstrap LRT), are effective for determining the number of profiles. Entropy can be used as a summary of the classification accuracy (see McLachlan & Peel, 2000; and Nylund, et al., 2007 for further details on profile enumeration). In addition, when estimating mixture models, researchers should be aware that they typically require large sample sizes and that multiple start values should be tested to ensure that the models converge on global rather than local solutions (McLachlan & Peel, 2000; Nylund et al., 2007). Alongside statistical criteria, it is important that researchers consider theory, the research question, parsimony, and the interpretability of the latent profiles (Bauer & Curran, 2003; Jung & Wickrama, 2008), as inferences made from incorrect models could cause ambiguity and erroneous conclusions (Duncan, Duncan, & Strycker, 2006; Nylund et al., 2007; Jung & Wickrama, 2008).

Traditionally, the specific type of mixture model invoked depended on the nature of the data (e.g., categorical or continuous) as well as the study design (e.g., cross-sectional or longitudinal). For example, within a cross-sectional design, latent class analysis (for categorical variables) and latent profile analysis (for continuous variables) can be used to examine unobserved subpopulations in observed variables (Muthén & Muthén, 2000).Within a longitudinal design, latent transition analysis (for categorical indicators) and latent profile transition analysis (for continuous indicators) can be used to examine change in class or profile membership, respectively, over time (Muthén & Muthén, 2000). Still within a longitudinal framework, latent class growth analysis can be used to examine one indicator over time to determine the number of different growth curves in a population (e.g., one class may have linear change and another may have quadratic change; Muthén & Muthén, 2000). However, in latent class growth analysis, only one mean growth curve is estimated for each latent class and for this reason, researchers have recently cautioned against its use given that it can lead to biased results caused by over-extracting spurious latent classes (Diallo, Morin, Lu, 2016). In contrast to the restricted latent class growth analysis, a growth mixture model can be estimated in which the mean growth curves are random and therefore, variation around the mean is permitted (Muthén & Muthén, 2000).

Other emerging types of mixture models include regression mixture models and factor mixture models. Regression mixture models can be used to examine if relationships between two variables differ across profiles of people (see Morin & Wang, 2016; Morin, Scalas, & Marsh, 2015). Factor mixture models combine a latent class (or profile) model with the common factor model (Lubke & Muthén, 2005). Therefore, in a factor mixture model, profiles are used to describe unobserved subpopulations whereas continuous latent factors are used to model the covariation among observed variables. Finally, generalized SEM (sometimes called general growth mixture modelling) is an extension of each of the above methods in that it allows researchers to integrate mixture modelling into a SEM framework. Therefore, using generalized

SEM researchers can examine antecedents or outcomes of profiles from any cross-sectional or longitudinal mixture model and also incorporate more than one type of mixture model into the same overall model (Morin & Wang, 2016; Muthén & Muthén, 2000).

Related Applications

Although variable centred analyses currently appear to be the *modus operandi* of SEP researchers, the advantages of mixture modelling and increasing ease of model estimation have led SEP researchers to employ mixture models to answer novel research questions (Morin & Wang, 2016). Table 2 provides an overview of select applications of mixture modelling in prominent SEP or related journals. For example, SEP researchers have used latent class or profile analysis to investigate if subgroups of athletes existed based on their perception of the talent development environment (Ivarsson et al., 2015) or if different profiles of exercise goal contents existed within the population (Lindwall, Weman-Josefsson, Sebire, & Standage, 2016). In the later application, Lindwall and colleagues (2016) uncovered five latent profiles of exercise goal contents that differed both quantitatively (i.e., one profile had low levels whereas another had high levels of goal contents) and qualitatively (i.e., the three remaining profiles had qualitatively different shapes/combinations of different types of goal contents).

Using a longitudinal design, Martinent and Nicolas (2016) first employed latent profile analysis to examine if there were different profiles of coping in sport and then conducted latent profile transition analysis to determine if athletes changed in their coping profiles over time. As a whole, they found evidence of four coping profiles over two separate time points and that there was some stability and change in these coping profiles over time (Martinent & Nicolas, 2016). Using latent class growth modelling, Gaudreau and colleagues (2009) examined if trajectories of positive and negative affect in elite adolescent hockey players changed over an 11-week period.

The authors found evidence of three trajectories of change in positive affect over time which they labelled as ‘high and decreasing’, ‘unstable’, and ‘medium and decreasing’. They also found three trajectories of change for negative affect over time which they labelled as ‘low and unstable’, ‘medium and unstable’ and ‘high and decreasing’.

Using a growth mixture model, Ventura and colleagues (2009) found four distinct trajectories in girls’ body mass index over ten years of childhood and adolescence. They also found that within each trajectory, there was individual variation such that each girl followed their own trajectory within their trajectory class. Finally, using general growth mixture modelling (or generalized SEM), Rodriguez and Audrain-McGovern (2004) identified four trajectories of change in sport participation from grade 9 to 11 and that participants in the ‘decreasing or erratic participation’ trajectory were almost three times more likely to be current smokers in grade 11 compared to those in the ‘high participation’ trajectory.

Future Directions

Mixture modelling is a rapidly developing area of statistics with advances being made annually. As mixture modelling becomes more accessible through further education (e.g., graduate student courses, workshops), developments in computer software, and advances in mixture modelling methods, we anticipate that SEP researchers will turn more frequently to mixture modelling to answer novel research questions.

Advances in Bayesian mixture modelling may be useful for SEP researchers dealing with complex models and small sample sizes. The current scarcity of Bayesian mixture modelling in SEP research could stem from unfamiliarity and the added complexity of Bayesian mixture modelling. For example, Bayesian mixture modelling can lead to issues associated with latent class labels switching during estimation (i.e., ‘switching labels’; Depaoli, 2013; Asparouhov &

Muthén, 2010), specifying priors, and violations of the assumption of conditional independence within mixture models (see Asparouhov & Muthén, 2010; Asparouhov & Muthén, 2011). Nevertheless, with further developments and Bayesian familiarity, SEP researchers may begin to take advantage of Bayesian mixture modelling. A fuller review of Bayesian Statistics in SEP is provided in the next section of this manuscript.

Four recent advances in mixture modelling involve modelling fully latent mixture models, auxiliary variables, multi-level mixture models, and examining profile similarity. First, rather than relying on manifest or observed variables in mixture modelling, researchers have begun to rely on mixture models based on latent variables (see Morin, Scalas, & Marsh, 2015), which is advantageous because latent variables remove measurement error. Second, when an external variable is added into a model to serve as a covariate, antecedent, or outcome, it can cause a shift in the meaning of the original latent profiles (Asparouhov & Muthén, 2014). New methods using auxiliary procedures in Mplus have been implemented to help researchers prevent these shifts in latent profiles (Asparouhov & Muthén, 2014; see Wang, Morin, Ryan, Liu, in press for an application of the auxiliary procedure in SEP research). Third, multi-level mixture models allow researchers to account for nested effects such as the effect of team membership on athletes. Finally, researchers have recently provided methods for examining if profiles obtained from mixture models are similar, a concept akin to testing measurement invariance in a variable-centred factor analysis approach (Morin, Meyer, Creusier, Biétry, 2016). Advances in mixture modelling with latent variables, auxiliary variables, employing multi-level mixture models, and examining profile similarities will likely gain momentum in the future and be incorporated into the mixture models SEP researchers employ.

Bayesian Statistics

An Overview

When thinking about a new project or idea, researchers often have some degree of prior knowledge (e.g., past research findings, theory) or expectation regarding the direction (e.g., positive or negative) and/or strength (e.g., small, moderate, large) of effects among the study variables. Armed with these expectations, a study is designed to test the idea (e.g., cross-sectional survey, experiment, intervention) and data are collected from the target population. Subsequently, these data are analysed with the view to ascertain the degree to which one's expectations or hypotheses are supported by the data. As is often the case in many scientific disciplines, including SEP (Buchanan & Lohse, 2016), the default approach to data analysis is to perform a significance test that is almost always summarised with a *p* value, and sometimes includes an associated effect size and/or confidence interval. Typically, the *p* value is the foundation of a dichotomous decision to reject or accept the null hypothesis (e.g., $p < .05$).

Despite the prominence in SEP research, the reliance strictly on p-values can be problematic and could lead to a misinterpretation of the results in several ways. First, using the *frequentist approach* (e.g., relying on p-values and ML), most researchers wish to know the probability that their hypothesis or theory is true given the data at hand; however, frequentist methods only provide insight into the probability of observing the same data or the probability of more extreme data, given a hypothesis or theory. Second, frequentist methods do not incorporate prior beliefs or expectations explicitly within the statistical model. Instead, frequentist methods rely on long-run frequency or a hypothetical infinite repetition of the same study in which the extremeness of the study data depends on data that were never observed. Such an approach limits the extent to which data are accumulated and synthesised over time because researchers essentially test the same null hypothesis repeatedly, while not explicitly incorporating results from previous research into their analyses (van de Schoot et al., 2014). Third, within a frequentist approach, interval estimates (e.g., confidence intervals) can be misinterpreted because they do not reflect the intuitive statements that most researchers wish to make from their data (Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2016); that is, within frequentist statistics, it would be incorrect to conclude that there is a 95% chance the effect of x on y (e.g., $\beta = .40$) lies between .30 and .50. Within a frequentist framework, a confidence interval is a 'numerical interval constructed around the estimate of a parameter' that is a property of a particular 'when used repeatedly across a series of hypothetical data sets (i.e., the sample space), yields intervals that contain the true parameter value in 95% of the cases' (Hoekstra et al., 2014, p. 1159). Finally, with regard to statistical significance tests, rejecting the null hypothesis (e.g., no difference between groups) does not provide support for the alternative hypothesis (e.g., differences between groups) because it is essentially undefined in frequentist statistics; nor does failing to reject the null hypothesis mean that the null hypothesis should be accepted (Greenland et al., 2016; Wasserstein & Lazar, 2016). As one approach for overcoming many of these issues, *Bayesian statistics* offer practical advantages for applied researchers who have an interest in parameter estimation or hypothesis testing (Wagenmakers, Morey, & Lee, 2016).

A summary of key differences between Bayesian and frequentist statistics is detailed in Table 3. A signature strength of Bayesian statistics is the formalisation of prior knowledge or beliefs into the statistical model through explicit statements regarding model parameters (e.g., mean, path coefficient). Prior knowledge can encompass past empirical work (e.g., pilot data, meta-analysis) or theoretical expectations (e.g., expert knowledge, direction of the effect). The degree of (un)certainty in this knowledge is modelled via the variance of the prior distribution, and includes three broad categories of expectations: (1) non-informative prior, which captures a substantial degree of uncertainty (e.g., equal probability of every value ranging from -∞ to +∞) and therefore may not strongly influence the results (i.e., data driven findings); (2) weakly informative prior, which reflects some degree of certainty (e.g., most likely value of the target parameter, though a wide range of plausible values ranging from -∞ to +∞) and therefore may minimally influence the final results; and (3) informative prior, which captures a substantial degree of certainty (e.g., most likely value of the target parameter, with a small variance; van de Schoot & Depaoli, 2014; van de Schoot et al., 2014) and therefore may substantially influence the final results. With regard to psychometric examinations of questionnaires, for example, researchers can use informative priors in a confirmatory fashion to model cross-loadings with mean of zero, small variance priors, and intended factor loadings with mean and variance values that are informed by previous factor analyses or guidelines for the meaningfulness of factor loadings (e.g., Howle et al., 2016; Niven & Markland, 2016). Non-informative priors could be utilised in cases where researchers want to capitilise on the strengths of Bayesian statistics (e.g., computationally cumbersome models with ML; Doron & Gaudreau, 2014; Tamminen et al., 2016) or where no prior knowledge exists. Finally, researchers could draw from theoretical expectations to propose weakly informative priors whereby the direction of an effect is expected alongside uncertainty regarding the strength of the association (e.g., Mahoney et al., 2014).

In Bayesian statistics, one's prior knowledge is combined or 'mixed' with new data to produce the posterior distribution, which provides a full summary of what is known about a parameter. For the purposes of hypothesis testing, the posterior distribution alone is unsuitable and therefore requires a comparison of the degree of belief for two competing models or hypotheses (Morey, Romeijn, & Rouder, 2016). For example, one can compare the relative plausibility of the null hypothesis (i.e., absence of an effect) with an alternative hypothesis (i.e., presence of an effect). Of interest here is the change in one's belief from before seeing the data to afterwards, which is captured in the *Bayes factor* (Morey, Romeijn et al., 2016; Wagenmakers, 2007). Using a Bayes factor, the researcher can test the degree to which the data at hand are most compatible with the null or alternative hypothesis. Using a Bayes factor (*B*), the researcher can test the degree to which the data at hand are most compatible with the null ($B < 1/3$) or alternative hypothesis ($B > 3$), or whether the data are insensitive ($1/3 < B > 3$) (Dienes, 2016). Thus, unlike *p* values, the Bayes factor can provide evidence for the null hypothesis (Dienes, 2016). For example, one might expect a zero correlation between athletic performance and the number of sporting themed movies one has watched (H_0), whereas the alternative hypothesis (H_1) relaxes this restriction to specify an equal probability of every value ranging from $\pm$ 1. A comparison of the likelihood of each hypothesis being correct, given the data at hand, indicates that the observed data are 5.65 times more likely under H_0 when compared with H_1. In other words, 'the data shift our prior beliefs about the relative plausibility of the competing hypotheses' by a factor of 5.65 (Wagenmakers et al., 2016, p. 171). Readers are referred elsewhere for user-friendly overviews of Bayesian statistics (Muthén & Asparouhov, 2012; van de Schoot et al., 2014; Wagenmakers et al., in press; Zyphur & Oswald, 2015), including those with a specific focus on SEP (Gucciardi & Zyphur, 2016; Stenling, Ivarsson, Johnson, & Lindwall, 2015). For a broader and comprehensive overview of the theoretical and practical underpinnings of Bayesian statistics, Etz and colleagues (in press) have produced a reading list to serve as a starting point for researchers who are new to the area.

Related Applications

The application of Bayesian statistics within the psychological sciences is on the rise (van de Schoot, Winter, Ryan, Zondervan-Zwijnenburg, & Depaoli, in press). Coinciding with this increased interest, there have been several applications of Bayesian statistics within the field of SEP over the past few years. An overview of some such applications is provided in Table 4. With the exception of one study, which employed Bayesian network analysis (Constantinou et al., 2014), SEP researchers have applied Bayesian statistics for the primary purpose of parameter estimation. The majority of this work has employed BSEM to examine the factorial validity of scores from questionnaires designed to assess constructs such as commitment (Jackson et al., 2014), sport motivation (Stenling et al., 2015), walking motivation (Niven & Markland, 2016), and movement skill competence (Barnett et al., 2016). Researchers have also employed BSEM to test theoretical sequences that encompass multiple antecedent, intermediary and outcome variables, such as the relations from self-efficacy beliefs to performance on endurance-based physical activity tasks via self-presentation motives and personal task goals (Howle et al., 2016); motivational pathways informed by self-determination theory (Chan et al., 2015); and the integration of basic psychological needs and the theory of planned behaviour (Gucciardi & Jackson, 2015). Other applications of Bayesian statistics include multilevel modelling (Doron & Gaudreau, 2014; Tamminen et al., 2016), latent growth modelling (Noordstar et al., 2016), and network analysis (Constantinou et al., 2014). Within and across each of the studies, researchers have drawn from theory and past empirical work to incorporate weakly informative and informative prior information, or employed the default non-informative prior. Readers are encouraged to consult Gucciardi and Zyphur (2016) for a didactic demonstration of the application of BSEM, and those papers listed in Table 4 where the authors made available their syntax.

Future Directions

Readers who completed their educational training in psychology or the sport and exercise sciences are likely familiar with the classical approach to statistical analysis that is founded on frequentist methods (e.g., *p* values). Despite being advocated as the preferred statistical approach for the psychological sciences over 50 years ago (Edwards, Lindman, & Savage, 1963), it is only in the past decade that Bayesian statistics have taken flight (van de Schoot et al., in press). With the rapid and continuous advancements in the computational capacities of computers, development of user-friendly statistical software packages (e.g., M*plus*, JASP), and publication of didactical and primer papers (e.g., Depaoli & van de Schoot, in press), we expect Bayesian statistics to (soon) play an important role in the evolution of SEP research and practice. In addition to the possibilities outlined in Table 4 and elsewhere (van de Schoot et al., in press), Bayesian statistics can offer new insights through a range of common and uncommon analytical approaches such as evidence synthesis via meta-analysis (Scheibehenne, Jamil, & Wagenmakers, 2016), sequential hypothesis testing (Schönbrodt, Wagenmakers, Zehetleitner, & Perugini, in press), analysis of single-subject designs (de Vries, Hartogs, & Morey, 2015), mixture modelling (Depaoli, 2013), and reproducibility efforts (Etz & Vandekerckhove, 2016).

Bayesian statistics are not without criticism. For most critics, the subjectivity of the prior is a critical concern with Bayesian statistics (e.g., Bowers & Davis, 2012). For example, two people (or research groups) may have different expectations of the study hypotheses and therefore specify different priors to be mixed with the data. As a result, these differing perspectives may result in different findings from the same data. There are at least two ways by which researchers who employ Bayesian statistics can minimise such concerns. First, as with any scientific endeavour, transparency with regard to the foundations of the priors is of central importance, both in terms of where they came from (e.g., past work, theoretical expectations) and their appropriateness to be mixed with the data to make inferences with posteriors (van de Schoot et al., 2014; Zyphur & Oswald, 2015). Second, it is important that researchers conduct sensitivity analyses to ascertain the degree of influence of the priors, that is, whether or not fluctuations in background knowledge influence the stability of inferences made with posteriors (Depaoli & van de Schoot, in press). There are two broad categories of sensitivity analyses (Depaoli & van de Schoot, in press). First, weakly informative or informative priors could be compared with uninformative priors to understand the degree of subjectivity and influence on the posterior distribution. Second, weakly informative or informative priors could also be compared with varied prior distributions in which the mean and variance values are adjusted upwards or downwards to examine the influence of small to large fluctuations in prior beliefs.

Concluding Remarks

Our intent with this manuscript was to provide a partial update to the seminal paper by Biddle and colleauges (2001) by outlining four emerging quantitative analyses that can be used by SEP researchers to answer novel research questions. Although we value the broad quantitative and qualitative approach taken by Biddle et al., we chose to review only four emergent quantitative methodologies in SEP research because we believe that our field – present authours included – may be in danger of at least occasionally ‘driving fast in reverse’ (Steiger, 2001) with regard to the application of advanced latent variable models. Most simply, we believe that while several user-friendly software programs have recently made it very easy to impose a wide variety of advanced latent variable models with a variety of estimators, an unfortunate by-product of these impressive technological developments is the increasing possibility that users may fit a complex model (and perhaps with an estimator) that they do not have a very deep understanding of (i.e., driving fast in reverse). It is hard to believe that such an approach is an optimal way to efficiently advance knowledge in any discipline, however elegant the model and/or the estimator is/are. For this reason we chose to devote more text to only a few emergent quantitative analyses in SEP in hope that readers will gain at least an increased awareness of one or more of the analyses that we have reviewed. Perhaps more importantly, however, we hope that readers will gain a broader appreciation of just how much preparation it likely will take to knowingly and thoughtfully apply any advanced latent variable model. Finally, we encourage all researchers in SEP to avoid the temptation to become dogmatic about the universal implementation of a specific facet (e.g., a particular model and/or an estimator) of advanced latent variable modelling.

In an effort to avoid ‘driving fast in reverse’ or becoming dogmatic in approaches, we offer a few final recommendations to accelerate knowing and thoughtful applications of advanced quantitative analyses in SEP. First, in recognizing that the statistical analyses outlined are complex and potentially daunting to implement, we have provided broad overviews of each method alongside tangible applications, while also referring readers to published SEP examples with accompanying syntax. We recommend that readers consult these resources to gain an in-depth understanding of the methods and how to model them using proper syntax. To this end, although software developers continue to implement accessible syntax, we encourage readers to avoid using syntax without a deep understanding of what each key command invokes. Learning syntax is similar to learning a new language; when one begins to master the basics, the foundation for further application and extension can be easily developed. Second, readers are encouraged to actively seek opportunities for further education. Resources for students and academics alike exist. For example, if advanced statistical courses are not offered within one’s department, opportunities to take courses in related departments (e.g., education, psychology) can be sought. There are also many accessible workshops offered around the world, conferences frequently offer workshops or symposiums, courses are available online, and text books represent an excellent resource for self-guided learning. It is our hope that our brief overview of emergent quantitative analyses in SEP has sparked a curiosity in readers and nurtured a sense of intrinsic motivation to initiate further, deeper, learning of quantitative analyses. In so doing, we are hopeful that more researchers will join the second or third profile of researchers who seek to maintain or who are at the forefront, respectively, of understanding and applying emerging quantitative analyses.

References

Alcaraz, S., Torregrosa, M., & Viladrich, C. (2015). How coaches' motivations mediate between basic psychological needs and well-being/ill-being. *Research Quarterly for Exercise and Sport*, *86*, 292-302. doi:10.1080/02701367.2015.1049691

Andruff, H., Carraro, N., Thompson, A., Gaudreau, P., & Louvet, B. (2009). Latent class growth modelling: a tutorial. *Tutorials in Quantitative Methods for Psychology*, *5*, 11-24. doi:10.20982/tqmp.05.1.p011

Appleton, P. R., Ntoumanis, N., Quested, E., Viladrich, C., & Duda, J. L. (2016). Initial validation of the coach-created Empowering and Disempowering Motivational Climate Questionnaire (EDMCQ-C). *Psychology of Sport and Exercise*, *22*, 53-65. doi:10.1016/j.psychsport.2015.05.008

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*, 397-438. doi:10.1080/10705510903008204

Asparouhov, T., & Muthén, B. (2010). Bayesian analysis using Mplus: Technical implementation. *Manuscript submitted for publication*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.310.3903&rep=rep1&type=pdf

Asparouhov, T., & Muthén, B. (2011). Using Bayesian priors for more flexible latent class analysis. In *Proceedings of the 2011 Joint Statistical Meeting, Miami, FL*. Retrieved from http://www.statmodel.com/download/BayesLCA.pdf

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 21, 329-341. doi:10.1080/10705511.2014.915181

Asparouhov, T., Muthén, B., & Morin, A. J. S. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances: Comments on Stromeyer et al. *Journal of Management, 41,* 1561-1577. doi:10.1177/0149206315591075

Barnett, A., Cerin, E., Reaburn, P., & Hooper, S. (2012). Identifying mediators of training effects on performance-related psychobiosocial states: A single-case observational study in an elite female triathlete. *Psychology of Sport and Exercise*, *13*, 541-549. doi:10.1016/j.psychsport.2012.02.010

Barnett, L. M., Vazou, S., Abbott, G., Bowe, S. J., Robinson, L. E., Ridgers, N. D., & Salmon, J. (2016). Construct validity of the pictorial scale of Perceived Movement Skills Competence. *Psychology of Sport and Exercise, 22*, 294-302. doi:10.1016/j.psychsport.2015.09.002

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models over-extraction of latent trajectory classes. *Psychological Methods, 8,* 338-363. doi:10.1037/1082-989X.8.3.338

Biddle, S. J. H., Markland, D., Gilbourne, D., Chatzisarantis, N. L. D., & Sparkes, A. C. (2001). Research methods in sport and exercise psychology: Quantitative and qualitative issues. *Journal of Sports Scie*nces, *19*, 777-809. doi:10.1080/026404101317015438

Bowers, J. S., & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, *138*, 389-414. doi:10.1037/a0026450

Buchanan, T. L., & Loshe, K. R. (2016). Researchers' perceptions of statistical significance contribute to bias in health and exercise science. *Measurement in Physical Education and Exercise Science, 20,* 131-139. doi:10.1080/1091367X.2016.1166112

Chamarro, A., Penelo, E., Fornieles, A., Oberst, U., Vallerand, R. J., & Fernández-Castro, J. (2015). Psychometric properties of the Spanish version of the Passion Scale. *Psicothema*, *27*, 402-409. doi:10.7334/psicothema2015.80

Chan, D. K. C., Ivarsson, A., Stenling, A., Yang, S. X., Chatzisarantis, N. L. D., & Hagger, M. S. (2015). Response-order effects in survey methods: A randomised controlled crossover study in the context of sport injury prevention. *Journal of Sport & Exercise Psychology*, *37*, 666-673. doi:10.1123/jsep.2015-0045

Chiu, W., Rodriguez, F. M., & Won, D. (2016). Revisiting the Leadership Scale for sport examining factor structure through exploratory structural equation modeling. *Psychological Reports*, *119*, 435-449. doi:10.1177/0033294116662880

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997-1003. doi:10.1037/0003-066X.49.12.997

Constantinou, A. C., Fenton, N. E., & Pollock, L. J. H. (2014). Bayesian networks for unbiased assessment of referee bias in association football. *Psychology of Sport and Exercise*, *15*, 538-547. doi:10.1016/j.psychsport.2014.05.009

Depaoli, S. (2013). Mixture class recovery in GMM under varying degrees of class separation: Frequentist versus Bayesian estimation. *Psychological Methods*, *18*, 186-219. doi:10.1037/a0031609

Depaoli, S., & van de Schoot, R. (in press). Improving transparency and replication in Bayesian statistics: The WAMBS-checklist. *Psychological Methods*. doi:10.1037/met0000065

de Vries, R. M., Hartogs, B. M. A., & Morey, R. D. (2015). A tutorial on computing Bayes factors for single-subject designs. *Behaviour Therapy*, *46*, 809-823. doi:10.1016/j.beth.2014.09.013

Diallo, T. M. O., Morin, A. J. S., & Lu, H. (2016). Impact of misspecifications of the latent variance–covariance and residual matrices on the class enumeration accuracy of growth mixture models. *Structural Equation Modeling: A Multidisciplinary Journal, 23,* 507-531. doi:10.1080/10705511.2016.1169188

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology, 72,* 78-89. doi:10.1016/j.jmp.2015.10.003

Doron, J., & Gaudreau, P. (2014). A point-by-point analysis of performance in a fencing match: Psychological processes associated with winning and losing streaks. *Journal of Sport & Exercise Psychology, 36*, 3-13. doi:10.1123/jsep.2013-0043

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and application* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference in psychological research. *Psychological Review, 70,* 193-242. doi:10.1037/h0044139

Etz, A., & Vandekerckhove, J. (2016). A Bayesian perspective on the reproducibility project: Psychology. *PLoS ONE, 11*. doi:10.1371/journal.pone.0149794

Etz, A., Gronau, Q., Dablander, F., Edelsbrunner, P., Baribault, B. (in press). How to become a Bayesian in eight easy steps: An annotated reading list. *Psychonomic Bulletin and Review.* Retrieved from http://www.springer.com/psychology/cognitive+psychology/journal/13423

Finch, W. H., & Bronk, K. C. (2011). Conducting confirmatory latent class analysis using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal, 18,* 132-151. doi:10.1080/10705511.2011.532732

Fitzpatrick, J., Gareau, A., Lafontaine, M. F., & Gaudreau, P. (2016). How to use the actor-partner interdependence model (APIM) to estimate different Dyadic patterns in MPLUS: A step-by-step Tutorial. *The Quantitative Methods for Psychology*, *12*, 74-86. doi:10.20982/tqmp.12.1.p074

Fogarty, G. J., Perera, H. N., Furst, A. J., & Thomas, P. R. (2016). Evaluating measures of optimism and sport confidence. *Measurement in Physical Education and Exercise Science*, *20*, 81-92. doi:10.1080/1091367X.2015.1111220

Gaudreau, P., Amiot, C. E., & Vallerand, R. J. (2009). Trajectories of affective states in adolescent hockey players: Turning point and motivational antecedents. *Developmental Psychology, 45*, 307. doi:10.1037/a0014134

Gerber, M., Jonsdottir, I. H., Lindwall, M., & Ahlborg, G. (2014). Physical activity in employees with differing occupational stress and mental health profiles: A latent profile analysis. *Psychology of Sport and Exercise, 15*, 649-658. doi:10.1016/j.psychsport.2014.07.012

Greenland, S., Senn, S.J., Rothman, K.J., Carlin, J.B., Poole, C., Goodman, S.N., & Altman, D.G. (2016). Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology*, *31*, 337-350. doi:10.1007/s10654-016-0149-3

Gucciardi, D. F., Hanton, S., & Mallett, C. J. (2012). Progressing measurement in mental toughness: A case example of the Mental Toughness Questionnaire 48. *Sport, Exercise, and Performance Psychology, 1*, 194. doi: 10.1037/a0027190

Gucciardi, D. F., & Jackson, B. (2015). Understanding sport continuation: An integration of the theories of planned behaviour and basic psychological needs. *Journal of Science and Medicine in Sport*, *18*, 31-36. doi:10.1016/j.jsams.2013.11.011

Gucciardi, D. F., Peeling, P., Ducker, K. J., & Dawson, B. (2016). When the going gets tough: Mental toughness and its relationship with behavioural perseverance. *Journal of Science and Medicine in Sport*, *19*, 81-86. doi:10.1016/j.jsams.2014.12.005

Gucciardi, D. F., Zhang, C., Ponnusamy, V., Si, G., & Stenling, A. (2016). Cross-cultural invariance of the Mental Toughness Inventory among Australian, Chinese, and Malaysian athletes: A Bayesian estimation approach. *Journal of Sport & Exercise Psychology*, *38*, 187-202. doi:10.1123/jsep.2015-0320

Gucciardi, D.F., & Zyphur, M.J. (2016). Exploratory structural equation modelling and Bayesian estimation. In N. Ntoumanis & N.D. Myers (Eds.), *An introduction to intermediate and advanced statistical analyses for sport and exercise scientists* (pp. 172-194). Chichester, UK: Wiley.

Gunnell, K. E., & Gaudreau, P. (2015). Testing a bi-factor model to disentangle general and specific factors of motivation in self-determination theory. *Personality and Individual Differences, 81,* 35-40. doi:10.1016/j.paid.2014.12.059

Hancock, G. R., & French, B. F. (2013). Power analysis in covariance structure models. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed.) (pp. 117-159). Charlotte, NC: Information Age Publishing, Inc.

Hancock, G. R. (2016). [Peer commentary on the book "*An Introduction to Intermediate and Advanced Statistical Analyses for Sport and Exercise Scientists"* by N. Ntoumanis & N. D. Myers (Eds.)]. Retrieved from http://www.wiley.com/WileyCDA/WileyTitle/productCd-1118962052.html

Hancox, J. E., Quested, E., Viladrich, C., & Duda, J. L. (2015). Examination of the internal structure of the behavioural regulation in sport questionnaire among dancers. *International*

Journal of Sport and Exercise Psychology, 13, 382-397. doi:10.1080/1612197X.2014.999347

Hodge, K., & Gucciardi, D. F. (2015). Antisocial and prosocial behaviour in sport: The role of motivational climate, basic psychological needs, and moral disengagement. *Journal of Sport & Exercise Psychology*, *37*, 257-273. doi:10.1123/jsep.2014-0225

Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E. J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, *21*, 1157-1164. doi:10.3758/s13423-013-0572-3

Howle, T. C., Dimmock, J. A., & Jackson, B. (2016). Relations between self-efficacy beliefs, self-presentation motives, personal task goals, and performance on endurance-based physical activity tasks. *Psychology of Sport and Exercise*, *22*, 149-159. doi:10.1016/j.psychsport.2015.06.010

Ivarsson, A. Stenling, A., Fallby, J., Johnson, U., Borg, E., & Johansson, G. (2015). The predictive ability of the talent development environment on youth elite football players' well-being: A person-centered approach. *Psychology of Sport and Exercise*, *16*, 15-23. doi:10.1016/j.psychsport.2014.09.006

Jackson, B., Gucciardi, D. F., & Dimmock, J. A. (2014). Toward a multidimensional model of athletes’ commitment to coach-athlete relationships and interdependent sport teams: A substantive-methodological synergy. *Journal of Sport & Exercise Psychology*, *36*, 52-68. doi:10.1123/jsep.2013-0038

Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modelling. *Social and Personality Psychology Compass*, *2*, 302-317. doi:10.1111/j.1751-9004.2007.00054.x

Kawabata, M., & Mallett, C. J. (2013). Examining discriminant validity issues of the Sport Motivation Scale-6. *Journal of sports sciences, 31*, 173-180. doi: 10.1080/02640414.2012.723132

Kelley, K., & Preacher, K. J. (2012). On effect size. *Psychological Methods*, *17*, 137-152. doi:10.1037/a0028086

Kline, P. (2000). *Handbook of Psychological Testing* (2nd ed.). London: Routledge.

Koh, K. T., Kawabata, M., & Mallett, C. J. (2014). The Coaching Behavior Scale for sport: Factor structure examination for Singaporean youth athletes. *International Journal of Sports Science & Coaching*, *9*, 1311-1324. doi:10.1260/1747-9541.9.6.1311

Lindwall, M., Weman-Josefsson, K., Sebire, S. J., & Standage, M. (2016). Viewing exercise goal content through a person-oriented lens: A self-determination perspective. *Psychology of Sport and Exercise*, *27*, 85-92. doi:10.1016/j.psychsport.2016.06.011

Locke, S. R., & Brawley, L. R. (2016). Development and initial validity of the Exercise-Related Cognitive Errors Questionnaire. *Psychology of Sport and Exercise*, *23*, 82-89. doi:10.1016/j.psychsport.2015.11.003

Lonsdale, C., Hodge, K., & Rose, E. (2009). Athlete burnout in elite sport: A self-determination perspective. *Journal of Sports Sciences*, *27*, 785-795. doi:10.1080/02640410902929366

Louvet, B., Gaudreau, P., Menaut, A., Genty, J., & Deneuve, P. (2007). Longitudinal patterns of stability and change in coping across three competitions: A latent class growth analysis. *Journal of Sport and Exercise Psychology*, *29*, 100. doi:10.1123/jsep.29.1.100

Louvet, B., Gaudreau, P., Menaut, A., Genty, J., & Deneuve, P. (2009). Revisiting the changing and stable properties of coping utilization using latent class growth analysis: A longitudinal investigation with soccer referees. *Psychology of Sport and Exercise*, *10*, 124-135.

doi:10.1016/j.psychsport.2008.02.002

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10,* 21-39. doi: 10.1037/1082-989X.10.1.21

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130-149. doi:10.1037/1082-989X.1.2.130

Mahoney, J. W., Gucciardi, D. F., Ntoumanis, N., & Mallett, C. J. (2014). Mental toughness in sport: Motivational antecedents and associations with performance and psychological health. *Journal of Sport & Exercise Psychology*, *36*, 281-292. doi:10.1123/jsep.2013-0260

Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J. S., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right-camouflaging misspecification with item parcels in CFA models. *Psychological Methods, 18*, 257- 284. doi:10.1037/a0032773

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*, 85-110. doi:10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, *49*, 1194-1218. doi:10.1037/a0026913

Martinent, G., & Decret, J. C. (2015). Motivational profiles among young table-tennis players in intensive training settings: a latent profile transition analysis. *Journal of Applied Sport Psychology*, *27*, 268-287. doi:10.1080/10413200.2014.993485

Martinent, G., & Nicolas, M. (2016). A Latent profile transition analysis of coping within competitive situations. *Sport, Exercise, and Performance Psychology*, *5*, 218-231. doi:10.1037/spy0000062

Massey, W. V., Meyer, B. B., & Mullen, S. P. (2015). Initial development and validity evidence for the processes of change in Psychological Skills Training Questionnaire. *Psychology of Sport and Exercise*, *16*, 79-87. doi:10.1016/j.psychsport.2014.09.009

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: John Wiley & Sons.

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E. J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review, 23,* 103-123. doi:10.3758/s13423-015-0947-8

Morey, R. D., Romeijn, J. W., & Rouder, J. N. (2016). The philosophy of Bayes factors and the quantification of statistical evidence. *Journal of Mathematical Psychology, 72,* 6-18. doi:10.1016/j.jmp.2015.11.001

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal, 23,* 116-139. doi:10.1080/10705511.2014.961800

Morin, A. J. S., & Maïano, C. (2011). Cross-validation of the short form of the physical self-inventory (PSI-S) using exploratory structural equation modeling (ESEM). *Psychology of Sport and Exercise*, *12*, 540-554. doi:10.1016/j.psychsport.2011.04.003

Morin, A. J. S., Maïano, C., Marsh, H. W., Nagengast, B., & Janosz, M. (2013). School life and adolescents' self-esteem trajectories. *Child Development*, *84*, 1967-1988. doi:10.1111/cdev.12089

Morin, A. J. S, Maïano, C., White, R. L., Owen, K. B., Tracey, D., Mascret, N., & Lonsdale, C. (2016). English validation of the short form of the Physical Self-Inventory (PSI-S). *Psychology of Sport and Exercise, 27,* 180-194. doi:10.1016/j.psychsport.2016.08.016

Morin, A. J. S., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In Hancock, G. R., & Mueller, R. O. (Eds.). (2013). *Structural Equation Modeling: A second course* (2nd ed.). Charlotte, NC: Information Age Publishing, Inc.

Morin, A. J. S., Meyer, J. P., Creusier, J., & Biétry, F. (2016). Multiple-group analysis of similarity in latent profile solutions. *Organizational Research Methods, 19,* 231-254. doi:10.1177/1094428115621148

Morin, A. J. S., Scalas, L. F., & Marsh, H. W. (2015). Tracking the elusive actual-ideal discrepancy model within latent subpopulations. *Journal of Individual Differences, 36*, 65-72. doi:10.1027/1614-0001/a000157

Morin, A. J. S., & Wang, J. C. K. (2016). A gentle introduction to mixture modeling using physical fitness performance data In N. Ntoumanis & N. D. Myers (Eds.), *An introduction to intermediate and advanced statistical analyses for sport and exercise scientists* (pp. 183-209). Chichester, UK: Wiley.

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modelling: A more flexible representation of substantive theory. *Psychological Methods*, *17*, 313-335. doi:10.1037/a0026802

Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*, 882-891. doi:10.1111/j.1530-0277.2000.tb02070.x

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 599-620. doi:10.1207/S15328007SEM0904_8

Myers, N. D. (2013). Coaching competency and (exploratory) structural equation modeling: Asubstantive-methodological synergy. *Psychology of Sport and Exercise*, *14,* 709-718. doi:10.1016/j.psychsport.2013.04.008

Myers, N. D., Ahn, S., & Jin, Y. (2011). Sample size and power estimates for a confirmatory factor analytic model in exercise and sport: A Monte Carlo approach. *Research Quarterly for Exercise and Sport*, *82,* 412-423. doi:10.5641/027013611X13275191443621

Myers, N. D., Ahn, S., & Jin, Y. (2013). Rotation to a partially specified target matrix in exploratory factor analysis: How many targets? *Structural Equation Modeling: A Multidisciplinary Journal, 20,* 131-147. doi:10.1080/10705511.2013.742399

Myers, N. D., Brincks, A. M., Ames, A. J., Prado, G., Penedo, F. J., & Benedict, C. (2012). Multilevel modeling in psychosomatic medicine research. *Psychosomatic Medicine: Journal of Biobehavioral Medicine, 74,* 925-936. doi:10.1097/PSY.0b013e3182736971

Myers, N. D., Celimli, S., Martin, J. J., & Hancock, G. R. (2016). Sample size determination and power estimation in structural equation modeling. In N. Ntoumanis & N.D. Myers (Eds.), *An introduction to intermediate and advanced statistical analyses for sport and exercise scientists* (pp. 267-284). Chichester, UK: Wiley.

Myers, N. D., Chase, M. A., Pierce, S. W., & Martin, E. (2011). Coaching efficacy and exploratory structural equation modeling: A substantive-methodological synergy. *Journal of Sport & Exercise Psychology*, *33*, 779-806. doi:10.1123/jsep.33.6.779

Myers, N. D., Feltz, D. L., Guillén, F., & Dithurbide, L. (2012). Development of, and initial validity evidence for, the referee self-efficacy scale: a multistudy report. *Journal of Sport & Exercise Psychology*, *34*, 737-65. doi:10.1123/jsep.34.6.737

Myers, N. D., Jin, Y., Ahn, S., Celimli, S., & Zopluoglu, C. (2015). Rotation to a partially specified target matrix in exploratory factor analysis in practice. *Behavior Research Methods, 47,* 494-505. doi:10.3758/s13428-014-0486-7

Myers, N. D., Martin, J. J., Ntoumanis, N., Celimli, S., & Bartholomew, K. J. (2014). Exploratory bi-factor analysis in sport, exercise, and performance psychology: A substantive-methodological synergy. *Sport, Exercise, and Performance Psychology*, *3*, 258-272. doi:10.1037/spy0000015

Myers, N. D., Park, S. E., Lefevor, G. T., Dietz, S., Prilleltensky, I., & Prado, G. J. (2016). Measuring multidimensional well-being with the I COPPE Scale in a Hispanic sample. *Measurement in Physical Education and Exercise Science*, *20*, 230-243. doi:10.1080/1091367X.2016.1226836

Nicholls, A. R., Levy, A. R., Carson, F., Thompson, M. A., & Perry, J. L. (2016). The applicability of self-regulation theories in sport: Goal adjustment capacities, stress appraisals, coping, and well-being among athletes. *Psychology of Sport and Exercise*, *27,* 47-55. doi: 10.1016/j.psychsport.2016.07.011

Niven, A. G., & Markland, D. (2016). Using self-determination theory to understand motivation for walking: Instrument development and model testing using Bayesian structural equation modelling. *Psychology of Sport and Exercise, 23,* 90-100. doi:10.1016/j.psychsport.2015.11.004

Noordstar, J. J., van der Net, J., Jak, S., Helders, P. J. M., & Jongmans, M. J. (2016). Global self-esteem, perceived athletic competence, and physical activity in children: A longitudinal cohort study. *Psychology of Sport and Exercise, 22,* 83-90. doi:10.1016/j.psychsport.2015.06.009

Ntoumanis, N., & Myers, N. D. (2016). *An introduction to intermediate and advanced statistical analyses for sport and exercise scientists*. Chichester, UK: Wiley.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*, 535-569. doi:10.1080/10705510701575396

Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*, 287-312. doi:10.1207/S15328007SEM0802_7

Payne, S., Hudson, J., Akehurst, S., & Ntoumanis, N. (2013). Development and initial validation of the impression motivation in sport questionnaire-team. *Journal of Sport and Exercise Psychology, 35*, 281-298. doi:10.1123/jsep.35.3.281

Perry, J. L., Nicholls, A. R., Clough, P. J., & Crust, L. (2015). Assessing model fit: Caveats and recommendations for confirmatory factor analysis and exploratory structural equation modeling. *Measurement in Physical Education and Exercise Science*, *19*, 12-21. doi:10.1080/1091367X.2014.952370

Preacher, K. J., & Coffman, D. L. (2006, May). Computing power and minimum sample size for RMSEA [Computer software]. Retrieved from http://quantpsy.org/

Prilleltensky, I, Dietz, S., Prilleltensky, O., Myers, N. D., Rubenstein, C. L., Jin, Y., & McMahon, A. (2015). Assessing multidimensional well-being: Development and validation of the I COPPE scale. *Journal of Community Psychology, 43,* 199-226. doi:10.1002/jcop.21674

Rathwell, S., & Young, B. W. (2016). An Examination and validation of an Adapted Youth Experience Scale for university sport. *Measurement in Physical Education and Exercise Science*, *20*, 208-219. doi:10.1080/1091367X.2016.1210152

Rodriguez, D., & Audrain-McGovern, J. (2004). Team sport participation and smoking: Analysis with general growth mixture modeling. *Journal of Pediatric Psychology*, *29*, 299-308. doi:10.1093/jpepsy/jsh031

Satorra, A., & Saris, W. E. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika*, *50*, 83-90. doi:10.1007/BF02294150

Scheibehenne, B., Jamil, T., & Wagenmakers, E. J. (2016). Bayesian evidence synthesis can reconcile seemingly inconsistent results: The case of hotel towel reuse. *Psychological Science*. doi:10.1177/0956797616644081

Schellenberg, B. J. I., Gunnell, K. E., Mosewich, A. D., & Bailis, D. S. (2014). Measurement invariance of the Passion Scale across three Samples: An ESEM approach. *Measurement in Physical Education and Exercise Science*, *18*, 242-258. http://doi.org/10.1080/1091367X.2014.942453

Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (in press). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods.* doi:10.1037/met0000061

Schweizer, G., & Furley, F. (2016). Reproducible research in sport and exercise psychology: The role of sample sizes. *Psychology of Sport and Exercise*, *23*, 114-122. doi:10.1016/j.psychsport.2015.11.005

Sparks, C., Dimmock, J., Lonsdale, C., & Jackson, B. (2016). Modeling indicators and outcomes of students' perceived teacher relatedness support in high school physical education. *Psychology of Sport and Exercise*, *26*, 71-82. doi:10.1016/j.psychsport.2016.06.004

Steiger, J. H. (2001). Driving Fast in Reverse: The Relationship between Software Development, Theory, and Education in Structural Equation Modeling. *Journal of the American Statistical Association*, *96*(453), 331-338. Retrieved from http://www.jstor.org/stable/2670370

Steiger, J. H., & Lind, J. C. (1980, May). *Statistically based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Stenling, A., Ivarsson, A., Hassmén, P., & Lindwall, M. (2015). Using bifactor exploratory structural equation modeling to examine global and specific factors in measures of sports coaches' interpersonal styles. *Frontiers in psychology*, *6*. doi:10.3389/fpsyg.2015.01303

Stenling, A., Ivarsson, A., Johnson, U., & Lindwall, M. (2015). Bayesian structural equation modelling in sport and exercise psychology. *Journal of Sport & Exercise Psychology*, *37*, 410-420. doi:10.1123/jsep.2014-0330

Tamminen, K. A., Gaudreau, P., McEwen, C. E., & Crocker, P. E. (2016). Interpersonal emotion regulation among adolescent athletes: A Bayesian multilevel model predicting sport enjoyment and commitment. *Journal of Sport & Exercise Psychology*, 1-49. doi:10.1123/jsep.2015-0189

Thomas, J. R., Salazar, W., & Landers, D. M. (1991). What is missing in $p < .05$? Effect size. *Research Quarterly for Exercise and Sport*, *62*, 344-351. doi:10.1080/02701367.1991.10608733

Tomás, I., Marsh, H. W., González-Romá, V., Valls, V., & Nagengast, B. (2014). Testing measurement invariance across Spanish and English versions of the physical self-description questionnaire: an application of exploratory structural equation modeling. *Journal of Sport & Exercise Psychology*, *36*, 179-88. doi:10.1123/jsep.2013-0070

Ullrich-French, S., Cox, A. E., & Rhoades Cooper, B. (2016). Examining Combinations of Social Physique Anxiety and Motivation Regulations Using Latent Profile Analysis. *Measurement in Physical Education and Exercise Science*, *20*, 63-74. doi:10.1080/1091367X.2015.1107571

van de Schoot, R., & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. *The European Health Psychologist*, *16*(2), 75-84. Retrieved from http://ehps.net/ehp/index.php/contents

van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. G. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child Development*, *85*, 842-860. doi:10.1111/cdev.12169

van de Schoot, R., Winter, S., Ryan, O., Zondervan-Zwijnenburg, M., & Depaoli, S. (in press). A systematic review of Bayesian papers in psychology: The last 25 years. *Psychological Methods.* Retrieved from http://www.apa.org/pubs/journals/met/

Ventura, A. K., Loken, E., & Birch, L. L. (2009). Developmental trajectories of girls' BMI across childhood and adolescence. *Obesity*, *17*, 2067-2074. doi:10.1038/oby.2009.123

Viladrich, C., Appleton, P. R., Quested, E., Duda, J. L., Alcaraz, S., Heuzé, J. P., ... & Ntoumanis, N. (2013). Measurement invariance of the behavioural regulation in sport questionnaire when completed by young athletes across five European countries. *International Journal of Sport and Exercise Psychology*, *11*, 384-394. doi:10.1080/1612197X.2013.830434

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review, 14*, 779-804. doi:10.3758/BF03194105

Wagenmakers, E. J., Morey, R. D., & Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Current Directions in Psychological Science*, *25*, 169-176. doi:10.1177/0963721416643289

Wagenmakers, E. J., Verhagen, A. J., Ly, A., Matzke, D., Steingroever, H., Rouder, J. N., & Morey, R. D. (in press). The need for Bayesian hypothesis testing in psychological science. In S. O. Lilienfeld & I. Waldman (Eds.), *Psychological science under scrutiny: Recent challenges and proposed solutions*. Wiley.

Wang, J. C., Liu, W. C., Chatzisarantis, N., & Lim, C. B. S. (2010). Influence of perceived motivational climate on achievement goals in physical education: A structural equation mixture modeling analysis. *Journal of Sport and Exercise Psychology*, *32,* 324-338. doi:10.1123/jsep.32.3.324

Wang, J. C., Morin, A. J., Liu, W. C., & Chian, L. K. (2016). Predicting physical activity intention and behaviour using achievement goal theory: A person-centred analysis. *Psychology of Sport and Exercise, 23,* 13-20. doi:10.1016/j.psychsport.2015.10.004

Wang, J. C., Morin, A. J., Ryan, R. M., Liu, W. C. (2017). Students' motivational profiles in the physical education context. *Journal of Sport & Exercise Psychology.* Advance online publication. doi:10.1123/jsep.2016-0153

Wasserstein, R.L., & Lazar, N.A. (2016). The ASA's statement on p-values: Context, process, and purpose. *The American Statistician, 70,* 129-133. doi:10.1080/00031305.2016.1154108

Zhu, W. (2012). Sadly, the earth is still round (p < 0.05). *Journal of Sport and Health Science*, *1*, 9-11. doi:10.1016/j.jshs.2012.02.002

Zyphur, M. J., & Oswald, F. L. (2015). Bayesian estimation and inference: A user's guide. *Journal of Management, 41*, 390-420. doi:10.1177/0149206313501200

Footnotes

[1] The 63 residual covariances represented three (i.e., past, present and future) method effects proposed by Prilleltensky et al. (2015). For example, allowing the residual term for each of the seven 'present' items to covary resulted in 21 residual covariances (and 21 multiplied by 3, the number of method effects, equals 63). The 70 covariances in the latent variable model resulted from 21 covariances between the scores derived from responses to the seven comparison instruments and 49 covariances between the scores derived from responses to the seven comparison instruments and the seven latent variables.

[2] Reference to specific page(s) and table(s) in Hancock and French (2013) was done with permission from Gregory R. Hancock (personal communication, September 16, 2016).

[3] Readers are referred to pp. 429-430 of the User' Guide for an example.

[4] The 21 residual covariances between pairs of error terms for present items that were freely estimated in Myers, Park, et al. (2016) were fixed to zero. This reduced model fit the data as well as the more complex model, $\Delta\chi^2_R(21) = 29$, $p = .106$, and was consistent with related findings from previous research (Prilleltensky et al., 2015). Model-data fit indexes reported in Myers, Park, et al. were: $\chi^2_R(119) = 175$, $p < .001$, RMSEA = .027 ($CI_{90\%}$ =.018-.035), $p = .999$, SRMR = .017, CFI = .99, and TLI = .98. Model-data fit indexes from the reduced model in this manuscript were: $\chi^2_R(140) = 205$, $p < .001$, RMSEA = .027 ($CI_{90\%}$ =.018-.035), $p = .999$, SRMR = .018, CFI = .99, and TLI = .98.

[5] In several cases input statements provided were not necessary but such input was retained for pedagogical purposes. A complete treatment of syntax writing in M*plus* is available in Muthén and Muthén (1998-2015).

[6] The owner(s) of the relevant real dataset had reservations about making their data publicly available. Consistent with some related methodological review papers (e.g., Myers, Brincks, et al., 2012) a simulated dataset that was nearly identical to the real dataset with regard to parameter estimates was created and provided as a compromise.

Table 1

Overview of Usage of Exploratory Structural Equation Modelling in Sport and Exercise Psychology Research.

Authors	Journal	Year	Study Objective	Rotation Method	Alternative Solutions Compared
Alcaraz, et al.	RQES	2015	Investigated how behavioral regulations mediated the relation between basic psychological needs and psychological well-being and ill-being in team-sport coaches; ESEM was used to test the factor structure of responses to each variable included in the model.	Not reported	ICM-CFA
Appleton, et al.	PSE	2016	Validated responses to coach-created Empowering and Disempowering Motivational Climate Questionnaire	Target rotation	Bi-factor ESEM; ICM-CFA
Chamarro, et al.	Psicothema	2015	Presented evidence of score validity for the Passion Scale in Spanish	Geomin rotation	None
Chiu, et al.	PR	2016	Explored the factor structure of scores of the shortened version of the Leadership Scale for Sport in a sample of collegiate swimmers	Geomin rotation	ICM-CFA
Fogarty, et al.	MPEES	2016	Examined the psychometric properties of scores from the Life Orientation Test-Revised, the Sport Confidence Inventory, and the Carolina SCI	Target rotation	ICM-CFA
Gucciardi, et al.	SEPP	2012	Reviewed mental toughness measurement issues and presented a psychometric examination of the most frequently used measure of mental toughness	Geomin rotation	ICM-CFA

Gunnell & Gaudreau	PID	2015	Tested the utility of the bi-factor model to examine motivation regarding physical activity and goal progress	Target rotation	Bi-factor ESEM
Hancox, et al.	IJSEP	2015	Explored the psychometric properties of scores from the Behavioral Regulation in Sport Questionnaire adapted to dance, as well as the tenability of different scoring protocols	Not reported	ICM-CFA
Kawabata & Mallett	JSS	2013	Re-assessed the factor structure of scores from the 24-item Sport Motivation Scale-6	Geomin rotation	ICM-CFA
Koh, et al.	IJSSC	2014	Assessed the factor structure of scores from the Coaching Behavior Scale for Sport for Singaporean youth athletes	Geomin rotation	ICM-CFA
Locke & Brawley	PSE	2016	Developed and demonstrated initial validity evidence for responses to the Exercise-related Cognitive Errors Questionnaire	Geomin rotation	None
Massey, et al.	PSE	2015	Provided validity evidence for responses to the Processes of Change in Psychological Skills Training Questionnaire	Not reported	ICM-CFA
Morin & Maïano	PSE	2011	Tested the psychometric properties of responses to the short form of the Physical Self-Inventory across French adolescents	Primarily geomin rotation; Several other rotations reported in the Appendix	ICM-CFA
Morin, et al.	PSE	2016	Examined the psychometric properties of scores on the English version of the short Physical Self-Inventory	Target rotation	ICM-CFA
Myers, et al.	SEPP	2014	Presented a general case for the possible utility of exploratory bi-factor analysis in sport and exercise psychology; tested	Target rotation; Considering other rotation criteria	ICM-CFA

			the factor structure of responses to the Psychological Need Thwarting Scale		
Myers, et al.	JSEP	2011	Developed a revised version of the Coaching Efficacy Scale for Head Coaches of youth sport teams	Geomin rotation	ICM-CFA
Myers, et al.	JSEP	2012	Developed and provided initial validity evidence for measures derived from the Referee Self-Efficacy Scale	Target rotation	ICM-CFA
Myers	PSE	2013	Measured athletes' evaluations of their coach's competency within conceptual models of effective coaching	Geomin rotation; Target rotation	ICM-CFA
Nicholls, et al.	PSE	2016	Investigated a model, informed by self-regulation theories from health psychology research; ESEM was used to test the factor structure of responses to each variable included in the model.	Not reported	None
Payne, et al.	JSEP	2013	Developed and validated responses to a measure of impression motivation in team sport athletes	Geomin rotation	None
Perry, et al.	MPEES	2015	Investigated the appropriateness of using the ICM-CFA approach in sport and exercise psychology research	Geomin rotation	ICM-CFA
Rathwel & Young	MPEES	2016	Developed and validated scores from an adapted Youth Experience Scale for University Sport	Geomin rotation	ICM-CFA
Schellenberg, et al.	MPEES	2014	Examined the invariance of scores from the Passion Scale across groups of athletes, exercisers, and sports fans	Target rotation	ICM-CFA
Sparks, et al.	PSE	2016	Explored a higher-order measurement model comprising distinct relatedness-supportive teacher behaviours in physical education	Not reported	None
Stenling, et	FP	2015	Used bi-factor exploratory ESEM to	Target rotation	Bi-factor

al.			examine the psychometric properties of responses to measures of coaches' need-supportive and controlling interpersonal styles.		ESEM; ICM-CFA
Tomás, et al.	JSEP	2014	Used ESEM as an alternative approach to evaluate the measurement invariance of scores from the Spanish version of the Physical Self-Description Questionnaire	Geomin rotation	ICM-CFA
Viladrich, et al.	IJSEP	2013	Examined the factorial validity of responses to the Behavioural Regulation Sport in Questionnaire when completed by young soccer players	Target rotation	ICM-CFA

Note. PSE = Psychology of Sport and Exercise; JSEP = Journal of Sport and Exercise Psychology; MPEES = Measurement in Physical Education and Exercise Science; IJSSC = International Journal of Sports Science & Coaching; JSMS = Journal of Science and Medicine in Sport; IJSEP = International Journal of Sport and Exercise Psychology; JSS = Journal of Sports Sciences; RQES = Research Quarterly for Exercise and Sport; SEPP = Sport, Exercise, and Performance Psychology; PSI = Psicothema; PR = Psychological Reports; FP = Frontiers in Psychology; PID = Personality and Individual Difference. None of the papers provided their syntax.

Table 2

Overview of Usage of Mixture Modelling In Sport and Exercise Psychology Research.

Authors	Journal/ Book	Year	Author Labelled Analysis	Syntax Available
Ivarsson, et al.	PSE	2015	Latent class analysis	No
Lindwall, et al.	PSE	2016	Latent profile analysis	No
Ullrich-French et al.	MPEES	2016	Latent profile analysis	No
Gerber, et al.	PSE	2014	Latent profile analysis	No
Wang, et al.	JSEP	2010	Structural equation mixture model (latent profile analysis combined with full SEM mixture model)	No
Wang et al.	PSE	2016	Latent profile analysis	No
Wang et al.	JSEP	2017	Latent profile analysis with auxiliary function	No
Martinent & Nicolas	SEPP	2016	Latent profile analysis, latent profile transition analysis	No
Martinent & Decret	JASP	2015	Latent profile analysis, latent profile transition analysis	No
Louvet, et al.	PSE	2009	Latent class growth modelling	No
Louvet, et al.	JSEP	2007	Latent class growth modelling	No
Morin & Wang	Book Chapter	2016	Latent profile analysis, mixture regression model, latent transition analysis, growth mixture model	Yes
Andruff, et al.	TQMP*	2009	Latent class growth modelling	Yes
Morin, et al.,	Child Dev*	2013	Growth mixture modelling	Yes
Morin et al.,	JID*	2015	Mixture structural equation modelling	Yes

Note. Select publications from prominent sport and exercise psychology journals. * Example taken from related field to showcase syntax for mixture modelling. PSE = Psychology of Sport and Exercise; MPEES = Measurement in Physical Education and Exercise Science; JSEP = Journal of Sport and Exercise Psychology; JASP = Journal of Applied Sport Psychology; TQMP = Tutorials in Quantitative Methods for Psychology; Child Dev = Child Development; JID = Journal of Individual Differences.

Table 3

Overview of the Similarities and Differences Between Frequentist and Bayesian Statistics (Reproduced with permission from van de Schoot et al., 2014).

	Frequentist Statistics	Bayesian Statistics
Definition of the *p* value	The probability of observing the same or more extreme data assuming that the null hypothesis is true in the population	The probability of the (null) hypothesis
Large samples needed?	Usually, when normal theory-based methods are used	Not necessarily
Inclusion of prior knowledge possible?	No	Yes
Nature of the parameters in the model	Unknown but fixed	Unknown and therefore random
Population parameter	One true value	A distribution of values reflecting uncertainty
Uncertainty is defined by	The sampling distribution based on the idea of infinite repeated sampling	Probability distribution for the population parameter
Estimated intervals	Confidence interval: over an infinity of samples taken from the population, 95% of these contain the true population value	Credibility interval: a 95% probability that the population value is within the limits of the interval

Note. With recent advancements in statistics and statistical software, there are cases in which prior knowledge can be incorporated as part of frequentist statistics such as using target rotation in exploratory structural equation modelling (e.g., Myers, Ahn, & Jin, 2013) and confirmatory mixture models (e.g., Finch & Bronk, 2011).

Table 4

Overview of Usage of Bayesian Statistics in Sport and Exercise Psychology Research.

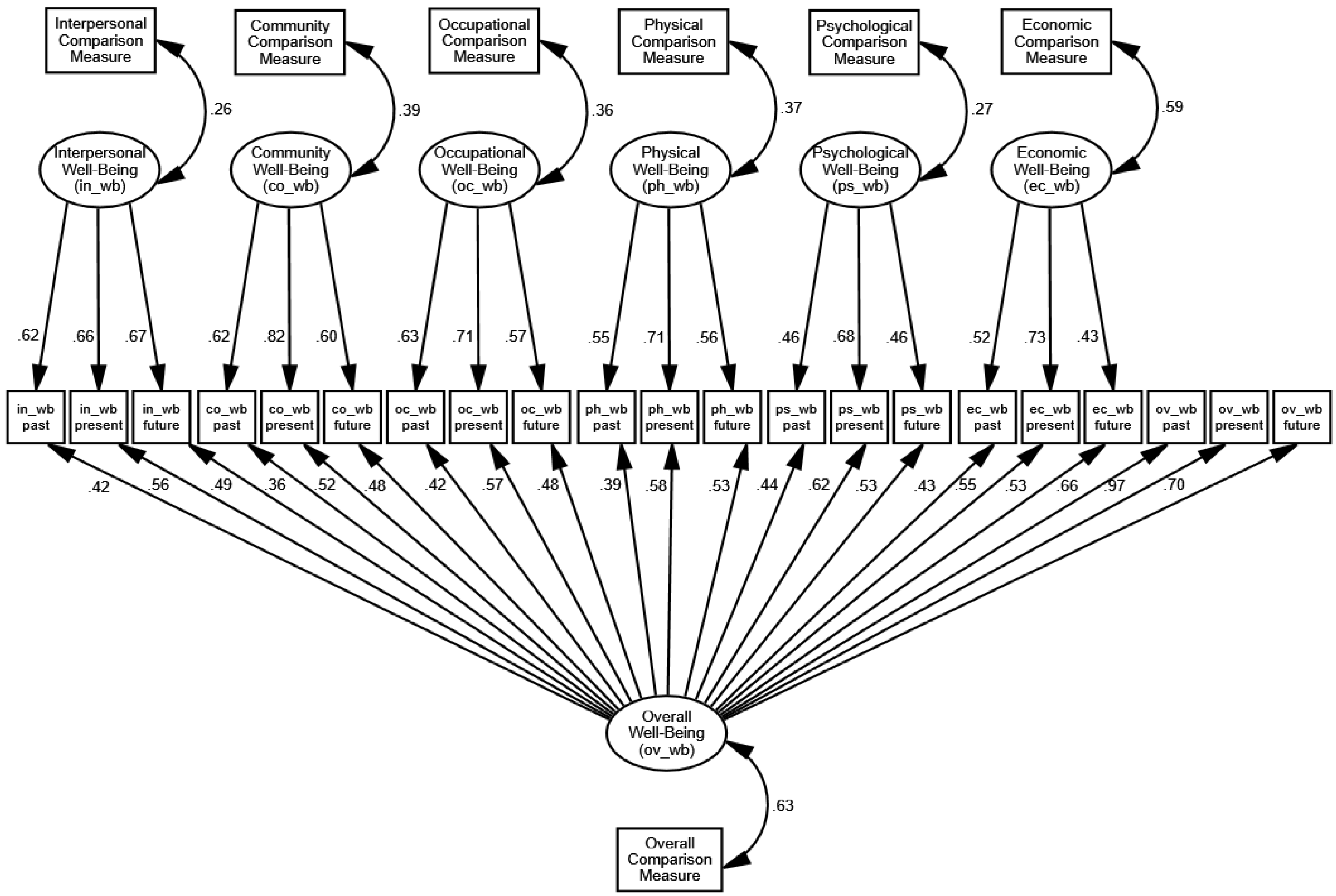
Authors	Journal	Year	Use of Bayesian	Analysis	Type of Prior	Syntax Available
Barnett et al.	PSE	2012	Estimation	Dynamic linear model	Combination of non-informative and weakly informative	No
Constantinou et al.	PSE	2014	Probabilistic graphical model	Network analysis	Weakly informative	No
Doron & Gaudreau	JSEP	2014	Estimation	Multilevel modelling	Non-informative	No
Jackson et al.	JSEP	2014	Estimation	Factor analysis	Informative	No
Mahoney et al.	JSEP	2014	Estimation	Path analysis	Combination of weakly informative and informative	No
Chan et al.	JSEP	2015	Estimation	Structural equation modelling Invariance analysis	Weakly informative	No
Gucciardi & Jackson	JSAMS	2015	Estimation	Structural equation modelling	Combination of weakly informative and informative	No*
Hodge & Gucciardi	JSEP	2015	Estimation	Path analysis	Combination of weakly informative and informative	No*
Stenling et al.	JSEP	2015	Estimation	Factor analysis	Combination of non-informative and weakly informative	No
Barnett et al.	PSE	2016	Estimation	Factor analysis	Comparison of non-informative and informative	No
Gucciardi, Peeling et al.	JSAMS	2016	Estimation	Structural equation modelling	Informative	Yes
Gucciardi, Zhang et al.	JSEP	2016	Estimation	Factor analysis Invariance analysis	Combination of non-informative and weakly informative	Yes
Howle et al.	PSE	2016	Estimation	Factor analysis	Informative	No

				Path analysis		
Niven & Markland	PSE	2016	Estimation	Factor analysis	Comparison of non-informative and informative	Yes
Noordstar et al.	PSE	2016	Estimation	Factor analysis Invariance analysis Latent growth models	No information reported	No
Tamminen et al.	JSEP	2016	Estimation	Multilevel structural equation modelling	Non-informative	Yes

Note. PSE = Psychology of Sport and Exercise; JSEP = Journal of Sport and Exercise Psychology; JSAMS = Journal of Science and Medicine in Sport. * = authors indicated that interested readers can contact them for a copy of the syntax.

Figure Captions (as a list)

Figure 1. Standardized parameter estimates commonly of primary interest (i.e., 39 pattern coefficients and 7 correlation coefficients) from Myers, Park, et al. (2016). Model parameters (e.g., variances; cross-loadings etc.) and identification constraints sometimes were omitted to reduce clutter.



Interpersonal Comparison Measure
Community Comparison Measure
Occupational Comparison Measure
Physical Comparison Measure
Psychological Comparison Measure
Economic Comparison Measure
.26
.39
.36
.37
.27
.59
Interpersonal Well-Being (in_wb)
Community Well-Being (co_wb)
Occupational Well-Being (oc_wb)
Physical Well-Being (ph_wb)
Psychological Well-Being (ps_wb)
Economic Well-Being (ec_wb)
.62
.66
.67
.62
.82
.60
.63
.71
.57
.55
.71
.56
.46
.68
.46
.52
.73
.43
in_wb past
in_wb present
in_wb future
co_wb past
co_wb present
co_wb future
oc_wb past
oc_wb present
oc_wb future
ph_wb past
ph_wb present
ph_wb future
ps_wb past
ps_wb present
ps_wb future
ec_wb past
ec_wb present
ec_wb future
ov_wb past
ov_wb present
ov_wb future
.42
.56
.49
.36
.52
.48
.42
.57
.48
.39
.58
.53
.44
.62
.53
.43
.55
.53
.66
.97
.70
Overall Well-Being (ov_wb)
.63
Overall Comparison Measure

Appendix A

Brief Demonstration 3: Software

Monte Carlo Methods: Step 1. Input for a real data analysis based on the Myers, Park, et al. (2016) example. Input file was written by the lead author of this manuscript in M*plus* 7.4 based on Example 12.7 in Muthén and Muthén (1998-2015). Annotations are in italics and denoted with a ! symbol.

```
TITLE: Demonstration 3, Step 1
! Provided a title for the analysis: Demonstration 3.

DATA: FILE = dem_3.dat;
! Specified the name of the data file: dem_3.dat.

VARIABLE:
NAMES = over_pr over_pa over_fu
        inter_pr inter_pa inter_fu
        comm_pr comm_pa comm_fu
        occup_pr occup_pa occup_fu
        physi_pr physi_pa physi_fu
        psycho_pr psycho_pa psycho_fu
        econo_pr econo_pa econo_fu
        ex_over_wb ex_int_a ex_comm
        ex_occup ex_physical ex_psych
        ex_econo;
! The columns (i.e., variables) in the data file are in the given order.

       MISSING ARE ALL (-9999);
! For all variables a value of -9999 indicates a missing value.

ANALYSIS:
      ITERATIONS=10000;
! Maximum number of iterations.

      ESTIMATOR=MLR;
! Maximum likelihood parameter estimates with a chi-square test statistic
! and standard errors that are robust to conditional non-normality.

      ROTATION = Target(orthogonal);
! Orthogonal Target rotation.

MODEL:
      Ov BY over_pr-econo_fu(*t);
! ...BY: provided name for latent variable.
! BY: "measured by".
! BY...: identified indicator variables, in this case over_pr through
! econo_fu, for identified latent variable.
! (*t): defines a set of factors.

      In BY over_pr-econo_fu
```

```
      over_pr~0 over_pa~0 over_fu~0
      inter_pr~1.25 inter_pa~1.25 inter_fu~1.25
      comm_pr~0 comm_pa~0 comm_fu~0
      occup_pr~0 occup_pa~0 occup_fu~0
      physi_pr~0 physi_pa~0 physi_fu~0
      psycho_pr~0 psycho_pa~0 psycho_fu~0
      econo_pr~0 econo_pa~0 econo_fu~0(*t);
! ~value: targeted value.

      Co BY over_pr-econo_fu
      over_pr~0 over_pa~0 over_fu~0
      inter_pr~0 inter_pa~0 inter_fu~0
      comm_pr~1.25 comm_pa~1.25 comm_fu~1.25
      occup_pr~0 occup_pa~0 occup_fu~0
      physi_pr~0 physi_pa~0 physi_fu~0
      psycho_pr~0 psycho_pa~0 psycho_fu~0
      econo_pr~0 econo_pa~0 econo_fu~0(*t);

      Oc BY over_pr-econo_fu
      over_pr~0 over_pa~0 over_fu~0
      inter_pr~0 inter_pa~0 inter_fu~0
      comm_pr~0 comm_pa~0 comm_fu~0
      occup_pr~1.25 occup_pa~1.25 occup_fu~1.25
      physi_pr~0 physi_pa~0 physi_fu~0
      psycho_pr~0 psycho_pa~0 psycho_fu~0
      econo_pr~0 econo_pa~0 econo_fu~0(*t);

      Ph BY over_pr-econo_fu
      over_pr~0 over_pa~0 over_fu~0
      inter_pr~0 inter_pa~0 inter_fu~0
      comm_pr~0 comm_pa~0 comm_fu~0
      occup_pr~0 occup_pa~0 occup_fu~0
      physi_pr~1.25 physi_pa~1.25 physi_fu~1.25
      psycho_pr~0 psycho_pa~0 psycho_fu~0
      econo_pr~0 econo_pa~0 econo_fu~0(*t);

      Ps BY over_pr-econo_fu
      over_pr~0 over_pa~0 over_fu~0
      inter_pr~0 inter_pa~0 inter_fu~0
      comm_pr~0 comm_pa~0 comm_fu~0
      occup_pr~0 occup_pa~0 occup_fu~0
      physi_pr~0 physi_pa~0 physi_fu~0
      psycho_pr~1.25 psycho_pa~1.25 psycho_fu~1.25
      econo_pr~0 econo_pa~0 econo_fu~0(*t);

      Ec BY over_pr-econo_fu
      over_pr~0 over_pa~0 over_fu~0
      inter_pr~0 inter_pa~0 inter_fu~0
      comm_pr~0 comm_pa~0 comm_fu~0
      occup_pr~0 occup_pa~0 occup_fu~0
      physi_pr~0 physi_pa~0 physi_fu~0
      psycho_pr~0 psycho_pa~0 psycho_fu~0
      econo_pr~1.25 econo_pa~1.25 econo_fu~1.25(*t);

      Ov with ex_over_wb;
      Ov with ex_int_a;
      Ov with ex_comm;
```

```
      Ov with ex_occup;
      Ov with ex_physical;
      Ov with ex_psych;
      Ov with ex_econo;
! with: "co-varies" with; covariance between pairs of variables.

      In with ex_over_wb;
      In with ex_int_a;
      In with ex_comm;
      In with ex_occup;
      In with ex_physical;
      In with ex_psych;
      In with ex_econo;

      Co with ex_over_wb;
      Co with ex_int_a;
      Co with ex_comm;
      Co with ex_occup;
      Co with ex_physical;
      Co with ex_psych;
      Co with ex_econo;

      Oc with ex_over_wb;
      Oc with ex_int_a;
      Oc with ex_comm;
      Oc with ex_occup;
      Oc with ex_physical;
      Oc with ex_psych;
      Oc with ex_econo;

      Ph with ex_over_wb;
      Ph with ex_int_a;
      Ph with ex_comm;
      Ph with ex_occup;
      Ph with ex_physical;
      Ph with ex_psych;
      Ph with ex_econo;

      Ps with ex_over_wb;
      Ps with ex_int_a;
      Ps with ex_comm;
      Ps with ex_occup;
      Ps with ex_physical;
      Ps with ex_psych;
      Ps with ex_econo;

      Ec with ex_over_wb;
      Ec with ex_int_a;
      Ec with ex_comm;
      Ec with ex_occup;
      Ec with ex_physical;
      Ec with ex_psych;
      Ec with ex_econo;

      ex_over_wb with ex_int_a;
      ex_over_wb with ex_comm;
      ex_over_wb with ex_occup;
```

```
ex_over_wb with ex_physical;
ex_over_wb with ex_psych;
ex_over_wb with ex_econo;

ex_int_a with ex_comm;
ex_int_a with ex_occup;
ex_int_a with ex_physical;
ex_int_a with ex_psych;
ex_int_a with ex_econo;

ex_comm with ex_occup;
ex_comm with ex_physical;
ex_comm with ex_psych;
ex_comm with ex_econo;

ex_occup with ex_physical;
ex_occup with ex_psych;
ex_occup with ex_econo;

ex_physical with ex_psych;
ex_physical with ex_econo;

ex_psych with ex_econo;

over_pa with inter_pa;
over_pa with comm_pa;
over_pa with occup_pa;
over_pa with physi_pa;
over_pa with psycho_pa;
over_pa with econo_pa;

inter_pa with comm_pa;
inter_pa with occup_pa;
inter_pa with physi_pa;
inter_pa with psycho_pa;
inter_pa with econo_pa;

comm_pa with occup_pa;
comm_pa with physi_pa;
comm_pa with psycho_pa;
comm_pa with econo_pa;

occup_pa with physi_pa;
occup_pa with psycho_pa;
occup_pa with econo_pa;

physi_pa with psycho_pa;
physi_pa with econo_pa;

psycho_pa with econo_pa;

over_fu with inter_fu;
over_fu with comm_fu;
over_fu with occup_fu;
over_fu with physi_fu;
over_fu with psycho_fu;
over_fu with econo_fu;
```

```
      inter_fu with comm_fu;
      inter_fu with occup_fu;
      inter_fu with physi_fu;
      inter_fu with psycho_fu;
      inter_fu with econo_fu;

      comm_fu with occup_fu;
      comm_fu with physi_fu;
      comm_fu with psycho_fu;
      comm_fu with econo_fu;

      occup_fu with physi_fu;
      occup_fu with psycho_fu;
      occup_fu with econo_fu;

      physi_fu with psycho_fu;
      physi_fu with econo_fu;

      psycho_fu with econo_fu;

OUTPUT: SAMPSTAT STANDARDIZED tech1;
! SAMPSTAT: requested sample statistics for data being analyzed.
! STAND: requested standardized parameter estimates and their standard
! errors.
! tech1: requested arrays containing parameter specifications and starting
! values for all freely estimated parameters in the model

SAVEDATA: ESTIMATES = final model estimates.dat;
! Specified the name of the file, final model estimates, in which parameter
! estimates will be saved.
```

Appendix B

Brief Demonstration 3: Software

Monte Carlo Methods: Step 2. Input for a Monte Carlo simulation study where parameter estimates saved from Step 1 (see Appendix B) are used for population parameter values for data generation (i.e., replications) and coverage. Input file was written by the lead author of this manuscript in M*plus* 7.4 based on Example 12.8 in Muthén and Muthén (1998-2015). Annotations are provided for commands not explained in Appendix B and are in italics and denoted with a ! symbol.

```
TITLE: Demonstration 3, Step 2

MONTECARLO:
! A Monte Carlo study ensues.
      NAMES = over_pr over_pa over_fu
      inter_pr inter_pa inter_fu
      comm_pr comm_pa comm_fu
      occup_pr occup_pa occup_fu
      physi_pr physi_pa physi_fu
      psycho_pr psycho_pa psycho_fu
      econo_pr econo_pa econo_fu
      ex_over_wb ex_int_a ex_comm
      ex_occup ex_physical ex_psych
      ex_econo;

      NOBSERVATIONS = 1000;
! Desired sample size for each replication.
! Only this line of code needs to be changed (i.e., NOBSERVATIONS = 500 or
! NOBSERVATIONS = 250) to reproduce the other two results.

      NREPS = 10000;
! Number of replications to be drawn.

      SEED = 82872;
! Provides a starting place for the random draws.

      POPULATION = final model estimates.dat;
! Names the data set that contains population parameter values.

      COVERAGE = final model estimates.dat;
! Names the data set that contains population parameter values.

ANALYSIS:
            ESTIMATOR=MLR;
            ROTATION = Target(orthogonal);

MODEL POPULATION:
```

```
! Provides the population model.

      Ov BY over_pr-econo_fu;
      In BY over_pr-econo_fu;
      Co BY over_pr-econo_fu;
      Oc BY over_pr-econo_fu;
      Ph BY over_pr-econo_fu;
      Ps BY over_pr-econo_fu;
      Ec BY over_pr-econo_fu;

      Ov with ex_over_wb;
      Ov with ex_int_a;
      Ov with ex_comm;
      Ov with ex_occup;
      Ov with ex_physical;
      Ov with ex_psych;
      Ov with ex_econo;

      In with ex_over_wb;
      In with ex_int_a;
      In with ex_comm;
      In with ex_occup;
      In with ex_physical;
      In with ex_psych;
      In with ex_econo;

      Co with ex_over_wb;
      Co with ex_int_a;
      Co with ex_comm;
      Co with ex_occup;
      Co with ex_physical;
      Co with ex_psych;
      Co with ex_econo;

      Oc with ex_over_wb;
      Oc with ex_int_a;
      Oc with ex_comm;
      Oc with ex_occup;
      Oc with ex_physical;
      Oc with ex_psych;
      Oc with ex_econo;

      Ph with ex_over_wb;
      Ph with ex_int_a;
      Ph with ex_comm;
      Ph with ex_occup;
      Ph with ex_physical;
      Ph with ex_psych;
      Ph with ex_econo;

      Ps with ex_over_wb;
      Ps with ex_int_a;
      Ps with ex_comm;
      Ps with ex_occup;
      Ps with ex_physical;
      Ps with ex_psych;
      Ps with ex_econo;
```

```
Ec with ex_over_wb;
Ec with ex_int_a;
Ec with ex_comm;
Ec with ex_occup;
Ec with ex_physical;
Ec with ex_psych;
Ec with ex_econo;

ex_over_wb with ex_int_a;
ex_over_wb with ex_comm;
ex_over_wb with ex_occup;
ex_over_wb with ex_physical;
ex_over_wb with ex_psych;
ex_over_wb with ex_econo;

ex_int_a with ex_comm;
ex_int_a with ex_occup;
ex_int_a with ex_physical;
ex_int_a with ex_psych;
ex_int_a with ex_econo;

ex_comm with ex_occup;
ex_comm with ex_physical;
ex_comm with ex_psych;
ex_comm with ex_econo;

ex_occup with ex_physical;
ex_occup with ex_psych;
ex_occup with ex_econo;

ex_physical with ex_psych;
ex_physical with ex_econo;

ex_psych with ex_econo;

over_pa with inter_pa;
over_pa with comm_pa;
over_pa with occup_pa;
over_pa with physi_pa;
over_pa with psycho_pa;
over_pa with econo_pa;

inter_pa with comm_pa;
inter_pa with occup_pa;
inter_pa with physi_pa;
inter_pa with psycho_pa;
inter_pa with econo_pa;

comm_pa with occup_pa;
comm_pa with physi_pa;
comm_pa with psycho_pa;
comm_pa with econo_pa;

occup_pa with physi_pa;
occup_pa with psycho_pa;
occup_pa with econo_pa;
```

```
      physi_pa with psycho_pa;
      physi_pa with econo_pa;

      psycho_pa with econo_pa;

      over_fu with inter_fu;
      over_fu with comm_fu;
      over_fu with occup_fu;
      over_fu with physi_fu;
      over_fu with psycho_fu;
      over_fu with econo_fu;

      inter_fu with comm_fu;
      inter_fu with occup_fu;
      inter_fu with physi_fu;
      inter_fu with psycho_fu;
      inter_fu with econo_fu;

      comm_fu with occup_fu;
      comm_fu with physi_fu;
      comm_fu with psycho_fu;
      comm_fu with econo_fu;

      occup_fu with physi_fu;
      occup_fu with psycho_fu;
      occup_fu with econo_fu;

      physi_fu with psycho_fu;
      physi_fu with econo_fu;

      psycho_fu with econo_fu;


MODEL:
! Provides the model to be fit to each replication that is generated.

      Ov BY over_pr-econo_fu(*t);

      In BY over_pr-econo_fu
      over_pr~0 over_pa~0 over_fu~0
      inter_pr~1.25 inter_pa~1.25 inter_fu~1.25
      comm_pr~0 comm_pa~0 comm_fu~0
      occup_pr~0 occup_pa~0 occup_fu~0
      physi_pr~0 physi_pa~0 physi_fu~0
      psycho_pr~0 psycho_pa~0 psycho_fu~0
      econo_pr~0 econo_pa~0 econo_fu~0(*t);

      Co BY over_pr-econo_fu
      over_pr~0 over_pa~0 over_fu~0
      inter_pr~0 inter_pa~0 inter_fu~0
      comm_pr~1.25 comm_pa~1.25 comm_fu~1.25
      occup_pr~0 occup_pa~0 occup_fu~0
      physi_pr~0 physi_pa~0 physi_fu~0
      psycho_pr~0 psycho_pa~0 psycho_fu~0
      econo_pr~0 econo_pa~0 econo_fu~0(*t);
```

```
Oc BY over_pr-econo_fu
over_pr~0 over_pa~0 over_fu~0
inter_pr~0 inter_pa~0 inter_fu~0
comm_pr~0 comm_pa~0 comm_fu~0
occup_pr~1.25 occup_pa~1.25 occup_fu~1.25
physi_pr~0 physi_pa~0 physi_fu~0
psycho_pr~0 psycho_pa~0 psycho_fu~0
econo_pr~0 econo_pa~0 econo_fu~0(*t);

Ph BY over_pr-econo_fu
over_pr~0 over_pa~0 over_fu~0
inter_pr~0 inter_pa~0 inter_fu~0
comm_pr~0 comm_pa~0 comm_fu~0
occup_pr~0 occup_pa~0 occup_fu~0
physi_pr~1.25 physi_pa~1.25 physi_fu~1.25
psycho_pr~0 psycho_pa~0 psycho_fu~0
econo_pr~0 econo_pa~0 econo_fu~0(*t);

Ps BY over_pr-econo_fu
over_pr~0 over_pa~0 over_fu~0
inter_pr~0 inter_pa~0 inter_fu~0
comm_pr~0 comm_pa~0 comm_fu~0
occup_pr~0 occup_pa~0 occup_fu~0
physi_pr~0 physi_pa~0 physi_fu~0
psycho_pr~1.25 psycho_pa~1.25 psycho_fu~1.25
econo_pr~0 econo_pa~0 econo_fu~0(*t);

Ec BY over_pr-econo_fu
over_pr~0 over_pa~0 over_fu~0
inter_pr~0 inter_pa~0 inter_fu~0
comm_pr~0 comm_pa~0 comm_fu~0
occup_pr~0 occup_pa~0 occup_fu~0
physi_pr~0 physi_pa~0 physi_fu~0
psycho_pr~0 psycho_pa~0 psycho_fu~0
econo_pr~1.25 econo_pa~1.25 econo_fu~1.25(*t);

Ov with ex_over_wb;
Ov with ex_int_a;
Ov with ex_comm;
Ov with ex_occup;
Ov with ex_physical;
Ov with ex_psych;
Ov with ex_econo;

In with ex_over_wb;
In with ex_int_a;
In with ex_comm;
In with ex_occup;
In with ex_physical;
In with ex_psych;
In with ex_econo;

Co with ex_over_wb;
Co with ex_int_a;
Co with ex_comm;
Co with ex_occup;
Co with ex_physical;
```

```
Co with ex_psych;
Co with ex_econo;

Oc with ex_over_wb;
Oc with ex_int_a;
Oc with ex_comm;
Oc with ex_occup;
Oc with ex_physical;
Oc with ex_psych;
Oc with ex_econo;

Ph with ex_over_wb;
Ph with ex_int_a;
Ph with ex_comm;
Ph with ex_occup;
Ph with ex_physical;
Ph with ex_psych;
Ph with ex_econo;

Ps with ex_over_wb;
Ps with ex_int_a;
Ps with ex_comm;
Ps with ex_occup;
Ps with ex_physical;
Ps with ex_psych;
Ps with ex_econo;

Ec with ex_over_wb;
Ec with ex_int_a;
Ec with ex_comm;
Ec with ex_occup;
Ec with ex_physical;
Ec with ex_psych;
Ec with ex_econo;

ex_over_wb with ex_int_a;
ex_over_wb with ex_comm;
ex_over_wb with ex_occup;
ex_over_wb with ex_physical;
ex_over_wb with ex_psych;
ex_over_wb with ex_econo;

ex_int_a with ex_comm;
ex_int_a with ex_occup;
ex_int_a with ex_physical;
ex_int_a with ex_psych;
ex_int_a with ex_econo;

ex_comm with ex_occup;
ex_comm with ex_physical;
ex_comm with ex_psych;
ex_comm with ex_econo;

ex_occup with ex_physical;
ex_occup with ex_psych;
ex_occup with ex_econo;
```

```
ex_physical with ex_psych;
ex_physical with ex_econo;

ex_psych with ex_econo;

over_pa with inter_pa;
over_pa with comm_pa;
over_pa with occup_pa;
over_pa with physi_pa;
over_pa with psycho_pa;
over_pa with econo_pa;

inter_pa with comm_pa;
inter_pa with occup_pa;
inter_pa with physi_pa;
inter_pa with psycho_pa;
inter_pa with econo_pa;

comm_pa with occup_pa;
comm_pa with physi_pa;
comm_pa with psycho_pa;
comm_pa with econo_pa;

occup_pa with physi_pa;
occup_pa with psycho_pa;
occup_pa with econo_pa;

physi_pa with psycho_pa;
physi_pa with econo_pa;

psycho_pa with econo_pa;

over_fu with inter_fu;
over_fu with comm_fu;
over_fu with occup_fu;
over_fu with physi_fu;
over_fu with psycho_fu;
over_fu with econo_fu;

inter_fu with comm_fu;
inter_fu with occup_fu;
inter_fu with physi_fu;
inter_fu with psycho_fu;
inter_fu with econo_fu;

comm_fu with occup_fu;
comm_fu with physi_fu;
comm_fu with psycho_fu;
comm_fu with econo_fu;

occup_fu with physi_fu;
occup_fu with psycho_fu;
occup_fu with econo_fu;

physi_fu with psycho_fu;
physi_fu with econo_fu;
```

```
      psycho_fu with econo_fu;

OUTPUT:     tech1 tech9;
! tech9: Print error messages related to convergence for each replication.
```

Appendix C

Brief Demonstration 3: Software

Monte Carlo Methods: Step 2. Truncated output identifying the power estimation value for each focal parameter when sample size equalled 250. The right-most column labelled, % Sig Coeff, provides power estimation values.

```
MODEL RESULTS

                              ESTIMATES              S. E.     M. S. E.  95%  % Sig
                 Population   Average   Std. Dev.   Average              Cover Coeff
 OV       BY
  OVER_PR             1.987     1.9532     0.2478     0.1390     0.0625 0.931 0.998
  OVER_PA             1.427     1.4010     0.1772     0.1326     0.0321 0.927 0.999
  OVER_FU             1.458     1.4324     0.1723     0.1237     0.0303 0.930 0.999

 IN       BY
  INTER_PR            1.456     1.4476     0.0944     0.0955     0.0090 0.947 1.000
  INTER_PA            1.325     1.3187     0.1063     0.1051     0.0113 0.943 1.000
  INTER_FU            1.346     1.3414     0.0925     0.0916     0.0086 0.944 1.000

 CO       BY
  COMM_PR             1.688     1.6772     0.0857     0.0851     0.0075 0.939 1.000
  COMM_PA             1.377     1.3722     0.1106     0.1092     0.0123 0.944 1.000
  COMM_FU             1.226     1.2227     0.0909     0.0913     0.0083 0.948 1.000

 OC       BY
  OCCUP_PR            1.930     1.9202     0.1128     0.1135     0.0128 0.947 1.000
  OCCUP_PA            1.653     1.6463     0.1289     0.1283     0.0167 0.947 1.000
  OCCUP_FU            1.382     1.3797     0.1075     0.1089     0.0116 0.950 1.000

 PH       BY
  PHYSI_PR            1.472     1.4620     0.0918     0.0904     0.0085 0.940 1.000
  PHYSI_PA            1.229     1.2252     0.1168     0.1172     0.0137 0.948 1.000
  PHYSI_FU            1.107     1.1050     0.0937     0.0929     0.0088 0.946 1.000

 PS       BY
  PSYCHO_PR           1.508     1.4993     0.1089     0.1135     0.0119 0.951 1.000
```

```
 PSYCHO_PA           1.018     1.0153     0.1198     0.1232     0.0144 0.950 1.000
 PSYCHO_FU           0.952     0.9500     0.0971     0.0996     0.0094 0.951 1.000

EC        BY
 ECONO_PR            1.833     1.8155     0.1086     0.1062     0.0121 0.936 1.000
 ECONO_PA            1.249     1.2458     0.1253     0.1275     0.0157 0.948 1.000
 ECONO_FU            0.981     0.9839     0.1072     0.1064     0.0115 0.946 1.000

OV        WITH
 EX_OVER_WB          4.757     4.7067     0.4399     0.4244     0.1960 0.936 0.999

IN        WITH
 EX_INT_A            4.609     4.5880     1.0991     1.1062     1.2083 0.951 0.986

CO        WITH
 EX_COMM             0.402     0.3967     0.0573     0.0568     0.0033 0.945 1.000

OC        WITH
 EX_OCCUP            2.679     2.6537     0.4307     0.4331     0.1861 0.946 1.000

PH        WITH
 EX_PHYSICA          3.811     3.7911     0.6919     0.6815     0.4791 0.945 1.000

PS        WITH
 EX_PSYCH            2.527     2.4959     0.5475     0.5488     0.3007 0.947 0.994

EC        WITH
 EX_ECONO            1.340     1.3306     0.1164     0.1179     0.0136 0.948 1.000
```

Appendix D

Brief Demonstration 3: Software

Monte Carlo Methods: Step 1. Readers can download the file named dem_3.dat (available on the online supplemental materials) and try running the syntax provided in Appendix A and Appendix B themselves. Changing the NREPS command from NREPS=10000 to NREPS=1000 should significantly reduce computational time (and should be sufficient for demonstration purposes) with only minor changes in the results.