On CTA-PLS corrections applied on sports performance

Mario Angelelli[1], Mattia Cefis[2], Enrico Ciavolino[1], Maurizio Carpita[2]

[1] *Department of Human and Social Sciences, University of Salento*

[2] *Department of Economics and Management, University of Brescia*

Abstract

This work explores a novel approach for assessing causal directions in measurement models and structural equation models with higher-order constructs. This extension of CTA-PLS incorporates different methods for controlling errors in multiple hypothesis testing, adapting them to the soft modeling context and highlighting their relevance during exploratory model construction. The *CTA-PLS corrections* method is applied to a second-order construct for performance assessment in sports analytics.

Keywords: CTA, PLS-SEM, Higher-order constructs, Multiple Hypothesis Testing

1 Introduction

Structural equation modeling (SEM) constitutes a class of graphical models that formalize structural connections between latent variables (LVs) and observed indicators. Estimation procedures in SEMs combine the structural (or inner) model, capturing cause-and-effect relationships between LVs, with a measurement (outer) model, relating manifest variables (MVs) to their corresponding constructs. Measurement modes distinguish how LVs and MVs interact: in reflective mode (Mode A), indicators are viewed as *effects* of the latent variable, akin to factor analysis, whereas in the formative mode, indicators are treated as *causes* defining the LV. The measurement models are formalized as follows:

$$\text{Mode A: } x_{pq} = \lambda_{pq} \cdot \xi_q + \varepsilon_{pq}, \quad \text{Mode B: } \xi_q = \sum_{p=1}^{p_q} w_{pq} \cdot x_{pq} + \delta_q \tag{1}$$

where λ_{pq} and w_{pq} are the loading and the weight linking the LV ξ_q and its p-th ($p = 1, \dots, p_q$) MV x_{pq}, respectively, while ε_{pq} and δ_q represent error terms.

Reflective and formative measurement models have distinct implications for SEM interpretation and application [1]. In reflective mode, the construct exists independently of its indicators, which are interchangeable; in contrast, formative mode defines the latent variable (LV) as a combination of its manifest variables (MVs), making each indicator essential to the construct's meaning since the omission of a formative MV alters the construct itself. Consequently, validity metrics such as composite reliability and AVE are unsuitable for formative models, where no measurement error is attributed to indicators, as is manifest in (1). Instead, validity is assessed through alternative methods, such as collinearity diagnostics. Furthermore, model misspecification can introduce bias into structural estimates [2].

The previous observation underscores the need to address measurement mode misspecification. While the choice of measurement modes is typically guided by theory, the occurrence of a common LV underlying MVs in reflective modes entails constraints or conditions, e.g., unidimensionality and relations between empirical correlations. Such correlation constraints do not apply to formative measurements, and this difference enables statistical testing for measurement modes [3]. The test statistics are derived from *tetrads*

$$\tau_{abcd} := \sigma_{ab} \cdot \sigma_{cd} - \sigma_{ac} \cdot \sigma_{bd}, \tag{2}$$

where $\Sigma = (\sigma_{ij})$ is the model-implied covariance matrix. For reflective measurements, all tetrads should *simultaneously* vanish in the population covariance matrix, which leads to the following null hypothesis H_0 : "for all $X \in \mathscr{J} : \tau_X = 0$", where $\mathscr{J}$ denotes the set of algebraically independent tetrads.

The statistical methods for testing H_0 should align with the estimation method, which is guided by data sample characteristics and research objectives. Factor- or covariance-based SEM (CB-SEM) relies on distributional assumptions, and it requires large datasets if they hold asymptotically. It focuses on confirmatory objectives, optimizing the fit between the model-implied and empirical correlation matrices. On the other hand, composite- or variance-based SEM (PLS-SEM) is a non-parametric approach that uses Partial Least Squares (PLS) algorithms to iteratively estimate both the inner and outer models until convergence. PLS-SEM is less sensitive to distributional assumptions and model misspecification; it does not require a large sample size and combines explanatory and predictive purposes. In line with these methodological differences, the original Confirmatory Tetrad Analysis (CTA) [1] is an *omnibus* test asymptotically distributed as a χ^2-distribution and aligns with the CB-SEM framework, focusing on the model-implied covariance matrix. CTA-PLS [4] uses bootstrapping with observed (partial) correlations, in agreement with the non-parametric nature of PLS-SEM, and tests *each* LV's measurement mode individually. While CTA-PLS is more flexible and suited for exploratory purposes, it involves *multiple hypothesis testing*, requiring correction methods to control errors and false discoveries. This extends beyond the measurement model and involves the structural model when it includes *higher-order* latent variables. Indeed, similar issues may arise in such models related to causal directions (reflective vs. formative) connecting lower- and higher-order latent variables.

This work addresses the role of correction methods in CTA-PLS for both first-order and second- or higher-order latent variables. Our contribution aims to support researchers in the selection of measurement modes that better align with empirical data and theoretical frameworks. Indeed, in human and social sciences, different uncertainty sources should be taken into account in moving from abstract constructs to proxies and latent variables and, finally, to the measured and observed indicators; in particular, there is no neat separation between different measurement models, so their selection should combine conceptualization, research objectives, and empirical evidence [5]. The present analysis uncovers differences between correction methods in multiple hypothesis testing, contextualizing the CTA-PLS corrections in an empirical application to provide insights for selecting measurement modes in first- and higher-order SEM.

2 The CTA-PLS corrections for First-order and Higher-order constructs

Here, we summarize the key concepts and steps in corrected CTA-PLS; for detailed discussions, see [6; 7]. Two primary approaches to controlling errors in multiple hypothesis testing are familywise error rate (FWER) and false discovery rate (FDR). FWER controls the probability of *at least* one type-I error, with the Bonferroni correction as a central method. Bonferroni applies the adjustment $\alpha \mapsto \frac{\alpha}{\#\mathscr{J}}$, or equivalently the scaling of individual p-values as $p_j \mapsto p_j \cdot (\#\mathscr{J})$. While Bonferroni correction aligns with the conjunctive form of the CTA-PLS null hypothesis (requiring *all* tetrads to vanish), it may not suit exploratory SEM analyses where the distinction between reflective and formative constructs is not sharp. As verified on simulated and real-world data [7; 6], this approach is overly conservative for CTA-PLS, often reducing test power in exploratory contexts. On the contrary, the Benjamini-Hochberg (BH) procedure is a widely used FDR method that ranks p-values $p_{(r)}$ and adjusts them as:

$$\tilde{p}_{(r)} = \frac{(\#\mathscr{J}) \cdot p_{(r)}}{r}, \quad \tilde{p}_{(\#\mathscr{J}+1)} = 1. \tag{3}$$

Hypotheses are rejected for $1 \leq r \leq \max\{s : \tilde{p}_{(s)} \leq \alpha\}$ in the ranked p-value list. To address dependencies between tests, the Benjamini-Yekutieli (BY) procedure modifies the BH adjustment by introducing a correction coefficient $c(r)$:

$$\tilde{p}_{(r)} = \frac{(\#\mathscr{J}) \cdot p_{(r)}}{c(r) \cdot r}. \tag{4}$$

2.1 First-order constructs

The proposed CTA-PLS corrections approach integrates the methodology from [4] with a comparative analysis of results from different correction methods.

Model validation should first be performed to identify potential issues in the hypothesized model, such as low reliability or AVE in reflective models and high collinearity in formative models. Then, the set $\mathscr{J}$ of relevant tetrads must be identified. Algebraic procedures can isolate non-redundant tetrads using software tools. Since each latent variable (LV) in CTA-PLS requires at least four manifest variables (MVs), researchers must establish criteria for handling LVs with fewer than four MVs. Options include borrowing additional items based on model-implied tetrads [4] or avoiding any extension [8, Ch. 3]. Bootstrapping is used to obtain an empirical distribution for each tetrad in $\mathscr{J}$. From this, we get a bootstrap t-statistics

$$t_j = \frac{\hat{\tau}_j}{\mathsf{se}_{\text{boot}}(\tau_j)}, \quad j \in \mathscr{J} \tag{5}$$

where $\hat{\tau}_j$ is the bootstrap estimate of the j-th tetrad, and $\mathsf{se}_{\text{boot}}(\tau_j)$ is the bootstrap standard error. Under the CTA-PLS null hypothesis, population tetrads vanish, and p-values derived from (5) are used to identify significant tetrads supporting rejection of H_0 in the subsequent correction phase. Such p-values are adjusted using methods such as Bonferroni, Benjamini-Hochberg, or Benjamini-Yekutieli. The Bonferroni correction of p-values (or the corresponding confidence intervals) adheres to the conjunctive form of H_0 and could be easily interpreted when the correction methods agree; however, it provides conservative control, while Benjamini-Hochberg is more powerful, particularly for small sample sizes or complex models [6]. So, when the methods disagree, researchers may prefer Benjamini-Hochberg when substantial evidence is needed to favor Mode A over Mode B measurement. Resampling procedures also offer insights into the consistency of corrections across different sample sizes, informing the agreement of the measurement model selection for different sampling procedures. Moreover, identifying significant tetrads through corrected p-values can guide revisions to the hypothesized model, improving its alignment with observed data.

2.2 Higher-order constructs

The CTA-PLS corrections procedure extends to structural models to select causality directions among multiple lower-order constructs (LOCs) and higher-order constructs (HOCs). As in Section 2.1, higher-order CTA-PLS corrections requires at least four LOCs to define HOC tetrads.

CTA-PLS correction is applied hierarchically, starting with the measurement models and proceeding to analyze connections between LOCs and HOCs until the highest-order LVs. After conducting CTA-PLS corrections for measurement models, HOCs need to be estimated; among the different approaches, we adopted the hybrid two-step method [9]. Following an initial estimation based on repeated indicators, the LOCs' scores are retrieved and used as proxy indicators for such LOCs. Furthermore, we get the estimated covariance matrix S between LOCs, which will be used in the next steps.

At this point, we can adapt the steps presented in Section 2.1 with such proxy indicators to formalize the null hypothesis for the second-order LV. We consider the higher-order tetrads obtained from the covariance matrix S between LOCs that was derived in the previous step

$$\hat{\tau}_{abcd} = S_{\hat{\xi}_a^{\text{LOC}},\hat{\xi}_b^{\text{LOC}}} \cdot S_{\hat{\xi}_c^{\text{LOC}},\hat{\xi}_d^{\text{LOC}}} - S_{\hat{\xi}_a^{\text{LOC}},\hat{\xi}_c^{\text{LOC}}} \cdot S_{\hat{\xi}_b^{\text{LOC}},\hat{\xi}_d^{\text{LOC}}}, \quad (abcd) \in \mathscr{K} \tag{6}$$

where $S_{\hat{\xi}_a^{\text{LOC}},\hat{\xi}_b^{\text{LOC}}}$ is the covariance between the scores $\hat{\xi}_a^{\text{LOC}}$ and $\hat{\xi}_a^{\text{LOC}}$ of two LOCs associated with the same HOC. Here, $\mathscr{K}$ denotes the label set for all HOCs with at least four LOCs. This allows us to formalize the higher-order null hypothesis:

$$H_0^{\text{II}}: \text{ for all } k \in \mathscr{K} : \tau_K = 0. \tag{7}$$

Finally, a bootstrapping procedure generates sample distributions for higher-order tetrads, including bootstrap estimates, standard errors, t-statistics, and, hence, p-values. These are then subjected to the correction procedures outlined above.

3 An application in sports performance

The CTA-PLS multiple test was applied within a second-order hierarchical model PLS-SEM framework to evaluate football players' performance, focusing on strikers. A consistent PLS-

SEM is proposed as an innovative composite indicator for measuring strikers' macro attacking performance in sports analytics. The dataset used for the analysis includes data collected in December 2024 from *sofifa* experts [7], assessing the performance of 2,598 strikers (i.e., the rows of the dataset) from 20 international leagues across 26 abilities (i.e., MVs, the columns of our dataset), on a discrete scale between 0 and 100 (0 = worst possible performance value, 100 = best possible performance value). The 5 LVs *Attacking*, *Skill*, *Movement*, *Power*, and *Mentality* serve as the LOCs in the model. From the CTA-PLS test output, after 5,000 bootstrap resamples, all the LOCs and the HOC are formative, which results in the path diagram shown in Fig. 1.

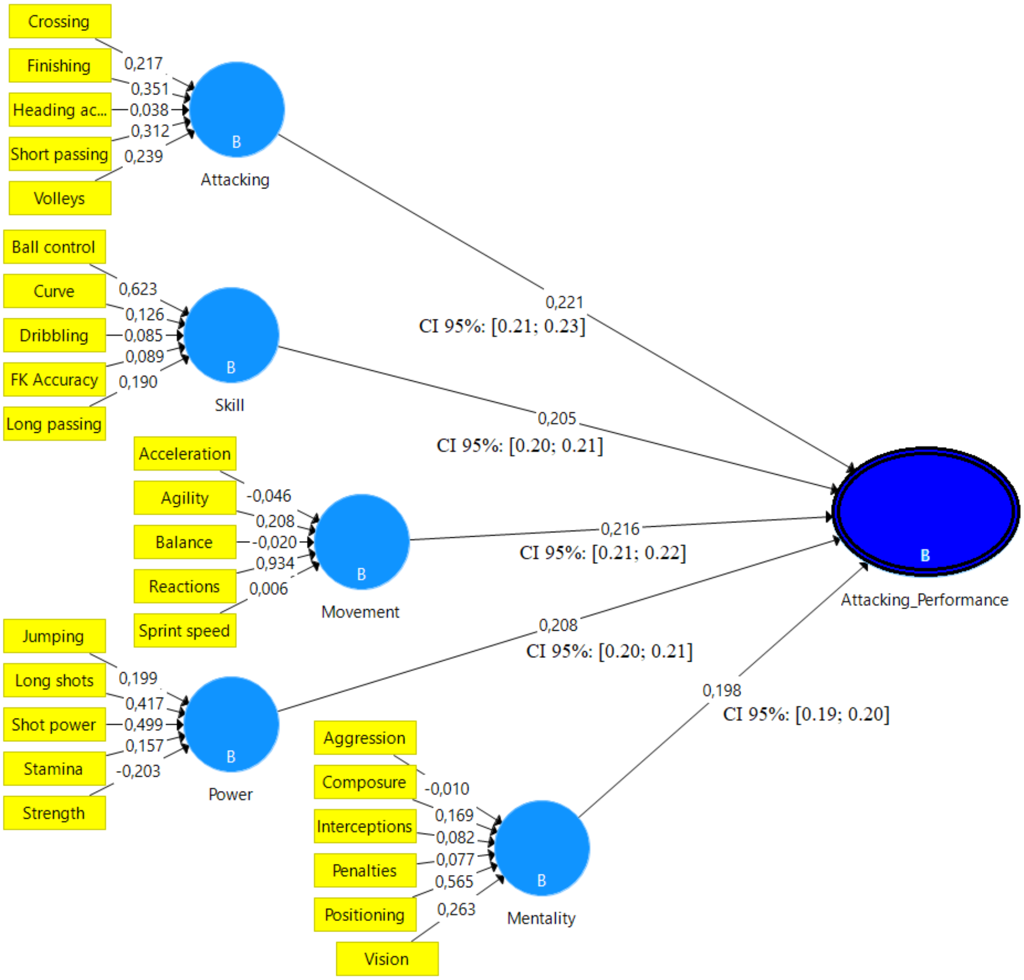

Figure 1: The specification and estimated parameters of the second-order model for measuring Strikers' Macro Attacking performance obtained with the CTA-PLS test (5,000 bootstrap resamples).

This application shows that CTA-PLS corrections can extract empirical evidence to support a proper choice of measurement modes, including higher-order constructs. By leveraging this evidence, researchers can adopt suitable validation methodologies to prevent misinterpretation of analysis results, e.g., assuming formative measurements as reflective and, consequently, evaluating them as invalid, unreliable, or poorly informative based on indices for assessing reflective validity [10]. In our context, correction methods for multiple hypothesis testing play a distinguished role compared with applications in other fields, such as biostatistics. While such domains generally deal with a large number of tests (typically in the thousands), in our setting, correction procedures mainly provide information to guide researchers in revising the structural model and improving its alignment with empirical data.

References

[1] Kenneth A. Bollen and Kwok-fai Ting. A tetrad test for causal indicators. *Psychological Methods*, 5(1):3, 2000.

[2] Kenneth A. Bollen. Latent variables in psychology and the social sciences. *Annual Review of Psychology*, 53(1):605–634, 2002.

[3] Kenneth A. Bollen and Kwok-fai Ting. Confirmatory tetrad analysis. *Sociological Methodology*, pages 147–175, 1993.

[4] Siegfried P. Gudergan, Christian M. Ringle, Sven Wende, and Alexander Will. Confirmatory tetrad analysis in PLS path modeling. *Journal of Business Research*, 61(12):1238–1249, 2008.

[5] Marko Sarstedt, Joseph F Hair, Christian M Ringle, Kai O Thiele, and Siegfried P Gudergan. Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research*, 69(10):3998–4010, 2016.

[6] Mattia Cefis, Mario Angelelli, Maurizio Carpita, and Enrico Ciavolino. Detecting causal relations among indicators with the CTA test: simulations and applications. *Social Indicators Research*, 2025. Accepted for publication.

[7] Mattia Cefis and Maurizio Carpita. On the CTA-PLS test for hierarchical models: an application to the football player's performance. *Computational Statistics*, pages 1–21, 2024.

[8] Joseph F. Hair, Marko Sarstedt, Christian M. Ringle, and Siegfried P. Gudergan. *Advanced issues in Partial Least Squares Structural Equation Modeling.* saGe publications, Thousand Oaks, CA, USA, 2017.

[9] Enrico Ciavolino and Mariangela Nitti. Using the hybrid two-step estimation approach for the identification of second-order latent variable models. *Journal of Applied Statistics*, 40(3):508–526, 2013.

[10] Kenneth A. Bollen and Richard Lennox. Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, 110(2):305, 1991.