



PLS Papers

J. Christopher Westland¹

Published online: 30 August 2023

© Springer Science+Business Media, LLC, part of Springer Nature 2023

1 Partial least squares

Partial Least Squares (PLS) is an embellishment of Principal Components Analysis that extended Sewall Wright's path analysis at a time when pencil, paper and adding machines were the researcher's computers. It appeared in the 1930s, applied to research in gene inheritance. A gene is an unobservable (latent component) expressed through observable, measurable traits (indicators or predictors). Inheritance is binary: you either inherit a gene or you don't. Thus the required statistical resolution for inheritance could be modest, keeping sample sizes relatively small.

PLS was superseded by more powerful methods in genetics, and would have been no more than a historical footnote had it not been resurrected in the 1980s for marketing survey research, where it was augmented with Likert-scaled survey instruments and a few rigid, dogmatic, but speculative psychological models to create the 'PLS Paper' (note the quotation marks). Electronic Commerce Research treats the "Convenience survey + Dogma + PLS" research paradigm (the 'PLS Paper') with skepticism. When such papers are rejected we include in the rejection letter a detailed set of guidelines for rigorous conduct of such studies.

Electronic Commerce Research (ECR) has sought to hold PLS Papers to a more rigorous standard than might be found in other journals. A critical reading of submissions is done prior to moving this genre into the review process. We do publish research using PLS statistical analysis when data collection, modeling and the method are used appropriately, and when motivation, conclusions and supporting analyses are rigorous.

We believe that PLS, PCA and unsupervised machine learning models have important roles to play in model specification search, especially with complex network models in the social sciences. They allow quick assessment of model fit with small datasets that may be expensive to acquire and which cannot be expanded in natural or quasi-experimental settings. PLS specifically allows competing network models to be quickly and cheaply tested with the available data to determine which is most likely to reflect reality, and can quickly identify the

✉ J. Christopher Westland
westland@uic.edu

¹ University of Illinois Chicago, Chicago, USA

dataset and model are incompatible. They are commonly the most efficient alternatives for network model specification searches in the social sciences. But we will desk reject when the ‘PLS Paper’ boilerplate substitutes for poor articulation of research questions and context and weak survey protocols.

I had been asked by several of ECR’s Associate Editors to articulate why we adopted these policies on PLS Papers. The following are the main technical reasons that PLS Papers are rejected for failing to meet reasonable standards for scientific rigor required at ECR.

1.1 Inability to accept or reject the research model

PLS and other structural equation models are network models that test a reality of multiple links between latent constructs. Such models need to be holistically tested, requiring sample sizes sufficient to accept a valid model, and also large enough to reject all invalid combinations of links [11]. Many PLS Papers submitted to ECR instead conduct only pairwise tests of links. In which case, they should have stated the model in terms of pairwise canonical correlations between behavioral constructs.

1.2 Questionable research models

Most PLS Papers we receive test one out of perhaps a half-dozen predefined psychological models (TRA, TAM, TAM2, etc.). PLS is an extension of PCA that allows researchers to pre-select ‘latent’ features from which the model is constructed, and these dogmas essentially do the pre-selection for the researcher. Researchers adopt these dogmatic boilerplates, rather than constructing models appropriate for their research, because by doing so, they think that they can reduce the amount of model criticism by reviewers. ECR feels that these dogmas undermine the accuracy of model explanation and prediction, and substantially reduce the rigor of research. Submissions would be better off implementing custom models suited to the submission’s specific research topic.

TRA [1] is the grandfather of dogmatic research models. It and derivative models have been criticized as being of questionable heuristic value, limited explanatory-predictive power, triviality, and lacking practical value [3, 6, 9, 10] which have diverted researchers’ attention away from other important research issues, created an illusion of progress, and contributed to theoretical chaos and confusion [2]. On a more technical note, the models demand that we perform Neyman-Pearson hypothesis tests on each combination of latent variable links—the analysis needs to accept link structures that are supported by the data, and reject all other structures that are not supported by the data. Since the links are not inherently binary—i.e., one latent variable can have varying degrees of influence on any of the other links—very large sample sizes are needed for validating any particular model at an acceptable statistical significance.

1.3 Proof of actual behavior always outperforms reporting of ‘intended’ behavior

Many of the PLS Papers submitted to ECR have an effect construct that is consumer “intention.” Since today we have access to actual consumer behavior in the form of purchases, reviews and other feedback, we feel that conclusions about “intentions” are unreliable. People will say anything in a consumer survey, and often do after the tenth question. 40 years ago, at the time of the birth of the PLS Paper, access to consumer data was difficult, and surveys were widely used. Today, most retailing and advertising has moved online, and consumer actions can be gleaned from APIs, scraping and other methods that measure actions, not what consumers happen to say on a questionnaire.

1.4 Poor survey protocols

The overwhelming majority of PLS Papers submitted to ECR suffer from poorly designed survey instruments. The six-latent construct models most often tested require on the order of 20 SEM predictors, and researchers often allocate only one question to a predictor. Good protocols require at least four questions per predictor to cross-check responses. Too often these are ‘trigger’ or ‘leading’ questions that coach respondents on the ‘correct’ answer. Obviously, such surveys will only replicate the research narrative; they will fail to capture the true beliefs of respondents. Earlier questions in the instrument will anchor later ones, leading to highly correlated responses and multicollinearity. Even 20 question surveys will be biased by fatigue and drop-offs which is a major problem in telephone, email, and convenience surveys [4, 7]. On average, after more than five questions, respondents become bored or distracted and drop out of it, or randomly fill out the remainder. This has an obvious relevance to question order, since later questions may contain mostly invalid responses [5]. Primacy and recency bias will overwhelm reliable responses in such surveys.

1.5 Convenience sampling

The overwhelming majority of PLS Papers submitted to ECR claim that their results are generalizable to the population, but then fail to obtain a representative sample from that population. Convenience samples are common. These email from a mailing list, and test only the responses that are mailed back to them. Software like SurveyMonkey can streamline the process, but may strongly bias the response. It is common for ECR submissions to see response rates less than 1%. Even with profiling of respondents, there is simply no way to credibly claim that such small response rates are free from bias. Indeed they may be biased towards respondents who have too much time on their hands.

1.6 The “rule of 10”

There is an odd but widely embraced myth among some researchers that a PLS model only requires a sample size of $10 \times X$ sample observations, where X is the number of latent constructs. This has no scientific basis, rather was an opinion rendered in an old textbook [8] that ‘a good rule is to have at least ten times as many subjects as variables.’ Jum Nunnally was not referring to SEM models or even any statistical model at all. Specifically, PLS models need to be holistically tested, forcing sample sizes not only to accept a valid model, but to be large enough to reject all invalid combinations of links. Subsequent publications have attempted to justify small sample sizes in PLS Papers, but a simple test of the data reveals their flaws. One can use cross-validation to test out the impact of small samples. Path coefficients in the PLS model will be highly unstable in cross-validation. This implies that each replication of the research will yield different, widely varying, conclusions.

1.7 Misuse of Cronbach alpha

Cronbach alphas and average variance extracted (AVE) are designed to assure that clusters of questions on the survey all provide information about the same construct. High alphas are not necessarily desirable, as this indicates that the items may be entirely redundant. Between latent constructs, or the factors that they are constructed from, alphas should not be large. Cronbach’s alpha measures internal consistency of observations with respect to a group, and is considered to be a measure of scale reliability. They were designed around large tests with a single construct like Intelligence Quotients. They might also be useful in segmenting small number of independent constructs, though a “high” value for alpha does not necessarily imply that the measure is uni-dimensional. PLS and other SEM models must be evaluated holistically, otherwise they should be separated into smaller independent models.

1.8 Data availability and replicability of research

Electronic Commerce Research requires the inclusion of a data availability statement as a condition of publication, which allows for data not to be publicly available, for instance when individual privacy could be compromised. Authors of PLS Papers, we find, typically rationalize that their survey data could potentially compromise privacy, that the data cannot be made available. Where submitted papers have made the data available, it has in several cases been observed not to support the research conclusions. Multicollinearity in survey responses, possibly for some of the reasons cited previously, is one of the main reasons for inability to reproduce results. Instead of a network of independent latent constructs, these models have been found to support only one common construct.

1.9 PLS is unsuitable for hypothesis testing

Path coefficients in PLS are sub-optimal canonical correlation measures. PLS lacks any distance measures beyond a simple linear distance metric. This is important for human subjects studies, because many human perceptual relationships are logarithmic or otherwise non-linear. Additionally, PLS lacks fit statistics to measure how well the dataset conforms to the researchers' hypotheses, and PLS cannot determine causal direction. The lack of fit statistics for model tests implies that PLS results are suggestive rather than conclusive. They are useful for specification searches, but further tests and models are required for hypothesis testing [12].

PLS like PCA is essentially distribution-free, and this is often cited by authors of PLS Papers as a justification for ignoring the statistical distribution of observations. This is useful as Likert-scaled data is integer-valued and truncated at zero. But additional understanding of the distribution of data would extract more information from the dataset, and would lead to stronger conclusions.

1.10 Poor motivation, research question articulation and technical execution

Because the the “Convenience survey + Dogma + PLS” research paradigm does not demand much technical expertise, and indeed is automated by tools such as SurveyMonkey, the method is accessible by researchers with weak technical skills, and poor understanding of the research context. Thus ECR submissions in this genre are generally of a lower quality than in other areas, tending to have problems in articulation, execution and claims.

1.11 Acyclic graphs

TRA, TAM and variants are acyclic, but some submitted papers test models with linkages containing cycles. Sewall Wright's Path Analysis dictated acyclic models be used; otherwise recursive influences could overwhelm any signals detected in the correlations that made up path model links. Such problems are widespread in network models, and acyclic graph models are necessary for ascertaining causal relations. Where models contain cycles, ECR tends to view any claims of causality with skepticism.

2 An invitation

PLS Papers have proliferated over the past two decades, and there is now a large body of research relying on the PLS Paper paradigm. I have colleagues who hold alternative positions concerning the appropriateness of PLS Papers, and we would like to hear their opinions. ECR welcomes alternative views, and we would like in a

future issue to compile any responses received to this commentary into a follow-up commentary on PLS Papers.

References

1. Ajzen, I., & Fishbein, M. (1975). A Bayesian analysis of attribution processes. *Psychological Bulletin*, 82(2), 261.
2. Benbasat, I., & Barki, H. (2007). Quo Vadis TAM? *Journal of the Association for Information Systems*, 8(4), 7.
3. Chuttur, M. (2009). Overview of the technology acceptance model: Origins, developments and future directions. *Journal of the Association for Information Systems*.
4. Fass-Holmes, B. (2022). Survey fatigue-literature search and analysis of implications for student affairs policies and practices. *Journal of Interdisciplinary Studies in Education*, 11(1), 56–73.
5. Groves, R. M. (2005). *Survey errors and survey costs*. Wiley.
6. Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen Jr, D. J., Hair, J. F., Tomas, G., Hult, M., & Calantone, R. J. (2014). Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013). *Organizational Research Methods*, 17(2), 182–209.
7. Lynn, P. (2008). The problem of non-response, chapter 3, 35–55. In *International Handbook of Survey Research Methods*.
8. Nunnally, J. (1967). *Psychometric methods*. McGraw-Hill Book Co.
9. Ronkko, M., McIntosh, C. N., & Antonakis, J. (2015). On the adoption of partial least squares in psychological research: Caveat emptor. *Personality and Individual Differences*, 87, 76–84.
10. Ronkko, M., McIntosh, C. N., Antonakis, J., & Edwards, J. R. (2016). Partial least squares path modeling: Time for some serious second thoughts. *Journal of Operations Management*, 47, 9–27.
11. Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. *Electronic Commerce Research and Applications*, 9(6), 476–87.
12. Westland, J. C. (2019). *Structural equation models: From paths to networks*. Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.