

《心理学报》审稿意见与作者回应

题目：潜变量交互效应标准化估计：方法比较与选用策略

作者：温忠麟，欧阳劲樱，方俊燕

第一轮

审稿人 1 意见：

文章对潜变量交互效应标准化估计的乘积指标法、潜调节结构方程法、及贝叶斯法进行随机模拟比较，讨论不同情形下各方法的优劣，并给出具体使用建议。本文的研究有助于心理学研究者加深对潜变量交互效应估计方法的理解，有助于心理学实证研究者恰当选用交互效应的估计方法，具有较好的创新性和重要的应用价值。文章脉络清晰，逻辑通顺，方法合理，结论妥当。文章在具体方法讨论、模拟指标报告等方面还可以进一步完善与提升。以下是具体意见和建议。

意见 1：建议作者进一步说明潜变量交互效应的原始估计和标准化估计的差异，交互效应的标准化估计能够回答哪些原始估计不能解决的问题，或说明标准化交互效应的应用场景。当前文章在引言第二段给出了一般参数标准化估计的用途，但未针对潜交互效应特定参数做具体说明。

回应：根据专家意见，在引言第二段段尾补充了潜变量交互效应的标准化估计的作用：“对于潜变量交互效应模型，标准化估计将模型中的潜变量都视为 Z 分数，不仅方便解释交互效应，还可以方便比较简单主效应。”（见 P1）

意见 2：作者在模拟研究设计中提到，文章使用固定负荷法指定测量单位。若是采用另一种指定测量单位的方法，将潜变量的方差设置为 1，那么公式(6)给出的标准化估计值是否等于这一设定下的参数估计值（注：公式(6)每一项中的 SD 都为 1）？那么是否可以省去交互效应标准化估计的计算过程。

回应：在潜变量交互效应模型中，当 ξ_1 、 ξ_2 的方差固定为 1 时，其交互项 $\xi_1\xi_2$ 的方差一般不为 1：当潜变量正态分布时， $\xi_1\xi_2$ 的方差等于 $1 + r_{12}^2$ ；潜变量非正态分布时，还没有公式计算其方差。这就说明很难设定公式（6）中的 SD 都为 1（尤其是外生潜变量的 SD ）。即使采用固定方差法，还是需要使用公式去计算标准化估计。而且，对于乘积指标方法，固定方差法不仅麻烦，还面临非正态带来的稳健性问题。

由于 LMS 和贝叶斯法不需要指定交互项的测量单位，可以使用固定方差法。但还是要 在 Mplus8.2 及以上版本中要求输出标准化解才能得到正确的标准化估计。

意见 3：作者考虑了潜变量正态和非正态的情形，建议明确这里的非正态是否包括分类变量，或仅为连续变量。文章 4.1 节模拟研究中设定非正态分布由两个卡方分布加权得到，为连续变量。但 3.4 节也提到其他研究者“在潜变量非正态时使用二到四个潜类别...”。那么，对于潜类别变量的交互作用，本文研究是否同样适用？

回应：只有可计算均值和标准差的连续变量才有标准分数，因而涉及标准化解的模型中的变量都是连续的。根据专家建议，在第 1 节最后一段说明“本文假设所论的潜变量及其指标都可以合理地当作连续变量。”（见 P2）

意见 4: 文章模拟研究结果部分, 重点关注了偏差 bias 指标, 而均方误差 MSE 指标并未直接报告或讨论。如表 1 中的 SD 越大, 是否说明该估计方法的表现越差? 建议作者加入 MSE 的讨论, 或说明不考虑 MSE 的原因。

回应: 谢谢审稿专家提出这个有意思的问题。用文字来表述, MSE 等于估计量的方差与偏差平方之和。分别考虑偏差与标准差(方差开方)比笼统地考虑 MSE 可以获得更多信息。MSE 是将偏差平方与等量的方差看作相同的东西了, 但实际上两者的统计意义可以是很不相同的。例如, 一个零售店用量具卖花生油, 每次卖出 10 两实际上都只有 9 两, 偏差为 1 两, 方差为零, 所以 MSE 是 1。另一个零售店用杆秤卖花生油, 由于所用的杆秤敏感性差, 卖出 10 两有时候多于 10 两, 有时候少于 10 两, 假设长年累月的结果是均值为 10 两(无偏), 标准差为 1 两, 那么 MSE 也是 1。虽然两家店的 MSE 相同, 但第一家店总是占了顾客的便宜, 即顾客总是吃亏(相当于 100 元买到 90 元的东西); 而第二家店, 长年累月下来不占便宜不吃亏, 常年累月去第二家店买油的邻居也不占便宜不吃亏(长年累月平均而言, 100 元买到了 100 元的东西)。

如果先考虑偏差大小, 就知道第一家店是奸商了。但用 MSE 却发现不了。所以应当先考虑偏差是否可以接受。不能接受的估计量就一票否决, 可以接受的情况下才比较 SD 大小。这就容易理解, 为何很多模拟研究(尤其是本文的同类研究)并不用 MSE 作为评价指标。

意见 5: 文章讨论的“有信息贝叶斯法”, 指代的是先验信息准确的情形。建议作者从三方面对有信息贝叶斯法作进一步讨论(如在 5.2 节中)。1). 当先验信息准确时, 先验方差对估计结果的影响。2). 当先验信息不准确时, 对估计结果可能的影响。3). 研究者应当如何根据实际问题设定先验。

回应: 按照审稿专家的建议, 在 5.2 中增加了有关讨论: “当先验信息准确时, 先验分布的方差越小, 估计结果越准确, 检验力越高(也见 Fang et al., 2019; Miočević et al., 2017)。但若设置的先验信息不准确, 先验分布的方差越小, 估计结果越不准确, 只有当样本量较大时, 才能弥补错误的先验信息带来的问题(Fang et al., 2019; Lee et al., 2007)。在实际的研究中, 可以基于前人的研究或元分析等得到的相关参数设定先验信息。若对这些参数值的把握不大, 可以对先验分布设定较小的精度, 例如方差大于 1 (Fang et al., 2019; Miočević et al., 2017)。而当较难获得参数的先验信息时, 可以首先使用其他方法如乘积指标法、LMS 对潜变量交互效应模型进行估计, 再基于得到的估计值来设置先验分布。”(见 P17)

意见 6: 书写方面。

1).图 1, 内生潜变量的指标标注有误, 不应为 x1。

回应: 感谢审稿专家的细心阅读与指正, 已修改(见 P3)。

2).公式(6)的第一项应为 γ_1 ”。

回应: 感谢指正, 公式(6)已修改(见 P3)。

3).第 2 节最后一段, “Wen 等人(2010)的恰当标准化估计...并提供了恰当标准化估计(Asparouhov & Muthén, 2020)”放在引言部分更合适。

回应: 第 2 节的主题为“潜变量交互效应模型的标准化估计”, 在介绍了恰当标准化解的有关理论知识后, 最后一段综述恰当标准化解的推广与应用, 是理论部分的延伸。而在引言部分, 只是简单地介绍恰当标准化解并引出研究问题。仔细考虑后, 没有修改。

4).第 9 页第二段, “交互效应项的 R^2 分别是...”。请明确这里的 R^2 是包括潜变量和交互项的

模型的 R^2 ，还是交互项的 R^2 增加量。

回应：此处应为交互项所对应的效应量 ρ^2 ，即交互项的 R^2 增加量。已修改（见 P7）。

5).在参考文献中，部分期刊名、书名中单词首字母未大写，部分文献标题中的单词不应大写。

回应：感谢审稿专家细致审读。已经根据《心理学报》的参考文献格式做了修改。

审稿人 2 意见：

作者通过模拟实验（分别选取了 100, 200, 500 三种不同样本量）对乘积指标法、潜调节结构方程法、贝叶斯法（包括有先验信息 and 无先验信息）等几种潜变量交互效应标准化估计方法的表现进行了比较，得出了非常实用、有巨大应用价值的创新研究结果。作者研究思路清晰、逻辑性强，选用的方法正确，选用比如，SE 偏差%、检验效能等作为评价指标，非常适恰。研究内容丰富，整个文稿文字数较大。可在字数上稍精简些会更完美。

回应：感谢审稿专家的意见和建议，已对文章中可有可无的句子进行删减。

第二轮

审稿人 1 意见：作者详尽的回复了我的所有问题，并对论文做了相应的修改。我没有更多意见，建议发表。

编委意见：同意两位审稿人的意见，修改后发表。

主编意见：该文章系统比较了潜变量交互效应标准化估计。该问题具有一定的理论价值和应用意义，文章整体质量较高，作者也很好地解决了审稿人的意见，本人同意编委和审稿人的意见，该文章已经达到发表标准，建议发表。