

## 《心理学报》审稿意见与作者回应

题目：二阶潜增长模型标度方法及其可比的一阶潜增长模型

作者：温忠麟 王一帆 杜铭诗 俞雅慧 张愉蕙 金童林

---

### 第一轮

审稿人 1 意见：

Based on my reading, I would comment this is a well written methodology work, applied researchers in Chinese mainland will surely benefit the modelling approach proposed in this manuscript for their own studies. Nevertheless, as an applied researcher, I was wondering some editing working might be needed to increase the readability of this manuscript, especially for people who had relatively weak math/statistics knowledge background. My comments are entirely my own points, for author(s) and editors' reference, with good wish to improve the manuscript readability for wider readers' community.

意见 1：自检报告 5. Solid evidence is needed here indeed, this response is a typical cultural-revolution style slogan, not a scientific statement.

回应：已经将自检报告 5 修改为：“目前二阶 LGM 还未有功效分析方法，因为它是二阶因子模型的特例，可以按通常因子分析的做法。因子分析中，通常要求被试人数是题目总数的 10 倍以上、自由参数的 5 倍以上（侯杰泰等，2021）。文中的例子，题目数 12，自由参数 25，按题目总数 10 倍计算是 120 个，按自由参数的 5 倍计算是 125，而例子的被试超过 1 千人，远远多于通常实证文章的被试人数。”

方法文章的例子，是为展示方法服务的，不像实证文章需要有功效分析才能保证不显著的结果是可信的。所以通常的方法文章的例子都不会要求功效分析，为节省篇幅，上面的内容就不写入正文了。

意见 2：自检报告 6. This point is redundant, feet for a snake drawing. If APA (or any other publication guideline) publication guideline was followed to prepare the manuscript, please check the writing against guidelines.

回应：谢谢意见，已经删除最后一句。

意见 3：Given LGM is not a generic term for wider applied statistics community and an English abbreviation, I would advise presenting the full name at its first time appearing in the text, for easy reading purpose and useful for literature reviews.

回应：LGM 的英文全称在正文中有，摘要字数限制只使用缩写，并且已经给出了对应的中文全称。

意见 4：For 固定负荷法？

回应：是的，修改稿已经写清楚。

意见 5：Some references here are needed to justify the statement.

回应：已经加上。

意见 6: Might be easier for readers to understand the LGM is just a CFA with repeated measure as manifest indicators. So applied researchers could not be panic with new terms as there are too much Greeks symbols in the 2nd order model, if they are familiar with CFA.

回应：完全同意，本文正是这样做的。对于单变量重复测量数据，从时间变量的回归方程开始引入，然后转到 CFA。

意见 7: This is kind of centering, to make the intercept meaningful, X variable could be centered at any value if needed. For example, the centered at last value which should the final status.

回应：这里不是对  $X$  中心化，而是对时间变量  $t$  中心化。虽然中心化可以有不同选择，但这里明确说明这样做是为了使得第一时间点的测量值为初始值。

意见 8: Generally speaking, the linear trend with equal time space is not very common in reality. Unequal time interval, continuous time and non-linear time effect, might be discussed later?

because this would be a pretty complex situation.

回应：已经根据建议将相对独立的议题放到该节最后讨论。

意见 9: Equation 2-5 could be integrated in one formula (equation 1) I would advise. For reference only. Same suggestion for all equations through the whole document.

回应：专家说的理论上没错。但本文要用到不同方程做推导运算，需要分开编号。例如，为了推出方程(7)，有这样一句“由方程(3)减去(2)后两边求均值”。

意见 10: Very intuitive information to aid reader understand 1st order LGM.

回应：是的，图示对读者理解一阶 LGM 和二阶 LGM 很有用，正文中已经加上“图 1 是一个 4 次测量的一阶 LGM 路径图。”接着是解释增长模型的参数及其含义。

意见 11: 图 1 Simple 1st order CFA model with 2 factors.

回应：参考同类方法文献上的做法，改为“一阶 LGM 的例子”。

意见 12: This is a very old term and was primarily common term used in mainland China. Please used Multilevel modelling.

回应：修改稿已经将“多层模型”改成“多水平模型”。

意见 13: This statement complicated the meaning of these four parameters. Something saying  $\mu_\alpha$  and  $\mu_\beta$  are intercept and slope parameter,  $\sigma_\alpha^2$  and  $\sigma_\beta^2$  are their relevance variance parameter at level 2.

回应：专家说的没错，LGM 中就是这样解释的，但多水平模型中是按固定效应和随机效应来解释的，这里只是说明多水平中的解释，方便熟悉多水平的读者理解这些参数。

意见 14: Looks not a complete sentence?

回应：“基于测量模型可以估计合成分数信度以及误差方差”这句话是陈述事实，是完整句

子。如何估计不是本文要讨论的事情，有兴趣的读者可以参考该句后面的参考文献。

**意见 15:** With reference on Leite's work, what is the point to play 2nd order LGM?

**回应:** 使用显变量建模的 LGM 是一阶 LGM (一阶 CFA 的特例); 使用潜变量建模的 LGM 是二阶 LGM (二阶 CFA 的特例), 因为潜变量本身已经需要一阶 CFA 模型。

**意见 16:** If there are >4 items for each  $\eta$ , by theory, those 4 items could used to measure a

2-factor structure at each time as what was shown in figure 1, then, these 2-factors measured at

each time could be used as indicators for 2nd order factors. If yes, how to interpret that structure?

**回应:** 图 2 所示的二阶 LGM 是描述单个潜变量随时间变化, 所以不管它有多少个指标, 在每个时间点, 只能用单因子测量模型去做 CFA。

**意见 17:** Just a simple 2nd order CFA with 2 2nd order factors, 4 1st order factors.

**回应:** 是的, 一个二阶 LGM 是二阶因子模型的特例。参考同类方法文献上的做法, 修改稿将图 2 的图注改为“二阶 LGM 的例子”。

**意见 18:** Statement + equation here I would advise, for a clear statement on which kind of 测量的跨时间不变性, 并且要求强不变性? And this assumption might be very strong in reality.

How about if these assumption was not met? I would argue it is very common for "测量的跨时间不变性, 并且要求强不变性" not met in reality.

**回应:** 是的, 这个是做 LGM 的基本要求 (理论上说, 一阶 LGM 也需要满足这个要求, 参见讨论部分的修改), 这个强不变性通常会比多组(如性别)比较的强不变性容易满足, 因为是同一组被试的重复测量。如果强不变性不满足, 文中这一段后半部分有提到会有什么后果。如同同类方法文章一样, 本文在强不变性满足的条件下进行讨论二阶 LGM。

**意见 19:** 方程(6')、(7')是 typos?

**回应:** 不是笔误, 这是为了区分显变量和潜变量的同类方程, 后者使用了一撇的编号。后面推理需要指出哪一个方程的。

**意见 20:** 效应编码法 How about give more details so people will know how setup loading constraint, without refer Yang 2021 and Little 2006 work?

**回应:** 赞同, 接着的内容就是专家所希望的细节了。修改稿已经加上“下面讨论文献上最常见的两种做法。”

**意见 21:** This is the methodology contribution made by author(s). Is this 两阶段建模 improved 潜变量标准化? If yes, the manuscript have 4 kinds of coding, Yang 等人(2021)介绍了三种标度方法, 包括尺度指标法、效应编码法和潜变量标准化法, + 两阶段建模, So table 2 should compare the performance of 4 approach, with aiming showing 两阶段建模 is the best, according to some evaluating criteria?

回应：Yang 等人(2021)提出的潜变量标准化法，需要知道题目指标或者合成分数的信度，因而缺乏可操作性。本文提出有可操作性的两阶段法，第一阶段做第一时间点潜变量标准化的 CFA，第二阶段使用第一阶段得到的参数估计值建模。这样做不仅解决了可操作性问题，而且避免使用信度。但如果他们的信度也是通过 CFA 去计算（使用其他方法得到的信度比较不精确），得到的结果其实跟我们提出的方法得到的结果是一样的，只是他们的方法绕了圈子。在这个点上，本文的贡献有两个：一是给出了途径实现他们的方法；二是避免了信度计算，简单了很多。

意见 22: I would recommend clearly differentiate the term of "一阶和二阶潜增长模型" and "一阶和二阶 part of a 2nd order 潜增长模型".

回应：赞同。本文说的一阶 LGM 和二阶 LGM 是不同的建模方法，前者使用显变量建模，后者使用潜变量建模。当说到二阶 LGM 不同部分的时候，文中用类似“二阶 LGM 中的一阶因子”、“二阶 LGM 中的二阶因子”进行区分。

意见 23: 一阶 LGM 和二阶 LGM, Which one do you recommend for researcher? Why?

回应：一阶 LGM 和二阶 LGM 的关系，就像是使用显变量（例如使用题目得分的均值）做回归和使用潜变量做回归（即结构方程）的关系。理论上说，后者是比较好的，但当显变量的信度高时，两者结果很接近。在 LGM 中，两者的差异如何受到信度的影响，需要另文讨论，我们在文末有提到。

意见 24: So, various coding approach could be used for both 1st order CFA and 2nd order CFA, which one is the best?

回应：这个问题在本文的讨论部分有说明：

*“除非对某个指标情有独钟而且该指标信度很高，否则不要使用尺度指标法，因为这样标度的二阶 LGM 结果只与尺度指标建立的单变量 LGM 有可比性，不同的尺度指标可能导致很不同的结果。效应编码法比较好，同时使用了全部指标的信息进行标度，结果与使用合成分数的单变量 LGM 有可比性。可以根据研究目的决定使用总分还是平均分相应的效应编码法，通常使用平均分的比较好，不受指标数量的影响。”*

对于潜变量标准化法，也有说明：

*“基于效应编码法的标准化明显好于基于尺度指标法。”*

意见 25: On top of how to do, how about presenting the pros & cons for each approach? Which one do you recommend and why? As for applied researchers with longitudinal data, the research focus is generally interpreting the estimates of growth factors. Your proposed 两阶段建模 will be the best approach, according to your statement in section 2.3, do you recommend your proposed 两阶段建模 and why?

回应：这个表给出的是一阶和二阶什么情况下可比，或者说二阶 LGM 的标度方法对应的一阶 LGM 是什么，而不是讨论一阶和二阶的优缺点。文末的讨论部分回答了专家关心的问题：

*“一阶 LGM 使用单变量建模，容易理解、操作简单，占有优势。”“单变量一阶 LGM 的不足之处是，忽略了测量误差、不能检验跨时间的测量不变性。”“对于多指标测量的变量，二*

阶 LGM 可以弥补上述不足。可以将测量误差分离出来；可以检验跨时间的测量不变性；在检验发展变化的个体差异方面有较高的统计检验力；可以分离题目指标的方法效应。”“一般情况下推荐使用效应编码法对二阶 LGM 进行标度和标准化，标准化使用二阶段法。”

意见 26: Where is the results of 王兴超、杨继平? Is there any difference of your estimates from various approach and their results? If any difference, why?

回应: 本文使用的是以王兴超、杨继平的问卷来采集的数据，但与他们的文章没有什么关系。

意见 27: Make sure you have the needed license (of MPLUS).

回应: 从 Mplus7.0 到 Mplus8.3 都有正式购买软件。

意见 28: How about if this assumption not met which is a most likely situation in reality?

回应: 我们的实例满足不变性假设。一般性，请参考对审稿意见 18 的回应。

意见 29: How about including the Greek symbol of each parameter in this table to aid reading and understanding the equations? Re contents, why different method yield different results from same dataset? which results is more precise and accurate?

回应: 赞同，修改稿的表头已经加上希腊字母。结果不同是可以预料的，不同标度方法使得潜变量有不同的尺度；一阶 LGM 和二阶 LGM 的差异是显变量建模和潜变量建模的差异，后者考虑了测量误差，而前者没有。精确性而言，在可比情况下，二阶 LGM 比一阶 LGM 的精确；效应编码法比尺度指标法的精确。文中已经有说明。

意见 30: I was wondering the terms of "1st order LGM & 2nd order LGM" were treated same as "1st order part and 2nd order part of a 2nd order LGM" here, right?

回应: 不是的，说的是两种建模方法。请参考对审稿意见 22 的回应。

意见 31: This justify the strength of 2nd order LGM against 1st order LGM. However, the strong assumption for 2nd order LGM might not be always met in real life, so a discussion of limits of 2nd order LGM and future works is needed here.

回应: 二阶 LGM 比一阶 LGM 精确，但模型也更复杂，这就是弱点。不变性假设是二阶 LGM 的前提条件。其实，多指标测量的潜变量使用合成分数做一阶 LGM，不仅要求跨时间的测量不变性，而且要求严格不变性。修改稿已经在讨论部分加上了如下内容：

“对于多指标测量的变量，使用合成分数做一阶 LGM 的优点是简单易理解，前提条件也是满足跨时间的测量不变性并且是严格不变性（即各指标的负荷、截距和误差方差都是跨时间不变的）。如果严格不变性不成立，使用合成分数做单变量 LGM 会引起参数估计偏差，因为这种 LGM 不能区分潜变量的真实纵向变化与测量上的变化(Leite, 2007; Sayer & Cumsille, 2001)。”

意见 32: “无论一阶还是二阶 LGM，合成总分和合成均分得到的标准化结果相同。” So, why play a complicate model when the simple model do the job? What is your suggestion for readers in their own work?

回应：这里不是比较一阶 LGM 和二阶 LGM，修改稿已经重新措辞。原意是：对一阶 LGM 而言，合成总分和合成均分得到的标准化结果相同；对二阶 LGM 而言，合成总分和合成均分对应的效应编码法的标准化结果也相同。

意见 33: Sharing code is very good practice, especially for applied researchers.

回应：完全同意，附录是有必要的，修改稿还增加了一个附录，用于潜变量标准化的第一阶段。

意见 34: If possible, please share your code for other two approach so readers could replicate what you present given that you agree share your data. reproducibility is essential for methodology study and basic ethnic requirement.

回应：实例的数据文档和用到的全部 Mplus 程序及结果文档已经发给了编辑部，审稿专家可以向编辑部索取。附录只给出本文推荐的方法的 Mplus 程序。

.....

**审稿人 2 意见：**

文章系统梳理了一阶和二阶潜增长模型的标度方法，并提出了简便、可操作性强的两阶段方法，通过一个实际例子说明了不同标度方法的估计结果，并对应用者提供了使用建议。文章针对目前追踪数据研究中常用的潜增长模型，关注了模型标度这一重要内容，具有较强的实际意义。所提出的方法科学、简便，具有一定的创新价值。下面提出一些建议，供作者参考。

意见 1：文章主要关注的是潜增长模型标度方法的比较，因此题目是否改为“一阶和二阶潜增长模型标度方法的比较”更为合适？同样，在摘要中“对一阶和二阶 LGM 做了系统比较”也应改为“对一阶和二阶 LGM 的标度方法做了系统比较”。

回应：谢谢专家对题目的疑问，已经将题目改为“二阶潜增长模型标度方法及其可比的一阶潜增长模型”，这样比较契合本文重点。将摘要中“对一阶和二阶 LGM 做了系统比较”改为“系统总结了二阶 LGM 标度方法及其可比的一阶 LGM 建模”。本文的重心不是比较一阶 LGM 标度方法（本文中是固定的）和二阶 LGM 标度方法（本文讨论了三种），而是：（1）二阶 LGM 的各种标度方法，与其可比的一阶 LGM 是什么；（2）提出了可操作性的二阶 LGM 潜变量标准化标度方法，并给出以其可比的合成分数的一阶 LGM 标准化模型。

意见 2：p.3, 引言部分“通常的标度方法（scaling method 也称为识别方法）是使用固定负荷法或者固定方差法指定测量单位”，建议先介绍标度方法包括两部分：设定测量单位和设定测量原点的方法。然后再分别介绍，这样结构更清晰一些。

回应：已经按建议做了修改。对设定测量单位和设定测量原点的方法分别做了介绍。

意见 3：建议将方法部分的结构改为“1 一阶潜增长模型及其标度方法”“2 二阶潜增长模型及其标度方法”“3 一阶和二阶潜增长模型标度方法的比较”。

回应：如果本文的重心是比较一阶 LGM 标度方法和二阶 LGM 标度方法，专家的建议是有道理的，但正如回应意见 1 中说的，本文的重心是二阶 LGM 标度方法。采用目前的结构，可以更好地突出重点。

**意见 4:** 公式 (1) 的介绍, 建议先将测量次数用  $T$  表示, 然后再以  $T=4$  举例。后文的例子都是  $T=4$  的情况。

**回应:** 已经按专家意见做了修改, 参见公式(1)前后内容。

**意见 5:** p.4, “ $t$  可能不是整数”, 建议举个例子便于理解。“当回归方程不是直线的时候”, 建议改为“当  $X$  随时间变化是非线性增长时”, 更为准确。最后一句话“这 5 个 LGM 参数后面会多次提到”建议改为“以下推导需要估计的这 5 个参数的公式”

**回应:** 根据审稿专家一的意见 8, 修改稿将间隔不等、非线性增长的说明放到 1.1 节末尾。举了一个间隔不等、 $t$  有小数的情形。已经将“当回归方程不是直线的时候”改为“当  $X$  随时间变化是非线性增长时”。此外, 在“这 5 个 LGM 参数后面会多次提到”后面, 增加了“下面根据方程(1) - (4)推导含有这 5 个参数的方程”。因为一阶 LGM 是超识别模型, 导出的方程还不是参数的估计公式, 而需要结合其他带有这 5 个参数的方程进行估计。

**意见 6:** 图 1 的名称建议改为“一个含有 4 个测量时间点的单变量一阶线性增长 LGM 路径图”。

**回应:** 两位审稿专家都对图一名称做了建议, 修改稿参考同类方法文献上的做法, 改为“一阶 LGM 的例子”。

**意见 7:** 文中有一些与所讨论主题关系不大的内容, 建议删除以帮助读者更加聚焦于主题。例如, p.6, 第四自然段关于自由估计负荷的内容, 第六自然段的内容前面提到过, 可以综合到前一段中; p.7, 第一自然段, 以及第二自然段中关于曲线因子模型的介绍(因为上一页介绍了用一组题目间接测量潜变量的概念, 与第一种因子曲线模型联系更紧密, 并且本文仅讨论这种模型); p.12, 第二自然段第一句话“正如使用总分与使用平均分做单变量 LGM 得到的参数估计很不相同”可以删除, 后面的内容放到上一段。p.13, 倒数第二段, “至于标准化的一阶和二阶 LGM 是显变量建模和潜变量建模的关系, 不会完全等同。”不属于本文探讨的内容, 建议删除。

**回应:** 确实, 上述建议中绝大多数段落可有可无, 已经删除。

**意见 8:** p.8, 最后一个自然段, 偏差大小与违反不变性的“种类”——改为“参数”, “多少”——改为“题目数量”, 和“大小”——改为“程度”, 似乎更加准确。

**回应:** 已经按建议修改。

**意见 9:** 建议在介绍标度方法的时候, 明确说明从设定测量单位和设定测量原点两个方面的方法, 会显得更加清晰。

**回应:** 在第 2 节第一段已经加上一句: “无论哪种方法, 都有一个等式限制负荷(设定测量单位)、一个等式限制截距(设定测量原点)。”

**意见 10:** 作者在总结尺度指标法的缺陷时提到, “一阶因子选用不同的尺度指标, 得到的 LGM 参数估计结果可能是很不相同的”。那么, 基于尺度指标法的潜变量标准化法(两阶段建模)是否也有这个问题?

**回应:** 理论上说没有这个问题, 因为不管基于哪个尺度指标, 目标都是将第一时间点的潜变量设定为标准化变量(均值为 0、标准差为 1)。但由于这种设定是间接的, 实际上会有误差, 即最后设定的潜变量与标准化变量有出入。

**意见 11:** p.10, 公式 (23) 下方文字像教科书上的内容, 且不够简洁。建议直接改为“研究者可以通过检查从理论上计算的自由度与程序运行结果得到的自由度是否一致, 检验潜变量标度是否正确。”

**回应:** 已经删除了“学会从理论上计算自由度很重要, 自由度是不依赖于具体数据、根据模型就可以计算得到的。”一句。但不可以“通过检查从理论上计算的自由度与程序运行结果得到的自由度是否一致, 检验潜变量标度是否正确。”

**意见 12:** p.10, 最后一个自然段第一句话, 建议改为“在方程 (24) 两边同时除以指标个数 (例如, 3)”。“如果限制”, 改为“则变为限制”, 并删除下一页的第一个“则”。

**回应:** 已经根据建议做了修改。

**意见 13:** 作者一直在关注标度方法的可比性问题, 什么是可比性, 满足可比性有什么好处, 满足可比性的表现是什么? 需要在前面首先进行解释。

**回应:** 为了更清晰说明可比性, 3.1 节第一段开头改为: “由 2.1 节可知, 对于使用尺度指标法的二阶 LGM, 与仅使用尺度指标 (而完全忽略其他指标) 建立的单变量一阶 LGM 有可比性。具体来说就是, 对 LGM 参数中的每一个, 以  $x_{t1}$  为尺度指标的二阶 LGM 与用  $x_{t1}$  建立的单变量一阶 LGM 可以得到共同的方程, 例如,  $\mu_{\alpha} = E(x_{11}), \mu_{\beta} = E(x_{21}) - E(x_{11})$ 。”

比较一阶 LGM 和二阶 LGM, 只能拿有可比性的两个模型来比较。文中 3.4 节已经提到, Leite (2007) 模拟比较的一阶 LGM 和二阶 LGM, 弄错了比较的对象。为了说明讨论一阶和二阶 LGM 可比性的作用, 已经在第 3 节第一段增加了一句话: “讨论一阶和二阶 LGM 的可比性, 不仅在需要比较两者的时候能找对比较的对象, 而且可以更好地理解二阶 LGM 标度方法。”

**意见 14:** p.11, 第三自然段, “同一种标度方法得到的结果也没有可比性”, 应改为“同一种标度方法下选用不同限制得到的结果也没有可比性”。同时, 建议只说这一个缺陷就可以了, 因为第一个缺陷“不同标度方法得到的结果没有可比性”, 作者所推荐的两阶段法也无法解决。

**回应:** 已经按建议改为: “前面介绍的两种标度方法可能会得到很不同的二阶 LGM 结果, 不仅不同标度方法得到的结果没有可比性, 同一种标度方法下选用不同限制等式得到的结果也没有可比性。”这一段是为寻求标准化解做铺垫。文中推荐的两阶段法, 是在实现前人提出的潜变量标准化法。理论上说, 无论从尺度指标法还是从效应编码法 (也无论用哪一种限制等式) 出发进行标准化, 都有唯一解 (Yang 等人 (2021) 文章就这样说的), 但实际上, 因为潜变量标准化的间接性, 结果可能有出入, 参见表 3 最后 3 行)。

**意见 15:** p.11, 最后一个自然段, “基于效应编码法也可以推出类似(30)的限制等式”, 如何推出? 另外, p.12, “显然, 无论基于合成总分或者合成均分进行标准化, 推出的限制等式都一样。”依据是什么? 这个结论是否对于两种两阶段建模方法都成立。建议作者阐述清楚, 或者把一些不太重要的推导过程放到附录。

**回应:** 关于第 1 个问题, 文中已经根据信度定义通过简单的代数推导了公式 (30), 效应编码法使用的是合成分数, 用合成分数代替尺度指标推导便可。



关于第 2 个问题，已经修改为“并将 $\lambda_1 + \lambda_2 + \lambda_3$ 限制为第一阶段得到的负荷之和（基于合成总分标准化）或者将 $(\lambda_1 + \lambda_2 + \lambda_3)/3$ 限制为第一阶段得到的负荷均值（基于合成均分标准化）。显然，无论基于合成总分或者合成均分进行标准化，推出的限制等式都一样。”

**意见 16:** 建议作者在介绍两阶段建模的方法时，分为两种方法介绍：基于尺度指标法的潜变量标准化法和基于效应编码法的潜变量标准化法。这样会更清晰一些。

**回应:** 已经按建议分开介绍。

**意见 17:** 公式 (32) 应该为  $x_{11}, x_{12}, x_{13}$ 。

**回应:** 谢谢审稿专家细致的审阅，已经将下标对调改正。

**意见 18:** 建议删除“LGM 模型”中的“模型”。M 即为 model。

**回应:** 已经全文修改。

**意见 19:** p.12, “3 一阶和二阶潜增长模型的比较”下面一段话，建议简单阐述一下如何建立一阶和二阶 LGM。一阶 LGM 是否是用合成分数，二阶 LGM 就是用潜变量建模？

**回应:** 是的，已经修改为“对于多指标测量的变量，既可以使用合成分数建立一阶 LGM，也可以使用潜变量建立二阶 LGM。”

**意见 20:** 表 1 中的“备注”建议改为“限制条件”。二阶 LGM 的潜变量标准化法，建议改为“基于尺度指标或者效应编码对参照点潜变量进行标准化”。

**回应:** 已经采纳后面一个建议修改。备注栏中的内容不止是限制条件。

**意见 21:** p.14, 测量不变性不用做 delta 卡方检验吗？RMSEA 的值略大。后面提到的单指标信度是采用什么方法计算的？

**回应:** 因为卡方是  $N-1$  乘以拟合函数的最小值，所以被试人数多（例如本文例子被试上千）的时候，即使模型拟合很好，只要不是完美拟合，卡方都可能很大，不适合用来做模型拟合检验，这就是结构方程模型要引进拟合指数的原因。Delta 卡方是两个卡方之差，等于两个模型的拟合函数之差乘以  $N-1$ ，被试人数多的时候，也是不合适用来检验的。就本文的例子而言，尝试过用一半或者三分之一的数据，进行 Delta 卡方检验，强不变性都成立，但用全部数据进行 Delta 卡方检验却不成立。如果需要，可以改用拟合指数的差异检验。单指标信度计算公式按信度定义，等于指标的真分数方差除以观测分数的方差。例如，注意到  $x_{11}$  测量的潜变量是标准化的， $x_{11}$  的信度公式  $\rho_{x_{11}} = \lambda_1^2 / \sigma_{x_{11}}^2$ ，推导公式 (30) 的时候使用过。

**意见 22:** 表 2, 建议将潜变量标准化中的“基于题目 1 标准化”改为“基于尺度指标法的潜变量标准化法（题目 1 为尺度指标）”，“基于题目 3 标准化”改为“基于尺度指标法的潜变量标准化法（题目 3 为尺度指标）”，“基于合成分数标准化”改为“基于效应编码法的潜变量标准化法”。

**回应:** 从标度方法的角度而言，专家建议的没错。但表格宜简洁，并且实际上就是表中列出的那么回事：基于尺度指标法的潜变量标准化法（题目 1 为尺度指标）实际上是通过题目 1

去推导标准化 $\eta_1$ 的限制, 参见用方程 (31) 推导公式 (30) 的过程。而基于效应编码法的潜变量标准化法实际上是通过合成分数去推导标准化 $\eta_1$ 的限制等式, 从方程 (24) 开始推导。

**意见 23:** 为什么三种标准化法得到的截距因子均值有明显差异, 是因为以不同指标为尺度变量标准化吗?

**回应:** 参见意见 14 的回应。按理, 无论用什么方法标准化 $\eta_1$ 后, 其均值应当是 0, 方差应当是 1。但因为潜变量标准化的间接性, 均值不一定刚好等于 0, 方差不一定刚好等于 1。此外, 参数估计还有一个误差来源, 就是 $\mu_\alpha = E(\alpha) = E(\eta_1)$ 不是真的成立, 因为从方程 (18)–(20), 可以得到有关 $\mu_\alpha$ 的其他方程, 例如, 由 (19) 可得 $\mu_\alpha + 2\mu_\beta = E(\eta_3)$ 。当线性增长不是很完美的时候,  $\mu_\alpha$ 与 $E(\eta_1)$ 会有出入。误差可正可负, 两种误差还可以叠加, 这就可以解释三种标准化法得到的截距因子均值有明显差异。显然, 基于效应编码法 (使用了所有指标的信息) 的结果是比较理想的, 截距因子均值很接近 0。

---

## 第二轮

**审稿人 1 意见:**

Thanks for the revised version of manuscript (xb23-053). It was a good reading indeed. I have gone through and was satisfied with the revisions and responses from author(s), I have no more comments on the revised version now.

**回应:** 感谢审稿专家的认可。

**审稿人 2 意见:** 感谢作者仔细的修改。建议在测量不变性部分使用 delta CFI, delta RMSEA 检验, 更加有说服力。

**回应:** 谢谢审稿专家的建议, 修改稿实例部分已经增加了 delta CFI、delta RMSEA 的跨时间测量不变性检验, 见第 4 节红色字体部分。

---

## 第三轮

**编委意见:** 该文有很强的应用价值, 也有一定的创新。作者根据两轮审稿意见做了详细修改和回应, 同意两位审稿人的意见, 建议发表。

**主编意见:** Yang, Luo, & Zhang (2021)介绍了三种二阶 LGM 的标度方法: 尺度指标法, 效应指标法和潜变量标准化标度方法。该论文作者在 Yang et al. (2021) 的工作基础上, 简化了潜变量标准化标度方法的流程, 并提出了二阶 LGM 标度方法与可比的一阶 LGM 建模, 最后

用实测数据将二阶 LGM 标度方法与可比的一阶 LGM 建模进行系统的比较。整篇论文是在测量跨时间强不变性（测量截距和负荷不随时间变化）满足的条件下展开的讨论。在这种测量跨时间不变性的情境下，公式和图表的展示也比 Yang et al. (2021) 的呈现更加简洁明了。

但论文还有如下小问题，建议小修改后发表。

**意见 1:** 图 2 的标注里面还是要明确“该二阶 LGM 例子基于测量不变性条件，即测量截距和负荷不随时间变化”。第 21 页中二阶 LGM 需要估计参数和自由度计算，要明确是这些计算是基于二阶线性 LGM 模型。对于线性估计，只需要考虑 5 个 LGM 参数 ( $\mu_\alpha$  和  $\mu_\beta$ ,  $\sigma_\alpha^2$  和  $\sigma_\beta^2$ ,  $\sigma_{\alpha\beta}$ )，但如果是考虑二次项 LGM 增长，LGM 参数就多了。也建议给出具体实例的自由度估计（基于图 2）：“根据公式 23，这里的二阶 LGM 的自由度为 65”。

**回应:** 谢谢主编的意见，已经按建议做了相应的修改。

**意见 2:** 作者在摘要里面提到“简化并改进了现有的二阶 LGM 标度方法，包括尺度指标法、效应编码法”，在讨论部分也提及“简化了各种标度方法的推演”。作者对前两种二阶 LGM 标度方法的介绍（尺度指标法和效应编码法）与 Jeon & Kim, (2021) 和 Yang et al. (2021) 一样，而对第三种标度方法（即标准化标度方法）进行了简化。作者要详细介绍一下，对前面两种常规的标度方法，简化和改进体现在哪些方面。

**回应:** 谢谢指出的问题，反映我们原来没有说清楚。已经在摘要和讨论部分修改了表述。其实，在有关二阶 LGM 的标度方面，本文做了下面两件事。第一件是简化了尺度指标法和效应编码法的推演方式（上面提到的两篇文献都是从研究者的视角，先将潜变量的均值表示为指标均值、负荷和截距的函数，然后讨论如何限制参数去标度潜变量；本文是从用户的视角，直奔主题，说明如何限制参数去标度潜变量及其相应的结果是什么）。第二件是提出了一个有可操作性的两阶段法去实施潜变量标准化标度方法（因而改进了文献上已有的方法）。