

《心理学报》审稿意见与作者回应

题目：预测视角下双因子模型与高阶因子模型的一般性模拟比较

作者：温忠麟，汤丹丹，顾红磊

第一轮

审稿人 1 意见：

文章基于徐霜雪等人(2017)在《心理学报》发表的《预测视角下双因子模型与高阶模型的模拟比较》，考察了分别以满足比例约束的双因子模型（等价于高阶因子模型）和不满足比例约束的双因子模型为真模型的前提下，两种模型预测效度的结构系数。整体来说，文章是对前人研究的补充，丰富和完善了模型比较的模拟研究设计和评价标准，文章的研究设计较为合理，考虑较为完善，得到的结论是前人研究的很好补充和完善，因此其结果也具有一定的价值。具体审稿意见如下：

意见 1：文章中某些表达和说法应当规范。例如，在引言部分，作者在解释双因子模型与高阶因子模型的关系时提到“一阶因子被高阶因子解释后的残差对应于局部因子，解释相应题组的特殊变异”，这时文中第一次出现题组的概念，最好解释一下题组的含义及特征。另外，此时一阶因子所对应的题目不一定是题组，也可以是一个分维度等，建议从共同变异的角度加以解释。

回应：感谢审稿专家的意见。原文“题组”的含义就是因子对应的一组题目，为明确意见，已改为“表示被全局因子解释后测量该因子的那些题目的共同变异”。

意见 2：文章 2.3 两种模型的关系部分，主要使用语言描述了两种模型之间的嵌套关系。建议作者采用公式，推导出两种模型的嵌套关系，会显得更加清晰。这部分的表述可以参考 Demars, C. E. (2006). Application of the bi-factor multidimensional item response theory model to testlet-based tests. *Journal of Educational Measurement*, 43(2), 145-168 的表达。

回应：审稿专家二认为第二节模型概述数学描述过于详细，考虑模型嵌套关系并非本文研究重点，我们采纳审稿专家二的意见，不再增加公式推导，将 Demars(2006)的文章作为参考文献，供有兴趣的读者进一步参考。

意见 3：文章 2.4 两种模型的预测效度中，只包含了预测变量是潜变量的情况，而徐霜雪等人的文章中同时考虑了预测变量是显变量和潜变量的情况，虽然作者在后面的讨论部分说明了在研究中同时考虑了两种情况，结果的一致性很高，但是这里第一次出现预测效度的研究设计，还是应当明确说明，考虑的是预测变量是潜变量的情况。

回应：谢谢提醒。采纳审稿专家二的意见后，我们删除了原来的模拟研究一，同时增加了下面的说明：

“我们首先重复了徐文的模拟研究，即真模型是满足比例约束的双因子模型。无论效标变量是显变量还是潜变量，在估计偏差上得到的结果与徐文高度一致，即满足比例约束的情形，高阶因子模型结构系数的相对偏差比较小。此外还发现，检验力方面，若真模型是全局因子和局部因子同时作为预测变量，整体上双因子模型的统计检验力较高。相应地，无论真模型是全局因子还是局部因子作为预测变量，都是高阶因子模型的第 I 类错误率比较小。为了节省篇幅，这里不报告细节。

下面报告真模型为不满足比例约束的双因子模型的模拟研究，这是徐文没有做的情形。虽然我们也考虑了效标为潜变量和显变量的两种设计，但因为两种设计的结果高度一致，所以下面只报告效标为潜变量的情形，与徐文的一样，效标潜变量有 3 个指标、标准化的负荷固定为比较有代表性的 0.7（也见 Gu et al., 2017b）。”

意见 4：作者在模拟全局因子的负荷时，考虑了 0.4, 0.5, 0.6 和 0.7 四种情况，这几种情况载荷都较高，为什么没有考虑载荷低于 0.4 的情况？这种设计的原因在徐霜雪等人的文章中有说明“本研究不予考虑路径系数为 0.3 及 0.3 以下的值，因为路径系数少于 0.3 的项目的路径系数不符合实证研究中好的结构效度模型应有的路径系数要求”。同理，效标变量负荷都是 0.7 的设置原因，也建议作者进行一定的说明。

回应：沿用徐文设计参数方面已经有说明。徐文说过的理由没有去重复，但采纳专家意见，修改稿增加了一个参考文献，说明前人设置效标变量的负荷也是 0.7(Gu, Wen, & Fan, 2017b)。

意见 5：文中多次提到“但一般情况下比例约束条件是不满足的”，事实上，这也是这个问题提出的主要原因，但是关于这一点是否有文献支持，现实中实际数据得到的情况是怎样的，作者并没有清楚的说明。希望对此予以补充。

回应：因为比例约束条件涉及多个等式，常识可以告诉我们，“一般情况下双因子模型是不满足比例约束的”。为了增强说服力，根据审稿意见，我们增加了下面一句：“例如，在 *Personality and Individual Differences* 期刊发表的关于双因子模型应用的文章中，有 33 篇论文报告了因子负荷，没有一个双因子模型是满足比例约束的。”（补充到 2.3 节）。

意见 6：在这个研究中，虽然整体上看双因子模型的结果较好，但是其第 I 类错误率较高仍然是一个值得关注的问题。作者在讨论部分对此的解释是“虽然第 I 类错误率略高，但从模拟结果可知增加样本容量可以降低第 I 类错误。”那么，样本量达到多大的时候，可以忽略双因子模型第 I 类错误率较高的问题，使其达到与高阶因子模型相近的结果。希望作者可以根据模拟结果提出建议，供实际应用者参考。

回应：谢谢建议，在讨论部分增加了下面内容：“本研究中发现若不满足比例约束条件，样本容量为 1000 时，双因子模型的第 I 类错误率基本上可以接受；样本容量为 500 时，近六成的处理中双因子模型的第 I 类错误率都在可接受的范围内。因此，使用双因子模型时样本宜大一些，比如不小于 500。”

综上，尽管这篇文章是基于已有研究的补充和拓展，其研究方法较为科学，论文撰写表述较为清晰，结果较为丰富，具有一定的说服力。因此建议退作者修改后再审。

回应：感谢审稿专家的肯定。

.....

审稿人 2 意见：

本文是根据徐文（徐霜雪、俞宗火和李月梅，2017）的延伸，有一点贡献，但新亮点不多，建议将全篇文章，大幅度缩减，作为一简短报告发表。第二节模型概述，数学描述过于详细，应指导读者参考已发表的文献，即缩短 2.1 到 2.4 的内容。

回应：谢谢审稿专家的建议，经过权衡，修改稿删除了原来的模拟研究一，大大压缩了篇幅。

徐文有两个缺陷：一是徐文研究目的和结论中所提到的双因子模型是一般的双因子模型，而在其模拟研究中所使用的双因子模型却是满足比例约束的双因子模型(此时等价于一个高阶因子模型)。二是徐文只是比较了结构系数估计的相对偏差，没有涉及检验方面的指标。

对于徐文的第一个缺陷，因为在一般情况下，徐文的结果是不对的，如果徐文读者采用其研究结果选用模型，就会被误导，在这个意义上，本文其实是在纠错了。因为在一般情况下得到的结果与徐文的结果差别大，只写出简要报告难以让读者信服，所以报告主要结果是合适的。

对于徐文的第二个缺陷，需要多一点说明。如所知，参数估计和假设检验是推断统计的两个主要方面，做模型参数（这里是结构系数）的比较时，需要同时考虑估计和检验两方面的指标，通常至少会考虑参数估计的相对偏差、检验的第一类错误率、检验力。因为本文使用了更多比较指标，呈现主要模拟研究结果是合适的。

至于模型概述部分，与通常的文献综述部分类似，详略长短需要一个平衡。审稿专家一建议“采用公式，推导出两种模型的嵌套关系”应当是没有必要的，因为嵌套关系的数学推导文献上已经有了。但也不是文献上有的都不要介绍，文章所用到的主要知识还是需要有一个基本的介绍，让已经有相关知识的读者，不必去翻查文献就能够顺利读完全文。

再次感谢审稿专家的意见，启发我们在修改文稿的时候做出权衡，删除了原来的模拟研究一，大大减少了篇幅。

第二轮

审稿人 1

作者较好地回答了第一轮审稿中提出的问题，也较为清楚地叙述了本研究 and 以往研究的关系，缩减后文章比较聚焦了。但是建议对下面几点进行进一步修改完善：

意见 1：目前文章中文献的引用有一些是来自二手文献，尤其是与已有研究对比部分，建议将原出处文献也进一步引用。

回应：感谢审稿专家的意见。在研究对比部分，修改稿已经尽量使用了一手文献。但为了避

免增加早期的且对读者参考意义不大的文献，并非所有地方都追溯到一手文献，这也是论文写作的惯例，否则许多论文都会充斥几十年甚至是几百年前的文献。

意见 2: 在讨论部分能够结合心理学的实际研究问题进行有针对性的讨论，并给出具体模型比较和选择上的建议。

回应: 感谢审稿专家的意见。修改稿强调了学科理论和研究目的对选用模型的作用。同时让读者明白，本文不是反对使用高阶因子模型，也不是推荐双因子模型。而是纠正徐文给人的误导，以为什么时候都是高阶因子模型优于双因子模型。

“……总之，当模型不满足比例约束条件时(通常的应用属于此种情况)，从统计角度不能说高阶因子模型比双因子模型还好。

那么，是不是在预测视角下，就优先考虑双因子模型呢？也不是，而是应当从学科理论出发、结合研究目的选用模型。与双因子模型相比，高阶因子模型不仅更加简洁，而且其中的一阶因子比局部因子更容易理解。如果研究者使用高阶因子模型进行预测，而整个模型已经拟合良好，且各项评价指标也达到要求，是可以接受的。但高阶因子模型可以接受，并不能说它优于双因子模型。

……”

意见 3: 参考文献格式不一致，需要修改。

回应: 已按 APA 格式修改了参考文献的格式。

审稿人 2 意见:

I have read through the revision. a)Yes, reducing one study makes the article more concise. b)The differences to the earlier studies have now been more clearly described. c)There is some reduction in length. Basically the article is alright, if really necessary ask the author(s) to reduce the length slightly further.

回应: 感谢专家的肯定。修改稿不超过一万字。