

## 《心理学报》审稿意见与作者回应

题目：有调节的中介模型检验方法：竞争还是替补？

作者：温忠麟，叶宝娟

### 第一轮

审稿人 1 意见：

**意见：** 本文估计是“有中介的调节模型检验方法：甄别和整合”的姊妹篇，应该可以发表。

**回应：** “有调节的中介”（moderated mediation）与“有中介的调节”（mediated moderation）有密切关系，从研究对象说，本文算是“有中介的调节模型检验方法：甄别和整合”（以下简称“甄别和整合”）的姊妹篇。非常感谢审稿专家对这个“姊妹篇”开了绿灯，可以填补方法缺块。

需要说明的是，这两“姐妹”其实很不相像。从本文可以看出，“有调节的中介模型”和“有中介的调节模型”不是对称或者对偶关系。从研究内容说，本文与“甄别和整合”一文没有平行关系，除了部分路径图相同外（部分路径图可以有两种解释），也几乎没有什么重合的内容。

具体一点说，“甄别和整合”一文主要工作是：（1）评介了文献上可以找到的五种检验有中介的调节模型的方法，逐一甄别，推荐好的方法；（2）提出了有中介的调节效应的一种新的表现形式。而本文的主要工作是：（1）将有调节的中介模型进行分类，通过分类有新的发现；（2）讨论了“有调节的中介模型”与“有中介的调节模型”的联系和区别；（3）本文最主要的贡献是弄清楚了各种方法之间是竞争关系、互补关系还是替补关系，并提出一个层次检验流程（参见“论文自检报告”）。

**意见 1：** 本文的英文摘要和中文摘要完全不同，请统一。

**回应：** 尽管做了点修改，使得中文摘要的内容都能出现在英文摘要中，但因为英文摘要是长摘要，所以两者出入还是很大的。

**意见 2：** 根据数据拟合的经验，这一类研究从模型和数据的检验到实际问题的解释有被夸大的倾向。继续研究的话，倒不如研究一下这方面的问题。（修改百分之几的数据，有没有功效的结论可能转换）

**回应：** 相信这是一个值得研究的问题。不过，在本文的例子中，即使使用另一组被试的数据，

检验结果仍然相同。至于解释被夸大的问题，是值得应用工作者注意的。

**意见 3:** 心理学报刊登的这一类文章的应用例子，过于简单化，缺乏实际应用价值，从而影响了文章的价值，值得继续研究时候注意改进提升。

**回应:** 方法文章的例子比较简化是一个事实。如果有比较专深的实际例子，无疑对实际领域的价值更大。不过，方法文章往往将重心放在方法和模型上，简化的模型可以更加明快地展示方法的应用，让读者容易理解，也容易控制文章篇幅。例如，温忠麟、侯杰泰和张雷等人2004-2006年在《心理学报》上先后发表了三篇有关中介和调节的文章，其中例子用的是同一个研究的数据，涉及的是儿童行为对同伴关系的影响，例子很简化，社科各领域的读者都能看懂，这或许是这几篇文章被社科各领域的研究者大量引用的一个因素。如果例子本身比较专深，很多读者会面临两重困难，一是方法本身的理解困难，二是对例子所涉概念不熟悉导致的困难。所以，期望审稿专家能从这个角度包容方法文章所含例子的简单性。诚然，我们认同专家的意见，在让广大读者容易理解的同时，提升例子本身的价值，是一个努力方向。

审稿人 2 意见:

**意见:** Moderation and mediation are very important concepts and their analyses have found wide applications in psychology and other social sciences. After introducing different types of moderated mediations, the manuscript discussed different testing methods and recommended a hierarchical procedure for testing moderated mediation. The paper is well written and deserves to be published.

**回应:** 非常感谢审稿专家对本文工作的肯定。

**意见 1:** In several places of the paper, the authors suggest treating parameter as zero when the statistic is not significant. This is not consistent with the principle of statistical testing. It is important to realize that mediation/moderation may still exist even if the statistics are not significant. In particular, statistical significance test is closely related to sample size. For example, there exists  $t = -1.807$  above figure 11 on page 15. It is very likely that another researcher will find that the statistic for the same parameter is significant if he/she has a sample with a larger size.

**回应:** 根据审稿专家的意见做了修改（用蓝色字体显示）。一般地说，审稿专家的意见在纯

粹验证性研究中是对的，但在探索性研究中，我们原稿中的做法是统计分析的通行做法。例如，在熟知的回归分析中，逐步回归建模属于探索性分析，回归系数不显著的自变量不会包含在模型中（相当于不显著的回归系数当作零了）。修改后，我们强调：“如果是纯粹验证模型，则不论系数是否显著，都保留在模型中，无需重新估计。”这样不仅可以解决审稿专家的疑虑，也让读者更加明确在不同范式的研究中的不同做法。

**意见 2:** The presentation of the paper closely depends on the order of the powers of the three test procedures. It is better to clearly state where the order of the powers is established. It is also very tutorial to have the exact significance levels of the three tests reported in the example. Just reporting  $p < .01$  or  $p < .05$  does not facilitate the understanding of the order of the power, as repeatedly noted in the paper.

**回应:** 赞同审稿专家的意见。在第 3 节中，增加了若干内容，清晰论述和比较三类检验方法的检验力，从而得出它们的检验力高低排序。还有，在例子中，已经将“ $p < .05$ ”类型的报告改成报告准确的显著性水平  $p$  值。

**审稿人 3 意见:**

**意见:** 作者在文中讨论了几种检验有调节的中介模型的方法，并对这些模型的优劣进行了比较，但是本文存在下述问题：

- 1) 文章缺少创新，所论问题国内外都有文章进行讨论；
- 2) 所论问题价值不大。

**回应:** 一般地说，期刊编辑部都会要求审稿专家根据文稿给出实质性的审稿意见。像“缺少创新、价值不大”这样的泛泛评语，不看文稿就可以写出来。如果这样的评语都能接受，其做法可能会传染：被这样对待的作者，可能在评审别人文稿的时候也采用同样的做法。我们觉得这样泛泛的评语应当不算数：一方面不应作为对所审文稿的意见，另一方面编辑部不应当支付审稿费。

是否真的“所论问题国内外都有文章进行讨论”？本文重点是讨论方法之间的关系，如果是竞争关系，应当弄清谁优谁劣；如果是互补关系，应当知道什么情况下用哪种方法较好；如果是替补关系，应当说明谁先谁后。目前文献上还未见从这个角度对有调节的中介模型检验方法进行讨论，我们依据方法之间关系提出的层次检验流程（审稿专家 2 称之为

hierarchical procedure), 更是首创。

## 第二轮

**意见 1:** The thesis of this paper is supposed to argue for the superiority of the combined model and to introduce the procedure for the analysis. However, the way it is written, it is not convincing to readers how this combined model contributes to better understanding of complex behaviors. To present a stronger argument, you need to illustrate what is the difference between this proposed procedure and the existing testing methods of the mediation-moderation model, and why this newly proposed procedure is better than existing testing method using your procedure.

**回应:** 第 4 节中, 在提出新的层次检验流程后, 增加了如下内容和表 1, 以说明新的检验流程与现有的检验方法的差别, 以及新的检验流程的优点:

“为了说明使用上面的层次检验流程与已有的检验方法的差别以及层次检验的优点, 不妨设想有 4 个研究者用同一组数据检验同一个有调节的中介模型, 甲使用传统的依次检验, 乙使用系数乘积的区间检验, 丙使用中介效应差异检验, 而丁使用本文提出的层次检验。根据 3.2 节的讨论可知, 就甲乙丙而言, 甲的检验力最低, 乙的次之, 而丙的检验力最高。因此, 如果甲的结果显著, 则乙的结果也显著 (不排除出现例外的可能, 但现有的模拟和实际数据都未发现例外); 如果乙的结果显著, 则丙的结果也显著 (从 3.2 节可知, 这里不会有例外)。因而, 可以推知甲乙丙的检验结果只有表 1 所列的 4 种情形, 使用层次检验的丁的检验结果可以根据甲乙丙的检验结果而确定。

情形 1: 甲的结果显著, 因而乙和丙的结果也显著; 丁做了一步检验就停止了, 和甲的结果相同。

情形 2: 甲的结果不显著, 但乙的结果显著, 因而丙的结果也显著; 丁做了两步检验后停止, 第一步得到甲的结果 (不显著), 第二步得到乙的结果 (显著)。

情形 3: 甲乙的结果都不显著, 但丙的结果显著; 丁做了三步检验, 第一步得到甲的结果 (不显著), 第二步得到乙的结果 (不显著), 第三步得到丙的结果 (显著)。

情形 4: 甲乙丙的结果都不显著; 丁做了三步检验, 第一步得到甲的结果 (不显著), 第二步得到乙的结果 (不显著), 第三步得到丙的结果 (不显著)。

丁的检验力和丙的一样, 高于甲和乙 (情形 3); 在各种情形, 丁获得的信息是最多的, 在情形 2 和 3 中都多于其他三人; 在检验结果解释方面, 丁是最强的一个, 在情形 1-3 中,

其解释都与甲乙丙中显著性解释最强的一个相同，详见表 1。”

表 1. 层次检验与依次检验、系数乘积的区间检验和中介效应差异检验比较

检验方法	情形 1	情形 2	情形 3	情形 4
甲：依次检验三对系数	检验结果 <b>显著</b> <b>信息：</b> 知道乙丙丁的结果也显著 <b>解释：</b> 明确知道中介路径上每个系数的显著性，以及前、后路径哪个受到调节	检验结果 <b>不显著</b> <b>信息：</b> 不知乙丙丁的结果 <b>解释：</b> 中介效应没有受到调节	检验结果 <b>不显著</b> <b>信息：</b> 与情形 2 相同 <b>解释：</b> 与情形 2 相同	检验结果 <b>不显著</b> <b>信息：</b> 与情形 2 相同 <b>解释：</b> 与情形 2 相同
乙：三个系数乘积的区间检验	检验结果 <b>显著</b> <b>信息：</b> 知道丙丁的结果也显著，但不知甲的结果 <b>解释：</b> 明确知道前、后路径哪个受到调节	检验结果 <b>显著</b> <b>信息：</b> 与情形 1 相同 <b>解释：</b> 与情形 1 相同	检验结果 <b>不显著</b> <b>信息：</b> 知道甲的结果不显著，但不知丙丁的结果 <b>解释：</b> 中介效应没有受到调节	检验结果 <b>不显著</b> <b>信息：</b> 与情形 3 相同 <b>解释：</b> 与情形 3 相同
丙：中介效应差异检验	检验结果 <b>显著</b> <b>信息：</b> 知道丁的结果也显著，但不知甲乙的结果 <b>解释：</b> 只知道中介过程受到调节，但不知哪个路径受到调节	检验结果 <b>显著</b> <b>信息：</b> 与情形 1 相同 <b>解释：</b> 与情形 1 相同	检验结果 <b>显著</b> <b>信息：</b> 与情形 1 相同 <b>解释：</b> 与情形 1 相同	检验结果 <b>不显著</b> <b>信息：</b> 知道其他三人的结果也不显著 <b>解释：</b> 中介效应没有受到调节
丁：层次检验	检验结果 <b>显著</b> <b>信息：</b> 与甲的情形 1 相同 <b>解释：</b> 与甲的情形 1 相同	检验结果 <b>显著</b> <b>信息：</b> 知道甲的结果不显著、乙丙的结果显著 <b>解释：</b> 与乙的情形 1 相同	检验结果 <b>显著</b> <b>信息：</b> 知道甲乙的结果不显著、丙的结果显著 <b>解释：</b> 与丙的情形 1 相同	检验结果 <b>不显著</b> <b>信息：</b> 知道其他三人的结果也不显著 <b>解释：</b> 中介效应没有受到调节
层次检验的优点	丁和甲的信息多于乙和丙，丁和甲的结果解释性也强于乙和丙	丁乙丙的检验力高于甲；丁的信息多于甲乙丙，丁和乙的结果解释性强于丙	丁和丙的检验力高于甲和乙；丁的信息多于甲乙丙	丁和丙的信息多于甲和乙

意见 2： Use the data set to provide more detailed illustration with more textual explanation than just using formulae.

**回应:** 在例子部分, 增加了对统计结果的文字解释:

“上述统计结果说明, 感恩 ( $X$ ) 经过复原力 ( $W$ ) 对学业成就 ( $Y$ ) 的中介效应受到生活事件 ( $U$ ) 的调节。当个体经历的生活事件属于平均水平时 ( $U=0$ ), 感恩 ( $X$ ) 经过复原力 ( $W$ ) 对学业成就 ( $Y$ ) 的间接效应占了总效应的三分之二左右。当个体经历较多的生活事件时, 这个间接效应变小, 直接效应变大; 当个体经历较少的生活事件时, 这个间接效应变大, 直接效应变小。从而, 对于经历较多生活事件的个体, 提高其感恩水平往往可以直接提高其学业成就; 但对于经历较少生活事件的个体, 应当重视提高其复原力, 才能更有效地提高其学业成就。”

**意见 3:** English abstract needs substantial editing.

**回应:** 已经重写了英文摘要。整段新写的如下:

“We discussed the different testing methods based on the most general type of model. The traditional testing method is the moderated causal steps approach, in which the regression coefficients are tested in sequence. Modern methods include the testing of the products of coefficients by using Bootstrap method or MCMC method, and testing of the difference between the maximum and minimum of the mediation effects. On the basis of the previous studies it can be summarized that the power of test with the moderated causal steps approach is the lowest among the three testing methods, whereas the testing of the difference between the maximum and minimum of the mediation effects has the highest power of test.

After comparing significant results of the three testing methods by reviewing the simplification, implication, information, and explanation, we concluded that the moderated causal steps approach should be recommended first; testing of the products of coefficients by using Bootstrap method or MCMC method should be treated as a backup; testing of the difference between the maximum and minimum of the mediation effects should be the final choice.

We proposed a hierarchical procedure for testing moderated mediation models as follows:

Step 1. Adopt the moderated causal steps approach to test the model. If the significant result is obtained from the test, the mediating effect is moderated. Otherwise, go to Step 2.

Step 2. Test the products of coefficients by using Bootstrap method or MCMC method. If any product is significantly different from zero, the mediating effect is moderated. Otherwise, go to Step 3.

Step 3. Test the difference between the maximum and minimum of the mediation effects. If the difference is significant, the mediating effect is moderated. Otherwise, the mediating effect is not moderated.”