

《心理学报》审稿意见与作者回应

题目：消极刻板印象对老年人医疗决策的影响及归因偏差的作用

作者：张宝山，金豆，马梦佳，徐冉

第一轮

审稿人 1 意见：

该研究通过实验一探究了老年人的刻板印象对医疗决策质量的影响，并且考察了在这个影响过程中归因偏差的中介作用。结果发现对老年人的刻板印象会对医疗决策质量产生负面影响，并且归因偏差起到了重要的中介作用；作者通过实验二进一步验证，对归因偏差进行干预可以有效缓解这种消极影响。本研究对于提高老年人的医疗决策质量具有一定的实践意义。但该文章写作逻辑比较混乱，方法和细节部分仍需做些调整，具体意见如下。

回应：感谢审稿专家的仔细审阅和对本文价值的认可，我们已经根据您的建议对文稿的写作逻辑、研究方法、讨论等部分进行了系统修改。同时修改过程中也关注了对具体细节的调整，希望修改后的论文能很好地解决了您的所有关注。

意见 1：归因偏差作为中介变量，是本研究的一个重要的内容之一，却没有体现在题目中。

回应：非常感谢审稿专家的宝贵意见。根据您的建议，我们已将文稿题目修改为“消极刻板印象对老年人医疗决策的影响及归因偏差的作用”。

意见 2：摘要中缺乏研究背景的适当描述，建议进行修改。

回应：感谢您的宝贵意见。根据您的意见，我们在摘要部分补充了对研究背景的描述，具体内容如下：

“随着年龄的增长，对老年人的消极刻板印象和医疗决策逐渐成为了与老年群体越来越相关的两个概念。在此背景下，很有必要系统地探讨消极刻板印象对老年人医疗决策的效应及其机制。”（详见中文摘要部分）

意见 3：引言部分缺乏逻辑性，建议适当调整一下这部分逻辑先后顺序。如，引言第二段，首先，第一句话“老年人相关的刻板印象对老年人广泛的心理行为结构都有着显著的影响。”，紧接着得出推论“因此，刻板印象很可能也是老年人医疗决策质量的一个重要影响因素。”然后再去描述刻板印象的定义，建议先描述刻板印象的定义，这里的逻辑顺序需要作者再自行斟酌和适当修改。其次，这部分只论述了刻板印象的定义，而没有展开论述以往研究中老年人相关的刻板印象对老年人的心理行为结构有什么样的显著影响，建议适当地增加相应的例子，使得这部分的推论更有理有据。第三，后面谈到“以往研究表明，刻板印象会通过刻板印象威胁对个体的决策过程和决策结果产生显著影响”，这句结论在这里显得太过单薄，建议增加相应的解释和例子，如，其中的刻板印象威胁是什么？提到的以往研究中的产生的显著影响是如何体现的？最后一句，“基于此，本研究将相关领域刻板印象威胁、归因偏差和老年人的决策质量整合到一个研究框架中，拟系统地探讨老年刻板印象对医疗决策质量的效应以及归因偏差的中介作用。”在此之前都没有和归因偏差相关的描述，更没有阐述归因偏差与主题的关系以及为什么要把归因偏差作为中介变量来研究的理由，读到这里会让读者

感到迷惑和不知所以。总而言之，引言部分的逻辑框架建议作者进行进一步的调整，优化这部分的逻辑性和可读性。

回应：非常感谢审稿专家的仔细审阅与宝贵意见。根据您的建议，为了确保阅读的连贯性以及文章逻辑的顺畅，我们在修改稿中对引言部分的逻辑框架进行了重新调整。

具体而言，在修改后的引言部分第二段，我们首先以“刻板印象是老年人行为决策的一个重要影响因素和预测变量”作为主题句，起到承上启下、连接上下文的作用。随后，明确了老年刻板印象的定义，接着说明了老年刻板印象存在的普遍性，并列举了一系列常见的老年刻板印象。在此基础上，我们引用多项研究结果佐证了刻板印象对老年人抑郁、焦虑、幸福感、认知能力等广泛的心理行为结构的显著影响。在此基础上，我们从一般领域聚焦到决策领域的相关研究，正式对消极刻板印象威胁与老年人决策的关系进行讨论。修改后的具体内容如下：

“大量研究表明，刻板印象是老年人行为决策的一个重要影响因素和预测变量。老年刻板印象指人们对老年人这一特定社会群体所持有的观念与预期(Levy et al., 2000)。世界上多数文化中都有着大量关于老年人的消极刻板印象。如人们普遍认为老年人认知功能差、缺乏能力、健康不良、反应迟钝，是需要被照顾的社会弱势群体(Coudin & Alexopoulos, 2010; Dionigi, 2015)。作为老年期最重要的社会心理结构之一，消极刻板印象对老年人广泛的心理结构都有着显著的影响。如以往研究显示，消极刻板印象的激活可以诱发老年人较高水平的抑郁和焦虑，以及较低水平的幸福感(Bryant et al., 2012)。同时，感知到较高水平的刻板印象还会加速老年人认知功能下降(Jordano & Touron, 2017)，增加生理功能的负荷(Levy et al., 2000)。除了对生理心理过程的效应外，刻板印象对个体的决策过程和决策结果具有显著影响也成为这一领域内众多研究者的共识。例如 Inzlicht 和 Kang (2010)的研究显示，刻板印象的激活降低了被试做出理性选择的能力，导致了更多的风险决策。Carr 和 Steele (2010)的研究也表明刻板印象激发了个体在决策中更高的损失厌恶倾向和更多的风险规避行为。”（详见第 2 页，第 2 段）

接下来在第三段，我们具体到医疗领域，总结了前人关于刻板印象与医疗决策关系研究存在的局限，并阐述了本研究的重要意义，具体内容如下：

“尽管已有研究关注了一般人群中刻板印象与决策的关系，也有研究探讨了消极刻板印象对特定老年病人决策行为的效应(Levy et al., 2000)，但关于消极刻板印象与健康老年人医疗决策关系的研究还相对较少，系统性也不强。诚如上所述，与其他领域的决策相比，医疗决策质量的水平无论对于老年病人还是健康老年人都有着尤其重要的意义。在这一背景下，消极刻板印象作为与老年群体密切相关的重要概念，是否也会对健康老年人医疗决策产生显著效应，是非常值得深入研究的一个问题。”（详见第 2 页，第 3 段）

然后，在第四段我们进一步提出对刻板印象效应机制的探讨，关注到了归因偏差在刻板印象和决策之间关系的重要作用。在该段落中，首先介绍了归因偏差的概念，然后补充了有关归因偏差在本研究中角色的简单讨论。具体内容如下：

“此外，消极刻板印象对老年人医疗决策效应的机制也需要相关领域的研究进一步关注。已有研究表明，归因偏差与刻板印象和决策行为都有着密切的联系(Koo & Yang, 2018; 王大伟, 刘永芳, 2008)。然而，尚未有研究检验归因偏差在刻板印象对行为决策效应中所扮演的角色。一般来说，归因偏差是指个体非理性地将事件结果归因为某个特定原因而忽视其真正原因的归因倾向(Casey et al., 2015)。研究表明，人们对某一群体固定的观念或预期能够影响个体对他人及自身行为的归因倾向，使个体在进行因果归因时存在一定程度的偏差(Fiske & Susan, 1992; Oakes et al., 1991)。另一方面，归因偏差又被证实是决策质量的一个显著的影响因素(王大伟, 刘永芳, 2008)。因此，归因偏差很可能是消极刻板印象对老年医疗决策质量效应的一个中介变量。”（详见第 3 页，第 2 段）

最后，我们对本文的研究框架做出了总结概括。具体内容如下：

“基于以上文献，本研究将相关领域刻板印象、归因偏差和决策质量整合到一个研究框架中，拟系统地探讨消极刻板印象对老年人医疗决策质量的效应以及归因偏差的中介作用。接下来，我们将就这三者的关系展开文献综述，并提出本研究的理论模型和假设。”（详见第3页，第3段）

意见 4：实验设计上描述不完整，如在归因偏差测量任务中，被试按键选择的时间有限制吗；医疗决策任务是只进行了一个回合吗？

回应：感谢专家的宝贵意见。在归因偏差测量任务中，被试按键选择的时间没有限制，直到被试进行按键反应后自动进入下一个试次。我们已经在修改稿中对这一问题进行了澄清，内容如下：

“考虑到老年人认知缓慢这一特点，被试按键选择没有时间限制。在被试完成选择后，实验程序会自动进入下一试次。”（详见第6页，最后一段）

其次，在医疗决策任务中，我们向被试提供了 A、B、C、D 四种治疗疾病的方案，对于每种方案来说，八条有关治疗方案属性的描述依次向被试呈现，最后由被试对每种方案进行评分。一个治疗方案评价后，会以相同的程序随机呈现下一个治疗方案的每条描述。因此，每个被试都按照随机顺序对四种方案进行了评分。我们将被试对最佳治疗方案（B 方案）与较差治疗方案（D 方案）的评分差值作为任务的指标，来衡量被试的医疗决策水平。医疗决策任务中，每个被试对四种方案都会进行评分。我们在修改稿中也加入了这部分细节，描述如下：

“在该任务中，主试首先向被试提供 A、B、C、D 四种治疗某种疾病的方案，然后由被试对这四种方案分别进行评分。”（详见第7页，第3段，第2行）

“被试对一个治疗方案评价结束后，电脑屏幕将以相同的程序随机呈现下一个治疗方案的描述。在完成对四种医疗方案的评价后，我们将被试对最佳治疗方案（B 方案）与较差治疗方案（D 方案）的评分差值作为该任务的指标。差值越大说明被试越能够区分最佳和较差的方案，医疗决策质量越高。”（详见第8页，第1段，最后4行）

意见 5：文中的“操作检验项目”不具有任务的概括性，读者不太容易从这个名字中联系起具体的任务，建议改成“刻板印象检验项目”或其他更具有辨识性的任务名称。

回应：感谢审稿专家的宝贵意见。我们已经在修改稿中将“操作检验项目”更改成为“刻板印象启动操作检验”。（详见第9页 2.2.1 标题和第12页 3.2.1 标题）

意见 6：图表部分，在行标题或图注应该把图表中的数字代表的含义写清楚，如表 1 中第三列和第四列的数字代表什么，括号中的是标准差还是标准误？

回应：感谢审稿专家的仔细审阅和宝贵意见。原稿表 1 第三列和第四列（修改后是第二列和第三列）的数字代表的是不同组别相关变量的均值和标准差，或是被试数量和相应的百分比率。我们已经在修改稿的表注中澄清了此信息，具体如下：

“注：自评健康为 5 点计分，1 代表非常不健康，5 代表非常健康，得分越高，代表自评健康越好。生活满意为 5 点计分，1 代表非常不满意，5 代表非常满意，得分越高，代表生活满意度越高。括号内是标准差(年龄、自评健康、生活满意度、归因偏差及决策质量)或百分比(性别和文化程度)。*表示 $p < 0.05$ ，**表示 $p < 0.01$ ，***表示 $p < 0.001$ ，下同。”（详见第10页表1表注部分）

同样地，我们也对其他表格的表注进行了完善，修改后，所有表格的表注都能够清楚地说明表内数据的含义，不再存在含混不清的问题。

意见 7: 本研究的创新点存在感较弱, 建议在总讨论中增加与以往研究对比, 本研究不同的研究意义是什么?

回应: 感谢审稿专家的建议。与以往相关领域研究相比, 我们研究的创新点主要体现在两个方面:

首先是检验了消极刻板印象对一般(或正常)老年人医疗决策的效应。以往尽管有些研究探讨了刻板印象对医疗决策或医疗行为的影响(Levy et al., 2000; Stewart et al., 2012), 但这些研究都只关注了正在接受治疗的老年病人群体, 研究的问题也只是探讨了消极刻板印象对一些非常具体、概括性比较低的单一医疗行为的效应(如是否选择继续治疗或是否选择进行复健活动等), 研究结果并不能推广到一般老年人群, 适用范围较小。研究结果的理论与实践价值都比较有限。

其次, 我们揭示了归因偏差在刻板印象威胁对老年人医疗决策效应中的中介作用。步入老年期后, 内归因偏差是老年个体对疾病等健康问题进行归因时非常容易出现的倾向, 对老年人后续的健康行为有着重要的意义。但是, 以往并没有研究重视归因偏差在老年人心理行为过程中的角色, 更没有研究关注归因偏差在刻板印象威胁与老年人医疗决策质量关系中的中介机制。

基于以上论述, 我们检验了消极刻板印象对一般老年人医疗决策的效应, 并进一步探讨了归因偏差的中介作用。

根据您的建议我们补充了相关领域前人研究的相关内容, 并与我们的研究进行了对比, 阐明了本研究方向的独特性、研究内容的必要性以及本研究的创新点, 具体修改内容如下:

“本研究丰富了老年刻板印象效应研究领域和行为决策研究领域的理论体系, 具有一定的创新性和理论价值。以往相关领域的研究表明, 刻板印象对个体的风险寻求、求职意向、交友偏好等多种领域的决策都有着显著的影响(Mather et al., 2012; 李爱梅等, 2009; 杨婷, 任孝鹏, 2016)。但是, 在老年心理研究领域, 仅有一些零散的研究探讨了消极刻板印象对老年病人一些比较具体、单一的医疗行为(如是否选择继续治疗或是否选择进行复健活动等)的效应(Levy et al., 2000; Stewart et al., 2012), 并没有研究关注在一般老年人群中, 消极刻板印象与医疗决策的关系。很显然, 由于研究对象局限于患病的老年人, 且决策任务过于简单、具体、仅适用于老年病人等问题, 因此, 以往研究结论的适用范围较小, 理论价值有限。除此之外, 在相关领域文献中, 也没有研究系统地揭示消极刻板印象对老年人医疗决策质量效应的具体机制。基于这些局限, 本研究使用了在医疗领域更具代表性的决策任务, 探讨了消极刻板印象对一般老年人医疗决策的效应, 并进一步检验了归因偏差的中介作用。本研究的结果除了证实消极刻板印象对一般老年人的医疗决策任务具有消极效应外, 更揭示了归因偏差的中介机制角色。

实践方面, 本研究结果对于减少消极刻板印象威胁效应和提高老年人决策质量具有一定的价值。本研究结果表明, 在刻板印象威胁情境中, 归因偏差的训练控制可以有效加强老年人对疾病的外归因偏差倾向, 减少老年人内归因偏差倾向。在此基础上, 消极刻板印象对老年人医疗决策的效应也得到了有效的改善。这一结果提示我们, 在减少消极刻板印象威胁效应或提高老年人决策质量的实践中, 相比于改变较为稳固的刻板印象水平, 恰当地使用归因训练对提高决策质量有着非常明显的优势。具体地说, 基于本研究的归因偏差训练任务较为简单、有效且容易操作, 对于环境、设备和实践者的专业水平都没有非常严苛的要求。基于这些优点, 归因偏差训练有着更加广泛的适用范围。因此, 在相关领域的实践中, 本研究结果有着较好的应用前景。”(详见第 17 页, 1-2 段)

意见 8: 细节部分, 文中有些句子的表达不够简洁和通顺, 如, 引言第一段的第二句话, “在此背景下, 老年人很容易需要在多种医疗服务中进行选择。”其中“很容易需要”表述冗杂

多余；2.3 讨论的第二句，“这可能是由于消极的刻板印象可能会导致老年人的决策表现偏向于与刻板印象相一致的结果”中的两个“可能”使得这个句子表达不通畅。

回应：感谢审稿专家的宝贵意见。按照您的建议我们做了如下修改：

将第 2 页，第 1 段，第 2 行的“很容易需要”改成了“经常需要”；

将第 11 页，第 2 段，第 2 行的“这可能是由于消极刻板印象可能会导致老年人的决策表现偏向于与刻板印象相一致的结果”改成了“这一结果与 Lockenhoff 和 Corinna (2017)的结论相一致，即在刻板印象威胁情境中，个体的决策表现可能会倾向于与刻板印象保持一致”。

此外，我们对文章整体语言都进行了仔细的检查，对表达不通畅的语句逐一做出了修正。在本轮修改中，我们对每个句子都进行了认真打磨，绝大部分的句子都进行了一定程度的完善。由于涉及到的语言表达方面的修改实在太多，这里就不再一一列举这方面的修改。希望通过本轮修改，文稿语言表达的流畅性和可读性能够得到显著提升。

.....

审稿人 2 意见：

该研究试图探讨老年刻板印象对老年人医疗决策质量的影响，作者从刻板印象、归因偏差及二者对医疗决策的影响切入，探讨三者之间的关系模式，发现归因偏差在老年刻板印象对老年人医疗决策质量的效应中起到中介作用。研究具有一定的创新性和应用价值，然而还存在一些问题与不足。

回应：感谢审稿专家的辛苦工作以及对本文价值的认可，我们已经根据您的建议对文稿进行了系统地完善，具体修改请见后面的回复。希望修改后的论文能够很好地解决了您的所有关注。

意见 1：研究中的被试群体均为老年人，所以本研究结论是否只适用于老年群体？对老年的刻板印象所引起的老年相关医疗决策效应，是否适用于中青年群体？请在文章标题及主要讨论、结论中阐明。

回应：非常感谢审稿专家的宝贵意见。本文关注老年刻板印象对老年人医疗决策质量的效应。根据刻板印象威胁理论(Steele & Aronson, 1995)，消极刻板印象主要对目标群体成员产生心理压力与消极影响。在当前研究中，我们启动了与老年相关的刻板印象，即关于老年人各项认知能力衰退，决策质量较差的消极刻板印象。因此就所启动的刻板印象领域而言，老年人是当前刻板印象启动的目标群体，而中青年群体并不是此类刻板印象的目标群体。所以，本研究的结果并不适用于中青年群体。

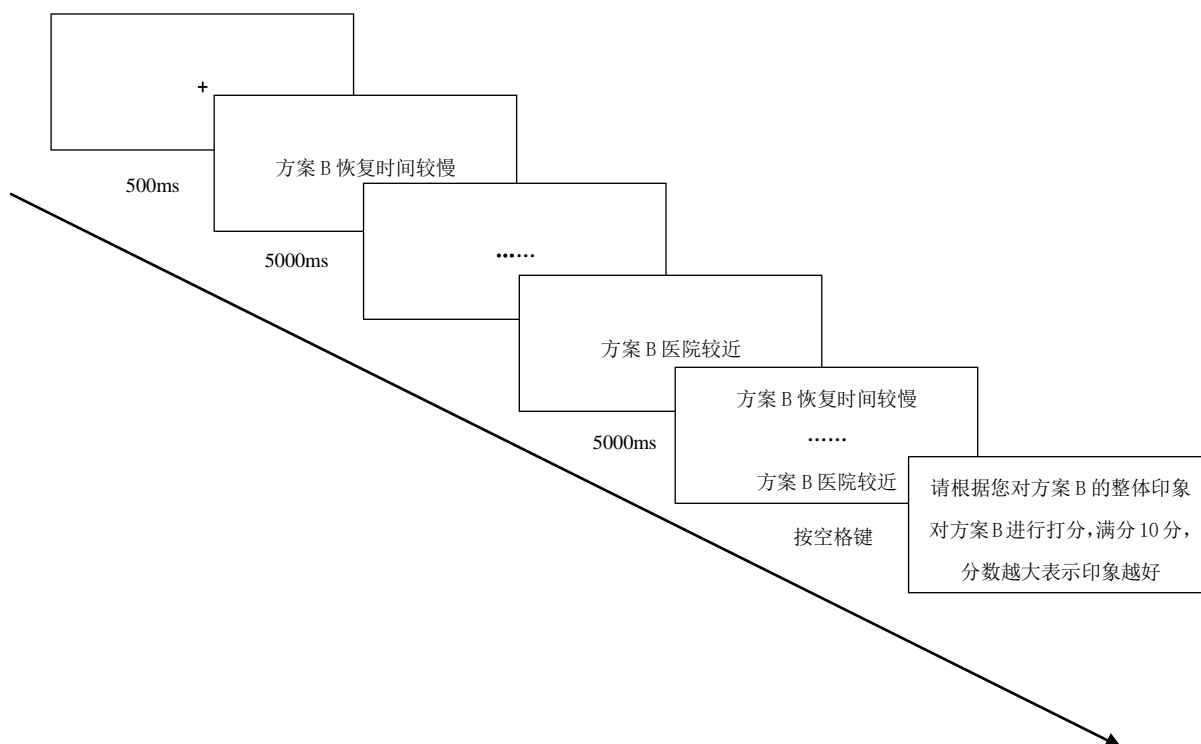
根据您的建议，我们已经在文章的标题做出了相应修改，修改后的题目为：

“消极刻板印象对老年人医疗决策的影响及归因偏差的作用”。（见修改版标题）

其次，我们在讨论中也做出了对应的修改，在相应的位置均强调了我们的研究结果仅适用于老年人群。（见修改版的讨论部分）

意见 2：医疗决策任务作为主要的任务，建议作者提供任务的流程图，方便读者尽快理解任务内容和指标。

回应：非常感谢审稿专家的宝贵意见。根据这一意见，我们已在修改稿中加入了医疗决策任务的流程图，以方便读者尽快理解任务内容。具体流程图如下所示。（详见修改稿第 8 页，图 2）



3. 实验组别进行哑变量编码为二分变量后，作者采用的哪种相关分析方法计算与其他变量之间的关系？请阐明具体方法及其合理性。

回应：非常感谢审稿专家的宝贵意见。在本研究中，将实验组别进行哑变量编码为二分变量后，我们采用了 SPSS 中皮尔逊积差的方法计算了它与其它变量之间的关系。我们查阅了《现代心理与教育统计学》（张厚粲，2015）等传统的统计教材发现，尽管严格来讲二分类别变量与连续变量之间的相关关系应该采用点二列相关来度量，但实际上，点二列相关仅是皮尔逊积差相关的特例，点二列相关的计算在本质上等同于一列变量取值为 0、1，另一列变量为连续数据情况下的皮尔逊积差相关。将这一类相关单列出来主要是为了在计算机技术不发达的时期简化手工计算过程，而在 SPSS 等统计分析软件高度发展的现今直接使用皮尔逊积差相关计算即可（Gravetter & Wallnau, 2009; Howell, 2013; 卢谢峰等，2011）。我们已在实验 1“结果”部分相应地补充了这部分内容，详细内容如下：

“首先，将实验组别进行哑变量编码，将刻板印象威胁组编码为“1”，将控制组编码为“0”。然后对组别、归因偏差和医疗决策质量三个变量进行皮尔逊积差相关分析。”（详见第 9 页，最后一段）

意见 4：实验 1 中刻板印象组和控制组在受教育程度上是否差异显著？

回应：非常感谢审稿专家的宝贵意见。我们在实验 1 使用了卡方检验来检验刻板印象组和控制组在受教育程度上是否存在差异，结果显示，两组的受教育程度不存在显著差异， $\chi^2 = 4.77$ ， $p = 0.19$ 。我们也在修改稿中报告了上述检验结果，具体内容如下：

“独立样本 t 检验及卡方检验的结果表明，刻板印象威胁组被试和控制组被试在自评健康($t = -0.56$, $p = 0.58$)、年龄($t = -0.55$, $p = 0.58$)、生活满意度($t = 0.34$, $p = 0.74$)、性别($\chi^2 = 0.10$, $p = 0.75$)、受教育程度上($\chi^2 = 4.77$, $p = 0.19$)不存在显著差异。”（详见第 6 页，第 1 段）

意见 5: 两个实验中的无关变量均未控制, 如果作为协变量控制后, 主要结论是否还能成立?

回应: 非常感谢审稿专家的宝贵意见。对于两个实验中的自评健康、年龄、生活满意度等无关变量, 我们都进行了 t 检验或卡方检验等统计方法来检验这些变量是否在实验组和控制组之间存在差异。实验 1 独立样本 t 检验及卡方检验的结果表明, 刻板印象威胁组被试和控制组被试在自评健康($t = -0.56, p = 0.58$)、年龄($t = -0.55, p = 0.58$)、生活满意度($t = 0.34, p = 0.74$)、性别($\chi^2 = 0.10, p = 0.75$)、受教育程度上($\chi^2 = 4.77, p = 0.19$)不存在显著差异(详见第 6 页, 第 1 段)。

而实验 2 独立样本 t 检验及卡方检验结果亦表明, 刻板印象威胁组被试和干预组被试在自评健康($t = 0, p = 1$)、年龄($t = -0.61, p = 0.54$)、生活满意($t = -0.31, p = 0.76$)、性别($\chi^2 = 1.92, p = 0.17$)、受教育程度($\chi^2 = 4.31, p = 0.37$)上不存在显著差异(详见第 12 页, 第 2 段, 4-6 行)。我们认为这些结果在一定程度上支持了无关变量对被试任务表现的影响在两组间无显著差异。

除了以上检验外, 我们也尝试将这些变量作为控制变量重新对研究的主要结果进行了分析。结果表明, 在控制了自评健康、年龄、生活满意度等无关变量后, 实验 1 中刻板印象威胁组被试的归因偏差指标 ($M = 3.33, SD = 1.38$)显著高于控制组 ($M = 1.57, SD = 0.93$), $F(1, 77) = 42.67, p < 0.001, \eta^2 = 0.38$, 刻板印象威胁组内部归因倾向显著高于控制组; 同时, 刻板印象威胁组被试的决策质量 ($M = 0.15, SD = 1.44$)显著低于控制组 ($M = 2.61, SD = 1.31$), $F(1, 77) = 57.13, p < 0.001, \eta^2 = 0.45$ 。中介效应检验仍然表明, 刻板印象正向预测被试的归因偏差, $b = 0.60, SE = 0.09, t = 6.53, p < 0.001, 95\% CI = [0.419, 0.787]$, 归因偏差负向预测被试的医疗决策质量, $b = -0.26, SE = 0.11, t = -2.35, p = 0.021, 95\% CI = [-0.474, -0.039]$ 。归因偏差在刻板印象对医疗决策质量的预测模型中起到中介作用, 直接效应值为 $-0.51, SE = 0.11, 95\% CI = [-0.720, -0.292]$; 间接效应值为 $-0.15, SE = 0.07, 95\% CI = [-0.314, -0.027]$ 。

相似地, 在将被试的年龄、健康状况、生活满意、性别和文化程度作为协变量纳入实验 2 分析后, F 检验结果也显示, 刻板印象威胁组被试的归因偏差指标 ($M = 3.96, SD = 4.03$)显著高于干预控制组 ($M = 0.16, SD = 0.26$), 威胁组被试呈现出更多的内归因倾向, 而干预组则呈现出更多的外归因倾向, $F(1, 79) = 35.74, p < 0.001, \eta^2 = 0.33$; 此外, 归因偏差干预组被试的医疗决策质量 ($M = 3.00, SD = 1.75$)显著高于刻板印象威胁组 ($M = 0.65, SD = 1.85$), $F(1, 79) = 34.46, p < 0.001, \eta^2 = 0.32$ 。

综合以上结果可知, 自评健康、年龄、生活满意度等变量对我们的假设模型拟合度、路径参数估计等结果没有实质的影响, 并未影响我们所关注变量之间的效应以及对于结果的解释。

然而, 一些研究者认为, 要谨慎地选取控制变量, 包含过多的控制变量的分析可能会被误解为“方法学上的投机取巧”, 而不适当地加入控制变量可能会产生误导性的结果。因此要确保每一个加入分析的控制变量都有理论上充足且合理的依据 (Becker, 2005)。综合以上理论与分析, 我们认为当前研究中加入年龄、健康状况、生活满意、性别和文化程度作为控制变量可能会削弱统计效力。因此我们本着“对控制变量的使用持保守态度”(Becker, 2005), 在最终分析中没有纳入自评健康、年龄、生活满意度等控制变量, 而是在被试信息部分进一步强调了两组被试在这些变量上不存在显著的差异(因为理论上讲, 如果两组在这些控制变量上面没有显著的差异, 则可视两组为同质组, 那么控制变量对最终的实验结果就不应该存在显著的影响), 详见修改稿两个实验“被试”部分。

意见 6: 作者“将实验 1 控制组被试操作检验的得分与实验 2 两组被试的得分进行独立样本 t 检验”是何目的? 为何进行多次 t 检验?

回应：感谢审稿专家的仔细审阅和宝贵建议。在实验 2 中，所有被试（包括归因偏差干预组被试和刻板印象威胁组被试）均需启动消极刻板印象。为了检验实验 2 的刻板印象启动操作的效果，最大限度保证实验 2 严谨性，我们将实验 1 控制组被试操作检验的得分与实验 2 两组被试的操作检验得分进行了独立样本 t 检验。尽管实验 1 被试与实验 2 被试可能不是同一个总体，但是两个样本实际上是来自同一个取样群体的独立样本。此外，实验 2 中两组被试与实验 1 中刻板印象启动组被试刻板印象威胁启动的方法和程序是完全相同的。因此，本着经济原则，我们在实验 2 中并未单独设置控制组。我们将实验 1 的控制组作为老年被试没有启动刻板印象时的基线水平，将实验 2 两个组的刻板印象操作检验项目的分数与实验 1 控制组的分数进行对比。尽管不够完美，但这种做法在很大程度上足以证明实验 2 刻板印象启动操作的有效性。

此外，由于涉及到三组操作检验分数的比较，多次 t 检验的做法确实有误。在本轮修改中，我们使用了单因素方差分析的方法来检验三组被试刻板印象操作检验的分数。基于以上回复我们也在修改稿的相应部分进行了澄清，具体内容如下：

“为了检验实验 2 两组被试刻板印象启动操作的效果，我们将实验 1 控制组刻板印象操作检验的分数作为老年被试没有受刻板印象影响时的基线水平，并将实验 2 两组被试操作检验项目的得分与实验 1 的控制组进行比较。这样做的理由有二：其一，实验 2 的被试与实验 1 的被试是来自同一个取样群体的独立样本，可以视为同质性样本。其二，实验 2 两组被试与实验 1 刻板印象威胁组被试消极刻板印象威胁启动的方法和程序是完全相同的。

单因素方差分析的结果表明三组被试的刻板印象操作检验分数存在显著的差异， $F(2, 115) = 13.18, p < 0.001, \eta^2 = 0.19$ 。多重比较（LSD）的结果表明，实验 2 归因偏差干预组刻板印象启动操作检验分数 ($M = 9.90, SD = 2.60$) 显著高于实验 1 控制组 ($M = 7.89, SD = 2.81$), $p = 0.001$ 。实验 2 刻板印象威胁组被试刻板印象操作检验分数 ($M = 10.85, SD = 2.35$) 显著高于实验 1 控制组 ($M = 7.89, SD = 2.81$), $p < 0.001$ 。此外，实验 2 两组被试刻板印象操作检验得分没有显著差异， $p = 0.10$ 。该结果表明在实验 2 中刻板印象启动有效。”（详见第 13 页，第 1 段和第 2 段）

意见 7：归因偏差干预后，是否还能在老年刻板印象对老年人医疗决策质量的效应中起到中介作用？

回应：感谢审稿专家的宝贵意见。根据中介效应因果链检验设计的范式，实验 2 是在实验 1 结果的基础上，对实验 1 中介变量进行干预控制，进而对实验 1 结果进一步验证。因此，就实验设计来说，实验 2 的自变量是归因偏差的干预与否，而不是刻板印象。其次，实验 2 的目的是通过归因训练来减少刻板印象威胁情境中老年人内归因偏差的倾向。这就要求两组被试都首先处于刻板印象威胁情境之中，因此两组被试同时启动了刻板印象。基于以上两点，在实验 2 数据分析中无法检验刻板印象对医疗决策的影响，以及归因偏差的中介作用。

尽管存在方法上的限制，为了能够最大限度解决专家的疑问，我们将从以下两方面回答这一问题。在实验 1 中，我们发现了刻板印象威胁情境和控制情境下被试归因偏差和医疗决策质量存在着显著的差异，并进一步揭示了归因偏差在消极刻板印象对医疗决策质量的效应中起到的中介作用。在这一研究结果的基础上，尽管无法再次通过统计模型检验归因偏差的中介作用，但是实验 2 可以从另一个角度支持归因偏差的中介作用。理论上讲，如果归因偏差能够中介刻板印象威胁对老年人医疗决策质量的效应，那么通过实验操作控制了归因偏差的水平后，刻板印象威胁对老年人医疗决策的效应就应该会受到影响。实验 2 的结果也正符合了这一假设。具体地说，在启动刻板印象威胁的情境下，与没有接受训练的老年人相比，接受归因偏差控制训练的被试内归因偏差倾向显著减少，外归因偏差倾向显著增强，其医疗

决策的质量也较高。这一结果在一定程度上进一步验证了实验 1 的结果，即归因偏差在老年刻板印象对医疗决策质量的效应中起到了中介作用。

另外，为了确保实验 2 中归因偏差与医疗决策质量之间也存在着密切的关系，进一步支持实验 1 的结果，我们将两组（干预控制组和刻板印象威胁组）被试归因偏差指标与医疗决策质量相关情况整合到了修改稿中，这些数据可以更加明确归因偏差与医疗决策质量的关系，为归因偏差的中介角色提供另一个证据。新增加的内容如下：

“为了进一步验证归因偏差与医疗决策质量之间的关系，我们进行了相关分析。结果表明，控制干预组老年人的归因偏差指标与医疗决策质量具有显著的相关关系， $r = -0.38$ ， $p = 0.016$ ；刻板印象威胁组老年人的归因偏差指标也与医疗决策质量具有显著的相关关系， $r = -0.40$ ， $p = 0.011$ 。这表明在刻板印象威胁情境下，无论干预与否，归因偏差始终与老年人的医疗决策质量有着密切的联系，归因偏差指标越高（代表内归因倾向越明显），老年人医疗决策质量越差。”（详见第 13 页，“3.2.2 相关分析”部分）

审稿人 3 意见：

文章采用实验法，关注消极的老化刻板印象对老年人医疗决策产生的影响，并验证了内部归因偏差的中介效应，在此基础上对归因偏差进行干预，以分析干预对老化刻板印象的消极影响，具有一定现实意义。但是，文章还需要对以下几点加以澄清。

回应：感谢审稿专家对文章意义的认可。我们已经根据您的意见对文章进行了较为系统地修改。希望修改后的文稿能够较好地解决了您的所有关注。

主要问题：

意见 1：正如文章提到本研究的不足之一是未探究积极的老化刻板印象对老年群体医疗决策质量的影响，那么文章仅研究了消极老化刻板印象，建议在题目中加以限定为“消极老化刻板印象”。

回应：感谢审稿专家的宝贵意见。在当前修改稿中，我们已经在文章题目限定了刻板印象性质，修改后的文章题目为“消极刻板印象对老年人医疗决策的影响及归因偏差的作用”。

意见 2：引言部分在介绍老化刻板印象时，还需要更明确地指出其具体含义，以及本研究关注某一领域（例如决策能力）的刻板印象，还是所有领域的刻板印象（还包括“老年人健康状况不佳”）。因为实验的刻板印象启动环节包含两个检验项目，其中之一是关于老年人身体健康不佳是由于年龄影响，是否与决策能力有关。

回应：非常感谢审稿专家的建议。

首先，我们在当前修改稿引言部分，加入了老年刻板印象的定义，即“老年刻板印象指人们对老年人这一特定社会群体所持有的观念与预期(Levy et al., 2000)。”（详见第 2 页，第 2 段，第 2 行）。

然后我们列举了关于老年人常见的刻板印象，并进一步论述了刻板印象对老年人广泛的心理行为结构的影响，最终落脚于决策领域，明确了我们的研究聚焦于决策领域的老年刻板印象而不是一般领域（详见第 2 页，第 2 段）。

其次，正如专家所言，我们在最初阶段除了启动决策的刻板印象，还启动了有关老年人身体健康的刻板印象。这主要是由于身体健康状况不佳不仅是一种典型老年刻板印象，也是一种对老年期生理性特点的客观阐述，即老年人在老年期身体健康状况不如以前好，更容易生病，有更多进入医院并对医疗方案进行决策的经历和机会。因此，这一表述在当前语境中

可以作为背景，顺利引出对于老年决策能力受损的描述。这样的内容逻辑更加严谨，指导语的可信度更高。

此外，Vailati Riboni 和 Pagnini (2021)的一项元分析表明，老年刻板印象启动方式主要可以分为依据指示（如使用来自口头或书面的指令信息操纵刻板印象），依据测试（如使用特定的测试激发对象的刻板印象）和依据心理意象（如通过操纵老年人的心态诱发刻板印象）三类，其中依据指示是目前最为常见且有效的刻板印象启动方法。本研究也采用了这种启动方式。而依据指示启动刻板印象也有不同的操纵形式，其中最常采用的手法之一就是针对更广泛的刻板印象，向老年人提供侧重于由于衰老过程而产生的各种生理差异的材料（如本研究中关于老年人身体状况不佳的描述），以此来启动老年刻板印象。以往的研究已经证实这样的启动方式在多个不同的领域都可以成功诱发相关的刻板印象威胁。另一种依据指示的主要操纵形式是伴随着不同的能力诊断性测试的实施而提供的口头或书面指示，直接指向目标行为领域（如本研究中关于老年人决策能力的描述）。本研究结合了这两种操纵形式，指示操纵从老年人的一般领域出发，落脚于决策相关领域。“身体健康状况不佳”作为背景描述出现，实质上属于一般领域的刻板印象，暗示了老年人在这一些方面的不足，与本研究启动的特定领域的刻板印象并不冲突，且存在一定的递进关系，可以流畅地将我们的关注点指向决策相关领域。

意见 3: 实验 2 旨在验证在干预内部归因偏差后，消极老化刻板印象对医疗决策质量的消极影响是否减弱，从而说明干预措施的有效性。问题在于，实验 2 仅通过单次测量内部归因偏差，无法证明干预组被试内部归因占比低于威胁组是干预措施造成的。实验 2 在逻辑上还需要前测证明在干预前，干预组与威胁组在归因偏差上可被视为同质群体。

回应：非常感谢审稿人的宝贵意见。在实验 2 中，我们确实没有对归因偏差进行前测，主要有以下几点原因：

首先，我们采用的测量归因偏差的任务不适合反复使用。在实验中，我们采用了 Trent 等人在 2019 年使用的 The Word Sentence Association (WSAP) 任务来测量被试的归因倾向。这个任务要求被试在每个试次中通过按键选择健康不佳症状对应的原因。WSAP 任务共包括 10 个练习试次和 70 个正式试次。在实验中，如果我们提前使用这个任务建立基线，后面再用这个任务训练（干预组）或测量（控制组）被试的状态性归因偏差，就会在短时间内反复使用这个任务，最终导致被试对这个任务产生厌烦等消极情绪，或者形成不当的归因定势，影响实验控制干预的效果，结果很可能被扭曲。

其次，如果使用其它任务测量归因偏差，也难以证明实验前两组被试的归因偏差一致。如果使用其它任务测量归因偏差，这样获得的归因偏差与我们关于归因偏差的操作定义不符，这和其他控制变量在本质上是一样的，只能在某种程度上说明两组可能一致，仍然无法完全保证两组在我们关注的变量上是同质的，对于支持两组归因偏差的同质性并无明显贡献。

第三，在理论上，我们对实验 2 的前期控制基本上可以证明两组是同质的，在归因偏差上没有显著差异，不会对后面的实验操作产生显著的影响。具体而言，在实验 2 中，我们在同一个群体中招募被试，并且将招募到的被试随机分配至归因偏差干预组和刻板印象威胁组。如果总体中的任意一位成员都有等同的机会被抽取并分配到任意一个处理组，那么可以预期这种随机分配形成的各种处理组的各种条件和机会是均等的（郭秀艳，2004）。此外，我们也根据统计检验力的要求，招募到了符合要求的样本的数量，这也能在一定程度上保证了本研究样本对于目标群体来说具有了充分的代表性。在这些控制的基础上，我们还对年龄、性别等所有主要背景变量进行了差异检验，结果证实两组不存在显著差异，因此，一定程度

上可以证明两组为同质群体，不存在结构性差异。也就是说，在其他背景变量都相同，两组唯独在特质性归因偏差存在显著差异的可能性应该也不大。

基于以上原因，我们认为两组被试在干预前是同质的，仅存在统计允许限度以内的随机误差。我们也在当前修改稿中对这一点进行了澄清：

“另外，考虑到干预组与对照组被试随机分配，不存在结构性差异，为了避免对后续任务造成干扰，并尽量减少老年被试认知压力，我们并未对归因偏差进行前测。”（详见第 12 页，第 5 段，7-9 行）

意见 4: 文章提到，在实验的归因偏差测量中，内部归因与外部归因的选择比例以 1 为基准，划分为内部归因偏差和外部归因偏差，文章并没有阐述这个划分标准和结果在研究中的意义。实验 2 的干预组的归因偏差得分远小于 1，应该是出现了外部归因偏差；也就是说，干预的结果不是“不偏差”而是让被试出现另一种“归因偏差”。那么讨论需要对此深入分析和解释，本文所提到的“归因偏差效应”就需要限定在“内部归因偏差效应”。

回应: 非常感谢审稿专家的仔细审阅。在本研究中，我们有两个关于归因的指标，包括内归因和外归因。如果在进行数据分析时将内部归因和外部归因作为两个指标可能会带来一些明显的局限性。首先，数据部分可能会太复杂。相较于当前的数据结果，分别使用内部归因和外部归因作为两个指标的话，数据量会增加一倍，结果的解释也会比较复杂。其次，两种归因倾向的指标会导致缺少对应的情况作为基线水平。没有比较的基线，就无法定义归因偏差。

基于以上原因，我们采用了当前的方式来定义归因偏差，即将被试内部归因比例与外部归因比例的比值作为归因偏差的指标。理论上，不存在归因偏差时，两者应该相同，比值为 1；当比值小于 1 时，则说明外部归因比例大于内部归因，存在外部归因偏差；当比值大于 1 时，则说明内部归因比例大于外部归因，存在内部归因偏差。这样，只使用一个指标就可以同时反映内部归因和外部归因的相对关系，以及归因偏差的方向，结果的呈现更加清晰简洁与直观。根据您的意见，我们在原文中加入了以下内容来阐明这一做法的意义：

“为了简化数据分析过程，清晰简洁地呈现研究结果，我们使用被试内部归因比例与外部归因比例的比值作为归因偏差的指标。若比值为 1，则表示不存在归因偏差；若比值小于 1 则表示被试存在外部归因偏差；若比值大于 1 则表示被试存在内部归因偏差。比值越高，内归因偏差水平越高，外归因偏差水平越低。”（详见第 7 页，第 2 段）

对于您的第二个问题，依据我们的归因偏差指标的含义，当比值小于 1 时，表示被试存在外归因偏差，实验 2 干预组的归因偏差得分也显示其出现了外归因偏差。为了更准确地体现出这一结果，我们修改了原稿中“威胁组被试的内部归因偏差程度显著高于干预组”这种笼统的说法，对归因偏差的结果进行了更精确的描述，并明确区分了内归因偏差和外归因偏差，修改后的具体内容如下：

“采用独立样本 t 检验比较两组被试归因偏差的水平。结果显示，干预组归因偏差指标显著低于刻板印象威胁组(平均数见表 2)， $t = -5.96$ ， $p < 0.001$ ， $Cohen's d = -1.33$ 。这表明，与刻板印象威胁组相比，干预组老年人的内归因偏差水平显著降低，外归因偏差水平显著升高，即对归因偏差干预有效地减少了被试的内归因偏差水平。”（详见第 14 页，第 2 段）

根据我们的归因偏差结果，确实如您所言，干预组出现了外部归因偏差。也就是说，干预后的结果不是归因偏差消除了，而是归因倾向转变了，即由内归因偏差转变成为了外归因偏差。这主要是由于归因具有高度主观性，在日常生活中，人们的归因往往是基于有限的信息基础且掺杂着个人的偏好，因而很容易产生归因偏差。完全没有偏差的客观归因实现起来非常困难，大部分情况下都是个体在自身的认知范围内搜索可能的因果关系尽力达到相对平衡的归因(Harvey et al., 2006)。我们也将此作为本研究的一个局限进行了深入分析和解释，具体内容如下：

“其次，本研究只证实了内归因偏差的局限和外归因偏差的优势，但无法确定无偏差归因是否与老年人医疗决策质量也存在着一定的关系。尽管对疾病或健康问题的外部归因优于基于衰老的内归因偏差，对老年人是有利的归因方式(Vassilopoulos et al., 2009)，但是，这种外归因偏差的优势是建立在一定的限度之内的，过度的偏差可能与一定的心理困扰相关(Sanjuán & Magallares, 2014)。因此，寻找一个相对平衡的无偏差的归因方式对个体的理性决策可能具有更关键的积极意义。然而，我们实验结果并没有训练出无偏差的归因方式，而是在归因干预控制后，被试表现出明显的外部归因偏差倾向。也即是说本研究结果并不能揭示消除归因偏差的效应。未来研究应尝试在多种情境下从更辩证的角度深入探讨内、外部归因偏差及平衡归因的积极与消极效应。”（详见第 17 页，第 3 段，4-13 行，和第 18 页）

除了以上修改，为了更准确地描述我们的结果和论述其含义，我们也对原文中涉及归因偏差的内容都做了相应修改与限定，明确了归因偏差的方向。

意见 5：研究的流程设计需要解释一下，为什么启动效应检验放在所有任务的后面，那么是否决策任务影响两个组的启动效应，且对不同组产生不同影响？（比如放大启动组的效应），因为决策任务呈现的刺激全部是老年人身体的不适情况。

回应：非常感谢审稿专家的宝贵意见。在本研究中，我们将启动效应的检验放在了所有任务的后面，是为了防止操作检验项目对实验变量产生无关的效应。关于操纵检验与因变量之间排序效应的讨论发现，在评估因变量之前进行操纵检验可能会改变被试对因变量的反应

（Kühnen, 2010）。如果在测量因变量之前进行检验，操纵检验很可能触发超出原本的启动刺激所带来的思维过程，从而掩盖启动刺激的真实效应，影响被试的反应。我们可以将其理解为操作检验本身所具有的启动效应。具体地说，刻板印象的操作检验中有关于年龄的表述也会成为启动刻板印象的一个线索。如果在刻板印象启动后马上呈现操作检验，就有可能出现实验操作和操作检验之间的交互影响，从而混淆实验操作的影响。因此，我们将启动效应检验放在了所有任务的后面，以避免操作检验对刻板印象启动操作的效应产生干扰。

另一方面，如专家所言，决策任务确实也可能会和实验启动检验产生交互作用。但就决策任务本身而言，两组的决策任务是一样的，由决策任务带来的影响可以视为相同的，两组唯一的不同在于刻板印象的激活与否。在这个背景下再考虑决策任务的影响时，一定程度上确实可能会放大启动组的效应。但即使产生上述影响也并不会对我们的研究结果造成混淆，因为从本质上来说决策任务中的体现的刻板印象与刻板印象启动所激活的刻板印象相一致，在此前提下，决策任务的这种影响也依然反映的是刻板印象威胁的效应，与我们的启动方向一致，仍然可以较好地反映了我们的实验操作的效应。

通过综合考虑，我们最终选择了在因变量任务之后进行操作检验的方式来尽可能减少变量间混淆与干扰。我们也在修改稿中对此进行了详细解释，具体内容如下：

“第四，进行刻板印象启动效果操作检验。由于刻板印象启动效果的操作检验可能会成为一种威胁性的社会线索，对因变量产生一定影响(Kühnen, 2010)，因此，在参考以往相关研究者做法的基础上(Brelet et al., 2016; 徐璐璐等, 2018)，我们在被试完成医疗决策任务完成之后进行刻板印象启动效果的操作检验。”（详见第 9 页，第 2 段，1-4 行）

意见 6：控制变量是否考虑，比如：参与者自身的归因偏好会对实验结果产生影响吗？

回应：感谢审稿专家的宝贵意见。根据以往文献，被试自身的归因偏好应该是会对实验结果产生一定的影响的。但是，就本研究而言，在被试量充足以及随机入组的前提下，实验中两组被试的归因偏好在理论上应该不存在显著的差异。即使对医疗决策任务有影响，两组被试自身归因偏好对实验结果的影响也应该是是一致的。两组之间在实验结果变量上的差异，更多

的应该是实验操作的不同造成的,而无关变量的影响对结果的影响应该在组间可以得到很好的平衡。

除了自身归因偏好,我们关注了实验中不同组被试的年龄、性别、自评健康、生活满意度、文化程度等控制变量的差异。在实验1中,独立样本 t 检验及卡方检验的结果表明,刻板印象威胁组被试和控制组被试在自评健康($t = -0.56, p = 0.58$)、年龄($t = -0.55, p = 0.58$)、生活满意度($t = 0.34, p = 0.74$)、性别($\chi^2 = 0.10, p = 0.75$)、受教育程度上($\chi^2 = 4.77, p = 0.19$)不存在显著差异(详见第6页,第1段,4-7行)。在实验2中,独立样本 t 检验及卡方检验结果表明,刻板印象威胁组被试和干预组被试在自评健康($t = 0, p = 1$)、年龄($t = -0.61, p = 0.54$)、生活满意($t = -0.31, p = 0.76$)、性别($\chi^2 = 1.92, p = 0.17$)、受教育程度($\chi^2 = 4.31, p = 0.37$)上不存在显著差异(见第12页,第2段,4-6行)。我们认为这些结果在一定程度上支持了这些控制变量对被试任务表现的影响在两组间无显著差异。

除了以上检验外,我们也尝试将这些变量作为控制变量重新对研究的主要结果进行了分析。结果表明,在控制了自评健康、年龄、生活满意度等无关变量后,实验1中刻板印象威胁组被试的归因偏差指标($M = 3.33, SD = 1.38$)显著高于控制组($M = 1.57, SD = 0.93$), $F(1,77) = 42.67, p < 0.001, \eta^2 = 0.38$,刻板印象威胁组内部归因倾向显著高于控制组;同时,刻板印象威胁组被试的决策质量($M = 0.15, SD = 1.44$)显著低于控制组($M = 2.61, SD = 1.31$), $F(1,77) = 57.13, p < 0.001, \eta^2 = 0.45$ 。中介效应检验仍然表明,刻板印象正向预测被试的归因偏差, $b = 0.60, SE = 0.09, t = 6.53, p < 0.001, 95\% CI = [0.419, 0.787]$,归因偏差负向预测被试的医疗决策质量, $b = -0.26, SE = 0.11, t = -2.35, p = 0.021, 95\% CI = [-0.474, -0.039]$ 。归因偏差在刻板印象对医疗决策质量的预测模型中起到中介作用,直接效应值为 $-0.51, SE = 0.11, 95\% CI = [-0.720, -0.292]$;间接效应值为 $-0.15, SE = 0.07, 95\% CI = [-0.314, -0.027]$ 。

相似地,在将被试的年龄、健康状况、生活满意、性别和文化程度作为协变量纳入实验2分析后, F 检验结果也显示,刻板印象威胁组被试的归因偏差指标($M = 3.96, SD = 4.03$)显著高于干预控制组($M = 0.16, SD = 0.26$),威胁组被试呈现出更多的内归因倾向,而干预组则呈现出更多的外归因倾向, $F(1,79) = 35.74, p < 0.001, \eta^2 = 0.33$;此外,归因偏差干预组被试的医疗决策质量($M = 3.00, SD = 1.75$)显著高于刻板印象威胁组($M = 0.65, SD = 1.85$), $F(1,79) = 34.46, p < 0.001, \eta^2 = 0.32$ 。

综合以上结果可知,自评健康、年龄、生活满意度等变量对我们的假设模型拟合度、路径参数估计等结果没有实质的影响,并不会影响我们所关注变量之间的效应以及对于结果的解释。

然而,我们也注意到并不是所有的研究者都认可通过纳入更多的控制变量来提高统计结果准确性的做法。有研究者认为,在统计分析中要谨慎地选取控制变量。研究者要确保每一个加入分析的控制变量都有理论上充足且合理的依据,不适当地加入控制变量可能会产生误导性的结果(Becker, 2005)。综上,无关变量对我们研究结果的影响有限,加入年龄、健康状况、生活满意、性别和文化程度作为控制变量除了削弱统计效力外,对于提高结果的准确性并没有明显的贡献。因此,在不同实验组没有明显差异的情况下,我们本着“对控制变量的使用持保守态度”(Becker, 2005),在最终分析中没有纳入自评健康、年龄、生活满意度等控制变量。

为了澄清控制变量对实验结果的可能影响,我们也在修改稿的被试部分进一步强调了两组被试在这些变量上不存在显著的差异(因为理论上讲,如果两组在这些控制变量上面没有显著的差异,则可视作两组为同质组,控制的无关变量对最终的实验结果就不应该存在显著的影响)。详见修改稿两个实验“被试”部分。

意见 7: 在医疗决策任务中, 因变量的衡量指标实际只用到 B 方案和 D 方案的评分, 设计 A 与 C 方案的目的是什么? 是否有更好的衡量指标, 从而能更加全面地反映参与者的决策模式?

回应: 感谢专家的宝贵意见。本研究医疗决策任务参照了 Dijksterhuis (2004) 在相关领域研究中所采用的决策任务范式。该任务范式是无意识决策领域的研究中被广泛采用的一种决策任务范式(Dijksterhuis, 2004, 2006; Rey et al., 2009; 陈虹, 2011)。在本研究中, 基于该范式的决策任务首先向被试提供四种针对某种疾病的治疗方案: 一种最佳治疗方案 B, 一种较差治疗方案 D, 以及两种一般治疗方案 A 和 C。然后, 要求被试对每种治疗方案进行评分。最后, 使用被试对最佳治疗方案 B 与较差治疗方案 D 的评分差值作为任务表现的指标。在该范式中, A 方案和 C 方案的作用是增加决策任务的复杂性与难度。换句话说, 如果去掉 A、C 两个一般治疗方案, 仅向被试呈现最佳治疗方案和较差治疗方案, 两者之间的对比就会过于明显。这样的决策任务过于简单, 很可能出现“天花板效应”, 无法获得对被试相关决策能力更真实准确的把握。我们在修改稿中进一步阐述了 A 方案和 C 方案在决策任务中的作用, 具体如下:

“方案 A 和方案 C 为一般的治疗方案, 各有 4 个正性描述, 作为填充内容加入, 以增加决策问题的复杂性与难度, 避免被试在任务表现中出现‘天花板效应’。” (详见第 8 页, 第 3 段, 3-5 行)

由于在很多情况下被试对方案 A 和方案 C 的评分都相同, 很难将两者的差值作为一种比较的基线, 反映到最终的决策质量指标中。因此, 在以往使用该任务范式的研究中, 普遍采用了 B 方案和 D 方案差值作为决策任务表现的指标, 这一指标也得到了相关领域研究者的普遍认可。Dijksterhuis (2004) 认为, 涉及到在多种选项进行选择的情境中, 能够认识到最佳方案显然是重要的, 但同时, 识别并拒绝一个较差的方案也同样重要。因此, 最佳方案与较差方案评分的差值能够很好地体现被试对最佳和较差治疗方案的区分能力, 可以既准确又简洁地反映被试的决策质量。

同时, 相较于对某一方案单一的态度评分而言, 两个相反态度评分之间的差异被认为具有更高统计检验效力(Dijksterhuis, 2004)。综合以上两点, 参照以往使用该范式的研究的一贯做法(Dijksterhuis, 2004, 2006; Rey et al., 2009; 陈虹, 2011), 我们也将两种方案评分的差值作为为了衡量被试对医疗方案决策质量指标。

意见 8: 刻板印象启动的材料问题: 控制组选择地理方面的材料, 刻板印象威胁组选择老化刻板印象的内容。两种材料的对比性不强, 前者完全无关年龄或“自我相关的负性体验”, 二者之间的启动效应是因为“年龄”本身激发, 还是因为诸如“负性自我”激发? 目前, 两种材料的对比启动效应差异似乎还不能确定为负性老化刻板印象所致。

回应: 感谢专家的仔细审阅。我们希望通过以下两方面说明解决您的疑惑。一方面, 我们研究中实验条件的启动方法是当前相关领域研究中广泛使用的成熟做法。在以往文献中, 通过阅读相关材料启动不同的实验条件是比较常用的一种做法。为了诱发实验组被试刻板印象威胁的状态, 研究者一般给被试呈现一段目标群体刻板印象相关的文字材料。同时, 相对应, 研究者会给控制组的被试呈现一段不诱发任何情绪状态的中性材料。然后将两种条件下被试的刻板印象激活水平作对比, 以检验刻板印象威胁操作的有效性。在我们的研究中, 控制组在启动阶段阅读地理材料主要是为了启动中性条件。地理知识材料不会激活控制组被试任何消极老年刻板印象或唤醒消极刻板印象相关情绪等无关变量, 从而避免其他无关变量对研究结果的干扰。在刻板印象研究领域, 这种向被试呈现中性短语或阅读材料的控制组处理方式已得到了广泛应用。大量研究都表明这种方法可以有效诱发中性情境, 启动对照条件(Stein, 2002; Levy, 1996; Bouazzaoui, 2015)。

另一方面,在当前实验背景下,老年刻板印象信息与由年龄信息本身激发的老年人负性自我体验并不存在严格的界限,也不涉及本质的冲突,可能没必要在实验处理中做进一步区分。具体而言,负性自我的激发本身是老年刻板印象内化过程中的产物以及刻板印象发挥消极影响的衍生物。不同于其他类型刻板印象,老年刻板印象随着年龄增长变得越来越与自我相关。也就是说,起初社会大众对老年群体的消极刻板印象最终将内化为老年人对自我的消极想法。以往证据也表明,于老年人而言,老年刻板印象可以在微妙的环境线索下自动被激活(Marques, Lima, Abrams, & Swift, 2014),并通过影响老年人对未来自我的看法内化为老年人现有的自我概念,并且这种联系随着年龄的增长而不断加强(Kornadt & Rothermund, 2012)。Rothermund 和 Brandtstätter (2005)也发现基线老年刻板印象内容预测了老年人自我概念的后续变化。换句话说,随着时间推移老年人的自我概念将逐渐同化为个体以往对老年人的普遍刻板印象。更直接的,Levy (2009)的刻板印象具体化理论认为,消极老年刻板印象通过内化为对自我的负面看法进而对老年人多种任务表现发挥消极影响。基于以上论述,在刻板印象威胁效应领域的研究中,负性自我的激发与老年刻板印象的激活本质上并不存在矛盾或混淆的问题。相反,这种负性的自我认知与体验是老年刻板印象发挥消极影响的重要中间机制与体现。因此,我们没有在本研究的启动任务中刻意区分老年刻板印象与消极负性自我的区别。

综上,我们已在修改稿对控制组材料的选取做了进一步的说明,具体内容如下:

“刻板印象威胁组被试的阅读材料强调个体在进入老年期后各项认知能力会衰退,喜欢在较少的选项中进行选择,决策质量较差,在一般决策任务和医疗情境决策任务上的表现都比年轻人差很多;控制组被试的阅读材料启动中性条件,主要介绍了青海盐湖的分布、含盐量及其所拥有的众多无机盐的类型。实验组材料及控制组材料的选取和编制主要参考了以往研究普遍采用的成熟范式。大量相关领域研究表明,使用阅读材料可以有效地诱发实验组被试刻板印象威胁状态,同时避免控制组被试受到消极刻板印象及其相关情绪等无关变量的干扰(Bouazzaoui et al., 2016; Lamont et al., 2015; Levy, 1996; Stein et al., 2002)。”(详见第6页,第3段,3-10行)

细节问题

问题提出部分

意见1:前言部分写到:“充分的研究证据也支持了低质量的医疗决策会对老年病人的生理和心理过程有着显著的消极影响”,但其后并未见到有关老年病人心理方面影响的文献。

回应:感谢审稿专家的建议。我们在修改稿中对这部分内容进行了优化,主要论述了低质量的医疗决策对老年人疾病康复过程的影响,以及低质量的医疗决策会加重老年人的经济负担。修改后的具体内容如下:

“此外,Tsoulos 等人(2019)的研究表明低质量的医疗决策会增加相关并发症出现的风险,延缓老年人的康复进程。与此相似,Ekstrom (2012)也发现低水平的医疗决策会加剧老年人在面对疾病时的身体负荷,更容易造成难以挽回的永久性伤痛。更有甚者,较差的医疗决策还会导致更多的额外治疗费用,增加老年病人的经济负担和死亡风险(Weiss et al, 2012)。”

(详见第2页,第1段,第9-14行)

意见2:作者在介绍完消极刻板印象以及医疗决策概念后,表明“尚未有研究关注消极刻板印象威胁对老年人医疗决策后果尤其是决策质量的效应及其机制。基于此,本研究将相关领域刻板印象威胁、归因偏差和老年人的决策质量整合道一个研究框架中……”在此前,并没有介绍过“归因偏差”的概念,应该在第一次引入时对概念进行简要介绍。对内外部归因概念的解释,应在第一次出现时(1.2 刻板印象与归因偏差)就给出”

回应：感谢专家的仔细审阅。我们已在修改稿的前言中介绍了归因偏差的概念，并补充了有关归因偏差在本研究中角色的简要讨论。修改后的具体内容如下：

“此外，消极刻板印象对老年人医疗决策效应的机制也需要相关领域的研究进一步关注。已有研究表明，归因偏差与刻板印象和决策行为都有着密切的联系(Koo & Yang, 2018; 王大伟, 刘永芳, 2008)。然而，尚未有研究检验归因偏差在刻板印象对行为决策效应中所扮演的角色。在以往文献中，归因偏差被界定为个体非理性地将事件结果归因为某个特定原因而忽视其真正原因的归因倾向(Working, 2015)。研究表明，人们对某一群体固定的观念或预期能够影响个人对他人及自身行为的归因，使个体在进行因果归因时存在一定程度的偏差(Fiske & Susan, 1992; Oakes et al., 1991)。另一方面，归因偏差又被证实是决策质量的一个显著的影响因素(王大伟, 刘永芳, 2008)。因此，归因偏差很可能是消极刻板印象对老年医疗决策质量效应的一个中介变量。”（详见第 3 页，第 2 段）

另外，我们也将对内外部归因概念的解释调整至了 1.2 部分，修改后的具体内容如下：

“对于自己和他人行为的结果，人们一般会从两个方面进行归因，包括内部归因（如能力、态度和努力）和外部归因（如任务特征、环境和运气）(Heider, 1958)。然而，由于通常会受到刻板印象等固有观念或认知的影响，人们的归因过程难以理性客观，常常存在偏差(Fiske & Susan, 1992; Oakes et al., 1991)。根据刻板印象解释偏差理论，当个体的行为结果与刻板印象预期相符时，人们倾向于将行为结果进行内部归因；当行为的结果与刻板印象预期不相符时，人们则更愿意将目标行为结果进行外部归因(Kulik, 1983)。”（详见第 4 页，第 2 段，1-6 行）

意见 3：前言部分同时也写到：“刻板印象会通过刻板印象威胁对个体的决策过程和决策结果产生显著影响”，刻板印象威胁是一种机制还是现象？请在论述过程中明晰其概念。

回应：感谢专家的意见。“刻板印象威胁”的概念最早由 Steel 等人（1995）在解释非裔美国学生和白人学生在学业成绩上的差异表现时提出的。在他们的研究中，当学业测验被正式描述为“智力诊断测验”时，就启动了美国文化中“黑人比白人智力更低”的负面刻板印象，这种社会层面的心理威胁干扰了黑人学生的能力表现。由此可见，从这种角度出发，刻板印象威胁是一种情境性困境，这种困境源于个体对自我及所属群体的负面刻板印象的意识和知觉，当个体担心这种消极刻板印象作用于自我时，个体的能力就会受到威胁（Schmader et al., 2008）。从这个角度出发，我们可以将其视为特殊情境孕育的一种现象。而如果从刻板印象的实质出发，我们又可以将其理解为刻板印象效应的一个作用机制——它导致个体自我概念、能力概念和所启动的消极刻板印象之间产生冲突，进而对其自我完善造成威胁。我们也在修改稿中前言部分进一步明晰了刻板印象威胁概念，具体如下：

“在社会情境中，当文化中的消极刻板印象对目标个体的心理行为产生影响时，刻板印象威胁就发生了。Steele 和 Aronson (1995)将刻板印象威胁定义为个体在对自我和所属群体消极刻板认识的基础上，因担心验证对自我或所属群体的刻板印象而经历的一种风险。刻板印象威胁是一种情境性困境同时也是消极刻板印象效应的一个重要机制。”（详见第 3 页，第 4 段，1-4 行）

意见 4：Masicampo 和 Baumeister (2008) 的研究表明老年人会因为受到刻板印象而对医疗决策更加消极生存意愿更低，而 Levy 等(2000) 的研究则表明老年人受刻板印象影响时在风险决策中更容易冒进，这两个研究的关联是什么，是否相互矛盾？

回应：非常感谢专家的仔细审阅与宝贵意见。在本研究中，我们试图通过引用这两个研究来论证消极刻板印象威胁情境中老年人更容易做出不合理的决策，决策质量更差。Masicampo 和 Baumeister (2008)以及 Schmader 等人(2008)的研究证明启动消极刻板印象会消耗老年人的

执行控制功能资源，这将导致他们做出更偏激的选择和比较冒险的决策。Levy 等(2000)的研究则阐述了在医疗情境中，受消极刻板印象影响的老年病人对治疗的态度较为消极，生存意愿更低，更容易拒绝治疗。这表明在刻板印象威胁影响下的老年人更倾向于做出不利于提高生存率的决策，这一现象也支持了刻板印象对老年人决策质量具有一定的消极效应。综上，这两个研究实质上印证了刻板印象威胁会降低老年人做出合理决策的能力。为了减少歧义，我们在修改稿中对这两条文献做了进一步的澄清，强调了两个文献都支持了消极刻板印象会削弱老年人做出合理决策的能力。修改后的内容如下：

“其他领域决策的研究表明，启动相关的消极刻板印象会对老年人造成某种社会心理威胁，削弱其合理决策的能力，最终导致老年人在刻板印象威胁情境中做出更偏激的选择和比较冒险的决策(Masicampo & Baumeister, 2008)。在医疗情境中，Levy 等(2000)的研究也发现，受消极刻板印象影响的老年病人对治疗的态度较为消极，生存意愿更低，更容易拒绝治疗。这表明在刻板印象威胁情境下，由于合理决策的能力降低，老年人更可能做出不利于提高其生存率的决策，即医疗决策质量较差。”（第 4 页，第 1 段）

意见 5：实验二前言“以往大量研究表明”部分，请引用大量相关研究或综述文献作为支撑。

回应：感谢审稿专家的宝贵意见。我们已在修改稿中这部分内容补充了更多相关参考文献。修改后的内容如下：

“以往大量研究表明，引导被试对结果进行恰当的归因并给予及时反馈是一种有效的归因训练方法(Stewart et al., 2010; Vassilopoulos et al., 2009; 廖全明, 郑涌, 2007; 黄小草, 2018)。”（详见第 11 页，最后一段）

方法与结果

意见 1：在心理学报，小于 1 的小数前的“0”无需省略；

回应：感谢审稿专家的意见。我们已经重新检查数据的书写格式，在修改稿中做出相应的修正（详见两个实验结果部分）。

意见 2：实验一报告了评分者一致性，起初报告的是“两名不了解实验目的的心理学专业人员对表述及匹配的原因进行评分”，后文又出现“三个评分者评分的内部一致性信度”，请检查；

回应：感谢审稿专家的意见。关于 WSAP 任务材料的评定，在搜集了老年人常见的健康不佳症状以及每个症状对应的可能的原因之后，主试首先从中挑选了一些合适的健康不佳的症状作为实验材料，然后从每个症状对应的原因中区分出内部原因和外部原因。之后，为了验证主试对原因区分的可靠性，我们又邀请了两名不了解实验目的的心理学专业人员独立地对所选症状对应的原因重新进行分类，区别内部原因和外部原因。因此，包括主试在内一共有三人参与了材料的评定，我们在文稿中汇报了三名评价者的内部一致性信度。我们已经在修改稿中对这个问题进行了澄清，内容如下：

“各种健康不佳的症状以及症状相对应的原因是通过发放开放式问卷和阅读相关文献来搜集的。根据搜集资料的内容，实验主试首先挑选出了一些最常见的健康不佳症状作为实验材料，然后从每个症状对应的原因中区分出内部原因和外部原因。为了验证主试对原因区分的可靠性，我们还邀请两名不了解实验目的的心理学专业人员独立对每个症状对应的两个原因进行内外因匹配。两人做出相同匹配的比例为 96.4%。此外，两人的匹配与主试的匹配结果显著正相关 ($r = 0.91, p < 0.001$ 和 $r = 0.92, p < 0.001$)，两人的匹配结果也显著正相关 ($r = 0.92, p < 0.001$)，三个评分者（主试与两个评分者）的内部一致性信度为 0.97。这表明 WSAP 任务的材料具有一定的可靠性。”（详见第 7 页，第 3 段）

意见 3: 实验二流程图应当保证前后步骤呈现的一致性，呈现两种选择所对应的两种反馈。
回应: 感谢审稿专家的意见。我们已经根据您的建议对实验二的流程图进行了修改，同时，我们也在流程图下面加入了一个图注，进一步说明了图中各内容的含义。修改后的流程图如下（详见第 13 页图 4）：

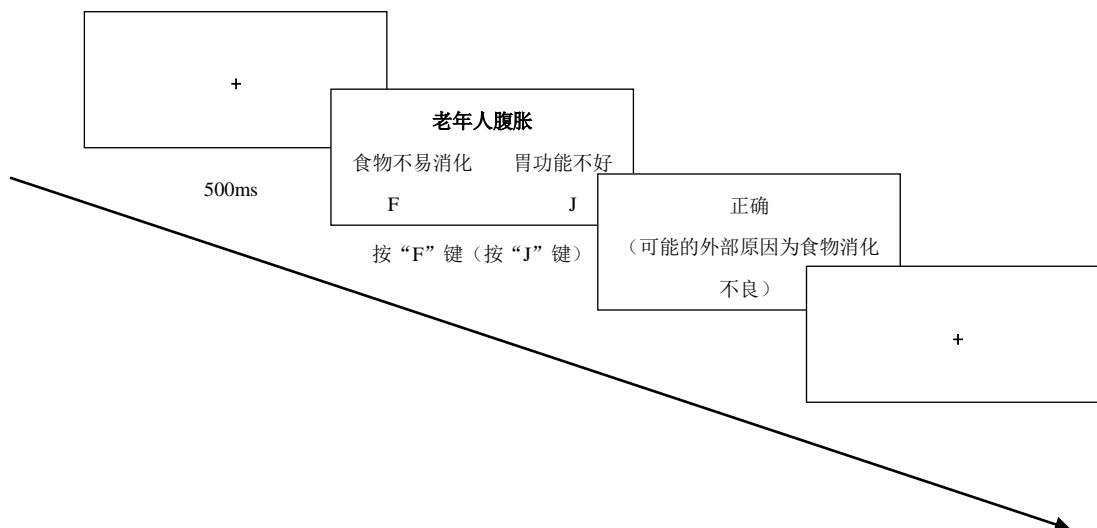


图 4 归因偏差干预流程

注：被试按键反应后进入反馈页面，当被试按“F”键时，呈现“正确”字样；当被试按“J”键时，呈现可能的外部原因，如图中括号内所示。

意见 4: 文章在被试特征信息（表 4）和实验结果（表 5）皆报告了被试的基本信息（年龄、生活满意度、自评健康、文化程度），建议合并为同一张表。并优化各个表图的标题，以确切表达其内容。

回应: 感谢审稿专家的意见。根据您的建议，我们在修改稿中对内容重合的表格进行了合并，并对各个图表的标题进行了优化，以便更明确清晰地表达信息。首先，我们将原文中的“表 1 背景信息”和“表 2 描述性统计及独立样本 t 检验”两表进行了合并，修改后表格如下所示：

表 1 实验一被试信息及相關变量差异检验结果

变量	威胁组 (<i>n</i> = 40)	控制组 (<i>n</i> = 38)	<i>t</i> / χ^2	<i>Cohen's d</i>
年龄 <i>M</i> (<i>SD</i>)	68.45 (7.08)	67.53 (7.69)	-0.55	-0.12
自评健康 <i>M</i> (<i>SD</i>)	3.50 (0.68)	3.42 (0.55)	-0.56	-0.13
生活满意 <i>M</i> (<i>SD</i>)	3.88 (0.69)	3.92 (0.49)	0.34	0.07
性别 <i>N</i> (%)				
男	14 (35.0%)	12 (31.6%)	0.10	
女	26 (65.0%)	26 (68.4%)		
文化程度 <i>N</i> (%)			4.77	
小学	5 (12.5%)	2 (5.3%)		
初中	15 (37.5%)	9 (23.7%)		
高中	6 (15.0%)	12 (31.6%)		
本科及以上	14 (35.0%)	15 (39.5%)		
归因偏差 <i>M</i> (<i>SD</i>)	3.33 (1.38)	1.57 (0.93)	-6.63***	-1.50
决策质量 <i>M</i> (<i>SD</i>)	0.15 (1.44)	2.61 (1.31)	7.87***	1.79

注：自评健康为 5 点计分，1 代表非常不健康，5 代表非常健康，得分越高，代表自评健康越好。生活满意为 5 点计分，1 代表非常不满意，5 代表非常满意，得分越高，代表生活满意度越高。括号内是标准差(年龄、自评健康、生活满意度、归因偏差及决策质量)或百分比(性别和文化程度)。*表示 $p < 0.05$ ，**表示 $p < 0.01$ ，***表示 $p < 0.001$ ，下同。

其次，我们将原文中的“表 4 背景信息”和“表 5 描述性统计及独立样本 *t* 检验”进行了合并，修改后的表格如下：

表 2 实验二被试信息及差异检验结果

变量	干预组 (<i>n</i> = 40)	威胁组 (<i>n</i> = 40)	<i>t</i> / χ^2	<i>Cohen's d</i>
年龄 <i>M</i> (<i>SD</i>)	65.30 (5.42)	66.13 (6.58)	-0.61	-0.14
自评健康 <i>M</i> (<i>SD</i>)	3.50 (0.68)	3.50 (0.68)	0.00	0.00
生活满意 <i>M</i> (<i>SD</i>)	3.80 (0.79)	3.85 (0.66)	-0.31	0.07
性别 <i>N</i> (%)			1.92	
男	18 (45.0%)	12 (30.0%)		
女	22 (55.0%)	28 (70.0%)		
文化程度 <i>N</i> (%)			4.31	
小学	1 (2.5%)	3 (7.5%)		
初中	7 (17.5%)	9 (22.5%)		
高中	8 (20.0%)	12 (30.0%)		
本科及以上	24 (60.0%)	40 (40.0%)		
归因偏差 <i>M</i> (<i>SD</i>)	0.16 (0.26)	3.96 (4.03)	-5.96***	-1.33
决策质量 <i>M</i> (<i>SD</i>)	3.00 (1.75)	0.65 (1.85)	5.83***	1.31

注：自评健康为 5 点计分，1 代表非常不健康，5 代表非常健康，得分越高，代表自评健康越好。生活满意为 5 点计分，1 代表非常不满意，5 代表非常满意，得分越高，代表生活满意度越高。括号内是标准差（年龄、自评健康、生活满意度、归因偏差及决策质量）或百分比（性别和文化程度）。

另外，我们也重新检查文章中的其它表格，考虑到原稿中的表 2 仅呈现了实验 1 中的三个主要变量之间的相关分析结果，内容太简单，我们在修改稿中移除了原来的表 2，并将对这部分相关结果的报告并入正文部分，具体内容如下：

“结果表明三个变量之间分别两两相关，刻板印象组别与归因偏差显著正相关， $r = 0.60$, $p < 0.001$ ；刻板印象组别和决策质量显著负相关， $r = -0.67$, $p < 0.001$ ；归因偏差和决策质量显著负相关， $r = -0.56$, $p < 0.001$ 。”（详见第 9 页，第 7 段和第 10 页，第 1 段）

意见 5：对年龄均值进行阐述时，小数位数需要保持一致，例如文中“被试的平均年龄为 66 ± 6.00 岁”

回应：感谢您的意见。我们已将此处修正为 65.71 ± 6.00（详见第 12 页，第 2 段，第 3 行）。同时，我们也对全文数据的表达格式进行了仔细检查，并对格式不一致或有误的内容进行了修改，最大程度避免了类似的问题。

意见 6：在报告中介检验的结果时，需要补充重采样的次数（Bootstrap=?）。

回应：感谢审稿专家的意见。我们已经根据您的建议在中介检验结果中补充了采样次数，具体内容如下：

“根据 Preacher 和 Hayes（2008）提出的检验小样本中介效应的分析方法，本研究采用

bootstrap 法（重复抽样 5000 次）来检验归因偏差在老年刻板印象和医疗决策质量之间的中介作用（见图 3）。”（详见第 10 页，第 2 段，1-3 行）

意见 7：统计量报告的格式不规范，如“刻板印象对医疗决策质量的直接效应为 -0.52 ， $SE = 0.10$ ， $t = -5.02$ ， $p < .001$ ， $95\%CI = [-0.725, -0.313]$ ”请核查期刊的用稿标准。”

回应：感谢审稿专家的宝贵意见。我们已参照期刊用稿标准对中介部分统计量报告的格式进行了修改，具体内容如下：

“根据 Preacher 和 Hayes（2008）提出的小样本中介的分析方法，本研究采用 bootstrap 法（重复抽样 5000 次）来检验归因偏差在老年刻板印象和医疗决策质量之间的中介作用（见图 3）。结果表明，刻板印象正向预测被试的归因偏差， $b = 0.60$ ， $SE = 0.09$ ， $t = 6.57$ ， $p < 0.001$ ， $95\% CI = [0.419, 0.784]$ ，归因偏差负向预测被试的医疗决策质量， $b = -0.25$ ， $SE = 0.10$ ， $t = -2.43$ ， $p = 0.018$ ， $95\% CI = [-0.457, -0.045]$ 。归因偏差在刻板印象对医疗决策质量的预测模型中起到中介作用，直接效应值为 -0.52 ， $SE = 0.10$ ， $95\% CI = [-0.725, -0.313]$ ；间接效应值为 -0.15 ， $Boot SE = 0.07$ ， $95\% CI = [-0.302, -0.033]$ 。”（详见第 10 页，第 2 段）

此外，我们也检查了文稿其他部分的数据，并对不符合规范的数据一一进行了修正。

其他

意见 1：实验一讨论部分“消极的刻板印象可能会导致老年人的决策表现偏向于与刻板印象相一致的结果”更像是描述，而不能对后文起到解释作用，请调整。

回应：感谢审稿人的宝贵意见。我们已经根据您的建议对相应内容进行了调整，修改后的内容如下：

“这一结果与 Lockenhoff 和 Corinna（2017）的研究结论相一致，即在刻板印象威胁情境中，个体的决策表现可能会倾向于与刻板印象保持一致。”（详见第 11 页，第 2 段，第 2-3 行）

意见 2：参考文献：请登录期刊官网下载参考文献著录要求，仔细核对参考文献格式。

回应：感谢审稿专家的仔细审阅。在本轮修改中，我们对参考文献进行了逐条检查，并对所有发现的问题进行了纠正，包括拼写错误、缩写、缺失信息等。修改后，所有参考文献的格式都已符合杂志的要求。

意见 3：英文摘要：部分句子不符合英文语法，如“*There are 78 participants were*”，请检查。

回应：感谢审稿专家的意见。目前该句修改如下：

“78 older adults (26 males and 52 females) were recruited as participants in Experiment 1.”
（详见英文摘要部分）

此外，我们邀请三名专业人士（包括一名英语专业的老师）对英文摘要进行了反复校对与修改，确保不存在类似的语法问题和其他类型的错误。

参考文献

- Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational research methods*, 8(3), 274–289.
- Bouazzaoui, B., Follenfant, A., Ric, F., Fay, S., Croizet, J. C., Atzeni, T., & Tacconat, L. (2016). Ageing-related stereotypes in memory: When the beliefs come true. *Memory*, 24(5), 659–668.
- Brelet, L., Moták, L., Ginet, M., Huet, N., Izaute, M., & Gabaude, C. (2016). Enhancing older drivers' safety: On effects induced by stereotype threat to older adults' driving performance, working memory and

- self-regulation. *Geriatrics*, 1(3), 20.
- Bryant, C., Bei, B., Gilson, K., Komiti, A., Jackson, H., & Judd, F. (2012). The relationship between attitudes to aging and physical and mental health in older adults. *International psychogeriatrics*, 24(10), 1674–1683.
- Carr, P. B., & Steele, C. M. (2010). Stereotype threat affects financial decision making. *Psychological Science*, 21(10), 1411–1416.
- Chen, H. (2011). *The affection of emotion and motivation on unconscious thought decision-making*. (Unpublished master's thesis). Ningbo University, Zhejiang.
- [陈虹. (2011). 情绪、动机对无意识思维决策的作用 (硕士学位论文). 宁波大学, 浙江.]
- Coudin, G., & Alexopoulos, T. (2010). 'Help me! I'm old!' How negative aging stereotypes create dependency among older adults. *Aging & mental health*, 14(5), 516–523.
- Dijksterhuis. (2004). Think different: The merits of unconscious thought in preference development and decision making. *Journal of Personality and Social Psychology*, 87(5), 586–598.
- Dionigi, R. A. (2015). Stereotypes of aging: Their effects on the health of older adults. *Journal of Geriatrics*, 2015, 1–9.
- Ekstrom, L. W. (2012). Liars, medicine, and compassion. *Journal of Medicine and Philosophy*. 37(2), 159–180.
- Fiske, & Susan, T. (1992). Thinking is for doing: Portraits of social cognition from daguerreotype to laserphoto. *Journal of Personality and Social Psychology*. 63(6), 877–889.
- Gravetter, F. J., & Wallnau, L. B. (2009). *Statistics for behavioral sciences* (8th ed). Belmont, CA: Wadsworth.
- Guo, X. Y. (2004). *Experimental Psychology*. Beijing, China: People's Educational Press.
- [郭秀艳. (2004). 实验心理学. 北京: 人民教育出版社.]
- Harvey, P., Martinko, M. J., & Gardner, W. L. (2006). Promoting authentic behavior in organizations: An attributional perspective. *Journal of Leadership & Organizational Studies*, 12(3), 1–11.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley.
- Howell, D. C. (2012). *Statistical methods for psychology* (8th ed). Belmont, CA: Wadsworth.
- Huang, X. C. (2018). *The influence and intervention of hostility attribution bias on aggression behavior*. (Unpublished master's thesis). Beijing Sport University.
- [黄小草. (2018). 敌意归因偏差对攻击行为的影响及干预 (硕士学位论文). 北京体育大学.]
- Inzlicht, M., & Kang, S. K. (2010). Stereotype threat spillover: How coping with threats to social identity affects aggression, eating, decision making, and attention. *Journal of Personality and Social Psychology*, 99(3), 467–481.
- Jordano, M. L., & Touron, D. R. (2017). Stereotype threat as a trigger of mind-wandering in older adults. *Psychology and Aging*, 32(3), 307–313.
- Koo, J. H., & Yang, D. (2018). Managerial overconfidence, self-attribution bias, and downwardly sticky investment: evidence from Korea. *Emerging Markets Finance & Trade*, 54(1), 144–161.
- Kornadt, A. E., & Rothermund, K. (2012). Internalization of age stereotypes into the self-concept via future self-views: A general model and domain-specific differences. *Psychology and Aging*, 27(1), 164–172.
- Kühnen, U. (2010). Manipulation checks as manipulation: Another look at the ease-of-retrieval heuristic. *Personality and Social Psychology Bulletin*, 36(1), 47–58.
- Kulik, J. A. (1983). Confirmatory attribution and the perpetuation of social beliefs. *Journal of Personality and Social Psychology*, 44(6), 1171–1181.
- Lamont, R. A., Swift, H. J., & Abrams, D. (2015). A review and meta-analysis of age-based stereotype threat: Negative stereotypes, not facts, do the damage. *Psychology and Aging*, 30(1), 180–193.
- Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology*, 71(6), 1092–1107.
- Levy, B. R. (2009). Stereotype embodiment: A psychosocial approach to aging. *Current Directions in Psychological Science*, 18(6), 332–336.
- Levy, B., Ashman, O., & Dror, I. (2000). To be or not to be: The effects of aging stereotypes on the will to live. *OMEGA—Journal of Death and Dying*, 40(3), 409–420.
- Li, A., M., Ling, W., Q., & Li, L., Y., (2009). The implicit facial attractiveness stereotype research in personnel interview circumstances. *Psychological Science*, 32(4), 970–973.
- [李爱梅, 凌文轻, 李连雨. (2009). 招聘面试中的内隐相貌刻板印象研究. 心理科学, 32(4), 970–973.]
- Liao, Q., M. & Zheng, Y. (2007). An experimental research on the influence of different training methods on pupils' sharing behavior. *Psychological Science*, 30(6), 1351–1355
- [廖全明, 郑涌. (2007). 不同训练方法对小学生分享行为影响的实验研究. 心理科学, 30(6), 1351–1355.]
- Lockenhoff, & Corinna, E. (2017). Aging and decision-making: A conceptual framework for future research a mini review. *Gerontology*, 64(2), 140–148.
- Lu, X., F., Tang, Y., H. & Zeng, F., M., (2011). Effect size: estimation, reporting and interpretation. *Psychological Exploration*, 31(3), 260–264.

- [卢谢峰, 唐源鸿, 曾凡梅. (2011). 效应量: 估计, 报告和解释. *心理学探新*, 31(3), 260–264.]
- Marques, S., Lima, M. L., Abrams, D., & Swift, H. (2014). Will to live in older people's medical decisions: immediate and delayed effects of aging stereotypes. *Journal of Applied Social Psychology*, 44(6), 399–408.
- Masicampo, E. J., & Baumeister, R. F. (2008). Toward a physiology of dual-process reasoning and judgment: Lemonade, willpower, and expensive rule-based analysis. *Psychological Science*, 19(3), 255–260.
- Mather, M., Mazar, N., Gorlick, M. A., Lighthall, N. R., Burgeno, J., Schoeke, A., & Ariely, D. (2012). Risk preferences and aging: The “certainty effect” in older adults' decision making. *Psychology and aging*, 27(4), 801–816.
- Oakes, P. J., Turner, J. C., & Haslam, S. A. (1991). Perceiving people as group members: The role of fit in the salience of social categorizations. *British Journal of Social Psychology*, 30(2), 125–144.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, 40(3), 879–891.
- Rey, A., Goldstein, R. M., & Perruchet, P. (2009). Does unconscious thought improve complex decision making? *Psychological Research PRPF*, 73(3), 372–379.
- Rothermund, K., & Brandtstätter, J. (2005). Age stereotypes and self-views in later life: Evaluating rival assumptions. *International Journal of Behavioral Development*, 27(6), 549–554.
- Sanjuán, P., & Magallares, A. (2014). Coping strategies as mediating variables between self-serving attributional bias and subjective well-being. *Journal of Happiness Studies*, 15(2), 443–453.
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological review*, 115(2), 336–356.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist*, 52(6), 613–629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811.
- Stein, R., Blanchard-Fields, F., & Hertzog, C. (2002). The Effects of Age-Stereotype Priming on the Memory Performance of Older Adults. *Experimental Aging Research*, 28(2), 169–181.
- Stewart, T. L., Chipperfield, J. G., Perry, R. P., & Weiner, B. (2012). Attributing illness to 'old age': Consequences of a self-directed stereotype for health and mortality. *Psychology & Health*, 27(8), 881–897.
- Stewart, T. L., Latu, I. M., Kawakami, K., & Myers, A. C. (2010). Consider the situation: Reducing automatic stereotyping through situational attribution training. *Journal of Experimental Social Psychology*, 46(1), 221–225.
- Trent, E. S., Viana, A. G., Raines, E. M., Woodward, E. C., & Zvolensky, M. J. (2019). Interpretation biases and childhood anxiety: The moderating role of parasympathetic nervous system reactivity. *Journal of Abnormal Child Psychology*, 48(3), 419–433.
- Tsoulos, I. G., Mitsi, G., Stavrakoudis, A., & Papapetropoulos, S. (2019). Application of machine learning in a parkinson's disease digital biomarker dataset using neural network construction (NNC) methodology discriminates patient motor status. *Frontiers in ICT*, 6, 1–7.
- Vailati Riboni, F., & Pagnini, F. (2021). Age-based stereotype threat: a scoping review of stereotype priming techniques and their effects on the aging process. *Aging & Mental Health*, Advance online publication.
- Vassilopoulos, S. P., Banerjee, R., & Prantzalou, C. (2009). Experimental modification of interpretation bias in socially anxious children: Changes in interpretation, anticipated interpersonal anxiety, and social anxiety symptoms. *Behaviour Research and Therapy*, 47(12), 1085–1089.
- Vassilopoulos, S. P., Banerjee, R., & Prantzalou, C. (2009). Experimental modification of interpretation bias in socially anxious children: Changes in interpretation, anticipated interpersonal anxiety, and social anxiety symptoms. *Behaviour Research and Therapy*, 47(12), 1085–1089.
- Wang, D. W., & Liu, Y. F. (2008). An experimental research on the effects of attributive style and time pressure on purchasing decision. *Psychological Science*, 31(4), 905–908.
- [王大伟, 刘永芳. (2008). 归因风格、时间压力对购买决策影响的实验研究. *心理科学*, 31(4), 905–908.]
- Weiss, B. D., Berman, E. A., Howe, C. L., & Fleming, R. B. (2012). Medical decision - making for older adults without family. *Journal of the American Geriatrics Society*, 60(11), 2144–2150.
- Casey, P. R., Psych, F. R. C., & Strain, J. J. (Eds.). (2015). *Trauma-and stressor-related disorders: A handbook for clinicians*. American Psychiatric Pub.
- Xu, L., L., Wu, P., P. & He, W. (2018). Meta-stereotype threat effects on intergroup relations among impoverished undergraduates: a moderated mediation model of intergroup anxiety and self-esteem. *Psychological Development and Education*, 34(4), 426–433.
- [徐璐璐, 吴佩佩, 贺雯. (2018). 贫困大学生元刻板印象威胁对群际关系的影响: 群际焦虑的中介和自尊的调节作用. *心理发展与教育*, 34(4), 426–433.]

- Yang, T. & Ren, X., P., (2016). The impact of gender orientation of names on female mate preferences. *Psychological Science*, 39(5), 1190–1196.
[杨婷, 任孝鹏. (2016). 不同性别定向的名字对女性择偶偏好的影响. *心理科学*, 39(5), 1190–1196.]
- Zhang, H. C. & Xu. J. P. (Eds). (2009). *Modern Psychology and educational statistics*. Beijing, China: Beijing Normal University Press.
[张厚粲, 徐建平. (编). (2009). *现代心理与教育统计学*. 北京: 北京师范大学出版社.]
-

第二轮

审稿人 3 意见:

作者针对一审意见做了仔细修改, 逻辑清晰了很多。但是, 本轮评审发现, 还有一个研究设计的逻辑问题, 需要作者进一步澄清。研究 2 的设计是启动消极刻板印象后, 将两组被试进行对比, 一组接受归因偏差干预, 另一组不干预。然后对比二者决策质量的差异。作者认为这个设计是用于检验归因偏差在消极刻板印象威胁和决策质量之间的中介作用。审稿人对此有不同意见, 认为这实际是调节效应而非中介效应的检验。据审稿人理解, 因果链式实验设计至少做两个实验: A 影响 B 和 C; B 影响 C。据此, 实验二应该以归因偏差为自变量, 决策质量为因变量, 无关乎消极刻板印象威胁这个变量。但是目前的设计本质仍然是以消极刻板印象为自变量看两种条件下, 对决策质量影响的差异。因此, 审稿人认为这是将归因偏差作为调节变量的研究设计。请作者进一步澄清因果链式实验设计的逻辑, 以免读者误解。

回应: 感谢审稿专家对我们修改工作的认可, 也非常感谢审稿专家在评审本文过程中的辛苦付出! 我们对您的意见进行了深入的思考。我们拟从以下两个方面对您的这一问题进行澄清:

首先, 从各自的目的看, 两个实验是符合因果链实验设计的逻辑的。实验 1 的目的是检验消极刻板印象对归因偏差的效应, 以及消极刻板印象对老年人医疗决策的效应。也就是说, 实验 1 是通过操纵消极刻板印象 (自变量), 检验自变量 (A) 对老年人的归因偏差 (中介变量 B) 和医疗决策质量 (因变量 C) 的影响。实验 2 的目的是在刻板印象威胁条件下, 检验归因偏差的控制干预是否能够改善老年人医疗决策的质量。实验 2 实质上就是在特定情境中 (刻板印象威胁情境), 检验归因偏差对老年人医疗决策的效应, 即变量 B 对变量 C 的影响。因此, 从研究目的看, 两个实验的目的符合因果链实验设计的逻辑。

其次, 从研究设计看, 实验 2 的自变量是归因偏差, 因变量是老年人的医疗决策质量, 消极刻板印象威胁只是一个无关的控制变量。在实验 2 中, 消极刻板印象操纵的本质就是人为地创设一种两组被试需要保持一致的实验情境。消极刻板印象是一个被控制的没有变化的量, 可以看作是实验 2 的一个无关变量。而对归因偏差的实验控制才是对实验 2 自变量的操纵。也就是说, 实验 2 是在控制了消极刻板印象的条件下, 操纵归因偏差 (实验 2 的自变量), 进一步检验归因偏差 (B) 对老年人的医疗决策质量 (因变量 C) 的效应。因此, 在实验 2 中, 自变量是归因偏差, 因变量是老年人医疗决策的质量。

综上, 实验 1 检验了消极刻板印象对归因偏差和老年人医疗决策质量的影响, 实验 2 在实验 1 的基础上, 进一步检验了归因偏差对老年人医疗决策质量的效应。尽管与以往研究者使用的因果链设计略有不同 (e.g., Spencer et al., 2005), 但本研究的设计基本符合因果链实验设计的逻辑。为了消除读者的误解, 我们在原文做了以下修改:

首先, 关于两个实验设计的逻辑, 我们在“引言”部分的最后一段, 作了进一步的澄清, 详细地描述了两个研究设计的逻辑。同时, 为了增加表述的准确性, 我们也在相关描述的措辞上作了修改 (如将原文中“本研究根据因果链实验设计中介效应检验的方法”修改成了“本研究借鉴因果链实验设计中介效应检验方法”, 降低了表述语气的绝对性)。修改后的内容如下:

“基于以上论述，本研究借鉴因果链实验设计中中介效应检验方法，拟采用两个实验来探讨消极刻板印象、归因偏差和老年人医疗决策质量之间的关系。实验 1 通过阅读文字材料操纵刻板印象情境和控制情境，比较两种情境中被试归因偏差和医疗决策任务表现的差异，考察消极刻板印象对归因偏差和老年人医疗决策质量的影响，并进一步检验归因偏差的中介作用。在实验 1 的基础上，实验 2 首先启动了所有被试的消极刻板印象，然后将被试随机分配至刻板印象威胁组（启动刻板印象后对被试不做任何实验处理）和干预控制组（启动刻板印象后对被试进行外归因倾向训练）。通过比较两组被试的医疗决策质量，验证归因偏差对老年人医疗决策质量的影响。在实验 2 中，自变量为归因偏差的干预水平，因变量为老年人的医疗决策质量，消极刻板印象为控制变量。”（见修改稿第 5 页，最后一段，蓝色字部分）

其次，在第 11 页，实验 1 讨论部分，为了消除可能的误解，删掉了原来“根据因果链实验设计中中介效应检验的方法”这一表述，替换成“根据实验 1 的结果”这一表述（见修改稿 11 页，第 3 段，第 1 行）。随后，同是在该部分的最后一段，又对实验 2 的变量关系做了澄清，明确了归因偏差是研究的自变量，医疗决策质量是因变量。修改后的内容如下：

“在实验 2 中，我们对老年被试的归因偏差进行控制训练，即训练被试对疾病做外部归因。在此基础上，我们拟检验在消极刻板印象威胁情境中，随着外归因偏差倾向的增加，老年人医疗决策的质量是否会得到显著的改善，以期进一步验证实验 1 的结果。”（见修改稿第 11 页，最后一段，蓝色字部分）

参考文献

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of personality and social psychology*, 89(6), 845–851.

第三轮

编委意见：感谢作者很好地回答了审稿意见并做出了修改。责编同意发表

主编意见：

这篇文章的选题具有现实意义，经过几轮修改研究质量有所提高。但从研究假设的写作方式看，问题不聚焦，似乎整个研究探讨的不是一个问题。建议重新凝练研究假设，并按照实验假设的写作方式阐述和写作，同时请注意两个实验与实验假设的对应关系，以及两个实验之间的关联。修后再审。

回应：感谢主编专家的建议。根据您的建议，我们在本轮修改中主要做了以下几个方面的修改。

首先，为了平衡两个实验假设的数量，减少两个实验研究不够聚焦的问题，我们重新凝练了研究假设。本研究的目的是检验消极刻板印象对老年人医疗决策质量的效应，以及归因偏差在消极刻板印象对老年人医疗决策质量的效应中所起到的中介作用。基于此，原文中的两个假设（“假设 2：在刻板印象威胁情境中，老年人会对自身疾病的归因产生偏差。具体

地说，受刻板印象威胁的老年人会对自身疾病做出更多的内部归因偏差。”和“假设 3：归因偏差影响老年人医疗决策的质量。具体地说，内部归因偏差消极影响医疗决策质量，外部归因偏差积极影响医疗决策质量。”）与研究目的无直接关系，因此为了减少无关研究假设对读者阅读过程的干扰，我们在本轮修改中删除了这两个假设。

其次，我们结合本研究的实验设计，对剩余三个实验假设的表述方式做了进一步的修改。修改后的假设如下：

“假设 1：与不受刻板印象威胁的控制组相比，启动消极刻板印象的老年人医疗决策质量较差。”（见修改稿第 4 页，第 3 段）

“假设 2：启动相关刻板印象会通过影响归因偏差进而对老年人的医疗决策质量产生消极效应，即归因偏差在消极刻板印象对老年人医疗决策质量的效应中起中介作用。”（见修改稿第 5 页，第 3 段）

“假设 3：在启动消极刻板印象的情境中，减少内归因偏差的训练控制可以有效缓解刻板印象对老年人医疗决策质量的消极效应。”（见修改稿第 5 页，第 5 段）

第三，在修改稿中，我们明确了两个实验与假设的对应关系。具体地说，实验 1 检验假设 1 和假设 2，实验 2 检验假设 3。我们已经将两个实验与实验假设的对应关系整合到了修改稿“引言”的最后部分，具体内容如下：

“基于以上论述，本研究拟采用两个实验来探讨消极刻板印象、归因偏差和老年人医疗决策质量之间的关系。具体来说，实验 1 检验假设 1 和假设 2，通过阅读文字材料操纵刻板印象情境和控制情境，比较两种情境中被试归因偏差和医疗决策任务表现的差异，考察消极刻板印象对老年人医疗决策质量的影响（假设 1），并检验归因偏差的中介作用（假设 2）。”（见修改稿第 5 页，第 6 段）

“在实验 1 的基础上，实验 2 检验了假设 3。……”（见修改稿第 6 页，第 2 段，第 1 行）

第四，我们也加强了两个实验之间关联的论述。具体来说，实验 1 的目的是检验消极刻板印象对老年人归因偏差和医疗决策质量的效应，并通过统计分析技术探讨归因偏差的中介作用。为了验证实验 1 中获得的归因偏差中介作用的稳定性和可靠性，本研究参照因果链实验设计中介效应检验的方法设计并开展了实验 2。实验 2 的目的是检验归因偏差对医疗决策的质量。实验 2 在实验 1 揭示了刻板印象对归因偏差存在显著影响的基础上，进一步检验归因偏差对老年人医疗决策质量的效应。结合这两个实验的结果，本研究相关变量间因果链的两个重要效应（刻板印象对归因偏差的效应，以及归因偏差对医疗决策质量的效应）可以得到完整、系统的检验。综上，关于两个实验的关系，一方面，实验 1 和实验 2 是因果链实验

设计必须包括的两个子实验,通过两个实验的结果来揭示归因偏差在刻板印象对老年人医疗决策质量效应中的中介作用。另一方面,实验2的结果是对实验1结果的进一步确认和验证。实验1通过统计分析手段揭示了归因偏差的中介作用,实验2则通过实验操作的方法进一步检验了归因偏差的中介作用,从实验设计和实验操作的角度验证了实验1的结果。我们在修改稿中对两个实验的关系作了进一步澄清,具体如下:

在引言部分倒数第二段,我们加入了如下内容:

“为了验证归因偏差的中介作用,深入地检验归因偏差中介作用的稳定性和可靠性,本研究借鉴因果链实验设计中介效应检验方法,设计了实验2。在实验1检验了消极刻板印象对归因偏差和老年人医疗决策质量影响的基础上,实验2则关注归因偏差对老年人医疗决策质量的影响。通过结合两个实验的结果,我们能够揭示本研究变量间因果链的两个重要效应:消极刻板印象对归因偏差的效应,以及归因偏差对老年人医疗决策质量的效应。在此基础上,实验2可以从实验设计和实验操作的视角探讨归因偏差在消极刻板印象对老年人医疗决策质量效应中所起的中介作用,从而进一步验证实验1的主要结果。”(见修改稿第5页最后一段和第6页第1段)

在实验1讨论部分,为了引出实验2,我们加入了以下内容:

“为了进一步检验归因偏差中介作用,同时验证实验1主要结果的稳定性和可靠性,本研究参考因果链实验设计中介效应检验的方法,设计并开展了实验2。在实验1揭示了消极刻板印象对老年人归因偏差存在显著影响的基础上,实验2则关注了归因偏差对老年人医疗决策质量的效应。根据实验1的研究结果,如果归因偏差确实发挥了中介作用,那么对归因偏差进行训练控制,减少刻板印象威胁条件下被试的内归因偏差水平,则应该可以有效降低刻板印象对医疗决策质量的消极影响。”(见修改稿第11页最后一段,和第12页第1段)

最后,为了确保文稿语言表达的准确性和流畅性,我们在本轮修改中还对文稿的语言又进行了逐句地打磨和修改,由于这些修改比较琐碎和零散,这里就不再一一列举,详见修改稿棕黄色字体部分。

希望我们的回复和对文稿所做的上述修改能够解决您的关注。再次感谢主编老师对本文的辛苦付出!