

《心理学报》审稿意见与作者回应

题目：生育年龄限制感提高女性的计划生育数量

作者：邢采，孟戡琦，林青青，秦子玉

第一轮

审稿人 1 意见：

意见 1：前言中提到“而在 9000 多万新增可生育二孩的目标人群里，有着最强烈生育意愿的，恰恰集中在这 4500 多万人中，尤其是城镇妇女。这一部分女性在进入育龄后，生育自由即受到当时执行的计划生育政策的限制，在生育限制的政策解除后，又面临年龄的限制，但是她们的生育意愿并没有被这些限制所削弱，反而明显高于生育自由几乎完全不受限制的年轻一代女性。”但文中作者没有引用任何证据来支持这些观点，例如“有着最强烈生育意愿的，恰恰集中在这 4500 多万人中，尤其是城镇妇女”以及“但是她们的生育意愿并没有被这些限制所削弱，反而明显高于生育自由几乎完全不受限制的年轻一代女性”。请表明这些观点来自于作者自身的猜测，还是有相关数据和证据支持？

回应：感谢审稿专家对文章认真严谨的审查。上述观点不是作者的自身猜测，有相关出处，来源于《时代周报》，作者：刘金环、傅明，该文于 2015 年 12 月 8 日在搜狐新闻转载（网址：<http://news.sohu.com/20151208/n430194421.shtml>），本文的观点引用自此转载的搜狐新闻。此观点来源已补充至正文第四段相应位置，并添加在文末的参考文献列表中。

作为补充说明，如文章中第四段所述，这一观点的原出处来源于国务院新闻办新闻发布厅于 2015 年 11 月 10 日举办的新闻发布会，召开此发布会的新闻可在多家大型网站查询到，发布会的具体通知如下：

“国新办定于 2015 年 11 月 10 日(星期二)上午 10 时在国务院新闻办新闻发布厅举行新闻发布会,请国家卫生计生委副主任王培安介绍坚持计划生育基本国策、实施全面两孩政策有关情况,并答记者问。”

意见 2：前言中，尽管作者提到了诸多理论，如时间知觉，心理阻抗。但这些理论概念在研究中并没有涉及或测量。因此，前言理论和实证研究有着较大的脱节。

回应：感谢审稿专家的建议。如审稿专家所指出的，我们在前言中所提到的这两个理论，在研究中没有直接测量。我们非常感谢审稿人提出的建议，并且针对此建议，在前言中删除了和社会情绪选择理论有关的内容，并将心理阻抗理论的内容进行了大幅删减（第 9 页）。

但是我们最初在前言中论述这两个理论也是出于一些考虑，以下我们解释这两个理论和本研究的关系。如果审稿人要求，我们也愿意把这两个理论补充回前言中。

首先，社会情绪选择理论提出个体剩余时间知觉的变化和动机状态的变化之间存在因果关系，当剩余时间知觉发生变化时，会导致动机状态发生变化。本研究所以提出的当女性认为自己距离女性最佳生育年龄的时间限制越接近则生育意愿越强烈这一观点，和基于社会情绪选择理论所做出的预测密切相关。

关于心理阻抗的理论，我们认为可以用于预测并解释本研究中所发现的结果的。但是出于避免被试猜测到研究目的这一考虑，本研究中没有直接测量心理阻抗。对此我们在总讨论部分的 6.2 中，将这一点作为文章的局限性加以说明，并同时引用发表在 *Journal of Personality and Social Psychology* 中提出类似观点的另一篇文章（Wisman & Goldberg, 2005）作为支持。

意见 3: 实验二中，作者并没有报告被试年龄这一重要变量。同时，在控制变量中也没有报告各组是否存在年龄差异。

回应: 感谢审稿专家的细致的建议。实验二均为在校大学生（含本科生和硕士生），没有收集被试的年龄数据，但是收集了被试的年级。对三个启动组的年级进行似然比卡方检验，结果发现，三组中每个年级的人数无显著差异， $\chi^2(6) = 7.037, p = 0.317$ 。这一结果补充在实验二的研究方法 4.1.1 中（第 12 页）。

意见 4: 实验二的结果和实验一的结果不一致，为何 32 岁和被试年龄相差更远（相比与 26 岁），被试对孩子的内隐态度反而越高。

回应: 感谢审稿专家提出的非常有建设性的问题。我们在按照审稿专家提出的第 6 条审稿意见对实验二的数据进行重新处理后，数据分析得到的结果和之前发生了变化。对于最佳生育年龄的启动对前测生育意愿高的被试对孩子的内隐态度影响显著， $F(2,145) = 3.726, p = 0.026, \eta_p^2 = 0.049$ ，且 26 岁组的被试对孩子的内隐态度最高，与控制组差异显著（事后多重比较 $p = 0.008$ ），32 岁组则居中。这一结果和实验一的结果保持了一致，当女性被试认为自己距离最佳生育年龄越接近时，其生育意愿越高（实验一），且对孩子的内隐态度越好（实验二）。感谢审稿专家对于两个实验之间结果的比较，这对于更有效地、连贯地理解本研究的三个实验的发现很有帮助。在此修改稿中，我们在实验二的研究假设提出部分进行了相应的修改，预期实验二的结果会进一步支持实验一的假设（第 10 页），并且在实验二的讨论部分对实验一和实验二结果的进行比较和讨论（第 13 页）。

意见 5: 第 9 页，报告效价并无显著差异。此处应为笔误？积极和消极词的效价应该为显著差异。并请给出积极和消极词的具体例子。

回应: 感谢审稿专家细致的审阅。此处确为笔误，积极词和消极词的效价差异显著，正文中已修改，并同时列出积极词和消极词的例子各两个（第 12 页）。

意见 6: IAT 的数据处理并不规范，IAT 测量的是相对态度，单独分析动物和小孩，并无根据，除非使用的是 single target IAT 这类 IAT 变式。

回应: 感谢审稿专家提出的非常专业的意见。我们之前单独分析动物和儿童的 IAT 分数是出于想了解被试对两者的内隐态度。在审稿专家的提醒下，我们意识到之前的处理方法是够规范的。在此修订稿中，我们重新对实验二的 IAT 数据进行了处理，按照 Greenwald et al.(2003) 所提出的方法重新计算了儿童相对于动物的 IAT 差异分数，并以此 IAT 差异分数重新进行了数据分析。并将数据分析的结果在实验二的结果部分进行汇报（第 13-14 页）。

意见 7: 作者将 IAT 分数代表生育意愿并无支持，首先请报告 IAT 的差异分数和生育意愿的相关，而不是动物 / 小孩和生育意愿相关的差异。其次从逻辑上来说，这里两组实验组被试的 IAT 分数是在实验操作后测量，因此受到了实验操作的影响，并不能有效作为生育意愿的相关指标。所以，请特别给出控制组条件下 IAT 的差异分数和生育意愿的相关。

回应: 感谢审稿专家提出的非常负责任的建议。我们在原稿中对这一问题的表述的确不够清晰准确，给读者的感觉是我们以 IAT 分数等同于生育意愿。我们的本意是以 IAT 分数作为推测被试真实生育意愿的一个参考值。因为内隐联想测验可以测量被试不愿意如实汇报或者不能够准确汇报的真实态度，因此，IAT 分数可以帮助我们了解被试对于儿童真实的态度，以此作为参考，进一步推测被试对于自己生育孩子的真实态度。除被试在研究一和研究三中的自陈量表中汇报的生育意愿外，研究二所测量的 IAT 分数可以作为一个重要的补充，帮

助我们更为准确地了解女性被试的真实生育意愿。我们在文章中对这一观点进行了更为清晰准确的表述和解释。

在本研究中，我们并没有对 IAT 分数是否会和被试在自陈量表中所汇报的生育意愿有显著相关进行预测。因为 IAT 测量的是被试不愿意或不能够准确汇报的态度，因此如果被试在自陈量表中对自已真实的生育意愿没有准确做答或者自己也说不清自己真实的生育意愿，那么 IAT 分数可能和自陈的生育意愿无相关；但是如果被试能够清楚地知道自己真实的生育意愿，并且在自陈量表中如实汇报，那么 IAT 分数则会和自陈的生育意愿有显著相关。研究二中被试所完成的问卷中主要为社会人口信息的问题，没有关于生育意愿的问题，因此无法计算 IAT 分数和自陈生育意愿的相关。但是，我们按照审稿专家的建议，单独分析了控制组中被试对于“孩子和生育意愿的相关程度”以及“小动物和生育意愿的相关程度”两个问题的差异，结果发现差异显著，控制组被试认为孩子比动物和生育意愿的相关更为显著， $t(49) = 8.02$ ， $p = 0.000$ （第 13 页）。

意见 8：在研究三中，在表一中请表明被试数量，是单独的控制组相关，还是控制组和实验组所有被试合在一起的相关。另外请作者单独报告实验组的年龄和计划生育个数的相关，因为该分析表示了是否在实验组被提醒了 40 岁为生育限制年龄后，年龄越大的被试（紧迫感越强）会有更高的生育意愿或计划生育个数。

回应：表一中的被试数量已补充在表一前面的文字描述中（ $N=113$ ，第 16 页），是所有被试合在一起的相关。另外，按照审稿专家的建议，我们对实验组的年龄和计划生育个数进行了相关分析（第 17 页），结果发现，年龄越大的被试计划生育孩子的个数越少（ $r = -0.315$ ， $p = 0.016$ ）。我们分析出现这一结果的原因可能有两个，并将此分析补充在总讨论中（第 18 页）。其一，实验二中通过提醒被试女性的最佳生育年龄对被试进行启动，而在实验三中，则是通过提醒女性生育孩子能力的转折点对被试进行启动，被试在阅读启动材料时被告知超过 40 岁的女性生育孩子时出现的健康问题明显增多，这一带有负面信息的启动材料有可能导致年龄接近 40 岁的女性因为对健康问题的担心出现生育意愿的下降。另一种可能性是由于实验三中的被试和实验一、二中的被试有明显的区别。实验三中被试均为已婚女性，而实验一和实验二中的被试均为未婚女性，且实验三中的被试年龄明显更大（实验三： $M=30.56$ 岁， $SD = 4.25$ 岁；实验一： $M=22.26$ 岁， $SD=2.11$ 岁，实验二未测年龄，被试均为在校大学生，年龄与实验一中被试的年龄接近）。有可能是启动女性生育年龄的时间限制对生育意愿的提升作用主要针对未婚女性效果更好。当然，这只是我们基于现有数据的一个合理的推测。澄清这一问题对于解释剩余限制对生育意愿的影响十分重要，并且对于之后的政策建议的提出也有重要的指导作用，所以在今后的研究中需要开展更为细致有针对性的实验，以更清楚地澄清这一问题。

.....

审稿人 2 意见：

意见 1：该论文作者采用问卷方法结合决策学中的心理阻抗和社会情绪选择理论对于女性生育意愿的影响做了相应研究，结论为当女性距离理自认为的理想生育年龄越接近时，生育愿望越强烈，在因果性地启动了生育年龄限制后，该时间限制能够明显提升未婚女性的生育意愿和已婚女性的生育计划。该研究的实验设计和实验假设都没有太多瑕疵，唯一令人担忧之处在于研究的实际应用价值。生育意愿和实际生育之间有非常明显的跨越，对于被试进行最佳生育年龄的启动唤起生育愿望和实际进行生育准备距离较远，实际生育涉及到的有关因素考虑实际非常庞杂。在市场调查中采取的“限时打折”营销策略也能够短暂提高被试的购买欲望，但是和实际购买行动之间仍存在差异。此项研究如果能够获得田野调查或者跟踪的实际

数据则可信度和实际应用价值能够得到明显的提升。目前看来，国家长期曾经对优生优育以及优育时间做过大范围的普及宣传，但是明显最后影响育龄妇女生育的明显因素并不是生育限制感，而是其他更多的物质因素。建议作者在讨论部分对该手段如何能够更好地对女性实际生育意愿造成影响作出更加符合实际的阐述。

回应：非常感谢审稿专家对本研究的肯定，并且也非常感谢审稿专家对本论文的讨论部分所提出的中肯的建议。我们完全同意审稿专家所提出的生育意愿和实际生育之间的差距。我们对审稿专家的建议进行了认真的思考，对总讨论的政策建议部分进行了大幅的重写。经过反复推敲本研究的发现，查阅相关资料，并咨询人口学研究专家，我们能提出的有较大把握的政策建议为：如果仅以放开生育限制作为唯一政策手段来提升生育率（最近的新闻里还提到有放开二胎的建议），那么很可能会适得其反。

虽然基于本研究的结果还可以进一步提出其他建议（如当年轻的育龄女性认为女性的最佳生育年龄较低时其生育意愿提升更加显著，所以可以考虑在宣传中强调女性的最佳生育年龄为二十几岁），但是其他的结果我们认为还需进一步研究的检验。所以从负责任的角度考虑，我们没有在文章中没有给出其他建议。

请审稿专家批评指正，我们希望有计划在审稿专家的建议下进一步修改完善本研究的政策建议部分的内容。

这一段内容具体改写为如下内容：（第 19-20 页）

作为一项探索性研究，本研究的最重要的贡献在于采用实证研究证实一定的生育限制感不仅不会降低生育意愿，反而会提升生育意愿。印度的计划生育政策与人口出生情况的变迁与此研究结果趋势一致（莫蕙芳，1986）。印度在 1941 年至 1951 年间，人口由 3 亿 1900 万增至 3 亿 6100 万，增加 4200 万人。印度政府在制定第一个五年计划时（1951-1956），明确必须开展计划生育工作，但仅以宣传鼓励为主，没有采取强制措施，结果印度人口在 1951 至 1961 年间由 3 亿 6100 万激增至 4 亿 3900 万，十年增长 7800 万；到 1971 年印度人口已达 5 亿 4800 万，十年增长 1 亿 1000 万人口。1976 年印度公布“计划生育法”，规定对不执行计划生育政策的人实行限制和处罚政策，之后人口出生率开始出现下降。印度的人口数据表明，生育限制的政策在未伴随强制措施的情况下，反而可能导致人口出生率的激增。根据本研究的结果推测，解除生育限制的政策不仅不会提升生育意愿，反而有可能降低生育意愿。与此思路一致，美国国家科学基金会(NSF)公布在取消某些项目的截止日期之后，项目的申请量减少了一半。因此，基于本研究的结果和其他相关的研究结果，我们能提出的有较大把握的政策建议为：如果仅以放开生育限制作为唯一政策手段来提升生育率，那么很可能会适得其反。并且，我们不建议以进一步放开生育指标的限制（如放开二胎或全面放开生育数量）作为促进生育意愿的主要手段，这反而可能导致育龄女性的生育意愿进一步下滑。

第二轮

审稿人意见：

意见 1：感谢作者的详细回复。文章有了非常大的改进。

回应：感谢审稿专家对一修稿的肯定和认可。

意见 2：我仍然对于最后一段有保留意见。作者提到了不应该做什么，但却没有明确指出应该做什么来提高生育意愿。另外，最后一句，即使是放开二胎，那么二胎仍然可以理解作为一种生育限制（相比于没有任何生育限制的国家/地区）。所以解除生育限制可能会降低生育

意愿用三胎的例子并不合理。而且文章的实验三发现启动生育限制增加了该组被试计划生育孩子的数量，这个结果也与最后的建议不符合。

回应：感谢审稿专家的建议。针对审稿专家提出的建议，我们进一步修改了政策建议这一部分的内容，删除了涉及到放开三胎的评论，并且结合实验三发现的启动生育限制增加改组被试的计划生育孩子的数量这一发现，补充了一条关于如何做可以提高生育意愿的建议。

具体而言，在此修改稿中，我们提出两条政策建议。“第一，在放开二胎的政策对提升生育率不理想的情况下，不建议匆忙以进一步放开人口生育指标（如全面放开生育数量的限制）来解决生育意愿低的问题。第二，创设令女性对生育年龄有限制感的环境，如在育龄女性每年的体检中，增加关于女性适宜生育所余时间的信息提示。”同时，我们将这一部分的子标题改为“6.3.2 提升育龄女性生育意愿的建议”。

并且，为了进一步突出强调本研究的主要发现和政策建议，我们将文章的题目修改为“生育年龄限制感提高女性的计划生育数量”，并对摘要进行了改写。

请审稿专家批评指正。

第三轮

审稿人 1 意见：

意见 1：修改稿大部分的改动比较好。

回应：感谢审稿专家对二修稿的肯定和认可。

意见 2：但我还有一个疑问希望作者可以解答。在实验二的 IAT 测试中，研究者发现在高意愿被试中，相比于控制组，26 岁启动组的 IAT D 分数提高。但这里的 IAT D 分数和低意愿被试相比并无显著差异。差异的来源在于控制组的 IAT D 分数异常小，这就提醒可能这里的操作并没有效果，仅仅是因为控制组的 IAT D 分数过低导致的差异？希望作者能对这点进行解释和讨论。

回应：感谢审稿专家这条细致、专业的意见。以下是我们对涉及到的变量的澄清以及对结果的解释。在实验二的内隐联想测验的结果汇报中，涉及到两个自变量，分别为前测生育意愿（两个水平：高生育意愿 vs.低生育意愿）和最佳生育年龄提醒（三个水平：控制组/无提醒组 vs.提醒 26 岁为最佳生育年龄组 vs.提醒 32 岁为最佳生育年龄组）。其中前测生育意愿不是我们操纵的自变量，是在招收被试时测量的；而生育意愿是在做实验时操纵的自变量。结果显示，在高生育意愿组中，26 岁启动组 IAT D 分数明显高于控制组；其实 32 岁启动组也高于控制组，但是差异尚未达到显著水平，所以未在文章中提出。这个结果正是我们所预期的。正如审稿专家所指出的，这个结果主要是由于控制组 IAT D 分数小导致的；而实验二的假设之一即是启动女性最佳生育年龄的时间限制会导致女性的生育意愿上升，在 IAT 任务中体现为两个实验组的 D 分数比控制组更高，所以实验二的预期就是控制组的 IAT D 分数明显小于两个启动组。并且实验二另一个假设是 26 岁启动组对 D 分数的提升更明显（因为距离被试的实际年龄更加接近），实验二的结果也和这个假设相符。

正如审稿专家指出，高生育意愿组的两个实验组（26 岁启动组和 32 岁启动组）的 IAT D 分数和低生育意愿组差异不显著。而且另一个情况是前测低生育意愿组的控制组被试 IAT D 分数并没有明显低（与预期不符），只有前测高生育意愿组的控制组被试 IAT D 分数明显低（与预期相符）。我们分析出现这种情况有两个可能的原因：第一种可能性是，启动女性的最佳生育年龄的时间限制对于女性的生育意愿的影响不受到前测生育意愿高低的调节。第二种可能性是，采用自陈汇报的方法测量的女性生育意愿更多反映女性的外显生育意愿，

而采用 IAT 测量的对孩子的喜爱更多反应的是女性内隐的生育意愿，两者有差异。有可能出于外部的（如现实中年轻女性生育面临的困难，如经济困难等）和内部的（如防御机制等）原因，内隐生育意愿高的女性在自陈汇报时汇报了较低的生育意愿；也有可能出于外部的（如传统习俗对女性必须生育的期望，等）和内部的（如自我保护等）原因，内隐生育意愿低的女性在自陈汇报时汇报了较高的生育意愿。这种内隐和外显生育意愿之间的差异有可能是被试自己也没有察觉的。正是出于对于如何准确测量女性生育意愿的考虑，本研究的实验二采用了内隐联想任务，实验三又补充了对与生育行为倾向的测量。这些尝试都是为了进一步探索能够反应女性真实生育意愿的测量指标。因为女性的生育意愿是影响女性生育水平的重要因素，也是预测国家生育水平发展趋势的最为有效的指标之一（陈卫，靳永爱，2011）。作为世界第一人口大国，准确地预测国家生育水平不仅具有重要的学术价值，同时对国民经济社会发展全局中也具有重要的意义（卢海阳，邱航帆，郑逸芳，2017）。

因为本研究所包含的三个实验均无法提供充分的证据支持上述两种可能性中的任何一种，也无法彻底排除其中一种，所以在前一稿中，我们没有深入探讨这一问题。但是正如审稿专家所指出的，不讨论这一问题会使读者对结果的疑问没法在文章中得到呼应。因此在本修改稿中，我们按照审稿专家的建议，针对这一问题对实验二的结果进行了更深入的解释和讨论，具体在第 17 页的“4.3 讨论”中。并且，在本修改稿的总讨论部分，我们结合本研究的发现，对女性的生育意愿的测量方法进行了深入的探讨（第 20—21 页的 6.3.2 中）。

第四轮

编委意见：

建议作者进一步精简语言，尤其是“研究假设”部分，将稿件全文控制在 1 万字左右。修回后可进行主编终审。

回应：感谢编委专家的意见。我们按照编委专家的意见，对全篇文章进行了进一步的凝练。主要的修改在两个地方，其一是研究假设部分，由之前的三段话 1238 字，精炼至一段话 663 字。其二是文章开始的引言部分第二段话，这段话中介绍了在 2016 年生育政策调整之后生育率不及当初预期的情况，因为最初投稿时为 2017 年，当时尚不知道 2017 年的生育率，本修改稿中对这一部分涉及的生育率的统计信息进行了更新和精炼。修改完后本文为 10638 字（含标题、摘要、关键字和所有图、表的字数）。