

《心理学报》审稿意见与作者回应

题目：认知模糊程度和动机强度对有意识和无意识自我欺骗的影响

作者：钟罗金 汝涛涛 范梦 莫雷

第一轮

审稿人 1 意见：

意见 1：总的印象：文章提出了一种改进的范式，使得定量地测量外显的和内隐的自我欺骗成为可能，这是一种很好的思路，且被证明是有效的。文章把着力点放在自我欺骗内隐和外显的区分和测量上对该领域的理论改进有较大的贡献。

回应：感谢审稿专家对本研究理论价值的认可！

意见 2：然而，遗憾的是，作者没有把这一方法很好地利用，这表现在实验二和实验三的设计都不够过硬，这主要体现在以下三个方面：一、这两个实验没有控制组，两个实验组互相比较是不符合设计逻辑的。二、操纵的效度存疑，实验二中通过“正印”和“反印”答案操作认知模糊性的方式有没有前人文献的支持？如有，需要举例说明，如果没有，需要进一步“从理论上”说明这种操作的意义，因为我们对此完全可以有其他的解读，如我们也可以认为“正印的答案增加了作弊时的流畅性，此时的认知更清晰（认知模糊性更小）”等等。实验三力图通过报酬操作动机，但对于被试，报酬的多与少可能只有相对的意义，在组间设计中这种操作的意义不大。三、三个实验作者都力图对被试做个体水平的区分，如“无意识自我欺骗”组和“有意识自我欺骗”组，但是并没有在这种区分的基础上尝试做其他的变量操作。因为被试的个体差异是普遍存在的，如果没有进一步的任务，这种区分的意义就不大了，它就是前面分析的一种重复。

回应：感谢审稿专家提出三个重要的意见！这三个意见让我们对本研究设计有了更深刻的思考。

第一，本研究实验二和实验三的确没有对照组，正如专家指出的若有对照组这两个实验会更完美。但是我们认为这并不会影响两个实验的效度。主要原因有三：1.本研究的实验一已经有了一个对照组，实验二和实验三如果设置对照组，因为材料、程序和实验情境都一样，结果也应该相近。如果将实验一的对照组纳入到实验二和实验三分析的话，得到的结论和现在的实验二和实验三的分析一致。因此，为节约成本我们在实验二和实验三没有设置对照组。2.实验二和实验三可以看成是实验组之间相互对比的实验。因为无论是对照组和实验组比较，还是实验组和实验组比较都是合理的。3.如果从单因素实验来看，如果两个水平可以证明结论没有必要再多设置一个水平。当然，正如专家所言，如果有对照组实验二和三的结果将更加可靠。

第二，关于实验二和实验三的实验操作确实没有人在同类研究中使用过类似的操作。这也是本研究的创新之处。首先，实验二正印和反印是基于 Sloman, Fernbach 和 Haggmayer 的研究(2010)。他们提出认知模糊性概念，我们在本实验中结合新的实验方法重新进行的操作定义。其基本假设是基于偷看反印的答案更困难，偷看行为更突出，作弊更不流畅，需要更多的认知资源，投入了更多的注意力，更难否认偷看答案的事实，因而是认知模糊性更小。我们也请了专家对这个操作的合理性进行论证，正如审稿专家指出的那样，论证专家也指出有一个可能会出现干扰是，正印的时候因其作弊更加流畅，被试会偷看更多的答案，偷看

经科学的证据。

回应：谢谢审稿专家这个专业的建议！关于两种自我欺骗的称呼是“内隐”和“外显”，还是“无意识”和“有意识”。我们有过思考，内隐更多是指无需意志力或注意力自动化的过程；而意识更重于个体对行为的觉察程度。本研究的自我欺骗并非是无需意志自动化的过程，而主要是指个体做出这个行为后是否觉察到。因此我们最后选择了无意识自我欺骗和有意识自我欺骗。另外，正如专家指出的，本研究仅探讨了两个变量对两种自我欺骗的影响，无法窥探到两种自我欺骗的内在构造、功能和原理及其相互关系，不足以称之为机制。但是目前确实没有找到合适的词语来概括这种变量之间关系的特点，而这种关系和特点确实也预示了两种类型的自我欺骗具有不同的构造和性质，所以我们才称之为机制。另外，感谢审稿专家给我们指出了两个未来研究两类自我欺骗很好的思路！我们会按照这个思路去进行更多的后续研究，尽量让自我欺骗的机制更加完整。

意见 5：从研究手段上看，大规模集体测量可能不是一个好的方式，你如何快速地得出被试的成绩？如何快速地发放报酬？这个都可能导致众多无关变量的混淆，如被试可能在你判卷子过程中有彼此交流，互相攀比等等，这些都可能影响实验的结果。

回应：谢谢审稿专家关于实验程序的专业建议！集体分组测试确实可能会存在审稿专家提到的上述问题。为此，我们在实验的过程中尽量做到严格控制：首先，为了快速得出成绩，每组测试尽量控制在 20 人左右，每次有 4-6 名非常熟悉测试题答案的主试。其次，被试每次答完题后将卷子交给主试批改，批改后不发报酬只给出分数和金额数，三个卷子最后一起上交再根据得分总和给被试计发报酬。测试中实行分组批改，即事先指定哪部分的被试归那个主试负责批改。批改的过程中，只有当主试面前没有被试在批改时，其他被试才可以起身前往主试处批改，否则在座位等候。另外，每次测试至少有一名主试在现场维持秩序。最后，我们在实验开始前对被试之间的距离，保持会场安静和做题的秩序会做出严格布置和要求，并在指导语中明确要求被试间不可以交流。总体而言，三个实验过程中的每次测试都比较顺畅、环境能够保持安静、在会场没有出现相互交流的情况。

意见 6：此外，前言中，作者对实验二和实验三的理论论证不够，且缺乏具体的假设，希望加以补充。

回应：已接受审稿专家意见。将前言中对实验二和实验三的理论论证放到实验二和实验三开始前进行论述。

意见 7：具体建议： p.3. “实验 2 的结果发现，认知模糊程度影响无意识自我欺骗和有意识自我欺骗；而实验 3 的结果发现，动机强度仅影响有意识的自我欺骗。”摘要应该更清晰地阐述，至少影响的方向应该说清楚。

回应：已按照审稿专家意见在摘要中修改。谢谢审稿专家的指出！

意见 8：p.4. “即个体能通过怎样的策略和手段使得自己能够相信正确观念的同时又相信错误观念...”，因为自我的观念是主观的，此处，用“正确”，“错误”似乎不妥，可改为“相信一种观念的同时又相信相反的观念...”

回应：非常认同审稿专家这一意见！已在原文修改。

意见 9：p.4 "Gur 和 Sackeim(1979) 的研究..."，这个实验中，为什么失败体验会否认自己的声音？这个实验的来龙去脉没有说清楚。

回应：已在原文添加说明。

意见 10: p.5. "即在第二次任务中每一次反应后对被试的反应速度给予精确的反馈或模糊的反馈...", 在这里, 什么是精确的反馈, 什么是模糊的反馈?

回应: 精确反馈是反馈本次任务速度在上一个实验中每次任务速度中的精确位置; 模糊反馈则只反馈与上一次实验相比是快了还是慢了而正常速度不反馈(快了, 在上一实验所有反应速度最快的 30% 区间内; 慢了, 在上一实验所有反应速度最慢的 30% 区间内; 不反馈, 落入中间 40% 区间内的速度)。已在原文添加说明。谢谢审稿专家的指出!

意见 11: p.5-p.6 "结果发现当给予精确反馈时被试的自我欺骗消失...", 是指“反应时加快的现象”消失吗?

回应: 是的。

意见 12: 又如 p.16.表 3 中“认知模糊程度大”, “认知模糊程度小”, 建议将“条件”改为“正立”和“倒立”合适, 不能用想当然的潜变量代替外显操作。

回应: 已按照审稿专家意见在文中修改。

意见 13: p.6. "Lopez 和 Fuxjager 的研究(2012)认为自我欺骗的个体会更加积极和正面的看待自己", 此处“的”应该为“地”。

回应: 已按照审稿专家意见在文中修改。

意见 14: p.6 "由此提示, 实验组被试自欺地将偷看答案的成绩归为个人能力...", 在这里, 如果被试不知道有没有答案就不能这么说, 如果测试一的答案帮助了被试“练习”, 或者提高了被试思维的流畅性(因为两次测验类型相似), 也不能归因于是自我欺骗。后面作者的实验采用类似的范式也存在同样的问题。

回应: 谢谢审稿专家指出! 首先, 自我欺骗的判断是根据试卷 2 的心理预估成绩来判断的, 并非真实的测试成绩来判断, 因此练习效应或思维流畅性对心理估计没有影响。其次, 本研究的三个测试题虽然是题型相同的等价测试, 但是这些测试主要考察是个人知识储备水平, 没有太多的做题技巧在里面。如果他的知识水平不够, 再怎么练习, 思维再流畅也没有办法做出答案。这可能是因为我们没有附上测试题目导致了审稿专家的误解, 这次修改稿我们将测试材料附上供审稿专家审核。

意见 15: p.6. "这种仍然存在的自我欺骗, 导致他们预估的准确性更差, 获得的奖励也更少...", 被试在完成测验时, 基于本能, 永远是以“得高分”为最终目的的, 而不是以真测验成绩和估计成绩一致为主要目的, 所以从被试的角度看, 这个“估计”可能根本就不那么重要。这个奖励可能也和社会生活中习得的“奖励”有很大不同, 所以“奖励无效”并不能说明太大的问题。

回应: 谢谢审稿专家的意见! 这可能是因为在原文中没有表述清楚。这里指的 Chance 等人研究中的另一个实验, 对预估准确性进行重大奖励。即使被试作对的题目不多, 但是如果预先估计出来的成绩和实际成绩没有差异的话, 也能额外获得丰厚的报酬。这样被试会比较认真预估, 尽量按照自己真实的知识水平来估计, 做到一题不差, 获取丰厚的估计准确性报酬。但是无意识的自我欺骗者却无法做到这个准确性, 因为他们真的把偷看答案的成绩当成了自己的真实水平。感谢审稿专家对这一点的质疑! 我们已在原文中修改, 做了更加清晰的表述。

意见 16: p.8. 作者阐述的实验逻辑清晰, 是一个很有趣的, 但是被试的真实成绩很重

要，这个需要作为协变量处理。

回应：谢谢审稿专家意见和建议！因为我们的三个实验都是单因素完全随机设计，假设两组被试的知识水平是相当的，而且测试的结果也表明两组的知识水平确实没有差异。因此，我们没有做协方差分析。这一数据处理方式也沿用了 Chance 等人(2011)的处理方式。

意见 17： p. 8. "就可以推断出该被试在试卷 2 的预估中是无意识的自我欺骗者（与卷 1 成绩接近）还是有意识（与卷 2 成绩接近）的自我欺骗者"，建议改为“就可以推断出该被试在试卷 2 的预估中是无意识的自我欺骗者（与卷 1 成绩接近）还是有意识的自我欺骗者（与卷 2 成绩接近）。”

回应：谢谢审稿专家指出！已在原文修改。

意见 18： p.9. "所有的被试除参加了本实验外，均没有参加本研究中的其他任何实验"，此句话语焉不详，建议删除，后面两个实验都有这句话，也可以删除。一个被试不能参加同一研究的多个实验应该是常识，不必指明。

回应：已按要求在原文中删除。谢谢审稿专家的建议！

意见 19： p.9. "这样形成了三份等价的试卷..."，只有 10 道题的常识性问卷，被试本身的知识结构会对测试结果，乃至自我评价结果有重大的影响：即使不管信度、效度如何，如何确保本研究的题目有足够的鉴别力？

回应：谢谢审稿专家这一专业的意见！确实如果题目鉴别力不够会对研究造成影响。这也是为什么我们前期要进行大量的预实验。在预实验中每道题有一百名被试预测，从而准确得出每道题的难度。然后，根据不同难度试题采取难度配对平均分配到三份试卷中，确保每一份试题的难度相近而且控制在 0.5 左右，以确保最大的鉴别力。

意见 20： p.10. "方差分析结果发现三者差异不显著（M 卷 1=4.3±1.45，M 卷 2=5.2±1.15，M 卷 3=4.4±2.18，F(2,59)=1.87，P>0.05）”，“检验不显著”不能认为就是“等价”，而且从平均分上看，差异还是挺大的：M 卷 1=4.3，M 卷 2=5.2。此外，为保证方法的可重复性，附件中或补充材料中是否应该加上这个实验材料？

回应：谢谢审稿专家意见和建议！此次修改稿中已经补充实验材料。另外，为了使三份实验材料等价，我们对大量的题目（100 道题）进行了大量的预实验，以确定每一道题的难度，然后按照难度相当和题目类型一致逐个分配到三份测试中。应该说仅这样的操作已经可以确保了三份试题的难度等价。但是为了更进一步的验证，我们还对三份试卷进行了预测试。预测试的结果也证实了我们的三份测试是差异不显著的。但是正如专家指出，第二份测试卷的平均分比其他两份都稍高一点。我们也注意到这个情况，我们认为这主要是因为预测试是验证性质，其被试量（每组 20 人）并没有预实验被试量（100 人）那么多。而且预测试结果必然会存在差异，但是只要差异在可控范围内，我们仍可以依据预实验的结果和程序认为三份试题是等价的。另外，本研究的三个实验主要关注的是被试估计成绩而非实际成绩，因此只要三份测试的难度差异不会太大，不至于被被试发现即可。

意见 21： p.10. "所有的被试首先完成试卷 1..."，完成这个测验给多少时间？被试是时间紧张还是时间宽松？

回应：谢谢审稿专家意见和建议！试卷的作答时间为 3 分钟，根据预实验的结果这个测试时间是 90% 以上的被试能够完成所有的题目。

意见 22: p.10. “然后要求被试快速浏览一下题目...”, 多快速? 只有 10 道题被试有没有可能全部看完? 这样对成绩的预估就没有效力了。

回应: 谢谢审稿专家意见和建议! 这个过程是非常短, 只有 30 秒。这个时间主要是为了让被试看清题目数量一致、有无答案和做出成绩预估的。其中一名主试念完这段指导语后会计时, 其他 4 名主试会监督本组的 4-6 个被试尽快完成预估。因此, 被试无法全部看完所有题目, 更没有足够的时间做完题目。

意见 23: p.10. “以确保实验组被试知道这是一份没有答案的试卷...”, 如何确保在第一次实验中实验组被试看到了答案, 或者让他们认为“印在下面的答案是? 如果是测试, 题目+答案的方式是不是很奇怪?

回应: 谢谢审稿专家意见和建议! 应该说肯定会有一点奇怪, 但这不影响实验效果, 因为我们主要关注试卷 2 的预估成绩; 而且被试可以自愿选择是否看答案。从实验过程来看, 并没有被试提出相关的问题。至于答案是否正确, 被试可以根据自己有把握的题目进行对照可判断; 而且即使不相信, 答案也会有指引提示作用。从两组被试的测试成绩的差异可以判断被试是否偷看了答案; 如果有答案的实验组的测试成绩远远高于没有答案的控制组成绩则表明实验组偷看了答案。

意见 24: p.10. “试卷 3 估计完后, 不需要进行作答 (被试事先并不知情)”为什么不作答? 这个成绩应该有, 而且很重要啊。

回应: 谢谢审稿专家意见和建议! 我们认为试卷 3 的测试成绩是一个真实成绩, 反应的是被试的真实水平。而这一成绩前面的测试中已经获得。如果试卷 3 再测也只是一种简单的重复。因此, 为避免测试时间过长, 导致被试厌烦以及从节约成本考虑, 没有进行试卷 3 的作答。

意见 25: p.11. 关于被试的“删除”: 这个删除方式有缺陷, 为什么不根据被试的自我报告结果删除被试?

回应: 谢谢审稿专家意见和建议! 这种集体分组的测试中被试不会承认是否偷看答案, 即使单独测试中被试也很难当面承认偷看了答案。因此我们没有设置自我报告, 即使有自我报告也不可信。我们只能通过后续的数据结果来“证明”被试是否偷看了答案。因此, 采用了现在的被试删除方式。

意见 26: 此外, 实验者怎么“事先”知道会删除这么多人, 而确保实验一中两组人数如此相当?

回应: 谢谢审稿专家意见! 确实无法知道实验组需要多少。一开始实验组和控制组都是 60 人左右, 后来发现实验组有较多诚实的被试, 不符合实验要求需删除。为了避免实验组和控制组人数相差太大, 后续再补了两组实验组数据 (每组 20 人)。

意见 27: p.12. “这个结果表明, 实验组被试不仅在卷 1 中偷看了答案, 同时还自欺地将偷看答案之后的结果归为个人能力, 认为自己在等价的卷 2 中也能取得好成绩...”, 这个结果也可能是由于有测验一的练习, 增加了被试思维的流畅性, 或者更“自信”, 这种可能如何解读, 或排除?

回应: 谢谢审稿专家意见! 我们的测试是一个知识的测试, 而非思维和能力等的测试。这种测试完全取决于个体的知识储备, 也不需要太多的应试技巧。如果被试不知道这个常识, 他再熟练测试技巧、思维再灵活, 也得不出正确答案。而过度的自信就是本研究认为的自我

欺骗，因为这种的自信不是建立在自己真实水平上的，而把偷看答案的成绩也归结于个人能力。

意见 28: p.12. "这一结果与 Chance et al (2011)的研究结果是一致的。正如前文所述，这一结果提示卷 2 的预估中可能既存在无意识自我欺骗者也存在有意识自我欺骗者..."，这个只从前两个测试可看不出来。

回应: 谢谢审稿专家建议！已在原文中进行修改，进行更为严谨的表述。

意见 29: p.12. "如果实验组卷 3 的预估成绩介于卷 1 的测试成绩和卷 2 的测试成绩之间，则说明卷 2 的预估存在两种形式的自我欺骗。这表明被试对卷 2 的成绩进行预估时，不仅存在无意识自我欺骗者，而且还存在有意识自我欺骗者。"这一段的比较不能只看实验组，而应该和对照组的成绩加以比对才能说明问题。

回应: 谢谢审稿专家建议！确实要以实验组和对照组的比较为前提。因为前面的分析已经进行过实验组和对照组的比较了。这里只关注实验组（自我欺骗）的被试中是否存在两种形式自我欺骗。此时，组内比较优于组间比较。因此，这里仅对实验组内进行分析比较。

意见 30: p.13. "对于介于两者中间，即卷 3 的预估成绩与卷 1 的测试成绩和卷 2 的测试成绩的距离相同，那么这类被试无法分清是属于哪一种自我欺骗者，所以这些被试暂不纳入后续分析，结果共发现 9 个这样的被试..."，怎么界定“相同”？是正好位于中数处呢？还是有一定的误差限度，需要说明。

回应: 谢谢审稿专家建议！这时是取中数处。已在文中添加说明。

意见 31: p.13. 在结果的分析中，必须考察被试的个人能力（即本身得分）这个先验的协变量，这种分析才有意义。

回应: 谢谢审稿专家意见和建议！因为我们的实验是单因素完全随机设计，假设两组被试的知识水平是相当的，而且测试的结果也表明两组的知识水平确实没有差异。因此，我们没有做协方差分析，数据处理方式仍沿用了 Chance 等人(2011)的处理方式。

意见 32: p.15. "将卷 1 中的答案以倒立的形式呈现在试题的下方。如果被试要偷看倒立的答案比较费力，这样被试在意识上能够更加明显地察觉到自己是否偷看了答案，因而对偷看答案的认知模糊性较小。"在这里，存在一个对“认知模糊性”的定义问题，即被试意识到作弊是认知模糊呢还是被试是否知道答案是认知模糊？我们甚至可以相反地理解，即如果被试看答案了，那么正立增加了操作的流畅性，也可以认为减少了认知的模糊。

回应: 谢谢审稿专家意见和建议！已在原文中做更加清晰的解释，并对认知模糊性进行定义。这里的认知模糊是根据 Sloman et al. (2010)定义，指的是被试对偷看答案行为的意识觉察程度，而非知道答案的流畅程度。

意见 33: p.17. "这个结果与前人发现模糊性会影响自我欺骗行为的结果是一致的 (Sloman et al., 2010)"，本实验对认知模糊性的操作的前人的方法不一致，能否直接对比呢？

回应: 这里确实和 Sloman et al. (2010)使用的实验方法不一样，但是我们对概念的定义和研究的对象是一致的。虽然方法不同，但是得出的结论是一致的。已将结果改成结论，做更准确表述。谢谢审稿专家意见和建议！

意见 34: p.18. "在低动机条件下，每答对一道题给 0.5 元的报酬；而在高动机条件下，

每答对一道题给 1 元的报酬..."，动机的高低是一个绝对的概念，而给钱的多少是一个相对的概念，故这种操作可能是无效的。

回应：谢谢审稿专家意见和建议！正如专家指出报酬的多少对不同被试而言可能是相对意义；对有的被试来说两种金额的报酬确实可能没有太大的差异。但是从随机选择两个组的群体来说，这样的被试在组内是可以平衡的。前人也有相关研究是通过报酬来操作动机的 (Mijović-Prelec & Prelec, 2010)。从我们的实验结果来看这种操作是成功的，高报酬组激发了高的欺骗动机，更多的被试做出了偷看答案的行为($\chi^2(1)=9.62, p=0.002, w=0.28$)。

意见 35：p.20. 关于讨论：一般的设想，如果被试想，看答案了，应该全对；被是不看答案，应该和控制组相当，不存在中间的条件。为什么在提供答案的情况下，被试仍不能回答对全部 10 道题？是给的时间不够，还是被试的主观原因？这个应该加以分析。另外，我很感兴趣的事，在几个实验中，有多少被试（在有答案时）是全部答对的？

回应：谢谢审稿专家这一宝贵建议！从实验结果来看全答对的人数并不多，实验一控制组 62 人中有 1 人，实验组 103 人中有 18 人；实验二正立答案组有 5 人（共 50 人），倒立答案组有 7 人（共 50 人）；实验三低动机组有 6 人（共 60 人），高动机组有 9 人（共 60 人）。被试在有答案的情况下不全部答对主要原因在于试题有一定的难度，对于一些被试来说全答对很明显不是个人能力所为，这样偷看一两道题而非全答对的就能给被试预留了辩解的空间。即为自我欺骗留下模糊的空间。从分析的结果也发现，全答对的人中没有一个是无意识自我欺骗的，全部都是有意识的自我欺骗者，他们完全觉察到自己的成绩源于作弊。至于承认部分偷看还是完全承认还留有自我欺骗的模糊空间，但是偷看了答案的事实已很少有模糊的空间了。

.....

审稿人 2 意见：

意见 1：研究通过改进 Chance 等（2011）的实验范式，验证了是否存在两种形式的自我欺骗，并进而探讨了两种自我欺骗的影响因素。研究通过三个行为学实验，为存在两种形式的自我欺骗提供了实证证据，具有一定的验证性价值。但同时文章仍存在一些问題，建议作者进一步完善：

回应：谢谢审稿专家对本研究价值的肯定！也感谢审稿专家对本研究提出的宝贵意见和建议！

意见 2：引言部分缺少近 5-10 年的实证性研究进展，文献综述存在断层；

回应：谢谢审稿专家意见和建议！我们在原文中添加了相关的最新文献。

意见 3：问题提出和开展这项研究的理由过于简略，强调了前人的研究和本研究要做什么，但是对于为什么要开展这项研究却一句话带过了，建议加强问题提出部分的逻辑推导和阐述；

回应：谢谢审稿专家意见和建议！按照专家建议，已在问题提出部分更详细地说明我们的研究逻辑。详见问题提出部分。

意见 4：同样假设部分，对于本研究为什么会提出“无意识自我欺骗”也缺少必要的说明，本研究所假设的自我欺骗两个诱发条件实际上是前人的研究发现，也建议作者进一步说明选取这两个诱发条件的理由；

回应：谢谢审稿专家意见和建议！按照专家建议，已在问题提出部分对为什么提出无意

识自我欺骗和有意识自我欺骗做了更加详细的说明。在实验二和实验三开始前也对为什么选取两个诱发条件进行研究做了补充说明。详见正文标红处。

意见 5: 讨论部分主要就是围绕 Chance 等 (2011) 这一个研究展开的, 除此之外就是对于结果和引言的重复叙述, 对于本研究创新性和理论意义的阐述尚不充分。

回应: 谢谢审稿专家意见和建议! 根据专家的建议, 已在讨论部分增加本研究的理论意义和创新性的论述, 并对讨论部分进行了重新的讨论。详见讨论部分标红处。

意见 6: 最后, 还有一些小问题建议作者加以注意: 摘要部分缺少实验设计、方法的简要介绍; 引言部分的连接词混乱, 特别是“第二”后面又分为“第一、第二”; 实验一的结果部分应避免表述上的循环论证; 进一步规范参考文献。

回应: 谢谢审稿专家意见和建议! 已对摘要、引言和参考文献进行规范。同时, 对文中的用词和论证进行了修正。

第二轮

审稿人 1 意见:

意见 1: 对于第一稿作者的解释和修改, 有如下反馈意见 (未提及的基本同意作者的解释): 关于第二点: 作为一个完整的实验, 实验二和实验三应该设立对照组, 鉴于实验条件接近, 用实验一对照组的成绩作为实验二和实验三对比条件也可以接受;

回应: 感谢审稿专家对我们第一次回应的认可!

意见 2: 关于正立/倒立和认知模糊性的关系, 鉴于这是该研究一个非常重要的变量, 还是谨慎处理为好, 不建议以本研究的结果作为反证的证据, 建议在指明“认知模糊性”操作定义的前提下, 通过适当的预备实验解决此问题;

回应: 感谢审稿专家给出的具有操作意义的建议! 按照审稿专家的建议, 我们重新检测了答案正倒立与认知模糊性大小的关系。我们选取了 60 名被试随机分到正立和倒立组中, 按照实验流程被试完成测试和领取报酬后, 提示被试在门口需填写一份无记名问卷调查, 总共有两个题目。1、你是否注意到试题下方的答案? (是或否)。2、下方的答案提示对你解题起到的帮助程度是多少? (完全没有帮助 0, 有一点点帮助 1, 有一定帮助 2, 有较大帮助 3, 有很大帮助 4, 非常有帮助 5)。被试在无监督的情况下填写问卷, 填写完后自己将问卷放到指定的盒子中。

60 个被试全部完成了试题和问卷填写。实验结果显示, 答案正立组和倒立组的被试在测试成绩上差异不显著($t(58)=0.31, P=0.75; M$ 正立=7.36, SD 正立=1.63; M 倒立=7.50, SD 倒立=1.68)。这个结果和实验一的控制组比较可以发现, 两组被试都偷看了答案。问卷调查的结果显示, 60 个被试都注意到了下方的答案; 而在答案对被试起到帮助作用回答的结果显示, 两组被试都不太愿意承认答案对自己的帮助作用, 即承认偷看了答案, 绝大多数被试的回答都是在 1 以下。但是相比较而言, 答案正立组更不愿意承认答案的帮助作用($t(58)=2.98, P<0.01; M$ 正立=0.47, SD 正立=0.63; M 倒立=0.97, SD 倒立=0.67)。这个结果说明答案倒立组被试在偷看答案的认知意识上更加清醒, 更多地承认了答案对其起到的帮助作用。根据认知模糊性的定义, 认知模糊性是指被试对作弊行为的觉察程度。因此, 答案倒立相对答案正立组被试认知模糊性更小。

意见 3: 同理, 对动机的操作是否有效也建议作者提供本研究以外的证据 (是否有类似

的组间研究用该法操纵动机？)不认可在被试间给 0.5 元和给 1.0 元是有效的动机诱发条件，如果事先规定一个基线暗示给钱数量的范围可能会是更好的操作。

回应：谢谢审稿专家的意见！通过金钱来诱发个体的动机是较为常见的方式。从自我欺骗研究领域来看，Chance 等人的研究（2011）和 Mijović-Prelec & Prelec (2010) 都通过金钱诱发过被试的自我欺骗动机。虽然他们的研究中没有组间操作，但是 Chance 等人的研究（2011）用实验 2 和实验 3 中分别答对一道题 1 美元和答对一道题 0.5 美元奖励。只是在这个研究中并没有关注动机大小这个因素，但从实验的结果来看 1 美元也确实让有答案的实验组偷看了更多的答案，诱发了更强的欺骗动机。而从我们的实验结果来看这种操作也是成功的，高报酬组激发了高的欺骗动机，更多的被试做出了偷看答案的行为($\chi^2(1)=9.62, p=0.002, w=0.28$)。在本研究的实验 3 中，虽然没有单独再做预实验确认动机高低操纵的有效性，但是在同一个实验中通过其他数据指标已确认自变量操纵是有效的。这种做法也常用其他研究中。因此，我们认为本研究实验 3 动机强度高低的操作是合理的。另外，关于提示发放报酬数范围的问题，我们也在这次增补的预实验中也对个别被试进行了调查。调查发现，虽然只告知被试答对一道题获得 0.5 元报酬，但是参加实验的被试都清楚地知道自己最后可能会获得报酬的数额范围。因为被试只需要进行简单的乘法即可得知。我们之所以在实验中，没有给钱数范围主要是基于两点考虑。第一，被试能清楚的了解；第二，为了表述更为准确，突出实验的目的，让被试认真回答每一道题。当然，如果事先考虑周全，两者都同时告知被试，那就更好。

意见 4：关于第五点：Chance 等人的实验是否也采用集体分组测试的形式？

回应：Chance 等人的研究（2011）和 Ren 等人的研究(2018)也是采用集体测试。

意见 5：按照作者的说法，在试卷二和试卷三时被试浏览题目是为了使其明确“这是一份没有答案的试卷”，但实际上，考虑到每张试卷的题目只有 10 道，30 秒的时间也足够做“前瞻性评估了”，这可能是一个重要的变量的混淆。

回应：感谢审稿专家的意见！不过我们认为这里应该不存在变量混淆问题。Ren 等人的研究（2018）中也是给被试 30 秒的预测时间。因为 30 秒的时间被试无法做完十道题，这个时间仅够给被试明确这份试卷的题目数量，有无答案以及做出成绩预估。而且指导语也强调说明先预估成绩，再做题目。同时，我们的实验虽然是集体测试，但是每五六个被试都有一位主试监督。因此，不存在被试先做题目，后预估的可能。

意见 6：除此以外，纵观当前的修改稿，尚有以下问题同作者商榷：1.前言中有大量实验证据的堆砌，但未形成一个和本研究相关的中心线，叙述尚欠逻辑，建议作更为精简的，重点突出的叙述；讨论部分也可作适量的精简；

回应：谢谢审稿专家的指出！已按照审稿专家的意见对前言部分和讨论部分进行精简，均控制在 3500 字内。另外，本研究将关于自我欺骗的心理学实验研究找出，并将其概况为三个方面。具体修改的地方请审稿专家查阅前言和讨论部分。

意见 7：2.虽然是个体水平的分析，但能否对“有意识自我欺骗”和“无意识自我欺骗”的卷入程度提出一个直接的量化指标？比较人数似有循环论证之嫌。

回应：谢谢审稿专家的意见！个体水平的区分确实是我们研究的一个小进步。但是目前确实还没有找到自我欺骗卷入程度的量化指标。感谢审稿专家给我们指出下一步研究努力的方向！在未来的研究中，我们希望能够更加精准地解决这个问题。

意见 8：其它问题： p.1.(该页码为正文中的页码，不包括修改说明，下同)：“3 个实验

采用新的前瞻性实验范式分别探讨了无意识和有意识自我欺骗的存在可能性”，在这里，“前瞻性范式”似乎是作者自己定义的，这样定义的概念不适合出现在摘要里；而“新的前瞻性范式”则更加语无所指；

回应：感谢审稿专家的指出！前瞻性范式是我们直译过来。Chance 等人研究（2001）原文表述是 forward-looking paradigm。新的前瞻性范式则是指对其进行改进后的范式。我们在摘要中，已按照专家的意见修改过来。

意见 9：p.1.“而实验 3 的结果发现，动机强度仅影响有意识的自我欺骗...”实际上，文章并未对“有意识的自我欺骗”和“无意识的自我欺骗”提出一个直接的量化指标，所比较的只是人数的差异，这样的比较用“动机强度影响...”似乎不合适；

回应：感谢审稿专家的指出！我们以按审稿专家的要求修改摘要部分，进行更为严格和规范的表述。

意见 10：p.4. Chance 等人所采用的“前瞻性范式”是否和被试的元认知能力或个体差异有关？本文沿用了这一范式，如果能将这一能力作协变量处理可能会使结果更加精确；

回应：感谢审稿专家的建议！确实不同被试的元认知能力、甚至智力水平和知识水平等存在个体差异，但是对照组和实验组的被试是随机分配的。因此，实验组和控制组的被试可以假定为等组。当然，如果能够将这些方面因素考虑在内，再做协变量分析，那么结果会更加可靠。而就目前已有的数据来说，无法做协变量处理确实是一个遗憾。这主要是因为我们参考了 Chance 等人研究（2011），他们在研究中也并没有做协变量处理。当然，在以后的研究中，我们会将审稿专家的意见考虑在内，做到更加严谨。但是，目前的结果基于实验组和控制组的设计，我们认为仍然有效。

意见 11：p.5. 从意向性/非意向性的区分转移到有意识/无意识自我欺骗尚不够清晰，是仅仅从哲学-心理学叙述上的差异还是有更深层次的含义？

回应：谢谢审稿专家提出的问题！关于意向性和非意向性的区分转移到有无意识自我欺骗讨论。主要是因为上一次审稿的时候，审稿专家指出本研究问题提出不够明确。我们重新回顾和梳理了我们选题由来后发现，我们之所以会提出这样的研究问题，是因为我们从哲学上对自我欺骗的意向性探讨中得到了启发。

意见 12：p.7. 实验材料部分应明确指出预实验的因变量。此外，无显著差异不能认为“难度是等价”的；

回应：谢谢审稿专家的指出！预实验的因变量是试卷得分。确实，正如审稿专家指出的那样，无显著差异不能说是等价的。然而，正如我们在第一次答复专家的意见中所说明的，三份试卷是否真的等价并不重要，重要的是被试不会轻易地怀疑试卷是不等价的。这可能是因为在正文中表述的不够明确。关于这点，我们已在原文中做更为准确的表述。

意见 13：p.13. “即表明实验 2 中无论认知模糊程度的大小，被试对卷 2 的预估均存在有意识与无意识两种形式的自我欺骗。”在这里，“存在两种欺骗”的说法值得商榷。一种类比是形状常性中，被试所绘制的椭圆总位于视网膜映像和正圆之间，但是并不能认为存在两种不同的常性机制；

回应：谢谢审稿专家的意见！首先，非常抱歉！可能是我们没有理解审稿专家关于形状常性的类比。我们认为两者之间还是非常不同的。我们这里主要指的是一组人中，可能存在两种类型的人；而不是一个事物从不同的角度来看。我们做出这个判断是基于前面提出了这

样的假设：如果实验组被试试卷 3 的预估成绩与试卷 2 没有答案的真实测试成绩接近，那么说明实验组被试先前在试卷 2 的成绩预估时是有意识的自我欺骗；因为被试在试卷 3 的预估时改变了对自己真实能力的看法。如果实验组被试试卷 3 的预估成绩仍和试卷 1 有答案的测试成绩接近，那么说明被试在预估试卷 2 的成绩时发生的是无意识自我欺骗；因为在知道自己真实水平的情况下，被试仍然坚持将偷看答案后的成绩当作自己的能力。如果实验组被试对试卷 3 的预估成绩介于试卷 1（偷看答案）的测试成绩和试卷 2（没有答案）的测试成绩之间，那么说明实验组被试在试卷 2 的预估中不仅存无意识自我欺骗，而且还存在有意识自我欺骗；因为有意识的自我欺骗者在知道真实成绩的时候，转而以真实成绩（试卷 2）为参考进行了成绩预估；这样导致试卷 3 的预估成绩远远低于卷 1 有答案的测试成绩；但是因为在无意识自我欺骗者仍然保持欺骗不变，因而试卷 3 的预估成绩又会远远高于卷 2 没有答案的测试成绩。

意见 14： p.14. “这个结论与前人发现模糊性会影响自我欺骗行为的结论是一致的 (Sloman et al., 2010)。”实际上，从表 4 中看，这个结果和前人的研究还是有矛盾的，答案倒立（作者认为模糊性小）的自我欺骗人数（30）和答案正立（作者认为模糊性大）的自我欺骗人数（31），并没有什么差异；

回应：感谢审稿专家的指出！我们这里表述确实不够严谨。事实上，这里指的是“这个结论与前人发现模糊性会影响无意识自我欺骗行为的结论是一致的(Sloman et al., 2010)”。虽然 Sloman et al.的研究(2010)没有特别指出其研究的是无意识的自我欺骗。但是根据我们关于无意识和有意识自我欺骗的区分标准，结合 Sloman et al.的研究(2010)中个体自我报告没有意识到自己欺骗自己的行为。我们认为 Sloman et al.的研究(2010)中虽然一直提的只是自我欺骗，但我们认为其应为无意识的自我欺骗。因此，这也说明我们提出有意识和无意识自我欺骗必要，将两种自我欺骗统称为自我欺骗确实存在问题。

意见 15： p.19, p.21. “本研究的实验 2 发现无意识自我欺骗随着认知模糊性的增加而增加，而有意识的自我欺骗随着认知模糊性的增加而减少”，这一叙述似乎和该实验的结果恰恰相反：在表 4 中，无意识自我欺骗的人数在倒立（模糊性小）时是 12 人，而在正立（模糊性大）时是 3 人，有意识自我欺骗的人数在倒立（模糊性小）时是 18 人，而在正立（模糊性大）时是 28 人...? 这可能意味着正立/倒立和认知模糊性的关系有相反的解释。

回应：谢谢审稿专家认真仔细！这是一处笔误。投稿时我们以“认知模糊性大”和“认知模糊性小”作为表中的条件栏。第一次审稿的时候，审稿专家建议我们还是用回原来的条件“答案正立”和“答案倒立”来做表中的条件栏。结果，修改的时候不小心把表 4 的条件写错，而事实应与表 3 的条件栏是一致。现已在表 4 中恢复过来。正立和倒立和认知模糊性大小的关系并没有相反的解释，表 4 只是修改时的笔误。最后，再次感谢审稿专家对本研究提出的意见和建议！

审稿人 2 意见：

意见 1：作者已针对上一轮提出的问题做出了较好的回答，但存在一些小问题：上一轮意见的最后已经提示作者认真规范参考文献列表，但是目前仍存在很多问题（比如期刊名称、期卷号页码等），请作者以对待实验一样严谨的态度重新检查公开出版物的细节。

回应：谢谢审稿专家批评指出！已对参考文献进行再规范和修改，详细内容请查阅参考文献部分。最后，再次感谢审稿专家对本研究的认可！

第三轮

审稿人 1 意见:

意见 1: 纵观全文, 文章的叙述是清晰和详尽的, 特别是在实验程序的叙述上。同意作者对上一稿建议中 1、4、5、6、8、9、11、12、15 所做的解释和修改, 部分同意作者对上一稿中 2、10、13、14 所做的解释和修改, 其他诸问题仍有可商榷之处。

回应: 感谢审稿专家对以上几点回应的认可! 我们下面将对还有待商榷的四处意见进行认真修改。

意见 2: 总体看, 目前存在的几个问题是: 1. 实验三用组间设计, 用给钱的多少操纵被试的动机, 可能是无效的操作, 因为没有被试间的比较 (除非被试“知道”给对方的钱数)。

回应: 谢谢审稿专家的指出! 首先, 非常抱歉! 可能是我们没有理解专家的意见。因为前两次审稿的过程中专家也就这一问题提出了意见, 我们也做了相应的回复。就该问题, 我们主要是没有明白为什么“用给钱的多少操纵动机是无效的”以及“为什么一定要让组间被试知道对方的钱数”。还请专家能够进一步解释说明以帮助我们进一步理解问题所在。如果实验三确实存在问题, 我们愿意补做实验或者删除实验三。

目前, 就我们现在的理解, 先做如下回复: 1. 实验三采用组间设计的原因。我们主要是考虑到实验条件之间的相互干扰问题。自我欺骗在进入无意识状态前主要受意识阈限影响。人的意识阈限是容易受到干扰的, 稍微一个提示阈限就提高了, 被试就无法进入无意识自我欺骗状态。例如, 被试在进行接受 0.5 元每道题的实验时, 本可以进入无意识自我欺骗状态。但是如果他先接受了是 1 元每道题的实验进入的是有意识的自我欺骗状态, 那么这种有意识的自我欺骗就会对后面接受 0.5 元每道题的实验起到干扰作用 (提示他确实偷看过答案), 使其进入无意识自我欺骗的意识阈限提高, 而不能进入无意识自我欺骗的状态。这就严重影响到实验结果的有效性。基于以上考虑我们采用了组间的实验设计。2. 关于金钱操纵动机的有效性问题。从我们的实验结果来看这种操作也是成功的, 高报酬组激发了高的欺骗动机, 更多的被试做出了偷看答案的行为 ($\chi^2(1)=9.62, p=0.002, w=0.28$)。在本研究的实验 3 中, 虽然没有单独再做预实验确认动机高低操纵的有效性, 但是在实验中通过其他指标已确认自变量操纵是有效的。这种做法也常见于其他研究中。因此, 我们认为本研究实验 3 动机强度高低的操作是有效的。3. 关于没有被试间比较的问题。实验三被试的确不知道给对方的钱数。这主要是因为, 我们当时认为组间设计没有必要让被试知道还有另外一组或多组被试在接受其他条件的实验。而且, 我们认为就本研究实验三来说, 让被试知道还有一组人在接受更少钱或更多钱的实验并不合理。因为这样至少会给被试带来自豪感、愉悦感或自卑感、愤怒感。无论是哪一种情绪对本实验来说都是额外的无关干扰。

意见 3: 2. 是否有必要区分“无意识自我欺骗者”和“有意识自我欺骗者”? 对于这个区分之后并没有新的实验操作, 这是否有重复论证之嫌?

回应: 谢谢审稿专家的意见! 有关无意识和有意识自我欺骗区分必要性, 我们在正文的讨论部分有详述 (详见 5.2)。我们想要补充的是: 第一, 现有自我欺骗的概念里确实包含着两种类型。特别是前人研究的操作性定义中确实含有两种类型的自我欺骗。第二, 关于自我欺骗的意向性问题早有争议。第三, 如果将二者混为一谈确实会影响自我欺骗机制的探明 (Trivers, 2011)。

关于“重复论证之嫌”我们的回复是: 首先, 应该来说本研究主要的贡献确实是只在于区分出了两种形式的自我欺骗, 确实没有进一步探明两种形式自我欺骗的内在机制。但是我们正在进行的后续研究确实发现了两种自我欺骗的机制不同: 无意识自我欺骗, 内心没有冲突,

实验中自欺反应时与正确反应时一致；而有意识自我欺骗，实验中自欺反应时与错误反应时相当。电生理指标也显示无意识自我欺骗者 P2 的波幅更大。前人研究发现，当视觉的词汇输入和内在记忆一致的时候，P2 的波幅更大(Evans & Federmeier, 2007)。由于本研究中的范式中无法使用反应时和电生理等指标，确实没有进一步的实验操作以探明自我欺骗的内在机制。然而，本研究中探讨了自我欺骗的主要影响因素（动机强度和认知模糊性）对两种形式自我欺骗的影响。这并非毫无价值，因为它们至少从外在影响机制的不同间接证明了两种形式自我欺骗区分的必要性。

意见 4: 3.用一些文字来阐述实验一与实验二，以及实验二与实验三的逻辑联系是必要的，但目前的阐述过于罗嗦，建议作者精简语言。

回应: 谢谢审稿专家！已在文中进行精简。

意见 5: 4. 建议作者在删除被试的时候，除了客观的指标，应该参考被试“自评”的结果。

回应: 谢谢审稿专家的意见！在被试删除的时候，的确最好是能够同时考虑主观指标和客观指标。但是我们在预实验的时候，发现主观的自评指标并不可靠。因为没有一个人当面承认自己是偷看了答案的。毕竟偷看了答案意味着带来更高的报酬。即使大量数据中会有个别被试报告偷看了答案，但是大多数被试是不会承认的。这样该指标在删选被试中所起到的作用并不大。因此，我们才采用多个客观的指标确认实验组的被试是否偷看了答案。

意见 6: 5.建议重新审查语言阐述，目前有一些赘语和缺失字之处。其它一些细节问题：“摘要”和“结论”中，“3 个实验”中，阿拉伯数字“3”应该为汉字的“三”。摘要中，“自我欺骗存在无意识的和有意识两种形式的自我欺骗...”，此处为赘语，可改为“自我欺骗存在无意识和有意识两种形式...”或“存在无意识的和有意识两种形式的自我欺骗...”。摘要中，“实验 2 的结果发现，随着认知模糊程度的降低，无意识自我欺骗明显减少，而有意识自我欺骗显著增加；而实验 3 的结果发现，随着动机强度的增加有意识自我欺骗人数显著增多，无意识的自我欺骗并没有显著变化。”此处，对两个实验的叙述要保持统一，如是用“自我欺骗显著增加”还是用“自我欺骗人数显著增多”。连接词“而”可去掉。问题提出中，“在冷水中忍耐的时间长，心脏会更好人会 longer”，“心脏会更好”后应增加“，”。建议文章标注英语的地方首字母不必大写。实验二结果部分“其中认知模糊程度小组有 4 人，认知模糊程度大组有 3 人属于此情况。”此处，应增加“的”，即“其中认知模糊程度小的组有 4 人，认知模糊程度大的组有 3 人”

回应: 感谢审稿专家严谨细致的指出！在原文中已按照审稿专家的意见修改，并用黄色标记。

意见 7: 实验二结果部分“卷 3 的预估成绩与卷 1 的作弊成绩接近还是与卷 2 的实际成绩接近来判断被试是无意识自我欺骗者还是有意识自我欺骗者。”此处，“接近”缺乏操作性定义。

回应: 谢谢审稿专家的意见！此处确实没有说明“接近”的操作定义。我们主要是考虑到实验一已经交代了，此处以免赘述。这里的接近指的是卷 3 的预估成绩与卷 1 和卷 2 的测试成绩比较，在分数上与谁的距离更近。