

《心理学报》审稿意见与作者回应

题目：三人问题解决中的惯例：测量及合作水平的影响

作者：张梅，辛自强，林崇德

第一轮

审稿人 1 意见：

总意见：文章考察合作问题解决中的惯例现象。作者应用加 10 游戏为实验任务，通过两个实验，分别对加 10 游戏中惯例现象的描述和测量及不同合作水平下惯例化程度的差异进行考察。结果发现，加 10 游戏中的惯例现象既可从行为、认知和情感等三个大类、七个方面指标进行描述和测量；也可以从个体与群体两个层面进行快速判断并进行惯例化程度的描述。同时，合作情景会调节惯例程度的高低。该研究构建了惯例及惯例化程度的描述指标体系，并发现惯例是合作不良情境下的最优选择，为未来合作问题解决中的惯例研究提供了思路。

文章的研究逻辑清晰，文字表达流畅，并且实验设计和统计分析合理。结果确凿。所得结果有理论意义和应用价值。

但是纵观全文，还有些细节需要进一步完善。

总回应：感谢审稿专家对本研究的积极评价，更感谢所提出的宝贵意见。我们已经按照要求逐一修改，并整体做了加工。下面分别对有关意见说明修改情况。

意见 1：“游戏中，被试分别在五个数字（0，2，3，4，10）中选择一个（ D_i ）……”作者在下文中提到“个体可以以不同方式贡献力量”，这里被试选择数字是否表示贡献力量。作者应该明确指出。

回应：遵照专家建议，在“个体可以以不同方式贡献力量”后增加，“即选择不同的数字”一句话明确指出贡献力量的涵义。

意见 2：加 10 游戏中，随着个体选择的数字不同，得到的收益也不同。作者可以更加清晰的描述一下个体的贡献如何影响其收益，以方便读者理解这个实验任务。

回应：对此，专门增加了有关获益的描述：（1）在实验 1“2.12”部分增加一段内容进行详细描述：“实验开始前，向被试简要说明实验流程，并重点说明实验的获益规则，即依据被试 28 局游戏的得分区间，给予被试 5~15 元不等的奖励，具体为：得分在区间[-112~196]得 5 元，在[197~504]得 8 元，在[505~812]得 12 元，在[813~1120]得 15 元。”（2）在实验 2“3.2.2”部分描述了获益的规则，即虽然实验时告知被试依据得分进行奖励，但由于实验操纵不同得分无法进行比较，最终均给予 10 元奖励。

意见 3：前言中提到的研究目的与要开展的两项研究有何关系。作者可以更明确地指出来。

回应：按照审稿专家的建议，做了两项改动和补充：（1）对“引言”部分问题提出的逻辑进行了重新表述，文字改动近 3000 字，详见文中标注。实际上，两项研究密切相关，实验 1

是实验 2 的基础,实验 2 是实验 1 以加 10 游戏为基础的惯例现象测量指标体系的具体应用。实验 1 证实了加 10 游戏中惯例现象的存在,修改并完善了惯例的测量指标,形成了完整的惯例及惯例化程度的测量思路。实验 2 在此基础上探讨了合作水平这一关键变量对惯例的影响,它应用了实验 1 提出的测量指标对不同实验处理下的惯例进行了比较。(2)在引言末段研究目的后更明确的说明了两个实验的目的。

意见 4: 电脑给出本局反馈页面(图 3),并要求被试对本局满意度和下一局三人相加所得的和进行预测。”“本局满意度”的评分范围和“预测得分”的预测范围应该指出来。

回应: 依据专家建议,本文在“2.12”部分分别对满意度和预测得分的范围进行了说明。

意见 5: “其累计得分得分 1026.67 分是最低组 305 分(第 14 组)的 3 倍多”删掉一个“得分”

回应: 按照审稿专家的建议,已进行删除,同时核查了文中其他笔误,具体见标注。

意见 6: “被试组每局得分逐渐增加($\beta = 0.12$, $t = 6.09$, $p < 0.001$)”文中的 t 检验结果需要报告自由度。

回应: 遵照审稿专家的提醒,已对相应结果进行了报告,自由度均为 2603。

意见 7: “上述惯例表现在同一实验处理条件下经卡方检验,均较其他策略的惯例更显著, $ps < 0.001$ 。”应该是 ps

回应: 遵照审稿专家的提醒,对相应部分进行了修改,详见标注。

意见 8: “但后两组之间无显著差异 $\chi^2_{(2)} = 0.27$, $p > 0.005$ 。这说明,不配合或有条件配合情况下,被试更容易形成惯例。”这里的 p 值大于 .005,作者需要核查一下。

回应: 经专家提醒,核查后发现数据结果没有问题,只是文字表述有歧义,故将最后一句话调整为“这说明,不合作或有条件合作间惯例化程度无差别,但两组与完全合作组相比,惯例化程度更高”

审稿人 2 意见:

总意见: 本文通过探讨三人问题解决情境下的惯例问题,构建了惯例及惯例化程度的描述指标体系、考察了不同合作水平对惯例化程度的影响。本研究使用“加 10 游戏”来探讨惯例问题是具有创新性的一次尝试,同时互联网技术的运用也使得被试操作更加简便,主试对实验的操控和分析也更加便利。

然而,本文从设计到分析却更多地套用前人的方法,这不仅严重削弱了自身研究的原创性,甚至导致了不同程度上的逻辑错误。另外,本文有关实验方法的描述过于简单,导致许多关键因素没有得到澄清。具体来说,主要有以下几个方面的缺陷:

总回应: 感谢审稿专家对本文创新性的肯定,更感谢专家对本文原创性表述及实验方法阐述上诸多问题的质疑、提醒和建议,这些推动了作者进一步挖掘和思考研究自身的逻辑和价值。本文已经逐一对照相应问题进行了修改,具体修改与回复如下:

意见 1:“加 10 游戏”是一种较为复杂的合作博弈情境，与“TTT 游戏”或是“汉诺塔游戏”的问题解决情境有本质的区别，而并不仅有规则和复杂程度的差异。本文无论是文献综述、实验设计，还是结果讨论都没有突出“加 10 游戏”背后的博弈情境的特征，而只是将其作为“TTT 游戏”较为简洁的版本，这不仅削弱了文章的创新性，还混淆了博弈情境与问题解决情境，导致了下面的一些问题：

回应:依据专家的提醒，研究沿着专家所述的将加 10 游戏视为较为复杂的合作博弈情境的思路重新表述了这一游戏的本质特征及其与其他博弈的区别（详见“引言”第 6 段），并逐一回复了专家提出的 3 个小问题：

意见 (1):一些关键变量没有得到说明和控制。例如，本文并未明确说明被试是根据集体得分还是个体得分来获得奖励。个体是追求集体利益最大化还是个人利益最大化可能会对他们的选择产生关键影响（例如，10, 0,0 和 4,4,2 两种策略的集体获益是一样的，个人获益却有较大差异）；本文并未明确说明被试进行的是静态博弈与动态博弈，即被试是同时选择还是依次选择，如果是依次选择他们在选择前能否看到已经做出选择的队友的答案。这些问题可能在“TTT 游戏”中不存在或是不重要，但在本研究中却会对结果产生重要影响，因此必须得到澄清。

回应:依据专家的提醒，文中一些关键变量的确说明不清晰，本文修改如下：1) 奖励是依据个人得分进行的，这一稿中在“2.12”奖励规则部分进行了说明。由于个人获益是在保证三人相加所得和大于或等于 10 基础上实现的，因此，确切地说，奖励是依据基于集体得分基础上的个人得分实现的。2) 对于动态还是静态的问题，理论上被试是同时做出选择的，属于静态博弈，即使有人先做出了选择，也需要等其他两人均做出选择后才完成一局，给出反馈界面。针对原文表述不清的弊端，本次修改通过对测验过程的详细说明解释了 3 人同时做选择的过程。详见实验 1“2.2”部分的描述。

意见 (2):使用以往研究中惯例化的测量指标时存在问题。例如，“重复的行动序列”是通过衡量被试是否会反复走出连续的某几步来衡量他们是否形成了组块化的思维，从而判断是否出现行为惯例化。而“加 10 游戏”中每局游戏只需一步，并不存在行为序列，只以选择某一数字的频率来衡量这一指标并不妥当；“偶尔的次优性”这一指标的则存在更大的偏差。从集体层面来说，三人的数字之和等于 10 的情况要优于大于或小于 10 的情况，但只要三人的数字之和等于 10，无论具体怎么选择（4,4,2；10,0,0；4,3,3）集体获益都是 110，并不存在优劣之分。而在个人层面，最优解是根据队友的选择来确定的，而且不存在严格优势策略（strictly dominant strategy，即无论队友怎么选择，我采取的这个策略都好于其它任何策略），因此，实验 1 中对最优和次优策略的描述存在前后矛盾和不合理的地方。

回应:依据专家的对两个指标可能存在问题的提醒，现答复如下：1) 重复的行动序列这一指标的确如专家所述适用于一局有几步的任务。但惯例涵义中包含了“行为”，不仅仅指“策略”，因此本指标对应修改为“重复行为”。加 10 游戏中的惯例不仅表现为群体用相同的策略完成任务，还表现为个体选择相同的数字完成任务，即使采用的是不同的测量。2) 偶尔的次优性这一指标本质上反映的是人们的“有限理性”，诚如专家所提醒，用 10,0,0 策略为例说明这一指标并不恰当，本文将其去掉，仅保留采用等于 10 策略的例子。详见“2.2.2”第

二段的标注。

意见(3):利用猜测率分析惯例时 also 存在问题。利用二项分布的原理的确可以证明当选择某个数的次数达到某个值的时候,重复选择这一行为不是偶然,而是刻意为之。但是,“加 10 游戏”并没有明显的正确答案,因此虽然我们能判断被试是有意重复选择某个数字,但却无法说明这样的行为是在合作中形成的(被试因为偏爱而重复选择,或是机械地重复选择的可能性无法排除)。

回应:专家所述的“偏爱”或“机械重复”的问题的确可能存在,但这涉及的是惯例的动机的问题。不论出于何种动机,惯例现象说明的就是表现出来的模式化行为或策略。另外,目前为止这是能想到最为“量化”的测量方法,它相对于前人提出的“速度提高”、“可靠性增加”等指标而言更为直接,推测的成分也更小。因此,本文继续采用这种方式。同时,经专家提醒,这样分析的原理介绍不够详细,前文铺垫也较少,所以本研究在引言部分进行了如下补充:“由于加 10 游戏每局仅能选择一个数字或采用一种策略,可分析所有 28 局游戏中选择某数字或某策略多少局就可视为形成了惯例,而不是偶然猜到本数字或策略?利用二项分布原理,可计算偶然选择某数字或策略的概率,利用概率加法可求得猜得某数字或策略 X 题以上概率超过 0.05(或 0.01)的临界点。”

意见 2:本文有两个研究目标:建立惯例的测量指标体系,考察合作程度对于惯例化程度的影响,它们之间缺乏内在逻辑关系,甚至后者使用的因变量仅仅是“选择次数”,而并未使用文中建立的指标体系。

回应:结合专家 1 提出的第 3 个问题,本文对引言部分问题提出的逻辑进行了重新表述,改动涉及近 3000 字,明确地说明了两个实验的密切联系。实际上,实验 2 对不同合作水平下惯例的分析完全依赖于实验 1 提出的惯例(简洁判定每种条件下均存在惯例的方法)及惯例化程度(个体和群体层面的测量方法)的测量指标。讨论中图 6 所示的“指标体系”是结合以往研究对当前研究结论的进一步提炼,并不仅仅是实验 1 的结论,实验 1 的指标体系同时包含了惯例化程度测量的指标体系。此外,实验 2 因变量虽然主要使用了“选择次数”,但本研究基于它同时分析了惯例化程度在行为和策略上的表现,详见表 8 和表 10,这是对实验 1 提出的惯例化程度测量指标的应用。

意见 3:本文主要探讨的是集体问题解决中的惯例,为什么使用猜测率分析时却分别讨论了个体层面和群体层面的惯例?这里的个体层面与个人情境所代表的个体层面有区别吗?群体层面的策略分类也存在疑问。同时,在使用其它几个指标分析时甚至出现了两个层面的混淆,有的从集体层面分析,有的从个体层面分析,有的将两个层面放在一起分析。另外,个体层面=行为模式,群体层面=认知策略这样的说法也不妥当。无论是个体层面还是群体层面应该都存在策略和行为模式。两个层面的策略应该都包含等于 10、小于 10 和大于 103 种,而个体层面的行为模式包含 0、2、3、4、10 五种,群体层面的行为模式包含 10,0,0; 4,4,3 等 125 种组合。因此,文中许多地方的划分是混乱而不合理的。

回应:感谢审稿专家的提醒,本文的思路是这样的:集体问题解决是惯例产生的背景,而个体和群体的区分是惯例分析的角度,任何团队任务的绩效和行为均可从个体和群体两个

层面分析，惯例也不例外。同一个团队中，个体的惯例化程度可能是不同，这反映了个体内差异；同时，团队间也可能存在整体上的差异，这反应了团队间的差异。所以本文采用了个体和群体两个层面。“个体层面=行为模式，群体层面=认知策略”这个说法是针对加 10 游戏来说的，别的任务中的确如审稿专家所述，个体和群体均包含认知和策略两个层面。但对加 10 游戏来说，个体层面就是选择数字，对于其选择数字背后的认知模式肯定是存在的，但目前的任务无法衡量，未来研究可以加以改进，例如每一局都追问其策略。

同时，从专家的质疑出发，本文对上述做法的逻辑在“引言”部分增加了如下文字表述使这一问题更加明确：“依据惯例的定义，它同时包含行为和策略两个层面，结合加 10 游戏本身特点，惯例个体层面体现为被试对不同数字的选择，这可能存在行为上的惯例；群体层面体现为对策略的使用，这反应了团队间的差异。相同的惯例化形成可能体现了不同的策略。例如，某人一直选择数字 3，可能源于集体基于 4,3,3 策略形成的惯例，也可能源于集体基于 4,4,2 策略形成的惯例。”同时，在“3.3.3”部分第一段也详细说明了为何不采用四个指标，而仅采用两个指标的原因：“惯例化程度也涉及个体与群体、行为和认知两个层次，共四个方面。结合加 10 游戏自身特点……”

意见 4：文中所构建的指标体系的创新意义并没有体现出来。

回应：感谢审稿专家的提醒，本文在“引言”部分重新阐述了本文提出的指标体系的依据，增加 1200 余字的内容，详见第 5 段“其次，惯例测量指标上的局限……”和第 8 段“针对惯例研究中测量指标上的局限……”。这既说明了本文问题提出的思路，又说明了本文对以往研究的超越和创新。

意见 5：对于被试信息的描述不够完整，性别、年龄、专业等信息都没有

回应：感谢专家的提醒，本文已在实验 1 和实验 2 的“被试”部分对这些信息进行了补充。

意见 6：实验流程描述过于简单，进行游戏的三人组是始终固定的还是会更换的？被试之间是否允许交流？是否所有被试都知道收益规则？这些问题也都没在文章中明确体现出来。建议进一步细化有关实验流程的描述。

回应：遵照专家的建议，本文在实验 1“2.1 实验方法”部分重新调整了所有表述（详见标注），并去掉了相应略显模糊的拷屏图片（将其作为附录供专家审阅），改为详尽的文字描述，使其更为清晰。调整后可以很明确的得知：三人小组是固定的，被试间不允许交流，所有被试在实验前均已知晓获益规则。

意见 7：在实验 2 中，个人获益方面，完全配合条件下的被试的最优策略仅比次优策略和最差策略分别多 2 分和 10 分，而有条件配合和不配合条件下这两个数字分别是 17 分&41 分和 20 分&24 分；集体获益方面，完全配合条件下各种策略没有差异。有条件配合下最优策略比次优策略和最差策略分别多 32 分和 118 分，不配合条件下这两个数字分别是 78 分和 86 分。因此，被试在不配合和有条件配合的情况下会更多地重复选择最优策略可能是因为选择其它策略的风险太高，而在完全配合的条件下几种策略差别不大（从某种程度上看甚至

可以说没有差别)。而如果这一因素没有得到控制,就难以说明被试在不同条件的选择是由于合作程度,还是出于对风险的回避。

回应: 专家的分析很有道理,本文“5.3”部分在阐述合作程度对惯例化程度的影响也提到惯例是“危机情境下的最优选择”。不合作情境下,个体是选择坚持现有策略以保证收益或将损失降到最小,还是持续探索可能存在的最优策略,这是一种决策,而任何决策都会有风险。实验 2 本身就是探索外在情境对惯例化程度的影响,风险本就是非合作情境下的必然产物,无法避免。

针对专家的提醒,本文进行了两方面处理:(1) 将专家的分析借鉴到讨论部分,进一步解释为什么非合作情境下惯例化程度更高,详见“5.3”部分第二段:“上述结果一方面源于对风险的规避……”。(2) 增加并修改了“引言”部分关于进行实验 2 的依据,更为清楚地说明实验 2 研究不同合作水平的依据。此处涉及 500 余字改动,详见第 9 段“通过更为简洁的加 10 游戏任务构建出惯例的简洁描述指标后……”。

第二轮

审稿人 1 意见:

作者对所有的问题均作出了清晰仔细的回复,并且对原文做了对应的修正。这次修正的质量很高。文章的研究问题提出合理,研究方法应用得当,结果详实可靠,有一定的理论意义和应用价值。建议发表。

审稿人 2 意见:

总意见: 本文经过修改之后,大部分问题得到了澄清,研究逻辑也更加清晰合理。还有一些细节问题已在原文中标注出来,供作者参考。另外,在格式方面仍有几处字体字号有不统一的现象,希望能再认真核查并修改。最后,两个实验中有关被试整体表现的描述其实并不都与研究目的直接相关,因此建议删除一些不必要的实验结果和图表,既节省篇幅又更能凸显文章主体。

总回应: 感谢审稿专家对本文第一次修改的肯定。本文统一核对并检查了所有文字表述,所做修改均已在文中用红色标出。依据专家本次的具体意见,本文逐一回复如下:

意见 1: 在格式方面仍有几处字体字号有不统一的现象,希望能再认真核查并修改。

回应: 已经仔细核查并对应修改,并考虑到了不同 word 版本兼容性的问题。

意见 2: 两个实验中有关被试整体表现的描述其实并不都与研究目的直接相关,因此建议删除一些不必要的实验结果和图表,既节省篇幅又更能凸显文章主体。

回应: 依据专家建议,删除了实验 1 图 2 的 4 个图表,实验 2 中的图 4 及相关叙述。此外,为将其他有关表现的描述与研究目的紧密相连,修改了部分标题,使所保留的描述作为惯例表现的描述。

意见 3: (文中标注)第 1 局时也有 6 组未形成明确策略,是同样的 6 组吗?如果是的话将是很有意思的现象,如能通过数据解释为何这 6 组始终未能形成明显策略就更好了。

回应：经过核查，6 个被试组第 1 局和第 28 局只有 1 组相同。很遗憾不能进行分析。但为了避免误解，加入“（但它们与第 1 局的 6 组不同）”备注。

意见 4：（文中标注）根据之前的建议，作者在使用猜测率进行整体描述时明确地区分出了个体和群体层面。然而，在这 5 个指标的使用上个体和群体层面仍有混淆之处（似乎 1-3 条是基于个体层面，4-5 条是基于群体层面）

回应：按照专家提醒，在本处添加“5 个指标中，前 3 个是从个体表现层面，后 2 个是从群体策略层面验证和描述惯例表现。”使前后呼应。

意见 5：（文中标注）这里的频率与表 3 中的频率不同，可能是基于不同的算法，应当说明，否则容易混淆。

回应：两个频率的确不同。此处的频率是客观、实际的对 5 个数字的选择频率。而表 3 的频率专指选择某个数字 10 次以上的频率。为避免混淆，表 3 中的“频率”二字前增加“10 次及以上”。相对应，表 4 中“频率”前增加“11 次及以上”几个字。此外，对表 3 和表 4 的标题进行了修改，更明确地指出这是在所选数字和所使用的策略上分别形成的惯例情况。

意见 6：（文中标注）证据不够充分，虽然第 6 组大多数时候选择了大于 10 的策略，但这也可能是他们多次探索失败的结果，而并非是在次优策略上形成了惯例。而且这仅是个案，多数小组还是坚持选择了最优策略。因此，这一指标仍待商榷

回应：文中所述 5 个指标每个均是惯例存在的必要条件但非充分条件，任一单独指标都不能证明惯例的存在。虽然所举例子的情况较少，但只要存在这种情况，就说明有惯例存在的可能。这一指标在以往研究中（Cohen & Bacdayan, 1994; 王建安, 张钢, 2008, 2010）被广泛认可，本文去掉它的理由更不充分。综合考虑，本文保留这一指标。

意见 7：（文中标注）从逻辑上看，通常是先粗略判断，再详细描述，所以结果呈现时两者顺序调换一下是否会好一些？

回应：本部分的逻辑是：先利用被广泛认可的 5 个指标确证惯例的存在，再构建出读者不熟悉的、但更简洁的指标，并说明其有效性。若颠倒过来，将不能很好体现我们创立简单指标的必要性，因此，我们倾向于维持现在的写作顺序。