

《心理科学进展》审稿意见与作者回应

题目：基于元分析的助推效果研究：“认知路径”与“透明性”的二维视角

作者：李燕；陈文进；张书维

第一轮

审稿人 1 意见：

本文针对行为科学干预类旗舰杂志 *Behavioural Public Policy* 上有关助推的 31 篇文章进行了元分析，主要从透明性（透明 vs. 不透明）和认知路径（系统 1 vs. 系统 2）这两个维度对不同助推手段进行划分，探讨透明性、认知路径及其交互作用对助推效应的影响。另外，本文也探究了助推的研究设计，行为特征与行为领域特征如何与助推的透明性和认知路径形成复杂交互，共同影响助推效应。文章采取的“透明性”和“认知路径”二维视角具有新意，但我对文章在理论动机、样本库、编码方法和结果分析这四方面都存在一定的疑惑，希望与作者进行探讨。

意见 1：作者在第七页提到，尽管以往的元分析文章通过依据助推措施所对应的认知偏误和心理特征来识别常见的助推类型，但这种分类方式无法涵盖所有可能的助推干预工具。然而，本文采用的 Hansen 和 Jespersen（2013）的分类框架可以包含实践中常见的助推干预手段。这是一个有价值的观点，但作者没有提供明确的论述以支持他们的观点。具体来说，应该说明哪些助推措施是以往分类方法未能覆盖到的，而本文所采用的方法却能够涵盖。

回应：根据审稿专家的意见，修改稿中，作者对表述不清的地方加强了解释与论证。

修改稿中，作者提供更明确论述来支撑本文提出的观点。一是以 MINDSPACE 框架为例解释为何过往的分类方式无法包括所有的干预措施，并补充了文献依据 (Michie 等 (2011) 综合研究了现有的 19 种行为干预分类框架，发现没有一种分类方式能够涵盖所有的干预手段)。二是说明为何 Hansen 和 Jespersen（2013）二维分类框架为何能够克服这种局限性，囊括所有的干预方式。详见修改稿“2 理论基础与分析框架” 第一段。

意见 2：本文的一个关键结论是“透明性与认知路径之间存在交互作用”。然而，在理论层面上，作者并未提供充分的推导和解释来说明为什么可能存在这种交互作用。在引言部分，作者仅讨论了透明性和认知路径分别如何影响助推效应，却未对两者的交互作用进行说明。在结果部分，作者仅简要描述了透明性与认知路径交互作用的结果，却没有很好得从理论层面解释为什么不透明的系统 1 助推最有效，而不透明的系统 2 助推最无效。

回应：修改稿从理论层面对结论进行了充分解释。首先，研究结论中“透明性与认知路径之间存在交互作用”是本文的一个探索性发现，4.2 章节的主要目的是比较这四类助推（透明型系统 1，不透明型系统 1，透明型系统 2，不透明型系统 2）效果的差异，并不是为了证明这种交互效应的存在，但考虑到数据结果所呈现出的规律可能具备普遍性才形成结论，修改稿中也重新解释了此结论仍有待更多经验证据的支持，因此在理论部分并未推导此交互作用。其次，在修改稿“5.1 研究结论”第二部分，作者进一步从理论层面对这种交互作用的生成原因进行了深入解释。

意见 3：本文的另一个重要结论是“助推发挥作用的认知路径、助推设计的透明性与行为特

征之间呈现出复杂的互动作用”。然而，与 1.2 部分的问题类似，作者并未从理论角度对这些复杂的互动作用进行推导和解释。

回应：再次感谢审稿专家的宝贵意见，修改稿扩充了理论依据，弥补了原文中的欠缺部分。不同认知路径助推适用性的影响因素包含三个方面：公民认知能力、行为偏好及助推行为领域 (de Ridder et al., 2022; Kahneman, 2011)；已有研究也从助推行为领域、公民感知自主权入侵程度两个角度探讨了透明型助推的适用情境 (Marchiori et al., 2017; Sunstein, 2015)。不同认知路径或透明性助推在不同情境中的适用性存在差异，而这种差异会通过认知路径、透明性与异质性影响因素间交互作用表现出来。以上研究表明，不能脱离具体的干预情境来比较不同认知路径或透明性助推效果。因此，分析认知路径、透明性与异质性影响因素可能存在的交互作用，有助于回应学界关于不同认知路径与不同透明性之下助推有效性的争议。详见修改稿“2 理论基础与分析框架”中“2.4 认知路径、透明性与异质性要素的交互作用”

意见 4：*Behavioural Public Policy* 杂志确实是一本行为助推领域的好杂志，但因为该杂志较为年轻，发表文章总量有限，因此最后仅有 31 篇文章纳入本文的元分析。对于一篇针对已有大量研究成果的领域的元分析文章而言，这个样本量显然不算多。此外，本文关注变量间的交互作用对助推效应的影响，而非单个变量对助推效应的影响，因此更大的样本库对结论可靠性的影响尤为关键。在行为科学与公共政策领域，还有一本旗舰杂志叫做 *Behavioral Science & Policy*，Richard Thaler 和 Cass Sunstein 均为该杂志的顾问委员会成员。*Behavioral Science & Policy* 与 *Behavioural Public Policy* 两本杂志的定位相似（关于类似话题、编委会成员均为行为科学领域顶尖学者、杂志学术声誉良好），纳入 *Behavioral Science & Policy* 的文章有助于扩大本文的样本量。

回应：感谢审稿专家意见，修改稿中文章纳入了 *Behavioral Science & Policy* 期刊中所涉及助推研究，进一步扩充了本文样本量，并对样本库的选择理由进行了更充分的说明。

Behavioral Science & Policy 期刊 2015-2022 年间共出版 8 卷 14 期，合计 107 篇文章，其中 39 篇定量研究，涉及行为干预实验研究共 16 篇，剔除 2 篇非助推干预研究与 5 篇文章无法计算效应量后，重新纳入 9 篇助推研究文章，合计 21 个样本量。扩充样本后，修改稿对 *Behavioural Public Policy*(2017-2022)与 *Behavioral Science & Policy*(2015-2022)中 40 篇助推研究的 108 个实证结果进行元分析。扩充样本量后，实证分析数据变化主要包括以下几个方面：

第一，在“4.1 助推效应量估计”中，效应量由 $d = 0.16$ ，变为 $d = 0.21$ ，效应量仍然较小。第二，在“4.2 基于透明性与认知路径的效应量分析”中，原文不透明型系统 1 效果大于透明系统 2，修改稿中这两类助推效果无显著差异。第三，在“4.3 助推干预效果的异质性分析”中的“4.3.1 助推效果异质性影响因素分析”，原文消费领域会影响助推效果，修改稿结果显著消费领域无显著影响。第四，在“4.3 助推干预效果的异质性分析”中的“4.3.2 不同类型助推与异质性因素的交互效应分析”出现新的交互影响变量。

整体来看，增加 *Behavioral Science & Policy* 期刊中所涉及助推研究样本后，修改稿实证部分数据的变化没有对原文结论造成影响，比如修改稿在 4.3.2 交互效应分析中出现更多的交互关系为结论提供了更充足的实证依据，进一步说明助推发挥作用的认知路径、助推设计的透明性与异质性影响因素间呈现出复杂的互动作用。

意见 5：本文需要给出更有说服力的样本库选择理由。例如，如果作者仅选取行为科学与公共政策领域的杂志，而不包含发表在其他类型杂志上的助推文章（如心理学领域的 PS、JPSP、JEPG，经济学领域的 AER、QJE 和 JPE），作者应该提供一些理由解释为什么这么做。此外，

作者还应该明确阐述自己样本库选取的优缺点。

回应：感谢审稿专家意见，修改稿对样本库选择理由进行了更充分的说明。本文聚焦“行为公共政策”领域的助推研究，相比心理学或经济学类期刊中助推干预研究，行为公共政策专业期刊更加关注助推在实际政策领域的效果，以公民在真实政策情境下做出的行为来衡量助推干预效果(大多为实地实验)，可以克服传统实验法存在的“外部效度困境”，研究结果更加能够体现出助推的实践价值。*Behavioural Public Policy*(2017–2022)与*Behavioral Science & Policy*(2015–2022)这两本专业期刊持续聚焦以助推为代表的行为公共政策理论与实践，汇聚高质量研究成果，在国际行为公共政策领域享有极高的学术声誉与学术影响力，本文认为以这两本旗舰期刊中的研究文献来分析助推在实际政策领域的效果更加具有代表性。详见修改稿“引言”倒数第二段。

意见 6：透明性和认知路径是本文的关键概念，因此希望作者在第 15 页将分类编码过程的指导语完整呈现（例如编码者如何定义透明与非透明，如果对透明性进行打分编码），以便帮助作者理解本研究是如何定义和归类透明性和认知路径的。

回应：照审稿人意见，修改稿对样本编码方式重新进行了解释说明。认知路径的分类取决于具体干预措施是否涉及反思性思维的运用，如果干预措施发挥效用以个体深思熟虑为前提，那么可归类为系统 2 助推，反之，则为系统 1 助推。根据助推实施过程中是否以目标群体能充分察觉的方式公开干预信息来区分透明型助推与不透明型助推，如果被试能够感知到干预手段、干预目的、或预期结果(任何一种均可)则为透明型助推，反之为不透明型助推。另外，编码仅对干预方式划分类型，不对透明性程度或者反思性程度进行打分。

修改稿在“4.2 基于透明性与认知路径的效应量分析”第一段中，介绍了此分类标准。

意见 7：关于样本量的编码。作者在第 10 页提到，作者做分析时，将样本量大于 1000 的编码为 1，小于 1000 的编码为 0。请问这一编码标准的理由是什么？这与作者在第 10 页关于样本量的论述好像并没有什么关系：“已有元分析的研究结果表明，当样本量在 250-500 之间时，实验干预效应量最大(Peterson et al., 1985)，当样本量大于 500 时，会降低干预效果(Nisa et al., 2019)。Jachimowicz 等(2019)对默认选项效应量异质性分析结果则显示，当样本量大于 1000 时，样本量对效应量不具有显著影响。”另外样本量究竟指的是什么？是一篇文章所有的被试数量，还是一篇文章中每个实验的平均被试数量，或者是单个实验每个组别中的被试数量？

回应：修改稿在理论部分“2.3 助推效果的差异及其成因”第三段中，重新就样本量的影响作用进行了分析，解释了为何选择 1000 作为编码标准的理由。

随机对照实验可以根据被试数量分为大规模实验(样本量 ≥ 1000)和小规模实验(样本量 < 1000)两类(Leloir et al., 1997)，现有文献在研究样本量(单个实验中控制组与对照组合计被试数量)对效应量的影响作用时，也通常以此作为分类标准来比较不同样本量实验研究之间效应量的差异(Kjaergard et al., 2001)。

另外，修改稿在理论部分和变量解释说明中对样本量进行了明确界定“单个实验中控制组与对照组合计被试人数”。

意见 8：虽然作者的结果分析较为全面（例如分析了 9 个关于助推异质性的指标和透明性及认知路径的交互作用），但结论部分较为零散，没有一个或者几个理论框架将这些零散结果串起来。作者大多使用一些描述性语言，没有基于这些描述性结果从理论角度上进行分析和归纳，这让本文的学术贡献显得有限。

回应：作者认真思考了审稿专家建议，重新撰写了结论部分，修改稿的结论分析以实证结果

为基础，与已有研究发现进行更有条理的对比分析、归纳总结，并从理论层面对研究结果进行了更充分的解释。

在“5.1 研究结论”第一部分，作者基于实验情境与实际政策情境的差异，进一步解释了与过往研究相比，本文研究结论中效应量较小的原因，并增加了两篇 *Nature Communications* 中研究文献来支撑本研究结论的合理性；其次，根据研究中存在助推失效的情况，总结了过往研究中对助推失败现象的理论分析。

在“5.1 研究结论”第二部分，作者主要围绕两个关键点进行分析，一透明性与有效性能够共存的实证依据以及重要意义；二是从理论上解释认知路径与透明性产生交互作用的深层原因（为何系统 1 助推更适合隐秘的干预，公开干预信息为何能强化系统 2 助推效果）

在“5.1 研究结论”第三部分，作者首先介绍了样本量对效应量具有显著影响的理论意义；其次，针对助推元分析对于不同行为领域结论通常出现差异的原因进行了深入分析。

在“5.1 研究结论”第四部分，作者重新分析 9 个关于助推异质性的指标和透明性及认知路径的交互作用，以“2 理论基础与分析框架”中“2.4 认知路径、透明性与异质性要素的交互作用”形成理论依据作为基础，对与理论部分相一致的实证研究结论进行归纳，对不一致的结果或无法涵盖的研究结果，进一步从理论层面进行解释。

详见修改稿“5.1 研究结论”

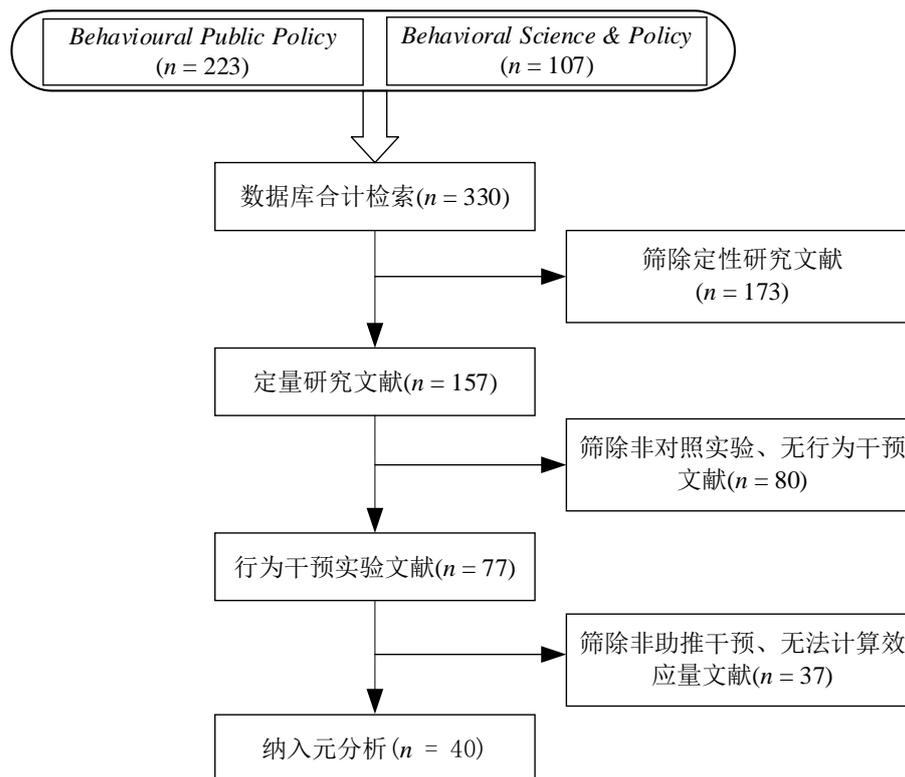
.....

审稿人 2 意见：本研究对 31 篇助推研究进行了元分析，并将多种助推方法归类到认知路径和透明度两个视角，比较了不同认知路径和透明度的助推方法的效果差异以及影响因素。整体上论文逻辑清晰、在文献研究中做了比较多的工作量，行为撰写也很流畅，是一篇不错的期刊论文。以下几点问题，请作者考虑。

意见 1：搜索文献时使用的关键词是什么需要汇报一下，是如何进行筛选的？

回应：修改稿从 *Behavioural Public Policy*(2017-2022) 与 *Behavioral Science & Policy*(2015-2022) 期刊合计发表的 330 篇文献中筛选 40 篇助推研究纳入元分析。文献搜索过程未利用关键词，而是将两本期刊发表的所有文章全部纳入检索范围(合计 330 篇)，再采用阅读全文的方式对文献进行逐一剔除，筛选流程如下图所示。

修改稿在“3.2 数据来源”章节中对此问题进行了说明。



意见 2: 4.3.1 部分中, 样本量的赋值方法是大于 1000 赋值为 1, 小于 1000 赋值为 0, 是根据 87 个研究的样本均值确定的吗? 文中只报告了样本中位数为 2219, 所以我比较好奇的是这个边界值是如何确定的, 这个标准也会直接影响后面的异质性分析结果。

回应: 感谢审稿专家意见, 修改稿在理论部分“2.3 助推效果的差异及其成因”第二段, 重新就样本量的影响作用进行了分析, 解释了为何选择 1000 作为编码标准。

随机对照实验可以根据被试数量分为大规模实验(样本量 ≥ 1000)和小规模实验(样本量 < 1000)两类(Lelorier et al., 1997), 现有文献在研究样本量(单个实验中控制组与对照组合计被试数量)对效应量的影响作用时, 也通常以此作为分类标准来比较不同样本量实验研究之间效应量的差异(Kjaergard et al., 2001)。

意见 3: 作者在对 87 个研究进行编码时, 是否有考虑到透明程度和使用系统 1/2 的程度? 图 3 中在坐标轴中标注的分类是按照所属维度的程度高低排列的吗?

回应: 感谢审稿专家意见, 修改稿对透明性(透明, 不透明)与认知路径(系统 1, 系统 2)的编码方式重新进行了解释说明。

本文研究的重点在于区分助推的类别, 编码仅对干预方式划分类型, 不对透明性程度或者反思性程度进行打分。认知路径的分类取决于具体干预措施是否涉及反思性思维的运用, 如果干预措施发挥效用以个体深思熟虑为前提, 那么可归类为系统 2 助推, 反之, 则为系统 1 助推。根据助推实施过程中是否以目标群体能充分察觉的方式公开干预信息来区分透明型助推与不透明型助推, 如果被试能够感知到干预手段、干预目的、或预期结果(任何一种均可)则为透明型助推, 反之则为不透明型助推。因此, 图 3 标注的分类只区分助推类型, 不涉及所属维度的程度高低。

修改稿在“4.2 基于透明性与认知路径的效应量分析”第一段中, 介绍了此分类标准。

意见 4: 在选取文献时, 仅收录了 *Behavioural Public Policy* 期刊上发表的文章有失偏颇,

为何没有考虑在 EBSCO, ProQuest 等收录全面的数据库中进行检索和选取文献?

回应:感谢审稿专家意见,本文聚焦“行为公共政策”领域的助推研究,相比心理学或经济学类期刊中助推干预研究,行为公共政策专业期刊更加关注助推在实际政策领域的效果,以公民在真实政策情境下做出的行为来衡量助推干预效果(大多为实地试验),可以克服传统实验法存在的“外部效度困境”,研究结果更加能够体现出助推的实践价值。因此,本文仅以行为公共政策专业期刊作为文献检索来源。另外,与 *Behavioural Public Policy* 期刊定位相似,*Behavioral Science & Policy* 期刊同样聚焦以助推为代表的行为公共政策理论与实践,在国际行为公共政策领域享有极高的学术声誉与学术影响力。修改稿进一步纳入了 *Behavioral Science & Policy* 期刊中所涉及助推研究,扩充了元分析样本量。本文认为以这两本旗舰期刊中的研究文献来分析助推在实际政策领域的效果更加具有代表性。

修改稿在“引言倒数第二段”对此问题进行了说明

Behavioral Science & Policy 期刊 2015-2022 年间共出版 8 卷 14 期,合计 107 篇文章,其中 39 篇定量研究,涉及行为干预实验研究共 16 篇,剔除 2 篇非助推干预研究与 5 篇文章无法计算效应量后,重新纳入 9 篇助推研究文章,合计 21 个样本量。扩充样本后,修改稿对 *Behavioural Public Policy*(2017-2022)与 *Behavioral Science & Policy*(2015-2022)中 40 篇助推研究的 108 个实证结果进行元分析。扩充样本量后,实证分析数据变化主要包括以下几个方面:

- (1) 在“4.1 助推效应量估计”中,效应量由 $d = 0.16$, 变为 $d = 0.21$, 效应量仍然较小。
- (2) 在“4.2 基于透明性与认知路径的效应量分析”中,原文不透明型系统 1 效果大于透明系统 2,修改稿中这两类助推效果无显著差异。
- (3) 在“4.3 助推干预效果的异质性分析”中的“4.3.1 助推效果异质性影响因素分析”,原文消费领域会影响助推效果,修改稿结果显著消费领域无显著影响。
- (4) 在“4.3 助推干预效果的异质性分析”中的“4.3.2 不同类型助推与异质性因素的交互效应分析”出现新的交互影响变量。

增加 *Behavioral Science & Policy* 期刊中所涉及助推研究样本后,修改稿实证部分数据的变化没有对原文结论造成影响,比如修改稿在 4.3.2 交互效应分析中出现更多的交互关系,这为本文结论提供了更充足的实证依据,进一步说明助推发挥作用的认知路径、助推设计的透明性与异质性影响因素间呈现出复杂的互动作用。

意见 5:在对不同类别的助推方法异质性分析中,模型 3 为什么是一次性放进去三种助推方法?表 3 中的助推类型为什么没有不透明性系统 2?在模型介绍部分可以再详细介绍一下。

回应:按照审稿专家意见,修改稿在“4.3.1 助推效果异质性影响因素分析”中对此问题进行了说明。下面分别就这两个问题进行解释:

一是模型 3 为什么是一次性放进去三种助推方法。本文设置模型 3 的目的是希望通过回归分析进一步检验不同助推类别间效果的差异,此模型是以不透明系统 2 作为基准组,与放入模型的三种助推类别进行比较。假如透明型系统 1 回归系数显著($\beta = 0.235$),说明透明型系统 1 比不透明型系统 2 助推效应量高 0.235。但本文中模型 3 拟合度较差, F 统计量不显著,无法进行有效分析。

二是表 3 中的助推类型为什么没有不透明型系统 2。由于助推类别全部编码为虚拟变量,在常数项模型中,如果定性指标共分为 M 类,为避免虚拟变量陷阱(多重共线性),回归分析中通常只会放入(M-1)个虚拟变量。因此,在模型设置中,本文以不透明型系统 2 作为基准组,只对另外三种助推类别进行定义。当某种干预措施不属于其中三种助推类别时,必然属于最后一种助推方式(即透明型系统 1 = 不透明型系统 1 = 透明型系统 2 = 0 时,则代表助推方式为不透明型系统 2)。因此,在表 3 中,未对不透明型系统 2 进行定义。

意见 6: 表 3 的变量解释与说明与前一段文字描述重复叙述, 建议删除一个。

回应: 感谢审稿专家意见, 修改稿已对与表 3 文字重复部分进行了删除。

意见 7: P18/P19/P20 结果汇报部分建议进行增补, 如金钱变动对两个系统助推效果的研究结果, 可具体阐述该结果说明了什么, 是否支持了前言部分中假设以及已有的理论等。现在的结果呈现更像是做了很多模型, 然后直接汇报了显著的模型就结束了, 缺少理论和假设的支持。

回应: 作者认真思考了审稿专家建议, 重新撰写了结论部分, 修改稿的结论分析以实证结果为基础, 与已有研究发现进行更有条理的对比分析、归纳总结, 并从理论层面对研究结果进行了更充分的解释。

在“5.1 研究结论”第一部分, 作者基于实验情境与实际政策情境的差异, 进一步解释了与过往研究相比, 本文研究结论中效应量较小的原因, 并增加了两篇 *Nature Communications* 中研究文献来支撑本研究结论的合理性; 其次, 根据研究中存在助推失效的情况, 总结了过往研究中对助推失败现象的理论分析。

在“5.1 研究结论”第二部分, 作者主要围绕两个关键点进行分析, 一透明性与有效性能够共存的实证依据以及重要意义; 二是从理论上解释认知路径与透明性产生交互作用的深层原因(为何系统 1 助推更适合隐秘的干预, 公开干预信息为何能强化系统 2 助推效果)

在“5.1 研究结论”第三部分, 作者首先介绍了样本量对效应量具有显著影响的理论意义; 其次, 针对助推元分析对于不同行为领域结论通常出现差异的原因进行了深入分析。

在“5.1 研究结论”第四部分, 作者重新分析 9 个关于助推异质性的指标和透明性及认知路径的交互作用, 以“2 理论基础与分析框架”中“2.4 认知路径、透明性与异质性要素的交互作用”形成理论依据作为基础, 对与理论部分相一致的实证研究结论进行归纳, 对不一致的结果或无法涵盖的研究结果, 进一步从理论层面进行解释。

详见修改稿“5.1 研究结论”

意见 8: 图 1 中包含了 2017 年的 13 篇文献, 所以选取的是 2017-2022 年还是 2018-2022 年?

回应: 感谢审稿专家指出此问题, *Behavioural Public Policy* 期刊是 2017 年起创刊, 因此本文检索文献是从 2017-2022 年(文献检索年份范围), 但最终纳入元分析的文献全部为 2018-2022 年(纳入元分析文献年份范围)。为便于理解, 修改稿中全部以文献检索年份范围为标准, 更正为 2017-2022 年。

意见 9: 英文摘要需要再润色一下; 参考文献还有个别不规范之处, 再认真检查一下。

回应: 感谢审稿专家意见, 修改稿重新对英文摘要进行了润色修改; 对参考文献不规范之处进行了更正。

第二轮

审稿人 1 意见:

作者对于我上一轮提出的大量问题进行了回答, 并对文章做出了相应的修改, 主要修改如下: 1) 将 *Behavioral Science & Policy* 纳入了元分析样本, 并解释了为什么没有将其他心理学相关杂志纳入分析样本; 2) 对文章的理论动机进行了更为详细的说明; 3) 对透明性与认知路径之间存在交互作用这一探索性发现的结果也做出了更为详细的说明和理论解释。

文章相对初稿有了很大的进步，但文章还存在以下一些细节问题，希望作者予以更好的回应和说明。

意见 1: 作者引用前人文章结论 (Michie et al., 2011) 说明现有 19 种行为干预分类无法涵盖所有干预手段。但为什么 Hansen 和 Jespersen (2013) 二维分类框架能够克服这种局限，作者依然没有做出相应的解释。我在前一轮问到，在本文所覆盖的助推手段中，“哪些助推措施是以往分类方法未能覆盖到的，而本文所采用的方法却能够涵盖？”作者没有对这一问题进行直接的回应。

回应: 感谢审稿专家意见。下文对“认知路径-透明性”二维分类框架与“MINDSPACE”框架的内在分类标准进行对比，结合具体助推措施，进一步解释为何过往的分类框架会存在局限性(无法囊括所有的助推措施)。

Hansen 和 Jespersen (2013) 以助推方式是否依赖直觉式反思和是否透明两个维度构建助推分类的分析框架，认知路径的分类取决于具体干预措施是否涉及反思性思维的运用，透明性是指助推实施过程中是否以目标群体能充分察觉的方式公开干预信息。由于助推措施的认知路径只涉及系统 1 或系统 2，透明性也可以明确划分为透明型或非透明型。因此，任何助推干预措施都能在“认知路径-透明性”的二维分类框架内进行明确区分。“MINDSPACE”框架是以个体行为背后的心理过程作为分类标准，但是其所包含的驱动行为改变的心理因素并不是全面的；此外，“MINDSPACE”框架所提供的心理驱动因素都更加侧重于激发个体的直觉式思维 (Automatic system)，常见的很多教育性助推 (Educative nudge) 也无法在此框架中进行归类 (Michie et al., 2011)。

下面以两种具体的助推措施为例进一步说明“MINDSPACE”框架的局限性。第一种助推措施：Keppeler 等 (2022) 通过激发公民对物品的心理所有权 (psychological ownership) 来引导个体接种新冠疫苗。这种干预措施包含的驱动行为改变的心理因素就超出了“MINDSPACE”框架包含的范围，因此无法在此框架进行归类；但在“认知路径-透明性”二维分类框架这种助推措施可以被归类为不透明型系统 1 助推。第二种助推措施：在医疗领域中常见的一种教育性助推：医院给病人提供包含不同诊疗方案的具体信息 (涉及医疗风险、金钱成本等)，病人可以根据偏好选择更加适合自己的诊疗方案。这是一种典型的教育性助推，同样无法在“MINDSPACE”框架中归类，但在“认知路径-透明性”二维分类框架可以被归类为透明型系统 2 助推。

详见修改稿“2 理论基础与分析框架” 第一段。

意见 2: 在将 *Behavioral Science & Policy* 的 9 篇文章纳入了元分析样本后 (新增文章数量并不多)，一些研究结论发生了变化，其中也包括主要结论“原文不透明型系统 1 效果大于透明系统 2，修改稿中这两类助推效果无显著差异。”由此看出，本文的一些元分析结果还是较为容易受到纳入样本数量的影响。希望作者将这一点作为文章局限进行明确说明。

回应: 感谢审稿专家意见。修改稿在“5.3 研究不足与未来展望” 本文研究局限性中对此问题进行了明确说明。

本研究局限性主要体现在：(1) 元分析样本文献涉及的干预行为大多集中在健康与消费领域，导致我们无法考察其他行为领域助推干预效果异质性。(2) 由于 *Behavioural Public Policy* 与 *Behavioral Science & Policy* 期刊中助推干预实验研究样本大多来自西方国家，以东方国家人群作为被试的研究相对比较缺乏，造成本文无法对助推效果的异质性进行跨文化分析。(3) 受样本文献数量的限制，过多的交互变量会导致模型估计误差偏大，因此本文未对异质性要素之间交互效应进行研究。(4) “认知路径-透明性”二维分类框架下四种不同类别助推的效果会受到纳入样本数量的影响，助推的认知路径与透明性所呈现的交互效应还有

待更多经验证据的支持。

未来研究应更多关注助推在其他行为领域的实际效果,进一步细化目标行为研究助推效果的异质性。已有研究结果表明,默认选项在西方文化背景中的效果要显著高于东方文化背景(赵宁等, 2022),未来可以进一步分析其他助推措施在不同文化背景下的效果差异及其原因。对助推效果的交互效应研究可以更好的揭示异质性根源,未来应设计更加多元的交互效应模型,探索更多异质性要素之间可能存在的交互效应。此外,不同透明性或认知路径助推的干预效果也还需要更大样本量数据的检验。

详见修改稿“5.3 研究不足与未来展望”

意见 3: 在第 32 页,作者提到“研究表明,不透明型系统 1 与透明型系统 2 助推更加高效,透明型系统 1 助推与不透明型系统 2 助推效应量明显偏小。”但这一句话缺乏相应的文献引用,请作者予以补充。

回应: 感谢审稿专家意见。此句话为本文研究所得出的结论,原稿中语言表达存在模糊不清之处,误导了审稿专家,为此,深表歉意。

为避免歧义,修改稿中将其改为“本文研究结果表明,不透明型系统 1 与透明型系统 2 助推更加高效,透明型系统 1 助推与不透明型系统 2 助推效应量明显偏小。”另外,由于尚未有研究从此角度对这四类助推效果进行比较,因此缺少文献直接支撑此结论。但通过理论层面的解释,同样可以论证此结论的合理性。文章在“5.1 研究结论”第二部分中基于认知路径的作用机制分别解释了“不透明型系统 1 与透明型系统 2 助推更加高效”的原因。

第三轮

审稿人 1 意见: 作者对我上一轮提出的细节问题进行了完善,我没有更多的意见了

编委 1 意见: 同意发表

编委 2 意见: 同意发表

主编意见: 2 位审稿人都是该领域真正在做事的专家,所提意见十分专业,到位,作者的回应也令人满意,我觉得修改后的版本应该会引起该领域研究者的关注。同意发表。