

公共政策中助推策略的伦理争议

付春野 吕小康 张雅睿*

【摘要】在公共政策中，对应用助推策略的伦理质疑可归纳为哲学批判和应用担忧两个层面。前者批判助推策略妨碍了公众的行为自主性；后者聚焦于分析助推策略在公共政策应用中的伦理风险，包括不同助推策略引发的不同程度的伦理风险、助推策略应用结果导致的分配公平问题以及应用助推策略的技术途径所引发的道德风险。助推支持者的回应则将助推的伦理问题置身于公共政策自身所在的伦理脉络中，以消解伦理争议的特异性，并力图证明助推自身具有伦理价值和能够对伦理风险提供解决途径。在公共政策中使用助推策略时，通过细分公共政策领域以增进伦理关切，平衡多元群体的不同伦理诉求，并以公共政策中的伦理框架为理论基础，以及立足中国政策实践开展助推伦理研究，可使助推在中国公共政策中的应用更具备伦理正当性。

【关键词】助推行为 公共政策 伦理 自主性

【中图分类号】D63

【文献标识码】A

【文章编号】1674-2486(2022)03-0179-17

一、引言

近年来，以助推（nudging）为代表的行为策略在公共政策中日益占据了重要地位，并发展出“行为公共管理”（behavioral public administration）这一交叉学科领域，有力地促进了公共政策的行为科学转向（代涛涛、陈志霞，2019；吕小康等，2018；吕孝礼等，2020；张书维、李纾，2018；Lodge & Wegrich，

* 付春野，天津师范大学心理学部博士研究生；通讯作者：吕小康，南开大学周恩来政府管理学院教授；张雅睿，南开大学周恩来政府管理学院本科生。感谢两位匿名评审专家对本文的意见与建议。

基金项目：教育部人文社科基金青年项目“医患共同体建设的社会协同机制研究”（19YJC840030），南开大学文科发展基金科学研究类项目“突发公共卫生危机时期的大众风险感知机制与应急心理服务体系研究”（ZB22BZ0109）。

2016; Paul et al., 2019)。作为一个杂糅概念,助推的理论内涵与具体策略仍在不断发展与演变,其个体行为改变层面的要义在于考虑个体非理性决策的基础上,使用轻微而隐蔽的干预策略引导个体行为向预期方向改变,如通过食物摆放位置来促进健康食物的选择倾向、设置默认选项以提高退休储蓄计划的参与率等(泰勒、桑斯坦,2009);而在公共政策的应用层面,对助推策略的基本共识则是它应区别于政府工具箱中的其他工具,如法律法规、激励或惩罚策略等(Grüne-Yanoff & Hertwig, 2016)。

毫无疑问,助推的理想原则是服务公众福祉,并已取得了丰富成果,但其背后的哲学理念与应用过程仍引发了诸多伦理争议(Ewert, 2020)。以往研究较多关注助推能否有效改变行为,但在公众政策中应用助推是否具备道德合理性同样值得重视(Hansen et al., 2016; Wilkinson, 2013)。详尽地理解助推策略所涉及的伦理风险,是确保其得以恰当实施的关键(Croson & Treich, 2014; Ledderer et al., 2020)。若缺少道德合理性这一前提,助推的丰硕成果必将因缺少稳固的道德基石而最终瓦解。

为此,本文拟先概述当前学界对助推策略的伦理担忧,并将之划分为哲学批判与实践反思两个层面进行论述;再归纳助推支持者对这些质疑的回应,以明确助推在伦理层面的可接受性;最后就达成助推应用于公共政策实践应具备的伦理共识提出建设性主张,以期推进学界及公共政策各利益主体的助推伦理意识,进而为助推策略在公共政策中的应用提供伦理考量框架。

二、对助推的两类伦理质疑

对助推的伦理质疑可分为哲学理念批判和实践策略担忧两个层面。其中,哲学层面的伦理质疑通常是一种话语批判,其可协商的空间较小,体现了在哲学层面对助推的坚决反对和摒弃的姿态。另一层面的伦理担忧则更为温和包容,它们避开对助推的哲学思辨,聚焦于助推策略在公共政策中的实践,并提供了一些建设性的主张。以下将对此分别进行归类论述。

(一) 对助推干涉个体行为自主性的哲学批判

第一种伦理担忧直接指向助推的哲学理念而非具体政策,主要由持自由主义哲学立场的西方学者提出。他们普遍支持19世纪英国自由主义哲学家密尔(2005)在《论自由》中提出的“家长制是对个体自由的干涉”的观点,认为家长制的特点是行动的手段具有操纵性或强迫性,而不是关注行动的结果是否有利。在他们看来,助推本身所推崇的是“自由主义家长制”(libertarian paternalism)风格,强调它具有不限制选择以及对决策者自身有利的特征,但这

两点并未改变“家长制”的本质。例如，Hausman 和 Welch（2010）认为，将助推定义为“以一种让决策者做出更好决策的方式来影响行为（泰勒、桑斯坦，2009）”，这本身就具有误导性，将助推的目标指向“更好的决策”回避了助推的伦理风险。在批判者看来，助推其实是通过“塑造”（shaping）或“操纵”（manipulation）决策者的选择而起作用（Vallgarda，2012；Wilkinson，2013），这种操纵必然以干涉决策者的自主性（autonomy）为代价。

批判者还指出，助推虽一直强调实现“更好的决策”，但这种“更好的决策”可能并非来自于决策者本人。如 White（2017）提出，若确如行为经济学家所言，人类普遍存在非理性的认知偏差或认知局限，那么基于理性偏好所形成的公众福祉的基础也同样可疑。诸多关于偏见的研究已表明，即使理性偏好真实存在，也难以准确识别，因此政策制定者的决策实际上总是替代了公众的决策（Riley，2017；White，2017）。更何况，政策制定者的偏好并不总是与个体偏好相一致（Hertwig & Ryal，2020）。例如，政策制定者是否知道公众的具体目标、是否知道目标在公众中的分布情况以及是否比公众更不容易出错等问题都值得反思。为此，政策制定者的目标应该是提高公众实现特定目标的能力，手段的选择则应由公众自己决定，而助推恰恰违反了这一基本要旨（Grüne-Yanoff & Hertwig，2016）。

对高度重视个体自主性的自由主义者而言，公共政策最应努力促成的结果是遵守并提升行为的自主性，这体现了个体对自身的评价和决策的控制力，即自主形成自身的信念、态度和决策。而助推只是实现了对规范的遵从，但未达到规范的内化（Mols et al.，2015）。这不仅妨碍了个体自主性，还剥夺了理性个体自主认识并纠正自身行为偏差的机会。其远期后果是抑制个体的思考或反思能力，妨碍其随着决策学习过程而变得更为理性（Rizzo & Whitman，2019；Vugts et al.，2020）。例如，本想控制体重的人在超市购物时自然会有意识地避开高热量食物的货架通道，但精心设计的助推策略反而减弱了个体对自身决策的控制程度（Dworkin，2019）。

简言之，此类伦理批判将个体理性与自主性置于不可侵犯的哲学高地，并由此立论抨击助推对个体自主权的限制，反对基于助推的公共政策，支持传统的公共政策。在他们看来，公众虽然不一定对传统公共政策感到满意，但仍可清晰地意识到自己的行为受到政策的影响（Hausman & Welch，2010；Weber，2017），仅凭“知情权”这一点就足以认定助推在伦理价值层面上处于劣势。可见，在此视角下助推天生即有“原罪”，助推是对个人主义、自由主义和理性主义传统的多重背叛。这些伦理批判旨在维护西方社会的价值观基石，代表了一种保守主义的公共政策立场。

（二）对实际助推策略的伦理担忧

按关注点的不同，此类伦理担忧可划分为不同助推策略的伦理风险、助推的应用结果以及技术途径的伦理考量这三类。它们之间可能有部分的重叠，但每种类型的主要质疑点并不相同，仍可进行相对独立的阐述。

1. 不同类型助推策略对应不同程度的伦理风险

助推策略的类型划分有多种可行方式。其中，以下两种划分方法主要以助推在公共政策应用中的伦理风险为出发点，考虑不同类型助推策略所对应的伦理风险程度的差异。

Baldwin（2014）提出，可根据助推策略对个体自主性的干涉程度，将助推策略划分为三个级别。第一级助推（first degree nudges）通过提供简单的信息或提醒来改善决策，如香烟盒上“吸烟有害健康”的警示标签；它是在反思层面上进行的，对自主性的干涉最小。第二级助推（second degree nudges）建立在对行为或自主性限制的基础上，以促使个体的决策偏向公共政策所预期的方向，如在默认选项中设置为默认同意器官捐赠。它依靠人类的习惯来推动决策，由于个体的行为或自主性的限制会导致其在非理性层面接受这种助推，因此第二级助推对个体自主性的影响高于第一级助推，但个体在反思后仍可意识到助推的实施并评估其影响。第三级助推（third degree nudges）涉及行为的操纵与动机的塑造，并威胁到个体偏好，如在香烟盒的设计上，用吸烟可能引发的消极健康后果的恐怖图片来唤醒个体对吸烟的恐惧情绪。它阻碍了个体独立思考的能力，且个体较难通过反思而察觉。据此，Ruehle 等（2021）提出，第一级助推可被广泛允许使用；第二级助推更类似于默认的社会规范，且提供给公众选择权，但类似默认选项这种策略也存在“忽略退出”的风险，因此它需谨慎使用；第三级助推存在对行为的操纵且无法通过其他策略抵消这种消极影响，应被严格限制。

另外，Hansen 和 Jespersen（2013）根据卡尼曼（2012）双加工理论中对系统 1 和系统 2 的区分，将助推分为第一类助推（type 1 nudges）和第二类助推（type 2 nudges）。前者通过影响系统 1 的自动思维来改善决策，而不涉及系统 2 的反思思维；后者则以影响系统 2 的反思思维为前提来改善决策。也就是说，第一类助推只影响行为而无须思考，而第二类助推是审慎思考后的结果。在此基础上，Hansen 和 Jespersen（2013）又引入透明性（transparent）的概念。透明性指让公众意识到助推的存在和目标，并促使公众理解助推为什么以及如何影响其决策（Loewenstein et al., 2014）。结合助推层级和透明与否，可将助推交互分为四种类型：无透明性的第一类助推（如默认选项）、有透明性的第一类助推（如用线索提示吸引注意）、无透明性的第二类助推（如改变信息框架）和

有透明性的第二类助推（如提供健康建议）。Hansen 和 Jespersen（2013）认为，有透明性的第一类助推引导自动化的行为，有透明性的第二类助推激发反思性的行为，两者均不存在伦理风险；而无透明性的第一类和第二类助推则会操纵行为，因此存在伦理风险。

这些伦理担忧主要通过对助推类型的理论划分而指出哪些类型的助推应得到伦理许可，但尚未达到可指导伦理实践的成熟水平。另外，助推类型的划分方式也存在一些争议。例如，Baldwin（2014）的第一级助推虽然较少干涉自主性，但是提供警示信息这种干预策略是否可称为“助推”仍有异议。总的来说，这些伦理思考将问题聚焦于具体的助推策略，在提出问题的同时也留下了解决空间，如强调增加透明性可增进公共政策的道德合理性。

2. 应用助推导致的分配公平问题

助推策略得以流行的基本背景是应对传统公共政策的部分失灵，其有效性一直是助推备受青睐的首要原因。但随着助推的应用越来越广泛，其所可能引发的负面后果也开始得到重视，其中一个重要的问题即分配公平（distributive justice）问题。Roberts（2018）提出了“防推”（nudge-proof）的概念，意指公众中特定群体是不受助推影响的；或者说，同一种助推策略不可能对所有群体都同等有效，对某些群体来说它其实具有隐性成本，即实现助推目标的“努力税”（effort tax）更高。例如，通过摆放位置促进健康食物购买的助推策略虽对部分群体有更多的收益，但对贫困人群来讲，食物的价格仍比摆放位置更能影响其购买行为。助推的提出者 Sunstein（2021）也意识到分配公平问题会使助推的效果无法具有普遍性，如默认登记这一助推策略虽会提高退休储蓄计划的参与率，但也可能使低薪工作者的生活状况更糟糕，因为对于他们来说，满足当下的生活需求其实更为迫切。因此，如果助推不能给贫困人口、老年人和残疾人等特定群体带来收益，这种风险就尤其令人担忧：只助推优势群体，却将弱势群体的利益抛置身后，有可能加剧不平等或创造新的不平等，这显然与助推立足于公众福祉的目标背道而驰。

这种伦理考量进一步揭示了助推应用于公共政策时的伦理复杂性。通常而言，公共政策的设计应遵从帕累托优化（pareto improvement）原则，即在没有使任何人境况变坏的前提下，使得至少一个人变得更好。助推往往能满足这一点，但这仅是从政策执行效率的角度论证它的伦理合理性。若将政策的伦理价值置于多元价值的思考框架中，就难有某种单一价值观能够具有支配地位。如此，也无法将政策的伦理合理性完全归因于其执行效率，而需兼顾其他价值诉求。其实，即使满足了帕累托优化，深层的不平等也会违背分配公平（Cohen，2008）。而且，帕累托优化本质上与公平无关（Quong，2010），一旦特定群体无法享受助推所带来的福利，助推的应用就会加剧社会不平等，这就构成了助

推应用于公共政策的最根本挑战。为此，当政府决定采用特定助推策略服务于公众时，需格外重视特定群体无法受益的问题。

3. 助推技术途径中蕴藏的伦理风险

随着网络技术的发展，助推借以施行的技术途径也愈发多样，相关伦理风险也随之产生。在网络时代，基于大数据的助推策略已普遍应用于社会生活。政府或相关组织已愈发便利地使用这一技术，以实现获取个体信息、了解个体偏好、提供定制化的助推（Sætra, 2019）。为此，有研究者提出了“数据助推”（digital nudges）的概念，特指通过数据环境的预先计算来实现塑造公众决策与行为的一系列算法治理策略（Schüll, 2016；Sætra, 2020；Yeung, 2017）。数据助推通常有两种形式：一是微助推（micro-nudges）或隐式助推，如点赞；二是显式助推，如在屏幕上弹出助推信息等（Petrakaki et al., 2021）。

但是，数据助推的流行也催生了新的伦理问题。大数据的信息获取过程具有不可避免的强迫性：一是信息收集通常是隐秘的，其工具和技术对被采集者而言并不透明，因此具有侵犯个人隐私的风险（Gandy & Selena, 2019）；二是公众对大数据的贡献并非总是自愿，且退出大数据往往会给个体带来高昂的经济成本和社会成本（Sætra, 2019）。此外，数据助推通常具有更强的隐蔽性和操纵性，它依靠低透明性、高复杂性和高自动化的系统来实现助推，其实已经把“助推”变成了“强推”（Sætra, 2020）。甚至有研究者将数据助推称为“超级助推”（hypernudge），因为与早期的助推相比，数据助推由于其网络化、持续更新、动态化和普遍性等的特征，实际上已经具有了强制性（Yeung, 2017）。

这一方面的伦理担忧主要来自于实现助推的技术途径。新技术途径将助推的伦理争议延伸到其他领域，它们通常涉及大数据技术的应用与个体隐私保护之间的复杂博弈。而有效的隐私监管本身存在两难：它不仅要使公共和个人监管系统切实透明，还须为相关主体的管理实践保留空间（Cohen, 2013）。助推技术途径本身蕴含的伦理风险，构成了助推伦理考量的外延部分，并将随着助推技术与实现方式的演变而不断更新。

三、助推支持者对伦理质疑的已有回应

对前述质疑，许多助推的支持者已作出了回应，认为将助推应用于公共政策并不存在伦理风险或可通过适当方式解决，并试图论证助推具有可接受的伦理价值（ethical values acceptably）。本部分拟对助推策略伦理交锋的正反两方做一结构性梳理，内容框架如图1所示。

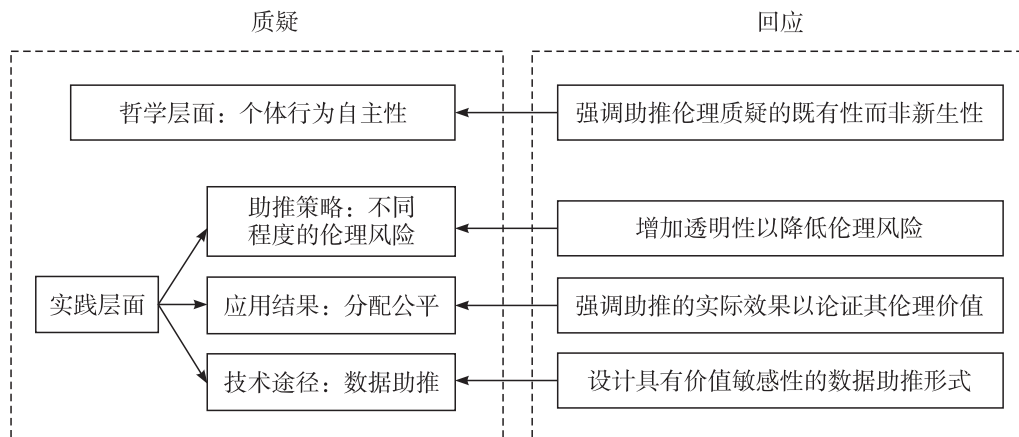


图1 公众政策中助推策略的伦理质疑与回应

资料来源：作者自制。

（一）强调助推伦理质疑的既有性而非新生性

助推的提出者早已预料到助推对个体自主性的干预并进行过辩解。泰勒和桑斯坦（2009，254-268）指出，政府和组织不可避免地在公共政策中承担着“建筑师”的角色，其许多强制性的政策必然影响公众选择。实际上，在法律与政策领域，不管这些规则是否叫做“助推”，这种机制都随处可见，它们几乎和人类的历史一样古老（桑斯坦，2021：3）。因此，公共政策对公众自主性的限制其实古已有之，而非新事。如此一来，相关的伦理争议自然也非助推所独有。由于政策制定者必然要发挥其管理职能，一定程度的“操纵”或“塑造”公众行为是不可避免的，因此对助推的这种伦理质疑其实并无意义。

与传统政策相比，基于自由主义家长制的助推策略更为温和且更具非干涉性。在Osman（2016）看来，这一策略的实质是某种行为受到国家政策的鼓励，国家只是提倡以某种价值观为基础的生活方式，但仍保留公众不这样做的权利。例如，通过改变选择架构来促使环保行为比其他行为更容易实现，但这并未减少个体可采用的选择集（Christian，2017）。同时，“自主性不可干涉”的观点是基于一种极端个人主义的立场，并未考虑到人类社会的公共性。在面临全球变暖等具有紧迫性的全球性问题时，重要的不是说服公众改变价值观（价值观本身很难改变），而是促使公众做出更加理性的行为（Nørnberg et al.，2016）。此时，沿用尊重个体自主性但于事无补的传统公共政策反而应该被认为是不道德的（Cookson，2013）。

这一回应思路的实质是将对公共政策中助推策略的伦理质疑置于公共政策本身具有的伦理问题中，强调两者之间的同质性而非异质性。由于公共政策本身并非纯粹的理性技术，而是事实领域与价值领域的结合体，具有内在的伦理

属性（刘舒杨，2020）。因此，这种回应有其合理性。尽管如此，公共政策本身的伦理争议与助推在公共政策应用中所带来的伦理风险仍是两个不尽相同的问题。如前所述，助推的部分伦理风险来自其创新性的政策目标的实现途径，而公共政策本身的伦理问题则更聚焦于政策目标自身的道德合理性。因此，仍需要从其他角度对前述伦理质疑作出更多的回应。

（二）增加透明性以降低伦理风险

实践层面的一种伦理担忧来自于部分助推策略因缺乏透明性而具有伦理风险。有不少助推支持者认为，若能确保助推的透明性，助推对自主性的威胁就会小很多（Sunstein，2015）。但是，增加透明性可能会削弱助推效果。因此，支持者开始寻找助推可以既有效又透明的证据。例如，有研究通过引入默认项来助推碳减排，并设置了三种不同类型的透明性，包括说明默认项的潜在影响、说明默认项设置的目的是同时说明上述两者，结果发现无论是哪种类型的透明化措施均未增加参与者的心理阻抗，也未降低助推效果（Bruns et al.，2018）。这表明助推可以是透明且有效的。另有网络安全决策研究在实验中对只提供基于助推的干预策略与提供助推和理性信息结合的干预策略，结果发现助推与理性信息相结合的干预策略比单纯的助推更有效。这说明理性信息不仅增加了助推的透明性，还提供了教育功能（Zimmermann & Renaud，2021）。还有研究检验了增加透明性是否会削弱默认参加卫生保健系统内学习活动的干预效果，结果发现默认参加增加了选择参加学习活动的人数，且在明确告知默认选项的设置后，参与者也未改变之前的选择（Dranseika & Piasecki，2020）。另有实验发现，默认选项助推和透明性信息均可提高政策依从性，但在二者结合使用的条件下效果最好（Paunov et al.，2020）。

以上这些研究都在一定程度上说明了提高透明性有助于缓解伦理质疑，但这一方式能否推广至所有助推策略仍有疑问。其实，并非所有助推策略都可同时做到透明且有效。如有研究发现，在应用社会规范的助推策略时，公开社会规范的作用（告知参与者：“请注意，向您提供大多数人的选择是为了影响您的决策。因为研究表明，人们会受到他人的行为和信念的强烈影响。”）会抑制助推效果（Kantorowicz-Reznichenko & Kantorowicz，2021）。因此，透明化虽可降低助推的伦理风险，但其作用范围仍有局限。

（三）强调助推的实际效果以论证其伦理价值

第三种回应立足于助推的实际效果来论证助推具备可接受的伦理价值，这一部分可与助推可能因分配公平问题引发负面结果的质疑相对应。首先，传统政策在特定领域存在失灵的情况，而在这些失灵的领域中，助推提供了较佳的解决方案。如在饮食、运动以及环保等领域，由于公众的行为习惯已根深蒂固

于公众心中，且深嵌于文化风俗和价值理念中，不是信息触发或简单纠正就能够预防的一次性事件，也不是对自身行为进行理性思考后的结果，而只是一种习以为常的行为习惯。此时，理性评估并不能驱动更好的决策（傅鑫媛等，2019；Kelly & Barker, 2016）。像在日常健康行为中，人们通常重视当下的愉悦感（摄入高热量食物），这既具有生理上的必然性，又与各地的饮食文化高度相关，通过理性说服很难达到控制热量摄入的效果。而类似将健康食物摆放在更加便利的位置等的助推策略则可避开这些阻力而有效地促进健康行为。因此，助推改进了公共政策的具体细节，提高了公众的政策遵从性并带来了更有效的结果（Paunov et al., 2020；李燕等，2021）。这些都应被视为助推具有合理性的理由（Halpern, 2015）。

其次，助推具有促进政策公平的潜力。研究发现，社会经济地位较低、领域知识较少和计算能力较差的消费者更容易受到助推的影响，进而能够作出有利于自身或社会的行为决策（Mrkva et al., 2021）。从这一视角来看，助推并未助长不平等，还可削弱不平等。此外，如果助推的目的在于对抗其他形式的操纵，那么采用助推也可视为是合乎道德的（DiSilvestro, 2012）。例如，向公众提供关于酒精危害性的图片助推信息，可解释为是为了抵抗来自酒厂的资金更充足、更具诱惑力的广告对消费者心理与行为的操纵（Wilkinson, 2013）。这种助推有益于抵制工商业领域不合乎伦理道德的助推策略，因此具有伦理正当性（Sunstein, 2015）。

上述内容反映出助推伦理价值的支持者和质疑者之间所看待问题的立场差异，且双方均可拿出证据以支持自身观点。这种辩论虽难达成完全共识，但仍有助于人们更深入地了解助推的优势及其局限。

（四）设计具有价值敏感性的数据助推形式

传统助推和数据助推的区别在于后者所应用的数字环境的动态性、信息化和自动化，导致适用于传统助推伦理困境的解决方法可能在数据助推的背景下效力不足。为此，还需进一步开发适用于数据助推伦理风险的解决途径（Ioannou et al., 2021）。这首先要求政府或相关组织在收集和使用数据来实现助推时，需充分认识到数据自身的公共属性（Prainsack, 2020），确保数据助推符合正确的价值观。例如，近年来“社会公益型人工智能”（artificial intelligence for social good）的理念不断被提及，提倡以人工智能技术解决社会问题（Hager et al., 2017）。依据这一理念，数据助推应遵循自主性（autonomy）、预防伤害（prevention of harm）、公平（fairness）和可说明性（explicability）四种基本价值观（Umbrello & van de Poel, 2021）。这些价值观可扩展为一系列具体的伦理规范，如助推效果应具有可证伪性，并验证其实验室效度和生态效度；确保其他干扰变量不会扭曲助推效果；建立基于用户的情境化决策系统，了解用户特点

并尊重用户忽视或修改助推的权利；提供具有说服力的论据为用户提供解释，确保助推的目标为用户可知；保护用户隐私；等等（Foridi et al. , 2020）。

在这些伦理规范的指导下，数据助推的实践可采用“价值敏感化设计”（value sensitive design）作为原则性的指导方法，即在整个设计过程中体现对个人价值的综合考量（Friedman et al. , 2013），以避免伦理风险。Capasso 和 Umbrello（2022）以亚马逊 Alexa 的医疗保健系统的案例来说明如何使用这一设计来指导数据助推。例如，在保护用户隐私的伦理规范下对用户数据进行匿名化处理，如此可将更抽象的价值观通过伦理规范转化为实际的设计需求，以降低数据助推的伦理风险。

当然，这种设计原则及伦理规范也为数据助推设置了严苛的执行标准，从而增加了技术落地的难度。此外，实践层面的设计需求总是因时因地而不同，因此任何设计案例都不能为其他数据助推策略提供详尽的模板，而只是为实现共同的目标打开了设计上的想象空间（Capasso & Umbrello, 2022）。

四、建立助推伦理共识的可能途径

在遵循恰当伦理指导的前提下，助推可以成为实现公共政策目标的有效工具（Grilli & Curtis, 2021）。但助推的伦理争议仍将伴随公共政策的制定、执行与评估的全过程。这种争议应被视为一种社会协商的过程，视之为合理的担忧而非绝对的反对（Engelen, 2019），从而不断反思公共政策的本质及其实现方式。只有在不断地积极地回应公众及学界的伦理诉求和质疑的过程中，助推策略才能以更得人心的方式取得并巩固其政策实践效果。以下将对缓解助推伦理争议的可能途径及其在中国情境下的应用提出一些方向性的思考与建议。

（一）细分公共政策领域推进伦理关切

前文虽已提及缓解伦理争议的探索性尝试，但其内容主要来自于研究者对助推的伦理思考，尚未涉及“被助推”对象自身如何看待助推等关键内容。根据张书维和李纾（2018）总结的“政府行为－公民体验”双轮模型，行为公共管理应聚焦于政府行为过程和公民体验两大领域。因此，公众对助推的态度理应构成伦理争辩的关键部分。

近年来，研究者已经就公众对助推应用于公共政策中的态度开展研究。已有研究表明，助推应用于公共政策中可能获得较高的公众支持（Evers et al. , 2018；Junghans et al. , 2015；Krisam et al. , 2021），但这一结论尚需结合应用领域因素细加考量。其实，公众对不同领域的助推策略的接受度并不相同。总体而言，公共健康和安全领域的助推最容易被接受（Junghans et al. , 2015；

Sunstein et al. , 2018)。即使是同一种助推策略,其公众接受度也存在领域差异。有研究比较了默认加入在器官捐献、碳排放和退休储蓄三个领域的接受度差异,结果发现在器官捐献和退休储蓄这两个领域,参与者认为默认加入比选择加入更有实际和社会效益,但在碳排放领域则并不认可(Yan & Yates, 2019)。还有实验研究发现,当要求参与者以第三方视角来评判一个虚拟人物进行注册器官捐献登记的自愿程度时,参与者认为默认加入和强制加入同样是非自愿的选择(Lin et al. , 2018)。这一结果体现了器官捐献领域的特异性,也提醒研究者在一些特定领域下公众对助推的态度并不统一。还有研究对比了助推在不同健康领域的接受度差异,结果发现在烟草、酒精和高热量零食的消费中,参与者对烟草消费中的助推策略接受度最高(Reynolds et al. , 2019)。这都表明助推在不同公共政策领域应用中可能面临不同程度的伦理风险。

虽然公共政策的统一目标是改善公众福祉,但从公众自身的视角来讲,公共政策仍可区分为“利己(如健康)”“利他(如捐赠)”和“利众(如环保)”三种类型。有实验研究发现,当为满足基本需求时,公众更倾向于强制性传统政策;但当涉及生活方式和价值观时,公众则更倾向于基于助推的自由家长式干预(Treger, 2021)。这说明公众对不同消费品的伦理立场是影响助推接受度的重要因素。此外,不同领域的公众动机和道德义务也存在明显的差异,如公众虽有保护环境的道德义务,其动机水平却往往较弱(Moser, 2007)。这会导致不同领域的伦理争议在关切形式和关切程度上存在差异。这些内容都应包含在伦理考量的基本框架中,为政策制定者在思考、设计和执行某一项助推策略时提供参照,以增强助推的伦理正当性。

(二) 平衡多元群体的不同伦理诉求

现实社会问题通常无法通过单一治理工具解决,需要不同治理工具的有效组合。而研究者为更鲜明地表明立场,往往会突出自身观点的特异性、单一策略的有效性而忽略了公共问题本身的复杂性。如前述的“助推的结果是助长还是削弱不平等”这一问题,双方均可发现支持自身立场的实证证据。以公共健康领域为例,通过将健康食品放置于“可得性”更强的位置,的确可以促进购买健康食品的行为,但前提是个体本身具有购买健康食品的意愿与支付能力。而针对健康素养较低或经济状况较差的群体来讲,这种助推策略显然就无法发挥作用。当公众本身对助推的行为目标持有抵触态度时,助推甚至还会起到反作用(Dewies et al. , 2021)。实际上,根据默认选项助推效应的元分析结果还发现:默认选项设置的助推效果受到情境领域的影响,其在金钱情境下的助推效果最好、环境情境次之、健康情境下的效果最弱(赵宁等, 2022)。这提示健康助推策略在健康行为的干预策略中只能扮演辅助角色。

为此,应更为客观地看待助推在实际应用中的局限性,也要全方面地理解

群体差异和公众需求。例如，当今世界有很多国家面临低生育率问题，个体层面的助推策略可能仅对部分群体发挥微弱的效果，而国家降低生育成本并提供相应的福利政策才是提高生育意愿更恰当的政策途径。此外，助推也应该加强与理性教育等政策途径的结合。在这方面，助力（boost）可与助推形成良性的互助关系。与助推将重心放置于目标行为不同，助力着重于公众的认知改善和能力培养（Grüne-Yanoff & Hertwig, 2016）。两者的结合可在适用的政策情境下充分发挥效用，共同促进公共政策的合理完善（张书维等，2019）。这也可避免部分群体因知识匮乏等因素难以被助推策略“推动”，进而不能享受助推的收益，并助长群体间的不平等。因此，由政府、权威机构与专业人士联合进行的政策沟通、健康教育、科学普及和素养提升等内容，仍是确保公共政策目标得以实现的关键步骤。此外，正确认识助推在公共政策中的位置及其与理性教育的兼容性关系，也可缓解助推质疑者的伦理质疑（Hanse et al., 2016）。这一立场可承接之前提出的将助推问题的伦理争议置于公共政策自身的伦理争议中，使质疑者和公众更为全面地意识到伦理问题自身的复杂性和历史性，从而避免将早已有之的伦理问题归结于助推策略引发的新问题。

（三）融入公共政策自身的伦理框架

有学者指出，公共政策的本质是规范统一性的，助推基于个体性的研究往往忽视了公共政策的普遍性行为机制（景怀斌，2021）。这一批判虽然并非指向助推在公共政策应用中的伦理问题，但却为缓解助推的伦理争议提供了有益的思路。作为连接价值规范与社会事实的治理工具，几乎任何公共政策都涉及伦理问题。因此，助推的伦理问题应融入公共政策领域现有的伦理框架中加以解决。这不仅可为助推伦理问题的解决提供更宽广的思维框架，也有助于进一步完善公共政策伦理理论。

以助推应用较为广泛的公共卫生领域为例。公共卫生伦理学大量借鉴了伦理学和生物伦理学开发的伦理工具，并创建了不同的理论框架来指导实践。这些理论多数规定需在尊重个体自主性、预防健康风险及促进健康义务之间取得平衡（Baum et al., 2007; Jaffe & Hope, 2010）。例如，为使理论框架具有实际应用性，纳菲尔德生命伦理学理事会（Nuffield Council on Bioethics）致力于在公共卫生伦理问题的讨论中增加社会层面的内容，如在哪些群体或子群体中寻求平等（Lee, 2012）。另有学者（Tronto, 1993; Tronto, 2010）提出关怀伦理理论（Care Ethics Theory），该理论以政策制定或施行者与公众之间的不平等关系为基础，即承认一方比另一方拥有更多权力、资源或知识，而另一方则依赖其获得利益，如父母与子女、医生与患者和教师与学生等（Engster, 2020）。关怀伦理理论提出应以“关注（attentiveness）”“响应（responsiveness）”“责任（responsibility）”和“胜任力（competency）”为核心进行互动（Tronto, 1993）。

以上公共政策伦理理论与助推的伦理争议密切相关，这些理论和指导框架都可作为助推在公共政策中的伦理应用提供实践思路。

（四）立足中国国情探索推进助推的伦理研究与干预实践

助推的伦理争议应超越非此即彼的哲学争辩，而回到政策实践中求得最优解。当前，关于助推在公共政策中的伦理可行性的实证研究大多基于西方政治体制或文化背景，尚缺乏在中国公共政策应用中伦理问题的系统性关注。但是，不同国家具有其特殊的社会结构与文化传统，公众对自身自主权的重视程度以及对政府的信任程度等都存在较大差异。Sunstein 等（2018）在调查中发现，中国公众对助推应用于不同的公共政策领域均表现了极高的热情。这可能源于中国公众的高政府信任，因此赞成政府的大部分政策；也可能由于中国公众较习惯硬家长式干预政策，因此更青睐助推这种自由主义家长式的干预模式（黄湛冰、刘磊，2020）。此外，中国等发展中国家对于隐私权和自主权的重视程度通常不如西方社会，这在手机 APP 信息采集或公共场合的摄像头安装等领域已有充分体现，使得中国公众可能并不热衷于讨论知识界看重的自主性问题，而更关心与其现实物质利益更密切相关的主题。但是，这并不意味着中国公众不关心伦理问题，而只是说明他们关心的伦理问题具有不同的价值序列与实现形态。如不能充分理解中国社会自身的伦理基础，在解决社会痛点、难点问题时就难以发挥助推策略应有的“以小博大”的优势。

当前，已有较多研究基于中国现实社会问题开展关于助推效果的实证检验，比如生育意愿（张书维等，2021）、健康饮食（耿晓伟等，2018）、善款捐赠（樊亚凤等，2019）、器官捐献（黄元娜等，2018）等。这些研究的关注点仍集中于“助推能否有效改变行为”的效率层面，而非“助推是否具备道德合理性”的伦理层面。同时，对助推效果的验证还多停留于小范围研究而未真正进入公共政策领域，尚未完成从实证研究到政策实践的转化，因而也未真正触及政策执行过程引发的切实伦理问题。此外，元分析研究还发现，基于默认选项效应的助推在西方文化背景下的效果好于其在东方文化背景下的效果（赵宁等，2022），这其中伦理因素扮演何种角色还有待进一步分析。

为此，可先探讨将助推策略应用于公共政策时可能引发的伦理争议，分析其中的影响因素并尝试提出解决建议，从而为可能的政策推行提供借鉴。而在此过程中，尤为重要的一点是应结合公众视角或常人视角（lay theories）而非专家视角或政府视角来理解其中的伦理关切，从而更全面地发现伦理交锋和政策执行阻碍的潜在根源。这样才能形成中国学者对助推伦理问题的独特贡献。此外，应加强中国背景下关于公共政策伦理问题的实证研究积累，可首先从当前发表的大量行为公共管理学实验（参见李晓倩，2018）为依据，从其研究领域和研究设计方面进一步归纳中国式助推的相关伦理议题，并采用实验、问卷调

查和元分析等验证公众自身的伦理立场对助推效果的实际影响（参见 Lemken, 2021; Michaelsen et al., 2021; Nilsson et al., 2020），从而超越抽象的哲学思辨、积累更为扎实的经验依据。在此基础上，可适时地寻找可能的政策领域与推行契机，结合中国自身的试点经验展开真正意义上的政策实践，以形成切实的政策影响力，从而为相关伦理争议的解决提供具有现实说服力的回应。

参考文献

- 代涛涛、陈志霞 (2019). 行为公共管理研究中的实验方法：类型与应用. 公共行政评论, 12: 166–203.
- Dai, T. T., & Chen, Z. X. (2019). Experimental Methods in Behavioral Public Administration: Types and Applications. *Journal of Public Administration*, 12: 166–203. (in Chinese)
- 丹尼尔·卡尼曼 (2012). 思考，快与慢. 胡晓娇，李爱民，何梦莹译. 北京：中信出版社.
- Kahneman, D. (2012). *Thinking, Fast and Slow*. (Hu, X. J., Li, A. M., & He, M. Y. Trans.). Beijing: China Citic Press. (Original Work Published 2011). (in Chinese)
- 樊亚凤、蒋晶、崔稳权 (2019). 网络公益平台默认选项设置对个人捐赠意愿的影响及作用机制. 心理学报, 51: 415–427.
- Fan, Y. F., Jiang, J., & Cui, W. Q. (2019). The Backfire Effect of Default Amounts on Donation Behavior in Online Donation Platform. *Acta Psychologica Sinica*, 51: 415–427. (in Chinese)
- 傅鑫媛、辛自强、楼紫茜、高琰 (2019). 基于助推的环保行为干预策略. 心理科学进展, 27: 1939–1950.
- Fu, X. Y., Xin, Z. Q., Lou, Z. X., & Gao, Y. (2019). Nudging Strategies for Pro-Environmental Behavior. *Advances in Psychological Science*, 27: 1939–1950. (in Chinese)
- 耿晓伟、张峰、王艳净、范琳琳、姚艳 (2018). 健康目标启动降低高热量食物消费. 心理学报, 50: 840–847.
- Geng, X. W., Zhang, F., Wang, Y. J., Fan, L. L., & Yao, Y. (2018). Health Goal Priming Decreases High-Calorie Food Consumption. *Acta Psychologica Sinica*, 50: 840–847. (in Chinese)
- 贺东航、孔繁斌 (2011). 公共政策执行的中国经验. 中国社会科学, 19: 61–79 + 220–221.
- He, D. H., & Kong, F. B. (2011). China's Experience in Public Policy Implementation. *Social Sciences in China*, 19: 61–79 + 220–221. (in Chinese)
- 黄元娜、宋星云、邵洋、李纾、梁竹苑 (2018). 以小博大：默认选项和反应模式效应助推中国器官捐献登记. 心理学报, 50: 868–879.
- Huang, Y. N., Song, X. Y., Shao, Y., Li, S., & Liang, Z. Y. (2018). Nudging: Default Option Effect and Response Mode Promote Organ Donor Registry Participation in China. *Acta Psychologica Sinica*, 50: 868–879. (in Chinese)
- 黄湛冰、刘磊 (2020). 轻推式干预未被中国广泛运用的原因：行为经济学分析. 公共管理与政策评论, 9: 55–68.
- Huang, Z. B., & Liu, L. (2020). Why Nudge Intervention Is Used Rare in China? An Analysis of Behavioral Economics. *Public Administration and Policy Review*, 9: 55–68. (in Chinese)
- 景怀斌 (2021). 行为公共政策研究的知识构建：三个层面及范式选择. 中国行政管理, 9: 56–63.
- Jing, H. B. (2021). The Knowledge Structure of Behavioral Public Policy: Three Perspectives and Its Paradigm. *Chinese Public Administration*, 9: 56–63. (in Chinese)
- 卡斯·R. 桑斯坦 (2021). 助推：快与慢. 王格非、路智雯译. 北京：中国人民大学出版社.
- Sunstein, C. R. (2021). *Nudge: Fast and Slow*. (Wang, G. F., & Lu, Z. W. Trans.). Beijing: China Renmin University Press (Original Work Published 2017). (in Chinese)
- 理查德·泰勒、卡斯·R. 桑斯坦 (2009). 助推：如何作出有关健康、财富与幸福的最佳决策. 刘宁译. 北京：中国人民大学出版社.
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. (Liu, N. Trans.). Beijing: China Renmin University Press (Original Work Published 2008). (in Chinese)
- 李晓倩 (2018). 行为公共管理学实验：基于 SSCI 期刊 (1978–2016) 相关文献的分析. 公共行政评论, 11, 37–61 + 219–220.
- Li, X. Q. (2018). Experimental Research in Behavioral Public Administration: An Analysis of Related Literature Based on SSCI-Indexed Journals (1978–2016). *Journal of Public Administration*, 11: 37–61 + 219–220. (in Chinese)
- 李燕、苏一丹、朱春奎 (2021). 公民政策遵从研究述评：基于“政策情境”与“行为特征”的二元视角. 公共行政评论, 14: 175–195 + 200.
- Li, Y., Su, Y. D., & Zhu, C. K. (2021). A Literature Review on Citizen's Policy Compliance: Based on the Perspective of "Policy Context" and "Behavior Characteristics". *Journal of Public Administration*, 14: 175–195 +

200. (in Chinese)
- 刘舒杨 (2020). 公共政策的功利主义伦理论析. *中国行政管理*, 8: 74 – 80.
- Liu, S. C. (2020). Research on the Utilitarian Ethics of Public Policy. *Chinese Public Administration*, 8: 74 – 80. (in Chinese)
- 吕小康、武迪、隋晓阳、汪新建、程婕婷(2018). 从“理性人”到“行为人”: 公共政策研究的行为科学转向. *心理科学进展*, 26: 2249 – 2259.
- Lyu, X. K., Wu, D., Sui, X. Y., Wang, X. J., & Cheng, J. T. (2018). From Rational Man to Behavioral Man: The Behavioral Turn of Public Policy. *Advances in Psychological Science*, 26: 2249 – 2259. (in Chinese)
- 吕孝礼、高娜、朱宪 (2020). 行为洞见与公共管理实践: 国际进展与启示. *中国行政管理*, 8: 125 – 133.
- Lyu, X. L., Gao, N., & Zhu, X. (2020). Behavioral Insights and Public Administration Practices: Progress and Implications. *Chinese Public Administration*, 8: 125 – 133. (in Chinese)
- 约翰·斯图亚特·密尔 (2005). 论自由. 许宝骙译. 北京: 商务印书馆.
- Mill, J. S. (2005). *On Liberty*. (Xu, B. K. Trans.). Beijing: The Commercial Press (Original Work Published 1959). (in Chinese)
- 张书维、李纾 (2018). 行为公共管理学探新: 内容, 方法与趋势. *公共行政评论*, 11: 7 – 36.
- Zhang, S. W., & Li, S. (2018). Exploring Behavioral Public Administration: Content, Method and Trend. *Journal of Public Administration*, 11: 7 – 36. (in Chinese)
- 张书维、谭小慧、梁歆佚、朱亚鹏 (2021). “助推”生育政策: 信息框架影响生育意愿的调查实验研究. *公共管理与政策评论*, 1: 42 – 54.
- Zhang, S. W., Tan, X. H., Liang, X. Y., & Zhu, Y. P. (2021). Nudging Fertility Policy: Two Survey Experiments of the Effect of Information Framework on Fertility Intention. *Public Administration and Policy Review*, (1): 42 – 54. (in Chinese)
- 张书维、王宇、周蕾 (2019). 行为公共政策视角下的助推与助力: 殊途同归. *中国公共政策评论*, 15, 20 – 38.
- Zhang, S. W., Wang, Y., & Zhou, L. (2021). “Nudge” and “Boost” from the Behavioral Public Policy Perspective: All Roads Lead to Rome. *Journal of Public Administration*, 15: 20 – 38. (in Chinese)
- 赵宁、刘鑫、李纾、郑蕊 (2022). 默认选项设置的助推效果: 来自元分析的证据. *心理科学进展*, 30: 1230 – 1241.
- Zhao, N., Li, X., Li, S., & Zheng, R. (2022). Nudging Effect of Default Options: A Meta-analysis. *Advances in Psychological Science*, 30: 1230 – 1241. (in Chinese)
- Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., ... & Wilson, S. (2017). Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. *ACM Computing Surveys*, 3(50): 1 – 41.
- Baldwin, R. (2014). From Regulation to Behaviour Change: Giving Nudge the Third Degree. *Modern Law Review*, 77(6): 831 – 857.
- Baum, N. M., Gollust, S. E., Goold, S. D., & Jacobson, P. D. (2007). Looking Ahead: Addressing Ethical Challenges in Public Health Practice. *The Journal of Law Medicine & Ethics*, 35(4): 657 – 667.
- Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Jonsson, M. L., & Rahali, B. (2018). Can Nudges Be Transparent and Yet Effective? *Journal of Economic Psychology*, 65: 41 – 59.
- Capasso, M., & Umbrello, S. (2022). Responsible Nudging for Social Good: New Healthcare Skills for AI-Driven Digital Personal Assistants. *Medicine, Health Care and Philosophy*, 25: 11 – 22.
- Christian, S. (2017). Green Nudges: Do They Work? Are They Ethical? *Ecological Economics*, 132(1): 329 – 342.
- Cohen, G. A. (2008). *Rescuing Justice and Equality*. Cambridge: Harvard University Press.
- Cohen, J. E. (2013). What Privacy Is For. *Harvard Law Review*, 126(7): 1904 – 1933.
- Cookson, R. (2013). A Response to Disney, Le Grand and Atkinson. In Oliver, A. (Ed.), *Behavioural Public Policy* (pp. 88 – 93). Cambridge University Press.
- Croson, R., & Treich, N. (2014). Behavioral Environmental Economics: Promises and Challenges. *Environmental and Resource Economics*, 3(58): 335 – 351.
- Dewies, M., Schop-Etman, A., Rohde, K. I. M., & Denktas, S. (2021). Nudging Is Ineffective When Attitudes Are Unsupportive: An Example from a Natural Field Experiment. *Basic and Applied Social Psychology*, 43(4): 213 – 225.
- DiSilvestro, R. (2012). What Does not Budge for any Nudge? *The American Journal of Bioethics*, 12(2): 14 – 15.
- Dranseika, V., & Piasecki, J. (2020). Transparent Defaults and Consent for Participation in a Learning Health Care System: An Empirical Study. *Journal of Empirical Research on Human Research Ethics*, 15(4): 261 – 270.
- Dworkin, G. (2019). Paternalism. In Zalta E. N. (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition). Retrieved September 6, 2021, from <https://plato.stanford.edu/archives/fall2019/entries/paternalism/>
- Engelen, B. (2019). Ethical Criteria for Health-Promoting Nudges: A Case-by-Case Analysis. *American Journal of Bioethics*, 19(5): 48 – 59.
- Evers, C., Marchiori, D. R., Junghans, A. F., Cremers, J., & Ridder, D. (2018). Citizen Approval of Nudging Interventions Promoting Healthy Eating: The Role of Intrusiveness and Trustworthiness. *BMC Public Health*, 18(1): 1182.
- Ewert, B. (2020). Moving Beyond the Obsession with Nudging Individual Behaviour: Towards a Broader Understanding

- of Behavioural Public Policy. *Public Policy and Administration*, 35(3): 337–360.
- Floridi, L., Cowls, J., King, T., & Taddeo, M. (2020). Designing AI for Social Good: Seven Essential Factors. *Science and Engineering Ethics*, 26: 1771–1796.
- Friedman, B., Kahn, P. H., Borning, A., & Hultgren, A. (2013). Value Sensitive Design and Information Systems. In Doorn, N., Schuurbijs, D., van de Poel, I. & Gorman, M. E. (Eds.), *Early Engagement and New Technologies: Opening Up the Laboratory* (pp. 55–95). Dordrecht: Springer.
- Grilli, G., & Curtis, J. (2021). Encouraging Pro-Environmental Behaviours: A Review of Methods and Approaches. *Renewable and Sustainable Energy Reviews*, 135: 110039.
- Grüne-Yanoff, T., & Hertwig, R. (2016). Nudge Versus Boost: How Coherent Are Policy and Theory? *Minds & Machines*, 26(1): 149–183.
- Hager, G. D., Drobni, A., Fang, F., Ghani, R., Greenwald, A., Lyons, T., ... & Tambe, M. (2017). *Artificial Intelligence for Social Good*. 2022–3–21. Retrieved from <https://doi.org/10.48550/arXiv.1901.05406>
- Halpern, D. (2015). *The Nudge Unit: Inside the Government Department that Changed Our Minds and Save*. New York, NY: Random House.
- Hansen, P. G. (2016). The Definition of Nudge and Libertarian Paternalism: Does the Hand Fit the Glove? *European Journal of Risk Regulation*, 7(1): 155–17
- Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the Manipulation of Choice: A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy. *European Journal of Risk Regulation*, 1(1): 3–28.
- Hausman, D. M., & Welch, B. (2010). Debate: To Nudge or not to Nudge. *Journal of Political Philosophy*, 18(1): 123–136.
- Ioannou, A., Tussiyadiah, I., Miller, G., Li, S., & Weick, M. (2021). Privacy Nudges for Disclosure of Personal Information: A Systematic Literature Review and Meta-Analysis. *PLoS ONE*, 16(8): e0256822.
- Jaffe, H. W., & Hope, T. (2010). Treating for the Common Good: A Proposed Ethical Framework. *Public Health Ethics*, 3(3): 193–198.
- Junghans, A. F., Cheung, T. T., & Ridder, D. D. (2015). Under Consumers' Scrutiny—An Investigation into Consumers' Attitudes and Concerns about Nudging in the Realm of Health Behavior. *BMC Public Health*, 15(1): 331.
- Kantorowicz-Reznichenko, E., & Kantorowicz, J. (2021). To Follow or not to Follow the Herd? Transparency and Social Norm Nudges. *Kyklos*, 74(3): 362–377.
- Kelly, M. P., & Barker, M. (2016). Why Is Changing Health-Related Behaviours Difficult? *Public Health*, 136: 109–116.
- Krisam, M., Maier, M., Janßen, R., & Krisam, J. (2021). What Do Germans Really Think about Health-Nudges? *BMC Public Health*, 21(1): 821.
- Ledderer, L., Kjr, M., Madsen, E. K., Busch, J., & Fage-Butler, A. (2020). Nudging in Public Health Lifestyle Interventions: A Systematic Literature Review and Metasynthesis. *Health Education & Behavior*, 47(5): 749–764.
- Lee, L. M. (2012). Public Health Ethics Theory: Review and Path to Convergence. *Journal of Law Medicine & Ethics*, 40(1): 85–98.
- Lemken, D. (2021). Options to Design More Ethical and Still Successful Default Nudges: A Review and Recommendations. *Behavioural Public Policy*, <https://doi.org/10.1017/bpp.2021.33>
- Lin, Y., Osman, M., Harris, A. J. L., & Read, D. (2018). Underlying Wishes and Nudged Choices. *Journal of Experimental Psychology Applied*, 24(4): 459–475.
- Lodge, M., & Wegrich, K. (2016). The Rationality Paradox of Nudge: Rational Tools of Government in a World of Bounded Rationality. *Law & Policy*, 38(3): 250–267.
- Loewenstein, G., Sunstein, C. R., & Golman, R. (2014). Disclosure: Psychology Changes Everything. *Annual Review of Economics*, 6(1): 391–419
- Michaelsen, P., Johansson, L. O., & Hedesström, M. (2021). Experiencing Default Nudges: Autonomy, Manipulation, and Choice-Satisfaction as Judged by People Themselves. *Behavioural Public Policy*, <https://doi.org/10.1017/bpp.2021.5>
- Mols, F., Haslam, S. A., Jetten, J., & Steffens, N. K. (2015). Why a Nudge Is not Enough: A Social Identity Critique of Governance by Stealth. *European Journal of Political Research*, 54(1): 81–98.
- Moser, S. C. (2007). More Bad News: The Risk of Neglecting Emotional Responses to Climate Change Information. In Moser, S. C., & Dilling, L. (Eds.), *Creating a Climate for Change: Communicating Climate Change and Facilitating Social Change* (pp. 64–80). New York: Cambridge University Press.
- Mrkva, K., Posner, N. A., Reeck, C., & Johnson, E. J. (2021). Do Nudges Reduce Disparities? Choice Architecture Compensates for Low Consumer Knowledge. *Journal of Marketing*, 85(4): 67–84.
- Nilsson, A., Erlandsson, A., Västfjäll, D., & Tinghög, G. (2020). Who Are the Opponents of Nudging? Insights from Moral Foundations Theory. *Comprehensive Results in Social Psychology*, <https://doi.org/10.1080/23743603.2020.1756241>
- Nørnberg, T. R., Houlby, L., Skov, L. R., & Pérez-Cueto, F. J. A. (2016). Choice Architecture Interventions for

- Increased Vegetable Intake and Behaviour Change in a School Setting: A Systematic Review. *Perspectives in Public Health*, 136(3): 132 – 142.
- Osman, M. (2016). Nudge: How Far Have We Come? *Oeconomia-History, Methodology, Philosophy*, 6(4): 557 – 570.
- Paul, B. R., Paolo, B., Nicola, B., & Paola, C. (2019). Behavioral Public Administration Ad Fontes: A Synthesis of Research on Bounded Rationality, Cognitive Biases, and Nudging in Public Organizations. *Public Administration Review*, 79(3): 304 – 320.
- Paunov, Y., Wänke, M., & Vogel, T. (2020). Combining Defaults and Transparency Information to Increase Policy Compliance. *Social Psychology*, 51(5): 354 – 359.
- Petrakaki, D., Hilberg, E., & Waring, J. (2021). The Cultivation of Digital Health Citizenship. *Social Science & Medicine*, 270: 113675.
- Prainsack, B. (2020). The Value of Healthcare Data: To Nudge, or Not? *Policy Studies*, 41(5): 547 – 562.
- Quong, R. B. J. (2010). Justice Beyond Equality. *Social Theory and Practice*, 36(2): 315 – 340.
- Reynolds, J. P., Archer, S., Pilling, M., Kenny, M., Hollands, G. J., & Marteau, T. M. (2019). Public Acceptability of Nudging and Taxing to Reduce Consumption of Alcohol, Tobacco, and Food: A Population-Based Survey Experiment. *Social Science & Medicine*, 236: 112395.
- Riley, E. (2017). The Beneficent Nudge Program and Epistemic Injustice. *Ethical Theory and Moral Practice*, 20(3): 597 – 616.
- Rizzo, M., & Whitman, G. (2019). *Escaping Paternalism: Rationality, Behavioural Economics, and Public Policy*. Cambridge University Press.
- Roberts, J. L. (2018). Nudge-Proof: Distributive Justice and the Ethics of Nudging. *Michigan Law Review*, 116(6): 1045 – 1066.
- Ruehle, R. C., Engelen, B., & Archer, A. (2021). Nudging Charitable Giving: What (if Anything) Is Wrong with It? *Nonprofit and Voluntary Sector Quarterly*, 50(2): 353 – 371.
- Sætra, H. S. (2019). When Nudge Comes to Shove: Liberty and Nudging in the Era of Big Data. *Technology in Society*, 59: 101130.
- Sætra, H. S. (2020). A Shallow Defense of a Technocracy of Artificial Intelligence: Examining the Political Harms of Algorithmic Governance in the Domain of Government. *Technology in Society*, 62: 101283.
- Schüll, N. D. (2016). Data for Life: Wearable Technology and the Design of Self-Care. *BioSocieties*, 11(3): 317 – 333.
- Sunstein, C. R. (2015). Nudges Do not Undermine Human Agency. *Journal of Consumer Policy*, 38(3): 207 – 210.
- Sunstein, C. R. (2021). The Distributional Effects of Nudges. *Nature Human Behaviour*, 6(1), 9 – 10.
- Sunstein, C. R., Reisch, L. A., & Rauber, J. (2018). A World-Wide Consensus on Nudging? Not Quite, but Almost. *Regulation & Governance*, 12(1): 17 – 39.
- Treger, C. (2021). When Do People Accept Government Paternalism? Theory and Experimental Evidence. *Regulation & Governance*, <https://doi.org/10.1111/rego.12443>.
- Tronto, J. (1993). *Moral Boundaries: A Political Argument for an Ethics of Care*. New York: Routledge.
- Tronto, J. (2010). Creating Caring Institutions: Politics, Plurality, and Purpose. *Ethics and Social Welfare*, 4(2): 158 – 171.
- Umbrello, S., & van de Poel, I. (2021). Mapping Value Sensitive Design onto AI for Aocial Good Principles. *AI and Ethics*, 1(3): 283 – 296.
- Vallgarda, S. (2012). Nudge-A New and Better Way to Improve Health? *Health Policy*, 104(2): 200 – 203.
- Veretilnykova, M., & Dogruel, L. (2021). Nudging Children and Adolescents Toward Online Privacy: An Ethical Perspective. *Journal of Media Ethics*, 36(3): 128 – 140.
- Vugts, A., van den Hoven, M., de Vet, E., & Verweij, M. (2020). How Autonomy Is Understood in Discussions on the Ethics of Nudging. *Behavioral Public Policy*, 4(1): 108 – 123.
- Weber, E. U. (2017). Breaking Cognitive Barriers to a Sustainable Future. *Nature Human Behaviour*, 1(1): 0013.
- White, M. D. (2017). Preferences All the Way Down: Questioning the Neoclassical Foundations of Behavioral Economics and Libertarian Paternalism. *Oeconomia*, 7(3): 353 – 373.
- Wilkinson, T. M. (2013). Nudging and Manipulation. *Political Studies*, 61(2): 341 – 355.
- Yan, H. Y., & Yates, J. F. (2019). Improving Acceptability of Nudges: Learning from Attitudes Towards Opt-In and Opt-Out Policies. *Judgment and Decision Making*, 1(14): 26 – 39.
- Yeung, K. (2017). “Hypernudge”: Big Data as a Mode of Regulation by Design. *Information Communication & Society*, 20(1): 118 – 136.
- Zimmermann, V., & Renaud, K. (2021). The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions. *ACM Transactions on Computer-Human Interaction*, 28(1): 1 – 45.

责任编辑：张书维