

全国中文核心期刊

中文社会科学引文索引(CSSCI)来源期刊

中国人文社会科学期刊AMI综合评价(A刊)核心期刊

理论粤军·教育部在粤人文社科重点研究基地建设资助项目

2024. 6

VOL.17 NO.6

公共行政评论

JPA
Journal of Public Administration

● 专栏：基层治理

专栏导语：基层治理中的政策执行及其影响

..... 肖 滨 1

基层政府负担的致因与消解：基于复杂适应系统理论

..... 刘元贺 周振超 4

基层治理中“比照式执行”的多重行为逻辑

..... 何 雷 唐 娜 23

压力聚合与结构利用：常规政策执行情景中基层政府何以逆向推动条块协作

..... 黄源源 唐银彬 40

调适性赋权：流动人口社区融入何以可能？

..... 蒋红军 张 东 60

可解释性何以影响基层公务员的算法决策采纳偏好

..... 段哲哲 焦嘉欣 78

●论文

模糊政策为何执行失败

——基于 A 市燃气企业安全生产标准化建设政策的个案研究

..... 祝 哲 林 婕 祝清军 张美莲 99

高质量发展背景下政府如何评价养老服务供给者？

——一个层次分析法模型的建构与检验

..... 彭 婧 杨 桥 肖 帅 118

县域城镇化进程中技术赋能如何促进整体性治理？

——基于 30 个案例的定性比较分析

..... 李燕凌 陈梦雅 139

人工智能技术应用如何影响街头官僚的自由裁量权使用意愿？

——一项调查实验研究

..... 孙 宇 邓 怡 孙琳之 158

●理论综述

公共创新实验室的作用功能、创新成效与关键因素：文献回顾及启示

..... 袁千里 刘新萍 176

●英文目录与摘要 196

●年度总结

《公共行政评论》2024 年评审人和专栏导语作者名单 201

《公共行政评论》2024 年总目录 203

可解释性何以影响基层公务员的 算法决策采纳偏好

段哲哲 焦嘉欣*

【摘要】以算法技术为辅助工具浸入公共部门决策已变得无处不在，人工智能渗透公共治理过程的新一代治理模式成为未来公共部门数字转型的重要方向。然而，当前算法系统“黑箱”属性受到越来越多的质疑，带来算法治理的“信任赤字”与“问责缺失”等问题。研究从可解释的人工智能（Explainable Artificial Intelligence, XAI）逻辑出发，通过一项 2（解释来源：规则 vs 数据）× 2（解释内容：程序 vs 结果）的调查实验，在税务部门应用算法决策系统的场景中，探索不同类型 XAI 对传统官僚采用算法决策建议的作用机制。实证结果显示：可解释的算法系统可以促进基层公务员对算法决策的采纳偏好，其中数据驱动的程序解释型算法系统或数据驱动的结果解释型算法系统，相较无解释组，均会促进基层公务员对算法决策的采纳偏好；数据驱动的 XAI 比规则驱动的 XAI 更能促进传统官僚对算法决策的采纳偏好；数据驱动结果可解释性算法系统能够增强传统官僚对于算法系统的公平感知和技术信任，从而更偏好采纳算法决策。这些发现表明，不仅要发展算法官僚的技术能力，也要建构传统官僚能够理解的算法系统，探索人机良性互动模式，才能推动算法系统在公共部门的顺利落地。

【关键词】算法决策 可解释性 人工智能 人机互动 算法官僚 基层公务员

【中图分类号】D63

【文献标识码】A

【文章编号】1674 - 2486 (2024) 06 - 0078 - 21

* 段哲哲，深圳大学政府管理学院、廉政研究院、全球特大型城市治理研究院院长聘副教授，深圳大学新质生产力研究院副院长，广东省习近平新时代中国特色社会主义思想研究中心深圳大学基地研究员。通讯作者：焦嘉欣，深圳大学传播学院博士研究生。感谢匿名评审专家、编委会专家和编辑部对论文的修改意见。

基金项目：国家社会科学基金青年项目“少数基层干部‘空忙’现象及矫治机制研究”（21CZZ03）；广东省普通高校特色创新类项目（哲学社会科学）“新时代超大城市社会治理现代化实践探索与长效机制研究”（2023WTSCX076）；深圳大学 2024 年研究生自主创新成果培育项目“公共部门算法应用的基层公务员接受度影响因素与提升路径研究”（868 - 000002020293）。

一、引言

随着人工智能的最新进展，算法嵌入公共治理过程改变了政府和基层公务员与民众互动的方式。基于深度学习、神经网络等技术的算法系统，以辅助行政裁量、自动化决策、预测调配资源的方式陆续嵌入各国司法判决、预测性警务、税务稽查、反腐、证券交易监管、行政审批与网络问政等公共治理领域的不同场景，增强或代替人类的决策（吴田、张帆，2020；许多奇、蒋龙宇，2023；Schmidt et al., 2020）。算法决策提升了行政系统的运行效率，创造了公共价值（Twizeyimana & Andersson, 2019）。基层公务员的行政行为越来越受到算法规则的指引，部分官僚自由裁量权逐步由人转移到机器上（Peeters, 2020；王贵，2021）。然而，许多算法应用因其不透明与“黑箱”属性影响了人们对算法决策的接受度，产生自动化偏见或算法厌恶的现象，带来算法治理的“信任赤字”与“问责缺失”等问题（Alon-Barkat & Busuioc, 2023；汤志伟、韩啸，2023；刘特等，2024）。

影响人类对人工智能信任的各种因素中，可解释性是核心且稳健的变量（Choung et al., 2023；Shin, 2020）。尤其是在高风险公共决策情境下，决策者必须清楚算法所扮演的角色，以便他们能够在考虑所有影响因素的情况下作出决定，而可解释的人工智能（Explainable Artificial Intelligence, XAI）可以推动人类理解算法系统的推理过程，成为传统官僚接受算法系统建议并作出决策的必要条件：一是确保决策公正，检测并纠正训练数据集中的偏见；二是不断改进系统的稳健性，以保持高水平的学习性能；三是作为保险机制，检验模型推理中可能存在的潜在因果关系，提升决策的合理性（Miller, 2019；Meske et al., 2022）。实务层面各国已将可解释性作为人工智能算法应用的核心原则，2021年11月，联合国发布的《人工智能伦理问题建议书》明确提出“人工智能系统的透明度和可解释性往往是确保人权、基本自由和伦理原则得到尊重、保护和促进的必要先决条件”。2022年3月，由国家互联网信息办公室等4部门联合发布施行的《互联网信息服务算法推荐管理规定》，提出要确保算法应用“规则的透明度和可解释性”。然而，现有研究对算法可解释性主要关注技术路径，较少从基层公务员视角出发（Giest & Grimmelikhuijsen, 2020）。基于此，本研究从XAI逻辑出发，基于算法可解释性的类型学，验证算法可解释性对基层官僚接受算法建议的社会心理机制，为推进数字政府在中国地域文化多样、队伍庞大的官僚机构落地提供可行的微观路径。

二、文献回顾与研究假设

（一）基层官僚算法决策采纳偏好的影响因素

算法决策（Algorithmic Decision Making, ADM）或自动化决策（Automated Decision Making, ADM）一般由算法或人工智能根据编码指令进行信息收集、加工、处理并自动作出决策，影响真实或虚拟环境的决策（Stone, 1971; Nagtegaal, 2021）。自动化在提升行政效率的同时也产生决策不透明的情况，算法决策的有意不透明性、技能不透明性与算法固有不透明性带来了“新的数字鸿沟”，使传统官僚产生隐形抵制、选择性遵从等不适应行为，在决策过程中，拥有自由裁量权的基层公务员可能拒绝采纳算法决策建议，以维护其专业自主性，从而影响了算法系统在公共部门部署进度（权一章、黄晖，2024; Alon-Barkat & Busuioc, 2023）。因此，如何促进基层公务员接纳算法系统并采纳其决策建议成为当前数字政府研究的热门议题之一。基于已有研究，可将影响基层官僚算法决策采纳偏好的因素分为“环境—技术—行动者”三类（Williams et al., 2015; Aysolmaz et al., 2023）：

一是环境维度。传统官僚处于“庙堂”与“江湖”双重情境之中，决策行为受到外部社会环境与内部组织环境的共同作用。一方面，社会规范对传统官僚进行伦理与道德规约，促使他们借助技术优化政府运作系统，外部的回应性任务压力可能会推动基层公务员积极采纳新技术（Twizeyimana & Andersson, 2019）。另一方面，组织内部环境是组织成员行为的“培养皿”，不同的组织内部因素会影响政府部署与个人采纳算法系统的意愿。如果组织部署的算法系统与组织价值契合，有利于提升个人绩效，基层公务员会理性选择，根据组织考核指标自动遵守算法建议（Mahmud et al., 2022）。当个人判断与算法建议冲突时，他们可能回避具有不确定性的算法决策，或出于“算法卸责”动机自动遵守算法建议来规避问责风险（Alon-Barkat & Busuioc, 2023）。此外，科层制内部的组织规模、资源分配、领导者风格、同侪压力与组织技术能力等要素会显著影响基层公务员的算法采纳行为（Ahn & Chen, 2022; Selten et al., 2023）。

二是技术维度。该维度指向算法系统使用技术的性质，如技术复杂性、可访问性、可解释性与透明度等特性。经典技术采纳模型（Technology Acceptance Model, TAM）指出技术的被采纳与否受到感知有用性与感知易用性的影响，而这两个因素受到技术复杂性、技术架构设计与人机协同安排等技术特性的影响。已有研究显示，基层公务员接纳算法决策偏好受到算法系统性能的影响，而FATE是算法服务设计和开发需要考虑的关键因素，FATE即算法系统的公平性（Fairness）、问责制（Accountability）、透明度（Transparency）和可解释性（Explainability）特征

(Meske et al., 2022; Madan & Ashok, 2023; 樊博、李晶晶, 2024)。

三是行动者维度。现实中传统官僚面对算法建议将基于专业能力对算法技术进行“裁剪”或“再造”，进而作出“决策中的决策”（Selten et al., 2023），这些潜在的个体特质包括外显层与内隐层。外显层的年龄、性别、教育水平等人口学意义的个性特征与算法素养、专业自主性、任务复杂性等胜任力条件均会影响传统官僚的技术感知，进而影响他们如何使用技术“工具箱”对算法决策作出“再决策”（Wang et al., 2023）。内隐层的心理因素则触发传统官僚在决策过程中的信息加工和利益判断，自我效能感、情绪态度、价值观念等心理属性可推演其历经历了何种动机推理活动而作出决策（Alon-Barkat & Busuioc, 2023）。

这些研究都关注到了算法系统这种新技术在公共部门应用的各类因素，但仍有两个方面有待提升：一是研究视角的局限性。当前研究对算法系统应用的内外环境的必要性有高度共识，且认为是算法系统“黑箱”特征影响了其在公共部门的应用，而可解释的人工智能成为缓解算法“黑箱”行政组织负面影响、弥合算法信任赤字的重要工具（Taeihagh, 2021; Meske et al., 2022），大量研究提出的 XAI 是技术专家的视角，而非基层公务员的视角。二是规范研究的局限性。大量研究提出了多个从 XAI 的维度来推动算法系统的部署，如技术层面实现算法系统可访问性、数据透明度、程序性解释等，实现良好的人机互动关系（Wang et al., 2023; 权一章、黄晖, 2024），但较少地从实证角度验证基层公务员如何看待不同类型的 XAI 及其对算法建议采纳的影响。

（二）算法决策可解释性内涵及其类型

XAI 是从委托人的初始命令出发，通过系统程式化推断产生结果的算法机器，向用户提供易于理解的推理链。《人工智能伦理问题建议书》将 XAI 定义为人工智能系统的结果可以被人们理解，XAI 可以提供阐释说明，包括算法系统的输入、处理与输出的过程。一般认为 XAI 的可解释性是机器向人类呈现可理解的术语的能力，这种能力将机器输出的参数表述，对应转换为人类生活世界的若干实体（苏宇, 2022; Doshi-Velez & Kim, 2017）。XAI 是基于这样一种理念的领域：如果算法系统给出的建议可以向人类解释，那么人类更容易接受这些建议（Shin, 2020; 2021; Choung et al., 2023; Schmidt et al., 2020）。

XAI 在人们对算法可自我论证的期待下应运而生，随着技术开发和理论研究的拓展出现了多种类型的 XAI。根据解释的来源，XAI 可分为规则驱动型（Rule-Driven）和数据驱动型（Data-Driven）（Janssen et al., 2022; Li et al., 2020; Wang et al., 2023）。规则驱动也被称为知识驱动（Knowledge-Aware），解释来自事先制定的一系列正式规则和专业知知识。此类系统基于自上而下的编程逻辑，将传统官僚决策依赖的法律法规、专业知识、组织公共价值权衡等显性知识通过代码转化为算法，机器通过接收到的指令复制人类行为，据此提供

解释 (Janssen et al., 2022; 段哲哲, 2021)。数据驱动 XAI 指自动从收集的数据中生成解释, 不需要纳入复杂的行政规则。该算法决策系统基于自下而上的编程逻辑, 以大量历史数据训练机器学习算法系统, 在不接收显式指令的情况下作出推断和预测行为, 据此提供解释 (Wang et al., 2023)。

根据解释的内容, XAI 可分为程序解释型和结果解释型 (Dodge et al., 2019; Janssen et al., 2022)。行政决策的正当性原则认为行政机关对其所作出的行政决定负有说明理由的义务, 以阐明其行为是基于充分的事实和理由而作出的, 包括对决策程序和决策结果的解释 (王贵, 2021)。决策程序强调行政机关作出影响行政当事人权益的行政行为时, 应当合理告知相对人根据法律法规、事实进行推理或裁量而得到结论的过程, 意在破解算法实现决策功能的过程逻辑, 使人们能够了解算法模型的工作原理。决策结果强调对结论与理由间的内在联系作出解释, 侧重用简单易懂的日常语言解释算法生成特定结果背后的因果关系, 以便人们可以识别显著影响输出结果的参数 (Oswald, 2018; 张凌寒, 2020; Janssen et al., 2022)。

现实算法应用案例中, 算法嵌入公共部门决策的基本流程是, 公共部门委托数据科学家编写算法决策代码, 经过若干设定好的处理步骤得到结果 (王贵, 2021; Janssen et al., 2022)。由此, 基于解释来源及解释内容两个维度提出 4 种可解释算法决策系统^①: 规则驱动的程序解释型、规则驱动的结果解释型、数据驱动的程序解释型、数据驱动的结果解释型。

第一, 规则驱动的程序解释型算法决策系统 (规则驱动 × 程序解释)。该算法系统基于一系列明确定义的正式规则和条件作出决策, 对其所遵循的行政程序作出解释说明, 旨在体现决策依照的正当程序和流程 (Binns et al., 2018)。对于导致决策结果的具体规则不提供明确解释, 最终解释权交予传统官僚。此类应用程序内嵌具体行政程序, 包含标准化的办理流程, 实现了算法系统对原有业务规则的承接。在海南省电子税务局办税业务板块的增值税专用发票票种核定环节, 系统主界面依据相关法律法规对纳税人税收风险程度、信用级别、实际经营情况等可能被限制办理的情形作出阐述, 为用户提供操作指引, 借助文字说明对业务办理进行程序解释, 对于业务办理过程中用户不能正常通过相关程序, 无法办理的原因则需要税务人员解释。

第二, 数据驱动的程序解释型算法决策系统 (数据驱动 × 程序解释)。此类算法系统主要对其数据来源和使用过程进行解释, 不出具导致特定结果的详细数据。数据驱动的方式通常将选定的公共事务数据集或其变体输入到“黑箱”模型中, 训练机器自主学习, 挖掘和匹配可用数据作出决策解释。江西省“电

^① 由于篇幅限制, 未呈现“不同类型的可解释算法决策系统”分类, 若读者想了解可以联系作者索取。

子警察”抓拍系统对车辆交通违法行为实施非现场自动抓拍,“电子警察”主要依托智能设备感应地理信息、路面压力等数据变化,判断是否存在违法情况。但其只能通过图片解释初步证明可能存在的违法事实,而无法提供详细参数或因果证据,要作出行政决定仍需要其他证据予以佐证。如“电子警察”拍摄的照片只能表明机动车发生剐蹭,而无法证明相对人明知情况但却“逃逸”的结论,最终需要传统官僚进一步复核判定并给出解释(张静静,2023)。

第三,规则驱动的结果解释型算法决策系统(规则驱动×结果解释)。该算法系统基于规则作出决策的同时,在交互界面对规则运行结果进行简洁且可以理解的阐述。无须展示详细的规则内容,重在提示用户“下一步应该如何改进”(Burton et al., 2020),较少需要传统官僚参与最终解释。当前,公共行政领域应用的此类算法系统往往只能给出初步意见,最终解释权仍由传统官僚控制,其尚且不足以支持算法官僚自主作出因果解释。

第四,数据驱动的结果解释型算法决策系统(数据驱动×结果解释)。此类算法决策系统通过自主学习对数据进行处理分析得出决策结果,对其检测到的数据相关性进行因果解释。该系统并不需要解释其背后的运作过程,而是通过说明关键参数的变动与决策结果的关联,解释特定参数何以引致结果(Schmidt et al., 2020; 樊博、李晶晶, 2024),解释更多依赖于算法官僚所提供的证据。2018年北京市信访办投入使用首都“智慧信访”大数据分析平台,通过机器学习、人工智能大数据分析技术,预测上访人员的行为倾向并提供量化结果与原因解释,如情感正负倾向、上访者行为轨迹、相关案例关联分析等(吴进进、何包钢, 2023)。

(三) 可解释性对基层官僚算法决策采纳偏好的影响

可解释性是公平、负责任和可信任的算法系统的先决条件(Fernandez et al., 2019; Miller, 2019),会影响人类与算法系统的互动方式(Alt, 2018; Meske et al., 2022),触发公务员使用算法系统的信心(Alon-Barkat & Busuioc, 2023)。技术接受模型认为个人的系统使用行为是由其使用意愿决定的,而意愿受到有用性感知的影响,XAI以可见的方式论证其作出推断的原因,会显著提高人们的技术有用性感知,进而促进人们的采纳意愿(Aysolmaz et al., 2023)。解释来源奠定XAI的论证基础,决定了人们通过哪一视角来理解算法提供的解释。基于规则的可解释性被认为是算法系统的“拟人化”(Anthropomorphism),会使人们从情感上倾向于相信算法能够作出合理决策(Shin, 2021; Choung et al., 2023)。数据驱动的算法决策系统代替人类快速地访问大量准确、组织良好的多维数据,根据观察到的和潜在的解变量相关概率作出推断和预测,使人们相信使用技术将获得更大收益(Selten et al., 2023)。

在解释内容方面,XAI划定了人们认识和理解算法决策的基本范围。程序解

释为“在环路”的人提供了与机器连接的接口，引导传统官僚确认决策证据，监督运行过程，提升自我效能感，从而提高他们采纳算法决策建议的意愿（Schmidt et al. , 2020；Alon-Barkat & Busuioc, 2023）。结果解释构造了一个事后解释机制，简化了理解过程，明确了因果关系，赋予 XAI 引导人们“找到答案”的角色，缩短人们的反应时间（Van Berkel et al. , 2021）。对于“为什么是这个结果”，人们不需要进行下一步推理，也不需要了解事实的所有原因，就可以知晓算法系统为何作出某个推断，帮助人们达到预期目标（Miller, 2019；Lee et al. , 2019）。基于此，本文提出如下假设。

H1：含解释来源和解释内容的算法系统会显著促进基层公务员对算法决策的采纳偏好。

H1a：规则驱动的程序解释型算法系统相较无解释组，会促进基层公务员对算法决策的采纳偏好。

H1b：数据驱动的程序解释型算法系统相较无解释组，会促进基层公务员对算法决策的采纳偏好。

H1c：规则驱动的结果解释型算法系统相较无解释组，会促进基层公务员对算法决策的采纳偏好。

H1d：数据驱动的结果解释型算法系统相较无解释组，会促进基层公务员对算法决策的采纳偏好。

解释是社会性的，包含不同解释来源与解释内容的组合 XAI 会对基层公务员算法决策建议的采纳意愿产生不同程度的影响（Li et al. , 2020；Wang et al. , 2023）。在解释来源方面，规则驱动的 XAI 通过建构人机兼容性呈现更高的解释水平（Burton et al. , 2020）。当人们认为算法决策结果存在反事实偏差，即不符合个人认知和预期时，会通过搜索人机一致性信息找寻答案（Miller, 2019）。正式规则和专业知识的匹配为人机兼容性提供了契机，减少人们的认知压力（Burton et al. , 2020）。解释水平理论认为，感知相似性是人机心理距离的有力解释。规则嵌入的“拟人化”解释能拉近人对机器的心理距离，更可能促进人们采纳算法决策的建议（沈鹏熠等，2023；Trope et al. , 2007）。实证研究显示，相较于数据驱动，公务员更偏好采纳规则驱动的算法系统的建议（Wang et al. , 2023）。

在解释内容方面，相对结果解释，程序解释提供了更为详细的决策依据和推断过程。在对行政决定作出解释时，人类决策者往往期望证据是充分而翔实的（Burton et al. , 2020）。程序解释能够使人们了解庞杂的规则和数据如何应用于算法决策过程，这可以强化决策的可追溯性，满足人们的询证需要。程序解释通常需要传统官僚参与最终决策解释，这一过程满足传统专业技术官僚以专业能力拥有自由裁量权的心理需求（Burton et al. , 2020；Peeters, 2020）。相关实证研究表明，人类参与算法决策的卷入度越高，感知决策控制越强，越可

能采纳算法决策的建议 (Omrani et al., 2022)。因此,解释来源与解释内容对基层公务员的算法决策采纳偏好产生交互影响。基于此,本文提出如下假设。

H2: 解释来源与解释内容的组合 XAI 会对基层公务员的算法决策采纳偏好产生不同程度的影响。

H2a: 规则驱动的 XAI 较数据驱动的 XAI 更能促进基层公务员对算法决策的采纳偏好。

H2b: 程序解释的 XAI 较结果解释的 XAI 更能促进基层公务员对算法决策的采纳偏好。

H2c: 基层公务员的算法决策采纳偏好中, XAI 的解释来源与解释内容存在交互效应。

(四) 算法可解释性影响算法决策采纳的作用机制

多种技术采纳理论模型指出技术特征会通过影响人类主观感知,从而影响人类的新技术采纳偏好 (Janssen et al., 2022; Wang et al., 2023)。在影响人类技术采纳的主观感知变量中,公平感知和技术信任得到了广泛讨论 (Williams et al., 2015; Venkatesh & Bala, 2008; Wang et al., 2023)。公平感知形塑了人们如何看待机器分配公共价值的能力,而信任描述了人机之间的互动关系 (Venkatesh & Bala, 2008)。公平感知指人们认为算法的判断是否平等和一致,这会影响公共决策的合法性 (de Fine Licht, 2011)。技术采纳模型认为,由于人们将社会规范的价值观念分配给人工智能,期望算法系统能够根据符合人类行为的概念对其决策进行解释,而公平是公共价值的核心要素之一,在公共部门内部,如果基层公务员认为算法系统能实现公平价值,会更愿意采纳该算法决策 (Korsgaard et al., 1995; De Graaf & Malle, 2017)。

不同领域的实证研究显示,基层公务员对算法系统公平性的感知会影响他们是否接纳算法系统的决策,如医疗健康、人事招聘与刑事司法等领域 (Dodge et al., 2019; Twizeyimana & Andersson, 2019; Schiff et al., 2022)。人们一般认为算法决策实质上是一种简化程序,其“黑箱”式的计算过程缺少必要信息的披露,导致人们怀疑算法的公平性,从而产生“算法厌恶”而拒绝接受算法决策建议。算法决策的透明度可以通过 XAI 提供有关系统结果和决策过程的解释来实现提高“黑箱”的可访问性 (Shin, 2021),理解算法决策原理及逻辑,为人们审查证据与干预纠偏提供了可能 (Meske et al., 2022),提升了决策公平性感知 (Miron et al., 2021),进而提高他们的算法决策建议的采纳意愿 (Dodge et al., 2019)。基于此,本文提出如下假设。

H3a: 算法的可解释性会正向影响基层公务员对算法决策的公平感知,从而促进基层公务员对算法决策的采纳。

技术信任指个人依靠技术实现任务目标,愿意承担技术后果的意愿,是人

机信任关系中重要的组成部分之一（Paravastu et al. , 2016）。算法的内部工作机制通常无法被直接访问，机器代理人如何学习和作出决策，人类行为者可能对此无法理解，而人们往往不会信任他们所不理解的系统（Giest & Grimmelikhuijsen, 2020）。XAI 有两条路径可以增强人类的技术信任。一是 XAI 提供一致性信息帮助人们比较和评估决策结果，降低人们对于技术后果的风险感知，增强人们对算法系统的技术信任；二是以认知契合为核心的解释可以提高人类的可理解性，加强技术有用性和决策质量的收益感知，使人们相信技术会产生积极结果（Meske et al. , 2022；McKnight et al. , 2011）。实证研究表明，在公私部门的算法系统中，使用者对算法系统的技术信任水平越高，越偏好采纳算法决策建议（Loh et al. , 2022；Selten et al. , 2023）。基于此，本文提出如下假设。

H3b：算法的可解释性会正向影响基层公务员对算法决策的技术信任，从而促进基层公务员对算法决策的采纳。

综合上述假设，图 1 展示了变量之间的关系。

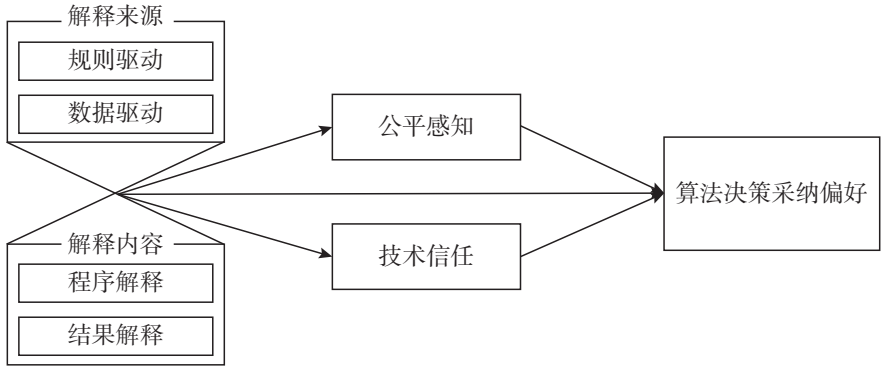


图 1 理论框架

资料来源：作者自制。

三、研究设计

（一）数据来源与实验设计

本研究通过见数平台进行在线析因调查实验。析因调查实验法常用于研究多个自变量对于因变量的影响，能够分析各自变量对因变量的主效应及各个变量间的交互效应（Wang & Zhang, 2023）。本研究将被试者随机分配到 5 个（1 个对照组和 4 个控制组）不同组，累计发放问卷 750 份，回收有效问卷 521 份，回收率 69.47%。受客观条件限制，本研究选取中东部城市进行问卷调查，虽无法获得覆盖全国区域范围样本数据，但抽样时考虑了性别、年龄、区域、层级、职级等因素（见表 1）。总体而言，样本具有一定的代表性。

表 1 描述性统计

分类变量	测量	比例 (%)
可解释性	对照组 (无干预)	20
	程序 * 规则	20
	程序 * 数据	20
	结果 * 规则	20
	结果 * 数据	20
性别	女性	48.21
	男性	51.79
领导岗位与否	是	20.53
	否	79.47
教育部门工作与否	是	15.27
	否	84.73
专业要求	无要求	32.94
	理工医学类	14.56
	文史哲类	21.48
	社会科学类	20.76
	其他	10.26
连续变量	取值范围	平均值 (标准差)
算法采纳偏好	0 - 1	0.521 (0.5)
公平感知	1 - 5	3.622 (1.031)
技术信任	1 - 5	3.609 (0.791)
年龄	23 - 65	33.743 (7.036)
教育水平	1 - 5	3.181 (0.603)
大学层次	1 - 7	3.714 (1.009)
行政级别	1 - 10	2.907 (2.02)
政治身份	1 - 5	1.907 (1.177)
公共部门工作年限	0 - 47	9.069 (7.499)
风险偏好	1 - 6	3.423 (3.345)

资料来源：作者自制。

(二) 变量可操作化

因变量为算法决策采纳偏好。该变量参考已有研究对算法决策态度的测量方法 (Alon-Barkat & Busuioc, 2023; Miller & Keiser, 2021)。设计税务系统的情景, 有两个考量: 一是理论层面。已有研究已检验税务部门算法应用的影响, 该部门基于庞大的税收业务和多元风险监管需求, 通常采用算法决策系统进行税务监管与服务 (Faúndez-Ugalde et al., 2020; 樊博、李晶晶, 2024)。二是实务层面。当前中国税务、海关、城管等监管部门正部署算法系统协助监管, 其中国家税务总局正自上而下部署治理虚开增值税发票的算法系统, 该系统

基于以往增值税发票的查处案例，对企业虚开增值税发票风险进行评分并定级（红、黄、蓝），协助税务稽查人员进行执法。因此，本研究以税务为场景，具有理论对话与现实对照的价值，具体而言以我国“金税工程”为背景，进行干预后采用五尺度李克特量表询问被试者“请问您在多大程度上愿意采纳‘金税五期’税务系统的决策？”。

核心自变量为算法可解释性。参考已有研究实验题（Wang et al., 2023; Burton et al., 2020），将算法可解释性分为四类情景：规则驱动的程序解释型、规则驱动的结果解释型、数据驱动的程序解释型与数据驱动的结果解释型。规则驱动算法的干预材料为“金税五期”的税收审批系统，该系统由软件开发人员依据法律法规、行政专业知识等正式规则编写代码。数据驱动算法的干预材料为“金税五期”的税务稽查系统，该系统依托大数据挖掘技术，基于各类公开和业务部门内部数据对特定行政行为进行数据挖掘。程序解释算法的干预材料为说明法律法规规定的内容，如“根据《中华人民共和国发票管理办法》第十五条、税总函〔2019〕64号第一条、第二条规定”。结果解释算法的干预材料为说明行政行为的具体原因，如“申请人存在‘法定代表人或财务负责人未实名’情形”。接着，采用2（规则驱动与数据驱动）×2（程序解释与结果解释）分组^①。随后针对四个控制组进行操纵有效性检验，对解释来源的操纵有效性项目，“请问您认为‘金税五期’税务系统的运作逻辑主要强调的是1. 规则；2. 数据”，以及对解释内容的操纵有效性项目，“请问您认为‘金税五期’税务【系统界面】给出的信息主要表达的是1. 对算法遵循的程序作出解释；2. 对算法输出的结果作出解释”。

中介变量为两个。一为公平感知。参考 Miller 和 Keiser（2021）开发的公平感知量表，“算法系统在评估时能公平对待不同地区和群体”“算法系统的建议是基于情况事实，而不是偏见或意见”。二为技术信任。参考 McKnight 等人（2011）开发的技术信任量表，测量题项为：“算法系统友好、可靠”“我对算法系统的意图、结果输出持怀疑态度”（反向测量）。采用五尺度李克特量表，求平均值，公平感知与技术信任变量均为连续变量。

控制变量分为三类。

一为人口统计学变量。个体特征会影响人们对算法技术的态度。男性和受教育水平高的人更能感知算法公平，对算法技术态度更加积极（Van Berkel et al., 2021）。1. 性别：男；女。2. 受教育水平，用两个题项测量。（1）学历：高中（中专）及以下；大专；本科；硕士；博士。（2）大学层次：清华大学或北京大学；“985”高校；“211”高校；普通本科院校；海外高校；大专及以下以及其他。3. 年龄，年轻人更易认为算法是不公平的（Lee et al., 2020），年

^① 由于篇幅限制，未呈现干预材料的完整内容，若读者想了解可以联系作者索取。

龄为连续变量。

二为社会学变量。1. 政治面貌：中共党员；非中共党员。2. 社会地位。更高社会地位的人更容易接受新技术（Shin, 2020），使用职务级别与岗位类型测量。（1）职务级别：办事员；科员；副科级；正科级；副处级；正处级；副（厅）局级；正（厅）局级；省部级及以上；其他。（2）岗位类型：领导岗位；非领导岗位。

三为个人风险偏好。风险偏好会影响个人创新行为（Alabede & Zainal, 2011），本研究采用中国家庭追踪调查（CFPS）的风险偏好量表，分6个等级，“1”为最低，“6”为最高。

四、实证结果

（一）主效应与交互效应检验：XAI与算法建议采纳偏好

首先，检验实验干预的有效性。本研究设计了两道问题检验受访者能否通过干预文本识别有效信息，问题1：“请问您认为‘金税五期’税务系统的运作逻辑主要强调的是规则，还是数据？”问题2：“请问您认为‘金税五期’税务【系统界面】给出的信息主要表达的是对算法遵循的程序作出解释，还是对算法输出的结果作出解释？”在对521份问卷进行清理后，有498份问卷正确填答了问题1，占比为94.54%；有480份问卷正确填答了问题2，占比为89.55%。两个问题的正确填答人数均远大于错误填答人数，这可视为实验干预有效，从而得到有效接受干预的样本480份。为检验问卷各组随机分配的有效性，本研究对有效接受干预样本控制变量进行了卡方检验，分析结果表明（见表2），各实验组样本在控制变量上不存在显著差异，证明了干预信息满足随机分配要求。

表2 实验组卡方检验

变量	卡方检验指标	p 值
性别	chi2 (4) = 0.222	0.876
年龄	F = 0.43	0.789
教育水平	chi2 (16) = 14.017	0.597
政治身份	chi2 (4) = 2.852	0.583
职务级别	chi2 (28) = 25.968	0.575
领导职位	chi2 (4) = 0.231	0.994
风险偏好	F = 0.68	0.606

资料来源：作者自制。

其次，检验算法可解释性对算法决策建议采纳的影响。表3显示了回归结果。实证结果显示，模型（1）—模型（4）的 R^2 的F检验显著，表明该模型整体具有解释力，稳健性说明算法可解释性对基层公务员的算法决策采纳偏好存在效应。其中，实验组3（数据驱动×程序解释）和实验组4（数据驱动×结果解释）的算法采纳偏好显著高于对照组，支持了H1b及H1d。实验组1（规则驱动×程序解释）和实验组2（规则驱动×结果解释）回归系数不显著，H1a及H1c未得到实证支持。这初步证明了H1得到部分支持，当算法提供可被理解的解释时，人们会偏好采纳算法决策的建议（Shin, 2020）。作为“决策助手”，数据驱动的解释方法能够从输入与输出之间的关系提供全面的解释，补充人类无法直接观察到的隐性知识，为传统官僚提供“中立客观”外部证据，符合公共行政的询证决策需求（Alon-Barkat & Busuioc, 2023）。Z市T区M街道投入构建以智慧指挥中心为中枢的基层智慧治理体系，渗透民生保障、街区管理、企业服务等多项事务，数据自动更新、自主解释。该街道办一位负责人表示：

“企业实时信息数据能根据企业内部需求，比对外部数据，为我们根据定量分析回应企业需求作出扶持。”（访谈资料 M20230813）

表3 基准模型结果

	(1) 采纳算法		(2) 采纳算法		(3) 采纳算法		(4) 采纳算法	
自变量：算法可解释性								
(对照组：无干预)								
实验组 1（规则驱动的程序解释）	0.11	(0.161)	0.121	(0.161)	0.119	(0.161)	0.123	(0.161)
实验组 2（规则驱动的结果解释）	0.12	(0.161)	0.131	(0.161)	0.133	(0.162)	0.129	(0.161)
实验组 3（数据驱动的程序解释）	0.3 *	(0.161)	0.31 *	(0.162)	0.32 **	(0.162)	0.299 *	(0.162)
实验组 4（数据驱动的结果解释）	0.38 **	(0.161)	0.422 ***	(0.163)	0.423 ***	(0.163)	0.419 **	(0.162)
女性（0 = 男性）			-0.192 *	(0.105)	-0.178 *	(0.106)	-0.151	(0.106)
教育水平			-0.077	(0.082)	-0.104	(0.085)	-0.09	(0.085)
年龄			-0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)
党员（0 = 非党员）					0.156	(0.118)	0.15	(0.118)
职务级别					-0.034	(0.058)	-0.039	(0.057)
非领导（0 = 领导）					-0.078	(0.159)	-0.069	(0.159)
风险偏好							0.11 **	(0.047)
常数	3.58 ***	(0.114)	3.952 ***	(0.291)	4.063 ***	(0.376)	3.668 ***	(0.411)
N	477		476		473		471	
R ²	0.035 **		0.055 **		0.129 **		0.14 ***	

注：括号内为标准误，*** $p < 0.01$ ，** $p < 0.05$ ，* $p < 0.1$ ，下同。

资料来源：作者自制。

规则驱动的 XAI 基于人类专家的知识系统建立，使算法能够模仿人类代理行为。作为“被模仿者”，传统官僚一般会认为自己处理公共事务的能力优于机器，因此将规则驱动的 XAI 归入“有限理性”的范畴，解释效力有所折损，相较于无解释组可能不会对基层公务员的算法采纳偏好产生影响（Alon-Barkat & Busuioc，2023）。因此，基层公务员可能更偏好采纳数据驱动算法系统的决策建议。

最后，检验 XAI 的解释来源与解释内容的交互效应。为了验证 H2，需分别对 H2a、H2b 与 H2c 进行检验。按照实验设计，本研究通过多因素方差分析对算法可解释性的解释来源和解释内容两个因素及其交互项的影响进行探究，结果如表 4 所示。解释来源方面，规则驱动和数据驱动的算法采纳偏好存在显著性差异（ $F = 4.12, p < 0.05$ ），实证结果推翻 H2a，表明解释来源于数据驱动的 XAI 比规则驱动的 XAI 更能促进传统官僚对算法决策的采纳偏好。数据驱动的 XAI 被认为具有超越人类的技术能力，源自数据的解释更加客观论证决策合理性，不需要复杂行政规则的推演（Dodge et al.，2019；Lee et al.，2019）。相较于规则驱动，数据驱动算法系统更能告知人类决策者隐藏的因果关系，提高决策透明度的同时也缩短了接受时间，降低了传统官僚可理解的“门槛”，有助于减轻反复排查的行政负担。Z 市 H 区税务局一位业务人员表示：

“（稽查）系统给出的仅仅是一个风险疑点的方向，稽查部门税务人员再根据系统提供的风险疑点进行人工检查，有时候疑点很不明确，我们比对排查起来很辛苦。”（访谈资料 P20240409）

因此，相较于规则驱动的 XAI，数据驱动的 XAI 对促进基层公务员对算法决策的采纳偏好的贡献更大。

表 4 解释来源和解释内容的作用

方差来源	SS	df	MS	F	p	Adjust R ²
解释来源	5.063	1	5.063	4.12	0.043	0.004***
解释内容	0.203	1	0.203	0.16	0.685	
解释来源 × 解释内容	0.123	1	0.123	0.10	0.752	
残差	486.79	396	1.229			

资料来源：作者自制。

解释内容方面，程序解释和结果解释的算法采纳偏好未表现出显著性差异（ $F = 0.16, p > 0.1$ ），H2b 不得证。关于交互效应，方差分析中的交互项不显著（ $F = 0.10, p > 0.1$ ），H2c 没有通过检验。因此，H2 部分通过实证检验，不同解释来源对于可解释性影响算法采纳偏好的贡献存在差异，不同解释内容及二者交互项的贡献水平并无显著差异。

（二）中介效应分析：公平感知与技术信任的中介作用

本部分对 H3a 和 H3b 进行检验，验证不同解释类型下公平感知和技术信任

是否为 XAI 影响算法决策采纳偏好的中介机制。将解释类型变为虚拟变量（无解释 = 0，特定类型 XAI = 1），然后对各路径分别进行 sobel 中介效应检验，结果如表 5 所示。

表 5 中介效应模型①

路径	系数 a	系数 b	系数 c (总效应)	系数 a × b (间接效应)	系数 c' (直接效应)	Sobel Z	作用 类型
1. 数据驱动 × 结果解释 - 公平感知 - 算法采纳	0.081 **	0.887 ***	0.402 **	0.072 **	0.330 **	1.925	部分 中介
2. 数据驱动 × 结果解释 - 技术信任 - 算法采纳	0.097 **	0.502 ***	0.402 **	0.049 **	0.354 **	1.933	部分 中介

注：人口统计学变量、社会学变量与风险偏好等因素皆被纳入模型。

资料来源：作者自制。

首先，对 4 个解释类型下公平感知的中介作用进行检验。实证结果显示，“数据驱动 × 结果解释 - 公平感知 - 算法采纳”的路径成立，即数据驱动的结果解释型 XAI，仅在提供结果解释时通过公平感知对基层公务员算法采纳偏好产生影响，且公平感知表现为部分中介效应（ $Z = 1.925$ ， $p < 0.05$ ）。而“规则驱动 × 程序解释 - 公平感知 - 算法采纳”“规则驱动 × 结果解释 - 公平感知 - 算法采纳”“数据驱动 × 程序解释 - 公平感知 - 算法采纳”的路径均没有得到验证。因此，H3a 得到部分验证，数据驱动的结果解释型算法系统会正向影响基层公务员对算法决策的公平感知，从而促进基层公务员对算法决策的采纳。人们普遍认为，数据驱动的算法决策具有不受情感因素影响的中立属性（Lee, 2018）。“技术中立”的数据是此类 XAI 作出解释的主要来源，人类专家的个人经验、直觉和专业知识等相对主观的因素在其中的作用很小，相较于人类参与制定的规则，减少了决策“不道德”的可能性（Dingelstad et al., 2022）。数据驱动的结果解释型 XAI 并不直接展露算法运行的背后逻辑，而是用简化的数据重要性特征或因果关系来证明决策合理性。基于有效性统计结果的解释在为用户理解模型时提供了更多的客观支持，强化用户对算法公平性的感知，从而促进基层公务员对算法决策的采纳偏好（Dodge et al., 2019）。

其次，对 4 个解释类型下技术信任的中介作用进行检验。实证结果显示，“数据驱动 × 结果解释 - 技术信任 - 算法采纳”的路径成立，即数据驱动解释类型下，XAI 在结果解释情况下通过技术信任对基层公务员的算法采纳偏好产生影响，且技术信任表现为部分中介效应（ $Z = 1.933$ ， $p < 0.05$ ）。而“规则驱动 × 程序解释 - 技术信任 - 算法采纳”“规则驱动 × 结果解释 - 技术信任 - 算法

① 由于篇幅限制，仅呈现存在中介效应的结果，无中介效应的没有呈现，若读者想了解更多可联系作者索取。

采纳”“数据驱动×程序解释－技术信任－算法采纳”的路径均没有得到验证。因此，H3b 得到部分验证，即数据驱动的结果解释型算法系统会正向影响基层公务员对算法决策的技术信任，从而促进基层公务员对算法决策的采纳。机器算法推断的过程非常复杂且不可见，一般业务人员即便了解决策过程是基于大量数据挖掘与比对进行的，通常也难以理解算法代码主导的结果究竟如何产生，仍然需要进一步寻找信息解释而作出决定（Alon-Barkat & Busuioc, 2023）。“数据驱动×结果解释”对数据结果的描述提供了更加清晰的线索，直接回答了“为什么是这个结果”，提供了更高水平的可见性解释，这种可见性的增加会促进用户对技术系统的信任（Lee et al., 2020; Shin, 2021），也满足了人类对于机器更快辅助决策达成的使用期望，强化感知有用性和使用信心（Burton et al., 2020），从而促进基层公务员对算法决策的采纳偏好。

五、总结

本研究回应了实务界和学界对于关注算法可解释性的呼吁，使用问卷实验验证了算法可解释性类型与人类采纳行为间的关系，得出以下结论。

第一，算法的可解释性能够促进传统官僚对算法决策的采纳偏好，H1 得到部分验证。H1b 与 H1d 得到实证支持，即规则驱动的结果解释型、数据驱动的结果解释型较无解释组，会显著提升基层公务员对算法决策建议的采纳偏好。本研究从算法技术本身出发，探究“人”与“机器”的互动过程与结果，以确定技术的哪些属性值得人们信赖。研究发现，要使基层公务员根据算法决策建议采取准确的行动，必须使其理解算法如何以及为何做出推断。具有可见性解释机制的模型，能够增强传统官僚对算法决策的理解，合理的解释说明不仅是行政决定正当性和合法性的重要来源，也是人机委托－代理关系建立良性互动的前提条件之一（Shin, 2021）。当前复杂的机器模型已达到了前所未有的性能水平，但如果只提升机器的技术性能，系统将变得越来越不透明，人类对此越来越难以理解（Alon-Barkat & Busuioc, 2023; Li et al., 2020）。人类决策者需要 XAI 来监督和控制算法系统，确保公共部门决策的合法性。为避免人们对算法系统越来越多的质疑，研究应更多关注可解释性纠正人工智能的不透明缺陷，为解释算法官僚决策构建完整的链条，建构“人”与“机器”（算法）的关系。

第二，不同解释来源的算法决策系统具有不同的解释效力，H2b 得到证否，即数据驱动的 XAI 比规则驱动的 XAI 更能促进传统官僚对算法决策的采纳偏好。由于规则驱动的算法决策系统泛化能力有限，加上近年来大规模数据的积累和深度学习等计算能力的发展，基于大语言模型的生成式人工智能广受关注，政府部门越来越多地使用数据驱动算法系统来支持公共决策（Wang et al., 2023）。数据驱动的 XAI 拥有大规模并行协作处理能力、强大的自我组织和学习能力，

能精准快速地在海量数据中发现各变量参数之间的相关性，模拟人类思维同时超越人类认知，自主作出决策并附以客观的证据解释，可能带来“算法欣赏”，从而促进基层公务员采纳算法决策的建议。本研究的实证结果表明，规则驱动算法系统并不必然影响传统官僚的算法采纳偏好，这可能存在两个方面的原因。一方面，解释有效性不仅依赖于技术使用，也取决于人类与机器之间共同知识的契合程度（Wang et al., 2023）。这一前提对使用者设定了一个潜在的“准入规则”，对于熟悉规则的传统官僚来说，他们会认为自己处理公共事务的能力要优于机器，并不愿意采纳算法建议。另一方面，规则驱动的 XAI 目前只能给出初步推断，现实应用场景仍需要传统官僚进一步审查并作出决策。由于人类的“认知惰性”，人们可能不愿意参与需要努力的认知心理过程（Alon-Barkat & Busuioc, 2023）。综上所述，基层公务员可能对数据驱动的 XAI 寄予更大的期待。

第三，H3a 与 H3b 得到部分验证，可解释性与算法决策采纳偏好间存在两个中介路径，即“数据驱动 × 结果解释 - 公平感知 - 算法采纳”与“数据驱动 × 结果解释 - 技术信任 - 算法采纳”。首先，数据驱动的结果解释型算法决策系统影响了基层公务员的公平感知，进而提升他们对算法决策建议的采纳偏好。相较于人类参与制定的规则，基层公务员可能认为数据驱动的算法系统会产生更客观、更公平的决策，算法系统能够克服人类的偏见和认知局限，作出平等、一致的决策，符合人们对于人工智能实现公共价值的社会期望（De Graaf & Malle, 2017；Ananny & Crawford, 2018）。其次，这也验证了 TAM2 技术信任变量的适用性，当 XAI 提供的推论可理解时，人们会倾向信任算法的技术能力，强化技术采纳意愿。相较于规则驱动的 XAI，数据驱动的 XAI 具有超越人类的数据处理和预测的能力，应用场景更加广泛，人们对该类算法系统表现出更高水平的技术有用性与技术信任感知，进而强化了他们的采纳意愿（Shin, 2020, 2021；Meske et al., 2022）。这些发现揭示了传统官僚对具有不同解释性能的技术感知的不同，人们在人机互动中的情感态度可能是未来研究的重要议题之一。

本研究对于推进数字政府建设具有一定的实践意义。一方面，公共部门在设计、实施和管理算法系统过程中，应建设以人为中心的可解释算法系统。本研究表明，算法透明并不等同于算法可知，算法透明不仅仅是公开源代码，更重要的是算法公开信息的可用性，应以用户可理解为基础，实现算法模型与人类认知的对应转化，否则不可理解的代码依然无法打开“黑箱”。当前《关于加强互联网信息服务算法综合治理的指导意见》《网络数据安全条例（草案）》等均强调算法公开透明作为指导，鼓励社会各界开放源代码。但在公共部门部署这些算法过程中，应更进一步创建基于具体场景的“用户友好型”政务服务系统，通过算法解释的不同层面，为决策结果的可视化作出努力，以提高决策的可解释性。如以文字、图表、动画等人类可直观理解的方式，阐明算法

决策系统以何种方式得到输出结果,通过“高清”解释引导人类决策者高效作出正确决策(王贵,2021)。另一方面,应关注基层公务员对算法系统的心理感知。新事物部署会给组织带来阵痛,实务中随着算法系统在不同部门的应用,算法出错的情况陆续出现,基层公务员开始隐形抵制算法官僚决策,要求控制算法系统的决策(Peeters,2020)。基于算法-人类二元决策视角,算法行政不仅要关注技术性能,还应关注机器规训人类过程与人机兼容性路径。如果算法决策能够被人类理解,算法决策在行政领域的渗透更容易被基层公务员接受(张会平,2021)。因此,人们需要关注传统官僚使用算法系统的心理感知,实现机器“智能决策”与人类“自由裁量权”的融合。

本研究还存在一定的局限性。一是样本容量和覆盖范围存在局限。本研究样本并未覆盖中国西部地区与农村地区,基层公务员对算法系统熟悉程度存在显著区域性差异,未来研究需扩大调查的地域范围,检验推论的普遍适用性。二是缺少结构化的机制分析。对人机互动的理解还需要进一步考虑个人、组织和文化因素。三是理论模型的延展度尚且不足。有关算法解释权的具体内容、公开程度、使用方式等相关问题仍然存在很大争议和讨论空间(张凌寒,2020)。四是实验操纵有待进一步完善。要做到正确干净的因果推断,需要对问卷材料采用最小的操纵,这种操纵在本研究中还存在改进的空间。未来研究可以设计更多情景化、最小操纵的实验设计,扩展可解释性以及人工智能技术的其他因果性问题,经过严格因果推断验证算法可解释性不同类型对传统官僚行为的影响。

参考文献

段哲哲(2021). 控制算法官僚:困境与路径. 电子政务, (12): 2-16.

Duan, Z. Z. (2021). Control Algorithmic Bureaucracy: Dilemmas and Paths. *E-Government*, (12): 2-16. (in Chinese)

樊博, 李晶晶(2024). 算法透明何以提升公众的政府信任? 公共行政评论, 17(1): 4-24+196.

Fan, B., & Li, J. J. (2024). Can Algorithmic Transparency Enhance Citizen Trust in Government? *Journal of Public Administration*, 17(1): 4-24+196. (in Chinese)

刘特, 郑跃平, 曹梦冰(2024). 算法行政:文献述评与研究展望. 公共行政评论, 17(1): 84-104+198.

Liu, T., Zheng, Y. P., & Cao, M. B. (2024). Algorithmic Administration: Literature Review and Research Prospects. *Journal of Public Administration*, 17(1): 84-104+198. (in Chinese)

权一章, 黄晖(2024). 从街头官僚到数字官僚:信息技术应用对一线行政效果的影响及机制. 公共行政评论, 17(4): 73-94+197.

Quan, Y. Z., & Huang, H. (2024). From “Street-Level Bureaucrat” to “Digital Bureaucrat”: The Impact of Information Technology Applications on Frontline Administrative Effectiveness and Its Mechanisms. *Journal of Public Administration*, 17(4): 73-94+197. (in Chinese)

沈鹏熠, 李金雄, 万德敏(2023). “以情动人”还是“以理服人”? 人工智能聊天机器人角色对顾客情感依恋的影响研究. 南开管理评论, <https://kns.cnki.net/kcms2/detail/12.1288.F.20230613.1755.004.html>

Shen, P. Y., Li, J. X., & Wan, D. M. (2023). “Moving People by Affection” or “Convince People by Reasoning”? Research on the Influence of Chatbots’ Role on Customers’ Emotional Attachment. *Nankai Business Review*, <https://kns.cnki.net/kcms2/detail/12.1288.F.20230613.1755.004.html>. (in Chinese)

苏宇(2022). 优化算法可解释性及透明度义务之诠释与展开. 法律科学(西北政法学报), 40(1): 133-141.

Su, Y. (2022). An Interpretation and Specification of the Obligations of Optimizing the Explainability and

- Transparency of Algorithm. *Science of Law (Journal of Northwest University of Political Science and Law)*, 40(1): 133–141. (in Chinese)
- 汤志伟, 韩啸 (2023). 动态能力视角下政府数字化转型的影响机制——来自混合研究的发现. *湖湘论坛*, 36(2): 102–113.
- Tang, Z. W., & Han, X. (2023). On Influencing Mechanisms of Government Digital Transformation from a Dynamic Capability Perspective: Findings from Mixed Methods Research. *Huxiang Forum*, 36(2): 102–113. (in Chinese)
- 王贵 (2021). 算法行政的兴起、挑战及法治化调适. *电子政务*, (7): 2–14.
- Wang, G. (2021). Rise, Challenges, and Legal Adaptation of Algorithmic Administration. *E-Government*, (7): 2–14. (in Chinese)
- 吴进进, 何包钢 (2023). 算法科层制的兴起及其形态. *社会学研究*, 38(6): 40–60+227.
- Wu, J. J., & He, B. G. (2023). The Rise of Algorithmic Bureaucracy and Its Typology. *Sociological Studies*, 38(6): 40–60+227. (in Chinese)
- 吴田, 张帆 (2020). 机关事务管理与数字政府建设——美国联邦总务署(GSA)实践研究及启示. *江苏师范大学学报(哲学社会科学版)*, 46(3): 90–101+124.
- Wu, T., & Zhang, F. (2020). Office Management and Digital Governance Construction: Practical Research and Enlightenment of General Services Administration (GSA). *Journal of Jiangsu Normal University (Philosophy and Social Sciences Edition)*, 46(3): 90–101+124. (in Chinese)
- 许多奇, 蒋龙宇 (2023). 多元共治视域下银行自动化决策中的个人信息保护. *贵州大学学报(社会科学版)*, 41(5): 77–86.
- Xu, D. Q., & Jiang, L. Y. (2023). Protection of Personal Information in Automated Banking Decision-Making from the Perspective of Multifaceted Co-Governance. *Journal of Guizhou University (Social Sciences Edition)*, 41(5): 77–86. (in Chinese)
- 张会平 (2021). 算法嵌入政府科层制的路径、影响与风险应对. *广西师范大学学报(哲学社会科学版)*, 57(3): 52–63.
- Zhang, H. P. (2021). Path, Influence and Risk Response of Algorithm Embedded in Government Bureaucracy. *Journal of Guangxi Normal University (Philosophy and Social Sciences Edition)*, 57(3): 52–63. (in Chinese)
- 张静静 (2023). 电子眼式自动化行政执法现状及其侵权救济. *东南大学学报(哲学社会科学版)*, 25(S1): 129–134.
- Zhang, J. J. (2023). The Current Situation of Electronic-Eye Automated Administrative Law Enforcement and Its Infringement Remedy. *Journal of Southeast University (Philosophy and Social Sciences Edition)*, 25(S1): 129–134. (in Chinese)
- 张凌寒 (2020). 算法自动化决策与行政正当程序制度的冲突与调和. *东方法学*, (6): 4–17.
- Zhang, L. H. (2020). How to Resolve the Conflicts of Algorithm Automated Decision-Making and Due Process. *Oriental Law*, (6): 4–17. (in Chinese)
- Ahn, M. J., & Chen, Y. C. (2022). Digital Transformation Toward AI-Augmented Public Administration: The Perception of Government Employees and the Willingness to Use AI in Government. *Government Information Quarterly*, 39(2): 101664.
- Alabede, J. O., Zainal Affrin, Z., & Md Idris, K. (2011). Individual Taxpayers' Attitude and Compliance Behaviour in Nigeria: The Moderating Role of Financial Condition and Risk Preference. *Journal of Accounting and Taxation*, 3(5): 91–104.
- Alon-Barkat, S., & Busuioc, M. (2023). Human-AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice. *Journal of Public Administration Research and Theory*, 33(1): 153–169.
- Alt, R. (2018). Electronic Markets and Current General Research. *Electronic Markets*, 28: 123–128.
- Ananny, M., & Crawford, K. (2018). Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability. *New Media & Society*, 20(3): 973–989.
- Aysolmaz, B., Müller, R., & Meacham, D. (2023). The Public Perceptions of Algorithmic Decision-Making Systems: Results from a Large-Scale Survey. *Telematics and Informatics*, 79: 101954.
- Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018). ‘It’s Reducing a Human Being to a Percentage’ Perceptions of Justice in Algorithmic Decisions. In *Proceedings of the 2018 Chi Conference on Human Factors in Computing Systems* (pp. 1–14).
- Burton, J. W., Stein, M. K., & Jensen, T. B. (2020). A Systematic Review of Algorithm Aversion in Augmented

- Decision Making. *Journal of Behavioral Decision Making*, 33(2): 220–239.
- Choung, H., David, P., & Ross, A. (2023). Trust in AI and Its Role in the Acceptance of AI Technologies. *International Journal of Human-Computer Interaction*, 39(9): 1727–1739.
- de Fine Licht, J. (2011). Do We Really Want to Know? The Potentially Negative Effect of Transparency in Decision Making on Perceived Legitimacy. *Scandinavian Political Studies*, 34(3): 183–201.
- Dingelstad, J., Borst, R. T., & Meijer, A. (2022). Hybrid Data Competencies for Municipal Civil Servants: An Empirical Analysis of the Required Competencies for Data-Driven Decision-Making. *Public Personnel Management*, 51(4): 458–490.
- De Graaf, M. M., & Malle, B. F. (2017). How People Explain Action (And Autonomous Intelligent Systems Should Too). In *2017 AAAI Fall Symposium Series*.
- Dodge, J., Liao, Q. V., Zhang, Y., Bellamy, R. K., & Dugan, C. (2019). Explaining Models: An Empirical Study of How Explanations Impact Fairness Judgment. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (pp. 275–285).
- Faúndez-Ugalde, A., Mellado-Silva, R., & Aldunate-Lizana, E. (2020). Use of Artificial Intelligence by Tax Administrations: An analysis Regarding Taxpayers' Rights in Latin American Countries. *Computer Law & Security Review*, 38: 105441.
- Fernandez, A., Herrera, F., Cordon, O., Jose Del Jesus, M., & Marcelloni, F. (2019). Evolutionary Fuzzy Systems for Explainable Artificial Intelligence: Why, When, What for, and Where to? *IEEE Computational Intelligence Magazine*, 14(1): 69–81.
- Giest, S., & Grimmelikhuijsen, S. (2020). Introduction to Special Issue Algorithmic Transparency in Government: Towards a Multi-Level Perspective. *Information Polity*, 25(4): 409–417.
- Janssen, M., Hartog, M., Matheus, R., Yi Ding, A., & Kuk, G. (2022). Will Algorithms Blind People? The Effect of Explainable AI and Decision-makers' Experience on AI-supported Decision-making in Government. *Social Science Computer Review*, 40(2): 478–493.
- Korsgaard, M. A., Schweiger, D. M., & Sapienza, H. J. (1995). Building Commitment, Attachment, and Trust in Strategic Decision-Making Teams: The Role of Procedural Justice. *Academy of Management Journal*, 38(1): 60–84.
- Lee, M. K. (2018). Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management. *Big Data & Society*, 5(1): 1–16.
- Lee, M. K., Jain, A., Cha, H. J., Ojha, S., & Kusbit, D. (2019). Procedural Justice in Algorithmic Fairness: Leveraging Transparency and Outcome Control for Fair Algorithmic Mediation. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW): 1–26.
- Lee, M. K., Grgić-Hlača, N., Tschantz, M. C., Binns, R., Weller, A., Carney, M., & Inkpen, K. (2020). Human-Centered Approaches to Fair and Responsible AI. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–8).
- Li, X. H., Cao, C. C., Shi, Y., Bai, W., Gao, H., Qiu, L., ... & Chen, L. (2020). A Survey of Data-driven and Knowledge-aware Explainable AI. *IEEE Transactions on Knowledge and Data Engineering*, 34(1): 29–49.
- Loh, H. W., Ooi, C. P., Seoni, S., Barua, P. D., Molinari, F., & Acharya, U. R. (2022). Application of Explainable Artificial Intelligence for Healthcare: A Systematic Review of the Last Decade (2011–2022). *Computer Methods and Programs in Biomedicine*, 226: 107161.
- Doshi-Velez, F., & Kim, B. (2017). Towards a Rigorous Science of Interpretable Machine Learning. *ArXiv preprint ArXiv*, 1702.08608.
- Madan, R., & Ashok, M. (2023). AI Adoption and Diffusion in Public Administration: A Systematic Literature Review and Future Research Agenda. *Government Information Quarterly*, 40(1): 101774.
- Mahmud, H., Islam, A. N., Ahmed, S. I., & Smolander, K. (2022). What Influences Algorithmic Decision-Making? A Systematic Literature Review on Algorithm Aversion. *Technological Forecasting and Social Change*, 175: 121390.
- Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust In a Specific Technology: An Investigation of Its Components and Measures. *ACM Transactions on Management Information Systems (TMIS)*, 2(2): 1–25.
- Meske, C., Bunde, E., Schneider, J., & Gersch, M. (2022). Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities. *Information Systems Management*, 39(1): 53–63.
- Miller, S. M., & Keiser, L. R. (2021). Representative Bureaucracy and Attitudes Toward Automated Decision Making. *Journal of Public Administration Research and Theory*, 31(1), 150–165.
- Miller, T. (2019). Explanation in Artificial Intelligence: Insights from the Social Sciences. *Artificial Intelligence*, 267: 1–38.

- Miron, M. , Tolan, S. , Gómez, E. , & Castillo, C. (2021). Evaluating Causes of Algorithmic Bias in Juvenile Criminal Recidivism. *Artificial Intelligence and Law*, 29(2) : 111 – 147.
- Nagtegaal, R. (2021). The Impact of Using Algorithms for Managerial Decisions on Public Employees' Procedural Justice. *Government Information Quarterly*, 38(1) : 101536.
- Omrani, N. , Rivieccio, G. , Fiore, U. , Schiavone, F. , & Agreda, S. G. (2022). To Trust or Not to Trust? An Assessment of Trust in AI-based Systems: Concerns, Ethics and Contexts. *Technological Forecasting and Social Change*, 181 : 121763.
- Oswald, M. (2018). Algorithm-Assisted Decision-Making in the Public Sector: Framing the Issues Using Administrative Law Rules Governing Discretionary Power. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2128) : 20170359.
- Paravastu, N. , Ramanujan, S. , & Ratnasingam, P. (2016). Role of Trust In Ecommerce: A Comprehensive Model of Interpersonal and Technology Trust Constructs. In *Encyclopedia of E-Commerce Development, Implementation, and Management* (pp. 1737 – 1747).
- Peeters, R. (2020). The Agency of Algorithms: Understanding Human-Algorithm Interaction in Administrative Decision-Making. *Information Polity*, 25(4) : 507 – 522.
- Schiff, D. S. , Schiff, K. J. , & Pierson, P. (2022). Assessing Public Value Failure in Government Adoption of Artificial Intelligence. *Public Administration*, 100(3) : 653 – 673.
- Schmidt, P. , Biessmann, F. , & Teubner, T. (2020). Transparency and Trust In Artificial Intelligence Systems. *Journal of Decision Systems*, 29(4) : 260 – 278.
- Selten, F. , Robeer, M. , & Grimmelikhuijsen, S. (2023). ‘Just Like I Thought’: Street-level Bureaucrats Trust AI Recommendations if They Confirm Their Professional Judgment. *Public Administration Review*, 83(2) : 263 – 278.
- Shin, D. (2020). User Perceptions of Algorithmic Decisions in the Personalized AI System: Perceptual Evaluation of Fairness, Accountability, Transparency, and Explainability. *Journal of Broadcasting & Electronic Media*, 64(4) : 541 – 565.
- Shin, D. (2021). The Effects of Explainability and Causability on Perception, Trust, and Acceptance: Implications for Explainable AI. *International Journal of Human-Computer Studies*, 146 : 102551.
- Stone, H. S. (1971). Parallel Processing with the Perfect Shuffle. *IEEE Transactions on Computers*, 100(2) : 153 – 161.
- Taeihagh, A. (2021). Governance of Artificial Intelligence. *Policy and Society*, 40(2) : 137 – 157.
- Trope, Y. , Liberman, N. , & Wakslak, C. (2007). Construal Levels and Psychological Distance: Effects on Representation, Prediction, Evaluation, and Behavior. *Journal of Consumer Psychology*, 17(2) : 83 – 95.
- Twizeyimana, J. D. , & Andersson, A. (2019). The Public Value of E-Government: A Literature Review. *Government Information Quarterly*, 36(2) : 167 – 178.
- Van Berkel, N. , Goncalves, J. , Russo, D. , Hosio, S. , & Skov, M. B. (2021). Effect of Information Presentation on Fairness Perceptions of Machine Learning Predictors. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1 – 13).
- Venkatesh, V. , & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39(2) : 273 – 315.
- Wang, G. , Guo, Y. , Zhang, W. , Xie, S. , & Chen, Q. (2023). What Type of Algorithm Is Perceived as Fairer and More Acceptable? A Comparative Analysis of Rule-Driven Versus Data-Driven Algorithmic Decision-Making in Public Affairs. *Government Information Quarterly*, 40(2) : 101803.
- Wang, Y. , & Zhang, J. (2023). Promoting Citizens' Willingness to Participate in Coproduction in Public Service Through Information Frames. *Public Administration*, 102(2) : 479 – 499.
- Williams, M. D. , Rana, N. P. , & Dwivedi, Y. K. (2015). The Unified Theory of Acceptance and Use of Technology (UTAUT): A Literature Review. *Journal of Enterprise Information Management*, 28(3) : 443 – 488.

责任编辑：张书维

英文目录与摘要

JPA Journal of Public Administration, Vol. 17 No. 6, 2024

●SYNOPSIS: Grassroots Governance

Introduction: Policy Implementation in Grassroots Governance and Its Impact Bin Xiao

The Causes and Resolution of the Burden of Grass-roots Government; Based on the Theory of Complex Adaptive Systems Yuanhe Liu & Zhenchao Zhou

Abstract Relieving grass-roots government's burden of an insufficient organizational adaptability to environmental changes is the key to enhancing its governance capacity. CAS theory applied in a view and structural interaction analysis framework explains this phenomenon. Modern governance demand is shaped by the modernized state governance concepts and resonates strongly with the public, thereby creating significant motivation for government action. As grassroots governance tasks increase, overlapping responsibilities and bureaucratic structures lead to incomplete resource allocations that fail to address unforeseen challenges. The proper governance approach is to establish axis-radiation governance. It facilitates the allocation of governance resources through a legalized mechanism, reduces the dominance of "Tiao" over "Kuai", enhances the integration capacity of "Tiao", and transforms grassroots government into a competent regional governance center.

Key Words the Burden of Grass-roots Government; Complex Adaptive Systems; View; Axis-Radial Governance

The Multiple Behavioral Logics of "Comparative Execution" in Grassroots Governance

..... Lei He & Na Tang

Abstract Comparative execution is a typical manifestation of policy implementation in grassroots governance; but it has been criticized for its "formalist" tendency. Previous studies have also adopted a critical attitude towards the comparative execution model, which does encompass multiple implications. This article examines the overall operational context of the bureaucratic system, provides a comprehensive description of the application categories of "comparative execution" in different scenarios, and explores the multiple behavioral logics of comparative execution. This article is based on a multi case analysis of the implementation of precision poverty alleviation policies in Y province. Under pressure situations, comparative execution with power game and ability matching as the core driving forces presents a behavioral logic of "comparative execution type = pressure type hierarchical system + power ability coupling form". According to the power ability coupling situation in a pressure type hierarchical system, comparative execution can be subdivided into four basic types. They are "adaptive comparative execution", "passive comparative execution", "active comparative execution", and "perfunctory comparative execution." They also explore the multiple behavioral logics of comparative execution in different scenarios to enrich the reality of policy execution research.

Key Words Grassroots Government; Comparative Execution; Power Game; Capability Matching

Pressure Extrusion and Structure Utilization: Can Grassroots Governments Promote Tiao-kuai Collaboration from the Bottom-up? Yuanyuan Huang & Yinbin Tang

Abstract Active cooperation between functional departments (“tiao”) and grassroots governments (“kuai”) is the key to policy implementation, despite the difficulties caused by the fragmentation between the two. Existing research tends to focus on high-incentive or high-risk policies as implementation scenarios and on tiao-kuai collaboration initiated by higher-level governments and departments. What’s missing is a view of the implementation scenarios of conventional policies and the proactive role of grassroots government. This article goes beyond the constraints of implementation scenarios and structures to retrieve the collaborative autonomy of grassroots governments in the conventional policy implementation scenario. By analyzing the implementation process of the old reform policy for Street A in County H, we found that the grassroots government has the motivation, space , and ability to promote tiao-kuai collaboration. First, higher government’s weak involvement in conventional policies and the non-exitability of the grassroots government make the latter the first to face the convergence of multiple pressures, creating the need for collaboration. Second, the responsibility linkage in a limited and segmented implementation structure and on-the-ground resources held by grassroots governments contribute to the construction of a space for cooperation. In this space , grassroots governments can leverage information, relationships, and community to enable effective collaboration. In fact, tiao-kuai collaborations are dominated by the grassroots government but hampered by contingent triggering conditions, the instability of the cooperation process, and the limited effectiveness of the cooperation. Exploration of tiao-kuai cooperation, with the grassroots government as the main body, can help to improve the path of tiao-kuai cooperation in the grassroots arena and to recover the “conventional narrative” of policy research in China.

Key Words Grassroots Policy Implementation; Tiao-kuai Relationship; Grassroots Autonomy; the Conventional Policy

Adaptive Empowerment: Community Integration for the Floating Population

..... Hongjun Jiang & Dong Zhang

Abstract China’s rapid economic development has led to a substantial influx of floating populations into its cities, frequently marginalizing them from the established community governance systems. Alleviating the tension arising from the interactions between the floating population and local residents in urban communities has become a crucial challenge for the modernization of grassroots community governance. Through a case study of Community S, this paper uses the logic of adaptive empowerment to show that new institutions and policies in the process of creating integrated communities are incrementally layered over existing systems such as household registration. Supported by three mechanisms—organizational empowerment, professional enablement, networked interaction—the interaction between new and old institutions achieves an institutional balance between the empowerment of community participation and the compromise with household registration status, thereby fostering a peaceful new form of community integration for the floating population. This study deepens our understanding of mobile urbanity and serves as a reference for exploring resilient community governance and the modernization of grassroots governance with Chinese characteristics.

Key Words Adaptive Empowerment; Floating Population; Community Integration; Historical Institutionalism

..... Zhezhe Duan & Jiaxin Jiao

Abstract Algorithms are becoming an omnipresent means of public sector decision-making. The new-generation governance model of artificial intelligence has emerged as a significant orientation for the future digital transformation of the public sector. Nevertheless, the “black box” attribute of the current algorithm system has been increasingly questioned, giving rise to the “trust deficit” and “lack of accountability” in algorithm governance. This paper explores the logic of Explainable Artificial Intelligence (XAI) through a 2 (interpretation source: rules vs. data) × 2 (interpretation content: procedures vs. results) survey experiment. In this scenario, the tax department applies the algorithm decision system, and the mechanism of different types of XAI on traditional bureaucrats' adoption of algorithm decision suggestions is examined. The empirical results indicate that, first, the interpretable algorithm system can boost the adoption preferences of grassroots civil servants; and the data-driven program interpretation algorithm system or data-driven result interpretation algorithm system can enhance the adoption preference of grassroots civil servants. Second, data-driven XAI can promote the algorithmic decision adoption preferences of traditional bureaucrats more effectively than rule-driven. Finally, the data-driven result-interpretable algorithmic system can heighten the traditional bureaucracy's perception of fairness and technical trust in the algorithmic system, thereby making them more inclined to adopt algorithmic decisions. These findings imply that we should develop the technical capabilities of algorithmic bureaucrats, provide interpretable types of algorithms to traditional bureaucrats, and explore favorable human-machine interactions to facilitate the smooth implementation of algorithmic systems in the public sector.

Key Words Algorithmic Decision; Interpretability; Artificial Intelligence; Man-machine Interaction; Algorithmic Bureaucracy; Grassroots Civil Servant

● ARTICLES

Why Ambiguous Policy Is Unable to Be Implemented: A Case Study of the Safety Production Standardization Construction Policy of Gas Enterprises in China's A City

..... Zhe Zhu, Jie Li, Qingjun Zhu & Meilian Zhang

Abstract In a risk society, understanding how to implement ambiguous policies is a crucial issue for enhancing national governance performance. Existing research has focused on implementing ambiguous policies in general contexts but has paid insufficient attention to their implementation mechanisms in the field of risk governance. Using process tracing, this single case study explores the failed process of implementing the safety production standardization policy for urban gas enterprises in A City. The findings reveal that policy ambiguity influences implementation strategies in the context of strong accountability. The capability of executors and the attention of leaders act as moderating factors, shaping policy outcomes. Motivated by blame avoidance when regulatory responsibilities and policy tools are clear, competent executors adopt selective implementation strategies to mitigate risks, even without leaders' attention. Conversely, when regulatory responsibilities and policy tools are ambiguous and leaders are inattentive, executors resort to inaction as a blame-avoidance strategy, leading to uncontrolled risks. When leaders are attentive and regulatory responsibilities are clear, but policy tools are ambiguous, incompetent executors adopt a “hot potato” strategy to pass off responsibility. When such shirking is no longer feasible, they resort to superficial implementation, resulting in risk

accumulation. This paper develops a theoretical framework for ambiguous policy implementation from a blame-avoidance perspective, offering insights into the implementation of ambiguous policies in the field of risk governance.

Key Words Ambiguous Policy; Blame Avoidance; Policy Implementation; Risk Management

How Does the Government Evaluate Providers of Elderly Care Services in the Context of High-Quality Development? Construction and Validation of an Analytic Hierarchy Process (AHP) Model

..... Jing Peng & Qiao Yang & Shuai Xiao

Abstract This paper uses general evaluation theories, such as 3E theory and 3D theory, and the AHP to construct a supplier evaluation model from three indicator dimensions—undertaking capacity, financial capacity, organizational construction and management capacity—to examine the technical support of the high-quality development of elderly care services. Through empirical verification of the practice of purchasing elderly care services by the Y City government, it was found that strong service acceptance capability plays a decisive role in becoming a high-quality supplier. The scale and professionalism of competitive social organizations are strengths that should be employed for targeted development measures. Therefore, effective evaluation of suppliers by the government is an important foundation for the high-quality development of elderly care services. The government should strengthen its evaluation responsibility and increase the number of high-quality service suppliers. The policy recommendation for the high-quality development of elderly care services is not only to enhance the government's ability to effectively evaluate suppliers, but also to create an institutional environment that promotes the sustainable development of social organizations and optimizes government social cooperation and governance.

Key Words Government Purchase of Elderly Care Services; Supplier Evaluation Model; Analytic Hierarchy Process (AHP); High Quality Development

Does Technological Empowerment Facilitate Holistic Governance in the Process of County Urbanization? A Qualitative Comparative Analysis Based on 30 Cases Yanling Li & Mengya Chen

Abstract Holistic governance is a new model and trend in urban-rural governance in the new era, providing strategic guidance for achieving high-quality development in China's counties. This study examined 30 typical cases of counties with advanced holistic governance practices. Grounded in the theory of holistic governance and aligned with the practical demands of new urbanization, this paper uses an analytical framework of "holistic governance-technological governance" from the perspective of integrated promotion across population, space, and industry. Using Qualitative Comparative Analysis (QCA), the paper identifies multiple conjunctural causations and effective pathways through which technological empowerment facilitates holistic governance in county contexts. The findings reveal that holistic governance in Chinese counties is the result of interactions among several conditional variables, including technological empowerment, population urbanization, spatial optimization, and industrial upgrading. Configurational analysis identifies four effective pathways: all-factor-driven, "technological empowerment-industrial upgrading" co-promotion, "technological empowerment-population urbanization" dual-driver, and balanced "technology-industry-population." Among these, the all-factor-driven pathway is the most prevalent, highlighting the critical role of technological empowerment in advancing holistic governance. The paper shows the importance of a people-centered development

philosophy, systematic thinking, and context-specific strategies. It also advocates for leveraging technological empowerment to drive high-quality county development, transitioning from fragmented intelligent governance to integrated intelligent governance.

Key Words County Urbanization; Technological Empowerment; Holistic Governance; Configurational Pathways; QCA

How Does the Application of Artificial Intelligence Affect the Willingness of Street-level Bureaucrats to Exercise Discretion? A Survey Experiment Yu Sun, Yi Deng & Linzhi Sun

Abstract The full exercise of discretion by street-level bureaucrats is an important guarantee for promoting the combination of education and punishment in the law enforcement process. However, long-term tracking surveys have found that as artificial intelligence (AI) is increasingly involved in traffic law enforcement traffic police officers are no longer inclined to exercise their discretion. Instead, the “penalties substituting for management” phenomenon has emerged in traffic law enforcement. To test these observations and explain the decrease in street-level bureaucrats’ willingness to exercise their discretion, this study conducted a survey experiment in H province’s L city, with a representative sample of 190 traffic police officers. The study found that: (1) AI intervention reduced the willingness of street-level bureaucrats to exercise discretion. (2) Perceived process complexity played a partial mediating effect. The results show that the application of AI technology has changed the objective conditions of street-level law enforcement. The room for action is being squeezed, and these bureaucrats are measurably reluctant to exercise their discretion. This research provides new insights into street-level bureaucrat discretion. It also provides empirical evidence of new mechanisms and shows how AI changes traffic law enforcement work practices.

Key Words Artificial Intelligence; Street-level Bureaucrat; Discretion; Willingness to Exercise; Survey Experiment

● THEORETICAL REVIEWS

Public Innovation Labs’ Roles, Impacts, and Determinants: A Literature Review

..... Qianli Yuan & Xinping Liu

Abstract Facing increasingly complex public management issues and public service demands, government departments need public innovation to enhance their effectiveness and legitimacy. Public innovation labs have become the new vehicles for the promotion of those innovative outcomes and capabilities. This article reviews research on public innovation labs, summarizing their characteristics, roles, outcomes, and influencing factors. These labs mainly play a role in the early stages of public innovation in problem identification and solution design. They also facilitate technological process innovation and management process innovation. This has led to increased efficiency and public participation in public management. As the existing research has shown, team building, organizational structure, and collaboration networks in public innovation labs are key micro-and meso-level drivers to public sector innovation. However, the macro administrative system and its traditions have imposed certain limitations on public innovation labs. This study provides insights for the practice and research of public innovation labs in China.

Key Words Public Innovation Lab; Innovation Space; Collaborative Innovation; Institutional Design; Literature Review

公共行政评论

双月刊，2008年创刊
第17卷，第6期（总第102期）
2024年12月15日出版

Journal of Public Administration
Bimonthly, Since 2008
Vol.17 No.6
Published in December 2024

主管单位 中华人民共和国教育部
主办单位 中山大学
协办单位 教育部人文社会科学重点研究基地-
中山大学中国公共管理研究中心
广东省行政管理学会

社长 肖滨
主编 朱亚鹏

联系电话 020-84113029 020-84038746

电子邮件 jpachina@163.com

编辑出版 《公共行政评论》编辑部
(广州新港西路135号; 邮编510275)

印刷 广州一龙印刷有限公司

国内发行 广东省报刊发行局

国外发行 中国国际图书贸易总公司
(北京399信箱)

传真 020-84111478

网址 <http://jpa.sysu.edu.cn>

Administrator Ministry of Education of the People's Republic of China

Sponsors Sun Yat-sen University

Supporters Centre for Chinese Public Administration Research, Sun Yat-sen University
Guangdong Public Administration Society

President: Bin Xiao

Editor in Chief: Yapeng Zhu

Tel: 86 20 84113029 86 20 84038746

Fax: 86 20 84111478

Email: jpachina@163.com

Website: <http://jpa.sysu.edu.cn>

Edited by Editorial Office of *Journal of Public Administration*

(NO.135 Xin Gang Xi Road, Guangzhou, China. 510275)

Distributed by China International Book Trading Corporation

(P.O.Box 399, Beijing, China)



刊号 ISSN1674-2486
CN44-1648/D

邮发 国内46-364
代号 国外BM8839

国内外公开发行
国内定价: 20.00元

ISSN 1674-2486

