

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

中文科技期刊数据库(全文版)收录期刊

RCCSE中国核心学术期刊(A)

2025. 4

VOL.18 NO.4

公共行政 评论

JPA
Journal of Public Administration

●专栏：人工智能与政府信任

当 AI 成为“公仆”：政务服务数字人类型对政府信任的影响研究

..... 李姝婷 邱 茜 1

公民更信任公共部门中的“人”还是“人工智能”：一项考虑任务复杂程度
与公共领域的调查实验研究

和思绮 刘炳胜 22

●论文

算法特征披露何以影响公众决策偏好：基于联合实验的实证分析

..... 王 戈 向 丽 张哲君 41

数字技术与多层级治理模式的变迁：以空气污染治理为例

..... 陈 靖 徐 灿 洪 伟 63

农村数字化转型何以悬浮：基于政策执行过程的案例研究	
..... 傅利平 康开荣 冯华超 董永庆	81
基层政府资源分配行为及其多重逻辑研究：基于山东省 A 县的案例分析	
..... 孙宗锋 林婧淇 张晓涵	102
政策组合强度何以影响政策效能：基于大气污染防治政策组合的分析	
..... 张晓杰 刘晓瑜	118
社会组织发展能否有效缓解居民收入差距：基于政府干预门槛效应的实证检验	
..... 刘 蕾 张新亚 何书仪	138
财政纵向失衡会降低地方政府公共服务注意力吗？	
..... 张 原 刘丽君	155

●理论综述

财政压力下的公共组织削减管理：一个文献综述	 陈 馨	173
-----------------------	-----------	-----

●英文目录与摘要		196
----------	-------	-----

公民更信任公共部门中的“人”还是“人工智能”： 一项考虑任务复杂程度与公共领域的调查实验研究

和思绮 刘炳胜*

【摘要】人工智能的快速发展正在重塑公共管理的运作模式与职能范畴，但相关理论建构与实证研究的发展相对滞后。近年来，人工智能在公共管理领域的广泛应用引发了大量争议，其应用价值与潜在风险成为学界关注的焦点。目前，针对公民对公共管理中人工智能应用态度的系统研究仍较为不足。论文通过2×2×2调查实验，考察了公民在公共部门场景中对人类工作者与人工智能的信任差异，并分析了任务复杂程度与公共领域的调节作用。结果显示，公民普遍对人类工作者表现出更高的信任水平。进一步分析发现，在生命、安全等具有庄严价值的情境下，人类工作者的信任度显著高于人工智能；而在金钱、物质等世俗价值领域中，信任度差异则不显著。相关发现为人工智能在公共管理领域的科学应用提供了理论依据，也为其实践中的有效引导提供了参考。

【关键词】人工智能 公民信任 任务复杂程度 公共领域

【中图分类号】D63

【文献标识码】A

【文章编号】1674-2486 (2025) 04-0022-19

一、引言

2022年，ChatGPT 3.5的发布成为人工智能（Artificial Intelligence, AI）发展的关键转折点。作为自主响应的智能系统，人工智能已深度融入公共治理，

* 和思绮，重庆大学公共管理学院博士。通讯作者：刘炳胜，重庆大学公共管理学院教授，河北工业大学经济管理学院教授。感谢匿名评审专家、编委会专家和编辑部提出的宝贵意见。

基金项目：国家自然科学基金国际合作项目“基础设施促进城乡融合可持续发展的驱动机理与决策机制研究”（T2261129477），国家自然科学基金重点项目“新基建项目驱动区域平衡充分发展的机制与政策研究”（72134002），教育部哲学社会科学研究重大课题攻关项目“新基建促进区域平衡充分发展的长效机制研究”（21JZD029）。

被应用于失业保险（Gil et al.，2019）、交通执法（Carnis，2007）、警务调度（Meijer & Wessels，2019）及司法决策（Angwin et al.，2022）等场景。自2017年《新一代人工智能发展规划》（国发〔2017〕35号）颁布以来，中国加速推进人工智能战略布局，典型实践如北京师范大学“AI好老师”项目（余胜泉，2018），通过智能教育促进资源公平。人工智能对公共管理的赋能体现在多个方面，例如，自动化提升决策效率（Zekić-Sušac et al.，2021），数据共享优化专业协同（Toreini et al.，2020），高效算法简化流程（Ingrams et al.，2022），去人格化特性有效抑制决策偏差（Battaglio Jr et al.，2019），等等。

然而，人工智能在公共管理中的应用引发了诸多担忧。作为新兴技术，人工智能的潜在风险和局限性尚未被充分认知（Gaozhao et al.，2022）。对公共部门而言，工作人员信息素养不足可能导致人工智能应用失误（Brundage et al.，2018），职位替代焦虑可能引发工作懈怠（吕荣杰、郝力晓，2018）。对公民而言，信息素养差异使其难以评估人工智能应用的合理性（Wachter et al.，2021），算法不透明导致的“黑箱效应”更会加剧公民的质疑（Busuioc，2021）。此外，人工智能在需要情境判断和情感考虑的公共决策中存在局限，因为它缺乏人类工作者的人格化特征和自由裁量能力（Bullock，2019）。

近期研究开始辩证地分析人工智能对公共部门的潜在影响，使人工智能应用前景更趋立体化。然而，现有研究多聚焦人工智能对宏观管理功能的影响，较少探讨公民对公共部门人工智能的信任态度（Sharma et al.，2020）。当前，人工智能正推动公共部门运作模式与人力资源结构的深度变革（余胜泉，2018），使公共管理处于新旧秩序交替的不稳定期。这种转型迫使公民改变与公共部门既有的互动模式，基于人际默契的传统服务关系面临重构。在此背景下，公民需要将信任从人类工作者转向人工智能，但“信任移交”的意愿尚未明确（刘特等，2024）。公民对人工智能的信任程度将直接影响公共部门人工智能应用的深度与广度，进而决定公共服务互动的质量。

公民对人工智能的信任与疑虑呈现并存的特征。在工具性层面，人工智能因其高效和客观获得青睐（Lyell & Coiera，2017；Young et al.，2019）。由于自动化偏见的存在，公民更偏好拟人化较低的人工智能（Sundar，2008）。在价值性层面，人工智能的“解释鸿沟”引发了合法性担忧（Bullock，2019），需通过算法透明和信息素养的提升来缓解（Giest & Grimmelikhuijsen，2020），导致公民对人工智能替代人类持谨慎的态度（Agarwal，2018；Araujo et al.，2020）。这种矛盾态度使信任问题更加复杂。

因此，本文提出的研究问题是公民更信任公共部门中的“人”还是“人工智能”。

此外,本研究基于“分而治之”的敏捷治理框架(贾开、赵静,2023),从任务复杂程度与公共领域双重维度考察公民的信任差异。公共任务可划分为简单任务(如信息咨询)与复杂任务(如政策制定),前者通过基础算法实现,如人工智能聊天机器人(Aoki,2020),后者依赖机器学习应对不确定性(Athey et al.,2019),但受限于历史数据依赖性和算法刚性(O'Neil,2016),在特殊案例处理中存在公平性风险(Binns,2022; Bullock,2019)。同时,根据禁忌权衡理论,公共领域可分为庄严价值领域(生命、健康)与世俗价值领域(物质、效率),公民在庄严价值领域因更高的心理成本和身份认同需求,更倾向信任具备专业经验的人类工作者(Fiske & Tetlock,1997; Tetlock et al.,2000)。这种分类治理模式突破了传统的同质化路径,为精准施策提供了理论依据。

本研究采用2×2×2调查实验,通过线上调查收集1040份样本,探究公共服务提供者身份、任务复杂程度和价值领域对公民信任的影响。研究尝试揭示公民对人工智能的信任机制,为公共部门人工智能应用面临的信任挑战提供实证依据,并就如何提升公民信任提出实践建议。

二、文献综述与假设模型

(一) 公共部门中的“人”与“人工智能”

公共部门与私营部门相对,承担公共事务管理和公共服务供给的职能。狭义的公共部门指政府机构,广义的公共部门还包括准公共组织。我国公共部门主要包括政权机关、事业单位、公共企业和民间组织(张义博,2012)。本研究将公共部门界定为从事公共事务管理和公共服务供给的组织,重点考察政府机构和事业单位两类。政府机构作为核心主体,履行政策制定、资源管理等职能;事业单位则提供教育、医疗、文化等专业公共服务^①。

公共部门作为抽象实体,其具体运作依赖于各层级工作人员(Lipsky,1980)。以公安局为例,虽然机构被视为整体,实则由局长、警官、警察等不同层级的人员构成。决策制定主要由高层负责,而公民日常接触的多为基层执行人员。实质上,公共部门通过其工作人员得以具象化(Hu,2020),他们不仅代表公共部门的形象,更直接参与公共服务提供,政策制定、执行与评估等全流程工作。

随着公共部门数字化水平的提升,人工智能正以“工作人员”身份深度融

^① 关于印发《“十四五”公共服务规划》的通知(发改社会〔2021〕1946号)。

入公共管理中。在舆论管理中，人工智能作为“管理员”规范社交媒体行为（Van Dijck et al.，2018）；在健康领域，虚拟健康管理师提供个性化方案（Bickmore et al.，2016）；在司法系统中，人工智能参与刑事判决（Angwin et al.，2022）和假释评估（Dressel & Farid，2018）。在国内，典型案例是北京海淀区“城市大脑”项目，该项目通过人工智能整合多维数据实现精准治理（张欣亮等，2023）。这些实践预示着公共部门正经历从人工执行到人机协同的转型，这一过程重塑了服务模式并模糊了人机角色边界。在此背景下，探究公民对公共部门中“人类工作者”与“人工智能”的信任差异具有重要现实意义，这既涉及技术接受度问题，也关乎公民对政府治理模式的信任基础。

（二）公民对“人”与“人工智能”的信任差异

信任作为公民与公共部门的核心纽带（张成福、边晓慧，2013），其内涵正随着人工智能的渗透而发生重构。公民对人机主体的差异化信任已成为影响公共服务效能的关键变量。学界对人工智能应用的争议性评价，恰恰折射出技术革新带来的新型信任挑战。

现有研究主要从四个维度论证了人工智能对公共管理的积极影响：第一，在效率提升方面，人工智能通过不间断工作的能力（Wilson & Daugherty，2018）、去官僚化特性（Ingrams et al.，2022）和自动化处理，能显著提高行政效能（Zekić-Sušac et al.，2021）；第二，在决策公平性方面，人工智能克服了人类决策中的认知局限（Kahneman et al.，1982）和自由裁量偏差（Tummers & Bekkers，2014）问题，通过去人格化设计减少主观偏见与人情影响（Young et al.，2019）；第三，在政民互动方面，人工智能构建了技术中介的沟通渠道，增强了公民与政府之间的互操作性与“共同语言”（Margetts & Dorobantu，2019；Vogl et al.，2020）；第四，在执行一致性方面，人工智能基于数据共享和系统互联，解决了信息不对称和专业断层问题，确保公共服务的标准化输出（Gaozhao et al.，2022）。

学界对人工智能应用的担忧主要聚焦于五个方面：一是技术成熟度方面，人工智能作为新兴技术仍存在未知领域（Gaozhao et al.，2022），操作失误可能被系统放大（Brundage et al.，2018）；二是成本效益方面，数字化改革需要时间、资金和人才等资源的大量投入（Nishant et al.，2020）；三是透明度方面，算法复杂性影响公众对决策合法性的认知（阮荣彬等，2023）；四是代表性方面，人工智能缺乏人类特征的多元性（Lünich & Starke，2020），难以体现群体利益（Krislov，2012）；五是执行灵活性方面，人工智能在个案处理和精细化决策中存在局限（Bannister & Connolly，2020）。

总体而言,人工智能在公共管理中的应用呈现出双重效应。一方面,人工智能显著提升了行政效率和决策速度;另一方面,人工智能透明度不足、代表性缺失及个性化处理等局限可能削弱公众信任。现有研究大多聚焦宏观层面的智能化转型,却忽视微观层面的公民体验。本研究旨在填补这一空白,特别关注公民对人工智能应用的感知与态度。值得注意的是,尽管对人工智能的评价存在争议,但受“自动化偏见”影响,公民可能对人工智能保持较高信任度。因此,本研究提出以下假设。

H1: 在公共部门中,与“人”相比,公民更信任“人工智能”。

(三) 任务复杂程度的调节作用: 简单任务与复杂任务

任务复杂程度是区分公共任务的关键维度。公共任务既包含基础性服务(如信息咨询),也涉及战略性决策(如政策制定)。人工智能应用亦具有由简单到复杂的连续性,从处理简单查询的聊天机器人(Aoki, 2020)到执行复杂分析的“城市大脑”系统(张欣亮等, 2023)。前者解决标准化问题,如询问人工客服“腹痛应该挂哪个科室的号”或“申请材料是否需要原件”;后者则需整合多维动态数据,进行系统性决策。基于敏捷治理原则,本研究将区分简单任务与复杂任务两种情境,以系统地考察公民对人工智能的差异化信任态度。

复杂任务要求人工智能具备更高的精确性、稳定性和适应性,需依赖高级算法实现(Athey et al., 2019)。但算法设计受限于人类的认知偏差(Kahneman, 2003),可能导致服务分配错误(Sun & Medaglia, 2019)。尤其涉及情感判断等人格化或情境化任务时,人工智能的自由裁量能力明显不足(Bannister & Connolly, 2020; Bullock, 2019)。相比之下,简单任务因目标单一更容易实现自动化(Bullock, 2019)。

公民对人工智能的信任可能呈现任务依赖性。面对复杂任务时,公民会对人工智能的失误风险和精细化处理能力表示担忧;面对简单任务时,公民则倾向使用具备高效性的人工智能。基于有限理性理论,公民在复杂情境中往往采取路径依赖策略(Kahneman, 2003)。由于人工智能作为新兴事物存在能力边界的不确定性(Gaozhao et al., 2022),公民在复杂任务中可能更倾向信任熟悉的人类工作者,以降低对不确定性的感知。

综上所述,公共任务的复杂程度可能影响人工智能的应用效果。在简单任务中,人工智能凭借其高效自动化优势提升服务效能;而复杂任务要求更高的准确性、一致性和情境判断力,人工智能当前的表现仍存在局限。这种技术边界不仅制约着人工智能在复杂决策中的应用,更因公民的路径依赖倾向而强化其对人类工作者的信任偏好。因此,本研究提出以下假设。

H2：任务复杂程度影响公民对公共部门中“人”与“人工智能”的信任差异。具体而言，在简单任务中，公民更信任“人工智能”；在复杂任务中，公民更信任“人”。

本研究选取两类典型任务为实验情境，一是以人工智能聊天机器人为代表的咨询任务（Aoki, 2020），其特点是基础性信息交互；二是以“城市大脑”为范例的决策任务（张欣亮等, 2023），涉及复杂信息处理并产生实质影响。这两类任务分别对应公共管理中的简单情境与复杂情境。

（四）公共领域的调节作用：庄严领域与世俗领域

公共领域是影响公民信任的另一关键变量。基于禁忌权衡理论（Taboo Trade-offs Theory），本研究将公共领域区分为涉及生命、健康等庄严价值的庄严领域，以及包含物质、娱乐等世俗价值的世俗领域（Fiske & Tetlock, 1997）。这种分类有助于系统考察公民对人机主体的差异化信任。

禁忌权衡理论阐释了人类在价值比较中的心理机制。该理论指出，跨维度比较面临认知不可通约性挑战（Payne et al., 1992），即不同价值在认知层面难以等同衡量（Fiske & Tetlock, 1997）。价值可分为庄严价值（如生命、基本权利）和世俗价值（如金钱、物质享受）两类（Tetlock, 2003）。大众文化普遍接受世俗价值间的交换，但庄严价值与世俗价值的权衡（如经济考量影响医疗决策）会触发心理防御机制，形成所谓“禁忌权衡”（Fiske & Tetlock, 1997）。

禁忌权衡理论指出，人类在价值权衡中存在层级偏好，将庄严价值置于世俗价值之上。2000年，Tetlock等人（2000）通过实验验证了该理论。具体而言，Tetlock等人将价值权衡区分为常规权衡（世俗价值间交换）、悲剧权衡（庄严价值间交换）与禁忌权衡（庄严价值与世俗价值交换）三类。实验采用医疗决策情境，悲剧权衡组需选择救治哪位病童，禁忌权衡组则需在救治成本与患者生命间抉择。结果显示，涉及禁忌权衡的决策引发了更强的负面评价，证实了庄严价值的优先性。

基于禁忌权衡理论，公共管理领域可划分为两类（De Fine Licht, 2011, 2014; De Fine Licht et al., 2014），即涉及生命、健康等庄严价值的领域（如医疗、治安），以及涉及物质利益的世俗价值领域（如税收、市政建设）。研究表明，在庄严领域，政策即便透明也难获得公众支持（如刑释人员管理、交通管制），而透明度甚至可能产生反效果（如医疗、环保）；但在世俗领域则无此负面效应（De Boer, 2020; De Fine Licht, 2014; Porumbescu et al., 2017）。这表明庄严价值领域更易触发公民的禁忌心理机制。

综上所述，公共领域可能影响公民信任。庄严领域（生命、健康等）的事

务往往使公众承担更高的心理成本，并激发更强的警惕性，而公众对世俗领域（物质利益等）的事务则更宽容。这种差异延伸至人机信任选择中则表现为：在庄严价值任务中，公民可能倾向信任熟悉可靠的人类工作者；在世俗价值任务中，则偏好高效便捷的人工智能。因此，本研究提出以下假设。

H3：公共领域影响公民对公共部门中“人”与“人工智能”的信任差异。具体而言，在世俗领域中，公民更信任“人工智能”；在庄严领域中，公民更信任“人”。

本研究选取公共医疗和税收作为两类价值领域的典型代表。公共医疗涉及生命、健康等庄严价值，引发公民高度审慎的态度；税收则关乎物质、分配等世俗价值，公民更关注其经济效率。

本研究的假设模型如图1所示。研究重点分析公共任务承担者身份（人类工作者或人工智能）对公民信任的影响机制。此外，本研究特别考察任务复杂程度（简单咨询或复杂决策）与公共领域（庄严领域或世俗领域）两个变量的调节作用。

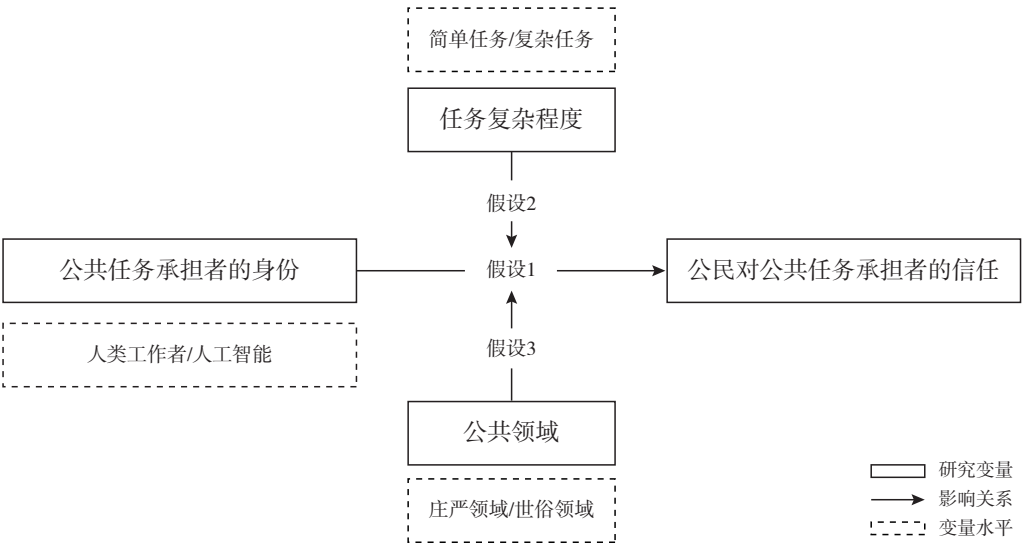


图1 假设模型

资料来源：作者自制。

三、研究设计

（一）实验材料与实验步骤

本研究采用2×2×2调查实验，考察公民对人机主体的信任差异，以及任务

复杂程度和公共领域的调节效应。调查实验通过文本操控构建不同情境，随机分配被试接受文本刺激，并测量其反应。该方法能有效控制干扰并确立因果关系（James et al., 2017；李晓倩，2018）。具体而言，本研究采用三因素实验设计：一是任务承担者身份（人类/人工智能），二是任务复杂程度（咨询/决策），三是公共领域（医疗/税收）。通过完全随机形成8个实验组，被试随机接受一个情境的文本刺激。各实验组的操控文本如下。

（1）“人类×简单任务×世俗领域”实验组：“您现在位于一个税务局。您准备向一位税务稽查员咨询纳税常识。具体而言，您先提出您的问题，这位税务稽查员会对您的问题进行了回答。”（2）“人类×简单任务×庄严领域”实验组：“您现在位于一家公立医院。您准备向一位医生咨询健康常识。具体而言，您先提出您的问题，这位医生会对您的问题进行了回答。”（3）“人类×复杂任务×世俗领域”实验组：“您现在位于一个税务局。一位税务稽查员将检查您的纳税情况并出具检查报告。首先，这位税务稽查员会根据税收相关的法律、法规与政策，获取您的收入与纳税信息；接着，他会对您的收入与纳税情况进行分析；最后，他会判定您是否存在违规或违法行为。”（4）“人类×复杂任务×庄严领域”实验组：“您现在位于一家公立医院。一位医生将诊断您的病症并开具处方。首先，这位医生会让您描述自己的病情；接着，在您描述的基础上，他会对您进行详细的询问与检查；最后，他会确定您的病症，并开出相应的治疗处方。”（5）“人工智能×简单任务×世俗领域”实验组：“您现在位于一个税务局。您准备向一位虚拟的人工智能税务稽查员咨询纳税常识。具体而言，您先提出您的问题，这位虚拟的人工智能税务稽查员会对您的问题进行了回答。”（6）“人工智能×简单任务×庄严领域”实验组：“您现在位于一家公立医院。您准备向一位虚拟的人工智能医生咨询健康常识。具体而言，您先提出您的问题，这位虚拟的人工智能医生会对您的问题进行了回答。”（7）“人工智能×复杂任务×世俗领域”实验组：“您现在位于一个税务局。一位虚拟的人工智能税务稽查员将检查您的纳税情况并出具检查报告。首先，这位虚拟的人工智能税务稽查员会根据税收相关的法律、法规与政策，获取您的收入与纳税信息；接着，它会对您的收入与纳税情况进行分析；最后，它会判定您是否存在违规或违法行为。”（8）“人工智能×复杂任务×庄严领域”实验组：“您现在位于一家公立医院。一位虚拟的人工智能医生将诊断您的病症并开具处方。首先，这位虚拟的人工智能医生先让您描述自己的病情；接着，在您描述的基础上，它会对您进行详细的询问与检查；最后，它会确定您的病症，并开出相应的治疗处方。”

此外，本研究以公民对任务承担者的信任为因变量，采用10点李克特量表测量，询问被试：“在刚才所描述的情境中，请问您有多信任对方？请根据您的

真实感受进行选择（‘1’表示‘非常不信任’，‘10’表示‘非常信任’）。”控制变量还包括可能影响人机信任差异的因素，包括人口学变量（性别、年龄、教育程度），人工智能接触程度（反映公民对人工智能的熟悉度），以及地区科技创新水平（衡量科技接受氛围）。

实证研究分为三个阶段进行。首先，控制变量采集。收集被试人口学特征和人工智能接触程度数据，以《中国区域科技创新评价报告 2024》划分的地区科技创新水平为依据，京沪粤苏津浙是科技创新水平高的地区，其他是低水平的地区^①，进行 IP（Internet Protocol，互联网协议）地址匹配。其次，实验操控。采用 2×2×2 完全随机设计，被试随机接受 8 种情境的文本刺激。实验设置 20 秒强制阅读时间确保被试的专注度。最后，因变量测量。评估被试对任务承担者的信任水平。

（二）操纵有效性检验

为检验实验操纵是否成功诱发被试在相应维度上的主观感知变化，本文通过预实验（ $N=200$ ），采用卡方检验（Chi-square Test）分析任务承担者身份与感知任务承担者身份、公共领域与感知公共领域、任务复杂程度与感知任务复杂程度之间的关联性。具体而言，在被试被随机分配至 8 个实验组之一，并阅读相应的操纵文本后，他们需要回答 3 个与情境相关的单选题。为避免顺序效应的干扰，3 个问题的呈现顺序被随机化处理。问题内容是：（1）在刚才的情境中，为您提供服务的是哪一类主体？（选项 1：真人；选项 2：人工智能）。（2）在刚才的情境中，您所面临的问题属于以下哪一类？（选项 1：与生命、健康等相关的问题；选项 2：与金钱、物质等相关的问题）。（3）在刚才的情境中，任务的性质是？（选项 1：比较简单的；选项 2：比较复杂的）。在进行操纵有效性检验前，我们对操纵变量的频数分布进行了统计，检验其水平分布是否均衡。结果显示，关于任务承担者身份的回答，选项 1 的频数为 104（52%），选项 2 的频数为 96（48%）。关于公共领域价值属性的回答，选项 1 的频数为 103（51.5%），选项 2 的频数为 97（48.5%）。关于任务复杂程度的回答，选项 1 的频数为 108（54%），选项 2 的频数为 92（46%）。操纵有效性检验的结果显示（见表 1），3 组变量均呈现出极显著的统计关联，表明实验操纵在感知层面成功诱导了被试的响应，达成了预期的操纵效度。

^① 由中国科学技术发展战略研究院发布。

表 1 操纵有效性检验

操纵变量	感知变量	卡方值	自由度
任务承担者身份	感知任务承担者身份	134.94 ***	1
公共领域	感知公共领域	173.54 ***	1
任务复杂程度	感知任务复杂程度	115.03 ***	1

注：*** 表示 $p < 0.01$ 。

资料来源：作者自制。

（三）数据来源与样本结构

本研究通过见数平台（Credamo）开展线上实验，最终获取 1040 份有效样本。样本的基本人口学信息见表 2。

表 2 样本人口学信息（ $N = 1040$ ）

变量	水平	百分比（%）
性别	男性	51.5
	女性	48.5
年龄	0 岁～20 岁	3.3
	21 岁～30 岁	28.9
	31 岁～40 岁	23.7
	41 岁～50 岁	26.6
	51 岁～60 岁	15.3
	60 岁以上	2.2
	小学及以下	0.3
教育水平	初中	1.6
	普高/中专/技校/职高	13.8
	专科	26.2
	本科	47.2
	硕士	10.0
	博士	1.0

资料来源：作者自制。

F 检验用于明确被试是否被随机分配到每个实验组。表 3 显示各实验组之间没有显著差异，意味着被试的分配实现了随机化。

表 3 平衡性检验

	性别 = 女性	年龄	教育水平	样本量
人类×简单任务×世俗领域	0.440 (0.498)	3.210 (1.090)	4.63 (0.873)	130
人类×简单任务×庄严领域	0.460 (0.500)	3.350 (1.160)	4.44 (0.996)	130

(续上表)

	性别=女性	年龄	教育水平	样本量
人类×复杂任务×世俗领域	0.550 (0.500)	3.280 (1.201)	4.47 (1.043)	130
人类×复杂任务×庄严领域	0.460 (0.500)	3.320 (1.264)	4.52 (0.990)	130
人工智能×简单任务×世俗领域	0.520 (0.502)	3.260 (1.168)	4.55 (1.053)	129
人工智能×简单任务×庄严领域	0.470 (0.501)	3.360 (1.220)	4.51 (0.856)	130
人工智能×复杂任务×世俗领域	0.450 (0.499)	3.280 (1.198)	4.60 (0.967)	131
人工智能×复杂任务×庄严领域	0.530 (0.501)	3.220 (1.239)	4.47 (0.864)	130
F 值	0.871	0.303	0.624	
p 值	0.529	0.953	0.736	

注：表格中数值为均值，括号内数值为标准差。虚拟变量的均值为比例。
资料来源：作者自制。

(四) 数据分析策略

本研究采用 R 语言进行方差分析，构建包含三阶交互效应的假设模型。自变量为任务承担者身份，因变量为公民信任，调节变量为任务复杂程度与公共领域，控制变量包括人口学特征、人工智能接触程度和地区科技创新水平。分析涵盖二阶交互，包括任务承担者身份×任务复杂程度、任务承担者身份×公共领域、任务复杂程度×公共领域，以及三阶交互，即任务承担者身份×任务复杂程度×公共领域。

四、实验结果与假设检验

(一) 三因素方差分析

表 4 呈现了方差分析的结果。结果显示，任务承担者身份与任务复杂程度（即交互项 1）的交互效应不显著，说明任务复杂程度不会显著改变任务承担者身份对公民信任的影响。换言之，无论任务是简单（咨询）还是复杂（决策），公民对“人”与“人工智能”的信任偏好基本一致。相反，任务承担者身份与公共领域（即交互项 2）的交互效应显著，表明不同领域会影响任务承担者身份对信任的作用。具体来看，在庄严领域（医疗）与世俗领域（税务）中，公民对“人”与“人工智能”的信任差异呈现出不同模式。此外，任务复杂程度与公共领域（即交互项 3）的交互效应不显著。公共任务承担者的身份和任务复杂程度和公共领域的三阶交互（即交互项 4）效应显著，说明“人”与“人

工智能”的身份效应会随任务复杂程度与公共领域的不同组合而变化，具体将在后文进一步探讨。

表 4 方差分析结果

	自由度	Ⅲ型平方和	均方	F 值
主要变量				
公共任务承担者的身份	1	111.8	111.81	58.349***
任务复杂程度	1	21.5	21.53	11.237***
公共领域	1	12.3	12.34	6.440*
交互项 1	1	0.0	0.00	0.002
交互项 2	1	59.1	59.08	30.830***
交互项 3	1	4.3	4.26	2.225
交互项 4	2	30.9	30.88	16.118***
控制变量				
性别	1	2.6	2.62	1.365
年龄	5	11.7	2.34	1.221
教育水平	4	14.8	3.71	1.935
人工智能接触程度	1	17.4	17.42	9.090**
地区科技创新水平	1	2.4	2.35	1.228
残差	1021	1956.4	1.92	

注：***、**、* 分别表示 $p < 0.01$ 、 $p < 0.05$ 、 $p < 0.1$ 。

资料来源：作者自制。

在控制变量中，性别、年龄、教育水平及地区科技创新水平对公民信任均无显著影响，人工智能接触程度的主效应显著，这表明日常生活中接触人工智能较多的被试，无论面对“人”还是“人工智能”，均表现出更高的信任（ $\Delta = 0.44$ ，置信区间 = $[0.152, 0.742]$ ）。进一步分析的结果显示，人工智能接触程度与任务承担者身份的交互效应显著（自由度 = 1，Ⅲ型平方和 = 12.8，均方 = 12.84，F 值 = 6.709， $p < 0.01$ ）。具体而言，人工智能更易获得频繁接触人工智能的被试的信任（ $\Delta = 0.85$ ，置信区间 = $[0.292, 1.426]$ ），而人类工作者的被信任水平则不受接触程度的影响（ $\Delta = 0.08$ ，置信区间 = $[-0.440, 0.618]$ ）。由此可见，人工智能接触程度调节了任务承担者身份与公民信任之间的关系。

（二）主效应分析与成对比较

图 2 用三个面板依次展示了任务承担者身份、任务复杂程度以及公共领域对

公民信任影响的主效应。成对比较结果表明，被试普遍更信任人类工作者，由人类转为人工智能会显著降低信任水平（ $\Delta = 0.65$ ，置信区间 = $[0.824, 0.487]$ ）。在任务复杂程度方面，无论任务承担者身份为何，公民在简单任务中的信任度更高（ $\Delta = 0.28$ ，置信区间 = $[0.456, 0.119]$ ）。在公共领域方面，公民在世俗领域中对任务承担者的信任也更高（ $\Delta = 0.21$ ，置信区间 = $[0.386, 0.049]$ ）。

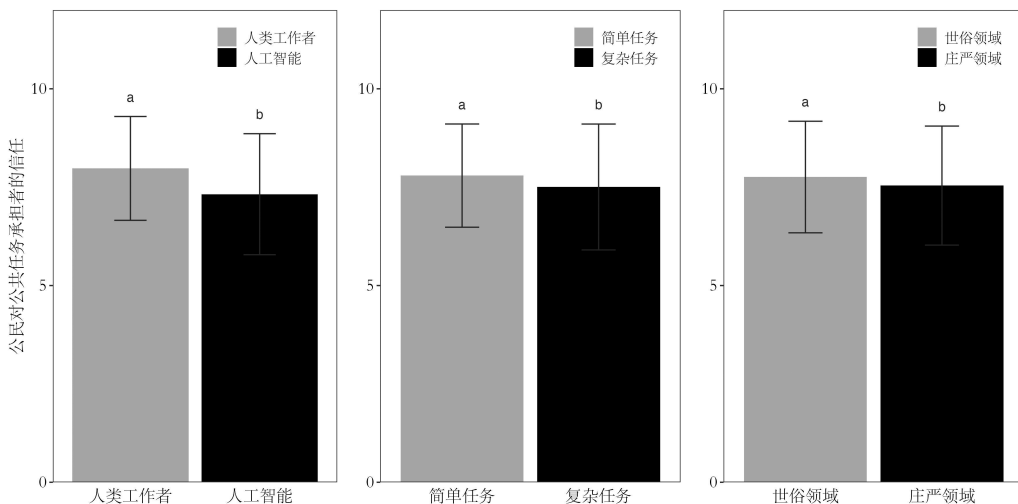


图2 主效应分析与成对比较结果

注：各直方图上方字母若相同，则说明相应组别之间无显著差异；若不同，则说明相应组别之间有显著差异。

资料来源：作者自制。

（三）交互效应分析与成对比较

图3展示了交互效应分析的结果，因涉及三阶交互，本文采用两个面板呈现：面板一按公共领域划分为子面板（1）和（2），面板二按任务复杂程度划分为子面板（3）和（4）。在子面板（1）中，公民在简单任务中更信任人类工作者；但在复杂任务中，公民对人类与人工智能的信任无显著差异。子面板（2）显示，无论任务复杂程度如何，公民均更信任人类工作者。在子面板（3）中，无论领域如何，公民普遍更信任人类工作者。子面板（4）显示，在庄严领域中，公民更信任人类工作者；而在世俗领域中，公民对人类与人工智能的信任无显著差异。

结合方差分析与成对比较结果可以发现，从总体来看，任务承担者身份显著影响公民信任，公民普遍更信任人类。任务复杂程度未调节该关系，但公共领域发挥了调节作用——在庄严领域信任差异显著，在世俗领域则不明显。因

此，H1 未得到验证，即公民在一般情况下更倾向信任人类。H2 同样未被支持，任务复杂程度并未调节身份与信任之间的关系，但其对整体信任水平存在直接影响。H3 获得支持，公共领域显著调节了身份与信任之间的关系。

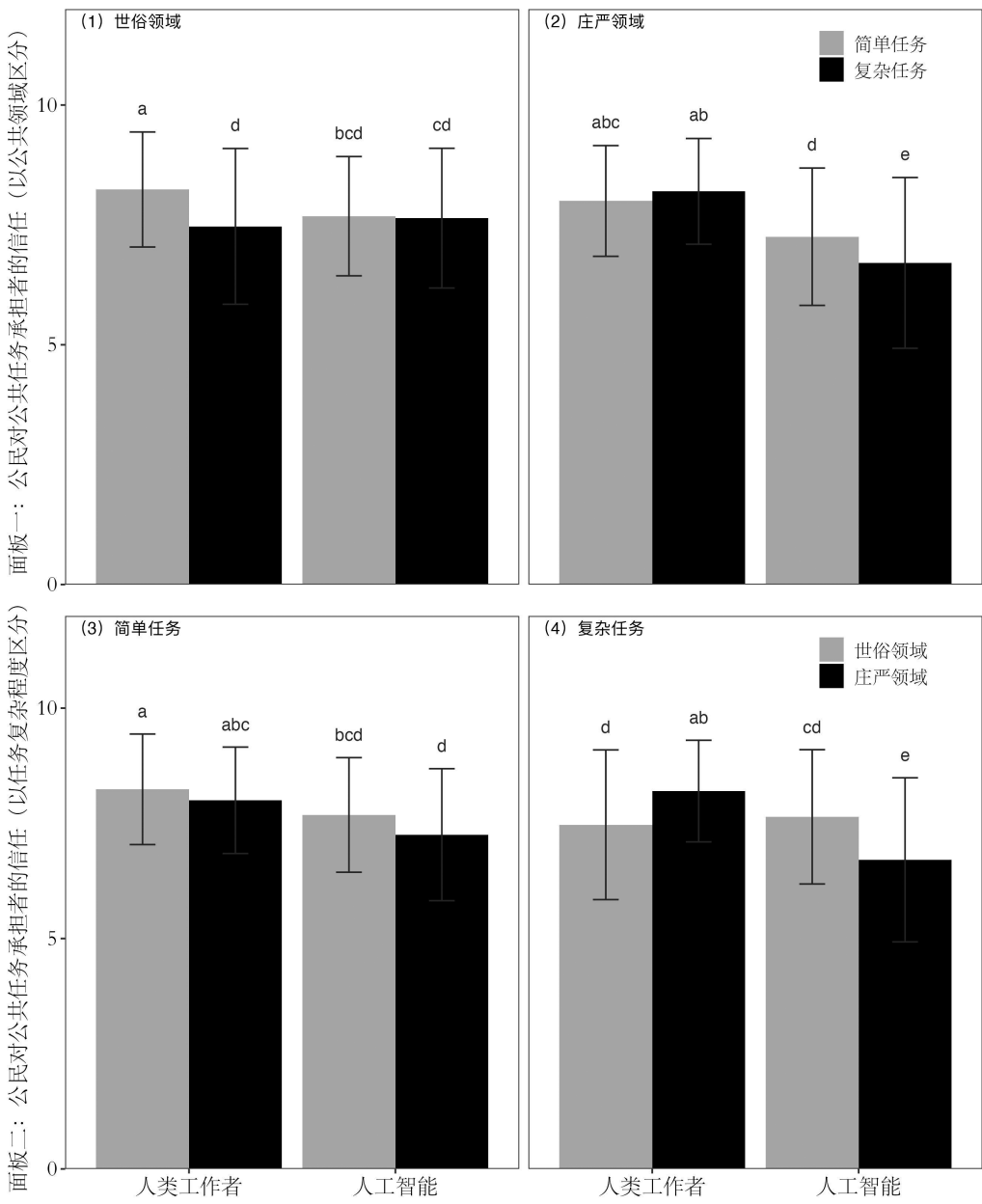


图3 交互效应分析与成对比较结果

注：各直方图上方字母若相同，则说明相应组别之间无显著差异；若不同，则说明相应组别之间有显著差异。

资料来源：作者自制。

五、讨论与结论

（一）研究发现

随着人工智能在公共管理中的广泛应用，学术界对其利弊展开了激烈讨论，认为公共部门应审慎推进人工智能应用，以公民接受度为关键考量（刘特等，2024）。当前研究的缺口在于缺乏公民对人工智能态度的实证证据，特别是在公共部门组织变革的背景下（余胜泉，2018），这种变革打破了公民原有的服务认知模式。核心问题在于公民对人机主体的信任偏好。若公民倾向于信任人工智能，则可延续现有实践；若公民偏好人类工作者，则需明确人工智能的适用边界。此外，人工智能的敏捷治理理念强调了分类应对的重要性（贾开、赵静，2023），但现有研究多聚焦宏观政策层面，忽视了公民对人机信任差异的微观分析。此外，公共管理具有多元性，不同任务类型塑造了差异化的部门形象（De Boer，2020），可能导致公民基于既有期望产生不同反应（Hayes，2005）。这凸显了采用分类思维分析人机信任的必要性，有助于政府在资源约束的情况下精准施策，维护公民信任。

本研究填补了公民对公共部门人工智能态度研究的空白，探讨了公民在公共部门中更信任“人”还是“人工智能”这一问题。实证发现，尽管人工智能应用日益广泛，公民仍更信任人类而非人工智能。这一结果挑战了这些观点，例如，自动化偏见（Logg et al.，2019）可能缘于人工智能在透明度、代表性等方面的缺陷（Busuioc，2021；Janssen et al.，2020；Wachter et al.，2021；Zuiderwijk et al.，2021；阮荣彬等，2024），公民存在路径依赖心理（Kahneman et al.，1982），以及对同类工作者具备身份认同等。未来研究需进一步验证这些机制。

为更有针对性地探讨公民在公共部门中对“人”与“人工智能”的信任差异，本研究考虑了任务复杂程度发挥的边界作用。结果表明，虽然公民在简单任务情境中对任务承担者表现出更高的信任水平，但人机信任并不存在差异。根据有限理性理论（Kahneman et al.，1982），复杂任务本应强化公民对熟悉路径（人类）的依赖，且人工智能作为新兴技术需额外的认知投入以建立信任（Brundage et al.，2018），但结果显示，无论任务复杂程度如何，公民对人类的信任始终显著高于人工智能。这一结果暗示，信任机制可能受更深层的主观因素驱动，例如身份认同或情感联结，而非单纯基于任务属性的理性计算。同时，任务复杂程度主效应的存在表明，复杂任务本身会加剧公民的信任焦虑，这与

有限理性理论关于复杂情境心理负担的论断一致。

此外，本研究还发现，公共领域不仅直接影响公民信任，还显著调节人机信任的差异。在涉及生命、健康等庄严价值的领域，公民更倾向信任人类工作者。禁忌权衡理论（Tetlock et al.，2003）为此提供了理论支撑——在庄严领域中，公民的认知投入显著增加；而世俗领域侧重效率，公民心理负担较低。这种分野与伦理学价值维度高度契合，即道德性价值（如公平）强调伦理正当性，工具性价值（如效率）关注手段合理性（Anderson，2009；Meynhardt，2009；包国宪、王学军，2012）。价值观的层级差异深刻影响着公民的信任选择，凸显公共价值考量在管理研究中的必要性。

综上所述，本研究采用2×2×2调查实验探讨了公共部门中的人机信任差异，发现公民整体上更信任人类，尤其在涉及生命、健康等庄严价值领域表现出显著偏好。任务复杂度虽直接影响信任水平，但未调节人机信任差异；而公共领域显著影响信任选择。研究结果表明，尽管人工智能具备效率优势，但公民对其在庄严领域的应用持审慎态度，凸显公共部门需要依据领域特性制定差异化的技术应用策略。未来研究应深入探究信任形成的主观机制（如身份认同、情感联结等），为人工智能治理提供更精细化的理论支持。

（二）对策建议

本研究为公共部门的人工智能应用提出三点对策建议：第一，在公共服务交互场景中，需评估公民的信任压力，并加强对人工智能技术的宣传与介绍；第二，依据禁忌权衡理论，在庄严领域（如公共医疗、公共安全等）应审慎引入人工智能，在世俗领域（如税收、市政管理等）可适度推进；第三，尽管任务复杂程度未显示出调节效应，但仍应该提升人类工作者的专业素养以缓解复杂任务中的信任风险。

（三）不足与展望

本研究仍存在三个方面的局限：一是实验情境的单一性，未来可采用刺激抽样法（Stimuli Sampling）扩展情境类型（如监狱管理、核安全等）；二是样本存在年龄与教育水平的偏差，需要提升群体的代表性；三是未控制时间因素、政民互动经历等潜在变量，后续研究可结合纵向追踪方法与多维度控制变量进行深化研究。

参考文献

包国宪、王学军（2012）. 以公共价值为基础的政府绩效治理——源起、架构与研究问题. 公共管理学报，9（2）：89-97+126-127.

- Bao, G. X., & Wang, X. J. (2021). Public Value-Based Government Performance Governance: Origin, Structure and Research Questions. *Journal of Public Management*, 9(2): 89-97+126-127. (in Chinese)
- 贾开、赵静 (2023). 技术嵌入、价值倾向与算法分类治理. *经济社会体制比较*, (4): 101-111.
- Jia, K., & Zhao, J. (2023). Technology Integration, Value-Oriented, and Categorized Algorithm Governance. *Comparative Economic & Social Systems*, (4): 101-111. (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(1): 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(1): 37-61+219-220. (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)
- 吕荣杰、郝力晓 (2018). 人工智能等技术对劳动力市场的影响效应研究. *工业技术经济*, 37(12): 131-137.
- Lü, R. J., & Hao, L. X. (2018). Research on the Influence of Artificial Intelligence and Other Technologies on Labor Market. *Journal of Industrial Technology and Economy*, 37(12): 131-137. (in Chinese)
- 阮荣彬、朱祖平、陈莞、李文攀 (2024). 政府科技伦理治理与人工智能企业科技向善. *科学学研究*, 42(8): 1577-1586+1606.
- Ruan, R. B., Zhu, Z. P., Chen, W., & Li, W. P. (2023). The Relationship Between Ethical Governance of Government in Science and Technology and Science and Technology for Social Good of AI Enterprises. *Studies in Science of Science*, 2024, 42(8): 1577-1586+1606. (in Chinese)
- 余胜泉 (2018). 人工智能教师的未来角色. *开放教育研究*, 24(1): 16-28.
- Yu, S. Q. (2018). The Future Roles of AI Teacher. *Open Education Research*, 24(1): 16-28. (in Chinese)
- 张成福、边晓慧 (2013). 重建政府信任. *中国行政管理*, (9): 7-14.
- Zhang, C. F., & Bian, X. H. (2013). The Rebuilding of Trust in Government. *Chinese Public Administration*, (9): 7-14. (in Chinese)
- 张欣亮、唐斯斯、李晶 (2023). 人工智能驱动政府敏捷治理的运行机制与实践探索——以北京海淀区城市大脑为例. *技术经济*, 42(11): 113-119.
- Zhang, X. L., Tang, S. S., & Li, J. (2023). Operation Mechanism and Practice Exploration of AI-driven Agile Governance in Government: A Case Study of Urban Brain in Haidian District, Beijing. *Journal of Technology Economics*, 42(11): 113-119. (in Chinese)
- 张义博 (2012). 公共部门与非公共部门收入差异的变迁. *经济研究*, 47(4): 77-88.
- Zhang, Y. B. (2012). The Change of Income Differentials Between Public and Nonpublic Sector in China. *Economic Research Journal*, 47(4): 77-88. (in Chinese)
- Agarwal, P. K. (2018). Public Administration Challenges in the World of AI and Bots. *Public Administration Review*, 78(6): 917-921.
- Anderson, E. (2009). *Democracy: Instrumental vs. Non-instrumental Value*. In *Contemporary Debates in Political Philosophy*. Hoboken: Wiley-Blackwell.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2022). *Machine Bias*. In *Ethics of Data and Analytics*. New York: Auerbach Publications.
- Aoki, N. (2020). An Experimental Study of Public Trust in AI Chatbots in the Public Sector. *Government Information Quarterly*, 37(4): 101490.
- Araujo, T., Helberger, N., Kruike-meier, S., & De Vreese, C. H. (2020). In AI We Trust? Perceptions about Automated Decision-making by Artificial Intelligence. *AI Society*, 35(3): 611-623.
- Athey, S., & Imbens, G. W. (2019). Machine Learning Methods that Economists Should Know About. *Annual Review of Economics*, 11(1): 685-725.
- Bannister, F., & Connolly, R. (2020). Administration by Algorithm: A Risk Management Framework. *Information Policy*, 25(4): 471-490.
- Battaglio Jr, R. P., Belardinelli, P., Bellé, N., & Cantarelli, P. (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.

- Bickmore, T. W. , Utami, D. , Matsuyama, R. , & Paasche-Orlow, M. K. (2016). Improving Access to Online Health Information with Conversational Agents: A Randomized Controlled Experiment. *Journal of Medical Internet Research*, 18(1): e1.
- Binns, R. (2022). Human Judgment in Algorithmic Loops: Individual Justice and Automated Decision-making. *Regulation & Governance*, 16(1): 197–211.
- Brundage, M. , Avin, S. , Clark, J. , Toner, H. , Eckersley, P. , Garfinkel, B. , & Amodei, D. (2018). The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. *ArXiv*, abs/1802.07228.
- Bullock, J. B. (2019). Artificial Intelligence, Discretion, and Bureaucracy. *The American Review of Public Administration*, 49(7): 751–761.
- Busuioc, M. (2021). Accountable Artificial Intelligence: Holding Algorithms to Account. *Public Administration Review*, 81(5): 825–836.
- Camis, L. (2007). The Automated Speed Enforcement System in Great Britain: Between a Technical Revolution and Administrative Continuity. *International Review of Administrative Sciences*, 73(4): 597–610.
- De Boer, N. (2020). How Do Citizens Assess Street-Level Bureaucrats' Warmth and Competence? A Typology and Test. *Public Administration Review*, 80(4): 532–542.
- 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.
- De Fine Licht, J. (2014). Policy Area as a Potential Moderator of Transparency Effects: An Experiment. *Public Administration Review*, 74(3): 361–371.
- De Fine Licht, J. , Naurin, D. , Esaiasson, P. , & Gilljam, M. (2014). When Does Transparency Generate Legitimacy? Experimenting on a Context-Bound Relationship. *Governance*, 27(1): 111–134.
- Dressel, J. , & Farid, H. (2018). The Accuracy, Fairness, and Limits of Predicting Recidivism. *Science Advances*, 4(1): eaao5580.
- Fiske, A. P. , & Tetlock, P. E. (1997). Taboo Trade-offs: Reactions to Transactions that Transgress the Spheres of Justice. *Political Psychology*, 18(2): 255–297.
- Gaozhao, D. , Wright II, J. , & Robinson, M. (2022). Bureaucrat or Artificial Intelligence: People's Preferences and Perceptions of Government Service. *Public Management Review*, 26(6): 1498–1525.
- Giest, S. , & Grimmelikhuijsen, S. (2020). Introduction to Special Issue Algorithmic Transparency in Government: Towards a Multi-level Perspective. *Information Policy*, 25(4): 409–417.
- Gil, O. , Cortés-Cediel, M. E. , & Cantador, I. (2019). Citizen Participation and the Rise of Digital Media Platforms in Smart Governance and Smart Cities. *International Journal of E-Planning Research*, 8(1): 19–34.
- Hayes, D. (2005). Candidate Qualities Through a Partisan Lens: A Theory of Trait Ownership. *American Journal of Political Science*, 49(4): 908–923.
- Hu, Y. (2020). Culture Marker Versus Authority Marker: How Do Language Attitudes Affect Political Trust? *Political Psychology*, 41(4): 699–716.
- Ingrams, A. , Kaufmann, W. , & Jacobs, D. (2022). In AI We Trust? Citizen Perceptions of AI in Government Decision Making. *Policy & Internet*, 14(2): 390–409.
- James, O. , Jilke, S. , & Van Ryzin, G. (2017). *Experiments in Public Management Research: Challenges and Opportunities*. Cambridge: Cambridge University Press.
- Janssen, M. , Brous, P. , Estevez, E. , Barbosa, L. S. , & Janowski, T. (2020). Data Governance: Organizing Data for Trustworthy Artificial Intelligence. *Government Information Quarterly*, 37(3): 101493.
- Kahneman, D. (2003). A Perspective on Judgment and Choice: Mapping Bounded Rationality. *American Psychologist*, 58(9): 697.
- Kahneman, D. , Slovic, S. P. , Slovic, P. , & Tversky, A. (1982). *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Krislov, S. (2012). *Representative Bureaucracy*. New York: Quid Pro Books.
- Lipsky, M. (1980). *Street Level Bureaucracy: Dilemmas of the Individual in Public Services*. New York: Russell Sage Foundation.
- Logg, J. M. , Minson, J. A. , & Moore, D. A. (2019). Algorithm Appreciation: People Prefer Algorithmic to Human Judgment. *Organizational Behavior Human Decision Processes*, 151: 90–103.
- Lünich, M. , & Starke, C. (2020). Artificial Intelligence for Political Decision-making in the European Union: Effects on

- Citizens' Perceptions of Input, Throughput, and Output Legitimacy. *Data & Policy*, 2: e16.
- Lyell, D. , & Coiera, E. (2017). Automation Bias and Verification Complexity: A Systematic Review. *Journal of the American Medical Informatics Association*, 24(2): 423-431.
- Margetts, H. , & Dorobantu, C. (2019). Rethink Government with AI. *Nature*, 568(7751): 163-165.
- Meijer, A. , & Wessels, M. (2019). Predictive Policing: Review of Benefits and Drawbacks. *International Journal of Public Administration*, 42(12): 1031-1039.
- Meynhardt, T. (2009). Public Value Inside: What Is Public Value Creation? *International Journal of Public Administration*, 32(3-4): 192-219.
- Nishant, R. , Kennedy, M. , & Corbett, J. (2020). Artificial Intelligence for Sustainability: Challenges, Opportunities, and a Research Agenda. *International Journal of Information Management*, 53: 102104.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishing Group.
- Payne, J. W. , Bettman, J. R. , & Johnson, E. J. (1992). Behavioral Decision Research: A Constructive Processing Perspective. *Annual Review of Psychology*, 43(1): 87-131.
- Porumbescu, G. A. , Lindeman, M. I. , Ceka, E. , & Cucciniello, M. (2017). Can Transparency Foster More Understanding and Compliant Citizens? *Public Administration Review*, 77(6): 840-850.
- Sharma, G. D. , Yadav, A. , & Chopra, R. (2020). Artificial Intelligence and Effective Governance: A Review, Critique and Research Agenda. *Sustainable Futures*, 2: 100004.
- Sun, T. Q. , & Medaglia, R. (2019). Mapping the Challenges of Artificial Intelligence in the Public Sector: Evidence from Public Healthcare. *Government Information Quarterly*, 36(2): 368-383.
- Sundar, S. S. (2008). *The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility: MacArthur Foundation Digital Media and Learning Initiative*. Cambridge: The MIT Press.
- Tetlock, P. , Kristel, O. , Elson, S. , Green, M. , & Lerner, J. (2000). The Psychology of the Unthinkable: Taboo Trade-offs, Forbidden Base Rates, and Heretical Counterfactuals. *Journal of Personality and Social Psychology*, 78(5): 853.
- Tetlock, P. E. (2003). Thinking the Unthinkable: Sacred Values and Taboo Cognitions. *Trends in Cognitive Sciences*, 7(7): 320-324.
- Toreini, E. , Aitken, M. , Coopamootoo, K. , Elliott, K. , Zelaya, C. G. , & Van Moorsel, A. (2020). The Relationship Between Trust in AI and Trustworthy Machine Learning Technologies. *Paper Presented at the Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 272-283.
- Tummers, L. , & Bekkers, V. (2014). Policy Implementation, Street-level Bureaucracy, and the Importance of Discretion. *Public Management Review*, 16(4): 527-547.
- Van Dijck, J. , Poell, T. , & De Waal, M. (2018). *The Platform Society: Public Values in a Connective World*. Oxford: Oxford University Press.
- Vogl, T. M. , Seidelin, C. , Ganesh, B. , & Bright, J. (2020). Smart Technology and the Emergence of Algorithmic Bureaucracy: Artificial Intelligence in UK Local Authorities. *Public Administration Review*, 80(6): 946-961.
- Wachter, S. , Mittelstadt, B. , & Russell, C. (2021). Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-discrimination Law and AI. *Computer Law Security Review*, 41: 105567.
- Wilson, H. J. , & Daugherty, P. R. (2018). Collaborative Intelligence: Humans and AI Are Joining Forces. *Harvard Business Review*, 96(4): 114-123.
- Young, M. , Bullock, J. , & Lecy, J. (2019). Artificial Discretion as a Tool of Governance: A Framework for Understanding the Impact of Artificial Intelligence on Public Administration. *Perspectives on Public Management and Governance*, 2(4): 301-313.
- Zekić-Sušac, M. , Mitrović, S. , & Has, A. (2021). Machine Learning Based System for Managing Energy Efficiency of Public Sector as an Approach Towards Smart Cities. *International Journal of Information Management*, 58: 102074.
- Zuiderwijk, A. , Chen, Y.-C. , & Salem, F. (2021). Implications of the Use of Artificial Intelligence in Public Governance: A Systematic Literature Review and a Research Agenda. *Government Information Quarterly*: 38(3), 101577.

责任编辑：张书维

英文目录与摘要

JPA Journal of Public Administration, Vol. 18 No. 4, 2025

●SYNOPSIS: Artificial Intelligence and Government Trust

When AI Becomes a “Public Servant”: The Impact of Digital Humans Types in Government Services on Government Trust Shuting Li & Qian Qiu

Abstract In the context of digital empowerment in government services, AI digital humans, as a new type of intelligent agent, are applied in diverse government service scenarios, contributing to the improvement of service quality and efficiency. However, the question of which type of government service digital human can foster greater public trust in the government, and how different types of digital humans influence public trust through various mechanisms, remains largely theoretical. Based on the theoretical models constructed by the theory of mind perception and trust transfer theory, this study explores the mechanism by which the type of government service digital person (emotion-oriented and task-oriented) influences public trust in government through three experimental studies. The study found that: (1) compared to task-oriented government service digital persons, emotion-oriented government service digital persons can lead to higher levels of public trust in government. (2) there is no significant difference in the public's perception of competence between the two types of government service digital persons; however, compared to emotion-oriented government service digital persons, the public has lower warmth perception of task-oriented government service digital persons. Warmth perception plays a mediating role in influencing public government trust through different types of government service digital persons. (3) the public's past experiences of government-citizen interactions (positive experiences and negative experiences) play a moderating role in influencing warmth perceptions and competence perceptions of different types of government service digital persons. This study not only enriches the research related to digital person types and government trust in government service scenarios, but also provides practical guidance for government departments to balance professionalism and emotional interaction in the process of deploying digital persons for government services, emphasizing the need to balance technological efficiency and humanistic care in the process of digital transformation of government services, to ensure the continuous optimization and effectiveness of government services.

Key Words Digital Humans; Government Trust; Warmth Perception; Competence Perception; Experiments

Do Citizens Trust “Humans” or “Artificial Intelligence” More in the Public Sector? A Survey Experiment Study Considering Task Complexity and Public Domains

..... Siqi He & Bingsheng Liu

Abstract The rapid development of artificial intelligence (AI) is reshaping the operational models and functional scope of public administration. However, theoretical frameworks and empirical studies lag behind practical applications. In recent years, the widespread adoption of AI in public administration has sparked considerable controversy, with growing debates on its value and potential risks. At present,

systematic investigations into citizens' attitudes toward AI applications in public administration remain limited. This study employed a 2×2×2 experimental survey to examine the differences in citizens' trust toward human officials and AI in public sector scenarios, while analyzing the moderating effects of task complexity and domain values. Results show that citizens' generally exhibits higher levels of trust in human officials. Further analysis reveals that in contexts involving solemn values such as life and safety, trust in human agents significantly exceeds that in AI; whereas in domains concerning secular values such as money and material interests, the trust gap is not statistically significant. These findings provide a theoretical foundation for the scientific deployment of AI in public administration and offer practical guidance for its effective governance.

Key Words Artificial Intelligence; Citizen Trust; Task Complexity; Public Domains

●ARTICLES

How Does Algorithmic Feature Disclosure Affect Public Decision-Making Preferences? An Empirical Analysis Based on a Conjoint Experiment Ge Wang, Li Xiang & Zhejun Zhang

Abstract Algorithmic decision-making has effectively enhanced the scientific and precise nature of government decision-making, but it has also aroused public concerns about decision-making bias, privacy invasion, ethical dilemmas, and monitoring failures. How to effectively implement algorithmic disclosure to enhance the public's preference for public sector algorithmic decision-making has become an important challenge for the current government. In this paper, we analyse the causal effect of different dimensions of algorithmic disclosure on the public's preference for algorithmic decision-making based on a representative sample of 1, 356 people and a conjoint experimental design around a public healthcare service scenario with the theoretical support of signaling theory and disclosure effect, and examine the mediating mechanism of perceived algorithmic trust and perceived algorithmic transparency in this process. It was found that the more critical the decision-making stage of disclosing algorithmic applications, the more pronounced its potential negative disclosure effect. In contrast, disclosing different levels of algorithmic engagement, algorithmic responsibility assignment, and algorithmic performance information produced a complex dual effect. Perceived algorithmic trust mediated high levels of algorithmic engagement, and algorithmic performance disclosure, while perceived algorithmic transparency mediated different levels of algorithmic performance information disclosure. This study adds to the research on the impact of algorithmic feature disclosure on the public's algorithmic decision-making preferences and provides countermeasure suggestions for the effective implementation of algorithmic disclosure in the public sector.

Key Words Algorithmic Decision-making; Disclosure Effect; Perceived Algorithmic Transparency; Perceived Algorithmic Trust; Conjoint Experiments

Digitalization and Changes in Multi-level Governance Modes: A Case Study of Air Pollution Governance Jing Chen, Can Xu & Wei Hong

Abstract The different distribution and combination of control rights at different administrative levels have shaped multi-level modes of governance in China. Based on field observations and interviews on air pollution control in DM City from 2018 to 2020, this study finds that the deep embedding of digital technology in air pollution governance has reduced the information asymmetry between multi-level governments, alleviated the difficulty of supervising the lower-level departments by the upper-level departments, compressed the negotiation space between the superiors and the subordinates. These changes have moved the control right up along the three dimensions. In the past, this tight-coupling

mode was usually accompanied by a highly mobilized governance mechanism. However, the application of digital technology normalizes the mobilized mode, making it possible to maintain the tightly coupled mode for a long time. As a result, the governance mode of air pollution has evolved from the traditional administrative contracting mode and loose-coupling mode to a routine mode of tight-coupling, which has greatly reduced the flexibility of governance to a certain extent. By observing the process of introducing digital technology into the governance mechanism, this study reveals the adjustments and changes of the multi-level modes of governance.

Key Words Multi-level Modes of Governance; Control Right; Digital Technology; Air Pollution Control

Why Is the Rural Digital Transformation Is Suspending? Case Studies Based on the Policy Implementation Process Liping Fu, Kairong Kang, Huachao Feng & Yongqing Dong

Abstract Rural digital transformation provides a pathway to modernise agriculture and rural areas. Previous mainstream research has attributed the suspension of rural digital transformation to the lack of farmers' digital literacy and weak rural digital infrastructure. The paper shows, through a stressful case study of goals, conditions and change, that the suspension of rural digital transformation from central policy design to local policy implementation to rural policy effectiveness is essentially due to the governance tension of the hourglass structure in the policy implementation process, which creates 'Softening of Grassroots Institutions'. This means that under the logic of balancing efficiency and effectiveness in policy implementation, the national institutional system has been expanded and alienated in the process of sinking down to the grassroots level and landing in rural areas, and the effectiveness of institutional constraints has been reduced. The study finds that the 'Softening of Grassroots Institutions' triggers the suspension of the digital transformation of the countryside through three logics, namely, the logic of goal advancement of the order of power, downward pressure on tasks, and the administrativeisation of governance; the logic of condition creation of the legitimacy of tasks, procedural compliance, and the marginalisation of small farmers, and the logic of reversal of pressure, conflict of pressures, and the tacticalisation of goals, and the logic of change outcome of reverse pressure, pressure conflict and goal strategisation. From the perspective of the whole process of policy implementation, the paper responds to the question of 'why rural digital transformation is suspending', which helps to expand and deepen the theoretical understanding of rural digital transformation in the local context of China.

Key Words Rural Digital Transformation; Digital Formalism; Policy Implementation; Softening of Grassroots Institutions; Basic Practical Conditions

The Multiple Logics Behind Resource Allocation by Grassroots Governments: A Case Study of County A in Shandong Province Zongfeng Sun, Jingqi Lin & Xiaohan Zhang

Abstract How grassroots governments allocate various fiscal resources is of great significance to regional industrial development and rural revitalization. This study builds a framework of multiple logics to explain these allocation behaviors, focusing on three dimensions: institutional pressure, interest structures, and social relationships. Based on field observations and comparative case studies in County A, the research identifies three distinct logics in resource allocation: a mandatory logic driven by rigid institutional rules, a performance-oriented logic influenced by administrative goals, and a relational logic shaped by local social hierarchies. These logics often overlap and interact in practice. Specifically, the mandatory logic plays a dominant role in allocating both universal and targeted

resources, especially the former. The performance-oriented and relational logics mainly affect how targeted resources are distributed.

Key Words Grassroots Resource Allocation; Multiple Logics; Project-Based System

How Does the Policy Mix Intensity Affect Policy Effectiveness? A Study Based on Air Pollution Control Policy Mix in China Xiaojie Zhang & Xiaoyu Liu

Abstract An appropriate arrangement of policy mix intensity is both the theoretical focus of policy design and a practical challenge in achieving high policy effectiveness. This paper takes China's air pollution control policy mix as a research sample, measuring the intensity of policy mix by using the adapted Index of Policy Activity. Based on the panel data of 31 provincial-level administrative regions from 2010 to 2022, it empirically analyzes the impact and mechanism of the policy mix intensity on policy effectiveness. The findings indicate that the intensity of the air pollution control policy mix exhibits a non-linear inverted U-shaped relationship with policy effectiveness. Both the resource effect and information effect serve as significant mechanisms through which the intensity of the policy mix influences policy effectiveness. The heterogeneity analysis reveals that the inverted U-shape relationship is more significant in the eastern and central regions, while a positive linear correlation is observed in the western region. In areas characterized by a lower degree of marketization, the effect of the intensity of policy mix on policy effectiveness is more remarkable. Moreover, the centrality of regional cooperation networks can significantly influence this relationship. Enhancing the centrality of these networks has the potential to reverse the inverted U-shaped relationship. The analytical results not only contribute to a better understanding of the effects and mechanisms of policy mix intensity on policy effectiveness across different contexts, but also offer valuable insights for improving policy mix design and ensuring the sustainable enhancement of policy effectiveness.

Key Words Policy Mix Intensity; Policy Effectiveness; Air Pollution Control Policy; Resource Effect; Information Effect

Can Social Organizations Development Effectively Alleviate the Income Gap among Residents? An Empirical Test Based on the Threshold Effect of Government Intervention

..... Lei Liu, Xinya Zhang & Shuyi He

Abstract Various types of social organizations play an active role in income distribution through diverse means, such as assisting disadvantaged groups, promoting employment and entrepreneurship, and protecting labor rights. However, most existing studies focus on the role of social organizations in the third distribution, with limited empirical examination of their actual effects on the income gap among residents. Using panel data from 31 provinces (excluding Taiwan Province, HongKong Special Administrative Region, and Macao Special Administrative Region) in China between 2005 and 2021, this paper empirically investigates the impact of social organizations development on the income gap among residents and explores the threshold effect of government intervention through panel benchmark regression and threshold regression models. It is found that the current development of social organizations has failed to curb the persistent widening of the income gap. However, increased government intervention can reverse this effect. When the intervention level is below the threshold of 0.2662, social organizations have failed to contain the widening income gap. Once this threshold is exceeded, social organizations contribute significantly in alleviating the income gap. Additionally, through heterogeneity analysis of the impact of social organizations on income gap by type and region, it is found that foundations, social service organizations, and social organizations in Northeast China have

not effectively curbed the widening income gap. The underlying reason lies in their early-stage development, functionally concentrated primarily on the third distribution while not yet engaging in primary distribution or in secondary distribution. The paper reveals the important role of government intervention in the process of income gap regulation by social organizations, further enriches polycentric governance theory, and provides new perspectives on the exploration of the co-governance model of “an active government and a dynamic society”.

Key Words Social Organizations Development; Income Gap; Government Intervention; Threshold Effect

Do Fiscal Vertical Imbalances Reduce Local Government Attention to Public Services?

..... Yuan Zhang & Lijun Liu

Abstract Achieving equalization of basic public services is a necessary step towards common prosperity, and the supply of basic public services is closely related to the fiscal situation. Based on the panel data of 31 provinces (excluding Taiwan Province, HongKong Special Administrative Region, and Macao Special Administrative Region) in China from 2007 to 2020, and based on the current situation of China’s fiscal decentralization system, this paper analyzes the text of the government’s work report and constructs a system of attention measurement indicators. It systematically examines the mechanism through which fiscal vertical imbalance affects local governments’ attention to public services, both theoretically and empirically. The study finds that the fiscal vertical imbalance has a significant inhibitory effect on local governments’ attention to public services. Furthermore, the higher the fiscal autonomy enjoyed by local governments, the weaker this inhibitory effect becomes. The negative impact of fiscal vertical imbalance on local governments’ attention to public services exhibits regional and environmental heterogeneity, as well as an increasingly prominent single-threshold effect. This paper provides new research ideas for clarifying local governments’ attention to public services through the new perspective of quantifying government attention, which is of practical significance for weakening the negative impact of fiscal vertical imbalance and promoting the realization of the goal of equalization of basic public services.

Key Words Fiscal Vertical Imbalances; Public Services; Attention; Fiscal Autonomy

●THEORETICAL REVIEWS

Cutback Management under Fiscal Pressure: A Literature Review

..... Xin Chen

Abstract When public organizations face reduced fiscal and budgetary resources, besides seeking traditional ways to increase financial revenues, how to achieve effective fiscal savings and control budget expenditures through cutback management has become an increasingly important research topic in the field of public administration. This article reviews the academic literature on cutback management both domestically and internationally, discussing the inherent characteristics, strategies, decision-making mechanisms, and impacts of cutback management, aiming to reveal the theoretical and practical progress in this field. The research finds that although the theory of cutback management has constructed a clear and content-rich system, there are still several areas that require further study.

Key Words Cutback Management; Budget Cutback; Public Financial Management; Fiscal Sustainability

公共行政评论

双月刊, 2008年创刊
第18卷, 第4期(总第106期)
2025年8月15日出版

Journal of Public Administration
Bimonthly, Since 2008
Vol.18 No.4
Published in August 2025

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



9 771674 248258