

智能政务服务的相对绩效信息 如何影响公众信任^{*}

杜亚斌^{**}

【摘要】与人工服务相比,算法驱动的智能政务服务既有优势,又有不足,一定程度上导致了算法欣赏与算法厌恶并存的现象。因此,为探究智能政务服务相比人工服务的相对绩效信息对公众信任的影响,本文基于个体工商户“秒批”改革这一现实案例,开展两项调查实验,以传统人工服务作为参照点,实证检验智能政务服务的相对绩效信息对公众信任(包括算法信任和政府信任)的影响及作用机制。研究发现如下。(1)信息面向会显著影响相对绩效信息对公众信任的作用,仅呈现智能政务服务相对优势的单面相对绩效信息不会显著影响公众算法信任和政府信任,而同时呈现智能政务服务相对优势与不足的双面相对绩效信息则显著降低了公众信任。(2)尽管双面相对绩效信息对公众信任的直接效应为负,但其仍能通过提升主观绩效认知间接提升公众信任,这表明理性认知对绩效信息作用具有重要影响。(3)在确认偏差与反确认偏差的双重影响下,公众对人工智能的乐观倾向增强了单面相对绩效信息的正向影响,削弱了双面相对绩效信息的负向影响。上述发现丰富了智能政务服务场景中的绩效信息研究,揭示了理性与非理性因素对公众绩效信息处理的共同作用,凸显出智能政务服务推广过程中信息沟通策略的重要性。

【关键词】智能政务服务;相对绩效信息;信息面向;认知偏差;政府信任

一、引言

近年来,随着大数据和算法技术在公共部门广泛引入,智能识别、智能问答、智

^{*} 本研究受国家社会科学基金青年项目“智能政务服务与政府信任的技术社会学研究”(项目批准号:24CSH011)的资助。感谢匿名评审专家和编辑部的宝贵意见。

^{**} 杜亚斌,湖南大学公共管理学院助理教授,duyabin@hnu.edu.cn。

引文格式:杜亚斌. 2026. 智能政务服务的相对绩效信息如何影响公众信任[J]. 公共管理评论,8(2): 147-172.

Cite this article: Du Y B. 2026. The impact of relative performance information on public trust in AI-enabled public services[J]. *China Public Administration Review*, 8(2): 147-172. (in Chinese)

能审批等各类智能政务服务蓬勃发展(李晓方等,2020;Chandra and Feng,2026),不仅深刻影响着政府自身的组织运作模式(Meijer et al.,2021;吴进进和何包钢,2023),还重塑着政民关系和公众对政府的信任(Miller et al.,2023;樊博和李晶晶,2024)。相关实证研究表明,智能政务服务的运用既可能提升公众信任(Kennedy et al.,2022;冉龙亚等,2022),也可能降低公众信任(Ingrams et al.,2022;Haesevoets et al.,2024),因此呈现出算法欣赏与算法厌恶并存的特征。

从委托-代理理论的角度看,向公众精准提供智能政务服务的绩效信息是消除算法厌恶、提升政府信任的有效手段(Yang and Holzer,2006;周豪和包国宪,2021)。然而,相较于传统人工服务,算法驱动的智能政务服务固然具有办事效率高、服务成本低、人为干预少等比较优势(Horvath et al.,2023;刘特等,2024a),但也可能带来决策失误、系统偏见、算法黑箱等各类风险和问题(Schiff et al.,2022;张成福和王祥州,2023)。因此,如何呈现智能政务服务相比于人工服务的相对绩效信息,可能会对公众信任产生重要影响。此外,公众对信息的处理并非完全理性无偏,而是深受各类认知偏差影响(Andersen and Hjortskov,2016),相同绩效信息对不同认知倾向公众的影响可能不同,因此也需考虑认知偏差在智能政务服务绩效信息与公众信任关系中的作用。

有鉴于此,本文基于个体工商户“秒批”改革这一现实案例,采用调查实验法开展两项独立子实验,以传统人工服务为参照点,实证检验智能政务服务相对绩效信息的不同呈现对公众信任的影响及作用机制,以期深化对智能政务服务场景中绩效信息与公众信任关系的认识。

二、文献综述

(一) 绩效信息与政府信任

公众对政府的信任基于两者之间的委托-代理关系(王正绪和赵健池,2021),反映的是作为委托人的公众对作为代理人的政府能否满足自身期望的基本评价(Hetherington,2004;高学德和王镇江,2024)。信息不对称是阻碍公众信任政府的关键因素(Bauhr and Grimes,2014)。自新公共管理运动以来,各国政府普遍通过公开绩效信息强化公众问责机制,以期提升政府信任(Yang and Holzer,2006;Meng and Li,2025)。相关实证研究表明,绩效信息的内容、形式和来源不同,对公众信任和态度的影响也存在差异(James and Moseley,2014;Schmidhuber et al.,2023;王思琦,2023)。

在智能政务服务研究中,绩效信息对公众信任和态度的影响也广受关注。相关实验研究表明,向被试呈现智能政务服务的准确性(Kreps and Jakesch,2023)、使用成本、时间效率(Hemesath and Tepe,2024)、透明性(Grimmelikhuijsen,2023;樊博和李晶晶,2024)、回应性及公平性(Schiff et al.,2022)等不同绩效信息,均会影响公众对

智能政务服务的信任和接受度。此外,联合实验研究还表明,不同方面的绩效信息对公众影响的程度存在差异,其中,智能政务服务结果的准确性对公众影响较大,而透明度、参与度等过程性特征对公众影响较小(Horvath et al.,2023;König et al.,2024)。

(二) 认知偏差与绩效信息处理

公众对绩效信息的处理并非完全理性无偏,其受各类认知偏差的影响(Andersen and Hjortskov,2016)。其中,确认偏差与反确认偏差尤其受到关注。根据动机性推理理论和认知失调理论(Kunda,1990),公众倾向于接受与既有认知及偏好一致的信息,而怀疑甚至拒绝与之冲突的信息,学界将前一种倾向称为确认偏差(confirm bias),将后者称为反确认偏差(disconfirmation bias)。相关实验研究表明,公众的政治倾向(James,2011;Porumbescu et al.,2023)、反公共部门偏见(Baekgaard et al.,2020;Brogaard et al.,2024)等都可能引发(反)确认偏差,进而影响公众对绩效信息的解读及对政府的评价。

在智能政务服务研究中,绩效信息处理的(反)确认偏差也得到了证实。孙源等(2024)基于志愿服务算法推荐开展的调查实验研究表明,当算法推荐逻辑与公众志愿服务动机一致时,公众算法信任度更高。Selten et al.(2023)基于警力部署预测的调查实验研究表明,当算法建议与警务人员自身专业判断一致时,后者对算法建议的信任度更高。Alon-Barkat and Busuioc(2023)基于教师续聘决策场景的调查实验研究表明,对少数族裔的负面刻板印象会削弱公众采纳算法建议的意愿。

(三) 现有研究不足与本研究进路

综上所述,围绕绩效信息对公众信任和态度的影响,现有研究已取得较为丰富的成果,但在智能政务服务领域,相关研究尚处于起步阶段,还存在较大的可拓展空间。具体而言,一方面,智能政务服务相关研究大多仅关注绩效信息的效价,即绩效的高低好坏对公众信任和技术接受度等的影响,却忽视了绩效信息其他特征的作用。另一方面,在探究认知偏差的影响时,现有研究往往忽视了公众认知的理性一面,未同时考察理性与非理性因素(认知偏差)对公众绩效信息处理的共同作用。

有鉴于此,本文基于个体工商户“秒批”改革这一现实案例,开展两项调查实验研究,以传统人工服务为参照点,实证检验智能政务服务的相对绩效信息及其信息面向对公众信任的影响,并探究理性认知和认知偏差在其中的共同作用机制,以期拓展智能政务服务场景中的绩效信息研究边界。

三、理论分析与研究假设

(一) 多重委托—代理视角下智能政务服务中的公众信任

随着政务服务智能化转型,传统的公众与政府之间的委托—代理关系转变为公

众、政府与智能算法三者间的多重委托-代理关系,公众信任的对象也从政府部门扩展到智能算法,从而形成政府信任与算法信任并存的双层信任结构(Grimmelikhuijsen, 2023; 樊博和李晶晶, 2024)。一方面,传统的公众与政府之间的委托-代理关系仍然存在,智能政务服务成效仍然会影响公众对政府的信任。另一方面,政府与智能算法之间形成新的委托-代理关系(汤志伟等, 2023),智能算法以强大的技术自主性扮演着行政代理人角色(刘玮和王锋, 2024),促使公众形成算法信任,即对基于智能算法的政务服务系统的信任。因此,本文同时将政府信任和算法信任作为因变量。

(二) 智能政务服务的相对绩效信息与公众信任

从委托-代理理论的角度看(Bauhr and Grimes, 2014),向公众披露智能政务服务的绩效信息,有助于缓解公众与政府及其所使用算法之间的信息不对称问题,降低公众的怀疑和不确定感,进而增进公众对智能算法和政府的信任。

然而,绩效信息的作用受到形式和内容影响。根据前景理论的参照依赖观点(Kahneman and Tversky, 1979),公众对公共服务绩效的评价并非基于绝对水平,而是基于其相对特定参照点的高低变化。相较于绝对绩效信息,相对绩效信息的作用更为突出(James and Moseley, 2014; Olsen, 2017)。因此,本文将传统人工服务视为智能政务服务的社会参照点,进而分析智能政务服务的相对绩效信息对公众信任的影响。

与传统人工服务相比,智能政务服务既有优势,又有不足(Schiff et al., 2022; 张成福和王祥州, 2023),从而导致正面与负面相对绩效信息并存。基于此,本文引入信息面向(message-sidedness)概念。该概念起源于市场营销领域,用以描述企业在宣传产品或服务时仅呈现其正面特征,还是同时呈现其正反两方面的特征,前者被称为单面信息(one-sided message),后者被称为双面信息(two-sided message)(Eisend, 2007; Han, 2025)。与之对应,本文将智能政务服务的相对绩效信息分为两类,一类仅呈现智能政务服务相对人工服务的优势(后文称之为单面相对绩效信息),另一类同时呈现智能政务服务相对人工服务的优势与不足(后文称之为双面相对绩效信息)。

不同面向的相对绩效信息分别会如何影响公众信任呢?要回答这一问题,需同时考虑信息效价和信息可信度两个方面。从信息效价的角度来看,既有研究表明,正面和负面绩效信息分别会提升和降低公众信任(Alessandro et al., 2021)。单面相对绩效信息凸显了智能政务服务的相对优势,因而可能会提升公众对智能算法和政府的信任。而受负面主导效应(negativity dominance)这一负面偏见的影响(Rozin and Royzman, 2001),双面相对绩效信息中负面信息的作用可能更强,因而可能降低公众的算法信任和政府信任。

从信息可信度的视角来看,根据归因-对应理论(attribution-correspondence theory)(Jones and Davis, 1965),在面对单面相对绩效信息时,公众倾向于进行非对应归因(外部归因),即认为这是政府出于推广智能政务服务目的进行的宣传,而非真

实绩效,故而公众对信息可信度的感知较低;而在面对双面相对绩效信息时,公众更倾向于进行对应归因(内部归因),即认为这是智能政务服务绩效的真实写照,故而对信息可信度的感知更高(Eisend,2007)。

综上所述,尽管智能政务服务的单面相对绩效信息有助于提升公众信任,但在现实情境中公众对其信息可信度的感知较低,因而其实际影响可能并不显著;受负面偏见影响,智能政务服务的双面相对绩效信息对公众信任的影响方向总体为负,且信息可信度较高,因而可能会显著降低公众对智能算法和政府的信任。据此,本文提出假设1和假设2。

H1: 与不提供信息相比,呈现智能政务服务的单面相对绩效信息不会显著影响公众的算法信任(H1-1)和政府信任(H1-2)。

H2: 与不提供信息相比,呈现智能政务服务的双面相对绩效信息会显著降低公众算法信任(H2-1)和政府信任(H2-2)。

需要指出的是,潜在的负面偏见并不会完全抹杀理性认知的作用。如前所述,在委托-代理理论的理性选择框架下,认知是信任的基础,只有当委托人(公众)对代理人(智能算法、政府)有充分了解时,才能建立起信任。双面相对绩效信息尽管暴露出智能政务服务的短板,但也使公众对智能政务服务绩效有了更完整、更准确的了解,从而有助于缓解信息不对称带来的不确定感,提升公众对智能算法和政府的信任。Grimmelikhuijsen(2012)在荷兰开展的一项实验研究也表明,绩效信息透明能通过提升公众的主观绩效认知,进而提升政府信任。据此,本文提出假设3。

H3: 智能政务服务的双面相对绩效信息能够通过提升公众的主观绩效认知,进而提升公众的算法信任(H3-1)和政府信任(H3-2)。

(三) 人工智能认知倾向的调节效应

不同于委托-代理理论对理性认知的强调,认知心理学认为公众的信息处理并非完全理性无偏,而易受各类认知偏差影响(Andersen and Hjortskov,2016)。根据动机性推理理论(Kunda,1990),公众对绩效信息的处理受到(反)确认偏差的影响,与既有观念一致的信息更容易被公众接受,而与既有观念冲突的信息更可能被公众质疑或拒绝(Baekgaard et al.,2020)。

基于此,本文考察人工智能认知倾向如何影响智能政务服务相对绩效信息与公众信任之间的关系。现有实证研究表明,公众对人工智能作用的认知越乐观,越支持使用智能服务(Ahn and Chen,2022);反之,越担忧人工智能带来的风险,公众对智能服务的信任和接受度就越低(Gesk and Leyer,2022)。因此,遵循确认偏差的观点(Kunda,1990),公众对人工智能的认知越乐观,越易认同并接受智能政务服务的正面绩效信息,同时越易质疑并忽视其负面绩效信息,从而增强单面相对绩效信息对算法信任和政府信任的正向影响,削弱双面相对绩效信息中负面信息的影响。据此,本文提出假设4和假设5。

H4: 公众对人工智能的乐观认知倾向会增强智能政务服务的单面相对绩效信息对算法信任(H4-1)和政府信任(H4-2)的正向影响。

H5: 公众对人工智能的乐观认知倾向会削弱智能政务服务的双面相对绩效信息对算法信任(H5-1)和政府信任(H5-2)的负向影响。

四、研究设计

(一) 实验设计

本研究基于个体工商户“秒批”改革这一现实场景,开展了两项单因素组间调查实验,通过操纵智能政务服务相对绩效信息的信息面向,对研究假设进行实证检验。调查实验是一种将调查法和实验法相结合的创新方法,将实验干预嵌入调查问卷之中,通过操纵向被试呈现的问卷内容实施干预。该方法简便易行、便于施测,是当前公共管理研究中应用最广泛的实验类型(李晓倩等,2022)。

1. 实验场景

本文选取个体工商户“秒批”改革作为实验场景。2019年3月20日,深圳市市场监督管理局在全国率先推出个体工商户设立、注销“秒批”服务,依托大数据、人工智能等技术,实现了网上申请、无人工干预自动审批。此后,不少地区相继推出了“秒批秒办”的政务服务,这种自动化和智能化的行政审批模式逐渐流行起来。

本研究以这一案例为蓝本进行实验设计,原因有二。其一,行政审批是智能政务服务的典型场景,在既有研究和现实中都广受关注(Grimmelikhuijsen, 2023; 刘特等, 2024b),且较易为公众理解。其二,以现实案例为实验场景,能在一定程度上避免调查实验中广泛存在的真实性不足问题。本研究实验材料中的表述大多基于新闻报道或学术论文,少数地方根据实验需要进行了调整和拓展,以兼顾研究的内外效度。

2. 实验流程

本研究遵循图1的流程开展调查实验。第一步,询问被试的背景信息,包括性别、年龄、受教育程度等人口社会学特征,以及对人工智能的使用情况和认知倾向。之所以先让被试填答背景信息,是为了避免后续实验影响被试对人工智能相关问题的回答。第二步,基于智能政务服务相对绩效信息的不同情况将实验情景分为三组(无绩效信息、单面相对绩效信息、双面相对绩效信息),随机向被试呈现一种实验情景,实施实验干预。第三步,在实验干预完成后,被试需再填答注意力检验和实验操纵检验问题,以及公众信任和主观绩效认知的相关问题。

3. 实验材料与实验操纵

1) 实验1

实验1首先向所有被试呈现如下基本信息:

为进一步优化营商环境,A市市场监管局探索开展个体工商户设立、注销“秒

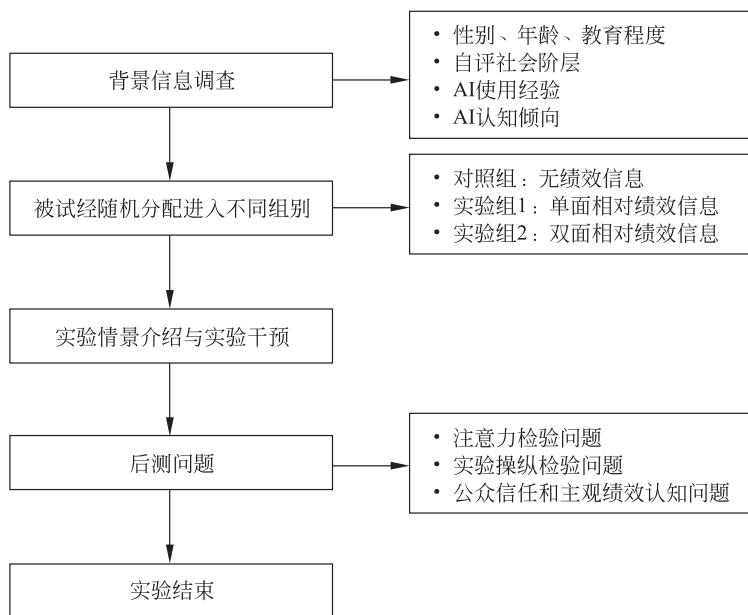


图1 调查实验流程

批”服务,利用人工智能技术,实现了网上申请、无人工干预的自动化审批^①。

在上述基本信息之外,本研究根据实验组别的不同,向被试呈现智能政务服务的不同绩效信息。对无绩效信息组,除基本信息外,不提供任何额外信息;而单面相对绩效信息组的被试还将看到下述内容:

“秒批”系统大大提升了审批效率,将个体工商户的业务申请办理时长从以往的2个工作日缩短至5分钟,推动了A市中小微企业数量井喷式增长^②。

上述文本仅强调了“秒批”系统相比人工服务的效率优势,未呈现其不足,因而属于单面相对绩效信息。

双面相对绩效信息组的被试则会看到内容类似但信息量更全面的文本^③,具体文本如下:

“秒批”系统大大提升了审批效率,将个体工商户的业务申请办理时长从以往的2个工作日缩短至5分钟,推动了A市中小微企业数量井喷式增长。但由于“秒批”系统存在“求简求快”的倾向,一些安全漏洞没有像人工审核那样被审慎思考和判断,虚假注册、虚开发票问题日益凸显。

上述文本在强调“秒批”系统审批效率高于人工审核的同时,也指出其不如人工审核审慎,结果存在出错风险,两者结合意味着与人工审核相比,“秒批”系统既有优

① 中国政府网:《深圳市市场监管局在全国率先开展个体工商户“秒批”服务》,2019年3月26日, http://www.gov.cn/xinwen/2019-03/26/content_5377778.htm,2025年2月13日访问。

② 这一实验操纵的内容来源与基本信息部分相同。

③ 双面相对绩效信息组的文本内容参考了吴进进和何包钢(2023)。

势,又有不足,从而构成兼具正负两方面特征的双面相对绩效信息。

2) 实验 2

尽管实验 1 大体实现了对智能政务服务相对绩效信息的操纵,但具体设计还存在一定瑕疵。就对单面相对绩效信息的操纵而言,实验 1 采用“业务办理时长从以往的 2 个工作日缩短至 5 分钟”的表述,同时混杂了“秒批”服务与人工服务的横向对比,以及当前与历史绩效的纵向对比。就对双面相对绩效信息的操纵而言,正面的审批时长信息采用精准的数字表述,而负面信息则以模糊化、情绪化的定性描述为主,未提供准确率、失误率等量化可比指标。这种操纵方式不一致的设计可能引发形式偏倚,导致实验结果可能源于信息表达风格的差异,而非绩效信息面向的变化^①。

基于此,本文对实验 1 进行了调整优化,进而设计开展了实验 2。在实验 2 中,各组获得的基本信息与实验 1 一致,但单面相对绩效信息和双面相对绩效信息的操纵方式发生了变化。

对于单面相对绩效信息组,除基本信息外,被试还将看到以下内容:

与传统人工服务相比,“秒批”服务的效率大大提升,通过传统人工服务办理个体工商户的业务申请平均时长为 2 个工作日,而通过“秒批”系统办理该业务平均仅需 5 分钟(如表 1 所示)。

表 1 人工服务与“秒批”服务的办事绩效比较

	传统人工服务	“秒批”服务
办事时长	2 个工作日	5 分钟

上述文本明确对人工服务与“秒批”服务进行了比较,凸显了“秒批”系统的效率优势,同时采用表格形式降低被试的信息加工成本。

双面相对绩效信息组的被试则会看到内容类似但信息量更全面的文本,具体如下:

与传统人工服务相比,“秒批”服务的效率大大提升,通过传统人工服务办理个体工商户的业务申请平均时长为 2 个工作日,而通过“秒批”系统办理该业务平均仅需 5 分钟。但与传统人工服务相比,“秒批”系统审批结果的准确率相对较低,传统人工服务审批结果的准确率约为 99%,而“秒批”系统审批结果的准确率约为 90%(如表 2 所示)^②。

表 2 人工服务与“秒批”服务的办事绩效比较

	传统人工服务	“秒批”服务
办事时长	2 个工作日	5 分钟
结果准确率	99%	90%

① 感谢匿名评审人对这一点的指正和建议。

② 此处对结果准确率的操纵并非来自真实统计数据,而是根据实验需要,参照已有研究(Horvath et al., 2023; König et al., 2024)进行的折中设定。

上述文本从办事时长和结果准确率两个方面对“秒批”服务与人工服务进行了比较,既体现了“秒批”服务的效率优势,又揭示出其在准确率上存在的不足,进而构成兼具正负维度的双面相对绩效信息。

(二) 变量测量

1. 因变量

本文的因变量为公众信任,包括公众的算法信任和政府信任。借鉴既有研究(Selten et al.,2023;樊博和李晶晶,2024)成果,采用被试对下列三个表述的同意程度衡量算法信任:(1)我相信“秒批”系统能够做出正确的审批决定;(2)我相信“秒批”系统能够公正处理所有审批材料;(3)我认为“秒批”系统是值得信任的。其中,前两个问题分别体现了被试对“秒批”系统能力与公正的信任,第三个问题则是总体信任评价。上述问题的选项均采用7点李克特量表(1=完全不同意,7=完全同意),且具有较高内部一致性(Cronbach's $\alpha=0.869$,实验1;Cronbach's $\alpha=0.837$,实验2),因此本文取三者均值衡量公众的算法信任。

参考既有研究(刘特等,2024b)成果,本文根据被试对下列三个表述的同意程度衡量政府信任:(1)我认为A市市场监管局的业务能力很强;(2)我认为A市市场监管局全心全意为老百姓服务;(3)我认为A市市场监管局切实履行了自身职责。上述三个问题分别反映了被试对A市市场监管局能力、亲民和诚信三个方面的评价,问题选项均采用7点李克特量表(1=完全不同意,7=完全同意),且具有较高内部一致性(Cronbach's $\alpha=0.875$,实验1;Cronbach's $\alpha=0.819$,实验2),因此本文取三者均值衡量公众的政府信任。

2. 自变量

本文的自变量为智能政务服务的相对绩效信息,即实验干预变量。如实验设计部分所述,智能政务服务的相对绩效信息分为三种情形,分别是无绩效信息、单面相对绩效信息和双面相对绩效信息,在实证分析中本研究以无绩效信息组为参照组,将其他两组分别与参照组进行比较。

3. 中介变量

本文的中介变量为公众的主观绩效认知,该变量仅在实验2中进行了测量。参考既有研究(Grimmelikhuijsen,2012),本文根据被试对下列两个表述的同意程度进行测量:(1)通过上述材料,我充分了解了“秒批”服务相比人工服务的优势与不足;(2)通过上述材料,我对“秒批”服务的办事绩效有了全面充分的了解。两个问题选项均采用7点李克特量表(1=完全不同意,7=完全同意),且具有较高内部一致性(Cronbach's $\alpha=0.712$,实验2),因此本文取两者均值衡量公众的主观绩效认知。

4. 调节变量

本文的调节变量为公众对人工智能(AI)的乐观认知倾向。参考既有研究(Ahn and Chen,2022),本文同时测量公众一般性的AI乐观倾向和对政府使用AI的乐观

倾向,其中对政府使用 AI 的乐观倾向仅在实验 2 中进行了测量。一般性的 AI 乐观倾向采用下述问题测量:“您是否认同人工智能技术将使人们的生活变得更美好?”。该问题选项采用 5 点李克特量表(1=非常不认同,5=非常认同),取值越大表明对人工智能技术的认知越乐观。

对政府使用 AI 的乐观倾向采用被试对下列四个表述的同意程度衡量:(1)政府使用人工智能有助于提升行政效率;(2)政府使用人工智能有助于改善公共服务;(3)政府使用人工智能有助于提升决策质量;(4)整体而言,政府使用人工智能利大于弊。四个问题选项均采用 5 点李克特量表(1=非常不同意,5=非常同意),并且量表内部一致性尚可(Cronbach's $\alpha=0.643$,实验 2),因此本文取四者均值衡量公众对政府使用 AI 的乐观倾向。

5. 控制变量

参照既有研究,本文主要控制了三类因素。(1)人口社会学变量,这是影响信任的基础背景变量,游宇和王正绪(2014)称之为现代化因素,包括性别(1=男)、年龄和教育程度,其中教育程度采用 5 点李克特量表测量(1=初中及以下,5=研究生及以上)。

(2)自评社会阶层,反映公众对自身利益所得的整体评价,是政府信任研究普遍纳入的变量(朱春奎和毛万磊,2017;樊博和李晶晶,2024)。本文通过下述问题对其进行测量:“您认为您本人的社会经济地位在社会上大体属于哪个层次?”该问题选项采用 5 点李克特量表(1=下层,5=上层)。

(3)技术使用经验,被视为影响技术信任的重要因素(Wang et al.,2021;Ahn and Chen,2022)。本文通过下述问题对其进行测量:“您是否经常使用人工智能应用,如智能语音助手(小爱同学、Siri 等)、通用大语言模型(文心一言、ChatGPT、DeepSeek、Kimi 等)、自动驾驶等?”该问题选项采用 5 点李克特量表(1=从不,5=总是)。

(三) 数据收集

与国内外多数调查实验研究类似,本研究通过在线调查平台“Credamo 见数”获得被试样本,于 2025 年 2 月下旬完成了实验 1 的问卷收集,5 月中旬完成了实验 2 的问卷收集。在此之前,研究采用 G*Power 软件,根据调查实验的分组情况、控制变量数量及自由度等信息,将效应量设定为 0.25,显著性水平设定为 0.05,统计功效设定为 0.9,计算得出实验 1 的理想样本量为 314,实验 2 的理想样本量为 338。

在正式实验阶段,考虑到问卷回收有效率,两个实验均实际发放问卷 500 份,并通过下述问题进行注意力检验:根据上述材料,“秒批”系统的审批模式是什么?该问题选项分为“人工智能辅助人工审核”和“无人工干预自动审批”,正确选项为后者,回答错误者即视为未通过注意力检测,其问卷予以剔除。此外,本研究还剔除了作答时间异常的问卷。最终,实验 1 回收有效问卷 337 份,问卷有效率为 67.4%;实

验 2 回收有效问卷 378 份,问卷有效率为 75.6%。

五、实证分析

(一) 描述统计分析

两个实验各变量的描述统计分析结果如表 3 所示。由表 3 可知,在两个实验中,被试的算法信任和政府信任都处于中等偏上水平,均值在 7 点李克特量表上取值接近甚至超过 5,且政府信任高于算法信任。被试对“秒批”服务的主观绩效认知较高,均值在 7 点李克特量表上超过 5,表明实验材料有效提升了其对“秒批”服务的认知程度。被试对 AI 和政府使用 AI 的认知均较为乐观,均值在 5 点李克特量表上均超过 4。被试的 AI 使用经验较为丰富,均值在 5 点李克特量表上接近 4。

表 3 描述统计分析

连续变量	取值范围	实验 1	实验 2
算法信任	1~7	4.907(1.338)	5.252(1.058)
政府信任	1~7	5.205(1.191)	5.586(0.962)
主观绩效认知	1~7	—	5.015(1.392)
AI 乐观倾向	1~5	4.196(0.701)	4.288(0.554)
对政府使用 AI 的乐观倾向	1~5	—	4.232(0.480)
年龄	18~63	31.282(8.935)	32.468(9.251)
自评社会阶层	1~5	2.635(0.831)	2.749(0.723)
AI 使用经验	1~5	3.763(0.781)	3.937(0.672)
分类变量	类别	实验 1	实验 2
相对绩效信息	无绩效信息	33.83%	33.07%
	单面相对绩效信息	32.94%	33.33%
	双面相对绩效信息	33.23%	33.60%
性别	男性	57.27%	58.47%
	女性	42.73%	41.53%
教育程度	初中及以下	1.48%	1.59%
	高中、中专或技校	3.56%	3.70%
	大学专科	9.20%	9.52%
	大学本科	65.58%	68.52%
	研究生及以上	20.18%	16.67%

注:在第 3 和第 4 列中,连续变量展示的是各变量的均值,括号内为标准差,分类变量展示的是各类别所占的比例;“—”表示该变量在实验中未测量。

在样本特征方面,两个实验中女性占比均高于男性,年龄均值介于 31~33 岁之

间,教育程度较高,大部分在大学本科及以上,自评社会阶层均值介于中层和中下层之间。上述样本特征与国内相关调查实验研究大体一致(樊博和李晶晶,2024;刘特等,2024b;孙源等,2024),但与我国人口总体特征存在一定差异,因此需审慎看待本文结论的可推广性。

(二) 随机分配与实验操纵检验

对于调查实验而言,样本随机分配至关重要,因此本文对两个实验样本的随机分配情况进行了检验,结果如表 4 所示。卡方检验结果显示,在两个实验中,各组被试在性别和教育程度分布上不存在显著差异。方差分析结果显示,在两个实验中,各组被试在年龄、自评社会阶层、AI 使用经验及 AI 乐观倾向等方面均不存在显著差异。上述结果表明,两个实验的随机化干预效果良好,通过了样本平衡性检验。

表 4 样本平衡性检验

	实验 1		实验 2	
	χ^2/F 值	p 值	χ^2/F 值	p 值
性别(男性=1)	$\chi^2 = 0.093$	0.955	$\chi^2 = 0.109$	0.947
年龄	$F = 0.15$	0.858	$F = 0.73$	0.480
教育程度	$\chi^2 = 12.894$	0.116	$\chi^2 = 7.198$	0.515
自评社会阶层	$F = 0.30$	0.738	$F = 1.54$	0.216
AI 使用经验	$F = 0.86$	0.422	$F = 0.76$	0.468
AI 乐观倾向	$F = 1.44$	0.239	$F = 0.40$	0.669
对政府使用 AI 的乐观倾向	—	—	$F = 0.01$	0.993

注:“—”表示该变量在实验中未测量。

为确定实验干预的效果,本文进行操纵检验。在实验 1 中,根据被试对下列表述的同意程度进行操纵检验:“秒批”系统的引入没有产生负面影响,问题选项采用 7 点李克特量表,主要用于检验双面相对绩效信息中负面信息的干预效果。方差分析结果显示,无绩效信息组和单面相对绩效信息组的均值均显著高于双面相对绩效信息组($p < 0.001$),但无绩效信息组与单面相对绩效信息组之间不存在显著差异($p > 0.1$),这表明在双面相对绩效信息条件下,负面信息的引入显著提升了被试对“秒批”系统负面影响的感知,实验干预有效。

在实验 2 中,根据被试对下列表述的同意程度进行操纵检验:上述材料既展示了“秒批”服务相较于人工服务的优点,也展示了“秒批”服务相较于人工服务的缺点,问题选项采用 7 点李克特量表。方差分析结果显示,双面相对绩效信息组均值显著高于无绩效信息组和单面相对绩效信息组($p < 0.001$),而无绩效信息组与单面相对绩效信息组之间不存在显著差异($p > 0.1$),这表明在双面相对绩效信息条件下,被试能更充分地认识到“秒批”服务相较于人工服务的优势与不足,实验干预有效。

(三) 假设检验

1. 实验 1

实验 1 的实证分析结果如表 5 所示。就智能政务服务相对绩效信息对公众算法信任的影响而言,由模型 1 和模型 2 可知,无论是否纳入控制变量,与无绩效信息相比,仅呈现“秒批”系统效率优势的单面相对绩效信息均无法显著影响公众的算法信任($\beta = -0.018, p > 0.1$, 模型 1; $\beta = 0.009, p > 0.1$, 模型 2),而同时呈现“秒批”系统优势与不足的双面相对绩效信息则显著降低了公众的算法信任($\beta = -0.475, p < 0.001$, 模型 1; $\beta = -0.452, p < 0.001$, 模型 2),研究假设 H1-1 和 H2-1 均得到了验证与支持。

表 5 相对绩效信息、AI 乐观倾向与公众信任(实验 1)

	算法信任			政府信任		
	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
相对绩效信息(无绩效信息为参照组):						
单面相对绩效信息	-0.018 (-0.327)	0.009 (0.182)	-0.866* (-2.474)	0.0004 (0.008)	0.022 (0.428)	-0.879* (-2.584)
双面相对绩效信息	-0.475*** (-8.941)	-0.452*** (-8.778)	-0.543* (-1.900)	-0.335*** (-5.797)	-0.306*** (-5.203)	-0.428 (-1.285)
单面相对绩效信息× AI 乐观倾向			0.589** (2.614)			0.540** (2.772)
双面相对绩效信息× AI 乐观倾向			0.055 (0.281)			0.068 (0.336)
性别(男性=1)		0.103* (2.315)	0.099* (2.207)		0.073 (1.480)	0.069 (1.394)
年龄		0.205*** (4.158)	0.222*** (4.471)		0.063 (1.150)	0.081 (1.499)
教育程度		-0.090+ (-1.741)	-0.098+ (-1.880)		-0.082 (-1.430)	-0.090 (-1.547)
自评社会阶层		0.005 (0.101)	-0.001 (-0.028)		0.140** (2.879)	0.135** (2.684)
AI 使用经验		0.179*** (3.382)	0.170** (3.151)		0.099+ (1.659)	0.089 (1.455)
AI 乐观倾向		0.176*** (3.445)	0.074 (1.156)		0.145** (2.760)	0.036 (0.490)

续表

	算法信任			政府信任		
	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
<i>F</i>	44.80 ***	23.42 ***	19.04 ***	18.71 ***	10.22 ***	9.08 ***
<i>N</i>	337	337	337	337	337	337
<i>R</i> ²	0.218	0.339	0.356	0.112	0.189	0.207

注：表中展示的是各变量的标准化回归系数，括号内为 *t* 值；⁺ $p < 0.1$ ，^{*} $p < 0.05$ ，^{**} $p < 0.01$ ，^{***} $p < 0.001$ 。

就智能政务服务相对绩效信息对公众政府信任的影响而言，由模型 4 和模型 5 可知，无论是否纳入控制变量，与无绩效信息相比，单面相对绩效信息对政府信任的影响均不显著($\beta = 0.0004, p > 0.1$ ，模型 4； $\beta = 0.022, p > 0.1$ ，模型 5)，而双面相对绩效信息则显著降低了公众的政府信任($\beta = -0.335, p < 0.001$ ，模型 4； $\beta = -0.306, p < 0.001$ ，模型 5)，研究假设 H1-2 和 H2-2 均得到了验证与支持。

在主效应分析基础上，本文进一步分析 AI 乐观倾向的调节效应，结果如表 5 模型 3 和模型 6 所示。由模型 3 和模型 6 可知，单面相对绩效信息自身对算法信任和政府信任的影响均在 5% 水平上显著为负($\beta = -0.866, p < 0.05$ ，模型 3； $\beta = -0.879, p < 0.05$ ，模型 6)，而 AI 乐观倾向与单面相对绩效信息的交互项均在 1% 水平上显著为正($\beta = 0.589, p < 0.01$ ，模型 3； $\beta = 0.540, p < 0.01$ ，模型 6)，这表明 AI 乐观倾向会削弱单面相对绩效信息对算法信任和政府信任的负向影响，公众越积极乐观地看待人工智能，单面相对绩效信息越可能提升而非降低政府信任，研究假设 H4-1 和 H4-2 得到了验证与支持。与之不同，在模型 3 和模型 6 中，AI 乐观倾向与双面相对绩效信息的交互项均不显著($\beta = 0.055, p > 0.1$ ，模型 3； $\beta = 0.068, p > 0.1$ ，模型 6)，研究假设 H5-1 和 H5-2 未通过验证。

为直观呈现 AI 乐观倾向对单面相对绩效信息与公众信任关系的调节效应，本文采用 Hainmueller et al. (2019) 开发的更为灵活的调节效应估计策略，绘制了调节效应图(图 2)。由图 2 可知，当 AI 乐观倾向取值在 4 及以下时，单面相对绩效信息对算法信任和政府信任的影响为负但并不显著；当 AI 乐观倾向取值为 5 时，单面相对绩效信息对算法信任和政府信任存在显著的正向影响，表明 AI 乐观倾向能够促进单面相对绩效信息的正向作用。

2. 实验 2

实验 2 的实证分析结果如表 6、表 7 所示。表 6 展示了智能政务服务相对绩效信息对公众算法信任和政府信任的影响，以及主观绩效认知在其中的中介效应。由表 6 模型 1 和模型 2 可知，与实验 1 结果一致，单面相对绩效信息对算法信任和政府信任的影响仍不显著($\beta = 0.041, p > 0.1$ ，模型 1； $\beta = 0.076, p > 0.1$ ，模型 2)，再次验证并支持了研究假设 H1-1 和 H1-2。与实验 1 结果不同，双面相对绩效信息对算法信任的影响虽仍为负，但仅在 10% 的水平上显著($\beta = -0.084, p < 0.1$ ，模型 1)，研究假设

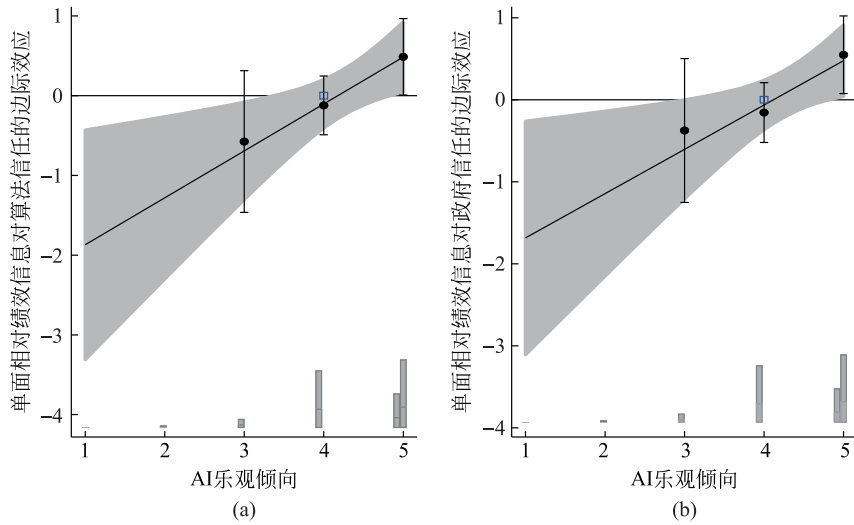


图2 AI乐观倾向对单面相对绩效信息与公众信任关系的调节效应(实验1)

注:图中的边际效应未经标准化处理;竖线和灰色区域代表95%水平上的置信区间;直方图为调节变量的取值分布。

H2-1 仅得到微弱支持;双面相对绩效信息对政府信任的影响甚至由负转正,且在5%水平上显著($\beta = 0.100, p < 0.05$, 模型2),未支持研究假设 H2-2。上述结果差异可能与实验2对负面绩效信息形式与表达风格的调整有关,也表明这一调整确有必要。

表6 相对绩效信息、主观绩效认知与公众信任(实验2)

	模型1 算法信任	模型2 政府信任	模型3 主观绩效认知	模型4 算法信任	模型5 政府信任
相对绩效信息(无绩效信息为参照组):					
单面相对绩效信息	0.041 (0.813)	0.076 (1.412)	0.082 (1.410)	0.011 (0.235)	0.045 (0.886)
双面相对绩效信息	-0.084 ⁺ (-1.687)	0.100 [*] (1.983)	0.507 ^{***} (10.453)	-0.273 ^{***} (-5.236)	-0.093 ⁺ (-1.756)
主观绩效认知				0.372 ^{***} (6.844)	0.381 ^{***} (6.589)
性别(男性=1)	0.075 ⁺ (1.681)	-0.021 (-0.449)	0.050 (1.123)	0.056 (1.346)	-0.040 (-0.925)
年龄	0.111 [*] (2.115)	0.032 (0.650)	0.092 ⁺ (1.965)	0.077 (1.530)	-0.003 (-0.055)

续表					
	模型 1 算法信任	模型 2 政府信任	模型 3 主观绩效认知	模型 4 算法信任	模型 5 政府信任
教育程度	-0.038 (-0.705)	-0.083 ⁺ (-1.725)	-0.102 [*] (-2.263)	-0.001 (-0.012)	-0.044 (-0.912)
自评社会阶层	0.113 [*] (2.208)	0.107 ⁺ (1.877)	0.153 ^{**} (3.092)	0.056 (1.140)	0.049 (0.911)
AI 使用经验	-0.034 (-0.658)	-0.054 (-0.970)	0.076 ⁺ (1.651)	-0.062 (-1.226)	-0.083 (-1.547)
AI 乐观倾向	0.138 ^{**} (2.699)	0.108 [*] (2.104)	0.030 (0.605)	0.127 ^{**} (2.631)	0.096 [*] (2.026)
对政府使用 AI 的乐观倾向	0.378 ^{***} (6.964)	0.413 ^{***} (6.517)	0.070 (1.415)	0.352 ^{***} (6.390)	0.387 ^{***} (6.544)
<i>F</i>	16.63 ^{***}	10.61 ^{***}	30.43 ^{***}	27.17 ^{***}	16.62 ^{***}
<i>N</i>	378	378	378	378	378
<i>R</i> ²	0.283	0.259	0.294	0.380	0.362

注：表中展示的是各变量的标准化回归系数，括号内为 *t* 值；⁺ $p<0.1$ ，^{*} $p<0.05$ ，^{**} $p<0.01$ ，^{***} $p<0.001$ 。

表 7 AI 乐观倾向的调节效应(实验 2)

	算法信任		政府信任	
	模型 1	模型 2	模型 3	模型 4
相对绩效信息(无绩效信息为参照组)：				
单面相对绩效信息	-0.119 (-0.267)	0.532 (0.977)	0.578 (1.296)	1.135 ⁺ (1.904)
双面相对绩效信息	-0.764 ⁺ (-1.917)	0.515 (1.219)	-0.017 (-0.037)	0.674 (1.104)
单面相对绩效信息×AI 乐观倾向	0.084 (0.367)		-0.239 (-1.146)	
双面相对绩效信息×AI 乐观倾向	0.353 ⁺ (1.716)		0.054 (0.262)	
单面相对绩效信息×对政府使用 AI 的乐观倾向		-0.260 (-0.921)		-0.509 ⁺ (-1.789)
双面相对绩效信息×对政府使用 AI 的乐观倾向		-0.317 (-1.421)		-0.276 (-0.956)

续表

控制变量	算法信任		政府信任	
	模型 1	模型 2	模型 3	模型 4
	是	是	是	是
F	15.18 ***	15.13 ***	10.16 ***	10.05 ***
N	378	378	378	378
R^2	0.289	0.287	0.265	0.270

注:表中展示的是各变量的标准化回归系数,括号内为 t 值; * $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ 。

上述分析未考虑主观绩效认知的作用,因此本文进一步分析主观绩效认知在智能政务服务相对绩效信息与公众信任之间的中介效应。由表 6 模型 3 可知,单面相对绩效信息对主观绩效认知的影响不显著($\beta = 0.082, p > 0.1$, 模型 3),而双面相对绩效信息则显著提升了主观绩效认知($\beta = 0.507, p < 0.001$, 模型 3)。由表 6 模型 4 和模型 5 可知,主观绩效认知对算法信任和政府信任均存在显著的正向影响($\beta = 0.372, p < 0.001$, 模型 4; $\beta = 0.381, p < 0.001$, 模型 5)。综上,双面相对绩效信息能够通过提升公众的主观绩效认知,间接提升算法信任和政府信任,研究假设 H3-1 和 H3-2 得到验证与支持。

同时,与表 6 模型 1 相比,模型 4 在将主观绩效认知纳入后,双面相对绩效信息对算法信任的负向影响增强,影响系数从 -0.084 变为 -0.273 ,显著性水平从 10% 变为 0.1% ($\beta = -0.273, p < 0.001$, 模型 4),表明双面相对绩效信息对算法信任的直接效应为负,与研究假设 H2-1 吻合。类似地,与表 6 模型 2 相比,模型 5 在将主观绩效认知纳入后,双面相对绩效信息对政府信任的影响系数由 5% 水平上显著为正转为 10% 水平上显著为负 ($\beta = -0.093, p < 0.1$, 模型 5),表明双面相对绩效信息对政府信任的直接效应为负,与研究假设 H2-2 吻合。^①

最后,本文检验人工智能认知倾向对智能政务服务相对绩效信息与公众信任关系的调节效应,结果如表 7 所示。由表 7 模型 1 可知,双面相对绩效信息对算法信任的影响显著为负 ($\beta = -0.764, p < 0.1$, 模型 1),而 AI 乐观倾向与双面相对绩效信息的交互项则在 10% 的水平上显著为正 ($\beta = 0.353, p < 0.1$, 模型 1),这表明 AI 乐观倾向削弱了双面相对绩效信息对算法信任的负向影响,研究假设 H5-1 得到验证与支持。由表 7 模型 4 可知,单面相对绩效信息对政府信任的影响显著为正 ($\beta = 1.135, p < 0.1$, 模型 4),而对政府使用 AI 的乐观倾向与单面相对绩效信息的交互项在 10% 的水平上显著为负 ($\beta = -0.509, p < 0.1$, 模型 4),这表明对政府使用 AI 的乐观倾向削弱了单面相对绩效信息对政府信任的正向影响,研究假设 H4-2 未得到验证,与预期相

① 由上述分析可知,双面相对绩效信息对公众信任的直接效应为负,通过主观绩效认知对公众信任产生的间接效应为正,因此主观绩效认知实际在双面相对绩效信息与公众信任之间发挥遮掩效应而非中介效应。但因为在统计分析意义上的区别不大,为避免表述混乱,此处仍使用中介效应的表述。

反。对此,一个可能的解释是,公众对政府使用 AI 的乐观倾向也反映出公众对智能政务服务的高期望,从期望失验理论的角度来看,在相同绩效水平下,高期望者更可能感到不满意,进而削弱了单面相对绩效信息的正向影响(van der Voet and van den Bekerom,2025)。

图 3 直观呈现了 AI 乐观倾向的调节效应。由图 3(a)可知,随着 AI 乐观倾向取值的提升,双面相对绩效信息对算法信任的负向影响逐渐减弱,并由显著变为不显著。由图 3(b)可知,随着对政府使用 AI 的乐观倾向取值的提升,单面相对绩效信息对政府信任的正向影响逐渐减弱,并由显著变为不显著。

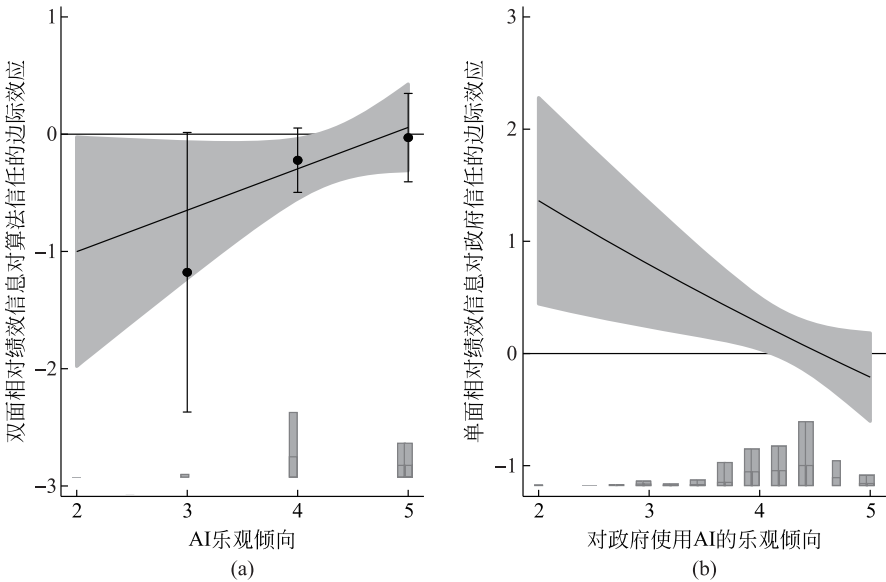


图 3 AI 乐观倾向对相对绩效信息与公众信任关系的调节效应(实验 2)

注:图中的边际效应未经标准化处理;竖线及灰色区域代表 95% 水平上的置信区间;直方图为调节变量的取值分布。

六、结论与讨论

政务服务智能化转型深刻影响着政民关系和公众对政府的信任。与传统人工服务相比,智能政务服务既有优势,又有不足。如何科学呈现智能政务服务相较于人工服务的相对绩效信息,可能会影响公众信任。鉴于此,本文遵循行为公共管理的研究范式,基于个体工商户“秒批”改革的现实案例,采用调查实验法开展两项独立子实验,以人工服务为参照点,实证检验智能政务服务的相对绩效信息及其信息面向对公众信任(包括算法信任和政府信任)的影响,以及理性认知和认知偏差在其中的共同作用。

研究得出以下几点发现。第一,相对绩效信息的信息面向对公众算法信任和政府信任存在显著影响。呈现智能政务服务相对优势的单面相对绩效信息容易被公众解读为政府的“选择性宣传”,因而说服力有限,难以有效提升公众对智能算法和政府的信任;同时呈现其优势与不足的双面相对绩效信息因具备更高的信息可信度与完整性,能显著影响算法信任和政府信任。第二,理性认知和认知偏差共同塑造双面相对绩效信息对公众信任的影响。一方面,受潜在负面偏见的影响,兼具正负维度的双面相对绩效信息对公众信任的直接效应为负。另一方面,在理性认知作用下,双面相对绩效信息能使公众形成对智能政务服务绩效更全面的认识,进而间接提升公众信任。第三,受确认偏差与反确认偏差的影响,公众对人工智能的乐观倾向增强了单面相对绩效信息对算法信任和政府信任的正向影响,削弱了双面相对绩效信息的负面影响。此外,公众对政府使用人工智能的乐观倾向削弱了单面相对绩效信息对政府信任的正向影响,这可能是因为高乐观倾向引发了高绩效期望,进而导致期望失验。

上述发现的理论贡献主要体现在以下两个方面。一方面,不同于既有研究更关注绩效信息的事实差异(正面绩效还是负面绩效),本文基于智能政务服务场景,从信息面向切入,揭示出绩效信息呈现策略的差异(呈现正面还是正负都呈现)对公众信任的重要影响,从而丰富了智能政务服务场景中的绩效信息研究。另一方面,不同于既有研究往往仅强调理性认知或认知偏差的单一作用,本文整合委托-代理理论与认知心理学理论,揭示出两者在公众绩效信息处理中的共同作用,深化了对智能政务服务中信任形成机制的理论解释。

本研究的发现对于进一步推进智能政务服务发展、提升公众信任具有一定实践启示意义。其一,在推广智能政务服务过程中,政府应摒弃“报喜不报忧”的简单化叙事,采取分阶段的信息披露策略。推广初期可重点宣传智能服务的技术优势,待公众初步接受后再逐步披露潜在局限性,深化公众理性认知,既防止片面信息引发公众抵触,也避免信息过载引发公众恐慌。其二,政府应根据不同群体的认知特征,采取差异化信息沟通策略。对于人工智能的高乐观群体,政府可侧重宣传智能政务服务的技术优势;而对低乐观群体,则应强调智能技术保障和迭代优化趋势,并通过科普和培训活动提升公众对人工智能的信心,缓解公众的算法厌恶心理。

本研究也存在不足和有待拓展之处。第一,本研究基于单一实验场景,这在一定程度上限制了研究的外部效度,未来研究可基于不同的政策领域、算法类型及任务特征等,进一步检验本研究结论的稳健性。例如,本研究中的算法属于全自动算法,未来研究可分析人机混合的半自动算法的绩效信息对公众信任的影响。第二,与其他通过在线调查平台收集数据的研究类似,本研究样本的代表性存在一定不足,男性、高学历群体及年轻人占比偏高,未来可尝试采用入户访问等方式收集数据,提升样本代表性。第三,智能政务服务相对绩效信息的操纵有待完善。本研究分别从办事时长和准确率两方面操纵“秒批”系统正面和负面的绩效信息,但由于两

者在公众心中的重要性可能不同,两者对公众信任的影响差异既可能源于正负性的不同,也可能源于重要性的差异。因此,未来可采用其他绩效属性进行实验设计和比较分析,以验证本文研究结论的可推广性。

参考文献

- 樊博,李晶晶. 2024. 算法透明何以提升公众的政府信任? [J]. 公共行政评论, 17(1): 4-24.
- Fan B, Li J J. 2024. Can algorithmic transparency enhance citizen trust in government? [J]. *Journal of Public Administration*, 17(1): 4-24. (in Chinese)
- 高学德,王镇江. 2024. 危机沟通策略对地方政府信任的影响研究[J]. 公共管理评论, 6(3): 76-98.
- Gao X D, Wang Z J. 2024. How does a government crisis communication strategy affect public trust in local government? Evidence based on a survey experiment [J]. *China Public Administration Review*, 6(3): 76-98. (in Chinese)
- 李晓方,王友奎,孟庆国. 2020. 政务服务智能化:典型场景、价值质询和治理回应[J]. 电子政务, (2): 2-10.
- Li X F, Wang Y K, Meng Q G. 2020. Intelligent government services: Typical scenarios, value inquiry, and governance responses[J]. *E-Government*, (2): 2-10. (in Chinese)
- 李晓倩,戴乐融,刘小雨. 2022. 实验方法在公共管理研究中应用的最新进展——基于国际期刊发表(2017—2021)的系统性文献综述[J]. 公共管理与政策评论, 11(4): 151-168.
- Li X Q, Dai L R, Liu X Y. 2022. The application of experimental methods in public administration research: A systematic literature review of international publications (2017—2021) [J]. *Public Administration and Policy Review*, 11(4): 151-168. (in Chinese)
- 刘特,郑跃平,曹梦冰. 2024a. 算法行政:文献述评与研究展望[J]. 公共行政评论, 17(1): 84-104.
- Liu T, Zheng Y P, Cao M B. 2024a. Algorithmic administration: Literature review and research prospects[J]. *Journal of Public Administration*, 17(1): 84-104. (in Chinese)
- 刘特,郑跃平,杨学敏. 2024b. 公共部门中的算法透明能否影响公众信任?——基于一项调查实验[J]. 电子政务, (8): 13-28.
- Liu T, Zheng Y P, Yang X M. 2024b. Can algorithmic transparency in the public sector affect public trust? Evidence from a survey experiment[J]. *E-Government*, (8): 13-28. (in Chinese)
- 刘玮,王锋. 2024. 政务服务智能化创新的演化、风险与图景——基于场域视角的分析[J]. 电子政务, (2): 79-88.
- Liu W, Wang F. 2024. Evolution, risks, and prospects of intelligent government service

- innovation: A field-based analysis[J]. *E-Government*, (2): 79-88. (in Chinese)
- 冉龙亚, 陈涛, 孙宁华. 2022. 人工智能参与公共决策对公众信任的影响——一项基于调查实验的实证研究[J]. *科学与社会*, 12(1): 103-121.
- Ran L Y, Chen T, Sun N H. 2022. The influence of artificial intelligence algorithm participating in public decision making on public trust—A survey experiment study[J]. *Science and Society*, 12(1): 103-121. (in Chinese)
- 孙源, 米加宁, 刘鲁宁. 2024. 志愿服务中算法透明度对感知可信度的影响——服务动机的调节效应[J]. *公共行政评论*, 17(1): 66-83.
- Sun Y, Mi J N, Liu L N. 2024. The impact of algorithm transparency on perceived trustworthiness in voluntary service: The moderating effect of service motivation[J]. *Journal of Public Administration*, 17(1): 66-83. (in Chinese)
- 汤志伟, 方録, 韩啸, 等. 2023. 基于委托代理视角的算法治理问题研究[J]. *情报杂志*, 42(7): 170-177.
- Tang Z W, Fang L, Han X, et al. 2023. Algorithmic governance issues: From a principal-agent perspective[J]. *Journal of Intelligence*, 42(7): 170-177. (in Chinese)
- 王思琦. 2023. 绩效信息理解难度与参考服务质量对满意度的影响[J]. *公共管理评论*, 5(4): 25-51.
- Wang S Q. 2023. The effects of difficulties in understanding performance information and the quality of reference services on satisfaction[J]. *China Public Administration Review*, 5(4): 25-51. (in Chinese)
- 王正绪, 赵健池. 2021. 民本贤能政体与大众政治心理: 以政治信任为例[J]. *开放时代*, (4): 139-156.
- Wang Z X, Zhao J C. 2021. People-based meritocracy and mass political psychology: Taking political trust as an example[J]. *Open Times*, (4): 139-156. (in Chinese)
- 吴进进, 何包钢. 2023. 算法科层制的兴起及其形态[J]. *社会学研究*, 38(6): 40-60.
- Wu J J, He B G. 2023. The rise of algorithmic bureaucracy and its typology[J]. *Sociological Studies*, 38(6): 40-60. (in Chinese)
- 游宇, 王正绪. 2014. 互动与修正的政治信任——关于当代中国政治信任来源的中观理论[J]. *经济社会体制比较*, (2): 178-193.
- You Y, Wang Z X. 2014. Interactive and revisionist political trust: A meso-theory on the sources of political trust in contemporary China[J]. *Comparative Economic & Social Systems*, (2): 178-193. (in Chinese)
- 张成福, 王祥州. 2023. 人工智能嵌入公共服务治理的风险挑战[J]. *电子政务*, (1): 37-51.
- Zhang C F, Wang X Z. 2023. Risks and challenges of embedding artificial intelligence in public service governance[J]. *E-Government*, (1): 37-51. (in Chinese)
- 周豪, 包国宪. 2021. 信息来源、信息一致性与公众对政府绩效信息的信任——基于调

- 查实验的发现[J]. 公共管理评论, 3(3): 5-32.
- Zhou H, Bao G X. 2021. Source, consistency and the credibility of government performance information: A survey experiment[J]. *China Public Administration Review*, 3(3): 5-32. (in Chinese)
- 朱春奎, 毛万磊. 2017. 政府信任的概念测量、影响因素与提升策略[J]. 厦门大学学报(哲学社会科学版), (3): 89-98.
- Zhu C K, Mao W L. 2017. Conceptual measurement, influencing factors and promotion strategies of government trust[J]. *Journal of Xiamen University (Arts & Social Sciences)*, (3): 89-98. (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[J]. *Government Information Quarterly*, 39(2): 101664.
- Alessandro M, Lagomarsino B C, Scartascini C, et al. 2021. Transparency and trust in government: Evidence from a survey experiment[J]. *World Development*, 138: 105223.
- Alon-Barkat S, Busuioc M. 2023. Human-AI interactions in public sector decision making: “Automation bias” and “selective adherence” to algorithmic advice[J]. *Journal of Public Administration Research and Theory*, 33(1): 153-169.
- Andersen S C, Hjortskov M. 2016. Cognitive biases in performance evaluations[J]. *Journal of Public Administration Research and Theory*, 26(4): 647-662.
- Baekgaard M, James O, Serritzlew S, et al. 2020. Citizens’ motivated reasoning about public performance: Experimental findings from the US and Denmark[J]. *International Public Management Journal*, 23(2): 186-204.
- Bauhr M, Grimes M. 2014. Indignation or resignation: The implications of transparency for societal accountability[J]. *Governance*, 27(2): 291-320.
- Brogaard L, Madsen J K, Petersen O H. 2024. Interpreting performance information: Motivated reasoning or unbiased comprehension? A replication and extension[J]. *Public Administration*, 102(1): 318-340.
- Chandra Y, Feng N K. 2026. Algorithms for a new season? Mapping a decade of research on the artificial intelligence-driven digital transformation of public administration[J]. *Public Management Review*, 28(3): 620-654.
- Eisend M. 2007. Understanding two-sided persuasion: An empirical assessment of theoretical approaches[J]. *Psychology & Marketing*, 24(7): 615-640.
- Gesk T S, Leyer M. 2022. Artificial intelligence in public services: When and why citizens accept its usage[J]. *Government Information Quarterly*, 39(3): 101704.
- Grimmelikhuijsen S. 2012. Linking transparency, knowledge and citizen trust in government: An experiment[J]. *International Review of Administrative Sciences*, 78(1): 50-73.
- Grimmelikhuijsen S. 2023. Explaining why the computer says no: Algorithmic transparency

- affects the perceived trustworthiness of automated decision-making [J]. *Public Administration Review*, 83(2) : 241-262.
- Haesevoets T, Verschuere B, Van Severen R, et al. 2024. How do citizens perceive the use of Artificial Intelligence in public sector decisions? [J]. *Government Information Quarterly*, 41(1) : 101906.
- Hainmueller J, Mummolo J, Xu Y Q. 2019. How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice[J]. *Political Analysis*, 27(2) : 163-192.
- Han Y. 2025. Message-sidedness in performance information disclosure and citizens' perceived accountability: An experimental study [J]. *Public Management Review*, 27 (9) : 2099-2125.
- Hemesath S, Tepe M. 2024. Public value positions and design preferences toward AI-based chatbots in e-government. Evidence from a conjoint experiment with citizens and municipal front desk officers[J]. *Government Information Quarterly*, 41(4) : 101985.
- Hetherington M J. 2004. Why trust matters: Declining political trust and the demise of American liberalism[M]. Princeton, NJ: Princeton University Press.
- Horvath L, James O, Banducci S, et al. 2023. Citizens' acceptance of artificial intelligence in public services: Evidence from a conjoint experiment about processing permit applications [J]. *Government Information Quarterly*, 40(4) : 101876.
- Ingrams A, Kaufmann W, Jacobs D. 2022. In AI we trust? Citizen perceptions of AI in government decision making[J]. *Policy & Internet*, 14(2) : 390-409.
- James O. 2011. Performance measures and democracy: Information effects on citizens in field and laboratory experiments[J]. *Journal of Public Administration Research and Theory*, 21(3) : 399-418.
- James O, Moseley A. 2014. Does performance information about public services affect citizens' perceptions, satisfaction, and voice behaviour? Field experiments with absolute and relative performance information[J]. *Public Administration*, 92(2) : 493-511.
- Jones E E, Davis K E. 1965. From acts to dispositions the attribution process in person perception[M]//Berkowitz L. *Advances in Experimental Social Psychology*. New York: Academic Press: 219-266.
- Kahneman D, Tversky A. 1979. Prospect theory: An analysis of decision under risk [J]. *Econometrica*, 47(2) : 263-292.
- Kennedy R P, Waggoner P D, Ward M M. 2022. Trust in public policy algorithms[J]. *The Journal of Politics*, 84(2) : 1132-1148.
- König P D, Felfeli J, Achziger A, et al. 2024. The importance of effectiveness versus transparency and stakeholder involvement in citizens' perception of public sector algorithms [J]. *Public Management Review*, 26(4) : 1061-1082.

- Kreps S, Jakesch M. 2023. Can AI communication tools increase legislative responsiveness and trust in democratic institutions? [J]. *Government Information Quarterly*, 40(3): 101829.
- Kunda Z. 1990. The case for motivated reasoning [J]. *Psychological Bulletin*, 108(3): 480-498.
- Meijer A, Lorenz L, Wessels M. 2021. Algorithmization of bureaucratic organizations: Using a practice lens to study how context shapes predictive policing systems [J]. *Public Administration Review*, 81(5): 837-846.
- Meng X, Li C P. 2025. Does public sector performance information impact stakeholders? Evidence from a meta-analysis [J]. *Public Administration Review*, 85(4): 1086-1097.
- Miller S M, Song M, Keiser L R. 2023. The effect of human versus automated interaction on willingness to participate in government programs: The role of representation [J]. *Public Administration*, 101(4): 1349-1366.
- Olsen A L. 2017. Compared to what? How social and historical reference points affect citizens' performance evaluations [J]. *Journal of Public Administration Research and Theory*, 27(4): 562-580.
- Porumbescu G, Moynihan D, Anastasopoulos J, et al. 2023. When blame avoidance backfires: Responses to performance framing and outgroup scapegoating during the COVID-19 pandemic [J]. *Governance*, 36(3): 779-803.
- Rozin P, Royzman E B. 2001. Negativity bias, negativity dominance, and contagion [J]. *Personality and Social Psychology Review*, 5(4): 296-320.
- Schiff D S, Schiff K J, Pierson P. 2022. Assessing public value failure in government adoption of artificial intelligence [J]. *Public Administration*, 100(3): 653-673.
- Schmidhuber L, Willems J, Krabina B. 2023. Trust in public performance information: The effect of data accessibility and data source [J]. *Public Administration Review*, 83(2): 279-295.
- Selten F, Robeer M, Grimmelikhuijsen S. 2023. "Just like I thought": Street-level bureaucrats trust AI recommendations if they confirm their professional judgment [J]. *Public Administration Review*, 83(2): 263-278.
- van der Voet J, van Den Bekerom P. 2025. Performance information, expectations and satisfaction with public service delivery in the context of co-production initiatives [J]. *Public Management Review*, 27(4): 1157-1178.
- Wang C L, Teo T S H, Janssen M. 2021. Public and private value creation using artificial intelligence: An empirical study of AI voice robot users in Chinese public sector [J]. *International Journal of Information Management*, 61: 102401.
- Yang K F, Holzer M. 2006. The performance-trust link: Implications for performance measurement [J]. *Public Administration Review*, 66(1): 114-126.

The Impact of Relative Performance Information on Public Trust in AI-enabled Public Services

DU Yabin

(School of Public Administration, Hunan University)

Abstract: The proliferation of AI-enabled public services has reshaped administrative processes and citizen-state interactions, yet public trust in such services remains ambivalent. This study examines how message-sidedness—i. e. , whether performance information about AI-enabled public services presents only advantages (one-sided messages) or both advantages and disadvantages (two-sided messages)—influences public trust in the services and government agencies. Drawing on the real-world case of the “instant approval” reform for individual business registration in China, we conducted two survey experiments using traditional manual service as the benchmark.

Our empirical findings are threefold. First, message-sidedness significantly affects the persuasive power of relative performance information. One-sided messages, which only highlight efficiency gains of AI services, fail to significantly alter trust in either the services or the government. In contrast, two-sided messages—reporting both shorter processing times and decreased accuracy—significantly reduce trust, suggesting a potential negative bias effect.

Second, we identify the dual role of rational cognition and cognitive biases in citizens’ performance information processing. Although two-sided messages directly reduce trust, they also improve perceived performance knowledge, which indirectly enhances trust. This mediation effect indicates that rational cognition may partly offset the detrimental impact of negative disclosures. Thus, public trust in AI-enabled services is jointly shaped by reasoned judgment and affective heuristics.

Third, we examine confirmation bias via citizens’ pre-existing attitudes toward artificial intelligence. Those with higher optimism toward AI are more receptive to positive one-sided messages and less reactive to negative content in two-sided messages. This confirms that prior beliefs systematically moderate the impact of performance communication, underscoring the relevance of tailored messaging. Interestingly, when AI optimism is operationalized specifically as support for government adoption of AI, it paradoxically weakens the trust-enhancing effect of one-sided messages—likely due to heightened expectations and greater disappointment in the face of incomplete information.

These results contribute to the growing literature on behavioral public administration

by integrating performance management, trust theory, and information processing perspectives. We extend prior research by unpacking how the format and framing of relative performance information, rather than its valence alone, condition citizens' evaluations of AI-enabled services.

Practically, the findings imply that governments should abandon one-sided, overly promotional communication strategies when introducing AI reforms. Instead, phased and audience-sensitive messaging—highlighting strengths while transparently acknowledging risks—can both foster trust and prevent future backlash. Moreover, enhancing digital literacy and addressing variation in AI attitudes across subgroups will be essential to securing public legitimacy in the era of algorithmic governance.

Keywords: AI-enabled public services; relative performance information; message-sidedness; cognitive bias; trust in government

投稿日期: 2025/3/31 送外审日期: 2025/4/9 首轮外审完成日期: 2025/5/3
录用日期: 2025/9/8 最终修回日期: 2026/2/14