

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

孙 源 米加宁 刘鲁宁*

【摘要】从党的十九大提出“推进志愿服务制度化”，到二十大强调“完善志愿服务制度”，志愿服务已经成为我国社会服务体系的重要构成。对志愿服务的重视和支持既是人民自治的体现，也是社会文明进步的标志。尽管经过三十余年的发展，我国在志愿者数量、服务领域等方面取得了显著成就，但公众参与率低下这一问题仍然是制约我国志愿服务可持续发展的主要因素。针对这一问题，学术界尝试应用算法推荐促进公众对志愿服务的了解和参与。然而，算法的使用过程中存在着广泛的用户信任问题，在志愿服务领域这一问题又格外突出。因此，论文聚焦于志愿服务活动的算法推荐，采用调查实验的方法，分析志愿服务动机如何影响算法透明度对于算法信任的作用。实验结果表明，（1）相对于不透明的算法，志愿者对透明算法的信任度更高；（2）志愿服务动机调节了算法透明度与算法可信度之间的关系，志愿者对与其志愿服务动机一致的算法信任度更高；（3）志愿服务动机与算法决策逻辑一致对算法透明的积极作用的加强在利己动机志愿者中更大。论文与算法信任及志愿服务领域的文献进行对话，通过评估志愿服务动机在算法透明度对算法信任的影响中的调节作用，扩展了已有研究在算法治理领域的发现，同时为相关管理人员在如何提高志愿者对于算法建议信任的方面提供了实践建议。

【关键词】志愿服务 志愿服务动机 算法信任 感知一致性

【中图分类号】 D63

【文献标识码】 A

【文章编号】 1674 - 2486 (2024) 01 - 0066 - 18

一、问题提出

作为社会服务体系不可或缺的供给构成，志愿服务填补了“市场失灵”和

* 孙源，哈尔滨工业大学经济与管理学院博士研究生；米加宁，哈尔滨工业大学经济与管理学院教授；通讯作者：刘鲁宁，哈尔滨工业大学经济与管理学院教授。感谢匿名评审专家和编辑部的修改意见。

基金项目：国家社科基金重大项目“大数据赋能共建共治共享的社会治理制度建设研究”（22ZDA100）。

“政府失灵”的空白，成为推动社会经济发展的重要力量。志愿服务不限志愿者年龄、职业、财富等基本特征，为每一位公民提供了参与社会服务的平等机会。这一举措促进了人民自治，是打造共建共治共享治理格局的核心要素（党秀云，2019；张瑞，2021）。同时，作为一项长期服务，志愿服务的持续、健康发展是满足社会需求、维护公共价值的基本前提（苗青、尹晖，2021）。从实践效果来看，虽然我国的志愿服务力量已然在脱贫攻坚、抗击疫情等活动中发挥了重要作用，但公众参与率低下仍是制约其可持续发展的关键问题（辛华等，2016）。民政部志愿服务官方网站“中华志愿服务网”显示，2022年我国实名注册志愿者数已达2.3亿人，但记录志愿服务时长的人数仅有8029万人，也就是说，志愿者群体的实际志愿服务参与率不足35%。

公众参与率低并不是由于缺乏志愿者，而是由于缺少科学的志愿服务全流程管理方法（苗青、尹晖，2021）。对于如何实现志愿服务的科学化管理目标，已有研究大多聚焦于制度层面。例如，针对志愿者的激励问题，学者们提出了精神激励、荣誉激励、物质激励等多种激励手段，并运用实验方法寻找最佳激励组合（苗青、尹晖，2021）。然而，相较于使志愿服务项目具备有足够的吸引力，更紧要的问题是令志愿者们找到合适的志愿项目。对于志愿者来说，搜索志愿服务组织和志愿服务项目具有很高的信息成本（Emrich & Pierdzioch，2016）。已有研究发现，“不知道去哪里找志愿服务活动”“不知道怎么参与”是影响公众参与率的主要原因（张网成等，2018）。信息成本的降低意味着志愿者可以投入更多的资源在真正的志愿服务工作中（Emrich & Pierdzioch，2016）。因此，如何通过大数据、云计算等新一代信息技术识别、匹配志愿服务的需求与供给，实现志愿服务的精准供给，使公众能够及时便捷地参与志愿服务已经受到了广泛的关注（党秀云，2019）。从实践层面来看，杭州、深圳、上海等地也已经推出了为志愿者精准匹配志愿服务活动的志愿服务APP。

随着算法推荐在志愿服务中的普遍应用，算法透明度和算法信任问题引起了广泛的关注。算法透明度是指用户能够理解算法如何工作及其决策过程的程度（Grimmelikhuijsen，2022）。通常而言，透明度高的算法能够增加用户的信任，进而提高他们的使用意愿。然而，目前存在的问题是算法透明度并不总是与用户信任成正比（Lehmann et al.，2022）。例如，当算法过于简单或者其决策逻辑不符合用户期望时，即便算法透明度高，也难以获得用户的信任。在志愿服务领域，动机的满足是提高志愿服务项目参与率、促进志愿服务可持续发展的元引力（辛华等，2016）。志愿者的服务动机可能是利己的（比如提升个人技能），也可能是利他的（比如帮助需要的人）（Stukas et al.，2016），这会影影响他们对于特定项目的选择和参与（辛华等，2016）。算法将志愿者与志愿服务项目进行匹配时，如果能够考虑到志愿者的不同动机，并通过提高算法透明度将匹配依据展现给志愿者，可能会提高志愿者的参与度和满意度。因此，在志

愿服务的背景下，当前研究需要提供对算法透明度与用户信任之间复杂关系的更细致的理解。具体来说，志愿服务推荐算法的设计和实现需要考虑志愿者的服务动机是如何对透明度与信任之间的关系产生影响的。

针对上述研究背景，本研究采用调查实验的方法，试图在志愿服务的特定情境下，回答已有的算法治理的相关研究尚未回答的问题。具体而言，本研究将评估算法透明度对算法可信度的影响，并从一致性的角度分析志愿服务动机的调节作用。通过对上述问题的回应，本研究可以揭示算法透明度、志愿服务动机、感知一致性、算法信任之间的关系，在丰富算法治理研究的同时，为公共部门在志愿服务领域正确使用算法提供参考。

二、文献综述与研究假设

（一）算法信任

算法通过学习行为和行政数据，了解公众的多元化需求，进而以精准推送的方式主动提供公共服务，显著降低了公众的遵从成本和学习成本，提升了服务体验（Keiser & Miller, 2020）。随着算法在政府治理，特别是在志愿服务领域的广泛应用，当前关于算法的研究呈现出两种路径，分别是算法的“技术赋能”正向作用及其“技术负外部性”负向作用。

“技术赋能”观点肯定了算法嵌入对志愿服务的赋能效用。在志愿服务情境下，寻找适合的志愿服务活动对于志愿者而言往往是一项高信息成本任务（Emrich & Pierdziuch, 2016；张网成等，2018），而算法的应用能够通过提高志愿活动的可及性和便利性，有效降低这一成本（李晓园等，2022）。

然而，算法的使用仍然存在广泛的争议。相对于前一种观点，后一种观点更加关注算法的不透明、歧视以及违反伦理等技术负外部性问题（昌诚等，2022；张雪帆、蒋忠楠，2022）。算法的“黑箱”性质使公众难以了解算法使用的数据、不同数据在决策中的权重以及决策的形成过程，从而动摇了算法决策的合法性、准确性和公正性，导致公众对于算法的信任度低，甚至产生算法厌恶。算法信任问题在志愿服务领域是一个关键议题，因为它直接影响到志愿者对于推荐结果的接受度。

（二）算法透明度

在当前对算法应用的研究中，提高算法透明度被普遍视为增强用户信任的关键机制。提高算法透明度包含两层含义，一方面是指提高算法的可及性，即允许技术或业务专家等外部参与者访问算法的底层数据和代码，以确保算法的客观合理性和可靠性。另一方面是指提高算法的可解释性，即将算法的决策结

果以公众可理解的方式呈现（Grimmelikhuijsen, 2022）。对公众而言，增强算法的可解释性是提高其透明度的有效路径，可以促进公众对算法决策过程的理解。所以本研究聚焦于可解释性，探究算法透明度与算法信任之间的关系。

然而，算法透明度并不总是对信任产生积极的影响。根据 Lehmann 等（2022）的研究，当公众发现算法的复杂性低于他们的预期时，透明度的增加反而会降低信任度。这种现象表明，算法透明度与算法信任之间的关系是复杂的，透明度的效果受到算法自身特性和公众预期的共同影响。因此，算法透明度在增强用户信任方面的作用需要根据具体情境进行细致评估。在志愿服务背景下，用户的目标是找到满足个人服务动机的活动（辛华等，2016），而政府相关部门开发算法的目标可能是帮助最需要帮助的人，考虑优先向弱势群体供给服务，也可能是满足志愿者动机，以增加用户黏性、保障志愿服务的可持续发展。因此，如何在匹配过程中考虑用户参与志愿服务的动机，进而提高算法信任，成为算法设计时的关键。

（三）志愿服务动机

志愿服务动机是一种来自于内在、存在于心理层面的驱动力，驱使志愿者奉献个人时间和精力参与志愿服务活动（Clary et al., 1996）。从个体层面而言，志愿服务动机的不足是导致志愿者参与度低下的关键因素（辛华等，2016）。已有研究从单因素、双因素到六维度框架对志愿服务动机进行了分类，但分类的底层逻辑都源于社会心理学中长期存在的、关于帮助行为是出于利他还是利己的争论（Stukas et al., 2016；张网成等，2018）。利己动机志愿者参加志愿活动的目的地包括积累工作经验、锻炼能力、保持积极正向的生活态度、自我实现、做有意义的事等等；利他动机志愿者主要是想通过志愿活动帮助他人、回报社会、改善社区环境等（张网成，2015）。志愿服务的功能性理论认为，当志愿者感知到自己的目标在活动中得以实现时，他们参与的可能性更高（Snyder & Omota, 2008；Clary et al., 1998）。

志愿者参与志愿服务的动机影响着其在志愿服务中的行为（辛华等，2016）。以他人为导向的志愿者通常服务时间更长，而以自身为导向的志愿者在无法实现个人利益的活动中难以维持长期的亲社会行为（Omoto & Snyder, 1995）。这些发现表明，志愿服务中算法推荐的应用需要考虑志愿者的动机，以提高公众参与率和促进志愿服务的可持续发展。换句话说，志愿服务动机是否得到满足会影响算法透明度对算法信任的作用。用户参与志愿服务的动机得到满足意味着用户志愿服务动机和算法推荐逻辑一致。算法决策逻辑与用户需求偏好间的一致性可能是影响算法可信度的重要因素，进而影响着志愿服务项目的参与率和志愿服务的可持续发展（辛华等，2016）。

（四）算法决策逻辑与志愿服务动机的一致性

一致性理论源于社会心理学中的相似性概念，最初被用于分析个体自我形象与外部实体间的匹配程度（Maille & Fleck, 2011）。研究发现，消费者更倾向于选择与其自我形象或价值观高度一致的品牌或产品，这种一致性促使消费者对相关品牌或产品形成更积极的态度，并进一步增强了他们的购买意愿（Paul & Bhakar, 2018）。算法推荐代理领域的一项研究也发现，用户对于性格或性别与自己匹配的推荐代理展现出更高的信任度（Reinkemeier & Gnewuch, 2022）。这些发现强调了一致性在个体评估和接受外部信息源时的重要性。

在志愿服务的背景下，一致性理论提供了一个重要的理论框架，可以用于分析志愿者的服务动机与算法推荐逻辑之间的匹配程度及其在算法透明度与算法信任的关系间发挥的作用。当志愿者的个人动机与算法所推荐的志愿活动高度一致时，这种一致性可能通过满足志愿者的动机，增强了算法透明度对算法信任的作用（辛华等，2016；Wang & Wang, 2019）。因此，一致性理论在分析志愿者如何通过算法透明度评估和建立对算法建议的信任方面发挥着关键作用，对于提升志愿者对推荐活动的参与度和满意度具有重要意义。本研究评估了算法透明度对算法可信度的影响，并从一致性的角度考察了不同动机类型（利己或利他）的调节作用。

（五）研究假设

在志愿服务领域，算法透明度与算法信任紧密相关。这种关系反映在两个维度。一方面，算法透明度的提高帮助志愿者理解算法决策过程，从而提高其对算法的信任程度；另一方面，志愿者服务动机与算法推荐逻辑之间的一致性会对算法透明度的作用产生进一步的影响。

算法推荐系统根据特定目标和其他相关因素，向用户推荐志愿活动（Benbasat & Wang, 2005）。通过数据分析和智能匹配，精准化推荐避免了传统人工搜索耗时长和精确性低的缺点，提高了志愿服务的效率和质量。在志愿服务情境下，算法的决策逻辑天然地分为利己逻辑和利他逻辑两种。然而，由于算法的“黑箱”特征，志愿者会对算法的决策过程和结果产生疑虑和不信任（昌诚等，2022），从而降低算法的应用价值。所以，从算法信任的角度来看，公布决策逻辑可以提高算法的可解释程度，使志愿者能够理解为何某些活动被推荐给他们（Grimmelikhuijsen, 2022），从而提高其对算法的信任（McKnight et al., 2002）。据此，本研究提出如下假设。

H1a：相对于不透明组，透明 - 利己逻辑组的算法感知可信度更高。

H1b：相对于不透明组，透明 - 利他逻辑组的算法感知可信度更高。

在志愿服务领域的算法推荐中，虽然透明度的提升有助于增进志愿者对算法推荐逻辑的理解，但仅提高透明度并不足以确保志愿者对推荐结果产生信任。一致性理论强调，公众在决策时会考虑自己的信念、态度和价值观，且仅当这些因素与接受到的信息内容高度匹配时，才会建立信任（Maille & Fleck, 2011）。人际互动和人机交互领域的研究发现，当信息源与个体的信念、态度和价值观高度一致时，个体更倾向于信任这些信息源（Byrne et al., 1967），从而证明了这一观点。换言之，感知一致对建立信任起着积极作用。

在志愿服务领域应用推荐算法的场景中，当志愿者的动机倾向于利他，且算法明确展示其推荐是如何服务于更广泛的社会利益时，志愿者对推荐结果的信任度可能增强。相应地，当志愿者的动机偏向利己，而算法展示出其聚焦于满足个人利益进行活动推荐时，信任度亦可能提升。因此，在提高志愿活动推荐算法的透明度时，仅当公布的决策逻辑与志愿者的服务动机高度一致时，才可能提高志愿者感知的算法可信度；若不一致，算法透明度所带来的积极效果可能有限（Coulter & Coulter, 2002）。

根据以上论证逻辑，提出以下研究假设。

H2a：当志愿服务动机是利己时，透明-利己逻辑的算法对感知算法可信度的积极作用更大。

H2b：当志愿服务动机是利他时，透明-利他逻辑的算法对感知算法可信度的积极作用更大。

此外，志愿者参加志愿服务的动机可能对感知一致性与算法可信度的关系产生进一步的影响。根据志愿服务的功能性理论，个体参与志愿服务是为了满足社交、自我实现、心理健康、学习和公民责任等需求，这些需求（即志愿服务动机）影响着公众的志愿行为（Clary et al., 1998）。志愿服务动机主要分为利己动机和利他动机两大类（张网成等，2018）。根据定义，利己动机志愿者主要受个人利益或自我满足驱动，通常更注重个人收益而非志愿服务的实际效果和影响（张网成，2015）。相反，利他动机志愿者更多是出于对他人和社会的责任感，关注参与志愿服务的社会价值和影响（Stukas et al., 2016）。这种动机差异使两类志愿者在对志愿活动的选择和接受度上存在显著差异。利己动机志愿者对于服务内容会有选择性，而利他动机志愿者则更加倾向于无差别服务，以服务对象的需求和利益为主（Omoto & Snyder, 1995）。因此，不同动机的志愿者对于算法推荐活动是否满足自身动机的敏感程度可能存在差异。据此，提出假设3。

H3：志愿服务动机调节了感知一致性与感知算法可信度的关系。

三、研究设计与方法

(一) 研究样本和实验分组

本研究选用调查实验方法，旨在通过实验操控，识别因果关系，清晰地揭示算法透明度与志愿服务动机是如何共同作用于算法信任的。此外，本研究通过开发模拟真实操作环境的 APP，提升了实验情境的真实感，进一步提高了结论的有效性，增强了研究发现的理论和实践价值。

实验数据来自见数平台。课题组在 2023 年 2 月 6 日到 3 月 3 日，将全国范围内的 885 位被试随机分配到不同的实验分组中，通过比对不同实验组被试感知的算法可信度，来测试不同的透明策略是否影响组间被试对于算法可信度的感知，以及影响程度在不同服务动机的志愿者之间是否存在差异。课题组通过问卷审核，剔除了 265 份无效问卷。无效问卷的审核标准包括：回答时长过长（大于 10 分钟）或过短（小于 4 分钟），所选择的实验 APP 主色调与实际的实验 APP 不符（红色以外的其他颜色），所选择的实验 APP 优先推荐活动类型与实际分组不符的样本（如分配为实验组 1 即“算法优先推荐满足利己动机的活动”，但被试在后测问题中选择的被优先推荐的活动类型为满足利他动机的活动）。最终，本研究的有效样本为 620 份问卷，足够满足在中等效应的前提下，设定 $\alpha = 0.05$ 并且检验效能为 0.85 时的样本要求。最终样本的人口统计学因素的描述性统计见表 1。

在实验开始时，研究人员根据对自变量的操控，将算法按照不同决策逻辑和透明度分成三组，并将被试随机分配到这三组中。共有 205 名参与者被分配到不透明组作为控制组，205 名参与者被分配到透明 - 利己逻辑算法实验组，210 名参与者被分配到在透明 - 利他逻辑算法实验组。参与者平均实验时间为 5.7 分钟，平均实验收入为 4.50 元。

(二) 实验流程

被试通过见数平台的随机模块功能被分配到三组中（如图 1 所示）。实验包含三个部分。第一部分是前测问题，调查了被试的人口统计学特征、参加志愿活动的动机以及其他可能与研究相关的控制变量（如上网经验）。第二部分是情境实验部分，要求被试充分阅读实验情境，通过实验情境材料分步骤引导被试首先扫描随机出现的二维码，进入实验性志愿活动推荐系统。接下来选择偏好的志愿活动地点、时间，最后点击开始智能匹配，并仔细查看推荐结果中是否有感兴趣的志愿活动。第三部分是后测题目，包括对于因变量的测量、操控检验以及注意力检查。本实验在自愿基础上进行，每位参与者参加实验的费用为

基础费用加上奖励金额。在参与者扫码进入推荐系统之前，告知他们额外奖励的金额将由他们的表现来决定，以此鼓励其认真对待实验。

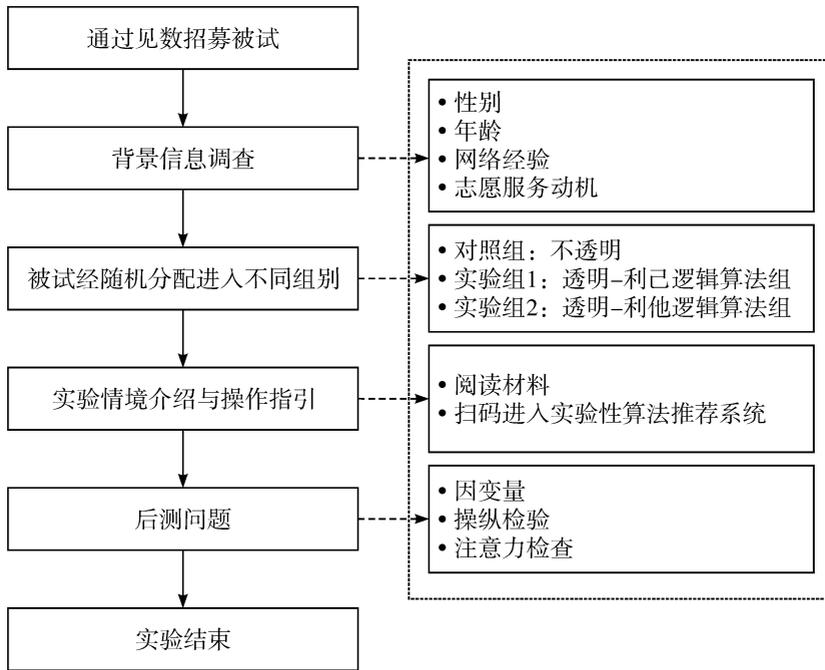


图1 实验流程

资料来源：作者自制。

（三）实验性算法推荐系统

根据本研究的核心问题（算法透明度是如何影响算法可信度的？志愿服务动机是如何影响算法透明度对算法可信度的作用的？），本研究设计了一个实验性算法推荐 APP，旨在模拟真实操作，帮助被试完成实验性决策任务，即选择他们感兴趣的志愿活动。实验性算法推荐系统的设计模仿了官方的社会公益活动管理应用程序——志愿汇。该设计在三位资深手机用户的协助下进行了验证，他们的反馈被纳入到设计的修改方案中，以确保应用程序设计的真实性和可靠性。另外，研究人员在正式实验之前进行了一系列的预测试，将预试验的反馈也纳入到实验设计中，以确保操控的有效性。

实验对自变量进行操控，将算法按照不同决策逻辑和透明度分成三组。在不透明组，参与者在选择了参与志愿活动的地点和时间偏好后，推荐系统直接显示了推荐活动列表，被试不会收到关于推荐算法决策逻辑的信息。在透明-利己逻辑算法组，参与者先选择了参与志愿活动的地点和时间偏好，然后在显示推荐活动列表前会出现一个额外的页面，该页面告知被试“算法在推荐活动时考虑了三个要素：距离、时间以及优先推荐有利于志愿者自身的活动”（如图

2 所示)。在透明 - 利他逻辑算法组, 告知被试“算法在推荐活动时考虑了三个要素: 距离、时间以及优先推荐有利于他人的活动”, 并保证三组之间的其他因素完全相同 (如图 3 所示)。



图 2 “透明 - 利己逻辑算法组” 界面设计
资料来源: 作者自制。



图 3 “透明 - 利他逻辑算法组” 界面设计

(四) 变量测量

1. 因变量

基于 McKnight 等 (2002) 对信任的定义, 信任包含能力、善意和诚实三个维度。这些维度从不同角度评估用户对算法的信任, 能力信任关注算法的效能和专业性, 善意信任涉及算法是否顾及用户的利益, 诚实信任则是对算法诚实性和承诺的评估 (McKnight et al., 2002)。这三种信任维度从不同角度评估用户对算法的信任, 对于理解公众如何看待和接受算法决策至关重要。

本研究调整了 Benbasat 和 Wang (2005) 的可信度量表以适应本研究的背景。三个题项分别是: 我认为中华志愿汇很好地发挥了活动推荐作用; 我觉得中华志愿汇在进行智能匹配时, 非常重视我参加志愿活动的目的; 我认为中华志愿汇提供的是公正的推荐。由被试按照七级量表打分。

2. 自变量

根据研究问题和实验设计部分的讨论, 研究的自变量是不同决策逻辑算法的透明情况 (1 = 不透明, 2 = 透明 - 利己逻辑算法, 3 = 透明 - 利他逻辑算法)。

3. 控制变量

为保证实验分组的科学性，根据已有文献测量了被试的性别、年龄和网络经验（包括通过网络看新闻、购物和查询信息三项活动的时长）（Wang & Benbasat, 2016; George et al., 2017; Grimmelikhuijsen, 2022）。

4. 其他变量

为测量用户的感知一致性，需要测量被试的志愿服务动机，并将其与所分配到的算法逻辑相比较。具体来说，先测量志愿服务动机，要求被试在问卷中选择参与志愿服务的主要目的，选择自我提升、获得志愿时长、其他有利于自身原因的为利己动机志愿者，选择帮扶弱势群体、回报社会、其他有利于他人原因的为利他动机志愿者。然后，当利己动机志愿者被分配至透明-利己逻辑算法组中，或利他动机志愿者被分配至透明-利他逻辑算法组中，则为一致，否则，则为不一致。

四、数据分析与研究发现

（一）实验操控和随机分组有效性检验

笔者先检验了实验操控的有效性，问卷中设置了两道七级量表问题。第一道题有关被试对算法决策逻辑的理解：我知道算法在推荐时的决策逻辑。第二道题有关被试对匹配活动的感知一致性：中华志愿汇优先匹配的活动类型（有利于志愿者自身的活动/有利于他人的活动）能够满足我的志愿服务动机。通过单因素方差分析发现，实验组和控制组对于第一道题的回答具有显著差异，透明-利己逻辑算法组对于算法逻辑的理解为 6.28，显著高于不透明组的 3.79 ($p < 0.001$)。透明-利他逻辑算法组的理解为 5.80，也显著高于不透明组 ($p < 0.001$)，但低于透明-利己逻辑算法组 ($p < 0.001$)。所分配算法逻辑与自身相一致的被试的感知一致性评分为 5.95，而不一致组为 4.07，两组间存在显著差异 ($p < 0.001$)，表明被试同样能感受到一致与不一致。从检验结果来看，本实验的操控是成功的。

然后进行随机分组检验，检验各组间个体的年龄、性别、网络经验的差异性。结果表明，除了“网络经验”在各组间存在显著差异外，其他变量均没有显著差异，说明实验的随机性得到了保证。网络经验变量在后续分析中被作为控制变量，所以对分析结果的准确性不会产生影响（George et al., 2017）。

（二）样本特征

样本分布主要集中在中青年群体，10-20岁人群占比 5.81%，21-30岁人

群占比 51.61%，31-40 岁人群占比 34.19%，41-50 岁人群占比 5.65%，50 岁以上人群占比 2.74%。性别略有差异，男性大约占 42.74%，女性占比稍高，约为 57.26%。在网上看新闻所用的时长主要集中在 1-4 小时，30 分钟以内的百分比为 8.55%，30 分钟到 1 小时的百分比为 15%，1-2 小时和 2-4 小时的占比相同，均为 22.26%，4-6 小时的占比为 19.52%，6 小时以上的约占 12.42%。在网上购物与在网上看新闻所用的时长分布非常相似，1-4 小时占比最高，大约为 50%。在网上查询信息所用的时间占比与以上两组略有不同，2-4 小时、4-6 小时、6 小时以上占比接近，约为 22%（如表 1 所示）。

表 1 样本人口学因素描述性统计

变量	变量序列	数量（人）	百分比（%）
性别	男性	265	42.74
	女性	355	57.26
年龄	10-20	36	5.81
	21-30	320	51.61
	31-40	212	34.19
	41-50	35	5.65
	50 以上	17	2.74
	每周在网上看新闻的时长	0-30 分钟	53
30-60 分钟		93	15
1-2 小时		138	22.26
2-4 小时		138	22.26
4-6 小时		121	19.52
6 小时以上		77	12.42
每周在网上购物的时长	0-30 分钟	45	7.26
	30-60 分钟	113	18.23
	1-2 小时	161	25.97
	2-4 小时	158	25.48
	4-6 小时	101	16.29
	6 小时以上	42	6.77
每周在网上查询其他信息的时长	0-30 分钟	21	3.39
	30-60 分钟	74	11.94
	1-2 小时	110	17.74
	2-4 小时	135	21.77
	4-6 小时	141	22.74
	6 小时以上	139	22.42

资料来源：作者自制。

(三) 主效应与调节效应分析

为检验 H1a 和 H1b，即不同决策逻辑算法的透明度对感知算法可信度的作用差异，本研究先通过方差分析对不同决策逻辑算法透明度的影响进行初步研究，图 4 展示了三组间感知算法可信度的差异。在透明-利己逻辑算法处理组中，感知算法可信度的评分为 5.81（七级量表），显著高于不透明组算法的感知可信度 5.36（ $p < 0.001$ ）。在透明-利他逻辑算法处理组，被试感知的算法可信度为 5.57，也显著高于不透明组算法的感知可信度（ $p = 0.012$ ）。

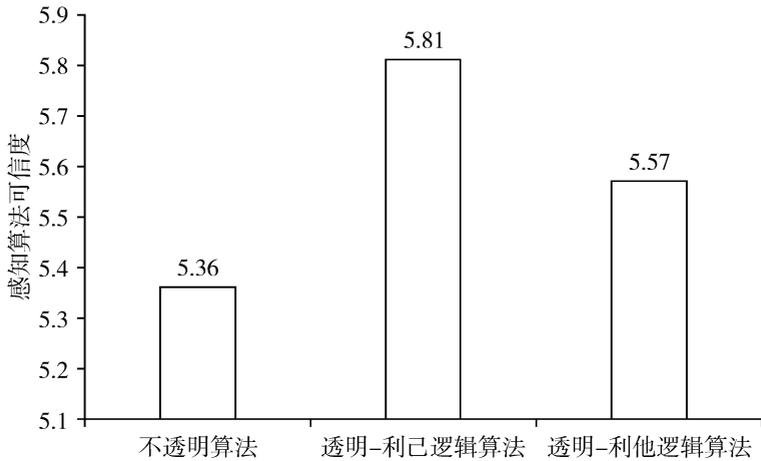


图 4 不同组别的感知算法可信度

资料来源：作者自制。

然后，基于普通最小二乘（OLS）回归对 H1a 和 H1b 进行具体检验（如表 2 所示）。将透明-利己逻辑算法和透明-利他逻辑算法作为自变量，不透明组作为常数项，将感知算法可信度作为因变量（模型 1）进行回归分析。统计结果表明，志愿者对于透明-利己和透明-利他算法的感知可信度均显著高于不透明算法，且对于透明-利己逻辑算法的感知可信度最高。当加入控制变量后，模型 2 的结果与在模型 1 观察到的是相同的。所以，H1a 和 H1b 得到了验证。

表 2 基准回归

自变量	因变量	
	感知算法可信度	
	(1)	(2)
透明-利己逻辑算法	0.45***	0.41***
透明-利他逻辑算法	0.20**	0.16*
常数项	5.36***	5.26***
样本量	620	620

(续上表)

自变量	因变量	
	感知算法可信度	
	(1)	(2)
控制变量	无	有
R ²	0.05	0.08

注：回归系数为标准回归系数。***、**和*分别表示相关系数通过0.01、0.05和0.10水平的显著性检验。

资料来源：作者自制。

上述分析主要是研究算法透明度对用户感知算法可信度的影响，接下来探究该影响在多大程度上受到用户志愿服务动机的调节（H2a和H2b）。本研究使用Stata进行调节效应分析，结果如表3所示。

表3 志愿服务动机的调节效应分析

自变量	因变量	
	感知算法可信度	
	(1)	(2)
透明-利己逻辑算法	0.71***	0.69***
透明-利他逻辑算法	-0.01	-0.03
透明-利己逻辑算法×志愿服务动机	-0.46***	-0.47***
透明-利他逻辑算法×志愿服务动机	0.48***	0.46***
志愿服务动机	0.18	0.16
常数项	5.25***	5.35***
样本量	620	620
控制变量	无	有
R ²	0.11	0.14

注：回归系数为标准回归系数。***、**和*分别表示相关系数通过0.01、0.05和0.10水平的显著性检验。

资料来源：作者自制。

分析结果表明，用户的志愿服务动机调节了不同决策逻辑算法的透明度对算法可信度的影响，并且在加入控制变量后结果依然相同，图5和图6更直观地展示了这一效应。具体来说，交乘项透明-利己逻辑算法×志愿服务动机的系数在1%的水平上显著为负，表明透明-利己逻辑对算法可信度的正向作用在利己动机志愿者中更为明显。交乘项透明-利他逻辑算法×志愿服务动机的系数在1%的水平上显著为正，表明透明-利他逻辑对算法可信度的正向作用主要存在于利他动机志愿者中。总的来说，本研究发现，当算法决策逻辑与志愿者服务动机相一致时，透明对于算法可信度的正向作用更大。因此，H2a和H2b得以验证。

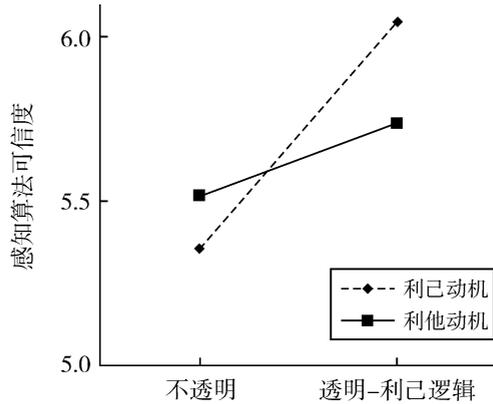


图5 志愿服务动机在不透明与透明-利己逻辑组中的调节效应

资料来源：作者自制。

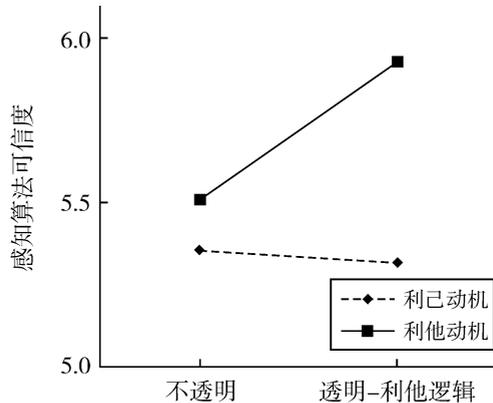


图6 志愿服务动机在不透明与透明-利他逻辑组中的调节效应

资料来源：作者自制。

由上述假设检验可知，相对于不透明算法，算法逻辑不同的两种透明算法（透明-利己算法和透明-利他算法）均对算法的感知可信度有正向作用。此外，研究发现，志愿者服务动机调节了算法透明度与感知算法可信度的关系。具体而言，当志愿服务动机与算法决策逻辑一致时，算法透明度对于感知可信度的正向作用更大。

（四）进一步分析

接下来，本研究进一步分析了在不同志愿服务动机的情况下，志愿服务动机与算法决策逻辑的一致性对算法可信度的影响。具体来说，笔者将实验数据按照算法决策逻辑与志愿服务动机的一致性划分为一致与不一致两组，构建了一致性这一变量（0 = 一致，n = 205；1 = 不一致，n = 210），并分析了志愿服务动机在一致性与算法可信度之间的调节效应（见表4）。

表 4 事后检验

自变量	因变量	
	感知算法可信度	
	(1)	(2)
一致性	0.72***	0.70***
一致性 × 公共服务动机	-0.51***	-0.52***
公共服务动机	0.44***	0.44***
常数项	5.24***	5.19***
样本量	415	415
控制变量	有	有
R ²	0.16	0.17

注：回归系数为标准回归系数。***、**和*分别表示相关系数通过0.01、0.05和0.10水平的显著性检验。

资料来源：作者自制。

分析结果如表4所示，交乘项一致性 × 志愿服务动机的系数在1%的水平上显著为负，这表明志愿服务动机在一致性对算法可信度的影响中具有调节作用。具体来说，一致性对算法可信度的正向作用在利己动机志愿者中更为显著，并且在加入控制变量后结果依然相同，H3得以验证。图7更直观地展示了这一效应。

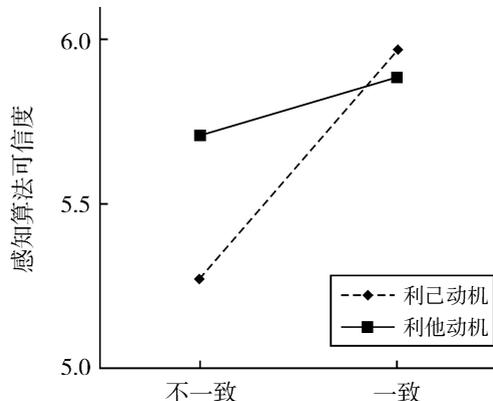


图 7 志愿服务动机在不一致与一致组中的调节效应

资料来源：作者自制。

五、结论与讨论

本研究基于一致性理论，以志愿服务为研究背景，通过调查实验方法探究了不同决策逻辑算法的透明度对其可信度的影响。研究表明，相较于不透明的算法，志愿者对于透明 - 利己逻辑和透明 - 利他逻辑的算法感知可信度更

高。并且，不同的志愿服务动机在其中发挥了调节作用。具体来说，当志愿者持有利己动机时，透明-利己逻辑算法对于感知算法可信度的正向作用更大；当志愿者持有利他动机时，对透明-利他逻辑的算法感知可信度也更高。总的来说，相较于不透明算法，志愿者对于透明的算法更加信任，并认为决策逻辑与自身动机相一致的算法可信度最高。此外，本研究还探索了感知一致性对算法信任的正向作用在不同动机志愿者中的差异，结果显示利己动机的志愿者感受到的一致性效应比利他动机志愿者高。一个可能的解释是，利己动机志愿者更加注重自身目的的达成（Omoto & Snyder, 1995）。

在算法治理领域的现有研究基础上，本文进一步探讨了算法透明度策略的有效性边界，在理论层面有两个贡献：首先，广泛的研究表明，提高透明度有助于增强信任；然而，当算法被感知为过于简单时，增加透明度反而会产生负面效果（Lehmann et al., 2022）。也就是说，算法透明策略的正向作用是有边界的。在此基础上，本文聚焦于志愿服务这一特定场景，发现只有当算法逻辑与志愿者的服务动机一致时，提升算法透明度才能有效增强志愿者对算法建议的信任，这印证了确认偏差的理论，也凸显了一致性在人与算法推荐交互中的核心作用（Taber & Lodge, 2006）。此外，这一结论进一步验证了 Selten 等人（2023）所发现的一致性在建立算法信任中发挥的关键作用。Selten 等人（2023）发现，在通过算法预测警务的场景中，警务人员对与其专业知识一致的算法建议表现出更高的信任度。而本文在志愿服务的场景中发现，志愿者对与其服务动机一致的算法建议同样表现出更高的信任度。同时，两项研究结果之间的对比揭示出一个重要观点：在不同的应用场景中，构建信任所需的一致性要素具有显著差异。在警务场景中，信任基于专业知识与算法逻辑的一致；而在志愿服务场景中，信任则基于志愿服务动机与算法逻辑的一致。这些发现为理解算法透明度在不同场景下如何影响信任提供了新的视角，并强调了在制定算法透明度策略时考虑特定情境的重要性。

其次，与将所有用户作为整体的研究逻辑不同，本文通过行为科学实验对个体的认知进行探究。实验结果发现，虽然从志愿者群体的角度来看，服务动机的满足会产生积极效果（Snyder & Omota, 2008），但利己动机志愿者和利他动机志愿者对于动机满足的敏感度有所不同，对于利己动机志愿者来说，透明一致对于感知算法可信度的正向作用更大。虽然在志愿服务领域对不同动机志愿者进行分类讨论的研究较多（Stukas et al., 2016），但是在算法领域，对用户进行分类研究的很少（Wang & Wang, 2019）。本研究在志愿服务领域讨论算法信任问题时，充分考虑到了不同动机志愿者的差异性。

本研究为志愿服务领域的算法治理提供了关键性的实证指导。研究发现，在提高算法透明度的过程中不仅要公开算法的决策逻辑，还需考虑到志愿者的服务动机，使得算法逻辑与志愿者的期望相一致，以最大化算法透明度对算法

信任的积极影响。进一步地,研究指出,算法推荐应考虑不同的用户需求,尤其是以个人目标为导向的利己动机的志愿者。因为利己动机的志愿者对志愿服务动机与算法逻辑间的一致性更为敏感。同时,本研究还强调了公共部门在算法设计时需要考虑场景的特异性。本研究建议,在算法设计和实施的各个阶段加强算法开发人员、管理部门和志愿者的共同参与,以确保算法推荐能够针对特定场景有效发挥其作用。

本研究仍存在一定的局限性。虽然利己动机与利他动机是最主要的志愿服务动机,但是可能还有介于中间状态的动机(Clary et al., 1998),比如对于利己和利他都没有足够兴趣,或者对两者都很在意的动机。后续研究可以在更加细分的动机基础上进行,但是无论动机是什么,只要能够满足主要动机,就应该会减少因动机类型而产生的结果差异。另一个局限是实验中包含的解释类型。人工智能领域认为如何解释算法取决于用户的认知和理解能力(Ahmad et al., 2018),本研究所提供的解释是对于算法决策逻辑的解释,未来的研究可以系统地考察不同表达方式和表达内容组合的效果。

参考文献

- 昌诚、张毅、王启飞(2022). 面向公共价值创造的算法治理与算法规制. *中国行政管理*, 448(10): 12-20.
- Chang, C., Zhang, Y. & Wang, Q. F. (2022). Algorithmic Governance and Algorithmic Regulation: A Perspective of Public Value Creation. *Chinese Public Administration*, 448(10): 12-20. (in Chinese)
- 党秀云(2019). 论志愿服务可持续发展的价值与基础. *中国行政管理*, 413(11): 118-123.
- Dang, X. Y. (2020). On the Value and Foundation of the Sustainable Development of Volunteer Service. *Chinese Public Administration*, 413(11): 118-123. (in Chinese)
- 李晓园、钟伟、滕玉华(2022). 互联网赋能如何影响政府相对贫困治理绩效?——开放式创新的中介作用. *公共行政评论*, 15(03): 161-178+200.
- Li, X. Y., Zhong, W. & Teng, Y. H. (2022). How Does Internet Empowerment Affect the Government's Performance in Relative Poverty Governance? The Mediating Role of Open Innovation. *Journal of Public Administration*, 15(03): 161-178+200. (in Chinese)
- 苗青、尹晖(2021). 从转移价值到放大价值:论慈善事业在第三次分配浪潮中的增长路径. *中国非营利评论*, 28(2): 8-12.
- Miao, Q. & Yin, H. (2021). From Transferring Value to Enlarging Value: On the Growth Path of Philanthropy in the Third Distribution Wave. *China Nonprofit Review*, 28(2): 8-12. (in Chinese)
- 辛华、周凌一、陈敏(2016). 志愿共同体:志愿服务的引力、能力与动力——基于美好社会咨询社的案例研究. *中国非营利评论*, 17(1): 197-210.
- Xin, H., Zhou, L. Y. & Chen, M. (2015). Volunteering Community: The Incentive, the Capability and the Motive. *China Nonprofit Review*, 17(1): 197-210. (in Chinese)
- 张雪帆、蒋志楠(2022). 公共行政的数字阴影:数字政府建设中的伦理冲突. *公共行政评论*, 15(05): 164-181+200.
- Zhang, X. F. & Jiang, Z. N. (2022). The Digital Shadow of Public Administration: Ethical Conflicts in Building a Digital Government. *Journal of Public Administration*, 15(5): 164-181+200.
- 张网成(2015). 当代中国企业雇员志愿服务解析. *中国非营利评论*, 15(01): 130-151.
- Zhang, W. C. (2015). Analysis of Voluntary Service by Chinese Contemporary Enterprise Employees. *China Nonprofit Review*, 15(1): 130-151. (in Chinese)
- 张网成、陈锋、刘小燕(2018). 青年缺位本社区志愿服务原因分析——基于北京某社区调查. *中国非营利评论*, 21(01): 180-199.
- Zhang, W. C., Chen, F. & Liu, X. Y. (2018). Reasons for Youth Absence in Community Volunteer Service: Based on a Community Survey in Beijing. *China Nonprofit Review*, 21(1): 180-199. (in Chinese)

- 张瑞 (2021). 公共危机中民间志愿服务的演变机理——以武汉市新冠肺炎疫情为例. *中国非营利评论*, 27(1): 197–217.
- Zhang, R. (2021). The Evolution Mechanism of Non-governmental Voluntary Services in Public Crisis: Take the COVID–19 Epidemic in Wuhan as an Example. *China Nonprofit Review*, 27 (1): 197–217. (in Chinese)
- Ahmad, M. A. , Eckert, C. & Teredesai, A. (2018). Interpretable Machine Learning in Healthcare. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 559–560.
- Benbasat, I. & Wang, W. (2005). Trust In and Adoption of Online Recommendation Agents. *Journal of the Association for Information Systems*, 6(3): 4.
- Byrne, D. , Griffitt, W. & Stefaniak, D. (1967). Attraction and Similarity of Personality Characteristics. *Journal of Personality and Social Psychology*, 5(1): 82–90.
- Clary, E. G. , Snyder, M. & Stukas, A. A. (1996). Volunteers' Motivations: Findings from a National Survey. *Nonprofit and Voluntary Sector Quarterly*, 25 (4): 485–505.
- Clary, E. G. , Snyder, M. , Ridge, R. D. , Copeland, J. , Stukas, A. A. , Haugen, J. & Miene, P. (1998). Understanding and Assessing the Motivations of Volunteers: A Functional Approach. *Journal of Personality and Social Psychology*, 74(6): 1516.
- Coulter, K. S. & Coulter, R. A. (2002). Determinants of Trust in a Service Provider: The Moderating Role of Length of Relationship. *Journal of Service Marketing*, 16: 35–50.
- Emrich, E. & Pierdzioch, C. (2016). The Internet and the Commitment of Volunteers: Empirical Evidence for the Red Cross. *Nonprofit and Voluntary Sector Quarterly*, 45 (5), 1013–1030.
- George, B. , Desmidt, S. , Nielsen, P. A. & Baekgaard, M. (2017). Rational Planning and Politicians' Preferences for Spending and Reform: Replication and Extension of a Survey Experiment. *Public Management Review*, 19(9): 1251–1271.
- Grimmelikhuijsen, S. (2022). Explaining Why the Computer Says No: Algorithmic Transparency Affects the Perceived Trustworthiness of Automated Decision-Making. *Public Administration Review*, 83 (2): 241–262.
- Keiser, L. R. & Miller, S. M. (2020). Does Administrative Burden Influence Public Support for Government Programs? Evidence from a Survey Experiment. *Public Administration Review*, 80(1): 137–150.
- Lehmann, C. A. , Haubitz, C. B. , Fügener, A. & Thonemann, U. W. (2022). The Risk of Algorithm Transparency: How Algorithm Complexity Drives the Effects on the Use of Advice. *Production and Operations Management*, 31(9): 3419–3434.
- McKnight, D. H. , Choudhury, V. & Kacmar, C. (2002). Developing and Validating Trust Measures for E-commerce: An Integrative Typology. *Information Systems Research*, 13 (3): 334–359.
- Maille, V. & Fleck, N. (2011). Perceived Congruence and Incongruence: Toward a Clarification of the Concept, Its Formation and Measure. *Recherche et Applications en Marketing (English Edition)*, 26(2): 77–113.
- Omoto, A. M. & Snyder, M. (1995). Sustained Helping Without Obligation: Motivation, Longevity of Service, and Perceived Attitude Change among AIDS Volunteers. *Journal of Personality and Social Psychology*, 68(4): 671.
- Reinkemeier, F. & Gnewuch, U. (2022, January). Match or Mismatch? How Matching Personality and Gender Between Voice Assistants and Users Affects Trust in Voice Commerce. In *Proceedings of the 55th Hawaii International Conference on System Sciences*.
- Wang, W. & Benbasat, I. (2016). Empirical Assessment of Alternative Designs for Enhancing Different Types of Trusting Beliefs in Online Recommendation Agents. *Journal of Management Information Systems*, 33 (3): 744–775.
- Wang, W. & Wang, M. (2019). Effects of Sponsorship Disclosure on Perceived Integrity of Biased Recommendation Agents: Psychological Contract Violation and Knowledge-based Trust Perspectives. *Information Systems Research*, 30 (2): 507–522.
- Paul, J. & Bhakar, S. (2018). Does Celebrity Image Congruence Influences Brand Attitude and Purchase Intention? *Journal of Promotion Management*, 24(2): 153–177.
- Snyder, M. & Omoto, A. M. (2008). Volunteerism: Social Issues Perspectives and Social Policy Implications. *Social Issues and Policy Review*, 2 (1): 1–36.
- Stukas, A. A. , Hoye, R. , Nicholson, M. , Brown, K. M. & Aisbett, L. (2016). Motivations to Volunteer and Their Associations with Volunteers' Well-being. *Nonprofit and Voluntary Sector Quarterly*, 45 (1): 112–132.
- Taber, C. S. & Lodge, M. (2006). Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science*, 50(3): 755–69.

责任编辑：郑跃平