

基于信任视角的消费者算法态度 研究述评与展望*



孙羽佳 苏 淞 唐红红

(北京师范大学经济与工商管理学院, 北京 100875)

内容提要:近期,随着以 ChatGPT 为代表的新型聊天机器人模型的应用,消费者通过与算法模型的高效智能对话,享受到了便捷、个性化、丰富有趣的服务。智能算法正快速融入到人们的日常生活中,使得消费者对于算法的态度也在不断发生变化,早期研究发现的人们对于算法的消极态度即算法厌恶已经不再适合当下的社会环境。通过收集和分析算法态度文献结论和发展脉络,本研究发现,消费者算法态度正从厌恶逐渐向信任转变,因此,亟需从算法态度形成的深层机制探索消费者算法态度发展脉络和方向。本研究旨在从信任的视角来解释这种行为的转变,并提出了消费者算法信任的三维框架,分别为:倾向性信任、情境信任和习得信任。三重信任相互影响,共同作用,最终从整体性的维度驱动了消费者递进式的行为表现:算法欣赏、算法选择和算法依赖。鉴于数字经济时代人机交互对消费者个人及企业的重要性,本研究将对消费者相关决策、企业实践,以及相关政策的制定提供有益的借鉴。

关键词:算法信任 三维框架 算法欣赏 算法选择 算法依赖

中图分类号:F270.7;F272.3 **文献标志码:**A **文章编号:**1002—5766(2023)10—0188—21

一、引言

在数字经济时代,数据成为了一种宝贵的资源,而算法(algorithm)是利用这一宝贵资源的关键工具。算法是根据特定的计算规则将输入数据转换为所需的输出,以完成某项任务的一系列编码程序(Castelo 等,2019)^[1]。算法的自然语言生成、预测能力,以及从经验中学习、模仿人类情感等特性,使得它在各个领域都得到广泛应用(Castelo 等,2019^[1];Dietvorst 和 Bharti,2020^[2])。在智慧城市、智能制造、智慧医疗等领域,算法被用来优化资源分配、提高生产效率和实现精准医疗;在智能购物、智慧出行、智能娱乐等日常生活中,算法被用来推荐商品、规划路线和提供个性化服务;在语音识别、图像处理、自然语言处理等人工智能领域,算法被用来解决复杂问题、改善用户体验。新近出现的交互式算法模型更进一步拉近了用户与人工智能之间的距离。美国人工智能公司 OpenAI 开发的大型 AI 模型——ChatGPT 通过构建用户与人工智能的交互式对话系统,使得“与 ChatGPT 对话”迅速成为国内外消费者和研究人员的关注热点(Kasneci 等,2023)^[3]。

收稿日期:2023-04-17

* 基金项目:国家自然科学基金项目“基于认知神经机制的消费者财务决策规避行为研究”(71872016)。

作者简介:孙羽佳,女,博士研究生,研究领域是数字营销和消费者行为,电子邮箱:202231030021@mail.bnu.edu.cn;苏淞,男,教授,博士生导师,研究领域是决策心理和数字伦理,电子邮箱:sus@bnu.edu.cn;唐红红,女,副教授,研究领域是决策和消费者心理与行为,电子邮箱:tang.h@bnu.edu.cn。通讯作者:唐红红。

理论方面,算法的广泛应用吸引了学术界对算法的探讨。其中,消费者算法态度是一个备受关注的新兴研究领域。这种态度不仅驱动着消费者行为,还深刻影响了社会关系和技术发展方向,因此具有极其重要的营销管理价值。现有研究已经取得了一些关于算法态度的重要成果,但仍存在三点空白之处。首先,现有关于算法厌恶和算法欣赏的研究主要集中在对这些现象的描述上,较少对行为倾向背后的深层内在机制进行解释或挖掘,因而缺乏对构建消费者对算法的信任以改变消费者算法态度的探讨。其次,在 AI 技术快速发展的时代背景下,现有文献主要从静态和单一的角度来审视算法态度,这不利于从整体上把握该研究热点的发展趋势。随技术发展,算法在生活中的角色和应用方式快速变化。因此,需要动态关注算法态度如何随着时间推移而演变,以及这些演变如何塑造行为和决策。最后,现有文献从个人、技术、任务等不同视角探讨了算法态度的影响因素,但缺乏系统性的归纳梳理。这些前置因素引起的消费者行为缺乏区分性整理,影响机制缺少归纳性说明。

实践方面,2022 年人工智能技术迎来拐点,算法出现颠覆性变革^①。在人工智能变革的背景下,构建算法信任具有重要的跨时代意义。从企业角度,利用算法对数据进行处理和深度挖掘,已经成为影响信息分发、服务提供、资源配置的基础性机制和力量,是在数字经济时代构建新发展格局的保障(Kumar 等,2019)^[4]。从国家角度,通过数字经济提升居民收入水平、优化产业结构及满足动态需求,能够驱动消费升级,为经济社会高质量发展贡献力量^②。自“十二五”规划以来,国家从宏观层面上对算法应用产业给予巨大支持,不仅如此,也加强了对数字技术研究的支持和引导^③。因此,研究消费者算法信任有助于政府和企业了解消费者对于算法的态度和期望,为制定相关政策和业务决策提供有力支撑,促进算法在商业和社会中的公正、可信、透明和负责任使用。

针对以上问题,本文首先对国内外消费者算法态度相关文献进行收集和梳理,并进行了科学计量分析,发现该研究主题呈现“算法模型—态度—行为”的趋势,分析结果纵向展示出人们对算法的态度转变。早期的研究发现,人们对算法持有消极态度,被定义为算法厌恶(algorithm aversion)(Dietvorst 等,2015)^[5]。但近期的研究证实了人们在特定条件下接受算法(Castelo 等,2019)^[1],或偏好算法建议(Logg 等,2019)^[6]。这为“消费者算法态度转变”提供了重要依据。随后,本研究对 20 多年来与算法相关的研究进行了梳理,基于动态研究视角发现,信任构建在连接本研究主题和促进消费者使用算法方面均占据重要地位。在此基础上,本研究从算法研究和人类信任的文献的交叉点出发,基于人际信任和人机信任的理论拓展,提出算法信任的概念,旨在探讨其作为内在机制以深入理解和改变消费者行为态度。再次,本文在已有文献基础上提供了一个包含三层次算法信任的一般框架(如图 1 所示),即包含倾向性信任、情境信任和习得信任。本研究将个人、算法变量和其他权变因素纳入算法信任的三维框架中,并基于双重加工理论分析了相对应的潜在机制。三维信任综合作用,引起不同程度的三种行为表现:算法欣赏、算法选择和算法依赖。

最后,本文基于整体脉络并结合技术爆发的时代背景,以前瞻性的动态视角,为未来研究提供了有趣的研究议题,帮助消费者和企业正确看待算法并在交互中充分利用算法提供的好处,促进良性人机关系的形成。未来,算法应用后续市场空间广阔,本研究为推动算法在各种社会情境中的有

① 中信建投,《从 CHAT—GPT 到生成式 AI(Generative AI):人工智能新范式,重新定义生产力》。

② 于凤霞. 以数字经济驱动消费升级[EB/OL]. (2023-06-23)[2023-07-13]. http://paper.ce.cn/pc/content/202306/21/content_276185.html.

③ 刘海军. 开展调查研究要用好数字化新手段[EB/OL]. (2023-04-14)[2023-09-25]. https://www.gov.cn/zhengce/2023-04/14/content_5751824.htm.

效应用提供基础,帮助市场实践者结合算法信任框架中的各类因素,以最大化算法的效率、生产力和用户满意度,引导算法信任,使得算法能够更好地造福社会多方。

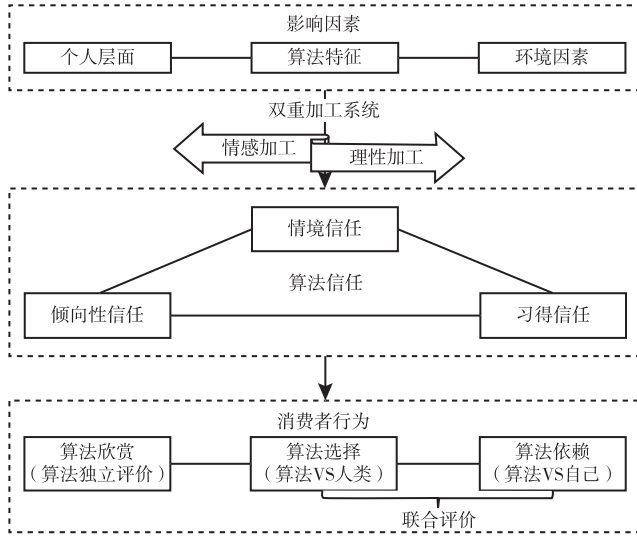


图 1 研究框架

二、消费者算法态度发生积极转变

为了深入研究算法态度的研究现状和转变趋势,本文借助 CiteSpace 软件采用文献计量方法 (Chen,2006)^[7],对 Web of Science 数据库中的相关研究进行了可视化分析。首先,在构建文献检索策略时,本研究明确定义了与算法态度相关的关键词和术语,以确保检索结果的准确性和相关性,包括:“algorithm”和“consumer attitude”;“algorithm”和“decision making”;选择 Business Economics、Psychology、Computer Science 和 Behavioral Sciences 类别主题的文献后,检索出 2000 年 1 月—2023 年 6 月收录的算法态度相关论文,类型涵盖期刊文章、会议论文和综述。然后对以上文献样本的摘要和内容进行人工筛选,剔除重复和排除了关于算法架构和模型的文章,最终对 615 篇相关文献进行了分析。

在进行可视化分析之前,本文首先对文献的年度发文量进行了分析(如图 2 所示)。某一主题的发文数量越多,说明该领域的研究越活跃(黄敏学和吕林祥,2022)^[8]。结果显示,从 2019 年开始,与算法态度相关的研究动态明显活跃,呈现指数增长的趋势(曲线拟合的 R^2 为 0.867)。这一变化可能部分归因于以人工智能为基础的算法技术的进步。这种技术的快速发展不仅使算法的应用变得更为普遍,还提供了更多研究机会。因此,算法态度的研究不仅反映了学术界对这一主题的浓厚兴趣,也与现实世界中的算法应用息息相关。

为了进一步探索和揭示这些文献之间的引用关系和关键词发展趋势,为下一步的研究和讨论提供基础,本文首先运用 CiteSpace 软件进行关键词共现分析得到了消费者算法态度的研究热点(如图 3 所示),然后进行文献共被引分析挖掘了消费者算法态度知识基础(如图 4 所示)。关键词概括了研究主题。在相关文献中,频次较高的关键词通常反映了研究领域内的研究方向和热点,中心性意味着关键词在领域内的影响力。图 3 列示了含有 156 个结点,251 条关联路径的关键词共现网络。“人工智能 artificial intelligence”“影响 impact”“采纳 adoption”“信任 trust”等关键词对应的节点最大,且与其他节点连接较多,表明它们在相关文献中经常共同出现,并具有重要的桥梁作用。频次较高的共引文献通常反映了该研究的前沿性。Logg 等(2019)^[6]发现了算法欣赏的现象,

即人们更倾向于依赖算法的建议而不是人的建议。图4显示该研究为算法态度领域的核心文献，也对应了算法态度发表文献量自2019年的激增趋势。截至2023年10月该研究在谷歌学术平台的引用次数为918次,为后续研究提供了指导。

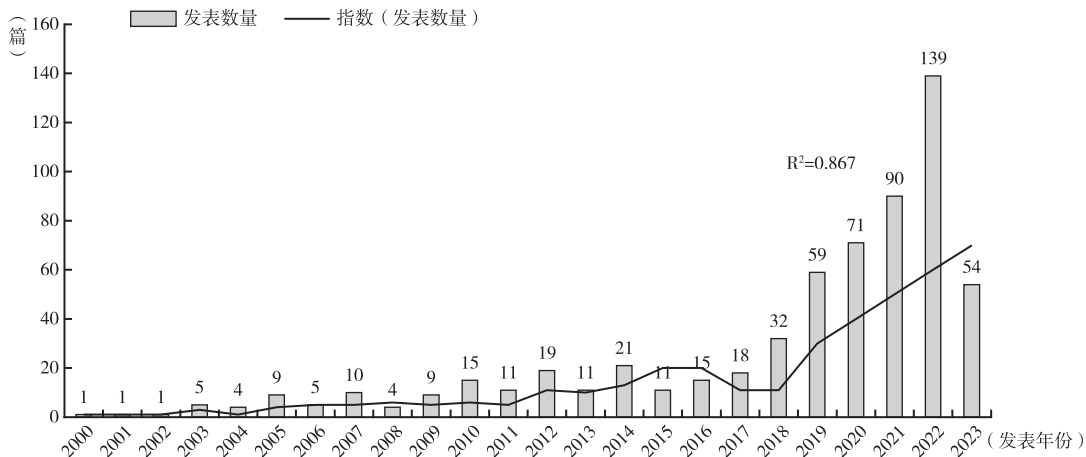


图2 消费者算法态度研究文献发表趋势 (2000—2023年)

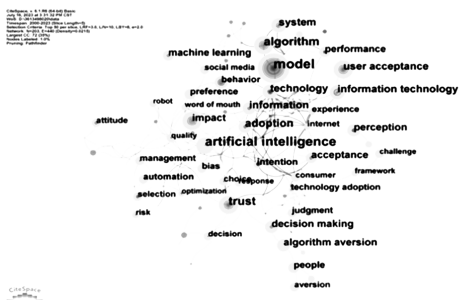


图3 消费者算法态度关键词共现图谱

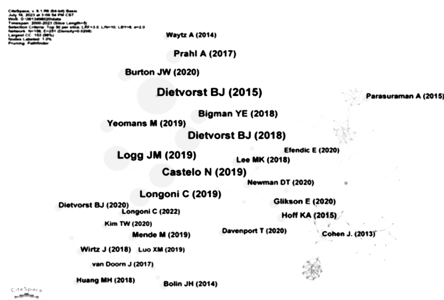


图4 文献共被引图谱

此外,本文还对高频关键词及其中心性进行了整理和统计(如表1所示)。高频关键词包括“模型 model”(频次 96)和“人工智能 artificial intelligence”(频次 84),这符合研究中将算法视为一种重要的人工智能技术的观点,其重要性在于模型的构建和应用。另外,“信任 trust”作为一个关键概念,涉及人们对算法的信赖和依赖程度,也高频出现(频次 51)。其他涉及消费者行为倾向的关键词包括“采纳 adoption”“用户接受 user acceptance”“决策 decision making”“行为 behavior”“选择 choice”等。在人工智能的发展背景下,这些关键词反映了研究关注消费者在与算法互动和使用过程中的态度和行为表现。

表1 消费者算法态度高频关键词统计

排名	频次	中心性	年份	关键词
1	96	0.11	2003	model
2	84	0.08	2019	artificial intelligence
3	52	0.07	2001	algorithm
4	51	0.12	2005	trust
5	38	0.57	2005	adoption
6	33	0.13	2010	information

续表 1

排名	频次	中心性	年份	关键词
7	32	0.23	2005	technology
8	32	0.1	2011	information technology
9	31	0.08	2012	user acceptance
10	29	0.16	2015	decision making
11	28	0.1	2020	algorithm aversion
12	27	0.05	2008	acceptance
13	27	0.03	2010	perception
14	25	0.07	2005	performance
15	22	0.07	2010	intention
16	20	0.03	2011	behavior
17	18	0.02	2016	preference
18	17	0.25	2005	technology adoption
19	17	0.16	2010	management
20	14	0.28	2010	choice

注:中心性指的是这一关键词对本领域和其他领域影响力的大小,中心性算法公式为: $C_D(N_i) = \frac{\sum_{j=1}^G x_{ij}}{N_i} (i \neq j)$,其值越大,说明该关键词被引次数越多;年份表示某关键词 Burst 的年份,代表该年有多少文章关注这一关键词

为了揭示不同时间段消费者算法态度研究领域的发展过程和关注点的变化,本研究对关键词进行纵向时间序列分析,使用“Time Zone”时区网络视图方法(如图 5 所示),从动态视角揭示研究趋势转变。2000—2023 年期间,该领域的研究呈现出三个阶段的发展脉络:算法表现探索阶段、人机关系探究阶段、算法态度探究阶段。

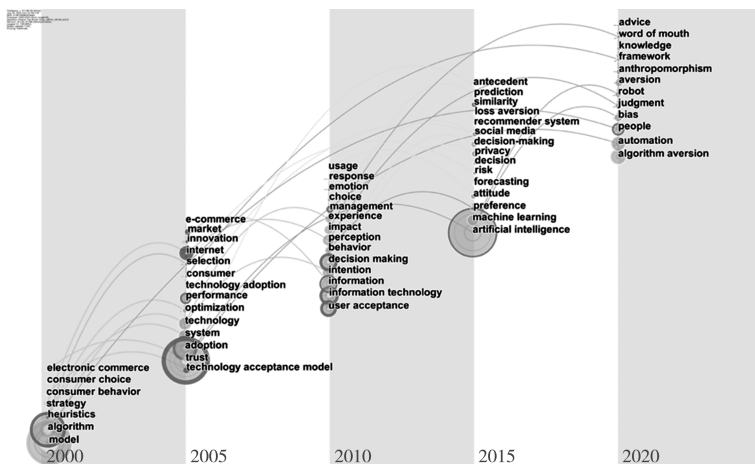


图 5 消费者算法态度关键词时区图谱

1. 算法表现探索阶段

在 2000—2004 年期间,研究主要关注算法的技术层面,探索算法在特定任务中的性能、效果和应用范围。研究者们致力于寻找最佳的算法策略和模型。例如, Grove 等(2000)^[9]通过临床进行了关于人类健康的元分析,发现算法的预测准确度平均比临床预测高约 10%。该阶段的研究还探讨了消费者对算法的启发式认知和态度,以了解消费者对这种新技术的接受程度。研究发现,消费者将算法视为一种“接近完美的自动化解决方案”(Dzindolet 等,2002)^[10],从而对算法设定了较高

的初始期望,导致对算法存在偏见。这种偏见增加了用户注意到算法错误的倾向,为后续的算法厌恶研究奠定了基础。

2. 人机关系探究阶段

在2005—2015年期间,学者们逐渐认识到算法不仅是一种技术工具,更是对消费者行为和决策产生影响的重要因素。研究发现,消费者对算法持有消极态度,表现为对算法生成的估计或预测结果持怀疑态度(Önkal等,2009^[11];Madhavan和Wiegmann,2007^[12]),这种低信任度导致消费者不选择算法建议(Madhavan和Wiegmann,2007)^[12]。这一现象被称为算法厌恶(Dietvorst等,2015)^[5]。在这一阶段,研究探索了不同主体(消费者、管理者、学生、患者)及不同环境(医疗保健、会计、金融、道德决策)中算法厌恶的情况,并将算法与不同人类主体进行了比较。同时,信任作为人机关系的重要机制开始受到关注(Hoff和Bashir,2015)^[13],推动进一步探讨信任与算法之间的关系。

3. 算法态度探究阶段

2016—2023年,随着研究的深入,学者们通过干预措施来改善消费者的态度,包括改变算法设计和展示算法表现等方面。例如,Cadario等(2021)^[14]在两项实验室研究中发现,增加对算法决策过程的简单解释可以增加人们利用算法医疗服务的意愿。此外,研究结果显示,消费者对算法态度发生了积极转变,包括在特定条件下的接受和纯粹的偏好。例如,Logg等(2019)^[6]发现,决策者支持算法而不是厌恶,他们更倾向于相信算法而不是人类代理的估计结果。同时,这一阶段的研究主题从信任扩展到行为领域。例如,随着对算法能力的信任提升,消费者更容易购买算法推荐的功利型产品(Longoni和Cian,2022^[15];Wien和Peluso,2021^[16])。

三、构建算法信任的理论框架及内涵解释

1. 算法信任的定义

算法作为数字技术的代表,其定义随着发展被不同的学者提出(Castelo等,2019^[1];Dietvorst等,2015^[5];Araujo等,2020^[17])。整合过去研究中对算法的解释,本文将算法的基本特点概括为:(1)处理大量信息;(2)基于规则或计算公式;(3)自动化;(4)可以满足预测目标。

算法信任指在数字经济背景下人与技术之间建立的信任,由人际信任推及人机信任。随着社会的发展,信任囊括的范围逐渐扩大。其意义从心理学定义的一种人格特质(Rotter,1971)^[18]和社会学定义的“认为值得信任的代理人特征”(Cook和Wall,1980)^[19]发展至人际关系层面(Lewis和Weigert,1985)^[20]。随后,Mayer等(1995)^[21]在组织信任模型中提出更广泛的定义:“一方愿意受到另一方行动的影响而不考虑监督或控制另一方的能力”。这一定义并没有将信任的概念局限于人与人之间的互动,允许考虑人工智能(包括算法)等技术方面的信任。学者们由此定义了人对自动化技术的信任(trust in automation),即“一种有意识或无意识的态度,认为计算机代理能够在具有不确定性和脆弱性的情况下帮助实现目标”(Hoff和Bashir,2015^[13];Lee和See,2004^[22])。

不同研究领域对信任的解释都包含三个共性:首先,必须有给予信任和接受信任的双方。算法信任关系中,消费者为给予信任的主体,受托人则是嵌入算法的各种应用。其次,受托人必须有某种执行这项任务的动机。对于算法,这种动机通常源于它们被设计来实现主体设定的特定目标。最后,情境中包含不确定性或风险。与此呼应的是,算法接受研究中,准确性与可解释性是影响消费者对算法推荐态度的两个关键维度(陈昌东和江若尘,2021)^[23],通过帮助消费者理解算法的决策过程来建立信任。通过以上分析,由于人类算法交互中包含人际关系和人对自动化技术信任的三要素,本文提出算法信任的概念。

算法信任与过去人际信任和人机信任存在的不同之处在于:算法为满足人类目标而设计,算法信任存在情境化、动态化的特点。现有的研究证据显示,在不同情境中,算法信任的水平不同,且受认知和情感加工过程影响。例如,在 Reich 等(2023)^[24]的实验中,参与者在短时间内由不信任算法产生的预测结果转变为接纳算法给出的答案。这种改变可以在有意识和无意识的层面发生(Hoff 和 Bashir,2015)^[13]。同时,Hoff 和 Bashir(2015)^[13]也指出,信任的形成是一个动态过程。当人们接触到新的信息时,信任感会发生巨大的变化。因此,在算法信任的维度中,整体信任的形成过程更加复杂,不仅需要考虑到人们静态的固有的信任,还应考虑随情境、交互过程变化的动态的信任,以及认知和情感加工过程的影响。

2. 算法信任的构成维度及内部关系

在明晰算法信任概念和动态性的基础上,本文接下来将算法信任作为一种态度和引发行为的动机,对其进行结构化的整合与分层。鉴于算法信任具有多方影响因素和静态与动态共同作用的内在复杂机制,本文借鉴 Marsh 和 Dibben(2003)^[25]从信息技术视角将信任划分的三个层次,即倾向性信任(dispositional trust)、情境信任(situational trust)和习得信任(learned trust),作为建立算法信任框架的基础(如图6所示)。其中,倾向性信任指个体具有信任或不信任的心理倾向或人格特质;情境信任强调情境线索的作用,信任倾向根据情境线索进行调整;习得信任是作为经验的结果。

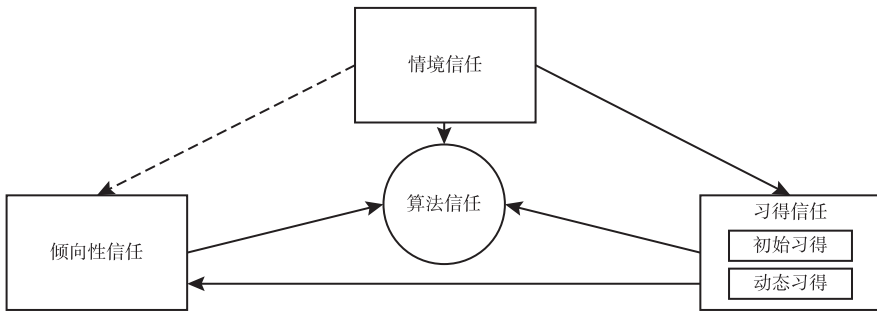


图6 算法信任的理论框架

注:箭头代表作用方向;虚线代表弱影响;实线代表直接影响;圆圈内心为核心变量;方框内为分层变量

本文构建算法信任三层框架有三个依据:首先,杜严勇(2022)^[26]在研究中呼吁探索算法互动与信任模式,并认同将 Hoff 和 Bashir(2015)^[13]提出的包含三层次的人与自动化信任框架延伸至算法系统互动过程。其次,自动化被定义为“主动选择数据、转换信息、做出决策或控制流程的技术”(Lee 和 See,2004)^[22]。而算法能够通过深度学习模型和大数据分析为自动化系统赋能。因此,可以将算法视为数字经济时代核心的自动化代理。最后,文献计量结果显示,在研究初期,许多算法建议任务中证明人们对算法的不信任倾向主要基于对人工智能的先验态度或信念(Dietvorst 等,2015^[5];Önköl 等,2009^[11])。这种态度和信念以人们已有的知识和经验为基础,体现了人们对算法的倾向性信任。而在之后的实证研究中,学者们通过引入干预因素,发现算法信任水平可以随外部或内部状态的改变而提升(Castelo 等,2019)^[1]。这种转变可以被纳入人们对算法的情境信任或动态的习得信任。近年来,习得信任被证明存在于无具体干预措施下(Logg 等,2019)^[6],可以在不断的学习和积累中形成倾向性信任,对后续行为持续产生影响(Chua 等,2023)^[27]。习得信任的这一转化过程进一步验证了算法态度的纵向动态变化。将上述三类算法信任进行层次结构化,能够清晰地展现出算法信任的初始状态、习得和演变过程。

这三层算法信任之间的逻辑关系体现为相互影响、共同作用。倾向性信任相对稳定,具有持久

性、可塑性低的特点。这一层次的算法信任有四个主要来源:启发式、人格、文化和人口统计学特征。在无情境线索干扰时,倾向性信任可以成为算法态度的决定性要素(Pearson等,2019)^[28]。情境信任可塑性强,不同情境下具有可变性,会受到消费者内部特征和算法、任务等外部因素的影响。这种信任受制于具体情境,虽然难以使得倾向性信任发生重大变化,但能够通过提高算法表现或帮助消费者积累经验,间接提高习得信任。习得信任产生于与特定算法交互的相关经验。广义上,习得信任可以被看作是过去情境信任集合所形成的经验(Marsh和Dibben,2003)^[25]。但考虑到人与算法的互动的即时关系,在交互过程中用户的信任程度可能会随着算法的实时性能而波动(Reich等,2023)^[24],研究者们将习得信任分为初始习得和动态习得两类。初始习得信任表示在与算法进行交互之前的信任,受到用户先验知识的影响。不同于倾向性信任,初始习得信任表现为不断积累提升的趋势。其与情境信任的区别在于引导信任的过去经验是与算法相关(习得信任),还是与环境相关(情境信任)。动态习得信任则反映了交互过程中信任的可变性,受到算法的整个相关系统的表现好坏的影响(Hoff和Bashir,2015)^[13]。积累一定程度的习得信任可以将用户的算法偏见倾向转变为算法欣赏倾向。

3. 算法信任的双重加工过程及整体理论框架

从驱动消费者行为的因素来看,已有研究从信任的两个维度区分了其中的解释机制:情感信任和认知信任(Komiak和Benbasat,2006)^[29]。认知信任指对代理人的绩效或可靠性的信心,而情感信任则主要基于个人的感受,并与代理人的表现相对独立(Johnson和Grayson,2005)^[30]。例如,在早期的自动化决策研究中,Madhavan和Wiegmann(2007)^[12]采用与感知可靠性相关的认知信任来解释用户对新手和专家的意见采纳程度差异。关于任务类型对算法态度的影响,Castelo等(2019)^[1]证明,面对客观任务,消费者由于对算法表现的有效性的信念(认知信任),从而愿意采纳算法结果;而在主观任务中,出于情感因素(如情感舒适),消费者更多采取人的建议而不是算法。类似地,消费者相信算法推荐更有能力评估功能属性价值,因此听从算法生成基于功能属性的推荐,而当激活享乐性目标时,消费者更有可能被情感信任驱使选择人工推荐(Longoni和Cian,2022)^[15]。这种解释机制的区分有利于从不同角度理解和干预消费者对算法的态度和行为选择。

基于以上分析,认知信任侧重于个体对算法性能和可靠性的评估,情感信任则更多地基于消费者个人的情感体验和情感连接。这分别对应双重加工模型假设中人们在决策过程中的两种加工方式:情感加工和理性加工(Shin,2021^[31];Shi等,2021^[32])。情感加工快速、直觉且基于联系,理性加工缓慢、经过深思熟虑且基于规则(Chaiken,1980^[33];Evans和Stanovich,2013^[34])。近年来,神经机制方面的研究也证实了人们在与算法的交互中对情感加工和理性加工两种加工方式的使用(Aw等,2023^[35];Yun等,2021^[36])。然而,现有研究缺乏对算法态度解释机制的系统性整理,尤其缺乏对两种加工方式的作用过程阐述。因此,本文聚焦于建构以双重加工模型为基础的算法信任理论框架。

在本文提出的算法信任系统中,虽然两种加工方式通常共同运作,但不同层次的信任对两种加工方式的依赖程度有所不同。倾向性信任主要依赖外围的情感加工路径,表现为基于心理特征或特定人格特质的情感信任。情境信任的内部环境因素通常基于当下个体的情感体验产生作用。其外部因素则可能因其性质的不同依赖于不同的加工方式。例如,算法的可解释性主要依赖分析理解的认知过程,因为当用户面对算法或与之交互时,需要采取类比或分析判断的方式进行信息加工(Reber,1989)^[37];面对不同任务和产品,消费者也经常通过认知处理判断是否信任算法。而一些外部环境因素也可以通过对内部因素产生影响,引发用户的情感反应。例如拟人化的算法界面设计会激发人们的内在联系感知,从而减弱对算法的外群体排斥(Wien和Peluso,2021)^[16]。习得信

任对加工方式的依赖差异体现在不同类别上。初始习得信任的形成由用户过去的经验和形成的算法素养决定,这一过程既包括认知处理也涉及情感处理。而动态习得信任产生于与算法的动态互动过程。这一过程中,用户主要依赖于认知加工过程对算法的表现如准确性、公平性等进行评估。不同层次的算法信任共同作用,最后整合影响消费者行为反应。

四、算法信任的影响因素及其机制

基于前文的算法信任理论框架,本文在界定算法信任三层关系的基础上,归纳总结过去研究中对个人因素、算法因素以及其他权变条件的讨论结果,利用双重加工理论提供机制解释,以搭建完整的算法信任模型(如图7所示)。本文以各因素非独立输入(Burton 和 Jensen, 2020)^[38],且三重信任相互影响为前提,深入分析了算法信任在非互动过程和动态互动过程的前置因素、加工机制及作用结果。

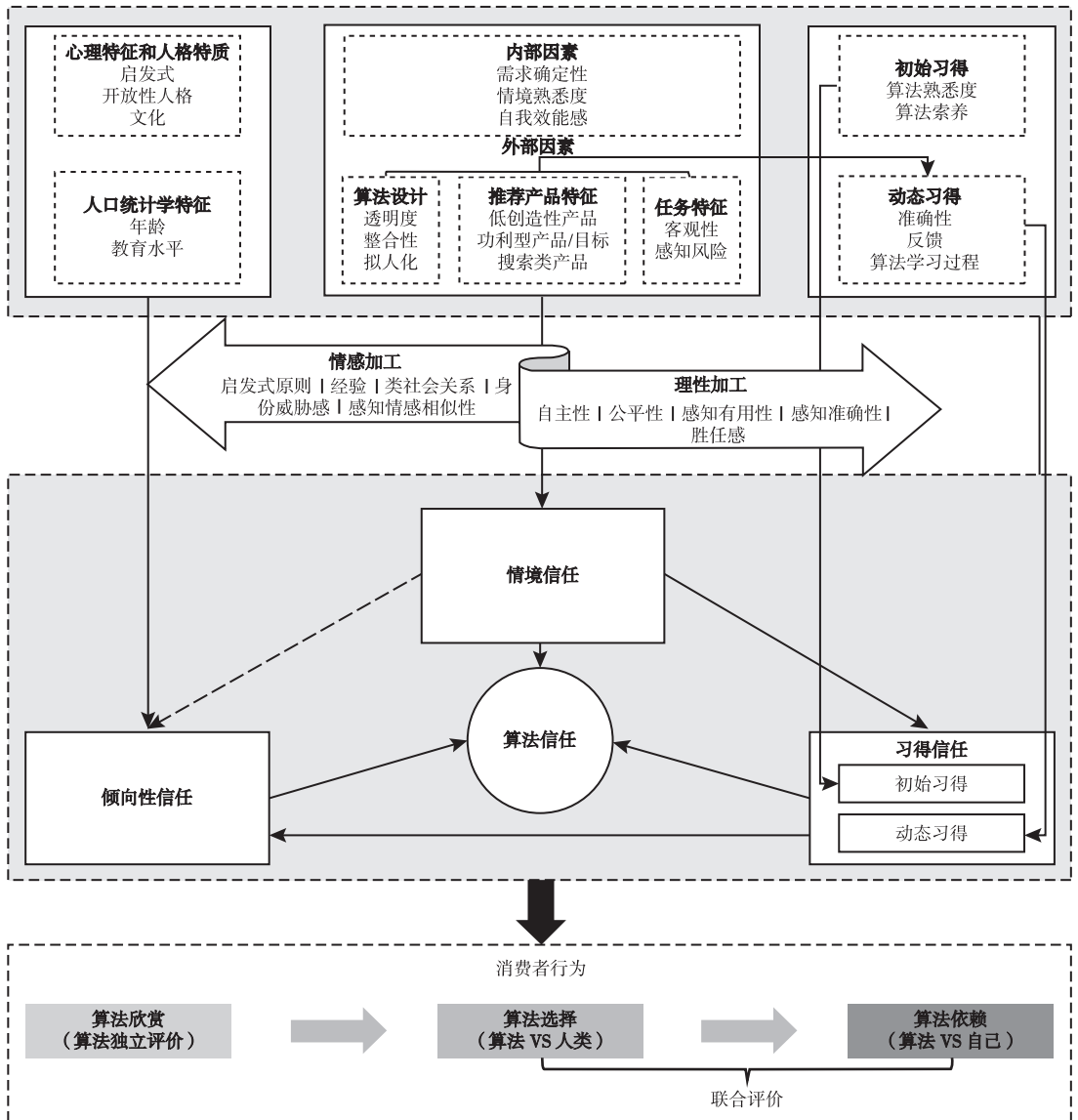


图7 消费者算法态度的影响因素及其机制

1. 倾向性信任

倾向性信任的输入因素为个体特征,代表个人信任算法的总体倾向,独立于环境且相对稳定。

(1)心理特征和人格特质。主要包括以下几个方面:

1)启发式(heuristic)。消费者对算法倾向性信任的一种表现为启发式的心理特征,即对算法感知特征相关线索的关注,包括来源可信度、相关讨论或一致性观点的数量等(Gunaratne等,2018)^[39]。Gunaratne等(2018)^[39]在一项模拟退休储蓄的实验中,发现算法提出的建议比众包(crowdsourced advice)提出的建议更有说服力,且参与者更愿意在算法的指导下承担风险(将更多的储蓄用来投资股票)。有趣的是,当将同样来自众包的参考信息被设计为源于算法时,参与者的信息采纳度比众包建议来源条件下更高,这反映了消费者面对算法时采用启发式线索。同样在金融决策研究中,对人工智能的初始态度直接影响参与者对算法推荐股票的信任度(Chua等,2023)^[27]。这种直觉的、自动加工的情感加工过程也反映了消费者对算法的启发式信任。

2)开放性。在众多的人格评估工具中,五大人格特质(Big Five Inventory, BFI)被广泛采用,包括:经验开放性(openness)、尽责性(conscientiousness)、外向性(extraversion)、亲和性(agreeableness)、神经质(neuroticism)(Roccas等,2002)^[40]。其中,经验开放性用于描述个体的创造性和对陌生情境的探索意愿。Oksanen等(2020)^[41]的研究结果表明,经验开放性是影响人工智能信任的一个重要预测因素。Sindermann等(2022)^[42]也揭示了中国受试者的开放性特质与人工智能态度之间的积极联系。Sharan和Romano(2020)^[43]对此现象进行了解释,指出具有更高水平外向性的人格对创新性技术的焦虑感更低,反应更快。这种由人格影响的倾向性信任涉及情感加工过程,即人们自主性地认为算法是可信的,从而持接受态度。

3)文化。文化维度是影响算法信任的重要因素。其中,Hofstede(1984)^[44]文化维度框架中的个人主义和集体主义经常被用来解释技术接受、信任和采用(Rau等,2009^[45];Liu等,2023^[46])。例如,相较于高个人主义的德国学生,中国学生更喜欢集体决策。因此,他们认为算法决策更加可信(Rau等,2009)^[45]。类似地,美国的个人主义维度更高,研究证据显示,当个性化被忽视时,美国人对算法的厌恶程度高于集体主义倾向的印度人(Liu等,2023)^[46]。

(2)人口统计学特征。主要包括以下几个方面:

1)年龄。年龄层面的差异导致消费者对算法技术持有不同倾向性的态度。基于一项大规模国际调查数据(N=53314),研究者发现年龄在算法推荐的个性化新闻的信任分布有显著差异(Thurman等,2019)^[47];相比人类编辑,年轻消费者更青睐算法的个性化推荐的新闻。Lourenço等(2020)^[48]在一项实证研究中证明,在其他条件相同的情况下,老年消费者对使用自动化工具生成养老金建议的满意度低于年轻消费者。导致这一情感性反应差异的原因则与老年人专业知识水平较低,无法理解算法决策这一认知过程显著相关。

2)教育水平。教育水平对算法信任的影响涉及情感和认知因素。从情感角度看,研究表明,当人们的教育水平较低,他们对算法的过程或结果感到不太舒服(Logg等,2019^[6];Thurman等,2019^[47])。从认知角度看,教育水平相对较高的群体对算法的高信任度是由对感知公平性和胜任感进行认知理解后的结果(Castelo,2023)^[49]。

2. 情境信任

虽然倾向性信任的输入因素能够影响消费者初始态度,但这种基本倾向会根据情境线索进行调整,即为信任三维框架中的情境信任。情境线索分为内部变量(受情境影响的个人心理状态)和外部环境因素。与倾向性信任涵盖的个人持久特征不同,情境信任的内部因素是在特定条件(如任务、特定类型的算法)下具有的,例如,对算法倾向性信任较高的消费者由于对某领域更加熟悉

(“专家型消费者”)而在相关情境下信任度降低。外部因素则分别从算法设计、营销产品分类和任务特点层面进行讨论。

(1)内部因素。主要包括以下三个方面:

1)需求确定性。在线购物场景中,即使在需求明确情况下,消费者也不会毫不犹豫地下单,因为在面对众多的选择时,他们可能需要算法提供准确的信息来帮助他们做出更好的决定。在此背景下,Zhu等(2022)^[50]模拟了一项在线实验,将被试随机分为两组:一组非常清楚自己计划购买的耳机品牌、类型和预算,另一组则需求不明确。结果表明,当需求确定性较高时,消费者更有可能选择算法推荐。这表明消费者信任算法提供准确信息。

2)情境熟悉度。目前的研究发现,在特定情境中,人们是否熟悉任务和对联合评价中的人类专家的熟悉度均显著影响算法信任,而人们对算法的熟悉程度则更多归为从经验中学习的过程(见下文对习得信任的总结)。与熟悉度相关的情境在医疗领域较为普遍,患者基于社会联系的情感加工,更熟悉人类医生作出医疗诊断,因此难以相信算法(Yun等,2021)^[36]。消费者更熟悉算法推荐情境因此信任度更高。例如,消费者认为由用户过去消费行为引导的算法选择优于编辑策划的新闻,从而选择算法的推荐(Thurman等,2019)^[47]。

3)自我效能感。接触算法时,自我效能感代表了人们对于自己使用这种技术的信心,其程度过低和过高均会导致消费者拒绝算法。前文在倾向性信任的影响因素中提及年龄较大的群体更难以信任算法的推荐。然而,该影响路径依据情境有所变化。例如,在一个药物管理系统的实验中,老年人比年轻人更依赖算法,因为老年人不太相信自己有能力做出正确的决定(Ho等,2005)^[51]。在线自我效能感衡量的是人们对网上个人信息和隐私的信念(Boerman等,2021)^[52]。调研和实验室数据结果证明在线自我效能感的情境变量正向促进了算法信任和后续使用(Araujo等,2020)^[17]。

(2)算法设计。主要包括以下三个方面:

1)透明度。算法的“黑箱效应”意味着用户无法了解或访问背后的推理过程和原理(Castelvecchi,2016)^[53]。因此,提高算法的透明度是增强用户对算法信任的关键。为了证明这一点,Yeomans等(2019)^[54]将算法与人工推荐系统进行比较,结果表明,虽然算法对笑话有趣性的预测更加精确,参与者倾向于拒绝算法建议。而在阅读了算法如何工作的解释后,参与者接受算法帮助的比例更高。此外,Cadario等(2021)^[14]在一项实地实验(谷歌广告)中利用广告点击率也验证了在解释算法决策过程后的信任转变。以上研究中,透明度的干预通过提高主观理解性来积极影响算法信任。

2)整合性。整合人类参与建议能够提高人们对算法的信任感。一种解释机制为数字技术的非人类特质容易使人们感到身份威胁(Zlotowski等,2017)^[55]。当在任务中用算法取代人类时,人们对算法产生更强烈的不信任感(Prahl和Van Swol,2021)^[56]。与这一情感加工过程不同,参与决策过程的自主性可以让人们相对更理智地看待算法的功能,从而提高信任。例如,当人们可以对算法预测的结果进行有限修改时,他们信任算法答案的次数更多(Dietvorst等,2018)^[57]。这种整合性也可以从算法扮演角色的角度来分析,即算法在决策中扮演支持型角色还是竞争型角色。将算法设计为辅助型决策者而非人类决策的替代者时,人们对算法的信任感可能会得到提高(Longoni和Cian,2022)^[15]。

3)拟人化。拟人化设计对消费者信任态度的影响尚未达成一致。一方面,具有类人特征的算法可以使人更容易与之建立类社会关系,激发情感联系(黄敏学和吕林祥,2022)^[8],从而增加对算法的接受度(Wien和Peluso,2021^[16];Aw等,2023^[35];Qiu和Benbasat,2009^[58])。然而,另一方面,有学者强调这种拟人化的设计也会引起“恐怖谷”(uncanny valley)的反作用(Arsenyan和

Mirowska, 2021)^[59]。为了缓解拟人化可能的负面作用,学者从情感处理的另一个角度出发:感知情感相似性。即强调算法具备人的情感能力而不仅仅在外观上贴近人类,这可以有效增加算法在主观任务中的使用(Castelo等, 2019)^[11]。类似地,最近关于人工智能的研究认为,与机械型AI和思考型AI相比,感觉型AI通过激发顾客的积极情绪加强信任从而增加消费(Schepers等, 2022)^[60]。由此推断,为了提高算法的情境信任,感知关系亲密度和感知情感能力均可以考虑融入拟人化设计中。

(3)推荐产品特征。在营销领域,算法主要以个性化推荐和嵌入式聊天机器人服务消费者。因此,产品相关的前置变量对算法情境信任的影响十分关键。其机制是消费者基于算法对特定类型的产品的推荐或介绍能力的考量,对能否满足自我需求进行判断,进而产生不同态度。

1)低创造性产品。出于对算法胜任力的感知,消费者更偏好由算法推荐的创造性较低、象征性价值少的产品(吴继飞等, 2020)^[61]。除了对算法推荐能力的认知分析,独特性动机降低了消费者对机械化产品的偏好:当消费者在购买有较高(相对于较低)象征价值的产品或在象征性消费语境中(例如,当表达个人的信仰和个性更重要时),他们更相信人类的劳动产出(Granulo等, 2021)^[62]。

2)功利型产品/目标(VS享乐)。近年来研究结合享乐消费(hedonic consumption)和功利消费(utilitarian consumption)属性和算法推荐,发现了算法推荐对功利型产品或功利目标激活时的有效性,并认为这是感知能力的认知处理结果(Longoni和Cian, 2022)^[15]。Wien和Peluso(2021)^[16]的实验结果则表明,算法推荐功利型产品和人工推荐同等有效。类似于推荐机制,在电商客服交流语境中,当强调的产品属性是功能性的时,人工智能比人类一线客服获得了更高的客户满意度(Ruan和Mezei, 2022)^[63]。

3)搜索类产品(VS经验)。搜索类产品是指在购买前容易通过一些客观标准对其质量进行评估的商品或服务(如行李箱、耳机),经验类产品是指只有在尝试或购买后才能很好地评价其质量的商品或服务(如旅行、餐饮)(Franke等, 2004)^[64]。在消费者需求明确的前提下,对算法推荐搜索类产品的接受度更高(Zhu等, 2022)^[50]。功利主义和明确搜索消费均为认知驱动,即基于功能和工具目标,并根据达到目的的程度进行评估。因此,基于算法胜任力的考量可以很好地揭示这一背景下的算法信任现象。

(4)任务特征。主要包括以下两个方面:

1)客观性。客观任务主要包括可量化和可衡量的事实,而主观任务则是开放的解释和基于个人意见或直觉的任务(Castelo等, 2019)^[11]。在任务客观性较高时,人们对算法的绩效或可靠性更有信心,对应表现出更高的认知信任(Johnson和Grayson, 2005)^[30]。这种感知有效性甚至能够抵消部分情感加工造成的不适感。因此,在实验干预中,Castelo等(2019)^[11]通过突出主观任务的客观性,增加了受试者对算法的整体信任。跟随该研究,Logg等(2019)^[6]和Reich等(2023)^[24]在不同实验任务类型中,也验证了受试者在参与客观性任务时对算法持有信任态度。

2)感知风险。感知风险是指消费者在决策过程中对风险程度的主观认知(Dietvorst和Bharti, 2020)^[2]。Chua等(2023)^[27]通过一组由两种高低风险的股票构成的实验,证明当不确定性较低时,倾向性的积极态度足以促使参与者接受自动化建议。然而,当决策的不确定性较高时,只有高水平的信任才能促使参与者接受算法的建议。这表明,在高风险情况下,人们更加依赖于信任感来做出决策。此外,风险的确定性程度也对算法信任产生影响。在确定的高风险情境下(如结果分布已知且相同的高风险彩票选择),消费者对算法计算结果的信任程度增加(Farjam, 2019)^[65]。

3. 习得信任

习得信任基于与特定算法(或当前交互)相关的过去经验,可分为初始的学习和交互过程中的动态学习。初始习得信任不体现在当下的互动中,而是基于过去的经验或学习,连同前文阐释的倾向性信任和情境信任,是经消费者情感或认知反应的交互前过程。而动态习得是一个交互过程。

(1)初始习得。主要包括以下两个方面:

1)算法熟悉度。算法熟悉度是指个体对执行任务的算法积累的经验(Castelo等,2019)^[1],通过时间长度、使用频率或体验满意度而建立。与具体任务中的情境熟悉度不同,这种熟悉过程与算法紧密相关而非环境。根据自我认知经验理论(Epstein,1994)^[66],基于过去经验的经验系统属于情感驱动过程。人类根据自己的内在感情建立熟悉度(Monin,2003)^[67]。在人工智能营销领域,机器学习算法以消费者的评论、先前的购买经历和产品使用情况进行个性化匹配,消费者基于过去良好的体验信任算法的个性化推荐(Gai和Klesse,2019)^[68]。

2)算法素养。与人口统计特征中的教育程度因素不同,算法素养(algorithmic literacy)侧重于人们过去接触并掌握的人工智能等计算机知识。人类决策者的算法素养有利于合理化用户对算法结果的期望,从而促进算法辅助的合理利用(Burton和Jensen,2020)^[38]。算法素养较低的人由于相对缺乏算法互动、统计输出解释的知识,容易以外显思维过程低估算法表现,因而表现为低信任感。认知加工可能会带来负向影响。例如,Önkal等(2009)^[11]的实验中,参与者为预测和经济相关领域的学生,他们的算法素养更高,因此基于过去掌握的知识理解统计预测的错误,表现出算法厌恶。然而,Prahl和Van Swol(2017)^[69]通过类似的实验设计没有复制该结论,作者推测的原因:招募的被试没有学习预测模型。

(2)动态习得:算法表现。动态习得与交互过程中的算法表现密切相关,消费者采取认知加工理念,权衡算法表现的情况,进而依据信任程度高低做出决策。

1)准确性。算法犯错误的概率受到情境因素的影响,如预测任务的不确定性使得用户更有可能在交互过程中收到不准确答案。基于信任视角,人们通常对于精算系统持有完美的渴望,因此,算法是否能够完美完成任务影响了参与者的信任轨迹(Glikson和Woolley,2020)^[70]。例如,参与者在—项预测任务中,在算法出错之前倾向于信任算法的预测(Prahl和Van Swol,2017)^[69],但当他们收到“糟糕”的建议之后,来自计算机的建议的利用率比人类建议的利用率下降更加明显(Dietvorst等,2015^[5];Prahl和Van Swol,2017^[69])。

2)反馈。研究发现,在消费者使用算法过程中提供反馈是有效的干预措施。在—项实证研究中,实验对象被要求在40轮中的每一轮预测—次股价的上涨或下跌趋势。每一轮预测后提供反馈是一个不断评估自我能力和算法表现并建立信任的过程(Filiz等,2021)^[71]。进一步地,校准反馈是根据算法的反馈修改输入数据,以便他们可以通过试错过程验证决策(Van Dongen和Van Maanen,2013)^[72]。这个互动过程可以反复进行,直到个人满意。反馈过程结合了前文中提到的算法设计因素——整合性,可以为人们带来自主性并减少过度信任,在动态过程中达到信任校准的目的。

3)算法学习过程。展示算法准确度提升的学习证据能够有效缓解对算法的厌恶。过去证据表明,算法犯错误导致厌恶(Dietvorst等,2015)^[5],但算法的机器学习能力已大大改善这种不准确现象。为了改变这种错误偏见,Reich等(2023)^[24]首先进行了两组实验,发现是否能从错误中学习是影响算法信任的主要因素。因此,他们在第三项研究中引入了—个关键的变量:学习证据,即展示算法动态的提升过程。结果发现,当提供算法不断学习的证据时,算法被视为具有与人类同等的学习能力,从而获得了与人类相当的信任水平。

五、算法信任与用户行为表现

1. 信任驱动的用户行为的三种行为表现

基于动态视角,算法技术的进步和有效条件下消费者信任度的提高,驱动了消费者行为的转变。Pearson 等(2019)^[28]在一个双重顾问决策任务中提出了信任偏见(trust bias),且证明了信任对依赖决策行为具有高度的预测性。因此,本文以信任的三个维度为基础,提出三种信任相互作用构建整体维度的算法信任,对消费者行为产生影响。消费者不仅在心理上信任算法,还因信任程度的加深对应不同层次的行为:算法欣赏、算法选择和算法依赖。具体而言,随着整体信任的加强,消费者先是在独立评价下青睐算法,进而在有选择的情况下使用算法,最后在自主选择前提下依赖算法。

(1)算法欣赏。算法欣赏指在独立评价机制的消费者表现,即在仅提供算法服务时对算法的表现持积极态度或正向评估的行为。倾向性信任和情境信任在用户行为之间的潜在联系具有双重作用。当用户对算法的情感信任占主导地位时,首先会对算法本身持有正向评价。例如,用户对拟人化算法的不适感降低,并主动接受(吴继飞等,2020^[61];Hildebrand 和 Bergner,2020^[73]),甚至相较于社会证明(social proof),更容易被专业算法的权威性启发式说服(Gunaratne 等,2018)^[39]。在 You 等(2022)^[74]的研究中,无论算法表现如何,信任均是用户欣赏并采纳算法建议的重要条件。情境条件下,通过认知处理的算法信任则主要基于人们对算法能力和回报结果的判断。人们欣赏算法高水平的客观分析能力,认为算法能够提供功利性价值(Logg 等,2019)^[6]。

(2)算法选择。算法选择指当算法和人工同时可供选择时,人们经比较后更容易信任和接受算法的建议,从而选择依从于算法的决策。相较于算法欣赏,算法选择需要更高层次的算法信任。只有在算法的表现明显优于人类时,人们才可能会选择算法决策。而低程度的信任会引起算法拒绝。例如,在 Dietvorst 等(2015)^[5]的实验中,被试若预测准确则可以获得现金奖励,在看到算法的表现(犯了一些错误)后,绝大多数(74%)的参与者选择了人类预测者,即使他们也观察到算法在整体上比人类预测者更准确。目前的研究主要关注情境因素,以证实算法信任的构建对于消费者选择算法的重要性。例如,在内容推荐、预测、估计等决策场景中,用户对算法的信任构建会促使他们坚持选择算法(Logg 等,2019^[6];Longoni 和 Cian,2022^[15];Thurman 等,2019^[47])。

(3)算法依赖。当用户容易接受计算机的推荐,而不是依赖他们自己的判断时,就会出现自动化偏见(automated bias)(Chua 等,2023)^[27]。算法依赖也可以理解为一种信任偏见,指当可以自己做决策时转而寻求算法帮助的遵从行为。算法依赖最早体现在人们对于专家的依赖(Dijkstra 等,1998)^[75]。随着对算法态度向行为的延伸研究,学者们发现将算法建议和自己可控的决策相结合时,也难以改变这种单纯对算法的偏好(Logg 等,2019)^[6]。例如,在模拟退休储蓄的任务中,虽然算法计算不透明(除了将其作为研究的结果框架之外,没有信息提供给用户),但受试者对算法建议的采纳率高于众包组(Gunaratne 等,2018)^[39],这体现了算法信任导致的依赖行为。

2. 用户行为测量方式

在实际情境中,算法信任导致的用户行为表现可以进行不同的操作化定义。具体而言,算法欣赏的测量以结果评价和代理人评价为主,算法选择的测量以人类和算法的两级偏好评价为主,算法依赖的测量以听从意愿和建议采纳为主。表2列示了近年来文献中对不同因变量(决策、判断、评价等行为)的界定和操作化定义。这些测量方式为未来算法信任的相关研究提供了测量方式和测量效果的参考。

表 2 算法信任与行为表现通过实验因变量的测量方法

行为表现	测量方法	研究者和年份	具体表述
算法欣赏	结果评价	Efendić等(2020) ^[76]	预测结果的准确性(-3=非常不准确,3=非常准确)
		Jago(2019) ^[77]	以7分量表评价任务结果的真实性
		Longoni等(2019) ^[78]	切换服务商的支付意愿
	代理人评价	Longoni等(2019) ^[78] Longoni和Cian(2022) ^[15]	对服务/推荐提供者的喜爱度(1=肯定是X(人类),7=肯定是Y(算法))
算法选择	人类和算法的 两极偏好评价	Castelo等(2019) ^[1]	使用算法还是使用人类来完成每项任务的意愿(0=人类,50=无偏好,100=算法)
		Logg等(2019) ^[6]	选择参考算法还是另一个参与者的建议
		Reich等(2023) ^[24]	选择算法还是专家服务
		Longoni和Cian(2022) ^[15]	以7分量表测量意见遵循意愿
算法依赖	听从意愿	Sharan和Romano(2020) ^[43]	建议采纳的反应时间
		Gunaratne等(2018) ^[39]	计算参与者对其最初估计的修改程度。
	建议采纳	Sharan和Romano(2020) ^[43]	答案中与算法建议相同的数量
		Efendić等(2020) ^[76]	将预测作为自己的答案的可能性(-3=非常不可能,3=非常可能)
		Logg等(2019) ^[6] Önkál等(2009) ^[11] Prahł和Van Swol(2017) ^[69] You等(2022) ^[74]	建议利用率:“每个参与者产生的最终估计数和初始估计数之间的差”除以“建议和初始估计数之间的差”

六、结论和未来研究展望

已有研究证实了部分有效干预条件下,人们表现出算法信任,即特定环境下的情境信任。另一方面,大数据调研以及实验类型研究均发现了人们对算法的偏好,这种信任可以归于倾向性信任或习得信任。但一些研究未对其中的因果关系或机制予以澄清,需要进一步探索和验证。人们在不同的算法使用情境中会因其对算法的初始态度或信念、综合环境因素和与算法的交互过程,显示出动态变化的算法信任。不仅如此,算法信任对用户的影响可以体现在决策、判断和评价等行为表现上,每种行为表现又可以进行不同的操作化定义。因此,在设计算法推荐系统或者制定市场营销策略时,需要综合考虑三个层次算法信任的影响因素及其对不同加工方式的选取,并设计合理有效的算法影响评价指标。本文在上述相关研究的基础上,基于信任视角构建了消费者的算法态度及行为框架,采用了分层方法来概念化影响信任的各种因素和机制。该理论框架可以作为对复杂的算法信任和用户积极行为反应的研究指南,为促进人工智能技术的可持续发展提供参考。

目前,业界采用的算法随着信息技术的发展进一步成熟,算法将通过各种形式越来越多地成为更多人日常生活的一部分。在算法应用中促进适当的信任是提高消费者满意度和企业生产力的关键。本文认为,未来算法信任的应用还需要更高层面的各方管理,继续完善基于算法信任的智慧型治理方案。基于以上研究评析,本文提出了以下几个值得进一步研究的问题,包括研究现状、未来研究方向和建议的研究方法。

1. 进一步探索算法推荐下产品类型、产品与消费者关系对算法态度的影响

文献可视化结果显示,在人机关系探究阶段,算法可作为推荐机制对消费者行为和决策产生影响。随后,在算法态度探究阶段,推荐系统的效果受到广泛关注并融入在实验场景设计中。市场实践中,随着算法推荐系统效率的提高,人们越来越依赖算法推荐的商品、服务或内容来做出决策(Puntoni等,2021)^[79]。这些决策会对消费者和市场产生深远的影响。从推荐的实践应用出发,未来研究可以着重探索以下几个方面:第一,更深入地研究算法推荐对于不同类型产品与消费者自我联系的影响。目前大部分研究集中在功利型产品、搜索型产品和象征性产品上,而其他产品属性和消费语境值得进一步思考,例如,小众品牌(Hoskins,2020)^[80]。深入研究算法推荐对不同产品类型的态度和行为变化,以及这种影响对市场和经济产生的潜在影响。第二,随着人们对算法的信任不断增加,越来越多的消费者愿意将自己的投资和理财决策交给机器顾问处理。这引发了对于产品所有权和决策权的重新思考。未来的研究可以从自主性角度探讨消费者将所有权和决策权交由算法代理时的心理状态和行为变化。第三,个性化算法推荐和消费者身份验证之间的关系变得更加紧密(陈昌东和江若尘,2021)^[23]。未来的研究可以探讨个性化算法推荐是否会影响到自我验证的准确性,以及如何在保护用户隐私的同时提供个性化算法推荐服务。

2. 进一步探究动态交互过程中算法信任的形成

前文基于文献计量分析和人机信任理论依据提出了算法信任的三维框架,并在进一步分析中囊括了在交互过程变化的信任感作用于整体信任的形成过程,如在总结习得因素时引入动态过程来区分于消费者个人或情境影响的静态过程。基于前文分析,目前大多数研究都是基于非交互的前提进行的,很少探索动态交互过程中算法信任的形成。未来还需要进一步弥补该研究空缺。研究可以从用户和算法之间的交互和反馈出发,探究动态习得信任的形成和加强。针对该研究问题,学者可以借助神经科学和心理学的理论和方法,探究动态交互过程中的信任形成机制。例如,已有研究借助眼动仪分析了视觉注意如何分配到决策的不同特征上的差异(Fisher,2021)^[81]。除此之外,在算法辅助场景下,人们决策的过程还可以结合神经机制进行测量来展示动态过程(Yun等,2021)^[36]。已有研究模拟领导小组的社会情境,发现拥有权力的算法技术比人类更主导性地改变了参与者决策(Hou等,2023)^[82]。未来研究可以考虑在多人协作或竞争的情境中,如何有效地传递算法信息和增加算法信任,以及其他参与者的动态加入如何影响到对算法的评价。学者可以使用神经影像技术,如fMRI、EEG等,来记录决策过程中的神经活动,以了解算法信任的加强机制。

3. 进一步探索存在人类影响源的情境下,社会关系对算法趋避倾向的影响

以算法信任作为整体性维度,前文在考察信任导致的行为意向时,发现算法趋避倾向表现出个体间的不一致性。现有文献在测量这种趋避倾向时通常采用对比方式:将人类和算法顾问、人类专家和算法顾问、个体自身和算法进行比较。为了进一步比较和理解这种倾向,本文鼓励更多的研究关注其他社会关系和表达方式,主要包括以下方面:第一,未来的研究可以考虑比较用户对内群体和外群体的建议提供者是否会与算法产生的建议有明显差异。第二,人类顾问与个体之间的亲密度和感知相似性可能会影响顾问的影响力和可信度。在算法信任研究中,这两个因素可能会从感知情感相似性角度影响到算法态度(Castelo等,2019)^[1]。如果一个人与某人类顾问有着亲密的关系,或感知自己与该顾问有许多相似之处,个体更可能认为顾问对自己有利并且值得信赖,从而轻视算法的作用。未来的研究可以通过设计实验,改变个体与顾问之间的亲密度和相似性水平,然后观察个体对顾问建议的态度变化,或采用社会网络分析(social network analysis)对不同节点(node)和边(tie)的性质进行区分,比较这种网络结构的差异。第三,社会关系中的输出方式也可能会影响到算法趋避倾向,例如算法口头表达的方式,可能会产生与文字表达不同的效应。如今交互式人工智能大模型已经支持语音输入输出,未来研究可以借助这种新技术进行实验设计,以探究不同表达

方式可能带来的算法趋避倾向的不同表现。

4. 从群体角度深入探讨算法在企业实践中的信任行为

前文综合已有文献发现, 针对算法倾向性信任和习得信任的前置因素的研究主要聚焦在个人层面。用户在不同情境中的信任程度也主要基于个人与算法(或算法执行的任务)之间的关系。组织和企业作为算法应用的主体, 其内部群体对算法的信任行为的研究还不够充分。为了更全面地理解信任与算法之间的关系, 未来的研究可以从群体角度深入探讨算法的信任行为, 并参考算法信任三维框架中的算法因素和任务因素, 塑造有利于情境信任的组织环境, 并在时间维度上动态提升群体整体性的习得信任。数据取得方面, 未来研究可以通过问卷、深度访谈、网络调查等方式, 从不同层次和维度上, 挖掘企业和组织对算法的信任结果和影响因素。例如, 为了提高算法在营销部门的应用效率, 已有研究借助 16 个董事会对高级营销经理的采访寻求减轻算法偏见的战略措施 (Plangger 等, 2022)^[83]。未来研究还可以进一步拓展到其他领域, 如在线问诊、职业教育等, 探究算法在不同领域中的应用效果和影响, 并提出相应的管理启示和政策建议。

5. 进一步研究跨文化和跨区域下的算法信任

根据前文分析结果, 文化维度是影响算法倾向性信任的重要因素。已有研究在个人主义和集体主义方面得到了有价值的结论。但现有研究多集中于单一文化或区域, 很少有研究跨越不同文化和商业背景进行比较。不同文化和商业背景可能会导致人们对算法信任产生不同的态度和行为 (Rau 等, 2009^[45]; Liu 等, 2023^[46])。通过深入研究不同文化和商业环境中人们对算法的信任和不信的表现, 可以发现一些共性和差异性。这将有助于更好地理解人们对算法的信任和不信的形成机制, 为中外企业制定跨国和跨文化的商业策略提供指导。具体而言, 跨文化和跨区域下的算法信任研究可以从以下几个方面展开: 第一, 未来研究可以比较不同文化或区域下人们对算法信任的水平和形成机制。除了已有对个人主义与集体主义文化差异预测不同程度算法厌恶的研究 (Liu 等, 2023)^[46], 还可以探讨其他文化维度, 如不确定性规避与算法风险决策之间的关系 (Chua 等, 2023)^[27]。或通过探究文化因素如何影响对算法的准确性或公正性的感知 (Grove 等, 2000)^[9], 从而预测消费者行为倾向。第二, 未来研究可以通过元分析探讨不同文化或区域下人们对特定类型算法的信任程度。例如, 有学者发现人工智能驱动的经济活动在美国各城市之间存在显著的异质性 (Makridis 和 Mishra, 2022)^[84]。第三, 未来研究可以分析不同文化或区域下人们对算法信任的变化趋势。随着人工智能技术的发展和运用, 智能化营销及管理领域均需要跨文化和跨区域地比较人们对算法信任的变化趋势以有针对性的更新营销管理策略。

参考文献

- [1] Castelo, N., M. W. Bos, and D. R. Lehmann. Task-Dependent Algorithm Aversion [J]. *Journal of Marketing Research*, 2019, 56, (5): 809 - 825.
- [2] Dietvorst, B. J., and S. Bharti. People Reject Algorithms in Uncertain Decision Domains Because They Have Diminishing Sensitivity to Forecasting Error [J]. *Psychological Science*, 2020, 31, (10): 1302 - 1314.
- [3] Kasneci, E., K. Sessler, S. Küchemann, et al. Chatgpt for Good? On Opportunities and Challenges of Large Language Models for Education [J]. *Learning and Individual Differences*, 2023, 103, 102274.
- [4] Kumar, V., B. Rajan, R. Venkatesan, and J. Lecinski. Understanding the Role of Artificial Intelligence in Personalized Engagement Marketing [J]. *California Management Review*, 2019, 61, (4): 135 - 155.
- [5] Dietvorst, B. J., J. P. Simmons, and C. Massey. Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err [J]. *Journal of Experimental Psychology General*, 2015, 144, (1): 114 - 126.
- [6] Logg, J. M., J. A. Minson, and D. A. Moore. Algorithm Appreciation: People Prefer Algorithmic to Human Judgment [J]. *Organizational Behavior and Human Decision Processes*, 2019, 151: 90 - 103.
- [7] Chen, C. Citespace Ii: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature [J]. *Journal of the*

American Society for Information Science and Technology,2006,57,(3):359-377.

[8]黄敏学,吕林祥.心理契合视角下智能产品营销研究的评述与展望[J].北京:经济管理,2022,(7):193-208.

[9]Grove,W.M.,D.H.Zald,B.S.Lebow,et al.Clinical Versus Mechanical Prediction:A Meta-Analysis[J].Psychological Assessment,2000,12,(1):19-30.

[10]Dzindolet,M.T.,L.G.Pierce,H.P.Beck,and L.A.Dawe.The Perceived Utility of Human and Automated Aids in a Visual Detection Task[J].Human Factors,2002,44,(1):79-94.

[11]Önkal,D.,P.Goodwin,M.Thomson,et al.The Relative Influence of Advice from Human Experts and Statistical Methods on Forecast Adjustments[J].Journal of Behavioral Decision Making,2009,22,(4):390-409.

[12]Madhavan,P.,and D.A.Wiegmann.Effects of Information Source,Pedigree,and Reliability on Operator Interaction with Decision Support Systems[J].Human Factors; The Journal of the Human Factors and Ergonomics Society,2007,49,(5):773-785.

[13]Hoff,K.,and M.Bashir.Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust[J].Human Factors The Journal of the Human Factors and Ergonomics Society,2015,57:407-434.

[14]Cadario,R.,C.Longoni,and C.K.Morewedge.Understanding,Explaining,and Utilizing Medical Artificial Intelligence[J].Nature Human Behaviour,2021,5,(12):1636-1642.

[15]Longoni,C.,and L.Cian.Artificial Intelligence in Utilitarian Vs. Hedonic Contexts: The “Word-of-Machine” Effect[J].Journal of Marketing,2022,86,(1):91-108.

[16]Wien,A.H.,and A.M.Peluso.Influence of Human Versus AI Recommenders: The Roles of Product Type and Cognitive Processes[J].Journal of Business Research,2021,137:13-27.

[17]Araujo,T.,N.Helberger,S.Kruikemeier,and C.H.De Vreese.In AI We Trust? Perceptions About Automated Decision-Making by Artificial Intelligence[J].AI & Society,2020,35,(3):611-623.

[18]Rotter,J.B.Generalized Expectancies for Interpersonal Trust[J].American Psychologist,1971,26,(5):443.

[19]Cook,J.,and T.Wall.New Work Attitude Measures of Trust,Organizational Commitment and Personal Need Non - Fulfilment [J].Journal of Occupational Psychology,1980,53,(1):39-52.

[20]Lewis,J.D.,and A.Weigert.Trust as a Social Reality[J].Social Forces,1985,63,(4):967-985.

[21]Mayer,R.C.,J.H.Davis,and F.D.Schoorman.An Integrative Model of Organizational Trust[J].Academy of Management Review,1995,20,(3):709-734.

[22]Lee,J.D.,and K.A.See.Trust in Automation: Designing for Appropriate Reliance[J].Human Factors; The Journal of the Human Factors and Ergonomics Society,2004,46,(1):50-80.

[23]陈昌东,江若尘.营销领域中算法推荐与消费者响应:研究评述与展望[J].北京:经济管理,2021,(10):193-208.

[24]Reich,T.,A.Kaju,and S.J.Maglio.How to Overcome Algorithm Aversion: Learning from Mistakes[J].Journal of Consumer Psychology,2023,33,(2):285-302.

[25]Marsh,S.,and M.R.Dibben.The Role of Trust in Information Science and Technology[J].Annual Review of Information Science and Technology (ARIST),2003,37:465-498.

[26]杜严勇.厌恶算法还是欣赏算法?——人工智能时代的算法认知差异与算法信任建构[J].上海:哲学分析,2022,(3):151-165.

[27]Chua,A.Y.K.,A.Pal,and S.Banerjee.AI-Enabled Investment Advice: Will Users Buy It? [J].Computers in Human Behavior, 2023,138,107481.

[28]Pearson,C.J.,M.Geden,and C.B.Mayhorn.Who’s the Real Expert Here? Pedigree’s Unique Bias on Trust between Human and Automated Advisers[J].Applied Ergonomics,2019,81,102907.

[29]Koniak,S.Y.X.,and I.Benbasat.The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents[J].MIS Quarterly,2006,30,(4):941-960.

[30]Johnson,D.,and K.Grayson.Cognitive and Affective Trust in Service Relationships[J].Journal of Business Research,2005,58,(4):500-507.

[31]Shin,D.The Effects of Explainability and Causability on Perception,Trust,and Acceptance: Implications for Explainable AI[J].International Journal of Human-Computer Studies,2021,146,102551.

[32]Shi,S.,Y.Gong,and D.Gursoy.Antecedents of Trust and Adoption Intention toward Artificially Intelligent Recommendation Systems in Travel Planning: A Heuristic-Systematic Model[J].Journal of Travel Research,2021,60,(8):1714-1734.

[33]Chaiken,S.Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion[J].Journal of Personality and Social Psychology,1980,39,(5):752-766.

- [34] Evans, J. S. B. T. , and K. E. Stanovich. Dual-Process Theories of Higher Cognition: Advancing the Debate[J]. *Perspectives on Psychological Science*,2013,8,(3):223–241.
- [35] Aw, E. C. – X. , T. Zha, and S. H. – W. Chuah. My New Financial Companion! Non-Linear Understanding of Robo-Advisory Service Acceptance[J]. *The Service Industries Journal*,2023,43,(3–4):185–212.
- [36] Yun, J. H. , E. J. Lee, and D. H. Kim. Behavioral and Neural Evidence on Consumer Responses to Human Doctors and Medical Artificial Intelligence[J]. *Psychology & Marketing*,2021,38,(4):610–625.
- [37] Reber, A. S. Implicit Learning and Tacit Knowledge [J]. *Journal of Experimental Psychology: General*, 1989, 118, (3): 219–235.
- [38] Burton, J. W. , and T. B. Jensen. A Systematic Review of Algorithm Aversion in Augmented Decision Making [J]. *Journal of Behavioral Decision Making*,2020,33,(2):220–239.
- [39] Gunaratne, J. , L. Zalmanson, and O. Nov. The Persuasive Power of Algorithmic and Crowdsourced Advice [J]. *Journal of Management Information Systems*,2018,35,(4):1092–1120.
- [40] Roccas, S. , L. Sagiv, S. H. Schwartz, and A. Knafo. The Big Five Personality Factors and Personal Values[J]. *Personality and Social Psychology Bulletin*,2002,28,(6):789–801.
- [41] Oksanen, A. , N. Savela, R. Latikka, and A. Koivula. Trust toward Robots and Artificial Intelligence: An Experimental Approach to Human-Technology Interactions Online[J]. *Frontiers in Psychology*,2020,11,568256.
- [42] Sindermann, C. , H. Yang, J. D. Elhai, et al. Acceptance and Fear of Artificial Intelligence: Associations with Personality in a German and a Chinese Sample[J]. *Discover Psychology*,2022,2,(1):8.
- [43] Sharan, N. N. , and D. M. Romano. The Effects of Personality and Locus of Control on Trust in Humans Versus Artificial Intelligence[J]. *Heliyon*,2020,6,(8),e04572.
- [44] Hofstede, G. *Culture’s Consequences: International Differences in Work-Related Values*[M]. London: sage, 1984.
- [45] Rau, P. L. P. , Y. Li, and D. Li. Effects of Communication Style and Culture on Ability to Accept Recommendations from Robots [J]. *Computers in Human Behavior*,2009,25,(2):587–595.
- [46] Liu, N. T. Y. , S. N. Kirshner, and E. T. K. Lim. Is Algorithm Aversion Weird? A Cross-Country Comparison of Individual-Differences and Algorithm Aversion[J]. *Journal of Retailing and Consumer Services*,2023,72,103259.
- [47] Thurman, N. , J. Moeller, N. Helberger, and D. Trilling. My Friends, Editors, Algorithms, and I: Examining Audience Attitudes to News Selection[J]. *Digital Journalism*,2019,7,(4):447–469.
- [48] Lourenço, C. J. S. , B. G. C. Dellaert, and B. Donkers. Whose Algorithm Says So: The Relationships between Type of Firm, Perceptions of Trust and Expertise, and the Acceptance of Financial Robo-Advice[J]. *Journal of Interactive Marketing*,2020,49:107–124.
- [49] Castelo, N. Perceived Corruption Reduces Algorithm Aversion[J]. *Journal of Consumer Psychology*,2023.
- [50] Zhu, Y. , J. Zhang, J. Wu, and Y. Liu. AI Is Better When I’m Sure: The Influence of Certainty of Needs on Consumers’ Acceptance of AI Chatbots[J]. *Journal of Business Research*,2022,150:642–652.
- [51] Ho, G. , D. Wheatley, and C. T. Scialfa. Age Differences in Trust and Reliance of a Medication Management System[J]. *Interacting with Computers*,2005,17,(6):690–710.
- [52] Boerman, S. C. , S. Kruikeemeier, and F. J. Zuiderveen Borgesius. Exploring Motivations for Online Privacy Protection Behavior: Insights from Panel Data[J]. *Communication Research*,2021,48,(7):953–977.
- [53] Castelvecchi, D. Can We Open the Black Box of AI? [J]. *Nature (London)*,2016,538,(7623):20–23.
- [54] Yeomans, M. , A. Shah, S. Mullainathan, and J. Kleinberg. Making Sense of Recommendations[J]. *Journal of Behavioral Decision Making*,2019,32,(4):403–414.
- [55] Złotowski, J. , K. Yogeewaran, and C. Bartneck. Can We Control It? Autonomous Robots Threaten Human Identity, Uniqueness, Safety, and Resources[J]. *International Journal of Human-Computer Studies*,2017,100:48–54.
- [56] Prah, A. , and L. Van Swol. Out with the Humans, in with the Machines? Investigating the Behavioral and Psychological Effects of Replacing Human Advisors with a Machine[J]. *Human-Machine Communication*,2021,2:209–234.
- [57] Dietvorst, B. J. , J. P. Simmons, and C. Massey. Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them[J]. *Management Science*,2018,64,(3):1155–1170.
- [58] Qiu, L. , and I. Benbasat. Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems[J]. *Journal of Management Information Systems*,2009,25,(4):145–182.
- [59] Arsenyan, J. , and A. Mirowska. Almost Human? A Comparative Case Study on the Social Media Presence of Virtual Influencers [J]. *International Journal of Human-Computer Studies*,2021,155,102694.

- [60] Schepers, J. J. L. , D. Belanche Gracia, L. V. Casaló, and C. Flavián. How Smart Should a Service Robot Be? [J]. *Journal of service research*, 2022, 25, (4) :565 – 582.
- [61] 吴继飞, 于洪彦, 朱翊敏, 张湘昀. 人工智能推荐对消费者采纳意愿的影响[J]. *哈尔滨:管理科学*, 2020, (5) :29 – 43.
- [62] Granulo, A. , C. Fuchs, and S. Puntoni. Preference for Human (Vs. Robotic) Labor Is Stronger in Symbolic Consumption Contexts [J]. *Journal of Consumer Psychology*, 2021, 31, (1) :72 – 80.
- [63] Ruan, Y. , and J. Mezei. When Do AI Chatbots Lead to Higher Customer Satisfaction Than Human Frontline Employees in Online Shopping Assistance? Considering Product Attribute Type[J]. *Journal of Retailing and Consumer Services*, 2022, 68, 103059.
- [64] Franke, G. R. , B. A. Huhmann, and D. L. Mothersbaugh. Information Content and Consumer Readership of Print Ads: A Comparison of Search and Experience Products[J]. *Journal of the Academy of Marketing Science*, 2004, 32, (1) :20 – 31.
- [65] Farjam, M. On Whom Would I Want to Depend; Humans or Computers? [J]. *Journal of Economic Psychology*, 2019, 72: 219 – 228.
- [66] Epstein, S. Integration of the Cognitive and the Psychodynamic Unconscious[J]. *American Psychologist*, 1994, 49, (8) :709.
- [67] Monin, B. The Warm Glow Heuristic: When Liking Leads to Familiarity[J]. *Journal of Personality and Social Psychology*, 2003, 85, (6) :1035 – 1048.
- [68] Gai, P. J. , and A. K. Klesse. Making Recommendations More Effective through Framings; Impacts of User-Versus Item-Based Framings on Recommendation Click-Throughs[J]. *Journal of Marketing*, 2019, 83, (6) :61 – 75.
- [69] Prah, A. , and L. Van Swol. Understanding Algorithm Aversion: When Is Advice from Automation Discounted? [J]. *Journal of Forecasting*, 2017, 36, (6) :691 – 702.
- [70] Glikson, E. , and A. W. Woolley. Human Trust in Artificial Intelligence: Review of Empirical Research [J]. *Academy of Management Annals*, 2020, 14, (2) :627 – 660.
- [71] Filiz, I. , J. R. Judek, M. Lorenz, and M. Spiwoks. Reducing Algorithm Aversion through Experience[J]. *Journal of Behavioral and Experimental Finance*, 2021, 31, 100524.
- [72] Van Dongen, K. , and P. P. Van Maanen. A Framework for Explaining Reliance on Decision Aids[J]. *International Journal of Human-Computer Studies*, 2013, 71, (4) :410 – 424.
- [73] Hildebrand, C. , and A. Bergner. Conversational Robo Advisors as Surrogates of Trust: Onboarding Experience, Firm Perception, and Consumer Financial Decision Making[J]. *Journal of the Academy of Marketing Science*, 2020, 49, (4) :659 – 676.
- [74] You, S. , C. L. Yang, and X. Li. Algorithmic Versus Human Advice: Does Presenting Prediction Performance Matter for Algorithm Appreciation? [J]. *Journal of Management Information Systems*, 2022, 39, (2) :336 – 365.
- [75] Dijkstra, J. J. , W. B. G. Liebrand, and E. Timminga. Persuasiveness of Expert Systems[J]. *Behaviour & Information Technology*, 1998, 17, (3) :155 – 163.
- [76] Efenđić, E. , P. P. F. M. Van De Calseyde, and A. M. Evans. Slow Response Times Undermine Trust in Algorithmic (but Not Human) Predictions[J]. *Organizational Behavior and Human Decision Processes*, 2020, 157:103 – 114.
- [77] Jago, A. S. Algorithms and Authenticity[J]. *Academy of Management Discoveries*, 2019, 5, (1) :38 – 56.
- [78] Longoni, C. , A. Bonezzi, and C. K. Morewedge. Resistance to Medical Artificial Intelligence[J]. *Journal of Consumer Research*, 2019, 46, (4) :629 – 650.
- [79] Puntoni, S. , R. W. Reczek, M. Giesler, and S. Botti. Consumers and Artificial Intelligence: An Experiential Perspective[J]. *Journal of Marketing*, 2021, 85, (1) :131 – 151.
- [80] Hoskins, J. D. The Evolving Role of Hit and Niche Products in Brick-and-Mortar Retail Category Assortment Planning: A Large-Scale Empirical Investigation of U. S. Consumer Packaged Goods[J]. *Journal of Retailing and Consumer Services*, 2020, 57, 102234.
- [81] Fisher, G. Intertemporal Choices Are Causally Influenced by Fluctuations in Visual Attention[J]. *Management Science*, 2021, 67, (8) :4961 – 4981.
- [82] Hou, Y. T. Y. , W. Y. Lee, and M. Jung. Should I Follow the Human, or Follow the Robot? ——Robots in Power Can Have More Influence Than Humans on Decision-Making[C]. *Association for Computing Machinery:New York*, 2023.
- [83] Plangger, K. , D. Grewal, K. De Ruyter, and C. Tucker. The Future of Digital Technologies in Marketing: A Conceptual Framework and an Overview[J]. *Journal of the Academy of Marketing Science*, 2022, 50, (6) :1125 – 1134.
- [84] Makridis, C. A. , and S. Mishra. Artificial Intelligence as a Service, Economic Growth, and Well-Being[J]. *Journal of Service Research*, 2022, 25, (4) :505 – 520.

A Literature Review and Prospects of Consumer Algorithmic Attitude Research based on Trust Perspective

SUN Yu-jia, SU Song, TANG Hong-hong

(School of Economics and Business Administration, Beijing Normal University, Beijing, 100875, China)

Abstract: Algorithms form the fundamental logic for managing information, facilitating automated decision-making through statistical analysis and precise predictions. The rise of Artificial Intelligence (AI) algorithms is increasingly reshaping human society, marking the advent of an intelligent era.

On a theoretical front, given the demonstrated superiority of algorithms in specific tasks, scholars are increasingly investigating their applicability across a wide array of fields. Analyzing the trajectory and findings of algorithmic attitudes literature through an econometric lens reveals that consumer attitudes towards algorithms have been dynamic, encompassing algorithm aversion, acceptance, or even pure preference under specific conditions. The evidence unequivocally supports the evolution of consumer trust. As algorithms continue to advance, consumer access and utilization will proliferate, necessitating a critical examination of how consumer attitudes toward algorithms will metamorphose.

From a practical standpoint, the application of algorithms in digital transformation empowers companies to enhance customer satisfaction, mitigate risks, bolster competitiveness, and ensure compliance with regulatory standards. Recent introductions of innovative chatbot models like ChatGPT into public consciousness hint at the integration of a new generation of intelligent algorithms into daily life. These interactive algorithmic models bridge the gap between consumers and AI, fostering increased consumer trust and potentially reshaping future attitudes and behaviors. Thus, systematically exploring and prospecting the positive evolution of consumer approaches to algorithms holds intergenerational significance.

Recognizing the limitations of holistic trust in exploring consumer attitudes comprehensively, this study introduces a three-dimensional framework of consumer algorithmic trust: dispositional trust, situational trust, and learned trust. The objective is to elucidate behavioral changes from a trust-based perspective. To the best of our knowledge, no existing studies have systematically consolidated positive algorithmic attitudes within this structured framework. Furthermore, prior research on the determinants of algorithmic trust has been dispersed and lacks explanatory mechanisms. Therefore, there is a pressing need for a systematic exploration of the conditions under which algorithms can garner positive user attitudes and the role of contingent factors in this process. Grounded in the establishment of the triadic relationship of algorithmic trust, this article summarizes the outcomes of discussions surrounding individual, algorithmic, and other influencing factors from previous studies. It offers a mechanistic explanation rooted in dual processing theory to construct a comprehensive model of algorithmic trust. With the premise that these factors are interdependent inputs, this article delves into an in-depth analysis of the antecedents, processing mechanisms, and outcomes of algorithmic trust in both non-interactive and dynamic interactive contexts.

Lastly, prior literature on downstream consumer behavior, such as decision-making and judgment, has been relatively limited, primarily focusing on one-dimensional investigations. Therefore, to address this gap, this paper posits that the three forms of trust exert influence and overlap, ultimately impacting consumer behavior through the overarching dimension of algorithmic trust. In terms of downstream behaviors, consumers not only place psychological trust in algorithms but also exhibit varying behaviors corresponding to algorithm appreciation, choice, and dependence.

In the future, the enhancement of our algorithm governance program should concurrently prioritize the promotion of technical trust and governance confidence, demanding increased levels of involvement from all stakeholders. In the realm of academic studies, AI-related topics remain at the vanguard of several research domains and hotbeds of inquiry. In this paper, we present five research avenues meriting further exploration based on both practical and theoretical considerations. Each research direction centers on algorithmic trust as its core theme, proposing viable, forward-looking, and innovative topics grounded in contemporary theoretical foundations and real-world applications.

Key Words: three-dimensional framework of algorithm trust; algorithmic appreciation; algorithmic choice; algorithmic dependence

JEL Classification: M10, M31, I81

DOI: 10.19616/j.cnki.bmj.2023.10.010

(责任编辑:张任之)