

• 第二十八届中国科协年会学术论文 •

隐私风险感知对自动驾驶汽车初始信任的影响： 从业者与非从业者的差异化反应*

孙一飞^{1,2} 李秀兰^{1,2} 杜峰^{3,4} 齐玥^{1,2}

(¹中国人民大学心理学系; ²中国人民大学心理学系实验室, 北京 100872)

(³中国科学院心理研究所认知科学与心理健康全国重点实验室, 北京 100101)

(⁴中国科学院大学心理学系, 北京 100049)

摘要 过往研究主要关注公众对自动驾驶系统的信任不足, 但普通消费者的过度信任同样可能导致系统误用, 进而增加使用风险。本文通过三项研究系统探索了从业背景对初始信任的影响, 聚焦于比较和校准普通消费者的信任水平, 使其更接近作为专家的从业者。研究 1 发现, 非从业者表现出过度信任倾向, 并且隐私风险感知和从业背景对初始信任的影响存在交互作用。研究 2 通过操纵隐私风险水平发现从业者与非从业者存在差异化反应, 风险水平提升显著增强了非从业者的隐私风险感知, 并降低其初始信任; 而从业者的初始信任较少受到风险变化的影响。研究 3 进一步揭示了非从业者对隐私风险信息而非对称反应: 在低风险情境下, 尽管隐私风险感知显著提升, 初始信任无显著变化; 而在高风险情境下, 隐私风险感知显著增加, 初始信任显著下降。这些结果揭示了从业背景与隐私风险感知对自动驾驶初始信任的交互作用, 凸显了从业者与非从业者在初始信任形成中的差异, 并提示自动驾驶系统的设计者应采用更具针对性的信任校准策略以应对从业者与非从业者的差异化反应。

关键词 自动驾驶汽车, 初始信任, 从业背景, 感知隐私风险, 人与 AI 信任

分类号 B849

1 前言

随着人工智能技术的快速发展, 自动驾驶汽车 (Autonomous Vehicles) 逐渐展现出明显优势, 为用户提供了更加轻松的出行体验 (Narayanan et al., 2020; Wu et al., 2023)。然而, 技术的成熟只是使用的前提, 人们是否信任自动驾驶系统才是影响接受与使用该技术的关键 (许为, 葛列众, 2020; Glikson & Woolley, 2020)。在人机交互领域, 信任被定义为个体在风险和不确定性条件下, 对代理系统能够帮助其实现目标的期望 (齐玥等, 2024; Lee & See, 2004)。信任不足会导致用户对技术的排斥 (Waung et al., 2021), 而过度信任同样可能导致接管不及时

等安全隐患 (黄心语, 李晔, 2024)。因此, 如何让用户持有适度的信任, 成为自动驾驶发展过程中的关键问题。

目前, 公众对自动驾驶汽车的安全性和可靠性仍存较大疑虑 (Abraham et al., 2017; Lee & Lee, 2022; Tan et al., 2022; Yuen et al., 2020)。德勤 (Deloitte China, 2019) 的消费者调研数据显示, 超过半数的受访者对自动驾驶技术持谨慎态度, 且 53% 以上支持政府加强监管力度。普华永道 (PricewaterhouseCoopers, 2023) 的调查进一步印证了这种谨慎立场, 数据显示 60%~70% 的受访者对自动驾驶汽车缺乏信任。这种普遍的信任危机不仅阻碍了自动驾驶汽车的商业化进程, 更延缓了全自

收稿日期: 2025-05-10

* 国家自然科学基金(32471130, 32000771)、中国人民大学科学研究基金(中央高校基本科研业务费专项资金资助)项目成果(21XNLG13)、2018 年度中央高校建设世界一流大学(学科)和特色发展引导专项资金(RUCPSY0007)支持。

通信作者: 齐玥, E-mail: qiy@ruc.edu.cn

动驾驶时代的到来(Pyte et al., 2017; Waung et al., 2021)。在此背景下, 过往研究多关注于信任不足(Wintersberger et al., 2018)。然而, 用户对自动驾驶的过度信任同样会带来严重风险(齐玥 等, 2024)。个体若高估系统能力, 就可能忽视潜在风险, 无法及时修正错误(高在峰 等, 2021)。

过度信任是指个体的主观信任水平高于客观可信程度。过往研究发现, 在紧急状态下, 个体往往难以准确评估智能系统的能力, 更容易出现过度信任(Christensen et al., 2019; Robinette et al., 2017)。也有研究提出, 低估风险同样会导致过度信任。Booth 等人(2017)在宿舍门口放置一台机器人, 让其请求路人帮助它进入宿舍。结果发现, 即使学生清楚机器人进入宿舍可能带来安全隐患, 但依然允许其进入。这反映了人们对机器人的过度信任。在现实生活中, 对自动驾驶系统的过度信任, 则可能直接导致车毁人亡的惨剧。

值得注意的是, 信任并非在所有用户群体中表现一致。现有研究往往将自动驾驶用户视为同质群体(Rupp et al., 2018; Tan et al., 2022), 但实际上, 用户既包括深耕该领域的从业者, 也包括缺乏专业知识的普通消费者(Walker et al., 2016)。二者在知识储备与信任机制上存在显著差异: 自动驾驶领域的从业者更熟悉自动系统, 能够更理性地评估系统性能与潜在风险(Jin et al., 2020; Lyons & Guznov, 2019; Rieger et al., 2021), 其信任水平更接近系统的真实可信度; 相比之下, 非从业者由于缺乏相应的知识和评估能力, 更容易形成不恰当的信任水平。

更重要的是, 忽视从业者与普通消费者的背景差异, 可能导致产品设计难以满足真实用户需求(Balfe et al., 2018; Sheridan, 2019)。调查显示, 两类群体在需求上存在明显分歧: 在中国市场, 从业者更关注自动驾驶汽车的信息娱乐系统, 而对按需服务功能的重视程度(67%)显著低于普通消费者(92%)(PricewaterhouseCoopers, 2023)。这种供需认知差异不仅可能降低用户的满意度(Merritt & Ilgen, 2008; Sheridan, 2019), 还可能进一步影响用户的信任水平与接受度, 甚至在特定情境下引发安全风险。一个典型案例是隐私风险问题: 从业者在沟通中往往弱化数据使用与二次利用边界, 导致非从业者过分信任隐私保护水平, 而一旦发生信息泄漏, 则会引发信任的剧烈下降(胡雅萍 等, 2025)。因此, 比较作为设计者的从业者与作为使用者的非从业者在信任机制上的差异, 是理解和校准信任水平、实现

以用户为中心的高可信度设计的关键。

综上所述, 本研究旨在探讨如何使非从业者在初始信任水平上更接近作为设计者的从业者, 从而形成更为合理、适度的初始信任。这一问题的回答对于提升人机协作的安全性和效率具有重要意义。基于此, 本研究聚焦于 SAE 3 级¹自动驾驶汽车, 系统考察初始信任的影响因素, 并比较从业者与非从业者在信任机制上的差异, 以揭示两类群体在初始信任方面的差异化反应。

1.1 自动驾驶汽车的初始信任

当前理论一致认为, 人们对自动驾驶汽车的信任是一个动态过程(齐玥 等, 2024; Hoff & Bashir, 2015)。高在峰等(2021)在动态信任模型(Lee & See, 2004)和两阶段信任模型(Merritt & Ilgen, 2008)的基础上, 结合信任的发展阶段与影响因素提出了动态信任框架。该框架将信任发展分为倾向性信任、初始信任、实时信任和事后信任四个发展阶段, 后三者被统称为历史性信任状态。在操作者与系统的交互过程中, 对系统的信任将由倾向性信任逐渐转变为以历史性信任成分为主(Merritt & Ilgen, 2008)。其中, 初始信任是该过程的基础(齐玥 等, 2024; Zhang et al., 2020)。Hoff 和 Bashir (2015)将初始信任定义为个体在缺乏直接经验的情况下, 基于已有知识对自动系统产生的信任感。研究表明, 初始信任在消费者对自动驾驶汽车的接受度和购买意愿中发挥着重要作用(Dirsehan & Can, 2020; Zhang et al., 2020)。因此, 初始信任是本研究的核心关注点。

1.2 自动驾驶汽车初始信任的影响因素

动态信任框架将信任的影响因素分为操作者特征、系统特征以及情境特征三个方面(高在峰 等, 2021)。对于自动驾驶汽车而言, 初始信任的形成尤其依赖于操作者的两类因素: 其一是相对稳定的固有特质, 如性别、年龄和教育背景等, 它们独立于系统与情境, 构成了个体的基本信任倾向。其二是通过与相似系统交互或社会学习所积累的先验经验, 基于系统与情境特征的感知以及他人的社会影

¹ 美国汽车工程师学会(Society of Automotive Engineers, SAE)提出的自动化分级标准已被广泛接受, 将自动驾驶车辆的自动化程度划分为从无自动化(L0)到完全自动化(L5)的 6 个等级(On-Road Automated Driving (Orad) Committee, 2023)。本研究主要以 SAE 3 (L2)级自动驾驶汽车为研究对象, 这一级别首次允许驾驶员在特定条件下短暂脱离驾驶任务, 但仍需在风险情境下及时接管。其高度不确定性和潜在风险使得对其信任问题成为当前学术界与产业界的焦点。

响均是先验经验的重要来源。

已有研究发现,个体对自动驾驶汽车的性能感知和风险感知会直接影响初始信任。例如,Zhang 等人(2019)基于技术接受模型(Technology Acceptance Model, TAM; Davis et al., 1989)发现,感知有用性和安全风险感知是影响自动驾驶汽车初始信任的关键因素。此外,社会影响也发挥重要作用(Wang et al., 2024),他人的评价、广告宣传以及社会舆论等外部信息均能改变用户的态度(Carlson et al., 2014; Celmer et al., 2018)。例如,Zhang 等人(2020)通过对 647 名中国驾驶员的调查发现,社会影响通过初始信任间接增强自动驾驶汽车的使用意愿。这些结果共同表明系统性能感知(如感知有用性、感知缺陷)、系统安全感知(如安全风险感知、隐私风险感知)和社会影响在塑造自动驾驶汽车初始信任中的核心作用。

然而,现有模型在解释群体差异时仍存在局限,尚未充分考虑从业背景的作用(Tan et al., 2022; Tao et al., 2018; Zhang et al., 2020)。从业背景反映了个体在专业知识和经验累积上的差异,这种差异可能导致其在面对相同系统时形成不同的主观感知与信任判断(高在峰 等, 2021; Balfé et al., 2018)。例如,业内从业者更可能基于专业知识识别出系统缺陷或隐私风险,从而形成与普通消费者不同的初始信任水平。

综上所述,本研究在动态信任框架的基础上,结合技术接受模型,提出了一个由四个因素组成的初始信任形成框架。该框架区分了直接经验路径与间接经验路径,系统揭示了先验经验对初始信任的多层次影响机制。具体包括:(1)系统性能感知(直接经验来源):感知有用性与感知缺陷反映个体基于直接操作经验形成的性能判断;(2)系统安全感知(直接经验来源):安全风险感知与隐私风险感知反映个体对系统潜在风险和安全性的主观评估;(3)社会学习(间接经验来源):他人态度与社会信息通过社会影响塑造个体信任;(4)从业背景(间接经验来源):作为关键调节因素,反映个体专业知识和经验的差异,并可能在上述变量与初始信任之间发挥差异化作用。通过整合上述四类因素,本文旨在通过先验经验系统解释不同群体在自动驾驶初始信任上的差异及其来源。

1.2.1 性能感知:感知有用性与感知缺陷

动态信任框架(高在峰 等, 2021)认为,个体对自动驾驶系统性能的主观感知是初始信任的重要

影响因素,其中感知有用性和感知缺陷是最受关注的两个核心维度。感知有用性是指个体认为使用自动驾驶系统能够提升驾驶表现的程度(Davis et al., 1989; Zhang et al., 2020),而感知缺陷则指对系统潜在缺陷的感知(Ma et al., 2020)。根据技术接受模型,感知有用性不仅能显著影响技术接受意愿,而且这种影响在很大程度上是通过信任实现的(Davis et al., 1989)。自动驾驶信任理论也同样强调,初始信任在很大程度上依赖于系统的感知性能(高在峰等, 2021; Hoff & Bashir, 2015)。尽管以往研究多数集中在对技术的感知有用性上,但在做出信任判断时,感知缺陷同样至关重要。

因此,感知有用性与感知缺陷作为基于性能的关键变量,会直接塑造个体对自动驾驶系统的初始信任。当消费者感知到更高的有用性时,他们通常对自动驾驶系统持有更高的初始信任水平(Dirsehan & Can, 2020)。相反,若消费者感知到更多的缺陷,他们往往会对自动驾驶系统的性能持消极的态度,进而降低初始信任水平(Ma et al., 2020)。

因此,提出以下假设:

H1:感知有用性对从业者和非从业者的自动驾驶汽车初始信任具有正向影响。

H2:感知缺陷对从业者和非从业者的自动驾驶汽车初始信任具有负向影响。

1.2.2 安全感知:安全风险感知与隐私风险感知

安全风险感知和隐私风险感知分别指个体对人身安全和信息安全的担忧,反映个体对潜在风险和安全性认知和感受(Ha et al., 2020; Zhang et al., 2019)。近期研究结果表明,人机信任存在安全维度,即人们会因为相信机器不会对其人身和隐私安全造成威胁而产生信任(由姗姗 等, 2025)。相反,当安全性无法得到充分保障时,消费者会对其潜在风险表现出持续担忧(Ahmed et al., 2020)。例如,德勤(Deloitte China, 2019)调查显示,42%的消费者对于生物数据共享机制存在顾虑。

安全风险感知是影响初始信任的关键因素。已有研究表明,当个体感知到较高的安全风险时,他们对自动驾驶汽车的初始信任水平显著下降(Hoff & Bashir, 2015; Zhang et al., 2019),并且这一趋势在跨年龄群体中表现出一致性(Ma et al., 2020)。

隐私风险同样是消费者关注的重要问题。消费者普遍担心个人数据被车企、保险公司滥用或越权调用(胡雅萍 等, 2025; Ahmed et al., 2020)。例如,一项对美国驾驶员群体的研究显示,约三分之一的

受访者对数据滥用导致的隐私风险表示严重担忧(Kyriakidis et al., 2015)。

然而, 隐私风险感知对初始信任的影响仍存在争议。Ma 等人(2020)发现, 对于知识储备丰富的成人群体, 隐私风险感知与对自动驾驶校车的初始信任呈负相关, 但在先验经验不足的儿童群体中未观察到类似效应。Zhang 等人(2019)对中国驾驶员的研究发现, 隐私风险感知与初始信任之间没有显著关系。这些研究结果表明, 隐私风险的影响可能受某些边界条件的调节, 跨年龄群体结果的不一致说明先验经验可能起到关键作用。

因此, 本研究将进一步探讨安全风险感知和隐私风险感知对自动驾驶汽车初始信任的影响, 并提出以下假设:

H3: 安全风险感知对从业者和非从业者的自动驾驶汽车初始信任具有负向影响。

H4: 隐私风险感知对从业者的自动驾驶汽车初始信任具有负向影响, 但对非从业者的影响较小。

1.2.3 社会影响

社会影响指个体对重要他人(如家人、朋友等)是否支持某一特定行为的感知(Fishbein & Ajzen, 1975)。在技术接受和信任的形成过程中, 他人的意见是先验经验的重要来源(高在峰 等, 2021)。在初期阶段, 个体往往缺乏直接的交互经验, 因此对自动驾驶系统的信任主要基于他人的描述或媒体报道(Hoff & Bashir, 2015)。研究表明, 当自动驾驶系统与知名品牌相关联时, 消费者对其信任度更高(Carlson et al., 2014; Celmer et al., 2018)。此外, 如果品牌形象被认为具有高温暖或高能力特征, 消费者的信任也会增强(Qin et al., 2024)。

实证研究进一步验证了社会影响在塑造自动驾驶汽车初始信任方面的关键作用。Zhang 等人(2020)发现, 社会影响会通过初始信任影响个体使用自动驾驶汽车的意愿。然而, 这一研究结果与其他技术接受研究有所不同(Rupp et al., 2018; Tao et al., 2018), 这可能是由于当时研究所涉及的自动驾驶汽车尚未全面商业化。在缺乏直接体验的情况下, 个体对自动化技术的认知评估可能不稳定, 并且更容易受到媒体或同伴意见的影响。因此, 本研究旨在重新评估社会影响在自动驾驶信任中的作用, 并提出以下假设:

H5: 社会影响对从业者和非从业者的自动驾驶汽车初始信任具有正向影响。

1.3 从业背景: 从业者与非从业者的信任差异

已有研究表明, 信任形成中最强的影响因素是个体对自动化系统的理解, 其影响强于系统的性能和可靠性, 而先验经验在促进这种理解方面起重要作用(Balfe et al., 2018)。自动驾驶行业的从业者代表了先验经验较高的群体, 他们对自动驾驶汽车更加熟悉, 能够形成合理预期并保持信任的稳定性; 而普通消费者由于相关知识和经验的局限, 更易受到表面线索影响, 往往形成更高但更脆弱的初始信任(Pop et al., 2015; Sheridan, 2019)。

在信任形成的初期, 从业者能更好地识别出系统的性能边界和潜在风险, 表现出更为谨慎的信任态度(Merritt & Ilgen, 2008)。例如, 经验丰富的驾驶员在使用自动驾驶系统时更倾向于保持对道路或仪表盘的注意, 而非转向手机等分心活动, 这反映出其对系统安全性与可靠性的持续关注(Jin et al., 2020)。相比之下, 非从业者缺乏先验经验以及对系统运作原理的理解, 更容易依赖系统的直观表现和直接反馈(Onnasch et al., 2023)。由于缺乏识别系统局限性和潜在风险(如隐私问题)的能力, 非从业者容易高估系统性能和安全性并表现出过度信任(Biassoni et al., 2016; Kahl, 2021)。然而, 当系统在关键情况下表现不佳时, 信任往往会迅速崩塌(Beggiato & Krems, 2013; Kraus et al., 2020), 且初始信任越高, 信任崩塌的幅度也越剧烈(Hoff & Bashir, 2015)。

在隐私风险评估时, 不同从业背景造成的个体差异可能更加明显, 但当前研究对此的关注仍然不足。根据动态信任框架, 客观的系统与情境信息需经过个体的认知加工转变为主观感知, 进而影响初始信任(高在峰 等, 2021)。不同经验和知识水平的个体对隐私风险的解读存在差异(Balfe et al., 2018)。例如, 曾经历过隐私泄露事件的人群对隐私风险的敏感性更高(胡雅萍 等, 2025)。因此, 本研究假设, 从业背景在隐私风险感知对自动驾驶系统的初始信任的影响中起调节作用: 非从业者由于先验经验不足, 可能对隐私风险感知较低, 但当隐私风险与其预期不符时, 初始信任的波动更为剧烈; 相比之下, 从业者由于对隐私风险有更合理的预期, 其初始信任水平较为稳定。在自动驾驶系统的设计和推广过程中, 认识到不同群体的差异对于确保用户获得适当的信息支持、建立合理的信任水平至关重要(Sheridan, 2019)。

基于上述讨论, 提出以下假设:

H6: 隐私风险感知与从业背景对自动驾驶汽车的初始信任存在交互作用。相比于从业者, 非从业者对隐私风险的感知较低, 并表现出更高的初始信任。

H7: 隐私风险水平可以通过隐私风险感知影响初始信任, 并且从业背景在其中起调节作用。

1.4 研究概述

综上, 本研究旨在探讨从业者和非从业者在自动驾驶汽车初始信任及其影响因素方面的差异。研究聚焦以下三个核心问题: 从业者与非从业者对自动驾驶汽车的初始信任存在哪些差异? 主观感知特征如何影响初始信任, 这种影响在从业者与非从业者之间有何不同? 是否可以通过干预手段来弥合从业者与非从业者在自动驾驶汽车初始信任方面的差异?

为回答这些问题, 研究1首先探讨影响自动驾驶汽车初始信任的因素, 并分析从业者与非从业者之间的差异; 研究2通过操纵隐私风险水平以验证研究1的发现; 研究3进一步通过实验干预探讨非从业者的信任波动是否可归因于隐私风险敏感性的增强。

2 研究1: 自动驾驶汽车初始信任的影响因素: 从业者与非从业者的比较

研究1进行了一项问卷调查, 以探讨从业者与非从业者在自动驾驶汽车初始信任方面的差异, 并考察感知特征如何影响初始信任。

2.1 方法

2.1.1 被试

本研究通过定向招募的方式从自动驾驶企业招募了200名从业者, 并通过Credamo平台招募了601名非从业者, 共收集到801份问卷。为了保证数据质量, 问卷中设置了注意力检测题, 共有735份数据通过检测, 通过率为92%。此外, 为了控制驾驶经验, 未获得驾照的被试(35人, 占4.8%)被排除在分析之外。学历为初中(1人, 占0.1%)及普高/中专/技校/职高(12人, 占1.6%)的被试由于数量过少, 也未纳入分析。最终获得689份有效样本, 总体有效率为86%, 其中从业者165人, 非从业者524人。

为检验两组样本在人口统计学特征上的差异, 我们对性别、年龄、学历和驾驶经验进行了比较。卡方检验结果表明, 两组在年龄上没有显著差异

($\chi^2_{\text{年龄}}(3) = 5.3, p_{\text{年龄}} = 0.152$); 但在性别($\chi^2_{\text{性别}}(1) = 82.6, p_{\text{性别}} < 0.001$)、学历($\chi^2_{\text{学历}}(2) = 111.0, p_{\text{学历}} < 0.001$)和驾龄($\chi^2_{\text{驾龄}}(4) = 20.7, p_{\text{驾龄}} < 0.001$)上均存在显著差异。在后续分析中, 我们将年龄、性别、学历和驾龄作为控制变量。两组被试的具体人口统计信息见附表2。

2.1.2 测量工具

本研究使用的量表全部翻译自先前研究中使用过的成熟量表, 采用5点李克特量表进行测量, 选项范围从1(非常不同意)到5(非常同意)(见网络版附表1)。

感知有用性。采用Davis等人(1989)开发的感知有用性量表, Zhang等人(2020)也使用了该量表。该量表使用5个题项测量, 本研究中该量表的内部一致性系数Cronbach's $\alpha = 0.75$ 。

感知缺陷。采用Ma等人(2020)开发的感知缺陷量表。该量表使用2个题项测量, 本研究中该量表的内部一致性系数Cronbach's $\alpha = 0.59$ 。

安全风险感知。采用Zmud等人(2016)开发的安全风险感知量表, Zhang等人(2019)也使用了该量表。该量表使用2个题项测量, 本研究中该量表的内部一致性系数Cronbach's $\alpha = 0.88$ 。

隐私风险感知。采用Kyriakidis等人(2015)开发的隐私风险感知量表, Zhang等人(2019)也使用了该量表。该量表使用3个题项测量, 本研究中该量表的内部一致性系数Cronbach's $\alpha = 0.94$ 。

社会影响。采用Zhang等人(2020)使用的社会影响量表。该量表使用3个题项测量, 其中2项来自Osswald等人(2012)开发的题目, 1项来自Madigan等人(2017)开发的题目。本研究中该量表的内部一致性系数Cronbach's $\alpha = 0.82$ 。

初始信任。采用Choi和Ji(2015)开发的信任量表, Zhang等人(2020)和Ma等人(2020)也使用了该量表。该量表使用3个题项测量, 本研究中该量表的内部一致性系数Cronbach's $\alpha = 0.90$ 。

2.1.3 实验流程

被试首先了解实验信息并签署知情同意书。随后, 阅读一份关于SAE 3级自动驾驶汽车的简要介绍。主试向被试强调其能够执行如车道保持和变道等驾驶任务, 但在复杂场景下仍需人为干预。接着, 被试填写感知有用性、感知缺陷、安全和隐私风险感知、社会影响及初始信任测量问卷。最后, 收集被试的人口统计信息, 包括年龄、性别、教育水平、

驾驶执照状态和驾驶经验。所有被试在完成实验后均获得预定的实验报酬。

2.2 结果

2.2.1 共同方法偏差(CMB)

采用 Harman 单因子检验测试共同方法偏差。结果显示, 具有特征值大于 1 的因子共有 2 个, 其中第一因子解释的方差占比为 45.80%, 低于 50% 的标准, 表明本研究不存在严重的共同方法偏差问题 (Podsakoff et al., 2003; Podsakoff & Organ, 1986)。

2.2.2 从业者与非从业者在各变量上的描述统计与差异检验

从业者与非从业者在各变量上的平均数与标准差见表 1。对两组被试的初始信任、感知有用性、感知缺陷、安全风险感知、隐私风险感知和社会影响进行独立样本 t 检验。由于两组被试样本量差异较大且方差不等, 因此采用 Welch's t 值进行检验。

结果表明, 从业者对自动驾驶汽车的初始信任 ($M = 3.14$, $SD = 1.00$) 显著低于非从业者 ($M = 4.04$, $SD = 0.74$), $t(223) = -10.71$, $p < 0.001$, Cohen's $d = -1.03$; 从业者的感知有用性 ($M = 3.51$, $SD = 0.96$) 显著低于非从业者 ($M = 4.08$, $SD = 0.56$), $t(200) = -7.30$, $p < 0.001$, Cohen's $d = -0.73$; 从业者的感知缺陷 ($M = 3.70$, $SD = 0.87$) 显著高于非从业者 ($M = 2.84$, $SD = 0.98$), $t(306) = 10.83$, $p < 0.001$, Cohen's $d = 0.94$; 从业者的安全风险感知 ($M = 3.98$, $SD = 1.03$) 显著高于非从业者 ($M = 2.69$, $SD = 1.20$), $t(317) = 13.44$, $p < 0.001$, Cohen's $d = 1.15$; 从业者的隐私风险感知 ($M = 3.88$, $SD = 1.19$) 显著高于非从业者 ($M = 2.44$, $SD = 1.21$), $t(279) = 13.56$, $p < 0.001$,

Cohen's $d = 1.21$; 从业者受到的社会影响 ($M = 3.34$, $SD = 0.97$) 显著低于非从业者 ($M = 4.09$, $SD = 0.72$), $t(224) = -9.20$, $p < 0.001$, Cohen's $d = -0.88$ 。

2.2.3 相关分析

进一步对各变量进行相关分析, 以探讨 6 个变量(感知有用性、安全风险感知、隐私风险感知、感知缺陷、社会影响和初始信任)之间的关系(见表 1)。结果表明, 感知有用性和社会影响与两组被试的初始信任呈显著正相关; 安全风险感知、隐私风险感知以及感知缺陷与两组被试的初始信任呈显著负相关。

2.2.4 回归分析

以初始信任为因变量, 感知有用性、感知缺陷、安全风险感知、隐私风险感知和社会影响为自变量, 从业背景(哑变量编码)为调节变量, 进行多层线性回归分析。模型中纳入了各自变量与从业背景的二阶交互项, 并在控制性别、年龄、学历和驾驶经验(均作哑变量编码)的基础上, 考察了各变量对从业者和非从业者初始信任的影响。回归模型的第一层放入了控制变量, 第二层放入了其他预测变量, 各变量及交互项对初始信任的预测作用见表 2。

结果表明, 隐私风险感知与从业背景的二阶交互作用显著 ($\beta = -0.06$, $p = 0.038$), 隐私风险感知对初始信任的预测作用在两类群体间存在差异。进一步的简单斜率检验结果表明(见图 1), 从业者的隐私风险感知显著负向预测初始信任 ($simple\ slope = -0.21$, $t = -3.59$, $p < 0.001$), 但是对于非从业者, 其预测作用不显著 ($simple\ slope = -0.06$, $t = -1.58$, $p = 0.114$)。

表 1 研究 1 中各变量的描述统计及相关分析结果

从业背景		M	SD	1	2	3	4	5	6
从业者组 ($N = 165$)	1 初始信任	3.14	1.00	—					
	2 感知有用性	3.51	0.96	0.50***	—				
	3 感知缺陷	3.70	0.87	-0.23**	-0.10	—			
	4 安全风险感知	3.98	1.03	-0.47***	-0.17*	0.31***	—		
	5 隐私风险感知	3.88	1.19	-0.36***	-0.10	0.41***	0.51***	—	
	6 社会影响	3.34	0.97	0.70***	0.51***	-0.13	-0.37***	-0.18*	—
非从业者组 ($N = 524$)	1 初始信任	4.04	0.74	—					
	2 感知有用性	4.08	0.56	0.61***	—				
	3 感知缺陷	2.84	0.98	-0.50***	-0.38***	—			
	4 安全风险感知	2.69	1.20	-0.69***	-0.49***	0.70***	—		
	5 隐私风险感知	2.44	1.21	-0.49***	-0.34***	0.67***	0.71***	—	
	6 社会影响	4.09	0.72	0.79***	0.57***	-0.46***	-0.62***	-0.38***	—

注: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

表 2 研究 1 中从业者组与非从业者组的回归分析结果

预测变量	β	t	p
性别(男)	-0.03	-1.31	0.191
年龄(18 至 30 岁)	0.12	1.82	0.069
年龄(30 至 40 岁)	0.17	1.84	0.066
年龄(40 至 50 岁)	0.06	1.53	0.126
学历(专科)	-0.01	-0.12	0.902
学历(本科)	0.03	1.12	0.261
驾龄(1 年及以内)	-0.01	-0.03	0.979
驾龄(1 至 3 年)	0.02	0.54	0.590
驾龄(4 至 6 年)	-0.01	0.40	0.691
驾龄(7 至 9 年)	0.04	1.24	0.216
从业背景	0.01	1.45	0.148
感知有用性	0.18	4.93	<0.001
感知缺陷	-0.01	-0.46	0.643
安全风险感知	-0.22	-4.94	<0.001
隐私风险感知	-0.10	-1.58	0.114
社会影响	0.51	13.42	<0.001
从业背景×感知有用性	-0.01	-0.33	0.738
从业背景×感知缺陷	-0.03	-1.11	0.267
从业背景×安全风险感知	0.01	0.27	0.785
从业背景×隐私风险感知	-0.06	-2.07	0.038
从业背景×社会影响	-0.01	-0.05	0.959

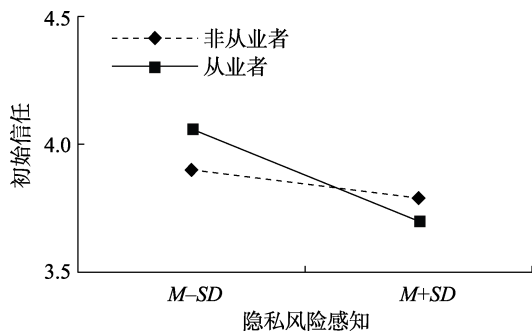


图 1 研究 1 中简单斜率检验: 隐私风险感知和从业背景

在主效应层面, 感知有用性($\beta = 0.18, p < 0.001$)和社会影响($\beta = 0.51, p < 0.001$)均显著正向预测初始信任; 安全风险感知($\beta = -0.22, p < 0.001$)显著负向预测初始信任; 感知缺陷($\beta = -0.01, p = 0.643$)和从业背景($\beta = 0.01, p = 0.148$)对初始信任的影响均不显著。

2.3 讨论

研究 1 旨在探讨从业者和非从业者在自动驾驶汽车初始信任上的差异及其影响机制。基于动态信任框架(高在峰 等, 2021)并参考技术接受模型(TAM, Davis et al., 1989), 研究 1 重点考察了四类因素: 系统性能感知(感知有用性、感知缺陷)、系统安全感

知(安全风险感知、隐私风险感知)、社会学习(社会影响), 以及操作者特征中的关键变量——从业背景。

总体结果显示, 非从业者的初始信任水平显著高于从业者。他们倾向于高估系统有用性、低估系统风险与缺陷, 并更容易受到社会影响。这表明, 在信任形成的早期阶段, 从业者通常持谨慎、保守的态度, 而非从业者可能因缺乏技术知识与风险意识表现出过度信任。

进一步分析发现, 隐私风险感知对初始信任的影响存在群体差异(支持假设 H6)。对于从业者, 隐私风险感知显著负向预测初始信任; 而对非从业者, 其作用不显著(支持假设 H4)。这一差异可能源于从业者对隐私风险更为敏感(Balfe et al., 2018)。相反, 非从业者对系统机制缺乏理解, 对潜在隐私风险感知不足, 导致该因素对其信任水平影响较弱。该结果也提示, 以往研究(如 Ma et al., 2020; Zhang et al., 2019)未发现隐私风险感知对初始信任的显著影响, 可能与当时自动驾驶汽车并未普及、被试对于自动驾驶汽车不够熟悉有关。

在其他预测因素方面, 社会影响、感知有用性和安全风险感知均能显著预测初始信任, 而感知缺

陷的影响不显著。与 Zhang 等人(2020)的研究结果一致, 本研究发现, 社会影响是最强的预测指标(支持假设 H5)。这一结果表明, 在信任形成的早期阶段, 个体缺乏对自动驾驶汽车的充分信息和直接经验, 因此媒体和他人意见是先验经验的主要来源。感知有用性正向预测初始信任(支持假设 H1), 符合技术接受模型(Davis et al., 1989)的核心假设, 即感知有用性是技术采纳与信任形成的关键驱动因素。安全风险感知对初始信任均具有显著负向预测作用(支持假设 H3), 与既有研究一致(Hoff & Bashir, 2015; Ma et al., 2020; Zhang et al., 2019), 说明即使缺乏实际的负面经验, 安全担忧也可能降低个体的初始信任。

相比之下, 感知缺陷未能显著预测初始信任(未支持假设 H2), 与 Ma 等人(2020)的研究结果不同。这可能与样本特征及人物场景差异有关, Ma 等人(2020)的研究聚焦于儿童家长对自动驾驶校车的信任, 其视角不同于普通乘客; 同时, 儿童群体的先验经验与成年群体不同, 这可能导致信任判断的差异。此外, 本研究中感知缺陷量表的信度较低, 这也可能导致其未能显著预测初始信任。

综上, 研究 1 验证了系统性能感知、系统安全感知与社会影响在初始信任形成中的关键作用, 支持了动态信任框架和技术接受模型的主要假设。同时, 本研究进一步揭示了隐私风险感知的群体差异, 说明从业背景在信任形成过程中具有调节效应。研究 2 将通过操纵隐私风险水平进一步检验这一群体差异。

3 研究 2: 隐私风险感知对初始信任的影响: 从业背景的调节作用

在研究 1 中, 我们发现隐私风险感知与从业背景对初始信任的影响存在显著的交互作用。为进一步验证隐私风险感知与从业背景之间的关系及其对初始信任的影响, 研究 2 采用在线实验, 招募从业者和非从业者作为被试。首先, 我们对两组被试的隐私风险水平进行操纵, 随后测量其隐私风险感知及其对自动驾驶汽车的初始信任。

3.1 方法

3.1.1 被试

根据 G*Power 软件(Faul et al., 2007)的计算结果, 以期望效应量 $f = 0.25$ 、统计检验力 $1 - \beta = 0.95$ 、显著性水平 $\alpha = 0.05$ 进行估算, 推荐的样本量为 $N = 106$ 。我们通过定向招募方式在自动驾驶

企业中招募了 114 名从业者, 并通过 Credamo 平台招募了 107 名非从业者, 共收集到 221 份问卷。为了保证数据质量, 问卷中设置了注意力检测题, 共有 192 份数据通过检测, 通过率为 87%。此外, 为了控制驾驶经验, 未获得驾照的被试(7 人, 占 3.6%)被排除在分析之外。学历为普高/中专/技校/职高(2 人, 占 1.0%)的被试由于数量过少, 也未纳入分析。最终获得 183 份有效样本, 总体有效率为 92%, 其中从业者 90 人, 非从业者 93 人。

为检验两组样本在人口统计学特征上的差异, 我们对性别、年龄、学历和驾驶经验进行了比较。卡方检验结果表明, 两组在性别($\chi^2_{\text{性别}}(1) = 34.9, p_{\text{性别}} < 0.001$)、年龄($\chi^2_{\text{年龄}}(3) = 31.8, p_{\text{年龄}} < 0.001$)、学历($\chi^2_{\text{学历}}(2) = 46.6, p_{\text{学历}} < 0.001$)和驾龄($\chi^2_{\text{驾龄}}(4) = 58.8, p_{\text{驾龄}} < 0.001$)上均存在显著差异。在后续分析中, 我们将对从业者与非从业者的年龄、性别、学历和驾龄进行控制。两组的具体人口统计信息见附表 3。

3.1.2 实验流程与测量

被试在了解实验信息后, 签署知情同意书。随后, 两组被试依次阅读低隐私风险和高隐私风险的实验材料。我们通过操纵隐私风险水平来影响被试的隐私风险感知。已有研究表明, 接触到潜在隐私风险较高的阅读材料可能会增强个体的隐私风险感知(Beggiato & Krems, 2013)。

在实验开始时, 被试被告知, 自动驾驶汽车需通过传感设备收集和共享大量数据, 以确保行车安全、系统可靠性及最佳性能。随后, 被试分别阅读了两家公司的数据安全案例——公司 Z (低隐私风险)和公司 H (高隐私风险), 并对其自动驾驶汽车进行评估。

在低隐私风险条件下, 案例描述如下:

“Z 公司非常重视数据安全的管控, 为提高数据使用的保密性、避免个人信息被不合理的利用, Z 公司建立了数据采集信息管理平台, 对数据传输采用专线加密通道, 核心业务在私有云上进行; 对于涉及到的人脸面部特征信息、用户车辆位置信息、驾乘习惯等个人用户隐私数据, 按照《中华人民共和国个人信息保护法》相关规定进行数据脱敏处理。”

在高隐私风险条件下, 案例描述如下:

“H 公司最近发布声明称公司遭遇了黑客攻击, 用户数据遭泄露, 黑客通过在线移动应用程序获取了部分驾驶员的个人信息, 包括姓名、电话、兴趣

点收藏、车辆行驶轨迹、个人社交账号、家庭信息、订阅的套餐、个人头像、搜索历史等。本次数据泄露事件反映出 H 公司的数据安全保障还存在提升空间。”

被试使用与研究 1 相同的测量题目, 使用 5 点李克特量表(1 = 完全不同意, 5 = 完全同意)评估其隐私风险感知, 得分越高表示隐私风险感知越高(本研究中该量表的内部一致性系数 Cronbach's $\alpha = 0.95$)。此外, 初始信任也采用与研究 1 相同的测量题目, 使用 5 点李克特量表进行评分, 得分越高表示初始信任水平越高(本研究中该量表的内部一致性系数 Cronbach's $\alpha = 0.95$)。最后, 被试填写人口统计信息, 并获得实验报酬。

3.2 结果

3.2.1 对隐私风险感知的操纵检验

各变量的描述统计结果见表 3。对隐私风险感知进行 2 (从业背景: 从业者 vs. 非从业者) $\times$ 2 (隐私风险: 高 vs. 低) 混合设计方差分析(ANOVA), 结果表明, 隐私风险水平的主效应显著, $F(1, 181) = 290.35, p < 0.001, \eta_p^2 = 0.62$ 。高隐私风险条件下($M = 4.36, SD = 0.82$)的隐私风险感知显著高于低隐私风险条件($M = 3.05, SD = 1.49$)。从业背景的主效应显

著, $F(1, 181) = 93.91, p < 0.001, \eta_p^2 = 0.34$ 。从业者组($M = 4.27, SD = 0.97$)的隐私风险感知显著高于非从业者组($M = 3.15, SD = 1.47$)。

隐私风险水平与从业背景之间存在显著的交互作用, $F(1, 190) = 201.30, p < 0.001, \eta_p^2 = 0.51$ 。简单效应分析表明, 对于非从业者组, 高隐私风险条件下($M = 4.31, SD = 0.79$)的隐私风险感知显著高于低隐私风险条件($M = 2.00, SD = 1.01$), $F(1, 181) = 473.07, p < 0.001$ 。对于从业者组, 高隐私风险条件下($M = 4.41, SD = 0.86$)的隐私风险感知显著高于低隐私风险条件($M = 4.14, SD = 1.06$), 但差异更小, $F(1, 181) = 6.28, p = 0.013$ 。

3.2.2 从业背景和隐私风险感知对初始信任的影响

为检验从业背景与隐私风险感知对初始信任的作用, 本研究采用线性混合效应模型(LMM)进行分析。考虑到不同个体可能存在不同的倾向性信任水平, LMM 能够将个体差异(随机效应)与实验操控效应(固定效应)加以区分, 从而获得更精确和更具推广性的推断。使用 jamovi.2.6.26 的 GAMLj 模块进行线性混合效应模型建模, 模型中将被试编号作为随机效应, 并将从业背景、隐私风险感知及其交互项(从业背景 $\times$ 隐私风险感知)作为固定效应, 同时控制年龄、性别、学历和驾龄。结果显示, 模型拟合的 AIC 值为 760.07, 各变量及交互项对初始信任的预测作用见表 4。

分析结果表明, 从业背景与隐私风险感知的交互作用显著($\beta = 0.73, p < 0.001$), 表明隐私风险感知对初始信任的影响在两类群体间存在差异。进一步的简单斜率检验结果表明(见图 2), 对于从业者,

表 3 研究 2 中各变量的描述统计

因变量		从业者		非从业者	
		低风险	高风险	低风险	高风险
隐私风险感知	<i>M</i>	4.14	4.41	2.00	4.31
	<i>SD</i>	1.06	0.86	1.01	0.79
初始信任	<i>M</i>	3.07	2.87	4.31	1.99
	<i>SD</i>	0.92	0.88	0.53	0.87

表 4 研究 2 中线性混合效应模型固定效应结果

预测变量	β	95% CI 下限	95% CI 上限	<i>df</i>	<i>t</i>	<i>p</i>
性别(女-男)	0.07	-0.17	0.32	168	0.59	0.556
年龄(30 至 40 岁-18 至 30 岁)	-0.25	-0.54	0.05	168	-1.65	0.101
年龄(40 至 50 岁-18 至 30 岁)	-0.11	-0.53	0.30	169	-0.53	0.595
年龄(50 及以上-18 至 30 岁)	0.09	-0.63	0.81	173	0.24	0.814
学历(本科-专科)	-0.68	-1.21	-0.14	168	-2.46	0.015
学历(研究生及以上-专科)	-0.63	-1.21	-0.05	168	-2.14	0.033
驾龄(1 至 3 年-1 年及以内)	-0.24	-0.84	0.36	168	-0.79	0.433
驾龄(4 至 6 年-1 至 3 年)	-0.44	-0.99	0.11	168	-1.56	0.120
驾龄(7 至 9 年-1 年及以内)	-0.38	-0.97	0.22	168	-1.24	0.218
驾龄(10 年及以上-1 年及以内)	-0.19	-0.83	0.45	168	-0.59	0.558
从业背景(从业者-非从业者)	0.34	0.05	0.64	180	2.33	0.021
隐私风险感知	-0.93	-0.97	-0.87	197	-35.98	<0.001
从业背景(从业者-非从业者) $\times$ 隐私风险感知	0.73	0.61	0.86	342	11.42	<0.001

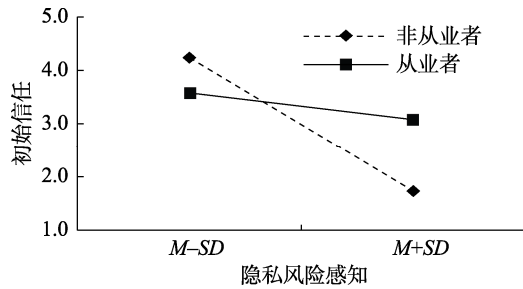


图 2 研究 2 中简单斜率检验：隐私风险感知和从业背景的交互作用

隐私风险感知提升时其初始信任显著降低, *simple slope* = -0.18 , $p = 0.002$; 对于非从业者, 隐私风险感知的提升导致其初始信任出现更加剧烈的下降, *simple slope* = -0.92 , $p < 0.001$ 。这表明从业者的初始信任水平更为稳定, 而非从业者在面临隐私担忧时信任下降更剧烈。

3.2.3 隐私风险水平对初始信任的影响：有调节的中介模型检验

为了进一步解释从业者和非从业者的信任机制, 参照温忠麟和叶宝娟(2014)提出的检验方法, 本研究考察隐私风险感知在隐私风险水平和初始信任关系中的中介作用, 以及从业背景对该中介作用前半路径的调节。采用 SPSS 宏程序 PROCESS 的模型 1 检验隐私风险水平对初始信任的影响是否受到从业背景的调节, 并控制性别、年龄、学历和驾龄。结果显示, 隐私风险水平对初始信任的影响显著($\beta = -2.32$, $p < 0.001$, 95%置信区间为 $[-2.55, -2.09]$), 隐私风险水平与从业背景对初始信任的交互作用显著($\beta = 2.12$, $p < 0.001$, 95%置信区间为 $[1.78, 2.45]$), 表明隐私风险水平对初始信任的影响受到从业背景的调节。

接下来建立有调节的中介模型, 采用 SPSS 宏程序 PROCESS 的模型 7 检验从业背景的调节作用, 并控制性别、年龄、学历和驾龄。首先检验隐私风险水平对隐私风险感知的影响。结果显示, 隐私风险水平对隐私风险感知的影响显著($\beta = 2.31$, $p < 0.001$, 95%置信区间为 $[2.05, 2.56]$), 隐私风险水平和从业背景对隐私风险感知的交互作用显著($\beta = -2.04$, $p < 0.001$, 95%置信区间为 $[-2.40, -1.67]$)。其次检验隐私风险水平通过隐私风险感知对初始信任的影响, 结果显示, 隐私风险水平对初始信任的直接效应显著($\beta = -0.60$, $p < 0.001$, 95%置信区间为 $[-0.79, -0.41]$), 隐私风险感知对初始信任的影响显著($\beta = -0.52$, $p < 0.001$, 95%置信区间为 $[-0.59, -0.44]$), 模型结果展示在图 3 中。

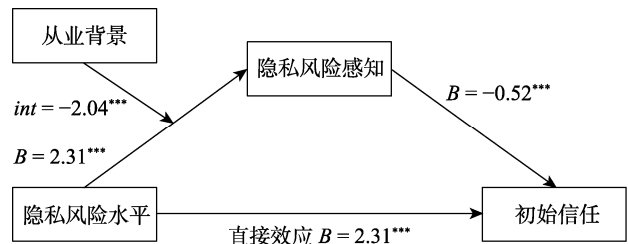


图 3 隐私风险水平对初始信任的影响：隐私风险感知的部分中介及从业背景的调节作用

综合来看, 隐私风险水平通过隐私风险感知对初始信任的部分中介效应受到从业背景的调节。对于从业者, 隐私风险水平通过隐私风险感知对初始信任的间接效应不显著, 间接效应值为 -0.14 , 95%置信区间为 $[-0.28, 0.01]$; 对于非从业者, 隐私风险水平通过隐私风险感知对初始信任的间接效应显著, 间接效应值为 -1.20 , 95%置信区间为 $[-1.42, -0.98]$, 结果见表 5。

表 5 在从业背景不同水平上隐私风险感知的部分中介效应

从业背景	间接效应值	Boot 标准误	Boot CI 下限	Boot CI 上限
0 (非从业者)	-1.20	0.11	-1.42	-0.98
1 (从业者)	-0.14	0.07	-0.28	0.01
有调节的中介效应	1.06	0.14	0.80	1.33

3.3 讨论

研究 2 通过操纵隐私风险水平, 进一步检验了隐私风险感知在初始信任形成中的作用机制, 并检验了从业背景在其中的调节效应。结果表明, 隐私风险感知的提升会降低个体的初始信任水平, 并且从业背景和隐私风险感知存在显著的交互作用(支持假设 H6)。

具体而言, 在低隐私风险条件下, 非从业者的初始信任较高, 可能是因为他们未能充分考虑数据收集、存储和使用所带来的潜在风险; 而从业者因具备相关专业知识与经验, 表现出更为谨慎的态度(Balfe et al., 2018)。然而, 当隐私风险水平升高时, 非从业者的初始信任显著下降, 甚至低于从业者。这一结果与 Kraus 等人(2020)的研究一致, 表明当实际系统表现与预期不符时, 普通用户的信任水平会出现明显波动; 相比之下, 从业者的信任水平较为稳定。

进一步, 有调节的中介分析结果支持了假设 H7, 隐私风险感知在隐私风险水平和初始信任间起中介作用, 从业背景调节了隐私风险水平对隐私

风险感知的影响。具体而言, 非从业者在隐私风险水平变化时, 其隐私风险感知显著波动, 并进一步影响初始信任; 而从业者的隐私风险感知较为稳定, 因此间接效应不显著。这一发现与 Onnasch 等人 (2023) 的观点一致, 即非从业者更多依赖外部提供的信息来形成对系统的判断, 从而导致其初始信任对情境信息更为敏感。

值得注意的是, 研究 1 中我们发现自然条件下隐私风险感知未能显著预测非从业者的初始信任, 而研究 2 通过操纵隐私风险水平发现, 非从业者在隐私风险水平提升时其初始信任显著下降。这一差异提示, 非从业者在自然情境下往往低估隐私风险, 但当风险因操纵而凸显时, 其隐私风险感知显著提高, 进而导致初始信任水平出现明显变化。

综上, 研究 2 验证了隐私风险感知在初始信任形成中的关键作用, 并揭示了从业背景在其中的调节机制。在研究 3 中, 我们将进一步探讨通过增强非从业者对隐私风险的敏感性, 能否有效调节其对自动驾驶汽车的过度信任, 从而促进更为理性的信任校准。

4 研究 3: 隐私风险感知对非从业者初始信任的影响

在研究 2 中, 我们发现在实验操纵条件下, 隐私风险感知的提升会显著降低非从业者的初始信任。然而在自然条件下, 隐私风险感知并不是非从业者初始信任的显著预测因素。这一变化可能源于隐私风险条件的操纵增强了非从业者对隐私风险的敏感性。为了进一步探究非从业者的隐私风险感知和初始信任的关系, 研究 3 依次向非从业者被试施加控制条件(不唤醒隐私风险意识)、低隐私风险条件和高隐私风险条件, 并测量他们对自动驾驶汽车的初始信任, 从而回答研究问题 3。

4.1 方法

4.1.1 被试

根据 G*Power 软件(Faul et al., 2007)的计算结果, 以期望效应量 $f = 0.25$ 、统计检验力 $1 - \beta = 0.95$ 、显著性水平 $\alpha = 0.05$ 进行估算, 推荐的样本量为 $N = 84$ 。我们通过 Credamo 平台招募了 200 名非从业者。为了保证数据质量, 问卷中设置了注意力检测题, 共有 156 份数据通过检测, 通过率为 78%。所有被试均持有有效驾驶执照。学历为普高/中专/技校/职高(1 人, 占 0.2%)的被试由于数量过少, 未纳入分析。最终获得 155 份有效样本, 有效

率为 78%。具体人口统计信息见附表 4。

4.1.2 实验程序与测量

实验分为两个阶段, 第一阶段重复了研究 1 的程序。在此部分, 未对被试的隐私风险意识进行唤醒(控制条件)。被试首先阅读了关于 SAE3 级自动驾驶汽车的介绍, 了解其能力及在特定情况下仍需驾驶员干预的必要性。随后, 被试基于自身经验评估他们的隐私风险感知和初始信任水平。

间隔 1 天后, 第二阶段开始, 实验程序参考研究 2 的设计。被试首先被告知, 为了确保安全性、可靠性和最佳性能, 自动驾驶汽车需要通过传感设备收集并共享大量数据。随后, 被试分别阅读了公司 Z (低隐私风险条件) 和公司 H (高隐私风险条件) 的数据安全管理实践, 并对其隐私风险感知和初始信任进行评估。

被试使用与研究 1 相同的量表对隐私风险感知进行评分(5 点李克特量表, 1 = 完全不同意, 5 = 完全同意), 得分越高代表隐私风险感知越高(本研究中该量表的内部一致性系数 Cronbach's $\alpha = 0.94$)。初始信任同样采用研究 1 的测量项目(5 点李克特量表, 1 = 完全不同意, 5 = 完全同意), 得分越高表明初始信任水平越高(本研究中该量表的内部一致性系数 Cronbach's $\alpha = 0.93$)。

最后, 被试填写人口统计信息, 并获得实验报酬。

4.2 结果

4.2.1 对隐私风险感知的操纵检验

研究 3 中各变量的描述统计结果见表 6。对隐私风险感知进行重复测量方差分析, 结果显示, 隐私风险的操纵显著影响了被试的隐私风险感知, $F(2, 308) = 686.00, p < 0.001, \eta_p^2 = 0.82$ 。进一步的 Bonferroni 事后检验表明, 低隐私风险条件下, 非从业者的隐私风险感知显著低于控制条件($M_{\text{低风险}} = 1.83, SD = 0.80; M_{\text{控制}} = 2.05, SD = 1.02$), $p = 0.005$ 。高隐私风险条件下, 非从业者的隐私风险感知显著高于控制条件($M_{\text{高风险}} = 4.40, SD = 0.55$), $p < 0.001$ 。

表 6 研究 3 中各变量的描述统计

因变量		控制组	低风险	高风险
隐私风险感知	M	2.05	1.83	4.40
	SD	1.02	0.80	0.55
初始信任	M	4.33	4.40	1.90
	SD	0.45	0.43	0.60

4.2.2 隐私风险感知对初始信任的影响

为检验三种风险情境下,隐私风险感知对初始信任的作用,本研究采用线性混合效应模型(LMM)进行分析。使用 jamovi.2.6.26 的 GAMLj 模块进行线性混合效应模型建模,模型中将被试编号作为随机效应,并将隐私风险情境(控制条件、低风险条件和高风险条件,对三种情境进行哑变量编码)和隐私风险感知作为固定效应,同时控制年龄、性别、学历和驾龄。结果显示,模型拟合的 AIC 值为 582.00,各变量对初始信任的预测作用见表 7。

分析结果表明,隐私风险感知显著影响非从业者的初始信任($\beta = -0.33, p < 0.001$)。在低风险情境下,非从业者的初始信任相对于控制条件并没有显著下降($\beta = -0.01, p = 0.942$),但是在高风险情境下,其初始信任相对于控制条件显著降低($\beta = -1.65, p < 0.001$)。进一步的 Bonferroni 事后检验表明,低隐私风险条件下,非从业者的初始信任与控制条件无显著差异($M_{\text{低风险}} = 4.40, SD = 0.43; M_{\text{控制}} = 4.33, SD = 0.45, p > 0.999$)。高隐私风险条件下,非从业者的初始信任显著高于控制条件和低风险条件($M_{\text{高风险}} = 1.89, SD = 0.60, ps < 0.001$ (见图 4))。

4.3 讨论

为了探究在隐私风险波动下非从业者初始信任的变化,研究 3 操纵了隐私风险感知水平。结果表明,相对于控制条件,非从业者在低隐私风险条件下隐私风险感知显著降低,但初始信任并未发生显著变化;但在高隐私风险条件下,隐私风险感知显著增加,并且初始信任显著降低。这一结

果揭示了非从业者对隐私风险信息敏感性的不对称性。

具体而言,在低隐私风险条件下,尽管非从业者的隐私风险感知显著降低,但其初始信任保持稳定。这表明,低风险信息不足以引发显著的信任变化,可能源于非从业者对自动驾驶汽车隐私风险的原有预期较低。换言之,他们对低风险信息不敏感,表现出一定程度的过度信任。

相比之下,在高隐私风险条件下,非从业者的隐私风险感知显著增加,并伴随初始信任显著降低。这一变化表明,非从业者对高风险信息的敏感性较强,尤其当隐私安全受到明显威胁时,他们倾向于急剧降低对自动驾驶汽车的初始信任(Beggiato & Krems, 2013)。这与以往研究的结果一致,当实际风险与使用者的预期不符时,信任往往会出现显著下降(Biassoni et al., 2016; Hoff & Bashir, 2015)。

这种“低敏感—高敏感”的非对称模式揭示了非从业者对隐私风险认知的不准确性及过度信任的倾向。本研究提示可以通过信息引导来改善这种信任偏差:一方面,在低风险情境下,通过适度的风险提示可增强非从业者的风险意识,防止其忽视潜在隐私威胁;另一方面,在高风险情境下,应帮助非从业者形成理性风险判断,避免因恐惧或不确定性而产生过度不信任。因此,通过合理的信息引导与风险提示,能够在不同风险水平下帮助非从业者形成合理的信任水平,从而促进人机关系的安全与稳定发展。

表 7 研究 3 中线性混合效应模型固定效应结果

预测变量	β	95% CI 下限	95% CI 上限	df	t	p
性别(男-女)	0.05	-0.06	0.15	138	0.86	0.392
年龄(30 至 40 岁-18 至 30 岁)	0.13	-0.01	0.26	138	1.78	0.079
年龄(40 至 50 岁-18 至 30 岁)	0.02	-0.20	0.25	138	0.20	0.839
年龄(50 及以上-18 至 30 岁)	0.32	<0.01	0.64	138	2.00	0.048
学历(本科-专科)	-0.08	-0.26	0.09	141	-0.95	0.345
学历(研究生及以上-专科)	-0.12	-0.32	0.09	138	-1.11	0.271
驾龄(1 至 3 年-1 年及以内)	0.29	-0.03	0.61	142	1.77	0.079
驾龄(4 至 6 年-1 至 3 年)	0.07	-0.25	0.38	145	0.41	0.680
驾龄(7 至 9 年-1 年及以内)	-0.03	-0.37	0.31	146	-0.17	0.863
驾龄(10 年及以上-1 年及以内)	-0.01	-0.38	0.36	147	-0.05	0.957
风险情境(低风险-控制)	-0.01	-0.09	0.08	305	-0.07	0.942
风险情境(高风险-控制)	-1.65	-1.80	-1.50	431	-21.85	<0.001
隐私风险感知	-0.33	-0.39	-0.28	451	-12.75	<0.001

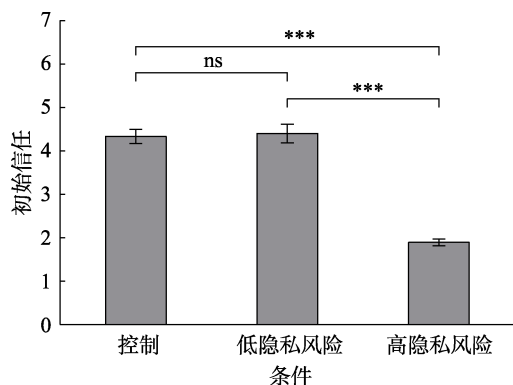


图 4 研究 3 中不同条件下初始信任的变化情况

注: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ 。误差棒为 $M \pm 1 SE$

5 总讨论

本研究基于动态信任框架,系统探讨了自动驾驶汽车的初始信任及其影响因素。参考技术接受模型,本研究构建了一个“四因素框架”,从系统性能、风险感知、社会影响及从业背景四个方面解释自动驾驶初始信任的形成机制与群体差异。通过三项研究,本研究重点考察了隐私风险感知在初始信任形成中的作用及从业背景的调节效应。

研究 1 基于在线问卷调查发现,非从业者的初始信任水平显著高于从业者,表现为高估系统性能、低估系统缺陷及风险,并且更容易受到社会影响。同时,回归分析显示,隐私风险感知与从业背景存在交互作用:隐私风险感知仅能显著预测从业者的初始信任。研究 2 通过操纵隐私风险水平进一步验证了隐私风险感知在初始信任形成中的中介作用,以及从业背景的调节作用。结果表明,隐私风险水平提升显著增强了非从业者的隐私风险感知并降低其初始信任;而从业者的信任水平较为稳定。这一结果揭示了从业者与非从业者在隐私风险变化时的差异化反应。研究 3 在研究 1 和研究 2 的基础上,进一步揭示了非从业者对隐私风险信息的非对称反应:在低隐私风险情境下,其隐私风险感知虽显著提升,但初始信任无显著变化;而在高隐私风险情境下,隐私风险感知显著增加,初始信任水平显著下降。这种“低敏感—高敏感”的反应模式反映出非从业者对隐私风险信息加工的不平衡性。总体而言,本研究发现隐私风险感知是解释不同群体初始信任差异的重要心理机制,而从业背景通过影响风险敏感性与信息加工方式,进一步塑造了信任形成的路径差异。

5.1 非从业者的过度信任倾向

本研究的第一个重要发现是揭示了自动驾驶初始信任水平的群体差异:相较于从业者,非从业者对自动驾驶汽车的初始信任更高,并展现出一定的过度信任倾向。动态信任框架认为,先验经验是初始信任的重要影响因素(高在峰 等, 2021)。研究 1 的结果印证了这一观点:从业者凭借丰富的先验经验,能够更清晰地识别自动驾驶系统的局限性和潜在风险(Balfe et al., 2018),因此在面对自动化技术时表现得更加谨慎(Jin et al., 2020; Sheridan, 2019)。相比之下,非从业者由于缺乏相关经验,往往高估系统性能、低估潜在风险和缺陷,从而更容易表现出过度信任(Booth et al., 2017; Christensen et al., 2019; Robinette et al., 2017)。这一发现为自动驾驶汽车用户群体的差异性提供了新的实证证据,表明在信任形成的早期阶段,从业者通常持更谨慎的态度,而非从业者则可能表现出过度信任。

5.2 初始信任影响因素的群体差异

研究 1 通过在线问卷调查,考察了从业者与非从业者在自然条件下初始信任的影响因素。基于动态信任框架并结合技术接受模型,研究 1 聚焦系统性能感知、系统安全感知、社会影响和从业背景四类变量。结果表明,不同群体在信任形成过程中存在显著差异。

首先,隐私风险感知仅能显著预测从业者的初始信任,而对非从业者无显著影响。这表明相较于从业者,非从业者缺乏必要的知识和经验,容易忽视隐私风险等关键但不易察觉的问题(Kahl, 2021)。这一发现与 Ma 等人(2020)的研究相吻合:成年人由于经验更为丰富,其隐私风险感知与初始信任之间存在显著的负相关关系;然而,在儿童群体中,并未观察到类似的关系。Zhang 等人(2019)发现隐私风险感知与自动驾驶初始信任之间并无显著关系,并将此解释为隐私议题对于中国被试的重要性相对较低。值得注意的是,该研究的样本主要通过方便取样(即在公共停车场面对面招募),且近半被试每周驾车次数不足一次,该样本特征可能是结果不显著的主要原因。此外,本研究结果表明,在自然情境下,普通消费者对隐私风险的敏感度较低,缺乏足够先验经验时往往忽视潜在的隐私侵犯及其带来的损失。综合来看,本研究验证了从业背景与隐私风险感知的交互作用,隐私风险对初始信任的影响依赖于经验差异,非从业者因经验不足而表现出“风险不敏感”。

此外,社会影响、感知有用性和安全风险感知共同预测了个体的初始信任。这表明,初始信任的形成是一个受多种因素影响的复杂过程(高在峰等, 2021; Choi & Ji, 2015; Kraus et al., 2020)。这一结果支持了社会影响(Waung et al., 2021; Zhang et al., 2020)和感知特征(高在峰等, 2021; Lee & See, 2004)在自动驾驶信任中的重要性,并强调了在分析信任形成时,应将从业背景、社会影响与主观感知等因素整合进统一的分析框架中。本研究揭示了从业者与非从业者在影响因素上的差异,为构建更具解释力的信任模型提供了实证依据。

5.3 隐私风险下的信任波动:从业者与非从业者的差异化反应

本研究的第二个重要发现是,隐私风险水平的变化显著影响个体的初始信任,并且从业者与非从业者在反应模式上存在明显差异。研究 2 的结果显示,隐私风险水平的提升会导致从业者的初始信任水平下降,但幅度较为稳定;而非从业者的信任变化更为剧烈。这一发现与先前研究一致。已有研究表明,当系统未明确披露其局限性时,人们往往对其形成较高的初始信任(Biassoni et al., 2016);而当系统表现未能满足用户预期时,信任水平可能会下降(Kraus et al., 2020)。更重要的是,初始信任水平越高,在系统表现未能达到预期时,信任下降的幅度也更大(Merritt & Ilgen, 2008)。研究 2 证明了隐私风险水平通过隐私风险感知对初始信任的影响,并揭示了从业背景的调节作用。

5.4 非从业者对隐私风险信息的非对称反应

研究 3 在前两项研究的基础上,进一步揭示了非从业者在隐私风险信息加工中的非对称性。在低隐私风险条件下,虽然非从业者的隐私风险感知显著降低,但其初始信任水平并未发生显著变化,表明低风险信息不足以引发其信任调整。这可能源于非从业者对隐私风险的预期较低,且对低风险线索不敏感,表现出一定程度的过度信任。

而在高隐私风险条件下,非从业者的隐私风险感知显著增加,并伴随初始信任显著降低。这一结果验证了既有研究:当实际风险超出预期时,个体信任会迅速下降(Biassoni et al., 2016; Hoff & Bashir, 2015)。尤其是在隐私安全受到明确威胁时,非从业者会明显降低对自动驾驶汽车的初始信任(Beggiato & Krems, 2013)。

这一非对称性具有重要实践意义。非从业者的信任水平可以通过提升隐私风险感知得到合理校

准,从而减少其过度信任。当缺乏隐私相关信息时,非从业者往往会对自动驾驶汽车的安全性做出过于乐观的评估(黄心语, 李晔, 2024; Hoff & Bashir, 2015)。因此,通过合理的风险提示,可以在社会层面促进公众的信任更贴近系统的真实水平。

5.5 研究意义与贡献

本研究在自动驾驶汽车初始信任的理解上做出了重要贡献。在理论层面,基于动态信任框架并结合技术接受模型,本研究提出了一个涵盖系统性能、系统安全、社会影响与从业背景的四因素信任框架。该框架揭示了先验经验在信任形成时的关键作用,并为分析不同背景个体在面对相同技术时的差异化信任反应提供了新的解释视角。

其次,本研究通过三项系统研究探讨了隐私风险感知对自动驾驶汽车初始信任的影响,揭示了从业背景在其中的调节作用。研究发现,从业者的初始信任较为稳定,而非从业者则对隐私风险变化更为敏感,尤其在高隐私风险情境下,初始信任显著下降。这些结果扩展了动态信任框架,强调隐私问题在技术接受过程中的关键作用,并为隐私风险管理提供了理论依据(Ma et al., 2020; Zhang et al., 2019)。

本研究还进一步探讨了风险提示在校准非从业者对自动驾驶汽车信任预期方面的潜在作用。研究结果表明,提高非从业者的隐私风险感知可以有效校准过度信任,缩小其与从业者之间的初始信任差距。该发现为信任管理策略,尤其是在提升隐私风险意识方面提供了实证支持。通过合理的风险提示能够有效地调整非从业者的信任预期,使其更加符合现实,这对公共信任的校准至关重要(胡雅萍等, 2025)。

本研究的发现为自动驾驶汽车企业制定提升消费者初始信任的策略提供了重要参考。研究强调,企业应考虑先验经验的差异,特别是在面临技术高度复杂、隐私问题突出的背景下(Zhu et al., 2022)。非从业者往往对隐私安全持乐观态度,然而在发生隐私泄露事件时,他们的初始信任水平会急剧下降。因此,自动驾驶汽车企业应优先加强数据安全和管理措施,以减少因数据隐私事件导致的公众信任下降(Ahmed et al., 2020)。此外,通过社会影响和媒体宣传强化消费者信任,尤其在初期推广阶段,可能成为有效的信任提升策略。

5.6 研究局限与未来方向

本研究仅针对隐私风险进行了干预,尚未考察

其他感知特征在从业者与非从业者初始信任差异中的作用。其中, 感知缺陷量表的信度较低, 未来研究可以开发更具信效度的测量工具, 以获得更稳健的结果。此外, 本研究仅测量了个体的初始信任, 而未涉及接受态度、实际使用行为等更深入的使用意愿指标。鉴于初始信任是技术接受的重要前因态度变量(Zhang et al., 2020), 未来研究可在此基础上, 探索如何更有效的迁移与应用技术接受模型于自动驾驶领域。

另一方面, 本研究并未对从业年限加以限制。研究的核心关注点在于不同先验经验的个体在自动驾驶汽车信任方面存在的差异(高在峰 等, 2021)。尽管从业者通常具备更多先验经验, 但不同工作年限的从业者在知识掌握和技术熟悉度方面可能存在差异。未来研究可进一步控制从业年限或考虑技术熟悉度等因素, 以更精细地揭示先验经验对信任的作用机制。

最后, 本研究的被试样本均来自中国, 这限制了研究发现的普适性。文化差异可能会影响不同国家个体对自动驾驶汽车的理解和信任评估(Gopinath & Narayanamurthy, 2022)。未来研究可进一步考察信任的文化差异, 并探索先验经验如何在不同文化背景下塑造对自动驾驶汽车的信任。

6 结论

本研究基于动态信任框架, 提出了一个包含系统性能、系统安全、社会影响和从业背景的四因素初始信任模型。本文通过三项研究验证了从业背景在初始信任形成中的关键作用, 揭示了隐私风险感知对初始信任的影响及其机制。研究1发现, 非从业者表现出较高的初始信任以及过度信任倾向。自然条件下, 社会影响、感知有用性和安全风险感知是个体初始信任的关键预测因素。隐私风险感知仅能预测从业者的初始信任, 而无法预测非从业者的初始信任。研究2通过操控隐私风险水平, 进一步验证了隐私风险感知在非从业者中起中介作用, 且在高隐私风险情境下, 非从业者的初始信任显著下降。研究3则揭示了非从业者对隐私风险信息的非对称反应: 在低风险情境下信任变化不显著, 而在高风险情境下, 信任水平显著下降。这些结果丰富了自动驾驶汽车初始信任的理论框架, 强调先验经验在信任形成中的核心作用, 尤其是在解释不同背景个体的信任差异方面。研究为自动驾驶企业设计信任校准策略提供了理论和实证依据, 强调了合理

的风险提示有助于在现实环境中建立更加安全和可靠的人机信任基础。

参 考 文 献

- Abraham, H., Lee, C., Brady, S., Fitzgerald, C., Mehler, B., Reimer, B., & Coughlin, J. F. (2017). *Autonomous vehicles and alternatives to driving: Trust, preferences, and effects of age*. Transportation Research Board 96th Annual Meeting.
- Ahmed, S. S., Pantangi, S. S., Eker, U., Fountas, G., Still, S. E., & Anastasopoulos, P. Ch. (2020). Analysis of safety benefits and security concerns from the use of autonomous vehicles: A grouped random parameters bivariate probit approach with heterogeneity in means. *Analytic Methods in Accident Research*, 28, 100134.
- Balfe, N., Sharples, S., & Wilson, J. R. (2018). Understanding is key: An analysis of factors pertaining to trust in a real-world automation system. *Human Factors*, 60(4), 477–495.
- Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour*, 18, 47–57.
- Biassoni, F., Ruscio, D., & Ciceri, R. (2016). Limitations and automation. The role of information about device-specific features in ADAS acceptability. *Safety Science*, 85, 179–186.
- Booth, S., Tompkin, J., Pfister, H., Waldo, J., Gajos, K., & Nagpal, R. (2017). Piggybacking robots: Human-robot overtrust in university dormitory security. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp.426–434). Association for Computing Machinery, New York, NY, United States.
- Carlson, M. S., Desai, M., Drury, J. L., Kwak, H., & Yanco, H. A. (2014). Identifying factors that influence trust in automated cars and medical diagnosis systems. In *2014 AAAI Spring Symposium Series on The Intersection of Robust Intelligence and Trust in Autonomous Systems*. Association for the Advancement of Artificial Intelligence, Washington, D.C., United States.
- Celmer, N., Branaghan, R., & Chiou, E. (2018). Trust in branded autonomous vehicles & performance expectations: A theoretical framework. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1), 1761–1765.
- Choi, J. K., & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction*, 31(10), 692–702.
- Christensen, A. B. H., Dam, C. R., Rasle, C., Bauer, J. E., Mohamed, R. A., & Jensen, L. C. (2019). Reducing overtrust in failing robotic systems. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp.542–543). IEEE.
- Davis, F., Bagozzi, R., & Warshaw, P. (1989). User acceptance of computer technology: A comparison of 2 theoretical models. *Management Science*, 35(8), 982–1003.
- Deloitte China. (2019). *Consumers pump the brakes on autonomous vehicle adoption* | Deloitte China | Newsroom. Deloitte. <https://www2.deloitte.com/cn/en/pages/about-deloitte/articles/pr-consumers-pump-the-brakes-on-autonomous-vehicle-adoption.html>
- Dirsehan, T., & Can, C. (2020). Examination of trust and sustainability concerns in autonomous vehicle adoption. *Technology in Society*, 63, 101361.

- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research* (Vol. 27). Addison-Wesley.
- Gao, Z., Li, W., Liang, J., Pan, H., Xu, W., & Shen, M. (2021). Trust in automated vehicles. *Advances in Psychological Science*, 29(12), 2172–2183.
- [高在峰, 李文敏, 梁佳文, 潘晗希, 许为, 沈模卫. (2021). 自动驾驶车中的人机信任. *心理科学进展*, 29(12), 2172–2183.]
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660.
- Gopinath, K., & Narayanamurthy, G. (2022). Early bird catches the worm! Meta-analysis of autonomous vehicles adoption – Moderating role of automation level, ownership and culture. *International Journal of Information Management*, 66, 102536.
- Ha, T., Kim, S., Seo, D., & Lee, S. (2020). Effects of explanation types and perceived risk on trust in autonomous vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 73, 271–280.
- Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434.
- Hu, Y., Xiao, J., & Liu, Q. (2025). Privacy protection from the perspective of risk perception: A meta-analysis study. *Journal of Intelligence*, 44(1), 198–206.
- [胡雅萍, 肖江艳, 刘千里. (2025). 风险感知视角下个人隐私安全保护——一项元分析研究. *情报杂志*, 44(1), 198–206.]
- Huang, X., & Li, Y. (2024). Trust dampening and trust promoting: A dual-pathway of trust calibration in human-robot interaction. *Advances in Psychological Science*, 32(3), 527–542.
- [黄心语, 李晔. (2024). 人机信任校准的双途径: 信任抑制与信任提升. *心理科学进展*, 32(3), 527–542.]
- Jin, M., Lu, G., Chen, F., & Shi, X. (2020). How driving experience affect trust in automation from level 3 automated vehicles? An experimental analysis. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 1–6. IEEE.
- Kahl, B. (2021). Human biases preventing the widespread adoption of self-driving cars. *ArXiv*, abs/2104.01022.
- Kraus, J., Scholz, D., Stiegemeier, D., & Baumann, M. (2020). The more you know: Trust dynamics and calibration in highly automated driving and the effects of take-overs, system malfunction, and system transparency. *Human Factors*, 62(5), 718–736.
- Kyriakidis, M., Happee, R., & de Winter, J. C. F. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, 32, 127–140.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80.
- Lee, J., & Lee, K. (2022). Polite speech strategies and their impact on drivers' trust in autonomous vehicles. *Computers in Human Behavior*, 127, 107015.
- Lyons, J. B., & Guznov, S. Y. (2019). Individual differences in human-machine trust: A multi-study look at the perfect automation schema. *Theoretical Issues in Ergonomics Science*, 20(4), 440–458.
- Ma, Y., Li, S., Qin, S., & Qi, Y. (2020). Factors affecting trust in the autonomous vehicle: A survey of primary school students and parent perceptions. *Proceedings - 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2020*, 2020–2027. IEEE.
- Madigan, R., Louw, T., Wilbrink, M., Schieben, A., & Merat, N. (2017). What influences the decision to use automated public transport? Using UTAUT to understand public acceptance of automated road transport systems. *Transportation Research Part F: Traffic Psychology and Behaviour*, 50, 55–64.
- Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors*, 50(2), 194–210.
- Narayanan, S., Chaniotakis, E., & Antoniou, C. (2020). Shared autonomous vehicle services: A comprehensive review. *Transportation Research Part C: Emerging Technologies*, 111, 255–293.
- Onnasch, L., Hoesterey, S., & Fahrner, V. (2023). To (dis)agree with automation: Effects of automation levels on trust attitude and trust behavior in high-risk situations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 67(1), 1393–1399.
- On-Road Automated Driving (Orad) Committee. (2023). *J3016 (WIP) Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. SAE International. <https://www.sae.org/standards/content/j3016/>
- Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012). Predicting information technology usage in the car: Towards a car technology acceptance model. *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp.51–58). Association for Computing Machinery, New York, NY, United States.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.
- Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12(4), 531–544.
- Pop, V. L., Shrewsbury, A., & Durso, F. T. (2015). Individual differences in the calibration of trust in automation. *Human Factors*, 57(4), 545–556.
- PricewaterhouseCoopers. (2023). *Digital Auto Report 2023*. PricewaterhouseCoopers. <https://www.strategyand.pwc.com/tr/tr/pdf/digital-auto-report-2023.pdf>
- Pype, P., Daalderop, G., Schulz-Kamm, E., Walters, E., & von Grafenstein, M. (2017). Privacy and security in autonomous vehicles. In D. Watzenig & M. Horn (Eds), *Automated driving: Safer and more efficient future driving* (pp. 17–27). Springer International Publishing.
- Qi, Y., Chen, J., Qin, S., & Du, F. (2024). Human-AI mutual trust in the era of artificial general intelligence. *Advances in Psychological Science*, 32(12), 2124–2136.
- [齐玥, 陈俊廷, 秦邵天, 杜峰. (2024). 通用人工智能时代的人与 AI 信任. *心理科学进展*, 32(12), 2124–2136.]
- Qin, S., Li, C., & Qi, Y. (2024). Trust transfer in autonomous vehicles: The role of warm brand image amid automation failure. In D. Harris & W.-C. Li (Eds), *Engineering psychology and cognitive ergonomics* (pp. 123–140). Springer

- Nature Switzerland.
- Rieger, T., Roesler, E., & Manzey, D. (2021). The imperfect automation schema—Evidence for increased trust in human support agents. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 65(1), 1020–1020.
- Robinette, P., Howard, A. M., & Wagner, A. R. (2017). Effect of robot performance on human–robot trust in time-critical situations. *IEEE Transactions on Human-Machine Systems*, 47(4), 425–436.
- Rupp, M. A., Michaelis, J. R., McConnell, D. S., & Smither, J. A. (2018). The role of individual differences on perceptions of wearable fitness device trust, usability, and motivational impact. *Applied Ergonomics*, 70, 77–87.
- Sheridan, T. B. (2019). Individual differences in attributes of trust in automation: Measurement and application to system design. *Frontiers in Psychology*, 10, 1117.
- Tan, H., Zhao, X., & Yang, J. (2022). Exploring the influence of anxiety, pleasure and subjective knowledge on public acceptance of fully autonomous vehicles. *Computers in Human Behavior*, 131, 107187.
- Tao, D., Yuan, J., Shao, F., Li, D., Zhou, Q., & Qu, X. (2018). Factors affecting consumer acceptance of an online health information portal among young internet users. *CIN: Computers, Informatics, Nursing*, 36(11), 530–539.
- Walker, G. H., Stanton, N. A., & Salmon, P. (2016). Trust in vehicle technology. *International Journal of Vehicle Design*, 70(2), 157–182.
- Wang, K., Wu, J., Sun, Y., Chen, J., Pu, Y., & Qi, Y. (2024). Trust in human and virtual live streamers: The role of integrity and social presence. *International Journal of Human-Computer Interaction*, 40(23), 8274–8294.
- Wang, M., McAuslan, P., & Lakshmanan, S. (2021). Trust and intention to use autonomous vehicles: Manufacturer focus and passenger control. *Transportation Research Part F: Traffic Psychology and Behaviour*, 80, 328–340.
- Wen, Z.L., & Ye, B.J. (2014). Different methods for testing moderated mediation models: Competitors or backups? *Acta Psychologica Sinica*, 46(5), 714–726.
- [温忠麟, 叶宝娟. (2014). 有调节的中介模型检验方法: 竞争还是替补? *心理学报*, 46(5), 714–726.]
- Wintersberger, P., Noah, B. E., Kraus, J., McCall, R., Mirnig, A. G., Kunze, A., Thakkar, S., & Walker, B. N. (2018). Second workshop on trust in the age of automated driving. *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp.56–64). Association for Computing Machinery, New York, NY, United States.
- Wu, M., Wang, N., & Yuen, K. F. (2023). Deep versus superficial anthropomorphism: Exploring their effects on human trust in shared autonomous vehicles. *Computers in Human Behavior*, 141, 107614.
- Xu, W., & Ge, L. (2020). Engineering psychology in the era of artificial intelligence. *Advances in Psychological Science*, 28(9), 1409–1425.
- [许为, 葛列众. (2020). 智能时代的工程心理学. *心理科学进展*, 28(9), 1409–1425.]
- You, S., Qi, Y., Chen, J., Luo, L., & Zhang, K. (2025). Safety trust in intelligent domestic robots: Human and AI perspectives on trust and relevant influencing factors. *Acta Psychologica Sinica*, 57(11), 1951–1972.
- [由姗姗, 齐玥, 陈俊廷, 骆磊, 张侃. (2025). 人与 AI 对智能家居机器人的安全信任及其影响因素. *心理学报*, 57(11), 1951–1972.]
- Yuen, K. F., Wong, Y. D., Ma, F., & Wang, X. (2020). The determinants of public acceptance of autonomous vehicles: An innovation diffusion perspective. *Journal of Cleaner Production*, 270, 121904.
- Zhang, T., Tao, D., Qu, X., Zhang, X., Lin, R., & Zhang, W. (2019). The roles of initial trust and perceived risk in public's acceptance of automated vehicles. *Transportation Research Part C: Emerging Technologies*, 98, 207–220.
- Zhang, T., Tao, D., Qu, X., Zhang, X., Zeng, J., Zhu, H., & Zhu, H. (2020). Automated vehicle acceptance in China: Social influence and initial trust are key determinants. *Transportation Research Part C: Emerging Technologies*, 112, 220–233.
- Zhu, A., Yang, S., Chen, Y., & Xing, C. (2022). A moral decision-making study of autonomous vehicles: Expertise predicts a preference for algorithms in dilemmas. *Personality and Individual Differences*, 186, 111356.
- Zmud, J., Sener, I., & Wagner, J. (2016). *Consumer Acceptance and Travel Behavior Impacts of Automated Vehicles*.

The impact of privacy risk perception on initial trust in autonomous vehicle: Differential responses of professionals and non-professionals

SUN Yifei^{1,2}, LI Xiulan^{1,2}, DU Feng^{3,4} QI Yue^{1,2}

(¹ The Department of Psychology, Renmin University of China, Beijing 100872, China)

(² The Laboratory of the Department of Psychology, Renmin University of China, Beijing 100872, China)

(³ State Key Laboratory of Cognitive Science and Mental Health, Chinese Academy of Sciences, Beijing 100101, China)

(⁴ Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract

This study investigates the role of professional background in shaping initial trust in autonomous vehicles (AVs), with a particular focus on how privacy risk perception influences trust differences between professionals and non-professionals. Previous research has primarily concentrated on the issue of insufficient trust in AVs and sought to enhance trust through improved design and communication. However, excessive trust among ordinary

consumers may equally lead to greater risks and hazards. Therefore, how to calibrate consumers' trust to be closer to that of industry professionals—who are also the system designers—has become a pressing issue. Although existing studies have examined various perceptual factors affecting trust in AVs, there remains a lack of systematic evidence on how prior experiential differences, represented by professional background, shape initial trust.

Drawing on three empirical studies with a total of 1,027 participants, this research systematically examines the mechanisms through which privacy risk perception influences trust formation. Study 1 employed an online survey to compare professionals and non-professionals. Results showed that non-professionals reported significantly higher initial trust in AVs, were more likely to overestimate system performance, underestimate potential risks, and be more susceptible to social influence. Regression analyses further revealed an interaction effect between privacy risk perception and professional background: privacy risk perception significantly predicted professionals' trust levels but had no significant effect on non-professionals. In addition, social influence, perceived usefulness, and safety risk perception jointly predicted initial trust, indicating that trust formation is a complex process shaped by multiple interacting factors.

Study 2 experimentally manipulated privacy risk levels to further explore the interaction between privacy risk perception and professional background. Results demonstrated that increased privacy risk significantly reduced professionals' trust, whereas non-professionals' trust fluctuated more dramatically. A moderated mediation analysis showed that privacy risk level predicted non-professionals' trust through privacy risk perception, but the effect was nonsignificant for professionals. This suggests that professionals' trust is relatively stable, while non-professionals—lacking accurate recognition of privacy risks—are more sensitive to contextual changes.

Study 3 examined the impact of enhancing privacy risk perception. Results indicated that increasing non-professionals' privacy risk awareness heightened their sensitivity to risks, thereby significantly reducing their initial trust. This finding suggests that enhancing privacy risk perception among non-professionals can effectively mitigate excessive trust and narrow the trust gap with professionals.

Taken together, this research reveals that professionals and non-professionals rely on different cognitive pathways in forming initial trust in AVs, and clarifies the interactive mechanism between privacy risk perception and professional background. The findings provide important implications for strategies to strengthen consumer trust in AVs. Specifically, enhancing non-professionals' privacy risk perception helps calibrate excessive trust and narrow the trust gap, while social influence and perceived usefulness also play critical roles in shaping trust. Leveraging social influence and emphasizing the usefulness of AVs may be effective approaches to promoting rational trust. Overall, this research deepens the understanding of trust formation mechanisms in human-machine interaction and offers practical insights for fostering more rational and well-calibrated public trust in autonomous vehicles.

Keywords autonomous vehicle, initial trust, professional background, perceived privacy risk, human-AI trust

附录 1: 实验问卷项目表

附表 1 问卷项目表

变量	项目	内容
感知有用性	问题 1	自动驾驶汽车能够有效地满足我的驾驶需求
	问题 2	在使用自动驾驶汽车时, 我能在车内同时做其它事情, 例如吃东西、看电影、使用手机等。
	问题 3	使用自动驾驶汽车可以降低事故率。
	问题 4	使用自动驾驶汽车可以缓解我的驾驶压力。
	问题 5	当我驾驶状态不好(例如疲劳驾驶、酒驾等)时, 自动驾驶汽车很有用。
感知缺陷	问题 1	自动驾驶汽车的驾驶表现相对保守, 速度较慢。
	问题 2	我担心自动驾驶系统出现故障, 维修难度较高。
安全风险感知	问题 1	我担心自动驾驶汽车技术的整体安全性。
	问题 2	我担心自动驾驶汽车的故障可能导致事故。
隐私风险感知	问题 1	我担心自动驾驶汽车收集太多我的个人信息。
	问题 2	我担心自动驾驶汽车未经我授权就把我的个人信息用于其他目的。
	问题 3	我担心自动驾驶汽车未经我授权就把我的个人信息共享给其他机构。
社会影响	问题 1	意见对我很重要的人也将会喜欢自动驾驶汽车。
	问题 2	总的来说, 我喜欢的人会鼓励我使用自动驾驶汽车。
	问题 3	如果我的朋友或家人使用自动驾驶汽车, 我更有可能使用它。
初始信任	问题 1	自动驾驶汽车是可靠的。
	问题 2	自动驾驶汽车是可信赖的。
	问题 3	总的来说, 我可以相信自动驾驶汽车。

附录 2: 各研究被试人口学信息表

附表 2 研究 1 中被试的人口学信息(N = 689)

基本信息	分类标准	从业者(N = 165)		非从业者(N = 524)	
		人数	比例(%)	人数	比例(%)
性别	男	138	20.0	226	32.8
	女	27	3.9	298	43.3
年龄段	18-30 岁	69	10.0	251	36.4
	31-40 岁	75	10.9	211	30.6
	41-50 岁	19	2.8	44	6.4
	51 岁及以上	2	0.3	18	2.6
	专科	2	0.3	40	5.8
学历	本科	80	11.6	420	61.0
	硕士及以上	83	12.0	64	9.3
驾龄	1 年及以下	16	2.3	27	3.9
	1-3 年	30	4.4	126	18.3
	4-6 年	40	5.8	188	27.3
	7-9 年	36	5.2	106	15.4
	10 年及以上	43	6.2	77	11.2

附表 3 研究 2 中被试的人口学信息($N = 183$)

基本信息	分类标准	从业者($N = 165$)		非从业者($N = 524$)	
		人数	比例(%)	人数	比例(%)
性别	男	78	42.6	42	23.0
	女	12	6.6	51	27.9
年龄段	18-30 岁	14	7.7	44	24.0
	31-40 岁	38	20.8	39	21.3
	41-50 岁	34	18.6	9	4.9
	51 岁及以上	4	2.2	1	0.5
学历	专科	2	1.1	5	2.7
	本科	38	20.8	80	43.7
	硕士及以上	50	27.3	8	4.4
驾龄	1 年及以下	1	0.5	6	3.3
	1-3 年	8	4.4	13	7.1
	4-6 年	12	6.6	41	22.4
	7-9 年	14	7.7	25	13.7
	10 年及以上	55	30.1	8	4.4

附表 4 研究 3 中非从业者被试的人口学信息($N = 155$)

基本信息	分类标准	人数	比例(%)
性别	男	58	37.4
	女	97	62.6
年龄段	18-30 岁	64	41.3
	31-40 岁	70	45.2
	41-50 岁	15	9.7
	51 岁及以上	6	3.9
学历	专科	13	8.4
	本科	119	76.8
	硕士及以上	23	14.8
	1 年及以下	4	2.6
驾龄	1-3 年	23	14.8
	4-6 年	56	36.1
	7-9 年	45	29.0
	10 年及以上	27	17.4