

信息接触与威胁情境:公众认知极化成因探析

韩亦坤,魏钰明,苏 竣,张 芳

(清华大学公共管理学院,北京 100086)

摘 要:伴随社会智能化转型,精准推荐等新技术引发的认知极化问题日益成为社会治理的重要议题。研究基于行为科学视角,探讨社会环境中的威胁情境如何与智能精准推荐下的信息接触产生交互作用,影响公众认知极化进程。通过设计2(无威胁情境 VS 威胁情境)×2(同质信息 VS 异质信息)的双因素被试间实验,研究发现接触同质信息会加剧公众认知极化,而接触异质信息则有助于缓解极化;存在威胁情境时,接触异质信息对公众认知极化的缓解作用会减弱。结果表明,对智能社会中公众认知极化问题的治理,不仅要破除精准推荐带来的信息茧房,还需重视社会环境中的威胁情境对公众认知的影响。

关键词:认知极化;精准推荐;信息茧房;威胁情境;智能社会治理

中图分类号:G301

文献标识码:A

智能技术的蓬勃发展推动人类迈向智能社会,引发生产生活方式、组织结构和文明秩序的深刻变革,催生出大量新问题、新挑战^[1]。治理智能社会中的新问题,已成为重要且紧迫的研究议题^[2]。现有研究聚焦于算法滥用、数字鸿沟、技术倦怠和技术信任^[3]等智能技术的风险治理。而随着精准推荐广泛使用,智能技术对公众的信息接触、心理认知和行为偏好造成的变革性影响愈加显著。精准推荐导致公众个体接触的信息高度同质,陷入信息茧房^[4],引发认知极化,加剧社会分裂风险,为智能社会治理带来巨大挑战^[5]。因此,深挖智能精准推荐下信息接触对公众认知极化的影响机制,探索有效治理对策,对智能社会的健康发展具有重要意义。

传统关于公众认知态度的分析多基于“理性人”假设,认为公众通过“成本-收益”的权衡取舍形成认知与态度^[6];许多智能技术情境下的研究也延续该假设,忽略心理机制的复杂影响^[7]。新近的一些研究从心理学与行为科学出发,开始关注情感偏好、身份意识、意识形态、人格特征等特质对公众信息接触与认知极化的影响^[8],但对公众外在社会环境感知的潜在影响关注较少。经典心理学研究表明,对社会环境的威胁感知是影响其认知态度的关

键因素^[9]。尤其在智能技术发达的网络空间,负面情绪与威胁内容更易被关注传播,加剧公众对社会环境的威胁感知和紧张情绪,导致公众认知和态度的极化^[10]。因此,当前研究亟需弥补对“威胁情境”与“信息接触”如何共同作用于公众认知的探讨,为有效治理精准推荐等技术引发的信息茧房和认知极化问题,提供更加科学全面的理论支撑。

本研究基于行为科学理论,通过实验方法分析社会环境中的威胁情境如何与智能精准推荐下的信息接触共同影响公众认知极化。研究以“行人闯红灯曝光”中人脸识别技术的应用为例,构建2(无威胁情境 VS 有威胁情境)×2(同质信息 VS 异质信息)的双因素被试间实验,通过线上渠道招募1281位实验者开展实验研究。结果显示,同质信息接触显著加剧公众观点的极端化程度,加剧认知极化;异质信息接触会降低公众观点的极端化程度,缓解认知极化。存在威胁情境时,异质信息接触对公众认知极化的缓解作用减弱,且男性、高收入水平的公众对威胁情境更加敏感。由此可见,对网络空间公众认知极化问题的治理,不仅需打破精准推荐带来的信息茧房,让公众接触到多元信息;也要防止渲染威胁情境、制造紧张和不安情绪的内容大规模传播发

收稿日期:2025-04-10;修回日期:2025-07-04

基金项目:国家重点研发计划项目(2022YFC3303100);国家自然科学基金资助项目(72104126,72204138)

作者简介:韩亦坤(2001-),女,博士研究生。魏钰明(1993-),男,助理研究员。苏竣(1965-),男,教授、博士生导师。张芳(1987-),女,副教授、博士生导师,通讯作者,E-mail:zhangfang2021@tsinghua.edu.cn。

群,塑造风清气正的网络环境,提升人民群众的获得感、幸福感、安全感。

1 理论基础与研究假设

1.1 精准推荐下的信息茧房与认知极化

极化概念起源于 20 世纪 50 年代心理学领域,原指群体讨论中意见趋向极端的群体极化现象。近年来,极化研究已扩展至政治议题、社会价值观念等领域,指人们观点的极端化与分歧的增大,并分为意识形态极化、认知极化、情感极化、社会互动极化等维度^[11]。其中西方学者的研究主要关注意识形态和政治议题上的两极对立;而在中国网络空间中,极化更多体现为公众观点的极端程度增大,即认知极化。随着精准推荐等智能技术发展,认知极化现象在网络空间愈发凸显,成为智能社会治理面临的一大挑战;对其机制的理解与治理方案的探讨,已成为当前社会治理的重要议题。

信息接触是影响认知极化的核心因素之一。公众接触的信息可以分为同质信息和异质信息两类:同质信息指与个体已有观点一致或相似的信息;异质性信息指与个体已有观点相悖或不同的信息^[12]。接触与自身原有观点相符的信息会加剧认知极化,而接触与观点相异的信息则有助于缓解极化^[13]。随着精准推荐等技术发展,网络媒体为增加用户黏度,会有选择地向公众推送符合其观点和兴趣的信息,将公众困在由同质信息构筑的“信息茧房”中,强化公众已有观点,导致认知极化^[14]。另一方面,接触与已有观点不同甚至相反的异质信息,通常被认为是缓解认知极化的重要方式。例如,一项基于美国脸书(Facebook)平台的实验研究向参与者随机提供不同党派的媒体订阅信息,发现接触与原有观点相反的新闻降低了参与者的认知极化水平^[15]。也有研究发现,接触更多异质信息的个体政治态度更温和,接触同质信息较多的个体政治态度更极端^[16]。因此,本研究假设:

H1-a:接触同质信息能够加剧公众认知极化。

H1-b:接触异质信息能够缓解公众认知极化。

1.2 威胁情境对信息接触与认知极化的影响

除了“信息茧房”与“回音室效应”,行为科学研究也关注公众对外部的威胁感知如何影响其情感、认知和行为。心理学中,威胁情境指可能损害个体的健康、财产、情感、自身认知、观点信念的情境,包

括自然灾害、疾病风险、经济危机、网络暴力、犯罪和移民问题等^[17]。威胁情境可能引发公众情绪波动,影响公众对信息的判断认知,促使公众支持更极端的意识形态、政策或政党^[18]。

威胁情境能够通过占用认知资源、引发负面情绪和带来不确定性,影响公众对同质或异质信息的消化、理解和吸收,进而影响认知极化。认知层面,有限认知资源理论表明,威胁情境对注意力的消耗会占用认知资源,导致个体思维简化,形成非黑即白的观点,使公众对原有认知过度自信,对不同观点容忍度降低,最终带来认知极化。情绪层面,威胁情境增强人们的不确定性感受,引发恐惧、焦虑等负面情绪^[19]。此时,不同观点将被视作对自身的挑战,公众更难容忍异质信息,甚至攻击意见不同者,以维护自身信念,满足确定性需求。这些研究表明,威胁情境带来的认知资源占用,不确定性与负面情绪,可能促使公众采取一系列防御和补偿措施,包括厌恶排斥异质信息、认可偏爱同质信息等,导致接触同质信息带来的认知极化进一步放大,接触异质信息对极化的缓解效果被削弱。基于此,本研究提出如下假设:

H2-a:与无威胁情境相比,公众在有威胁情境下接触同质信息后,观点极端程度的增幅更高。

H2-b:与无威胁情境相比,公众在有威胁情境下接触异质信息后,观点极端程度的降幅更低。

2 研究方法和数据

2.1 实验设计

本研究通过线上情境实验的方法检验上述假设。如表 1 所示,研究采用 2×2 组间设计,参与者随机分为 4 组,接受不同的威胁情境和信息接触干预。根据现有文献,威胁情境能够影响公众对异质、同质信息的理解,进而影响认知极化。因此本实验中的 4 组被试先分别接受有、无威胁情境干预,再接受同质或异质信息接触干预。

实验流程如图 1 所示,共分为 4 个步骤:第一,采集被试的人口学信息,包括年龄、性别、受教育情况、年收入水平、职业类型等。第二,使用八分量表对被试的观点极端程度开展前测。第三,将被试随机分配至 2(无威胁情境 VS 威胁情境)×2(同质信息 VS 异质信息)条件组合的 4 组。在威胁情境操纵上,随机分配至有威胁情境组的被试,需要观看

45 – 60 秒的自然灾害和人为意外视频,构建威胁情境,激发外部威胁感知;随机分配至无威胁情境组的被试则需要观看同等时长的自然风景和娱乐活动视频。在信息接触操纵上,随机分配到同质信息组的被试需接触同质信息,即阅读支持其原有观点倾向

的 3 条评论,每条评论的阅读时间不少于 5 秒;随机分配到异质信息组的被试需接触异质信息,即阅读反对其原有观点倾向的 3 条评论,阅读时间要求相同。第四,再次使用八分位量表测量各组被试在威胁情境操纵与信息接触操纵后的观点极端程度。

表 1 实验条件与分组
Table 1 Experimental conditions and grouping

	实验组 1	实验组 2	实验组 3	实验组 4
威胁情境	无威胁	无威胁	有威胁	有威胁
信息接触类型	同质信息	异质信息	同质信息	异质信息
样本数	327	316	314	324

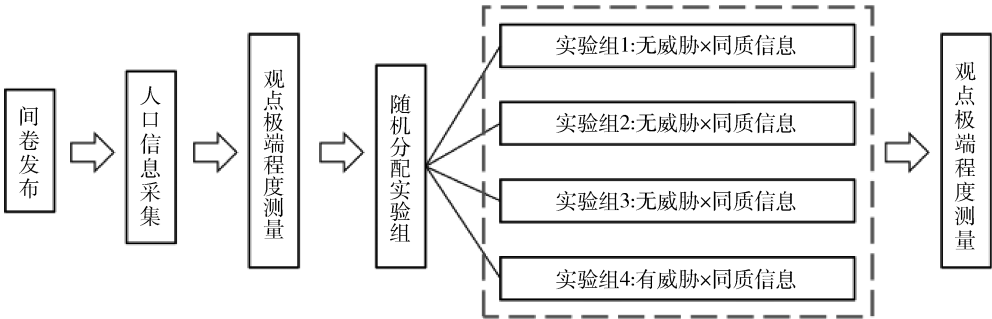


图 1 实验流程示意图
Figure 1 Experimental procedure flowchart

本研究以中国普通公众作为研究对象,通过“Credamo 见数”在线调查平台的“在线随机实验”面向全国招募实验参与者,收集实验数据。与传统的线下便利抽样方法或线上高校学生样本收集方式不同,线上网络平台收集的数据更具普遍性和样本多样性,在一定程度上解决了样本代表性和外部效度的问题。

本研究于 2024 年 6 月 4 日至 2024 年 6 月 13 日期间进行,共发放问卷 1499 份。剔除作答时间异常、多题目选项一致与极端取值的无效问卷后,共获得有效问卷 1281 份,有效回收率 85.46%。样本描述性统计结果如表 2 所示。被试性别分布基本均衡(男性占 48.2%,女性占 51.8%),平均年龄 30.88 岁($SD = 10.50$)。私企工作者占据半数(50.2%),其次是学生(23.3%)。本科及以上学历占比 86.2%,受教育程度总体较高。年收入低于 10 万元以下的样本占比 48.2%,平均主观社会阶层(1 = 最底层,10 = 最高层)均值为 5.26($SD = 1.52$)。

本研究运用方差分析中的 One – way ANOVA

和事后 LSD 多重比较以及卡方检验进行平衡性检验,验证样本分配的随机性。表 3 为平衡检验结果,4 个实验组的协变量均值间没有显著差异($p > 0.253$),4 组样本实现了随机化分配。

2.2 实验背景:“智能行人闯红灯取证系统”

本研究以“智能行人闯红灯取证系统”为实验背景,以被试对该系统的观点极端程度测量公众认知极化水平。2018 年以来,该系统在全国多地试行,通过人脸识别技术获取闯红灯者的图像信息,识别提取违法者的人脸图像,匹配数据库中的信息,在路口设置大屏幕,曝光闯红灯者的面部图像与部分身份信息,起到警示作用。然而,该类曝光惩罚方式引发了广泛争议。支持者认为,曝光车牌号以惩罚机动车交通违法行为存在执法先例,这种曝光也有助于提高行人守法意识,维护了公共安全。反对者认为,该惩罚的法律依据不足,人脸识别技术可能出错,公开曝光个人信息更涉嫌侵犯个人隐私。

本研究选择该技术应用作为实验背景,主要基于以下考虑:第一,该技术曾在多城市广泛应用并引

发讨论,易于被试理解并带入实验情景。第二,人脸识别技术在社会治理中的探索体现了隐私保护与安全保障的价值取向冲突,公众态度分化明显,适合作为认知极化的研究材料。第三,智能技术在城市治

理领域的应用与研究逐渐增多^[20],但有关技术使用中隐私和安全如何平衡的公众认知研究依然较少^[21],本研究可为相关政策制定提供参考。

表 2 主要人口统计学变量描述性分析结果
Table 2 Descriptive analysis results of major demographic variables

人口统计学变量(定类)		频数	百分比
性别	男	618	48.20%
	女	663	51.80%
职业	政府单位	38	3%
	事业单位	94	7.30%
	国企	147	11.50%
	私企	643	50.20%
	教育科研单位	14	1.10%
	学生	298	23.30%
	无业/失业	8	0.60%
	其他	39	3.00%
	小学及以下	0	0%
	初中	6	0.50%
最高学历	普高/中专/技校/职高	47	4.10%
	专科	123	13.70%
	本科	903	70.50%
	硕士	190	14.80%
	博士	12	0.90%
	0-5000 元	107	8.40%
	5000-1 万元	79	6.20%
	1 万-3 万元	100	7.80%
年收入水平	3 万-5 万元	75	5.90%
	5 万-10 万元	254	19.80%
	10 万-20 万元	394	30.80%
	20 万-50 万元	224	17.50%
	50 万-100 万元	44	3.40%
	100 万元及以上	4	0.30%
人口统计学变量(定量)		均值	标准差
年龄	整数年龄	30.878	10.496
主观社会阶层	1 = 最底层,10 = 最顶层	5.263	1.516

2.3 实验干预

(1) 威胁情境

实验采用短视频材料操纵威胁情境,该方法在心理学中常用于引发被试的认知或情绪改变,特别适用于线上实验,能够增强沉浸体验,提高数据

质量^[22]。

威胁感知通常包括两个维度:严重性(威胁造成的影响程度)与易感性(威胁发生的概率大小)^[23]。自然灾害(如洪水、地震)伤害大,严重性高;人为意外(如车祸、火灾),更贴近日常,易感性

高。因此,研究同时采用这两类视频材料,以实现
对威胁情境的多维度操纵。考虑到“行人闯红灯曝
光”议题与交通安全、个人隐私高度相关,为避免干

扰,本实验在视频选择上规避了此类内容。如图 2
所示,自然灾害视频内容展示洪水损毁房屋、道路的
场景;人为意外视频为电动车爆燃引发火灾的片段。

表 3 平衡性检验
Table 3 Test for baseline balance

变量均值	实验组 1	实验组 2	实验组 3	实验组 4	F / Chi2	P
性别	0.47	0.47	0.48	0.5	0.3	0.826
年龄	30.4	30.78	31.31	31.03	0.434	0.729
受教育水平(最低年限)	16.15	16.35	16.21	16.19	1.008	0.389
年收入	145672.78	149050.63	164148.09	163989.2	1.097	0.349
主观社会阶层	5.17	5.35	5.19	5.35	1.362	0.253
职业	-	-	-	-	15.0	0.823

注:实验组 1 为“无威胁-同质信息”组;实验组 2 为“无威胁-异质信息组;实验组 3 为“有威胁-同质信息”组;实验组 4 为“有威胁-异质信息”组。受限于分类变量属性,未展示各组“职业”变量的均值。

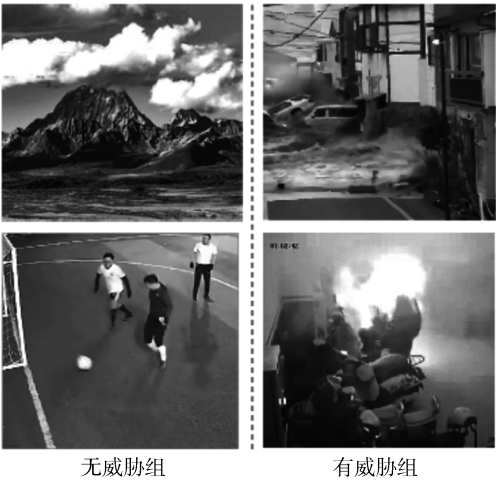


图 2 有威胁组、无威胁组使用的视频材料
Figure 2 Video materials presented to the threat and no-threat groups

对照组则观看无威胁情境视频,包括自然风景
(中国自然风光纪录片选段)与人类活动(比如滑滑
板、踢足球),不涉及危险情境,营造轻松氛围。所
有视频内容来源于网络公开材料。

(2)信息接触

为避免框架效应并提升内部效度,实验中呈现
的支持与反对“行人闯红灯曝光”观点,均涵盖法律
合规、安全、隐私保护、技术效用等维度,并经预实验
验证其倾向性。同质信息指支持或认同被试自身原
有观点的信息,异质信息指挑战或反对被试原有观
点的信息。例如,若被试原本支持曝光措施,支持该

措施的观点为同质信息,不支持该措施的观点为异
质信息。具体内容如表 4 所示。每位被试需观看三
条同质(或异质)信息,信息按随机顺序呈现,均为
文字形式,要求至少观看 5 秒。

为验证信息接触材料的态度倾向,研究进行了
预实验。共招募 79 名被试,测量其对“行人闯红灯
曝光”的原始态度,采用 8 分位量表(1 = 非常反对,
8 = 非常同意)对呈现的支持与反对观点进行评分。
结果显示,支持曝光者对支持评论的平均评分为
6.67,反对评论为 3.50,差异显著($t(77) = 9.42$,
 $p < 0.001$);反对曝光者对支持评论的平均评分为
3.63,对反对评论为 6.20;同样差异显著($t(77) =$
 $6.82, p < 0.001$)。说明支持、反对信息材料确实存
在显著的观点差异,可用于呈现对应观点。

2.4 变量测度

本研究用被试观点极端程度作为公众认知极化
的衡量指标,具体通过其观点数值与中立值的绝对
偏离值计算^[24],范围在 0.5 到 3.5,数值越高表示立
场越极端。被试需在实验前后分别回答“对于曝光
行人闯红灯行为这一举措的态度如何”,并在 8 分
量表上作答(1 = 非常反对,8 = 非常支持)。

核心自变量包括威胁情境与信息接触,均为哑
变量。威胁情境表示个体是否受到威胁情境的操
纵,1 = 有威胁情境,0 = 无威胁情境。信息接触表示
个体接受的信息类型,1 = 接触同质信息,0 = 接触异
质信息。

表 4 信息接触材料的内容

Table 4 Content of the information exposure materials

支持类	A1:那些闯红灯的人就是图方便,把自己的生命安全,别人的生命安全都不当回事。交警又没空天天蹲路口抓闯红灯的人,这种人早该被好好教育教育了!我看曝光人脸,让他们丢丢脸,是当前最实用的方法。
	A2:交通法中对于交通信号灯的规则有明确规定,红灯停绿灯行是法定规则,遵纪守法是公民基本要求。本来就应该根据交通法规走路行车,这个系统在降低闯红灯行为频率上效果确实不错。违法行为被公检系统发现并曝光,蛮正常的。想避免丢脸就改正,法律是用来规范人的行为的。
	A3:支持!我家这边闯红灯的人特别多,正需要这种系统好好警示一下!听说身边安装了这个设备的路口确实闯红灯的人变少了,公开个人信息确实比较社死,但这样才能有效教育这些不守交规的人呀!
反对类	B1:虽然这个方法确实有用,但是把个人的脸放上去循环公示过分了吧。教育个人遵守交规还有很多手段,比如定位到个人,短信通知或者罚款,没必要让人“社死”。这样公示人家的脸根本没考虑会对这个人有什么恶劣影响啊。
	B2:罚款可以,但是在大街上公示人脸抓拍,是不是太侵犯隐私了……毕竟闯红灯的行人个体对公共安全的伤害还是很小的,那些危害性更大的闯红灯的大车也只是罚款而已,凭什么单单要把一个闯红灯拿出来曝光给大庭广众呢?
	B3:也就是闯个红灯,这样的惩罚有必要吗?人脸和个人信息是公民的隐私,怎么还有交警把闯红灯的人的信息记录下来发到网上接着曝光啊?在当地大屏曝光还不够是吗?

此外,研究控制了多项人口学变量,包括年龄、性别、受教育程度、年收入水平、职业类型和主观社会阶层。已有研究指出,女性威胁感知更敏感,反应更温和,较少支持强硬政策^[25]。教育水平高的人面对威胁更理性,极端观点更少。社会经济地位较低的人群更可能支持极端或民粹观点。主要变量描述性分析结果如表 5 所示。

表 5 主要变量描述性分析结果

Table 5 Descriptive analysis results of key variables

变量名称	测量方式	样本量	均值	标准差	最大值	最小值
信息接触前观点极端程度	信息接触前的观点态度取值减去中立值(4.5)的绝对值	1281	2.254	0.96	3.5	0.5
信息接触后观点极端程度	信息接触后的观点态度取值减去中立值(4.5)的绝对值	1281	2.132	1.029	3.5	0.5
信息接触	1 = 同质信息;0 = 异质信息	1281	0.5	0.5	1.0	0.0_
威胁情境	1 = 有威胁;0 = 无威胁	1281	0.498	0.5	1.0	0.0

3 研究结果

3.1 信息接触对公众认知极化程度的影响

实验结果表明,信息接触对公众观点极端程度影响显著。如图 3 所示,信息接触前,同质信息组与异质信息组的观点极端程度不存在显著差异;信息接触后,异质信息组的观点极端程度($M = 1.64$, $SD = 0.95$)显著低于同质信息组($M = 2.62$, $SD = 0.86$, $t(1279) = 19.48$, $p < 0.001$)。整体而言,接触异质信息后,观点的极端程度($M = 1.64$, $SD = 0.95$)显著低于接触异质信息前($M = 2.24$, $SD = 0.97$, $t(1279) = 11.21$, $p < 0.001$);接触同质信息后($M = 2.62$, $SD = 0.86$),观点的极端程度显著高于接触同质信息前($M = 2.27$, $SD = 0.96$, $t(1279) = 11.57$, $p < 0.001$)。

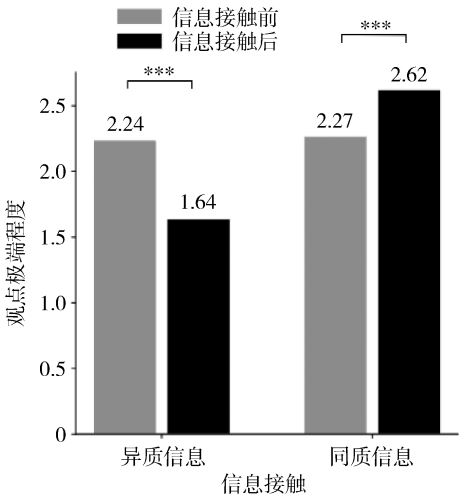


图 3 信息接触对观点极端程度的影响

Figure 3 The effect of information exposure on opinion polarization

说明接触同质信息能够显著加剧公众认知极化,接触异质信息能够显著缓解认知极化,H1-a 与 H1-b 得到验证。

3.2 不同威胁情境下信息接触对认知极化的影响

在异质与同质信息接触影响公众认知极化的过程中,威胁情境产生的作用不同。如图 4 所示,在接触异质信息后,有威胁组(组 4)的观点极端程度($M = 1.76, SD = 0.96$)显著高于无威胁组(组 2)(M

$= 1.52, SD = 0.93, t(638) = 3.27, p = 0.001$);但在接触同质信息后,有威胁组(组 3)的观点极端程度与无威胁组(组 1)没有显著差异($p = 0.157$)。说明在有威胁情境下,异质信息接触缓解公众认知极化的效果被削弱了,验证了假设 H2-b。但在有威胁情境下,同质信息接触加剧公众认知极化的效果与无威胁情境下没有显著差异,假设 H2-a 未能得到验证。

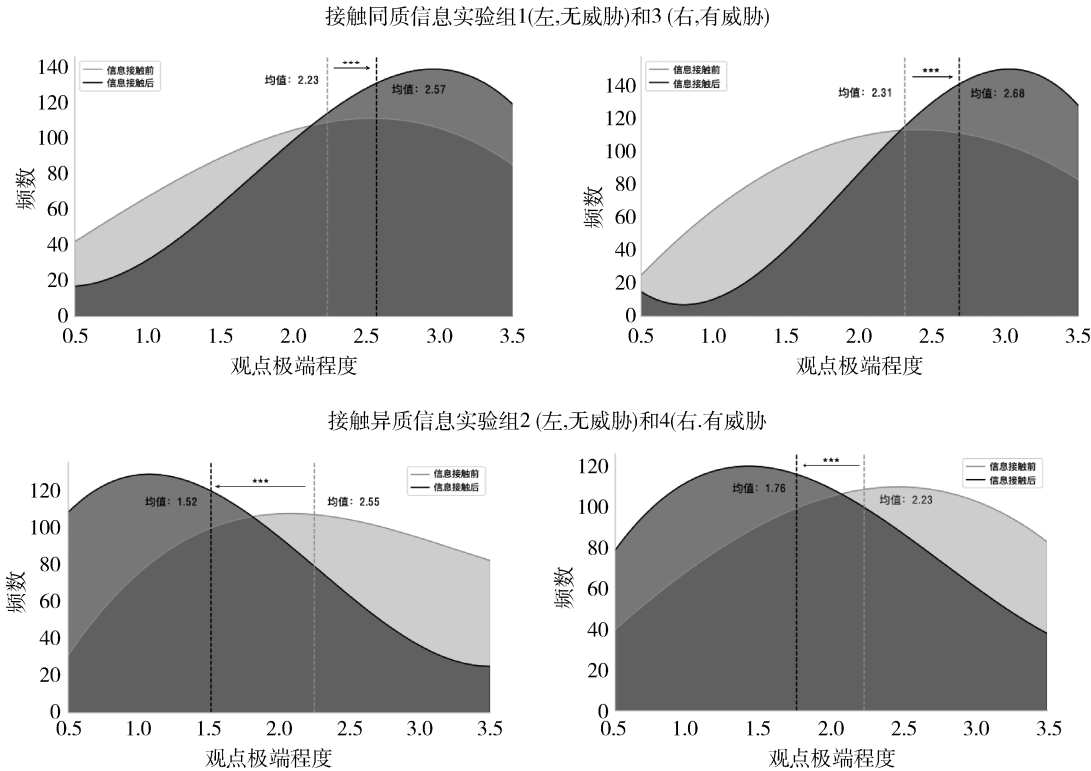


图 4 不同分组接触信息前后观点极端程度频率分布
Figure 4 Frequency distribution of opinion polarization before and after information exposure across groups

3.3 稳健性检验

为检验实验结果的稳定性,研究采用不同因变量构造 OLS、Logic 与 Probit 回归模型,分析威胁情境与信息接触的交互效应是否影响观点极端程度的变化。在 OLS 模型中,以观点极端改变值(接触信息后减去接触前的极端程度)为因变量,正值表示极化增强,负值表示极化减弱。在 Logic 与 Probit 模型中,将因变量进行二分处理,极化增强为 1,极化不变或减弱为 0。自变量包括信息接触(0 = 异质信息,1 = 同质信息)和威胁情境(0 = 无威胁,1 = 有威胁)两项哑变量,并引入二者交互项。模型同时控制性别、年龄、主观社会阶层、年收入水平(对数)、受教育水平(年限)等变量。

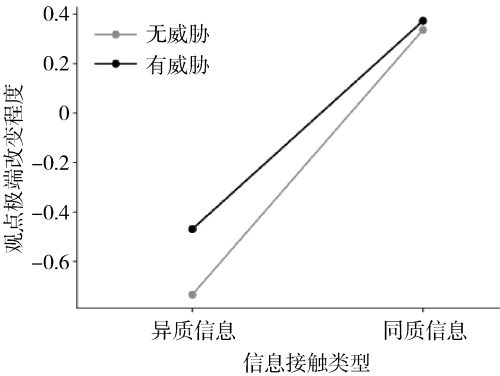


图 5 不同威胁情境与信息接触下观点极端改变程度
Figure 5 Changes in opinion polarization across threat contexts and types of information exposure

如表 6 所示,3 种模型的分析结果均显示,威胁情境、信息接触与其交互作用都能影响观点极端改变程度。威胁情境使信息接触后的观点极端改变程

度显著增大;显著削弱异质信息对认知极化的缓解作用,但不影响同质信息加剧极化的效果。这与前文研究结果相符,验证了研究结论的稳健性。

表 6 威胁情境与信息接触类型对观点极化改变程度的交互作用

变量:观点极端改变程度	模型 1:OLS 回归			模型 2:Logit 回归			模型 3:Probit 回归		
	coeff	LLCI	ULCI	coeff	LLCI	ULCI	coeff	LLCI	UI
威胁情境(0 = 无威胁,1 = 有威胁)	0.265 ***	0.133	0.397	1.039 **	0.424	1.654	0.512 **	0.217	0.8107
信息接触(1 = 同质,0 = 异质)	1.071 ***	0.939	1.202	2.519 **	1.955	3.084	1.370 ***	1.096	1.645
威胁情境 * 信息接触	-0.229 *	-0.415	-0.043	-0.920 **	-1.611	-0.228	-0.438 *	-0.792	-0.084
常数项	-0.734 ***	-0.828	-0.640	-2.999 ***	-3.518	-2.481	-1.670 ***	-1.907	-1.433
N	1281	1281	1281						
R ² / Presudo R ²	0.248	0.136	0.136						
Adjust - R ² / LLR p	0.246	< .001	< .001						

注:*** 表示 P<0.001,** 表示 P<0.01,* 表示 P<0.05。括号中的数据为标准误,LLCI = 95% 置信区间的下限,ULCI = 95% 置信区间的上限。

3.4 异质性分析

基于人口统计学数据的异质性分析揭示了,对于不同性别和收入水平的公众,威胁情境与信息接触对认知极化的影响不同^①。与以往研究类似,本研究发现女性的观点极端程度相较于男性更低^[25]。相较于女性,男性对威胁情境更为敏感,更容易在威胁情境下排斥异质信息,加剧认知极化(t(309) = 2.97, p = 0.003)。意外的是,社会经济地位更高(年均收入等于或高于 10 万元)的群体更容易受到威胁情境的影响,抗拒异质信息,加剧极化(t(342) = 3.30, p = 0.001)。这可能由于收入较高的公众更易感受到外界环境对自身社会经济地位的威胁,具有更高的确定性需求与更保守的价值观念^[26]。

通过对公众原有态度进行分析发现,支持曝光的群体初始观点更极端,但在面对异质信息时,观点也容易缓和(t(638) = 2.60, p = 0.009),更不易极化(t(639) = 3.41, p = 0.001)。然而,如果在威胁情境下接触异质信息,原本支持曝光者的观点缓和效果被显著削弱(t(564) = 3.445, p < .001);而原本反对曝光者的态度不受威胁影响(t(72) = 0.539, p = .592)。这一现象可能有两方面解释。一是“天花板效应”^[27],原本支持者的观点更极端,也就更易缓和而非极化。二是不同态度群体的技术

风险感知差异。关于技术风险认知的文章指出,对新兴技术应用的风险认知会显著降低公众的技术接受度^[28]。原本支持曝光的公众技术风险感知较弱,更易在接收技术负面信息后动摇立场;但本就技术风险敏感,反对曝光的公众则在威胁情境下态度更稳定。

4 结论与讨论

4.1 结果讨论

随着智能推荐等新技术迅速发展,信息茧房引发的认知极化已成为智能社会治理的重要议题。本研究以“行人闯红灯曝光”这一智能技术的应用为切入,探讨社会环境中的威胁情境如何与不同类型的信息接触相互作用,推动公众认知极化的形成和演化。研究发现,接触同质信息加剧认知极化,接触异质信息有效缓解认知极化。威胁情境则在影响过程中发挥关键的调节作用,能够显著削弱异质信息对极化的缓解效果。此外,不同性别和收入水平的公众在作用机制中表现出明显差异。

本研究表明,威胁情境作为关键情景变量,能够干扰信息接触效果,削弱异质信息缓解认知极化的作用。因此,智能社会治理认知极化不能仅依赖信

① 因文章篇幅限制,更详细的分析结果省略。

息多样性,还需充分重视公众对外部社会环境的感知。在社交媒体高度发达、智能推荐技术广泛应用的背景下,带有威胁性的信息更容易传播发酵,营造线上的威胁情境。威胁情境与高同质化的信息相互叠加,可能进一步加剧公众对不同观点的抵触,加剧认知极化。基于此,行为干预策略应更关注对威胁感知的缓解,通过情感管理、环境安全感提升等手段间接降低极化风险。

4.2 治理启示

本研究结果表明,公众认知治理应更重视情境适配性和个体心理特点。在治理网络空间的观点极化时,除了推动公众接触更多异质信息,鼓励多元思考,还需关注公众所处的外部环境。注重传递积极、稳定的信息,缓解焦虑情绪和威胁感知,增强异质信息缓解认知极化的作用。基于此,本研究从算法治理、内容审核以及长效机制建立三方面提出以下治理启示。

第一,优化算法推荐机制,提升算法透明度。中央网信办等有关部门应继续推动平台优化算法参数,减少极端或争议性内容的推送,提高异质信息推荐权重,帮助用户接触多元观点。同时,督促平台提升算法推荐透明度,让用户理解推荐逻辑,自主选择推荐内容。并要求平台加强信息标识和预警机制,防范极端内容误导群众。

第二,强化内容审核,净化网络环境。持续打击煽动性、暴力性内容的发布与传播,巩固“清朗”行动成果。平台应继续深化人机协同审核机制,提升对网络黑话与隐喻表达的识别能力。同时,建立跨平台协作机制,实现对违规高风险账号的联合封禁,遏制违规内容的流窜传播。

第三,完善长效风险防控机制,实现高效治理。目前中央网信办等部门针对网络平台的治理仍以“专项治理”“运动式治理”为主,缺乏长效风险防控和监管治理机制。未来,政府与平台应探索建立跨平台的全过程预警机制,实现透明化、常态化、长效化的审核干预。

4.3 研究局限

尽管研究取得了一些重要发现,仍存在以下局限。第一,本研究的威胁情境操纵通过自然灾害与人为意外视频材料完成,侧重模拟生命安全类威胁情境,缺少对网络中围绕社会事件的威胁情境的探讨。未来可进一步细分威胁类型,比较不同威胁情境的影响差异。第二,虽然采用现实案例与材料进

行情景模拟,线上实验仍无法真实还原现实中的网络互动讨论,可能影响外部效度。并且,受限于网络问卷形式,研究对低受教育程度、老年群体等群体代表性不足。未来可通过线下实验访谈和网络行为数据采集,增强研究的外部效度和解释力。第三,本研究仅测量公众对技术的整体态度,未区分技术风险感知、收益感知等技术态度的不同维度。已有研究发现,对技术风险更敏感的公众更易受到隐私泄露、伦理冲突等威胁情境的影响^[28]。技术风险感知对公众认知极化的影响值得未来继续讨论。

参考文献:

- [1] 肖峰. 大模型与智能社会:基于历史唯物主义的探索[J]. 中国社会科学, 2024(7): 71-89, 205-206. Xiao F. Big models and the intelligent society: An exploration based on historical materialism[J]. Social Sciences in China, 2024(7): 71-89, 205-206.
- [2] 苏竣,魏钰明. 迈向智能社会:现实图景、发展趋势与治理使命[J]. 西北大学学报(哲学社会科学版), 2025, 55(1): 78-88. Su J, Wei Y. Towards an intelligent society: Realistic landscape, development trends, and governance mission[J]. Journal of Northwest University (Philosophy and Social Sciences Edition), 2025, 55(1): 78-88.
- [3] 宋艳,陈琳,李琴,等. 人工智能伦理风险感知、信任与公众参与[J]. 科学学研究, 2022, 40(7): 1153-1162, 1171. Song Y, Chen L, Li Q, et al. AI ethical risk perception, trust, and public participation[J]. Studies in Science of Science, 2022, 40(7): 1153-1162, 1171.
- [4] Piao J, Liu J, Zhang F, et al. Human-AI adaptive dynamics drives the emergence of information cocoons[J]. Nature Machine Intelligence, 2023, 5(11): 1214-1224.
- [5] 苏竣,宋立夫,魏钰明,等. 智能社会的舆论操控:理论机理、手段特征与社会影响[J]. 电子政务, 2025(1): 29-38. Su J, Song L, Wei Y, et al. Public opinion manipulation in the intelligent society: Theoretical mechanisms, characteristic methods, and social impacts[J]. E-Government, 2025(1): 29-38.
- [6] 郭跃,何林晟,苏竣. “工具-叙事-反馈”:一个行为公共政策的研究框架[J]. 中国行政管理, 2020(5): 71-78. Guo Y, He L S, Su J. Instruments, narrative, and feedback: A framework of behavioral public policy[J]. Chinese Public Administration, 2020(5): 71-78.

- [7] 霍治方,刘刚,樊正德. 中国公众对新兴技术采纳的影响机理研究——基于自动驾驶汽车社会实验调查的微观证据[J]. 科学学研究, 2025, 43(2): 311 – 24. Huo Z, Liu G, Fan Z D. Research on the mechanisms of public adoption of emerging technologies in China: Micro evidence based on social experiments on automated driving vehicles[J]. Studies in Science of Science, 2025, 43(2): 311 – 324.
- [8] Renström E A, Bäck H, Carroll R. Intergroup threat and affective polarization in a multi – party system [J]. The Journal of Social and Political Psychology, 2021, 9 (2): 553 – 576.
- [9] Jonas E, McGregor I, Klackl J, et al. Threat and defense: From anxiety to approach [A]. Advances in Experimental Social Psychology[C]. Elsevier, 2014: 219 – 286.
- [10] McLoughlin K L, Brady W J, Goolsbee A, et al. Misinformation exploits outrage to spread online [J]. Science, 2024, 386(6725): 991 – 996.
- [11] Arora S D, Singh G P, Chakraborty A, et al. Polarization and social media: A systematic review and research agenda [J]. Technological Forecasting and Social Change, 2022:183.
- [12] Stroud N J. Polarization and partisan selective exposure [J]. Journal of Communication, 2010, 60 (3): 556 – 576.
- [13] 易妍,叶淑兰. 社交媒体竞选如何强化美国政治极化:基于美国新冠议题的框架化考察[J]. 社会科学, 2023 (2): 114 – 127. Yi Y, Ye S L. How political campaigns on social media strengthen U. S. political polarization: A case study of framing the COVID – 19 issue [J]. Journal of Social Sciences, 2023(2):114 – 127.
- [14] 李龙飞,张国良. 算法时代“信息茧房”效应生成机理与治理路径——基于信息生态理论视角[J]. 电子政务, 2022 (9): 51 – 62. Li L F, Zhang G L. The generation mechanism and governance path of the "information cocoon" effect in the algorithmic era: Based on the perspective of information ecology theory[J]. E – Government, 2022(9):51 – 62.
- [15] Levy R E. Social media, news consumption, and polarization: Evidence from a field experiment [J]. American Economic Review, 2021, 111(3): 831 – 870.
- [16] Garrett R K, Stroud N J. Partisan paths to exposure diversity: Differences in pro – and counterattitudinal news consumption [J]. Journal of Communication, 2014, 64 (4): 680 – 701.
- [17] Stevens D, Vaughan – Williams N. Citizens and security threats: Issues, perceptions and consequences beyond the national frame [J]. British Journal of Political Science, 2016, 46(1): 149 – 175.
- [18] Van Prooijen J – W, Krouwel A P. Psychological features of extreme political ideologies [J]. Current Directions in Psychological Science, 2019, 28 (2): 159 – 163.
- [19] Gollwitzer A, Olcaysoy Oken I, Pizarro A O, et al. Discordant knowing: A social cognitive structure underlying fanaticism [J]. Journal of Experimental Psychology: General, 2022, 151(11): 2846.
- [20] 朱依娜,何光喜. 信任能降低公众对人工智能技术的风险感知吗? [J]. 科学学研究, 2021, 39 (10): 1748 – 1757, 1849. Zhu Y N, He G X. Can trust reduce the public's risk perception of artificial intelligence technology? [J]. Studies in Science of Science, 2021, 39(10): 1748 – 1757, 1849.
- [21] 廉翔鹏,魏钰明,苏竣. 渗透与理解:智能技术的公众信任机制研究[J]. 科学学研究, 2025, 43(7):1354 – 1363. Lian X P, Wei Y M, Su J. Penetration and understanding: A study on the public trust mechanism of intelligent technology [J]. Studies in Science of Science, 2025, 43(7):1354 – 1363.
- [22] Anderson C A, Carnagey N L, Eubanks J. Exposure to violent media: The effects of songs with violent lyrics on aggressive thoughts and feelings [J]. Journal of Personality and Social Psychology, 2003, 84(5):960 – 971.
- [23] Witte K. Predicting risk behaviors: Development and validation of a diagnostic scale [J]. Journal of Health Communication, 1996, 1(4): 317 – 342.
- [24] Schedler A. Rethinking political polarization [J]. Political Science Quarterly, 2023, 138(3): 335 – 359.
- [25] Stevens D, Bulmer S, Banducci S, et al. Male warriors and worried women? Understanding gender and perceptions of security threats [J]. European Journal of International Security, 2021, 6(1): 44 – 65.
- [26] Kraus M W, Piff P K, Keltner D. Social class, sense of control, and social explanation [J]. Journal of Personality and Social Psychology, 2009, 97(6): 992.
- [27] Ho A D, Yu C C. Descriptive statistics for modern test score distributions: Skewness, kurtosis, discreteness, and ceiling effects [J]. Educational and Psychological Measurement, 2015, 75(3): 365 – 388.
- [28] 朱旭峰,楼闻佳. 公众的风险感知与新兴技术接受度——面向人脸识别技术的调查实验研究[J]. 管理世界, 2024, 40(10): 118 – 132. Zhu X F, Lou W J. Risk perception of the public and acceptance of emerging

technologies: A survey experiment on facial recognition (10): 118 – 132.

technology[J]. Journal of Management World, 2024,40

Information exposure and threat situation:

An exploratory analysis of the causes of public cognitive polarization

HAN Yi – kun, WEI Yu – ming, SU Jun, ZHANG Fang

(School of Public Policy and Management, Tsinghua University, Beijing 100086, China)

Abstract: The rapid development of intelligent technologies is propelling humanity toward an intelligent society, bringing about profound changes in production, lifestyle, organizational structures, and social order. These transformations have given rise to numerous new challenges, making the governance of intelligent society an urgent research agenda. While existing studies have addressed key risks associated with intelligent technologies—including algorithmic abuse, the digital divide, and issues of trust and fatigue—relatively less attention has been paid to the transformative impact of intelligent recommendation systems on public information exposure, cognitive patterns, and behavioral preferences. The widespread use of personalized recommendations often results in individuals being repeatedly exposed to homogeneous content, leading to the formation of information cocoons, exacerbating cognitive polarization, and increasing the risk of societal fragmentation. This presents a significant challenge for intelligent social governance.

Traditional analyses of public cognition and attitude formation are often based on the rational actor model, assuming that individuals make decisions through cost – benefit calculations. However, such perspectives tend to overlook the psychological and emotional mechanisms involved in information processing. Recent studies from the fields of psychology and behavioral science have begun to explore how emotional preferences, identity, ideology, and personality traits influence the relationship between information exposure and cognitive polarization. Nevertheless, little attention has been given to the role of perceived external social environments, especially perceptions of societal threats. Psychological research indicates that perceived threats in the social environment significantly shape individual cognition and attitudes. In today’s digital landscape, where intelligent technologies dominate, threat – inducing and emotionally charged content is more likely to attract attention and be widely disseminated. This can amplify public anxiety and perceptions of insecurity, thereby intensifying cognitive polarization. Accordingly, this study seeks to bridge the gap by examining how threat perceptions and information exposure interact to influence public cognitive polarization.

This paper introduces the perspective of external threat situations and explores how they, together with information exposure, jointly influence public cognitive polarization. Using the application of facial recognition technology in “pedestrian red – light violation exposure” as a case study, this research constructs a 2 (no threat situation vs. threat situation) × 2 (homogeneous information vs. heterogeneous information) between – subjects experiment. A total of 1,281 participants were recruited online for the experiment. The results indicate that exposure to homogeneous information strengthens the public’s original views, thereby intensifying cognitive polarization, while exposure to heterogeneous information weakens the public’s original views, reducing cognitive polarization. Compared to the absence of external threat situations, threat situations significantly weaken the depolarization effect brought about by exposure to heterogeneous information. Additionally, for individuals of different genders and average annual incomes, threat situations play varying roles in the process by which information exposure influences polarization.

The findings demonstrate that the external environment in which the public is situated, including threat perceptions caused by social unrest, economic crises, crime, disease risks, and unexpected events, may have a significant impact on changes in cognitive polarization during information exposure. This deserves further attention and research. Moreover, when the government addresses viewpoint polarization in cyberspace, it should not only promote the public’s exposure to more heterogeneous information that challenges their original views and encourages diverse thinking, but also pay attention to the external environment the public faces. This can help reduce public anxiety and threat perception, thereby enhancing the role of heterogeneous information in mitigating viewpoint polarization.

Key words: cognitive polarization; intelligent recommendation algorithms; threat situation; information cocoon; intelligent society governance