

DOI: 10.16538/j.cnki.fem.20210818.103

## AI算法决策能提高员工的程序公平感知吗?

裴嘉良, 刘善仕, 钟楚燕, 谌一璠

(华南理工大学工商管理学院, 广东广州 510641)

**摘要:** 人力资源决策程序的公平性始终是组织管理的一个关键问题。在数智化时代, 越来越多的管理者期望引入人工智能算法优化决策环境并提高决策效率。但是, 考虑到人力资源决策的直接对象是员工, 他们如何感受、认知和评价人工智能算法决策程序公平性的问题十分重要。本文基于信息透明度的视角, 以程序公平理论为研究框架, 探究了人力资源决策中不同决策主体(人工智能算法vs.上级主管)对员工程序公平感知的影响机理。两个情境实验结果如下: 实验1发现, 在组织人力资源决策情境下, 与上级主管决策相比, 员工认为人工智能算法决策的信息透明度更低, 并由此对人工智能算法决策产生更低的程序公平感知; 实验2发现, 组织包容型氛围在人工智能算法决策对员工程序公平感知的主效应中起调节作用, 即当员工感知到具有较低水平的包容型氛围时, 人工智能算法决策对程序公平感知的影响更加明显。本研究识别出人工智能算法决策潜在的负面影响, 对于辩证地认识组织决策中人工智能的应用具有理论意义和实践价值。

**关键词:** 人工智能; 算法; 人力资源决策; 信息透明度; 程序公平

**中图分类号:** F270 **文献标识码:** A **文章编号:** 1001-4950(2021)11-0041-15

### 一、引言

人工智能、机器学习和大数据基础设施的发展正在对传统的管理模式发起挑战。许多企业开始拥抱数字化技术, 运用人工智能(artificial intelligence, AI)持续赋能组织, 为企业高质量且具有竞争力的发展提供协调市场资源、提升管理效率、优化决策环境和重塑商业模式等机遇和动力(Duan等, 2019; 徐鹏和徐向艺, 2020)。其中, 管理决策贯穿组织管理全过程, 许多具有前瞻性的管理者青睐于使用以计算机深度学习算法为基础、大数据驱动的人工智能来集成组织内外部信息, 协助管理者进行复杂且关键问题的决策(Wilson和Daugherty, 2018)。例如, 亚马逊使用混沌存储算法实时计算可分配空间以实现大型仓库的智能化管理(Danaher, 2016); 谷

收稿日期: 2021-03-29

基金项目: 国家自然科学基金重点项目(71832003); 国家自然科学基金面上项目(71772067)

作者简介: 裴嘉良(1997—), 男, 华南理工大学工商管理学院博士研究生;

刘善仕(1966—), 男, 华南理工大学工商管理学院教授, 博士生导师;

钟楚燕(1996—), 女, 华南理工大学工商管理学院硕士研究生(通讯作者, 1107821643@qq.com);

谌一璠(1997—), 女, 华南理工大学工商管理学院硕士研究生。

歌使用基于自然语言处理算法技术帮助管理者分析员工邮件中的文本信息以预测哪些员工具有离职倾向,从而预先采取防范措施(Strohmeier和Piazza,2015)。

从实践层面来看,人工智能的底层技术是算法(algorithms),即基于数理统计模型和决策程序的一系列指令组合,能告诉计算机在精准的步骤和规则内如何完成目标任务(Schildt,2017)。人工智能算法(以下简称AI算法)通过机器、系统、数据和网络构建更高逻辑层次的复杂系统,随着数据资源的持续输入而不断重写升级(Li和Du,2017)。具备高效、优化和数据驱动以实现决策的能力和优势的AI算法已被越来越多的管理者应用于管理决策的实践活动中(刘善仕等,2021)。此外,在管理实践过程中,缺乏公平感是导致人才流失的重要因素之一(Shih和Susanto,2011)。而人力资源决策(以下简称HR决策)(例如,员工招聘、工作分配和绩效考核等)的直接对象是员工,决策过程是否一致、基于准确的信息或不受决策者主观偏见的影响对于员工公平感的体验十分重要,管理者有强烈的动机去寻找能够确保决策过程是公平的方法。因此,许多人力资源专家尝试在实践中利用AI算法来保障决策的科学有效性。例如,领英(LinkedIn)等职业社交平台上的推荐算法会对数以千万职业人员的资料进行排序,结合招聘需求和求职者个人信息向客户个性化地推荐最匹配的人才(Cho和Lam,2021);施乐服务公司(Xerox Services)则实践应用了一种招聘算法来支持HR管理者的招聘决策,为他们提供申请人与工作匹配程度的分数(Peck,2013)。可见,实践界十分关注AI算法在提升HR决策准确性、客观性和公平性方面的作用。

从理论层面来看,虽然AI算法技术在决策公平领域的研究十分火热,但现有文献针对“AI算法决策是否无偏公正”问题的研究结论仍存在较大争议。先前研究表明,当算法被视作重要的信息来源时,由于算法每次运算都遵循相同且固定的程序,不会受到人类情感因素的影响,因此相比人类决策者而言存在更少的决策偏见(Schildt,2017;Wilson和Daugherty,2018)。Edwards等(2000)指出,算法打破了人类对于知识、时间和精力限制,可以基于数据信息创造出多种备选方案,将管理决策的思维从“有限理性”过渡到“极限理性”,有利于辅助决策者做出更加科学合理的决策。但是,后续研究则发现算法不仅客观上可能会持有偏见,并且其决策程序可能被视为是不公平的。一方面,现有研究表明,由于在所有的案例中,算法的设计与开发都有人的参与,并且AI算法具备自主学习的能力。因此,用于训练和强化算法的机器学习的历史数据模型存在人类偏见的可能性。例如,许多在线招聘公司使用的推荐算法从具有结构性偏见(例如,男性优于女性等)中的历史案例中获取数据用于强化学习,这导致招聘过程中出现歧视问题(Lambrecht和Tucker,2019)。另一方面,人机交互领域的文献表明,人们对于算法的感知十分重要,不论实际决策结果的质量如何,由于这些决策出自算法而非人类自己,这一事实可能会影响人们对算法决策的信任和态度(Sundar和Nass,2001)。例如,Lee(2018)针对员工如何感知算法决策的问题,通过一项在线实验研究发现,在招聘的情景任务中,由于员工倾向于认为算法缺乏人类的直觉以及主观判断能力,因而导致更低的公平性和可信度;Newman等(2020)的基于HR决策的情景实验研究也发现,在组织借助算法进行HR决策的情境下,员工认为算法决策的过程是被简化的,算法缺乏考虑语境化的定性信息,故其决策的程序公平性难以被信服。

综上所述,在AI算法越来越多地被企业引入到管理决策的新兴背景下,关于算法决策公平性问题的探讨受到实践界和理论界的广泛关注和热议。员工如何感知AI算法决策公平性,对于他们后续采取何种工作态度和行为以及组织数字化管理实践能否取得预期效果具有重要影响。人力资源(以下简称HR)管理领域中关于决策公平问题的既有研究表明,员工参与、上下级沟通以及领导成员交换关系等因素能够影响员工体验到的公平感(Jablin,1979;Hollenbeck等,1998;Kacmar等,2003)。但少有学者基于决策主体差异的视角对上述问题作出

解答。随着越来越多地管理者运用AI算法进行HR决策,员工对于决策程序的一致性、准确性、无偏性以及决策主体能否接纳不同意见并进行调整等信息即程序公平问题的担忧日益普遍(Colquitt等,2013),员工认为AI算法决策是否比上级主管决策更加公平,其影响机理是什么,以及组织应该采取何种措施予以干预这一具有理论价值和实践意义的问题尚未得到解决。因此,本研究围绕“决策主体差异对员工程序公平感知影响”的研究主题,在回顾已有理论文献的基础上,基于信息透明度的视角,以程序公平理论为基础构建了理论模型,通过设计2个实验研究检验HR决策中不同决策主体(AI算法vs.上级主管)对员工程序公平感知的影响,并揭示这一影响的作用机理和边界条件。本研究关于AI算法决策与员工程序公平感知之间关系的探讨,将补充已有决策公平和人工智能领域的研究空白,并为企业管理者科学借助人工智能等数字化技术进行管理实践创新提供借鉴。

## 二、理论基础与研究假设

### (一)程序公平理论与AI算法决策

程序公平是指产生结果的过程具有公平性(Lind和Tyler,1988)。初期的组织公平性的研究关注分配公平的影响,但近年来许多研究发现决策的过程比最终结果更加重要。程序公平能够影响员工对决策者的反应,会进而影响其后续的工作态度和行为。当员工认为上级决策有失公平时,会降低管理者行为的有效性并且导致员工针对工作环境的消极情绪(Cropanzano等,2001)。此外,Bal等(2011)发现,员工的程序不公平感知显著增强了其离职意愿。这些对决策程序不公平的感知会产生各种负面结果,会严重降低组织效能。因此,挖掘何种潜在因素导致以及缓解甚至扭转员工对决策程序不公平的看法始终是组织管理中的一个重要研究议题。Leventhal(1980)在组织情境中应用了程序公平理论,并指出影响员工程序公平感知的主要因素包括能够对决策过程产生影响的控制性、决策程序的一致性、决策程序的无偏性、决策信息的准确性、决策程序的代表性以及决策程序的道德性。后续研究围绕Leventhal(1980)提出的6个判断程序是否公平的标准展开研究,发现决策的不确定性程度、公开透明程度以及组织结构特征(个体主义或集体主义、正式化程度和员工参与度)等因素影响员工的程序公平感知(Schminke等,2002;刘亚等,2003;Thau等,2009)。

伴随大数据、人工智能和机器学习等数字技术的发展,许多组织尝试将AI算法引入到决策过程中,规避人类决策过程中固有的偏见,程序公平问题的一个潜在方案可能是更多地使用大数据驱动的数据辅助决策。算法被定义为根据统计模型或决策规则自主决策,告诉计算机在一系列精准的步骤和规则内应该如何运作的计算程序,其过程无需人工干预(Schildt,2017)。算法客观上可以依靠数据驱动的数字逻辑来消除人的主观性偏见从而改进决策,组织期望算法能够在时间和资源的有限条件下依靠大数据进行理性客观决策,使员工对决策的程序感到满意和公平。但是,在围绕人们如何判断算法决策公平的研究中, Lee(2018)通过实验研究发现,通过算法进行招聘的决策被员工视作不如人类主管的决策公平。此外,Newman等(2020)基于简化论视角的实验发现,由于算法决策的程序是被简化的,缺乏考虑语境化的定性信息,员工产生更多程序不公平感。上述研究结果表明员工并不认为由算法进行决策的程序更加公平。程序公平取决于决策程序是否给予准确的信息,是否透明,是否一致适用等。然而既有文献表明,由于算法决策的过程本身不具有透明性,加之复杂的运算过程不易被理解并且缺少必要信息的披露,这导致许多工作者的沮丧情绪,认为决策程序是不公平的(Orlikowski和Scott,2014;Lee等,2015)。在程序公平理论的框架内,信息透明度是影响个体公平感知的重要因素。本研究试图基于信息透明度的视角阐释不同决策主体(AI算法vs.上级主管)对员工程序公平感知的影响。

## (二) AI算法决策对员工程序公平感知的影响

在数智化的组织环境中,越来越多地管理者试图借助大数据驱动的AI算法进行科学决策。客观上, AI算法比人类更能够理解和解释决策的过程,大数据资源为算法提供了众多的模型和架构从而提高决策效率(Schildt, 2017)。但是,由于现阶段人们并不能真正地了解机器学习算法是怎样在不断变化的市场环境下学习适应并且自行优化的,算法运作的过程存在着“黑箱”问题,即基于神经网络的深度学习,将抓取的数据细分为数以千万计的变量自动加权计算与输出结果之间存在不可观察的空间,算法内部运作的程序往往是不透明的(Burrell, 2016)。人类的智慧能够针对某一具体的结论进行推理和论证,但无法解释得到某一特定结论的复杂和隐含的过程。因此,人们会对算法决策的结果或得出的结论表示担忧和困惑。与此同时, AI算法的不透明性与实时动态变化的特点,也使一系列因算法产生的问题难以被发现和监督。其中,基于机器学习的算法受到来自于社会大数据资源的训练,机器很容易习得既有的偏见,类型化、标签化的算法设计思想很容易在社会群体中固化社会偏见,会对某个社会群体作出误判或伤害。例如Miron等(2021)在刑事司法中使用风险评估算法对犯罪率进行比较,由于该算法是从本身就具有结构化歧视的案例中汲取数据进行强化学习,因此它报告的结果是非裔美国人的犯罪风险系数高于白人。Burrell(2016)指出,在商业环境下,由于算法对于企业而言是作为核心竞争力的关键,企业会有意隐藏由一系列源代码组成的算法以保护可能涉及的关键知识产权。但对于算法输出结果的接受方而言,他们似乎并不理解和确定为何由算法计算后会得到一个特定的结果。

在程序公平理论的研究框架下,本研究认为,在组织进行HR决策时,当员工得知决策的结果是由AI算法制定时,面对所属群体的不同属性(AI和人类),员工会通过将AI算法与拥有人类自身特征的上级主管进行严格对比和区分,以获得既定条件下相应的心理认知。尽管员工可能倾向于认为AI算法在决策程序执行上具有相对客观中立的特点,能够有效规避人类决策过程中固有的偏见。但是,在程序公平感知形成的其他重要影响因素上(例如:决策参与、决策申诉和决策不确定性),AI算法决策比上级主管决策可能让员工产生更低水平的程序公平感知。具体而言:第一,在决策参与方面,员工倾向于认为上级主管在HR决策过程中会积极接纳他们不同的声音与建议,并根据他们提供的信息而进行调整(Colquitt等, 2013)。但AI算法是“冷冰冰”的程序公式,员工无权干预其决策过程,他们认为自己的意见并不会被预先设定好、标准化且重复运行的AI算法所采纳。第二,在决策申诉方面,当组织通过算法进行HR决策时,员工倾向于认为自己难以获得申诉的机会,面对AI算法的决策结果,他们只能选择被迫接受,并对申诉无门而感到困惑和不满(Kuhn和Maleki, 2017)。与之相比,当决策主体是上级主管时,通常情况下员工对于决策结果的不满可以找到上级主管予以申诉,充分表达自己的意见。第三,在决策不确定性方面,由于AI算法考虑了复杂的多种因素,其进行HR决策的结果难以被员工所预测。而上级主管通常会追踪、记录和参考员工的绩效表现来进行HR决策,员工通常会对自己的绩效表现有一个心理预判标准,因此, AI算法决策会给员工带来更多的不确定性。例如, Rosenblat和Stark(2016)进一步的研究中也指出, Uber等网约车平台基于AI算法的派单机制除了考虑乘客的位置信息外还考虑了各个区域的运力情况、司机往返途中的其他业务、司机的顾客评价等诸多复杂的因素,但司机显然对此并不了解。此外,现有心理学研究表明,即使算法决策优于人类决策,但人们似乎对AI算法的容错度更低,这可能导致他们对算法输出的结果更加严苛。人们必须了解算法推理过程和逻辑关系的心理需求使其更倾向于对算法表现出不信任(Dietvorst等, 2015)。综上所述,本研究提出以下假设:

假设1:在组织进行HR决策情境下, AI算法决策比上级主管决策让员工产生更低的程序公平感知。

### (三)信息透明度的中介作用

透明度往往与信息分享、信息披露、信息交换等词义相关,被广泛地界定为相关主体间信息的公开程度(Liu等,2015)。信息透明度作为决策者建立公平感和信任感的纽带,能够帮助个体理解决策过程中的相关信息以便于针对决策结果予以判断并采取行动(O'Toole和Bennis,2009)。在决策程序公平领域,信息透明度具有三个标准:首先,决策标准的清晰度反映了决策者作出决策的规则和方法是否公开且易被理解。在组织情境下,当缺乏决策标准的沟通时,员工可能会对决策过程表示质疑或产生不公平感,例如,当上级主管在绩效评估的过程中没有明确评定规则时,员工体验到的公平感就会下降(Kim和Mauborgne,1998)。其次,决策标准的有效性反映了在既定的决策环境下用以决策的规则是否合理。如果用以决策的标准是基于不合理的推断或毫无根据的假设前提,则被视作无效。例如,在派发薪酬奖励过程中用于评判优秀与否的标准如果仅考虑个人简历信息而非业绩信息时,员工会认为决策标准是无效且决策程序是不公平的(Brockner和Tyler,1992)。最后,决策标准的代表性反映了决策信息或规则多大程度上能够考虑决策参与者所关注的内容。决策标准具有代表性能够充分考量和回应决策过程中个体的意见,避免他们产生被孤立或排斥的心理(Miller,1989;Lee,2018)。

近来大量文献开始关注人工智能中“黑箱”问题,并以此描述AI算法如何在“暗处”运作的景象,指出AI算法内部运作的过程往往是不透明的(Orlikowski和Scott,2014;Lee等,2015;Burrell,2016)。Burrell(2016)系统归纳了AI算法信息不透明的三种不同形式包括:首先,企业或机构对核心算法有意的保护和隐瞒:算法的不透明性可以归因为企业以竞争优势的名义有意防止算法泄露,但也可能是一种新的回避监管或操纵受众群体的形式。因此,企业或机构可能会利用算法的不透明性做出有失公平的机会主义行为(Sandvig,2015)。其次,编写和读取算法代码具有专业性:尽管算法背后的代码是使用特定的R或Python等语言编写的,但编写和读取这些代码需要严格遵守逻辑规则和精确语法以便于机器读取,这与人类语言具有显著的差异,复杂众多的代码难以被理解,需要强大的专业知识和计算思维(Schildt,2017)。最后,机器学习算法的智能复杂性:机器学习算法会在运作过程中随着数据资源持续大量的输入而不断重写和升级,通过对海量数据的集成,利用构建情境模拟等方式协助复杂问题的分析、推导和处理,因此,机器学习算法的更高层次逻辑已然超越了人类现阶段的思维。此外,机器学习算法是基于类型化和标签化的设计,可能在大数据集成学习的过程中进行不具有广泛代表性和合理性的决策,甚至出现歧视等问题(Miron等,2021)。

程序公平理论始终强调决策过程的信息准确透明(Thau等,2009)。因此,基于信息透明度的三个标准,本研究认为,在组织进行HR决策时,与上级主管决策相比,员工认为AI算法决策具有更低的信息透明度,进而产生更低的程序公平感知,理由如下:第一,在决策标准的清晰度方面,由于AI算法运作机制的“黑箱”属性(Burrell,2016),其难以被理解的运算规则通常会使用员工感知到决策标准的清晰度远低于上级主管,这不仅是因为AI算法其本身需要强大的专业知识和计算思维去理解,同时上级主管的决策标准通常情况下会向员工公开,并不会刻意隐瞒所有的信息。但当AI算法决策的标准和上级主管的决策标准都未被完全公开时,相比于复杂的AI算法,员工更容易猜测和理解上级主管做出决策的相关依据。第二,在决策标准的有效性方面,尽管大数据驱动的AI算法的决策过程客观上具有更高的效率和准确性,但这与员工的主观感知具有一定的差距(Lee,2018),即员工倾向于认为AI算法只是将部分可以量化的指标纳入到HR决策的标准,而一些非量化且无从记录的内容会被AI算法所简化(Newman等,2020)。与之相比,员工倾向于认为上级主管在决策标准的设定上,不仅会考虑可以量化的指标,还会参考日常工作过程中一些隐性且只能被人类观察的内容,例如人情世故、精神面貌等因素,因此,员工会认为AI算法决策标准并非基于完整的信息,因而对其有效性的评价更低。

第三,在决策标准的代表性方面,由于AI算法决策程序的编写来自组织中的技术团队,员工会质疑这些并不了解他们基本信息的编程人员设计的AI算法所做出的决策结果,员工不可避免地会认为AI算法并不会充分考量他们的意见,并且与上级主管相比,AI算法也不会回应他们的意见,AI算法决策更容易导致员工产生被孤立或排斥的心理,因而对AI算法决策标准的代表性表示出不信任(Dietvorst等,2015)。综上所述,考虑到信息透明度对决策程序公平的重要影响以及针对AI算法运作“黑箱”不透明性的担忧,本研究认为信息不透明度能够解释AI算法决策对程序公平感知的作用机理,提出以下假设:

假设2:信息透明度在AI算法决策对程序公平感知的影响中起中介作用,即与上级主管决策相比,员工认为AI算法决策具有更低的信息透明度,进而产生更低的程序公平感知。

#### (四)包容型氛围的调节作用

随着组织边界被逐渐打开,多元化的各类要素流入组织。在强调平等管理方式和理念的基础上,为了避免阻碍群体差异性 or 个性发挥的问题(Miller,1998),学者们在多样化管理的基础上提炼出既平等对待差异化个体,又尊重个性化的存在,鼓励差异化个体表现自我,提升自我价值的包容型管理(许梅枝和张向前,2018)。其中,包容型氛围作为实现包容型管理的重要手段,强调组织鼓励员工保留个体特质并尊重其独特性的存在,群体成员能够共同感知到来自组织的尊重并体验到归属感的体验(Shore等,2011)。Nishii(2013)构建了包容型氛围的内涵,指出包容型氛围是群体成员对“组织是否平等对待员工、鼓励员工参与、尊重员工在工作中的差异性”等包容特征的心理感知。许多组织行为学者重点探讨了包容型氛围的调节作用。例如,Nishii(2013)发现包容型氛围负向调节了性别差异与冲突、冲突与满意度之间的关系;David等(2016)研究表明包容型氛围调节了上下级之间存在的残疾(身体、认知或心理健康水平偏离正常人)状态差异对领导成员交换关系质量和绩效的影响。过去研究强调组织包容型氛围在影响和塑造员工在人际关系中相关心理感知的积极作用,却忽视了上述员工直接或间接感知到工作环境中具有包容性的组织特质对于构建和谐的人—技术关系的潜在作用。组织氛围作为组织成员对于其所处组织环境的一种认知诠释(James和Jones,1974),会影响和塑造组织成员后续各个方面的认知判断和行为表现。

自古以来,新技术的发展都要经历质疑、理解、接受再到成熟的过程,当人们对新技术持有包容的态度时更有益于缓解对新技术的排斥心理,愿意接受新技术现阶段存在的问题并且相信其发展的前景(Orlikowski和Scott,2014)。因此,本研究认为,当员工感知到不同水平的组织包容型氛围时,自然会形成差异化的认知判断,进而导致决策主体差异对员工程序公平感知产生不同的影响效应,理由如下:第一,在决策参与方面,在良好包容型氛围中的员工通常会感知到组织积极鼓励员工参与决策,并且激励员工商讨并提出有益于组织的建议(Nishii,2013)。因此,不论决策主体是AI算法还是上级主管,在日常工作环境中,员工都能够感知到组织决策对于员工自身参与的重视。虽然决策主体是AI算法,但员工同样会体验到良好的决策参与感,因为他们倾向于认为AI算法背后的组织会充分考虑他们的声音和意见。第二,在决策申诉方面,在良好包容型氛围中的员工通常会感知到组织关注实践过程中的公平性和多元性,并强调决策过程中对于员工主体地位的认可和尊重。换言之,在包容型氛围的熏染下,员工更容易体验到公平给自身带来的主体地位(Shore等,2011)。因此,当员工对于决策结果有异议时,不论决策主体是AI算法还是上级主管,员工对于向上申诉都更有信心。虽然决策主体是AI算法,但员工并不会因为申诉无门而感到困惑和不满,因为他们感知到的主体地位会促使他们积极寻找各方途径予以申诉,并相信组织会正面回应他们的申诉意见。第三,在决策不确定性方面,在良好包容型氛围中的员工通常会感知到组织对于不确定性的因素更加具有包容性和开放性(Nelissen等,2017)。因此,虽然员工可能对AI算法决策持有更低的容错度,但具有更高包容型

氛围感知的员工可能会对AI算法决策带来不确定性持有更加开放的态度,承认并且接受这种不确定性的存在。综上所述,包容型氛围本质上强调兼容并蓄(张凯丽和唐宁玉,2017),因而在良好组织包容型氛围熏染下的员工对于决策主体差异可能表现出更高的兼容和接受度,他们可能会认为不论决策主体是AI算法还是上级主管,在高水平组织包容型氛围下的HR决策的程序是相对公平的。综上所述,本研究提出以下假设:

假设3:组织包容型氛围在AI算法决策对程序公平感知的影响中起调节作用,即当员工感知到具有较高水平的包容型氛围时,AI算法决策相比于上级主管决策对员工程序公平感知的影响效应没有显著差异。

### 三、实验1:AI算法决策对员工程序公平感知的影响机制检验

#### (一)实验设计

实验1的目的是检验在组织进行HR决策情境下,AI算法决策对员工程序公平感知影响的主效应以及信息透明度的中介作用。本次实验调查周期为3日,采用单因素两水平(AI算法决策vs.上级主管决策)被试间设计。本研究实验被试源于华南地区3家农牧企业,涉及养殖、饲料加工等领域。在正式开始实验之前,首先向每家企业的人力资源经理介绍本次实验的目的及概况,随后秉持自愿参与的原则,由人力资源经理通过企业内部邮件系统征集实验被试。最后,共计244名员工同意参与本次实验,调研结束后给予纪念礼品作为奖励。被试均为具有一定工作经验的企业员工,被随机分配到AI算法决策组和上级主管决策组,每组122人。在调查周期结束后,剔除漏填、乱填等无效样本,以及决策主体感知量表两题项回答前后不一致的问卷,共获得220份有效问卷,有效率达到90.16%。其中,AI算法决策组99人,男性被试占58.6%,被试平均年龄约34岁;上级主管决策组121人,男性被试占49.6%,被试平均年龄约为34岁。两组被试在人口统计学特征上除学历存在显著差异外,其余不具有显著差异。借鉴Newman等(2020)的实验设计,利用HR决策情境模拟的方法操纵决策主体,要求被试针对两个决策主体之一(AI算法vs.上级主管)进行的人力资源裁员决策进行评估。

#### (二)实验问卷和变量测量

1.描述性情景材料。本研究所采用的描述性情景材料在正式实验之前经过材料评定选取。在正式实验之前,本研究对Lee(2018)和Newman等(2020)所用的实验材料根据国内的现实情境进行改编,得到关于晋升、裁员、政策实施、奖励、绩效评估和面试共六种人力资源决策的描述性情景材料,每种情景材料均包含AI算法决策和上级主管决策两种情况。本研究通过线上平台招募20位在职的工作者对六种描述性情景材料进行评定,被试需回答“该材料描述的决策者是谁?”“该材料描述的是何种人力资源情景?”等关于主题的开放性问题,以及“该材料描述的场景多大程度上是合理的?”“该材料描述的场景在现实生活中发生的可能性有多大?”等关于场景真实性的Likert7点评分问题。本研究最终根据主题吻合度(对开放性问题的回答包含关键词的被试比例)、场景真实性得分两个标准评定出吻合度及真实性最高的两份材料——裁员、奖励决策情景材料。

2.操纵决策主体。本研究邀请被试阅读由材料评定选取出的关于人力资源裁员决策的描述材料,目的在于操纵解释变量决策主体:AI算法vs.上级主管。为确保实验条件的真实性,并且避免由于裁员这一相对敏感情境下被试的紧张反应,实验材料已将裁员部门名称和对象姓名模糊化处理。两组被试阅读的材料前半部分相同:“由于2020年公司受到新冠肺炎疫情的影响,公司经营受到前所未有的挑战,大面积业务量快速缩水,预计公司业务全面复苏仍需相当长的时间,为控制公司经营成本,公司决定对××部门进行裁员。”接下来,AI算法组被试阅读的

情景描述为：“管理层借助AI(人工智能)算法对××部门的××进行分析并决定他/她是否在裁员名单中。”上级主管组被试阅读的情景描述为：“××部门的上级主管对××进行分析并决定他/她是否在裁员名单中。”随后，请两组被试回答：“①根据上述情景描述，是谁分析并决定裁员的最终名单？②谁来决定受到疫情影响被迫进行的裁员名单？”采用Likert7点评分法，1为肯定是上级主管，7为肯定是AI算法。决策主体感知的信度系数为0.870。

3.信息透明度和程序公平感知。信息透明度的测量工具选取Liu等(2015)开发的4题项量表来测量。代表题项有“决策依据的信息是清晰的”“决策依据的信息是容易理解的”“决策依据的信息是容易获得的”等。采用Likert7点评分法，1为非常不同意，7为非常同意，该量表在本研究中的信度系数为0.838；程序公平感知的测量工具选取Colquitt(2001)开发的7个题项量表来测量，该量表的信效度在中国样本中得到有效检验，代表题项有“决策过程没有偏见”“决策过程基于准确的信息”“决策过程符合伦理道德标准”等。采用Likert7点评分法，1为非常不公平，7为非常公平，该量表在本研究中的信度系数为0.862。

4.控制变量。根据当前学术界的建议，应在统计分析中保留能够得到理论解释的控制变量，为最终研究结论的可靠性提供严格和准确的信息(Bernerth和Aguinis, 2016)。参考以往研究，除性别、年龄、工作年限等人口统计特征变量可能影响最终结果外，Newman等(2020)还将决策的典型程度(指代表性或常见性，典型性越高代表AI算法或上级主管在日常HR决策中是越常见的，越能代表日常工作中的情况)作为控制变量。本研究借鉴Newman等(2020)的测量方式，采用单一题项“您认为上述材料中由AI(人工智能)算法/上级主管做出裁员决策的典型性程度有多高？”对决策的典型性进行测量，采用Likert7点评分法，1表示非常不典型，7表示非常典型。Lee(2018)指出，决策主体考虑因素的数量也可能影响个体公平性的感知，因此，本研究请两组被试回答：“在上述情景材料中，AI算法或上级主管在决策过程中可能考虑哪些因素？”随后将考虑的因素数量进行统计并在数据分析过程中予以控制。

### (三)实验结果

1.描述性统计及相关性分析。表1描述了本研究实验1主要变量的均值、标准差以及各变量间相关分析结果。由表1可得，实验1的被试平均年龄34岁、平均工作年限3.6年，男性118名，占53.6%，本科及以上学历170名，占70.5%。相关分析结果显示，决策主体与信息透明度、程序公平感知存在显著的负相关关系( $r=-0.328, p<0.001$ ;  $r=-0.411, p<0.001$ )；信息透明度与程序公平感知存在显著正相关关系( $r=0.432, p<0.001$ )。以上分析结果符合本研究的理论预期，为研究假设提供了初步支持。

表1 实验1描述性统计与相关系数矩阵

变量	<i>M±SD</i>	1	2	3	4	5	6	7	8	9	10
1.性别	1.464±0.500										
2.学历	3.395±1.172	-0.026									
3.年龄	33.773±4.130	0.113	-0.036								
4.工作年限	3.555±1.680	-0.014	-0.045	0.102							
5.职位层级	1.555±0.498	-0.010	0.045	-0.116	-0.069						
6.决策典型性	3.982±1.801	0.009	-0.105	-0.039	0.000	0.027					
7.决策考虑因素的数量	2.155±0.851	-0.040	-0.066	0.000	-0.073	-0.149*	-0.049				
8.决策主体感知	4.348±1.492	-0.110	0.022	0.001	0.022	-0.033	-0.022	0.109			
9.决策主体	0.450±0.499	-0.090	0.202**	-0.026	-0.016	0.002	-0.037	0.115	0.338***		
10.信息透明度	4.081±1.157	0.077	-0.043	-0.005	-0.014	-0.030	0.168*	0.044	-0.039	-0.328***	
11.程序公平感知	4.169±1.053	-0.058	-0.106	0.020	0.041	0.013	0.068	0.005	-0.019	-0.411***	0.432***

注：*N*=220；\*表示 $p<0.05$ ，\*\*表示 $p<0.01$ ，\*\*\*表示 $p<0.001$ 。



2.操纵检验。单因素方差分析结果表明,针对决策主体感知的AI算法组中被试的评分显著高于上级主管组中的被试,AI算法组的均值 $M=4.904$ ,上级主管组的均值 $M=3.893$ ,两组被试的均方比值 $F(1,218)=28.148$ ,显著性系数 $p<0.001$ ,效应量系数 $\eta^2=0.114$ 。表明本研究对决策主体的操纵成功。

3.主效应检验。单因素方差分析结果表明,AI算法组的程序公平感知显著低于上级主管组被试,AI算法组的均值 $M=3.693$ ,上级主管组的均值 $M=4.560$ , $F(1,218)=44.197$ ,显著性系数 $p<0.001$ ,效应量系数 $\eta^2=0.169$ 。数据显示,在组织进行HR决策的情境下,AI算法决策与上级主管决策相比显著的给员工带来更低的程序公平感知。因此,假设1得到验证。

4.中介效应检验。首先,单因素方差分析结果表明,AI算法组被试评估的信息透明度显著低于上级主管组被试,AI算法组的均值 $M=3.662$ ,上级主管组的均值 $M=4.424$ , $F(1,218)=26.338$ ,显著性系数 $p<0.001$ ,效应量系数 $\eta^2=0.108$ ,表明员工更倾向于认为AI算法决策的信息透明度更低,符合本研究的理论预期。其次,在决策的典型程度上,AI算法组的均值 $M=3.909$ ,上级主管组的均值 $M=4.041$ , $F(1,218)=0.293$ , $p=0.589$ 不显著,效应量系数 $\eta^2=0.001$ 。在决策考虑因素的数量上,AI算法组的均值 $M=2.263$ ,上级主管组的均值 $M=2.066$ , $F(1,218)=2.926$ , $p=0.089$ 不显著,效应量系数 $\eta^2=0.013$ 。数据显示,上述2个其他可能的解释变量在AI算法组和上级主管组被试打分上不存在显著差异,排除其可能的潜在影响,后续中介效应检验中作为控制变量予以控制。最后,利用Bootstrap分析中的模型4检验信息透明度的中介作用,以程序公平感知为被解释变量,以决策主体为解释变量,决策主体中上级主管取值为0,AI算法取值为1,以信息透明度为中介变量。Bootstrap分析结果显示,信息透明度在AI算法决策对程序公平感知的影响中起部分中介作用,95%CI=[-0.335,-0.093],不包含0。为进一步提高研究结论的可靠性,本研究参考Bernerth和Aguinis(2016)的建议,汇报不含控制变量时的中介效应检验的分析结果。数据结果同样表明信息透明度所起的部分中介作用,95%CI=[-0.390,-0.135],不包含0。因此,与上级主管决策相比,员工认为AI算法决策具有更低的信息透明度,并产生较低的程序公平感知,假设2得到验证。

综上分析,实验1利用两组组间实验的方法验证AI算法决策对员工程序公平感知的影响,同时还验证了信息透明度的中介作用。数据结果揭示了在组织进行HR决策的情境下,AI算法决策对员工程序公平感知的影响机理,接下来重点检验在什么边界条件下AI算法决策对员工程序公平感知的影响可能减弱。

#### 四、实验2:组织包容型氛围的调节作用检验

##### (一)实验设计

实验2的目的是重复检验实验1得出的结论,并进一步检验组织包容型氛围的调节作用。本次实验调查周期为2日,采用AI算法和上级主管的组间设计。此次实验被试源于华南地区1家小微企业,涉及人力资源外包服务领域。在正式开始实验之前,首先向企业负责人介绍本次实验的目的及概况,随后秉持自愿原则,由企业负责人通过企业内部邮件系统征集实验被试。最后,共计264名员工同意参与本次实验,调研结束后给予纪念礼品作为奖励。被试均为具有一定工作经验的企业员工,被随机分配到AI算法决策组和上级主管决策组,每组132人。在调查周期结束后,剔除漏填、乱填等无效样本,以及决策主体感知量表中两题项回答前后不一致的问卷,共获得252份有效问卷,有效率达到95.45%。其中,AI算法决策组112人,男性被试占55.4%,被试平均年龄约为34岁;上级主管决策组140人,男性被试占47.9%,被试平均年龄约为34岁。

两组被试除学历存在显著差异外,其余人口统计学特征不具有显著差异。实验2对决策主体的操纵设计与实验1类似,借鉴Newman等(2020)的实验设计,利用HR决策情境模拟方法操纵决策主体,要求被试评估两个决策主体之一(AI算法vs.上级主管)所做的人力资源奖励决策。

### (二)实验问卷和变量测量

1.操纵决策主体。邀请被试阅读由材料评定选出的关于人力资源奖励决策的描述材料,目的在于操纵解释变量决策主体:AI算法vs.上级主管。为确保实验条件的真实性,避免社会赞许性问题,实验材料将奖励部门名称和对象姓名模糊化处理。两组被试阅读的材料前半部分相同:“在2020年公司积极应对新冠肺炎疫情的影响,抓准灵活用工市场快速扩张的机会,公司的业务往来同比去年增加10个百分点,为激励公司员工继续努力再创辉煌,公司决定对××部门进行10万元现金奖励。”接下来,AI算法组被试阅读的情景描述为:“管理层借助AI(人工智能)算法对××部门的××进行分析并决定他/她是否在现金奖励名单中。”上级主管组被试阅读的情景描述为:“××部门的上级主管对××进行分析并决定他/她是否在奖励名单中。”随后,请两组被试回答:“①根据上述情景描述,是谁分析并决定奖励的最终名单?②谁来决定由于业务增长进行的奖励名单?”采用Likert7点评分法,1为肯定是上级主管,7为肯定是AI算法。决策主体感知的信度系数为0.854。

2.程序公平感知、信息透明度、包容型氛围和控制变量。程序公平感知和信息透明度的测量工具与实验1一致,针对具体情境进行适当调整,两个量表在本次实验的信度系数分别为0.855和0.831。调节变量包容型氛围的测量工具选取Nishii(2013)开发的15题项量表来测量,本研究主要关注员工个体感知层面,代表题项有“在公司中,员工经常平等地分享观点、互相学习”“公司的绩效评估流程是公正的”“公司的员工收入是按劳分配的”等。在本研究中,该量表的信度系数为0.908。实验2选择的控制变量及其测量方式也与实验1一致。

### (三)实验结果

1.描述性统计及相关分析。表2描述了实验2主要变量的均值、标准差及相关分析结果。由表2及统计结果可得,被试平均年龄约34岁、平均工作年限3.6年,男性129名,占51.2%,本科及以上学历190名,占75.4%。相关分析结果显示,决策主体与信息透明度、程序公平感知存在显著的负相关关系( $r=-0.315, p<0.001$ ;  $r=-0.408, p<0.001$ );信息透明度与程序公平感知存在显著正相关关系( $r=0.425, p<0.001$ );包容型氛围与程序公平感知存在显著正相关关系( $r=0.377, p<0.001$ )。以上分析结果为研究假设提供了初步支持。

表2 实验2描述性统计与相关系数矩阵

变量	<i>M±SD</i>	1	2	3	4	5	6	7	8	9	10	11
1.性别	1.488±0.501											
2.学历	3.357±1.191	-0.006										
3.年龄	33.861±4.108	0.103	-0.038									
4.工作年限	3.560±1.672	-0.013	-0.055	0.134*								
5.职位层级	1.548±0.499	-0.038	0.045	-0.091	-0.073							
6.决策典型性	3.988±1.771	0.029	-0.079	-0.040	0.002	-0.011						
7.决策考虑因素的数量	2.179±0.872	-0.009	-0.050	-0.014	-0.050	-0.162*	-0.022					
8.决策主体感知	4.327±1.501	-0.099	0.057	-0.018	-0.008	-0.065	0.004	0.059				
9.决策主体	0.444±0.498	-0.075	0.228***	-0.044	-0.056	-0.021	-0.044	0.119	0.356***			
10.信息透明度	4.073±1.109	0.066	-0.052	0.011	-0.024	0.001	0.153*	0.019	-0.054	-0.315***		
11.包容型氛围	4.175±0.942	0.040	-0.158*	-0.038	0.008	0.052	0.109	0.059	-0.064	-0.197**	0.189**	
12.程序公平感知	4.125±1.022	-0.079	-0.109	0.023	0.040	0.055	0.072	-0.042	-0.027	-0.408***	0.425***	0.377***

注:*N*=252;\*表示 $p<0.05$ ,\*\*表示 $p<0.01$ ,\*\*\*表示 $p<0.001$ 。

2.操纵检验。单因素方差分析结果表明,针对决策主体感知的AI算法组中被试的评分显著高于上级主管组中的被试,AI算法组的均值 $M=4.924$ ,上级主管组的均值 $M=3.850$ ,两组被试的均方比值 $F(1,250)=36.369$ ,显著性系数 $p<0.001$ ,效应量系数 $\eta^2=0.127$ 。包容性氛围对决策主体感知没有影响, $p=0.899$ 不显著;决策主体操纵与包容型氛围的交互对决策主体感知没有影响, $p=0.463$ ,表明本研究对决策主体的操纵成功。

3.假设检验。首先,本研究通过实验2进一步地重复检验实验1得出的结论。单因素方差分析结果表明,决策主体对程序公平感知的影响显著,并且AI算法组的程序公平感知显著低于上级主管组被试,AI算法组的均值 $M=3.661$ ,上级主管组的均值 $M=4.497$ , $F(1,250)=49.808$ ,显著性系数 $p<0.001$ ,效应量系数 $\eta^2=0.166$ 。假设1再次得到验证。利用Bootstrap分析中的模型4以程序公平感知为被解释变量,以决策主体为解释变量,以决策典型性、决策考虑因素的数量及人口统计学变量为控制变量,再次检验信息透明度的中介作用。Bootstrap分析结果显示,信息透明度在AI算法决策对程序公平感知的影响中起部分中介作用,95%CI=[-0.360,-0.120],不包含0。不含控制变量的分析结果同样表明了信息透明度所起的部分中介作用,95%CI=[-0.329,-0.123],不包含0。假设2再次得到验证,证明本研究具有良好的外部效度。

其次,利用Process分析中的模型1检验包容型氛围的调节作用,以程序公平感知为被解释变量,以决策主体为解释变量,以包容型氛围为调节变量。Process分析结果表明,决策主体与包容型氛围的交互对程序公平感知产生显著影响, $t(241)=4.693$ , $p<0.001$ 。图1给出两个不同组中被试对程序公平感知的评分均值统计结果:当员工包容型氛围感知较低时,AI算法组的均值

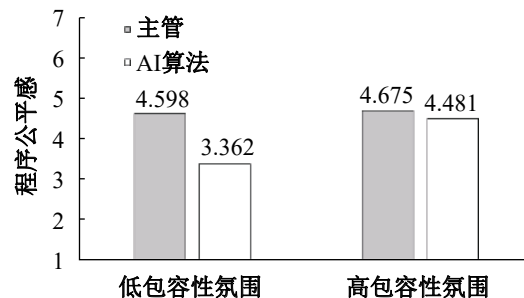


图1 研究2结果

$M=3.362$ ,上级主管组的均值 $M=4.598$ , $t(241)=-7.856$ ,显著性系数 $p<0.001$ ;当员工包容型氛围感知较高时,AI算法决策对程序公平感知的消极影响减弱,AI算法组的均值 $M=4.481$ ,上级主管组的均值 $M=4.675$ , $t(241)=-1.201$ ,显著性系数 $p>0.05$ 。数据表明,包容型氛围调节了AI算法决策对程序公平感知的主效应,当员工包容型氛围感知较高时,决策主体差异对程序公平感知的影响效应不显著,当员工包容型氛围感知较低时,这种效应更加明显。因此,假设3得到验证。进一步地,不含控制变量的分析结果同样表明包容型氛围调节AI算法决策对程序公平感知的影响:决策主体与包容型氛围的交互对程序公平感知产生显著影响, $t(248)=4.820$ , $p<0.001$ ;当员工包容型氛围感知较低时, $t(248)=-8.019$ ,显著性系数 $p<0.001$ ;包容型氛围感知较高时, $t(248)=-1.095$ ,显著性系数 $p>0.05$ ,进一步支撑了本研究的假设3。

综上所述,本研究在人力资源奖励决策情境下再次验证了AI算法决策对员工程序公平感知的影响机制,并且验证了组织包容型氛围对主效应的调节作用。实验结果发现,在具有高包容型氛围感知的员工群体中AI算法决策的主效应减弱,说明组织包容型氛围是缓解AI算法决策消极影响的边界条件。

## 五、结论与讨论

### (一)研究结论

本研究旨在围绕数字化背景下员工如何看待和感知组织借助AI算法驱动决策的公平性

问题,运用实验方法在2个不同的人力资源决策情境下探讨AI算法决策对员工程序公平感知的影响,并揭示这一影响的作用机理和边界条件。结果表明,在组织进行人力资源决策情境下,与上级主管决策相比,员工认为AI算法决策的信息透明度更低,由此对AI算法决策产生更低的程序公平感知;组织包容型氛围能够调节AI算法决策的主效应强度,即当员工在组织感知到较高水平的包容型氛围时,AI算法决策对程序公平感知的作用会减弱。本研究挑战了现有主流观点,即认为人工智能算法决策不存在或不具有与人类决策相比更低的偏见。研究结论不仅为全面、辩证地理解和科学地应用人工智能提供了新的视角,也为企业系统地构建数字化管理创新的组织环境提供实践启示。

### (二)理论贡献

本研究的理论贡献在于:第一,基于辩证的视角分析并论证了决策主体(AI算法vs.上级主管)对员工差异化体验程序公平感的影响。响应了Kuhn和Maleki(2017)强调关注AI算法决策和员工程序公平感知问题的呼吁,并挑战了现有认为与人类决策相比,由于AI算法决策能够剔除人类决策过程中的主观偏见而更加客观公正的主流观点(Waytz和Norton,2014)。本研究从理论阐释和实验分析上检验了AI算法决策为何以及如何让员工体验到更低水平的程序公平感,从决策主体视角拓展了程序公平感知的前因变量。第二,强调了信息透明度在员工如何理解HR决策程序公平性问题中的重要作用,积极拓展了程序公平理论在人工智能决策领域的价值和应用。本研究提出并论证了信息透明度的中介作用,即在引入AI算法辅助进行决策的情况下,员工认为AI算法并不能提供具有可获得、可理解、可辨认以及准确清晰等特征的决策信息,进而导致他们产生更低的程序公平感知。第三,在程序公平理论的整体框架和现有文献的基础上,引入组织包容型氛围作为AI算法决策效应的边界条件,解释了何种条件下AI算法决策导致员工体验到更低的程序公平感的负面影响会被削弱。不仅积极响应了Newman等(2020)在未来展望中提到的关于阐明算法决策影响个体公平感的边界条件的建议,同时,也将组织包容型氛围的影响效应由“人际关系”拓展到“人—技术关系”的层面。研究发现,当员工感知到较高水平的包容型氛围时,AI算法决策对其程序公平感知的影响会减弱。基于不同情境下的权变研究有益于进一步拓展本研究的外部效度,帮助学者们更加全面系统地理解人工智能背景下员工决策程序公平感知的作用机理。

### (三)实践启示

本研究的实践意义在于:第一,本研究的发现揭示了通过AI算法进行HR决策可能存在的一种负面影响,即由于员工认为AI算法复杂的决策过程信息不透明,进而导致他们体验到更低的程序公平感。因此,企业在进行数字化转型并引入人工智能协助组织决策时,需要关注在人机交互过程中员工对于诸如人工智能之类的数字化技术的感知及看法,管理者可以尝试向员工公开和共享AI算法决策过程的相关信息,增加员工对AI算法的好感甚至信任。第二,对管理者而言,让员工体验到更高水平的公平感是降低人才流失率、提高工作效率和组织承诺的保障(Lee,2018;Newman等,2020)。了解采取何种措施有利于在引入AI算法优化决策环境并提升决策效率的同时,也能够抑制或减弱AI算法决策的负面影响,这对于组织数字化转型过程中构建积极的人机交互关系并实现管理决策的预期效果十分重要。考虑到本研究的发现,即当员工感知到较高水平的包容型氛围时,决策主体差异对其程序公平感知的影响会被减弱,管理者应当致力于在组织构建尊重差异、决策兼容以及雇佣公平的包容型氛围,公平对待每个人,接纳或重视不同见解,并鼓励每个人的决策地位(Nishii,2013),同时鼓励员工参与决策,设置申诉渠道,并努力增强决策过程的确定性。第三,值得注意的是,对处于数字化转型阶段的企业

而言,过分依赖大数据驱动的AI算法进行决策可能存在潜在的风险,一方面,员工感知到的公平感持续下降会导致一系列职场偏差问题逐渐凸显。另一方面,对AI算法的过度依赖可能会增加组织的道德风险,以最优决策为标准的AI算法会忽视管理伦理问题。因此,管理者需要考虑管理效率和道德伦理之间的平衡。

#### (四)研究局限与展望

本研究的局限性在于:第一,仅仅通过情景材料编写人为地创造实验条件可能会使假设检验远离现实情境中的自然状态,影响本研究的外部效度以及对管理实践指导的针对性。未来研究可以尝试田野实验、二手数据、情感分析等多种方法检验本研究的各项假设。例如,可以尝试选择分别由AI算法决策和上级主管决策的两家企业,利用网络爬虫技术抓取员工邮件中与决策内容相关信息,并结合自然语言处理等方法进行语义情感分析,检验不同决策主体带来的程序公平感知之间的差异。第二,本研究基于程序公平理论选取了信息透明度作为员工理解人工智能背景下决策主体差异影响程序公平性问题的关键因素,未来研究可以跳出程序公平理论的框架,进一步地拓展新的作用机理。例如,社会认同理论强调个体倾向于通过社会分类而对自己的群体产生认同,并通过维持积极的群体认同来提高自尊,群体认同会影响个体对公平感的评价(Taylor等,1995)。因此,员工在针对人工智能认知评价的过程中,可能会将自己划分为与上级主管一致的人类群体,相应地将人工智能划分为外部群体,进而产生内群体偏好和外群体偏见,这会影响他们对AI算法决策程序公平感的体验。第三,尽管本研究检验了包容型氛围在AI算法决策对员工程序公平感知的主效应中的调节作用,但并不能减弱员工对AI算法决策产生更高信息不透明性的感知。未来可以探讨何种干预措施能够对中介效应产生调节作用。此外,在AI算法进行人力资源决策的新情境下,本研究在针对包容型氛围的调节作用进行阐述时,部分内容可能超出了包容型氛围的内涵。例如,在良好包容型氛围中的员工可能对新技术更加具有包容性和开放性,这似乎与Nishii(2013)对于包容型氛围概念内涵的界定相对不符。未来研究可以寻找更加贴切的调节变量,进一步论证本研究提出的理论假设。第四,本研究只考察了目标对象针对员工进行人力资源决策的情境,但当人工智能的组织决策内容针对业务、设备等客观主体时,则AI算法决策可能被员工视为相对公平的。未来可以将AI算法决策的具体情境拓展到组织战略、市场营销、财务会计等其他商业环节,进一步考察AI算法决策相较于上级主管决策的差异化影响。

#### 主要参考文献

- [1]刘善仕,裴嘉良,钟楚燕.平台工作自主吗?在线劳动平台算法管理对工作自主性的影响[J].外国经济与管理,2021,43(2):51-67.
- [2]徐鹏,徐向艺.人工智能时代企业管理变革的逻辑与分析框架[J].管理世界,2020,36(1):122-129.
- [3]Bal P M, de Lange A H, Ybema J F, et al. Age and trust as moderators in the relation between procedural justice and turnover: A large-scale longitudinal study[J]. *Applied Psychology*, 2011, 60(1): 66-86.
- [4]Bernerth J B, Aguinis H. A critical review and best-practice recommendations for control variable usage[J]. *Personnel Psychology*, 2016, 69(1): 229-283.
- [5]Burrell J. How the machine “thinks”: Understanding opacity in machine learning algorithms[J]. *Big Data & Society*, 2016, 3(1): 1-12.
- [6]Colquitt J A, Scott B A, Rodell J B, et al. Justice at the millennium, a decade later: A meta-analytic test of social exchange and affect-based perspectives[J]. *Journal of Applied Psychology*, 2013, 98(2): 199-236.
- [7]Dietvorst B J, Simmons J P, Massey C. Algorithm aversion: People erroneously avoid algorithms after seeing them err[J]. *Journal of Experimental Psychology:General*, 2015, 144(1): 114-126.

- [8]Duan Y Q, Edwards J S, Dwivedi Y K. Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda[J]. *International Journal of Information Management*, 2019, 48: 63-71.
- [9]Edwards J S, Duan Y, Robins P C. An analysis of expert systems for business decision making at different levels and in different roles[J]. *European Journal of Information Systems*, 2000, 9(1): 36-46.
- [10]Hollenbeck J R, Ilgen D R, LePine J A, et al. Extending the multilevel theory of team decision making: Effects of feedback and experience in hierarchical teams[J]. *Academy of Management Journal*, 1998, 41(3): 269-282.
- [11]Kacmar K M, Witt L A, Zivnuska S, et al. The interactive effect of leader-member exchange and communication frequency on performance ratings[J]. *Journal of Applied Psychology*, 2003, 88(4): 764-772.
- [12]Kim W C, Mauborgne R. Procedural justice, strategic decision making, and the knowledge economy[J]. *Strategic Management Journal*, 1998, 19(4): 323-338.
- [13]Kuhn K M, Maleki A. Micro-entrepreneurs, dependent contractors, and instaservers: Understanding online labor platform workforces[J]. *Academy of Management Perspectives*, 2017, 31(3): 183-200.
- [14]Lambrecht A, Tucker C. Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads[J]. *Management Science*, 2019, 65(7): 2966-2981.
- [15]Lee M K. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management[J]. *Big Data & Society*, 2018, 5(1): 1-16.
- [16]Lee M K, Kusbit D, Metsky E, et al. Working with machines: The impact of algorithmic and data-driven management on human workers[A]. Proceedings of the 33rd annual ACM conference on human factors in computing systems[D]. Seoul, Republic of Korea: Association for Computing Machinery, 2015.
- [17]Leventhal G S. What should be done with equity theory?[A]. Gergen K J, Greenberg M S, Willis R H. Social exchange: Advances in theory and research[M]. Boston: Springer, 1980.
- [18]Li D Y, Du Y. Artificial intelligence with uncertainty[M]. Boca Raton: Chapman & Hall/CRC, 2007.
- [19]Liu Y Y, Eisingerich A B, Auh S, et al. Service firm performance transparency: How, when, and why does it pay off?[J]. *Journal of Service Research*, 2015, 18(4): 451-467.
- [20]Miller F A. Strategic culture change: The door to achieving high performance and inclusion[J]. *Public Personnel Management*, 1998, 27(2): 151-160.
- [21]Miron M, Tolan S, Gómez E, et al. Evaluating causes of algorithmic bias in juvenile criminal recidivism[J]. *Artificial Intelligence and Law*, 2021, 29(2): 111-147.
- [22]Newman D T, Fast N J, Harmon D J. When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions[J]. *Organizational Behavior and Human Decision Processes*, 2020, 160: 149-167.
- [23]Nishii L H. The benefits of climate for inclusion for gender-diverse groups[J]. *Academy of Management Journal*, 2013, 56(6): 1754-1774.
- [24]Orlikowski W J, Scott S V. What happens when evaluation goes online? Exploring apparatuses of valuation in the travel sector[J]. *Organization Science*, 2014, 25(3): 868-891.
- [25]Schildt H. Big data and organizational design—The brave new world of algorithmic management and computer augmented transparency[J]. *Innovation-Organization & Management*, 2017, 19(1): 23-30.
- [26]Shore L M, Randel A E, Chung B G, et al. Inclusion and diversity in work groups: A review and model for future research[J]. *Journal of Management*, 2011, 37(4): 1262-1289.
- [27]Strohmeier S, Piazza F. Artificial intelligence techniques in human resource management—A conceptual exploration[A]. Kahraman C, Onar S Ç. Intelligent techniques in engineering management[M]. Cham: Springer, 2015.
- [28]Taylor M S, Tracy K B, Renard M K, et al. Due process in performance appraisal: A quasi-experiment in procedural justice[J]. *Administrative Science Quarterly*, 1995, 40(3): 495-523.
- [29]Thau S, Bennett R J, Mitchell M S, et al. How management style moderates the relationship between abusive supervision and workplace deviance: An uncertainty management theory perspective[J]. *Organizational Behavior and Human Decision Processes*, 2009, 108(1): 79-92.

# Can AI Algorithmic Decision-making Improve Employees' Perception of Procedural Fairness?

Pei Jialiang, Liu Shanshi, Zhong Chuyan, Chen Yifan

(*School of Business Administration, South China University of Technology, Guangzhou 510641, China*)

**Summary:** With the development of digital technologies such as big data, artificial intelligence and machine learning, many organizations try to introduce AI algorithms into the decision-making process to avoid the subjective inherent biases. Although AI algorithms improve the scientificity of decision-making, they also raise the issue of fairness in the decision-making process. Considering that employees are the direct subject of HR decision-making, how employees view and evaluate the fairness of AI algorithmic decision-making process is very important. However, researches on the above questions are rather limited.

Based on the procedural fairness theory, this study discusses the impact of different decision-making subjects (AI algorithms vs. supervisors) on employees' perception of fairness in the HR decision-making process. Through two experiments, we find that: (1) The decision-making of AI algorithms makes employees have a lower perception of procedural fairness than that of supervisors. (2) Information transparency mediates the relationship between AI algorithmic decision-making and employees' perception of procedural fairness. Employees believe that AI algorithmic decision-making has lower information transparency, which in turn produces a lower perception of procedural fairness. (3) Inclusive climate moderates the relationship between AI algorithmic decision-making and employees' perception of procedural fairness. When employees perceive a high level of organizational inclusive climate, the effect of AI algorithmic decision-making on the perception of program fairness will be weakened.

Overall, this study responds to the call for more research that bridges AI algorithmic decision-making and individual perception of procedural fairness by examining information transparency as a mediator and inclusive climate as a moderator in this relationship. The findings not only challenge the existing mainstream view that AI algorithmic decision-making does not exist a lower bias than human decision-making, but also offer useful advice to managers and provide implications for organizational HR decision-making practices in the digital and intelligent era.

**Key words:** AI; algorithms; HR decision-making; information transparency; procedural fairness

(责任编辑:宋澄宇)