

DOI: 10.16538/j.cnki.fem.20210818.104

## 算法决策、员工公平感与偏差行为： 决策有利性的调节作用

魏 昕, 黄鸣鹏, 李欣悦

(对外经济贸易大学 国际商学院, 北京 100029)

**摘 要:** 随着数字化时代的到来, 算法决策在企业管理中的应用越来越广泛, 算法决策对员工造成的心理及行为层面的影响, 成为一个具有重要理论意义与实践意义的问题。本文结合算法管理、公平等领域研究, 采用归因视角来分析决策者类型(算法决策、领导决策)和决策有利性的交互作用对员工公平感, 以及对员工的组织偏差行为和社会偏差行为的影响。我们开展了一个质化的预研究和一个实验研究来检验受调节的中介模型。研究表明: 在面临不利决策时, 员工觉得算法决策比领导决策更公平, 从而展现出更少的组织偏差行为和社会偏差行为; 而这些差异在面临有利决策时不显著。本研究是对近来兴起的关于算法如何影响员工感知与行为的研究的重要发展, 也为算法时代的管理决策提供了实践启示。

**关键词:** 算法决策; 决策有利性; 公平感; 偏差行为

**中图分类号:** F270 **文献标识码:** A **文章编号:** 1001-4950(2021)11-0056-14

### 一、引 言

随着科学技术的进步和发展, 我们迎来了数字化时代。而算法作为数字化时代的产物之一, 成为帮助企业快速进行决策及判断的重要工具, 被广泛应用于企业管理中, 例如为网约车司机计算最有效的路线、为外卖骑手规划多个订单的取送餐顺序等。由于算法决策在现实中日益普遍的应用, 并对企业和身处其中的员工产生重要影响, 研究者们也开始越来越关注员工对算法决策的感知和反应。关于员工对算法的感知, 现有文献中存在两种总体观点。一种是算法欣赏(algorithm appreciation), 认为与人类的判断相比, 算法更加高效、准确、客观, 因此人们会更加信任算法决策(如, Logg等, 2019)。另一种则是算法厌恶(algorithm aversion), 认为相比于算法决策, 人工决策可以允许员工在决策过程中有更高的参与和控制感, 因此人们会更倾向于选择人工决策(Dietvorst等, 2015, 2018)。

收稿日期: 2021-03-31

基金项目: 国家自然科学基金项目(71672035, 72072032)

作者简介: 魏 昕(1986—), 女, 对外经济贸易大学国际商学院教授(通讯作者, weixin@uibe.edu.cn);

黄鸣鹏(1988—), 男, 对外经济贸易大学国际商学院副教授;

李欣悦(1998—), 女, 对外经济贸易大学国际商学院硕士研究生。

与这两种总体观点紧密相关的是算法决策(相对于人工决策)对员工公平感的影响。公平感是一种重要的工作感知,对员工的诸多工作行为和结果会产生重要影响(Colquitt等,2001)。因此,近年来,算法决策如何影响员工的公平感也越来越受到研究的关注,但并未呈现一致的研究结果。具体来说,一些研究发现,相对于人工决策,算法决策会带来更高的公平感,因为算法决策具有更高的一致性、无偏性、使用的信息更加准确等(如,de Melo等,2017;Schlicker等,2021)。然而,另一些研究则发现算法决策未必会导致更高的公平感,因为人们认为算法没有考虑质化信息、不够情境化、无法进行主观评价(如Lee,2018;Newman等,2020)。这些不一致的研究发现意味着,在决策者类型(算法决策vs.人工决策)与员工公平感的关系中可能存在重要的调节因素,使得两者的关系在不同情境下呈现出不一样的结果。探讨这些潜在的调节因素对于我们更全面、具体地理解算法决策对员工公平感及其后续行为的影响非常重要。因此,本研究的第一个目标是探讨对决策者类型与员工公平感之间关系起到调节作用的情境变量。

为了探讨这一问题,我们将采用归因理论(Kelley和Michela,1980)的视角。根据归因理论,员工会对涉及自身的决策进行解释,而这种解释会影响他们对于该决策的反应。例如,已有研究表明,个体的归因过程会对其公平感知产生影响(如Ployhart和Ryan,1997)。特别地,人们面对有利和不利的结果,会进行不同的归因。关于自我服务归因偏差(self-serving attribution bias)的研究发现,当面对有利的决策结果时,员工更倾向于归因于自身因素,因此其公平感较少受外部因素(如决策者特征)的影响。在这种情况下,算法决策或人工决策对员工的公平感不会有差异化的影响。相反,当决策结果不利时,员工更倾向于归因于外部因素,如决策者(Bradley,1978)。此时,由于员工对于算法和人工这两类决策者的无偏性、一致性、准确性等方面有不同的感知(如de Melo等,2017;Lee,2018;Newman等,2020;Schlicker等,2021),会产生不同程度的公平感。因此,基于归因理论,我们认为决策结果的有利性可能是影响决策者类型与员工整体公平感之间关系的一个重要调节变量。

决策者类型与决策有利性的交互作用不仅会影响员工的公平感知,还会影响员工后续的行为结果。其中特别值得关注的是员工的偏差行为。员工在应对算法的过程中表现出的针对组织或社会的偏差行为已经引起了广泛关注。例如,2020年的一篇引起社会广泛讨论的报道——《外卖骑手,困在系统里》提到,骑手们可能因为抱怨系统派单不合理、时间紧急等原因,以超速、逆行、闯红灯等违反社会规则的做法,或者采取在没有到达送餐地点之前就提前点击“送达”按钮等一系列对抗系统的行为,来达到系统要求的时效(赖祐萱,2020)。关于公平的研究也发现,当员工感到不公平时,会展现出更多的偏差行为(Colquitt等,2013)。因此,本研究将进一步探讨决策有利性如何调节决策者类型通过员工公平感对员工偏差行为产生的间接影响。需要指出的是,在本研究中,我们聚焦于员工的整体公平感,即员工对组织各个方面公平程度的总体感知(Jones和Martens,2009)。尽管研究者详细区分了分配公平、过程公平、互动公平等不同类型的公平感,但在现实生活中,人们对与公平相关的信息进行感知时,往往很难清楚区分公平的各个子维度,而是形成一个整体的公平感(Lind,2001)。

本研究将对算法决策的相关文献做出几个方面的重要贡献。第一,本研究揭示了决策有利性的调节作用,拓展了现有文献中关于算法决策(相对于人工决策)如何影响员工公平感的理解。目前既有一些研究发现算法决策(相对于人工决策)能够带来更高的员工公平感(如,de Melo等,2017;Schlicker等,2021),也有研究发现相反的结果(如,Lee,2018;Newman等,2020)。因此,要更全面地理解决策者类型对员工公平感的影响,就需要探讨潜在的情境因素。本研究基于归因理论,引入并揭示了决策有利性的调节作用,因而提供了更加精细的视角来探索算法决策对员工公平感的影响。第二,更广泛地说,我们的研究也对算法决策如何影响员工感知的

文献有所贡献。人们究竟是欣赏算法(如,Logg等,2019)还是厌恶算法(Burton等,2020;Dietvorst等,2015,2018),目前的文献中并没有一致的观点或结论。我们的研究结果意味着,人们对算法的欣赏或厌恶是随着情境发生变化的。在决策结果不利时,人们更加“欣赏”算法;而在决策结果有利时,算法或人工决策带来的感知上的差异并不显著。因此,本研究的结果有助于深化我们理解人们对算法决策的感知。第三,本研究进一步分析了员工如何以偏差行为来应对算法决策或人工决策对其公平感带来的影响,将关于算法对员工影响的研究视域(scope)扩大到员工的行为应对策略,是对这一研究领域的重要拓展。

## 二、理论基础与研究假设

### (一)决策者类型与决策有利性对公平感的交互影响

在组织中,当面临某个决策结果时,员工会探究其原因。如果员工将决策结果归因于内部,即自身因素(如过往业绩),那么其公平感取决于对自身因素的解读,而较少受外在因素的影响。相反,如果员工将决策结果归因于外在因素(如决策者类型),其公平感则取决于员工对该外部因素的解读。以往研究表明,当面临不利的结果时,人们倾向于进行外部归因(Wolosin等,1973)。因此,当决策结果对员工不利时,员工会试图从决策者的无偏性、其决策的一致性、准确性等外部因素去解释不利的决策结果,由此产生不同程度的公平感(Colquitt,2001;Shin和Park,2019)。如果决策者在这些方面存在问题,员工就会认为不利的结果是由决策者的错误导致的,进而认为不公平。

本文认为,当不利决策是由人工(如领导)做出的,员工更容易对决策者的无偏性、其决策的一致性、准确性等方面产生质疑,因而更容易将不利结果归因于决策者的问题,导致更低的公平感。相反,算法决策相比起人工决策来说,在这些方面具有优势,员工更难将不利结果归因于决策者(算法)的问题,因而其公平感受到的负面影响较小。具体来说,第一,算法不同于人类,不受情绪、偏好等因素带来的个人偏差的影响(Gray和Wegner,2012)。关于公平的研究指出,即便决策结果是不利的,例如员工得到了较低的绩效评估,但如果决策过程是无偏的,则有助于提升员工的公平感(Taylor等,1995)。因此,在面临不利决策时,算法决策过程的无偏性使得员工认为算法(相对于人工)的决策过程更加公平。近来一项对优步司机的访谈发现,当算法参与到管理过程当中分配任务时,司机们肯定了算法作为决策者的无偏、中立的特征(Lee等,2015)。这一发现间接地支持了我们的观点。第二,由于算法决策每次都遵循相同的程序,员工会认为算法决策的一致性更强,从而认为其决策更加公平。研究者们在多种不同的决策场景中发现,人们认为自动系统(算法)做决策的过程和结果都具有更高的一致性,这些场景包括招聘面试(Langer等,2020)、日常管理(Schlicker等,2021)、任务分配(Ötting和Maier,2018)、职业培训(Ötting和Maier,2018)等。而决策的一致性是影响人们公平感的重要因素——一致性意味着,对所有潜在的决策接受者,按照稳定的标准进行决策。这样的机会均等,带来了较高的公平感知(Leventhal,1980)。第三,人们倾向于认为,和人类决策相比,算法在决策时更加准确。随着算法技术的进步,人们对算法的可靠性和准确性的信心逐渐提高,认为算法具有远远高于人类的信息处理能力和效率(Lee,2018)。如果某个信息的来源是计算机,人们认为该信息质量更高、更加客观(Sundar和Nass,2001)。近来的一项质性研究的结果也发现,人们认为算法决策具有更高的准确性,因此感到更加公平(Shin和Park,2019)。

尽管有些研究者指出,相比于算法决策,人工决策存在一些优势,例如能考虑到一些无法被算法量化的质化信息(Newman等,2020)、具有算法所不具备的主观判断能力(Lee,2018),进而对员工的公平感产生更积极的影响。但同时,质化信息的模糊性、与主观判断相伴的人为

偏差,又不可避免地会削弱员工感知到人工决策的公平性(Lee,2018)。与这一逻辑相关的是,研究发现,在施以员工一些对他们并不友好的举措时——例如以非建设性的方式给他们反馈——如果是主管施以这些举措(和数字化技术施予的相比),其中所蕴含的情感内涵和人际敏感性使得员工的公平感更低(Alder和Ambrose,2005)。或者当员工面临工作中的监控这一可能侵犯自己隐私的举措时,比起主管的观察,数字化监控让员工感到更公平(McNall和Roch,2007)。因此,我们认为,尽管人工决策存在一些优势,但在决策结果不利、员工倾向于归咎于决策者时,人工决策者(相比于算法决策系统)更有可能被归咎,进而导致更低的员工公平感。

然而,当决策结果对员工有利时,由于人们倾向于将有利的结果进行内部归因(Woloshin等,1973),他们更有可能将这一有利结果归功于自己(如自身业务能力、过往业绩),而很少归因于决策制定者或其他外部原因。在这种情况下,决策者类型的差异不会对员工公平感产生显著的影响。

综上,我们认为,在不利决策的情况下,相对于人类领导者决策,算法决策下员工的公平感更高。在有利决策的情况下,决策者类型不会对员工的公平感产生显著的影响。因此,本文提出如下假设:

假设1:决策者类型与决策有利性交互影响了员工的公平感。具体来说,当面临不利决策时,与领导决策相比,算法决策下的员工公平感更高;当面临有利决策时,决策者类型对员工公平感没有显著影响。

## (二)公平感与偏差行为

当员工感知到不公平时,可能会寻求改变,如通过改变组织制度的方式来达到组织公平。但是,如果员工认为他们不能改变制度,只能通过调整自己的行为来降低这种不公平感(Aquino等,1999)。一些关于公平的研究基于社会交换的视角,认为员工与组织的雇佣关系本质上是一种交换关系,员工通过衡量自身付出和回报是否对等,来改变自己在工作上的投入程度(Colquitt等,2013)。当员工感知到不公平时,会减少对组织的付出,使其与感知到的组织回报相匹配。比如他们可能展现出更多的组织偏差行为,即自发地违反组织规范,损害组织及其他成员利益的行为。研究发现,当组织中对公平的感知较低时,员工会以破坏性的行为来报复(Colquitt等,2001)。

此外,也有一些研究者从情绪的视角来解释不公平感对于员工行为的影响——不公平感会导致员工出现负面情绪,例如愤怒。这些负面情绪驱使他们做出报复性举动,包括反生产行为、偏差行为等(Colquitt等,2013)。结合以上两种视角(社会交换视角、情绪视角),我们认为,员工会做出组织偏差行为来回应他们感知到的不公平。

假设2a:员工的公平感对其组织偏差行为呈负向影响。

除组织偏差行为之外,公平感还会对员工的社会偏差行为产生影响。社会偏差行为是指一系列违反社会规范、破坏社会秩序,给社会造成不良影响的行为(Hay等,2010)。不同于组织偏差行为,社会偏差行为的实施对象是在组织之外的。当员工遭受到管理者的伤害时,可能会将不满发泄到其他非伤害源上(Marcus-Newhall等,2000)。例如,有研究发现,当员工遭受不公平的待遇时,他们会觉得自己受到了管理者的侵害,这时他们不仅会对组织实施发泄,还会将这种发泄行为渗透至非工作领域,如家庭领域(Hoobler和Brass,2006)。我们认为非工作领域不仅局限于家庭,当员工觉得自己受到组织的侵害时,产生的转移攻击行为也可能会溢出至社会领域,出现一系列违反社会规则的行为。

假设2b:员工的公平感对其社会偏差行为呈负向影响。

### (三)受调节的中介效应

根据假设1,在不利决策的条件下,与领导决策相比,员工对算法决策的公平感知更高。而根据假设2,当员工公平感知较低时,会做出对抗组织的行为,即组织偏差行为,甚至蔓延到对社会规则的对抗,展现出社会偏差行为。综合以上两部分的效应,我们推论,算法决策对员工公平感的影响会进一步传递到他们的偏差行为上,并且,这些间接影响在有利决策和不利决策的条件下呈现出不同的强度。即决策者类型与偏差行为之间存在受调节的中介效应。

假设3a:比起面临有利决策,当员工面临不利决策时,决策者类型通过员工公平感对其组织偏差行为形成的间接效应更强。

假设3b:比起面临有利决策,当员工面临不利决策时,决策者类型通过员工公平感对其社会偏差行为形成的间接效应更强。

综上所述,本文提出如图1所示的研究模型。为了检验这一模型,我们开展了一个质化的预研究,以及一个正式的实验研究。在预研究中,我们对30位骑手开展了半结构化访谈,了解在他们的工作中算法或人工是如何进行决策的、他们对算法决策和人工决策的公平感知,以及他们如何应对(不)公平感。根据预研究的结果,我们设计了一个研究情境,对关键变量进行实验操纵,开展了一个2×2组间设计的实验,来检验各个假设。

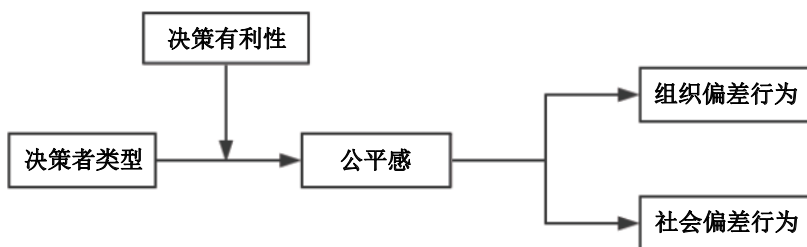


图1 研究模型

### 三、预研究

#### (一)样本选择与实施过程

为了更好地设计实验情境,使其具有较高的心理现实性(psychological realism),我们在正式研究之前对30名外卖骑手进行了访谈,了解他们的日常工作场景、算法决策与领导决策在他们工作中的应用,他们对不同决策的感知和应对。受访骑手的平均年龄为28岁(标准差=5.45),其中男性占100%,他们在该平台的平均工作年限是1.07年(标准差=0.41),平均每天工作时长9.41小时(标准差=1.82)。

我们采取半结构化访谈的方式,访谈内容包括人口统计学信息、与平台管理相关的问题、与站点管理相关的问题、与客户相关的问题、与商家相关的问题。每位骑手的访谈时间平均持续30分钟,在征得被访谈者同意的情况下,我们对所有访谈进行了录音并转为文字底稿。

#### (二)数据整理与分析

由于本研究的目的是提取典型工作场景、感知和行为,并非开创新的理论构念(constructs),我们对Gioia等(2013)的方法进行简化,仅仅从访谈内容中提取一级目录(first-order categories)和二级主题(second-order themes),并结合量化的方法进行分析。首先,本文作者各自独立地对1/3的访谈进行编码,产生若干个一级目录。接下来对这些一级目录与对应的访谈内容展开讨论,对一级目录进行修改,以确保编码的方式一致。在此基础上,本文作者独立地对剩下2/3的访谈进行编码,并在完成后基于一致性原则(consensus)对所有有争议的编码进行讨论和修改,最终形成10个一级目录。同时,我们也对每个一级目录出现的频次进行了统计。在此过程

中,当没有更多新的一级目录产生时,我们认为一级目录达到饱和,进入下一步,即根据一级目录之间共同的属性形成二级主题。我们并未将二级主题进一步归类为聚类维度(aggregate dimensions),因为本研究关注的是现有构念在新场景下的表现形式。

### (三)研究结果

我们整理访谈内容(参见表1)发现,外卖骑手的订单分配是一个同时存在算法(平台系统)决策和人工(配送站点的站长)决策的典型场景。一方面,算法根据订单距离、商家距离、骑手位置、骑手所持订单数量等因素,决定订单分配到哪个站点的哪位骑手。另一方面,站长有时会根据具体情况对订单进行协调分配,例如有骑手的电动车在路上出了故障,无法送单;极端天气导致骑手送单时间超出算法的预估值,无法按时完成其他订单的配送;系统中的距离、位置等没有及时更新;等等。

表 1 骑手关于算法/人工决策的公平感、偏差行为

二级主题	一级目录	内容举例	比例
算法(系统) 派单公平	规则公开透明	系统派单规则是显而易见的,按app中的骑手等级派单,有青铜、白银、黄金、王者,等级越高,派单量就越高,还有专属的派单特权。跑的单越多,很快就能成为王者,跑不够单就只能青铜	16.67%
	综合考虑各种因素	系统派单是根据骑手与商家的距离、骑手账号准时率、送达时长、差评百分比等指标综合考量的	10%
算法(系统) 派单不公平	时间安排过于紧张	送餐高峰时,按照系统规定的时间根本送不过来	6.67%
	不会考虑到特殊情况	如果是那种特别热的大夏天和暴雨暴雪的天气,其实我们的订单数还是会增加,最后一爆单我们也送不过来,但又不得不去送	10%
	算法失灵	系统有时会出现数据更新不及时的情况	3.33%
人工(站长) 派单公平	根据情况需要进行调单	后台可以人工调整,人工可以把分配得不顺的单调顺一点,比较公平	20%
人工(站长) 派单不公平	规则不公开透明	有的时候会不公平吧,因为看骑手社区里有人说站长把好的单子派给他送礼的人了,派单制度不透明	6.67%
	根据私人关系决定派单	如果说你是站长,我是你朋友,我是个骑手,你很可能把近的单子先给这个人派,把远单子让别的骑手跑,我们是没有办法发现的	23.33%
组织偏差行为	提前点击“送达”后,关机、断网或卸载app	提前点送达站长会让你关机,这样系统就无法捕捉到骑手的轨迹,配送完这单回到站点再开机,有时候可以避免处罚	13.33%
社会偏差行为	通过违反交规(闯红灯、逆行)的方式来保证时效	工资相比于同水平的工作算是比较高的,不满意就是太危险了,敢问有几个骑手为了准时没有闯红灯、没有逆行过	20%

63.33%的骑手提到,在日常工作中曾经有过被分配到不顺路的订单这样的不利决策。例如,在订餐高峰时段或者在恶劣天气的情境下,订单数量暴涨(23.33%);或者在突发情况下,区域内交通异常,导致算法失灵(10%)。在这些情况下系统会分配给骑手一些不顺路的订单。在站长进行人工派单的场景中,他们有时也会出于某些个人的原因,将一些远距离或不顺路的订单分配给骑手(13.33%)。

26.67%的骑手认为系统派单是公平的,因为系统派单的规则透明(16.67%)、能综合考虑各种因素(10%)。16.67%的骑手认为系统派单也会出现不公平的情况,例如派单时间安排过于紧张(6.67%)、不会考虑到特殊情况,如恶劣天气(10%)。

有20%的骑手认为站长的人工派单决策是公平的,因为会根据情况需要进行调单。30%的

骑手认为站长派单决策会存在不公平的情况,例如派单规则不公开不透明(6.67%)、根据私人关系决定派单(23.33%)。

骑手送单超时被惩罚,是他们在工作中经常面临的不利结果。23.33%的骑手会采取提前点击“送达”后关机、断网或卸载app等方式来“欺骗”系统,即展现出组织偏差行为;20%的骑手曾经为了能按时完成订单发生过违反交通规则的行为,即展现社会偏差行为。

基于以上结果,以及骑手对其工作场景的描述,我们在实验研究中设计了一个情境。

#### 四、实验研究

根据预研究的结果,我们采用2(算法决策vs.领导决策)×2(有利决策vs.不利决策)的实验设计,研究决策者类型和决策有利性对公平感,以及组织偏差行为和社会偏差行为的影响。

##### (一)实验对象与实施过程

我们通过问卷星开展了线上实验,参与者为来自互联网、金融、教育、制造等不同行业的388名企业员工。参与者的平均年龄为30岁(标准差=7.5),男性占51.03%,本科学历及以上占60.31%。16.46%的实验参与者正在从事或者有从事骑手工作的经历。

我们将被试随机分配至4组实验情境的其中一组,每组实验条件下的人数如表2所示。在该情境中,我们要求所有参与者扮演一个外卖平台骑手的角色。参与者

表2 各种实验条件下的人数

	有利决策	不利决策
算法决策	95	103
领导决策	92	98

首先阅读了一段模拟情境的描述,然后按照在模拟情境中的真实想法回答关于公平感及偏差行为的相关问题。情境材料的背景信息如下。

A平台是一家外卖平台,主营即时配送业务。该平台拥有多个配送站点,每个配送站点有若干名骑手和一名管理骑手的站长。平台依据骑手业绩将其分为青铜、白银、黄金、钻石四个等级。骑手的收入来自于送单的配送费,每单收入在5元左右。

平台开发了一套系统,在收到顾客订单后,系统会将订单分配给距离商家4公里以内的某一名骑手,由该骑手在指定时间内前往商家取餐,并将餐品送至顾客处。平台要求商家在规定的时间内出餐,若超过规定时间未出餐,骑手可以投诉商家,从而影响商家的信誉度。

如果骑手超过了系统规定的时间才将产品送达顾客,将被视为超时,会受到惩罚。另外,系统可以监测骑手的实时位置。如果骑手在超过顾客200米外的距离提前点“送达”,将被视为违规,也会受到惩罚。惩罚包括罚款、降等级等。

同时,站长也有一定的权限,可以对所在站点骑手的订单进行分配调整。根据以往的情况来看,骑手手头的订单不会同时超过4单。

现在请你想象,自己是这个平台上某个站点的骑手,你的等级是白银。一天下午你去一家快餐店取餐,恰巧在店里遇到了三位骑手朋友,他们和你都在同一个站点。在等单过程中,你通过闲聊了解到,他们手中都各有一个东南方向的订单,而你有三个西北方向的订单。

##### (二)实验操纵

接下来,我们对决策制定者类型及决策有利性进行了操纵。

###### (1)决策者类型操纵

算法决策:就在你们收到餐品准备开始送餐时,你突然接到系统新派的订单。

领导决策:就在你们收到餐品准备开始送餐时,你突然接到站长新派的订单。

###### (2)决策有利性操纵

有利决策:你突然接到……(系统/站长)新派的订单,让你去距离自己1公里左右的商家处

取餐,然后送往某西北方向的顾客处。凑巧的是,这个……(系统/站长)新派的订单和你现在手头所有的订单方向正好相同,你觉得自己可以将所有单子都准时送达。

不利决策:你突然接到……(系统/站长)新派的订单,让你去距离自己3公里左右的商家处取餐,然后送往某东南方向的顾客处。不巧的是,这个……(系统/站长)新派的订单和你现在手头所有的订单方向正好相反,你觉得自己很难将所有单子都准时送达。

### (三)变量测量

公平感:参照Konovsky和Folger(1991)在实验研究中测量公平感的方式,参与者阅读完情境材料后,在量表上回答“我认为这一单指派给我是公平的”(1=强烈不同意,6=强烈同意)。

社会偏差行为:接下来,我们给参与者描述一个涉及社会偏差行为的情境:“离开餐厅后,你以最快的速度赶往顾客处,按照这个速度应该可以在规定时间内将站长(系统)新指派的订单准时送达。但是在通过最后一个路口前,信号灯突然由绿变红,你环顾四周后发现周围并没有车和交警。这时你会怎么做?”参与者回答他们在多大程度上“会闯红灯,快速通过路口”(1=根本不可能,6=非常可能)。

组织偏差行为:进一步地,我们描述了一个涉及组织偏差行为的情境:“按照平台规定超过顾客所在地200米外点送达属于违规行为,但你在距离顾客500米的时候,眼看就要超时了。这时你会怎么做?”参与者回答他们在多大程度上“会提前点送达,并采取一些手段避免被惩罚(如在提前点送达之后马上关机,这样系统就无法检测到你的行动轨迹)”(1=根本不可能,6=非常可能)。

另外,以往研究表明,性别、年龄等个体因素对员工的公平感和偏差行为具有一定的影响(Jones和Martens,2009)。因此,我们将性别、年龄等人口统计学变量作为控制变量。是否具有骑手工作经历可能会影响实验参与者的代入感,进而影响研究结果。因此,我们将参与者是否具有骑手工作经历也列为控制变量。

### (四)研究结果

#### 1. 操纵检验

在实验参与者完成了所有问题后,我们对实验操纵进行了检验。针对决策者类型,我们请实验参与者回答“你现在手上的订单是系统派的还是站长派的?”(1=系统,2=站长)。针对决策有利性,请他们回答“这个订单与你之前送的三个订单方向相同还是相反?”(1=相同,2=相反),以及“新派的订单距离你多远?”(1=非常远,6=非常近)。独立样本t检验的结果表明,领导决策组( $M=1.98, SD=0.144$ )和算法决策组( $M=1.02, SD=0.141$ )在派单者的选择得分上有显著差异( $t=66.26, p<0.01$ ),说明决策者类型变量操纵有效。除此之外,有利决策组( $M=1.01, SD=0.10$ )和不利决策组( $M=1.98, SD=0.14$ )在派单任务方向的选择得分上有显著差异( $t=-77.181, p<0.01$ ),有利决策组更倾向于认为派单任务方向和目前的订单方向相同,不利决策组更倾向于认为派单任务方向和目前的订单方向相反。在派单任务距离上,有利决策组( $M=4.45, SD=0.91$ )和不利决策组( $M=2.71, SD=1.18$ )的选择得分也存在显著差异( $t=16.22, p<0.01$ ),相比于有利决策组,不利决策组认为派单任务距自己的距离更远。这说明本实验对决策有利性的操纵是有效的。

#### 2. 描述性统计分析

各变量的均值、标准差和相关系数情况如表3所示。决策者类型与公平感之间的相关性不显著;公平感则与组织偏差行为( $r=-0.13, p<0.01$ )、社会偏差行为( $r=-0.12, p<0.05$ )显著地负相关,为假设2a和假设2b提供了初步的证据。



表3 描述性统计分析和相关系数

变量	均值	标准差	1	2	3	4	5	6	7
1.性别	0.49	0.50							
2.年龄	30.08	7.55	-0.13**						
3.骑手经历	0.68	0.47	0.41***	0.19**					
4.决策者类型	0.49	0.50	-0.01	0.06	0.01				
5.决策有利性	0.48	0.50	-0.10	0.10*	-0.05	0.00			
6.公平感	4.02	1.33	-0.08	0.05	-0.16**	-0.06	0.58**		
7.组织偏差行为	3.24	1.41	-0.11*	-0.00	0.02	-0.02	-0.05	-0.13**	
8.社会偏差行为	2.72	1.55	-0.08	-0.04	0.00	-0.00	0.01	-0.12*	0.36***

注： $N=388$ ；决策者类型：1=领导，0=算法；决策有利性：1=有利决策，0=不利决策；性别：0=男性，1=女性；骑手经历：0=有骑手工作经历，1=无骑手工作经历；\*表示 $p<0.05$ ，\*\*表示 $p<0.01$ ，\*\*\*表示 $p<0.001$ 。

### 3. 假设检验

我们运用PROCESS(Hayes, 2017)对所有假设进行检验,分析结果如表4所示。首先,我们以公平感为因变量,构建了一个交互效应模型,将性别、年龄、学历水平、骑手经历作为控制变量。结果显示,决策者类型与公平感之间的关系不显著( $\beta=-0.08, ns.$ ),但决策者类型与决策有利性之间的交互项与公平感之间显著相关( $\beta=0.16, p<0.01$ )。条件效应分析的结果表明,在不利决策的条件下,决策者类型对公平感的影响是显著的( $\beta=-0.23, 95\% \text{ Boot CI}=[-0.38, -0.08]$ ),即算法决策会导致更高的公平感;而在有利决策的条件下,决策者类型和公平感之间的关系不显著( $\beta=0.09, 95\% \text{ Boot CI}=[-0.07, 0.24]$ ),假设1得到支持。具体的交互效应如图2所示。

表4 假设检验结果

变量	公平感	组织偏差行为	社会偏差行为
性别	0.07	-0.22*	-0.17
年龄	0.01	-0.01	-0.01
骑手经历	-0.21**	0.10	0.06
决策者类型	-0.08	-0.03	-0.01
决策有利性	0.76***		
决策者类型×决策有利性	0.16**		
公平感		-0.14**	-0.14*
$F$	36.95***	2.95*	1.93
$R^2$	0.37***	0.04*	0.02

注： $N=388$ ；决策者类型：1=领导，0=算法；决策有利性：1=有利决策，0=不利决策；性别：1=男性，2=女性；骑手经历：0=有骑手工作经历，1=无骑手工作经历；\*表示 $p<0.05$ ，\*\*表示 $p<0.01$ ，\*\*\*表示 $p<0.001$ 。

接下来,我们分别以组织偏差行为、社会偏差行为为因变量,构建了受调节的中介模型,对假设2a、假设3a,以及假设2b、假设3b分别进行检验。关于组织偏差行为的分析结果显示,公平感与组织偏差行为呈显著负相关关系( $\beta=-0.14, p<0.01$ ),假设2a得到支持。并且,基于5000次Bootstrap抽样的条件中介效应(conditional indirect effects)分析显示,当面临不利决策时,决策者类型通过员工公平感对组织偏差行为发挥了显著的间接影响(间接效应为0.03, 95% Boot CI=[0.01, 0.08]);当面临有利决策时,间接效应不显著(间接效应为-0.01, 95% Boot CI=[-0.04, 0.01])。这两个条件下的间接效应之间的差异显著(index of moderated mediation=-0.05, 95% Boot CI=[-0.10, -0.01])。以上结果支持了假设3a。

关于社会偏差行为的分析结果显示,公平感与社会偏差行为呈显著负相关关系( $\beta=-0.14, p<0.05$ ),假设2b得到支持。并且,基于5 000次Bootstrap抽样的条件中介效应分析显示,当面临不利决策时,决策者类型通过员工公平感对社会偏差行为发挥了显著的间接影响(间接效应为0.03,95% Boot CI=[0.01,0.08]);当面临有利决策时,间接效应不显著(间接效应为-0.01,95% Boot CI=[-0.04,0.01])。这两个条件下的间接效应之间的差异显著(index of moderated mediation=-0.04,95% Boot CI=[-0.10,-0.02])。以上结果支持了假设3b。

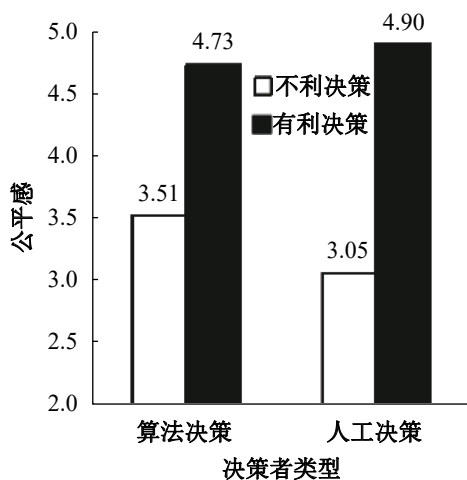


图2 决策者类型与决策有利性对公平感的交互效应

## 五、结论与讨论

本文探讨了决策者类型与决策有利性的交互作用对员工的公平感以及组织偏差行为和社会偏差行为的影响。研究发现,当决策不利时,与领导决策相比,算法决策更有助于维护员工公平感,进而减少员工的组织偏差行为和社会偏差行为;当决策有利时,决策者类型对员工公平感及其偏差行为均没有显著影响。

### (一)理论贡献

本文对于算法决策和公平感的相关文献具有三个方面的理论贡献。第一,通过揭示决策有利性的调节作用,本研究对关于算法决策与员工公平感之间关系的文献进行了重要拓展,进而为算法决策与员工公平感之间的关系提供了更加细致的理解。关于算法决策(相对于人工决策)如何影响员工公平感这一越来越受到关注的问题(Newman等,2020;de Melo等,2017;Lee,2018;Schlicker等,2021),现有的不同研究发现了不一致的结论。一些研究的发现表明,算法决策(相对于人工决策)会带来更高的公平感(如,de Melo等,2017;Schlicker等,2021)。另一些研究则发现了相反的结果(如,Lee,2018;Newman等,2020)。本研究采用归因理论的视角,揭示了决策有利性在决策者类型与员工公平感之间的调节作用。我们发现,当决策结果对员工不利时,算法决策下的员工公平感更高;当决策结果有利时,算法决策与人工决策在影响员工公平感方面则没有发现差别。以上结果表明,算法决策对员工公平感的影响取决于一些情境因素(如决策有利性),从而为两者的关系提供了更具体的理解,也协调了现有研究中发现的不一致的结果。

第二,从更加广义的文献来看,本研究的发现也拓展了关于算法如何影响人们感知的研究。具体来说,关于算法决策对员工感知的影响,文献中存在两个不一致的观点:算法欣赏(Logg等,2019)和算法厌恶(Burton等,2020;Dietvorst等,2015,2018)。算法欣赏的观点认为,与人类的判断相比,人们会更加相信算法的判断(Logg等,2019),因为算法的决策过程是无偏的(Lee,2018),决策依据信息是全面的(Sundar和Nass,2001),决策结果是准确高质量的(Sundar和Nass,2001)。相反,算法厌恶的观点则认为,尽管算法预测比人工预测准确的可能性更高,但是当面对两种预测时,人们还是会更倾向于选择人工预测,因为后者意味着人们决策过程中更高的控制力和更多的参与(Dietvorst等,2015,2018)。通过揭示算法决策(相对于人工

决策)影响员工公平感的边界条件,本研究的结果表明,算法欣赏或厌恶都不是普遍规律,而是随着情境而发生变化的。在一些情况下(如决策结果不利时),人们欣赏算法,在另一些情况下(如决策结果有利时),算法或人工决策不会带来差异化的感知。因此,本研究也在一定程度上协调了算法欣赏和算法厌恶这两种对立的观点,从而拓展了我们对于算法决策如何影响人们的感知的理解。

第三,我们揭示了决策者类型与决策有利性通过公平感对员工偏差行为的影响,拓展了关于员工如何应对算法管理的研究领域。现有的研究关于员工对算法管理的应对主要聚焦于员工认知或情感反应方面(Claggett和Karahanna,2018;Lee,2018),而对于员工的行为应对的研究并不多。少量关于这一问题的研究采用了控制或资源依赖的视角,探讨员工如何对算法控制进行合作或反抗、改变组织与自身之间的权力动态(如Karanovic等,2020)。而我们的研究采用公平的视角,探讨员工如何以偏差行为来应对算法决策可能带来的不公平感。尽管实践界已经对员工针对算法的偏差行为有诸多讨论(如赖祐萱,2020),但组织管理领域针对这一问题的研究还有待发展。在本研究中,我们发现当决策对员工不利时,算法决策(相对于人工决策)更有助于维护员工的公平感,进而减少员工的偏差行为。这些结果进一步丰富了我们对于算法管理与员工结果之间的关系的理解。

## (二)实践启示

本研究也具有两个方面的实践启示。第一,在需要做出一些员工可能难以接受的决策时,特别是在涉及需要机械化技能的任务中,比如工作分配与规划(Lee,2018),管理者可以更多考虑使用算法决策。虽然算法决策已经越来越多地被运用于企业管理中,但在许多企业决策情境中,还是由算法和领导共同决定。那么究竟在何种情况下应该运用算法决策,何种情况下应该运用领导决策呢?本研究发现,在不利决策的情况下,算法决策更有利于维护员工的整体公平感,进而抑制员工的偏差行为。因此,当决策可能对员工不利时,比如,在给员工布置难度较大的任务时,我们建议更多地引入算法决策,在一定程度上减少员工的抗拒心理和行为。

第二,为了减少员工的偏差行为,组织及其管理者应当关注员工的整体公平感。本研究发现,员工的整体公平感对其后续的偏差行为负相关。而且,整体公平感不仅能够降低组织偏差行为,也能够抑制社会偏差行为。因此,我们建议除了利用算法决策之外,管理者也应该从一些其他方面来提高员工的公平感。例如,目前有研究发现,尽管算法做出的不利决策会减少员工对其的信任,但如果管理者能对算法决策的机制进行公开透明的解释,这种负面影响会得到缓解(Kizilcec,2016)。

## (三)研究局限与未来研究方向

本研究也存在一些不足之处,值得未来研究改进。第一,本文采用了实验研究的方法对提出的理论模型进行检验,实验的情境来源于我们对骑手们的半结构化访谈,具有较高的心理现实性(psychological realism)。并且,实验方法使得我们可以对关键的因果关系进行验证,具有较高的内部效度。但实验研究在外部效度方面有所不足。并且,对于关键变量——如组织偏差行为、社会偏差行为——的测量方式,受到了情境的限制。我们基于预研究的结果,选取了适合该情境的具体行为来对这些变量进行操作化,但需要指出的是,算法决策已经被应用在许多不同的组织场景中(Kellogg等,2020),本研究以外卖配送的工作配送作为一个典型场景开展研究,但在其他场景下,员工会有其他不同形式的偏差行为。未来研究可以在其他使用算法决策的场景中,采用多种不同的研究方法,包括田野调查、实地实验等方法,进一步提高研究结论的外部效度。

第二,本文的预研究和实验研究都关注的是算法决策在工作分配(如派单)中的应用。而实际上,算法决策已经在组织的许多任务场景中得到了应用,既包括需要机械化技能(mechanical skills)的任务,如工作分配、工作规划,也包括需要与人相关技能(human skills)的任务,如选拔晋升、绩效评估等。有研究认为,在需要机械化技能的任务中,员工对于算法决策和人工决策的公平感没有显著差别;而在需要与人相关技能的任务中,认为人工决策更加公平,因为算法决策缺少直觉和主观判断能力(Lee, 2018)。在本研究关注的场景下,算法决策应用在派单中(工作分配的一种形式)。尽管在质化研究中我们发现,骑手在解释他们为什么觉得算法/人工决策更加公平/不公平时提到,决策者是否根据情况需要来调整派单会影响骑手对该决策的公平感,但总体来看派单场景对与人相关技能的要求较少。因此,我们的研究结果在需要更多与人相关技能的场景中的普适性有待进一步检验。

第三,未来研究还可以从个人、情境等多个角度探讨更多潜在的调节变量。例如,从个人角度来看,员工的专业度可能是一个重要的调节因素。有研究发现,和专家相比,业余人士会更相信算法的预测,这可能是由于业余人士缺少相应的知识,认为算法代表了“专家系统”(Logg等, 2019)。我们推断,这种对算法的准确性的认可,同样会导致业余人士感到更高的公平性。又如,与第二点相关的是,任务特征也可能发挥调节作用。任务的复杂程度、对机械化技能/与人相关的技能的要求程度等,可能与决策者类型形成二重交互效应,或者与决策者类型、决策结果有利性之间形成三重交互效应,影响员工的感知与行为。我们鼓励更多这方面的研究进一步丰富文献中关于算法决策与员工感知和行为之间关系的理解,也为算法管理这一新兴而蓬勃的管理方式提供更多实践启示。

### 主要参考文献

- [1]赖祐萱. 外卖骑手,困在系统里[EB/OL]. <https://zhuanlan.zhihu.com/p/225120404>, 2020-09-09.
- [2]Alder G S, Ambrose M L. An examination of the effect of computerized performance monitoring feedback on monitoring fairness, performance, and satisfaction[J]. *Organizational Behavior and Human Decision Processes*, 2005, 97(2): 161-177.
- [3]Aquino K, Lewis M U, Bradfield M. Justice constructs, negative affectivity, and employee deviance: A proposed model and empirical test[J]. *Journal of Organizational Behavior*, 1999, 20(7): 1073-1091.
- [4]Bradley G W. Self-serving biases in the attribution process: A reexamination of the fact or fiction question[J]. *Journal of Personality and Social Psychology*, 1978, 36(1): 56-71.
- [5]Burton J W, Stein M K, Jensen T B. A systematic review of algorithm aversion in augmented decision making[J]. *Journal of Behavioral Decision Making*, 2020, 33(2): 220-239.
- [6]Claggett J L, Karahanna E. Unpacking the structure of coordination mechanisms and the role of relational coordination in an era of digitally mediated work processes[J]. *Academy of Management Review*, 2018, 43(4): 704-722.
- [7]Colquitt J A, Conlon D E, Wesson M J, et al. Justice at the millennium: A meta-analytic review of 25 years of organizational justice research[J]. *Journal of Applied Psychology*, 2001, 86(3): 425-445.
- [8]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.
- [9]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.
- [10]Dietvorst B J, Simmons J P, Massey C. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them[J]. *Management Science*, 2018, 64(3): 1155-1170.
- [11]Gioia D A, Corley K G, Hamilton A L. Seeking qualitative rigor in inductive research: Notes on the Gioia methodology[J]. *Organizational Research Methods*, 2013, 16(1): 15-31.
- [12]Hay C, Meldrum R, Mann K. Traditional bullying, cyber bullying, and deviance: A general strain theory approach[J]. *Journal*

of *Contemporary Criminal Justice*, 2010, 26(2): 130-147.

- [13] Hayes A F. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach[M]. 2nd ed. New York: Guilford Press, 2017.
- [14] Hoobler J M, Brass D J. Abusive supervision and family undermining as displaced aggression[J]. *Journal of Applied Psychology*, 2006, 91(5): 1125-1133.
- [15] Jones D A, Martens M L. The mediating role of overall fairness and the moderating role of trust certainty in justice-criteria relationships: The formation and use of fairness heuristics in the workplace[J]. *Journal of Organizational Behavior*, 2009, 30(8): 1025-1051.
- [16] Kelley H H, Michela J L. Attribution theory and research[J]. *Annual Review of Psychology*, 1980, 31: 457-501.
- [17] Kellogg K C, Valentine M A, Christin A. Algorithms at work: The new contested terrain of control[J]. *Academy of Management Annals*, 2020, 14(1): 366-410.
- [18] Kizilcec R F. How much information?: Effects of transparency on trust in an algorithmic interface[A]. Proceedings of the 2016 CHI conference on human factors in computing systems[C]. San Jose: ACM, 2016.
- [19] Konovsky M A, Folger R. The effects of procedures, social accounts, and benefits level on victims' layoff reactions[J]. *Journal of Applied Social Psychology*, 1991, 21(8): 630-650.
- [20] Langer M, König C J, Sanchez D R P, et al. Highly automated interviews: Applicant reactions and the organizational context[J]. *Journal of Managerial Psychology*, 2020, 35(4): 301-314.
- [21] 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.
- [22] 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[C]. Seoul: ACM, 2015.
- [23] Leventhal G S. What should be done with equity theory? New approaches to the study of fairness in social relationships[A]. Gergen K J, Greenberg M S, Willis R H. Social exchange: Advances in theory and research[M]. New York: Plenum, 1980.
- [24] Lind E A. Fairness heuristic theory: Justice judgments as pivotal cognitions in organizational relations[A]. Greenberg J, Cropanzano R. Advances in organization justice[M]. Stanford: Stanford University Press, 2001.
- [25] Logg J M, Minson J A, Moore D A. Algorithm appreciation: People prefer algorithmic to human judgment[J]. *Organizational Behavior and Human Decision Processes*, 2019, 151: 90-103.
- [26] Marcus-Newhall A, Pedersen W C, Carlson M, et al. Displaced aggression is alive and well: A meta-analytic review[J]. *Journal of Personality and Social Psychology*, 2000, 78(4): 670-689.
- [27] McNall L A, Roch S G. Effects of electronic monitoring types on perceptions of procedural justice, interpersonal justice, and privacy[J]. *Journal of Applied Social Psychology*, 2007, 37(3): 658-682.
- [28] 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.
- [29] Ötting S K, Maier G W. The importance of procedural justice in human-machine interactions: Intelligent systems as new decision agents in organizations[J]. *Computers in Human Behavior*, 2018, 89: 27-39.
- [30] Schlicker N, Langer M, Ötting S K, et al. What to expect from opening up "black boxes"? Comparing perceptions of justice between human and automated agents[J]. *Computers in Human Behavior*, 2021, 122: 106837.
- [31] Shin D, Park Y J. Role of fairness, accountability, and transparency in algorithmic affordance[J]. *Computers in Human Behavior*, 2019, 98: 277-284.
- [32] Sundar S S, Nass C. Conceptualizing sources in online news[J]. *Journal of Communication*, 2001, 51(1): 52-72.
- [33] Wolosin R J, Sherman S J, Till A. Effects of cooperation and competition on responsibility attribution after success and failure[J]. *Journal of Experimental Social Psychology*, 1973, 9(3): 220-235.

# Algorithmic Decision-making, Employees' Justice Perception and Deviant Behaviors: The Moderating Role of Decision Favorableness

Wei Xin, Huang Mingpeng, Li Xinyue

(*International Business School, University of International Business and Economics, Beijing 100029, China*)

**Summary:** Algorithms have been increasingly applied in decision-making in organizations. It is of theoretical significance and managerial relevance to explore how algorithmic decision-making affects employees' psychological states and behaviors. Integrating the literatures of algorithmic management and justice, we adopt the attribution perspective to investigate how the type of decision-maker (algorithm vs. leader) and decision favorableness interact to affect employees' justice perception and their subsequent organizational and social deviant behaviors. We propose that decision favorableness moderates the effect of the type of decision-makers on employees' overall justice perception and the indirect effect of the type of decision-makers on employees' deviant behaviors. Specifically, when the decision is unfavorable for employees, employees will perceive higher levels of overall justice if the decision is made by algorithms rather than by human leaders. In contrast, when the decision is favorable, employees' overall justice perception will not differ significantly between the algorithmic decision situation and the leader decision situation. In addition, we propose that employees' overall justice perception is negatively related to their deviant behaviors in both organizational and social settings. Furthermore, decision favorableness moderates the indirect effects of the type of decision-makers on employees' organizational and social deviant behaviors via their overall justice perception, such that these indirect effects are stronger when the decision is unfavorable rather than favorable for employees.

We conduct a qualitative pilot study and an experimental study to test the theoretical model. The results provide substantial support for our hypotheses. This research contributes to the emerging literature on the relationship between algorithmic decision-making and employees' justice perception by identifying the moderating role of decision favorableness. In a broader scale, it also enhances the understanding on how algorithms affect employees' perception. Moreover, by revealing the indirect effect of the type of decision-maker on employees' deviant behaviors via justice perception, this research extends the scope of literature on how employees respond to algorithmic management. The findings also provide practical implications for management decision-making in the algorithm era.

**Key words:** algorithmic decision-making; decision favorableness; justice perception; deviant behavior

(责任编辑: 宋澄宇)