

论作为美好生活之人际伦理的信任偏好

龚天平

摘要: 作为人际关系的一种纽带,信任偏好意指个人相信他人、维护行为规则的亲社会情感、态度和行为。它既是积极态度,也是积极行为。作为一种积极态度,信任偏好是个体基于认知和情感对信任对象表现出的正向心理期待;作为一种积极行为,信任偏好既涉及信任偏好显示者内在心理倾向的外显行为,也关乎信任对象予以信任偏好显示者的积极反馈行为。信任偏好既能促进人际合作,也能增进人际友善,因而对实现美好生活具有重要意义。促进信任偏好显示以至信任偏好涌现以助力美好生活,应遵循以下逻辑:坚守自由平等原则,为信任偏好显示以至涌现提供精神基础;提升个体诚信,为信任偏好显示以至涌现提供内生驱动力量;构建正当性奖惩机制,为信任偏好显示以至涌现提供外在约束手段。

关键词: 人际伦理;信任偏好;美好生活

中图分类号: B82 **文献标识码:** A **文章编号:** 1003-0751(2026)01-0104-10

当代行为和实验经济学提出了一种创新性的人性理论即社会偏好理论,社会偏好理论的核心范畴就是社会偏好。社会偏好又被当代行为和实验经济学家们分解为三种典型意义的偏好:利他偏好、公平偏好和互惠偏好。这三种社会偏好其实还蕴含着信任偏好。信任偏好与这三种社会偏好是相辅相成的关系。作为人类的一种亲社会情感和行为,信任偏好对于美好生活具有极为重要的价值。后者作为一种理想的生活状态,是人类孜孜以求的价值目标。就两者的关系来看,信任偏好是美好生活得以维系的人际伦理,美好生活的实现有赖于信任偏好显示以至涌现。本文试图探讨这种作为人际伦理的信任偏好对于美好生活的重要价值。

一、信任偏好:积极态度与积极行为相统合的社会偏好

实验经济学家乔伊斯·伯格等^[1]与行为经济

学家阿明·福尔克及其同事迈克尔·科斯菲尔德^[2]¹²⁷⁻¹²⁹所设计的信任博弈实验都表明,信任也是一种社会偏好。所谓偏好,一般是指人所固有的、稳定的行为选择习惯;所谓社会偏好,是指个人对他人的福利的关心和维护伦理规范的愿望。信任作为一种社会偏好,就表现为信任偏好,意指个人相信他人、维护行为规则的亲社会情感、态度和行为,其对他人的行为有极强的正向激励作用,是一种重要的人际伦理。

值得指出的是,人是一个集自利性与利他性于一身的统一体,社会偏好是人基于利他本性而显示的偏好,因而信任偏好也是人出于利他而显示的偏好。当然,人出于自利也可以信任他人,但这种信任属于自利偏好。自利偏好是理性的,社会偏好则是理智与情感的有机融合,是一种打上了理性印痕的情感理智。因而社会偏好实质上是将自利偏好包含在内的,是自利偏好和利他偏好的统一体。信任偏好是由利他、公平和互惠三种社会偏好衍生出来的

收稿日期:2025-10-06

基金项目:国家社会科学基金项目“社会偏好理论与美好生活实现的伦理学研究”(21BZX023)。

作者简介:龚天平,男,中南财经政法大学哲学院教授、博士生导师(湖北武汉 430073)。

社会偏好。所谓利他偏好,是指人们既关心自身利益也关心社会总福利,特别是社会最弱势群体之利益的偏好;公平偏好意指在利益上人们都具有的既不愿领先于他人也不愿落后于他人的偏好;互惠偏好意指人对我善、我对人善,人对我不善、我对人不善的偏好。因此,信任偏好是以利他、公平、互惠为基础的。只有以利他、公平、互惠为前提而显示出来的信任偏好才可称为社会偏好。

(一) 信任偏好是个体对信任对象的积极态度

讨论信任偏好不能不联系信任,两者当然不能直接等同。信任偏好是社会偏好理论所揭示的一种人性,是个体行为;信任则是由信任偏好所形成的结果。但是,两者也有着密切关系,因为信任偏好毕竟是因信任而生发并由信任所限定的一种社会偏好和亲社会情感、行为。因此,界定信任偏好只有联系信任的含义才能获得恰当理解。

关于信任,在有的学者看来,它是个体“相信某人的行为或周围的秩序符合自己的愿望”^{[3]14}的体现。由此可见,信任首先体现为一种态度,但又并不只是一般的态度,而是一种积极的态度,因为其建立在符合自己的愿望的基础上。在心理学中,态度是指“个体基于过去经验对其周围的人、事、物持有的比较持久而一致的心理准备状态或人格倾向”^[4],它以三种内在成分,即由“对态度对象的知识或信息构成”^[4]的认知成分、反映“个体对态度对象的评价”^[4]的情感成分、体现“个体对态度对象的行为倾向”^[4]的行为意向成分为基础。信任作为一种态度,同样内在包含这三种成分。

第一,信任是一个认知过程,是个体对信任对象的理性选择。一般而言,认知过程就是人脑对信息的加工即一个由感性认识到理性认识的过程。著名政治学家罗素·哈丁曾将信任与认知关联,认为“信任与知识和信念同属认知范畴”^[5],A之所以信任B,无非是A了解到B的某种情况,让A认为B值得信任。这一过程就是个体的认知过程,即在了解某种情况(信息)的基础上作出理性判断,从而选择信任。社会学家詹姆斯·S.科尔曼以更为清晰的数学表达式解释了信任的认知过程。他将影响委托者是否给予信任的因素归结为三个方面:获得成功的概率(P)、可能的损失(L)及可能的收获(G)。他认为当个体估计 $PG > L(1-P)$ 时,才会理性地选择信任,其中“信息影响人们对成功机会的估计”^[6]。也就是说,个体是否选择信任,一定程度上取决于对他人信息的掌握程度,个体通过加工处理已知信息,

进而详细评估成功的概率、损失及收获,并作出理性的判断。而作为理性判断的重要依据,信息的类型错综复杂,它可以是一个人的言辞、学历、身份等在信息,如有学者在做以“人格特质论”为基础的信任测量研究时,将信任定义为“个体对他人的言词、承诺及口头或书面的声明之可靠性的一般性期望”^{[7]27};也可以是一个人的内在价值观,经济学上常常用偏好来考量人的价值观,如有学者将能力、利他主义、同理心和责任心、公平意识、互惠意识、尊重等个人偏好视为可信任之人的特点^{[8]29}。他人的客观信息会促成个体的理性思考和评价,这是由人的特殊性决定的。作为一种政治动物,人理应具有理性思考的能力。就此而言,可将信任视为个体依据客体信息而进行理性评价的结果。上述从理性角度出发来解释信任的方式,在经济学上最为突出,经济学家普遍认为“信任少感情,多计算,信任的双方——托主与受托人都是理性的”^{[3]58},且将信任看作可降低交易成本的“理性计算”。

第二,信任是一种情感态度,是个体对信任对象的非理性选择。从认知意义上看,信任是一种理性选择,但也并不总是如此。它还表现为一种情感,“因为没有足够的客观依据,信任属于主观的倾向和愿望”^{[3]14},这是由非理性的主观“情感态度”决定的。而这种情感态度在熟人社会和陌生人社会又有不同的表现、倾向,因此信任的方式也存在差异。熟人社会中的情感态度多表现为“天然地相信”,此时的信任非理性所致,而是本性使然。心理学家约翰·鲍比于1958年提出具有进化意义的“依恋理论”,依恋系指个体在成长过程中与他人之间形成的“心理关联”,如婴儿与照顾者之间、情侣之间都会产生依恋感,而“这种与某个特殊的个体形成强烈的情感纽带的倾向是人的本性之一”^{[7]14}。约翰·鲍比认为正是由人的本性所产生的依恋,才延伸出信任。鲍比所说的依恋主要指“婴儿对照顾者的依恋”,这种童年时期对照顾者的依恋构成信任形成的生物基础。这样看来,信任是个体的一种自然而然的情感、态度,而不是出于理性的计算。试想一个人连信任自己的父母都要经理性的算计,这似乎有些可怕。当然,无须理由的信任并不意味着个体完全没有任何想法,相反,这实质上是个体赋予信任对象的积极期待,是在特殊关系下,个体天然地认为他们信任的对象不仅不会伤害自己,甚至会为自己带来益处。这也引申出另一个问题,即在血缘、婚姻等亲密关系的陌生人社会,信任将基于何

种情感而生?一般而言,人们对自己的内群体有天然的偏爱,因此更容易产生信任。但在外群体(陌生人社会)彼此互不相识无法获取信息时为何还会存在信任?例如,你会在不做过多思考的情况下购买路边摊的小吃;会在路边随手拦下一辆出租车。显然,对路边摊主及出租车司机的信任既不是基于理性的信任,也不是基于特殊关系的信任,而是个体主观呈现的积极情感,表现为个体认为他人对自己存有善意。这一观点在心理学领域得到相对明显的体现。例如,有学者将信任界定为“一种期望或信念,使个人能够依赖他人的言行,并认为对方对自己有善意”^{[7]6};有学者认为信任是“个体特有的对他人的诚意、善良及可信性的普遍可靠性的信念”^{[7]5}。正是个体的积极情感,才使得在没有直接依据条件下信任陌生人的情况得以发生,即“我认为你对我存有善意,所以我选择信任你”。

第三,信任是一种行为意向,是个体对信任对象的正向期待。心理学上将态度中的行为意向解释为一种心理意向,而非外显行为本身。作为一种态度,信任同样表现为非外显行为的心理意向。从理性角度看,信任是个体对信息的认知,是理性思考后认为他人的行为符合自己的利益;从非理性角度看,信任是个体的本性、倾向及积极情感的结果,认为他人对自身存有善意。二者从本质上都表现出信任的心理意向特征化,即个体认为信任对象的行为会符合自己意向的正向期待。将信任看作一种正向期待,在不同学科领域都有体现。例如,社会学家伯纳德·巴伯将“行动者彼此寄予的期望”作为研究信任意义的起点,并提出有三种期望概念内在包含着信任的基本含义。第一种是“对维持和实现自然秩序和合乎道德的社会秩序的期望”^[9]。这属于一般性的期望,因而建构的是一般性的信任,如“我相信天不会塌下来”或者“我相信我的朋友是善良的”。第二种是“对同我们一道处于社会关系和社会体制之中的那些人的有技术能力的角色行为的期望”^[9]。其建构的是对他人能力的信任,如“我相信修理师能帮我修好电脑”或者“我相信医生的技术”。第三种是“期望相互作用的另一方履行其信用义务和责任,即在一定情况下把他人利益摆在自己利益之上的义务”^[9]。其构建的是对他人善良、道德的信任,如国际关系中的托管制度。这三种期望正是巴伯认为的信任的三层含义,同时体现了信任的正向期待特征。心理学更为明显地表现出对信任的“正向期待”特征的解读。例如,有学者认为

“信任是一种状态,涉及对他人动机所持的有信心
的正向期望”^{[7]7};也有学者认为“信任是一种个人
在有风险的情境中对他人动机所持有的信赖性的
正向期望状态”^{[7]8}。在经济学领域,信任的定义同样
包含着正向期待,行为和实验经济学家恩斯特·费
尔认为“信任伴随着一种期望,即该行为将以达到
委托人目的的形式得到某种回报”^{[10]13}。

综上所述,社会学偏向从社会秩序、社会道德等
宏观视角定义信任,心理学更倾向于从个人特质的
微观角度研究信任问题,经济学以理性经济人假设
为前提,更习惯从理性角度理解信任。尽管视角各
异,但我们从中可以发现,不同学科对信任内涵的界
定都有一个共同点,即信任内在包含着对信任对象
的正向期待。正如有学者所言:“海量的讨论信任
的文献有着或明示或隐晦的不同,但它们都大致同
意信任是一种信念,相信对方即使在能够伤害你的
情况下也不会伤害你。”^[11]有鉴于此,由信任所定
义信任偏好,作为一种态度,同样内在包含着认
知、情感、意向三种成分并以积极样态得以展现,因
而我们可以将其视为一种个体对信任对象报以积极
态度的社会偏好。

(二)信任偏好是个体对信任对象的积极行为

在信任问题上,多数研究者倾向于从个体内在
心理角度将信任理解为一种态度、信念,但事实上,
态度也需要一些行为来表现,如果信任仅仅是一种
心理状态,似乎就无所谓信任崩塌或由此产生的利
益损失。“信任”一词由“信”和“任”组成。《常用
汉字源流字典》将“信”解释为“言语真实,诚实而有
信用。引申为信任、相信、信奉(宗教)、凭据、消息、
书信或文书、听凭、随意”^{[12]464}。据此可知,“信”
既是一种道德品质,也是一个知识和信息。个体通
过信息选择是否相信,因此“信”的基本概念恰好构
成信任建立的客观依据,一旦信任建立,个体将会对
被信任者产生正向期待,这正是作为一种态度的信
任所具有的基本特征。而“任”在字典中的初始含
义为“抱”,后引申为“负担、怀孕(后作‘妊’)、担
子、职责、使用、任用、听凭”^{[12]349}等词义。作为一
个动词,“抱”在本质上是一种行为,而“任”的引申
意义如职责、使用、听凭等词同样也表现出行为特
征。当信任者对被信任者产生正向期待时会使用相
应的行为方式发出信任信号,表达自己的信任,而后
听凭被信任者的行为。

其中,信任作为一种行为,主要包含信任者对被
信任者作出的“我相信你”的外在行为和被信任者

向信任者显示的自己值得相信的行为。在经典的信任博弈实验中,委托人将一定金额交予代理人,是委托人认为代理人可信而产生的信任行为;而代理人将一定金额返还于委托人,是代理人基于委托人的信任行为作出的可信任行为,表示自己值得信任。“金钱交易”是一种非常明显的外显行为,但从更细微的视角看,人的微表情同样可以传递信任信号。信任博弈实验证明,人们更容易对面带微笑的人产生信任,因为微笑行为是传递可信任行为的重要方式。但“正如微笑是一种表示微笑者值得信任的方式一样,它也可以用来发出信任邀请”^{[8]51}。婴儿从出生开始,一般要经过长达一年的时间才能走路,但却可以在短时间内作出微笑行为。婴儿对父母微笑,是婴儿对父母发出的信任邀请;而父母回以微笑,则表示其值得信任。就此而言,可以说微笑行为就是婴儿与他人建立信任的入口,进而构成与他人之间的亲密关系。提及微笑与信任,其实质是对信任与行为关系的论证。虽然内在的心理倾向是信任的第一表现,但心理倾向总是会以外在行为展现出来,如一句话、一个点头、一个手势等。换言之,如果没有外在行为的表现,信任就无法体现出来。或许科技可以实现,如通过核磁共振成像来检测信任前后的变化,但对于社会的普遍信任而言,时刻依靠科技来判断信任,显然有些令人匪夷所思。

实质上,单从信任的内涵看,信任的行为特征似乎仅仅涉及信任者一方的信任行为,因为不管被信任者是否采取可信任行为,信任已经产生。但就整个信任关系的构建而言,信任并不是一个单向过程,而是“信”与“任”的双向循环。信任以信任者的“信”为始,信任者基于认知和情感对被信任者产生信赖,并敢于向被信任者予以“任”的托付,此时,就信任者而言,已完成从“信”到“任”的转换。但信任过程并未真正完成。既然信任者向被信任者已发出信任邀请,那么被信任者需要对信任者进行反馈,这一过程也就是从“任”到“信”的逆向回归。当被信任者的行为符合信任者的期待和预期时,信任才算真正构建。就此而言,信任是双方共同交互实践的结果。当然也应注意,由“任”到“信”的回归过程充满着未知与冒险,一旦被信任者的行为违背信任者的期待,那么信任就会遭遇破坏,整个信任关系将无法构建,因此也有很多研究者将信任定义为一种冒险行为。

基于上述分析我们可以认为,作为人际关系的一种纽带,信任除了表现为态度外,也表现为个体对

自己内心“相信他人的行为符合自己期待”的积极行为。作为一种积极行为,信任既涉及信任者内在心理倾向的外显行为,也关乎信任对象予以信任者的积极反馈行为。同积极态度一样,积极行为也构成信任的基本要素。基于此,由信任所定义信任偏好,作为一种习惯性的行为选择,我们就可以将其视为一种个体对信任对象发出积极行为的社会偏好。

二、促进合作与增进友善： 信任偏好对美好生活的意义

美好生活是一种安定、幸福且和谐的生活状态。在人类历史长河中,人际合作、社会和谐的美好生活一直是人们亘古不变的追求。而要实现美好生活,作为一种人际伦理的信任偏好,具有举足轻重的作用。因为一旦“离开了人们之间的一般信任,社会自身将变成一盘散沙”^[13],美好生活也就成为无源之水、无本之木。

(一) 信任偏好促进人际合作

人际合作是美好生活的重要规定。人际合作离不开信任,信任偏好构成人际合作的内生动力。但信任偏好之于人际合作是借助信任本身具有促进人际合作的功能而发挥影响的。

在人际合作形成前,有一个特殊环节构成合作的必要前提即分工。分工就是社会成员根据自身内在属性及特点从事相对独立的社会工作,主要有自然分工和社会分工。最早的自然分工当属以两性为基础的男女分工,氏族部落的“男狩猎女备饭”、农业社会的“男耕女织”是其典型。这种分工既是社会分工的前身,也是人类持续进化的基础。恩格斯的《家庭、私有制和国家的起源》在论述自然分工的基础上,进一步阐释了野蛮时代向文明时代转化的因果机制,即社会的三次大分工。他认为,分工导致了阶级压迫,但不可否认,作为一种生产过程的催化剂,分工不仅提高了生产效率,也加速了社会的新旧交替,是社会发展的进步因子^[14]。亚当·斯密认为,“劳动生产力上最大的增进,以及运用劳动时所表现的更大的熟练、技巧和判断力,似乎都是分工的结果”^[15]。分工已然成为社会生产的重要机制。然而,以分化的方式提高生产率,仅仅是分工作用的冰山一角。社会学家埃米尔·涂尔干的分工理论强调,分工的最大作用并不是生产率的提高,而是加强了联合。他认为,“有了分工,个人才会摆脱孤立的

状态,而形成相互间的联系;有了分工,人们才会同舟共济,而不一意孤行”^[16]。随着社会生产力的不断提高,当代社会的分工更为精细化、专业化,充分表现出涂尔干意义上的“有机团结”。这是一种分工高度发达,个体相对独立,但却必须相互依赖才得以生存的生活模式。这也意味着,于当代社会而言,分工是基础,合作才能实现互利共赢。

从生物进化的角度看,人类独特的演变方式决定了人类生来就是合作的物种。在“适者生存”的“生存游戏”中,人类只有相互合作,“游戏”才能持续。尤其是在当今时代,陌生人之间的大规模合作机会得以创造出来,若要持续发展,合作已成为必然趋势。而要达成合作,就须以个体显示信任偏好并形成信任偏好涌现为前提,否则,互利合作就没有可能。那么信任偏好何以促进人际合作?这要归因于信任的黏合及简化作用。

作为一种黏合剂,信任使合作成为可能。依照信任的定义,其最典型的特征就是个体对他人的一种心理认可、行为支持。一旦信任产生,将会对对方产生强烈的依赖感。但黏合效果的达成并非一方显示信任偏好所能及,而是双方互相显示信任偏好即互信的结果,这构成了合作形成的基础。互信产生合作,合作创造共赢,这就是博弈论中“猎鹿模型”(Stag Hunt Model)所反映的内在逻辑。“猎鹿模型”源自法国思想家卢梭的《论人类不平等的起源和基础》中的一则故事,该模型以猎人A和B、猎物兔和鹿为主角,以竞争和合作两种博弈形式展开(A和B都不知道对方选择),并产生以下博弈结果:若A和B彼此不信任,那么二者为竞争捕猎模式,每人每天各猎得4只兔,可维持两人4天的生活;若A和B互相信任,那么二者为合作捕猎模式,二人共同捕获一头鹿,可维持二人10天的生活;若A因信任选择合作捕鹿,而B因不信任选择单独捕兔,则A一无所获,B获得4只兔,反之亦然。就猎鹿模型的结果分析而言,只有互信产生的合作才能达到“帕累托最优”(资源分配的一种理想状态),互不信任或仅有一方显示信任偏好都无法达成合作,无法实现利益最大化。以猎鹿博弈视角看人际交往可以发现,不同的人、不同的群体都面临着信与不信的问题:信则可以互利合作,不信则意味着相互冲突。当前人际交往发展的趋势也显示交往合作、互利互惠仍是主流模式,分工越精,依赖性越强,越要求互信。

互信使合作顺理成章,但面对复杂的人际交往,如何选择正确的合作关系以促进人际交往的正向发

展,是一个不可回避的问题。可以说,人类自古以来就在与复杂的社会环境作斗争,在不确定中找确定,在迷茫中寻明朗,这几乎是人类进化的必修课。为了在复杂的社会系统中生存,人类以语言表达所思所感,以文字记录万事万物,以金钱代替烦琐的交换媒介等,简化复杂成为人类的一种生存策略。与语言、文字、金钱类似,显示信任偏好也是一种生存策略,其所带来的信任既是合作的前提,也是复杂合作关系的简化机制。

将信任看作一种简化机制,最早出自社会学家尼克拉斯·卢曼。在他看来,信任可以超越现有的有用信息,通过排除一些行动的可能性达到降低社会复杂性的目的。具体而言,降低社会复杂性是以时间维度展开的,时间具有连续性,涉及过去、现在和将来。过去确定不移,现在迟疑不定,将来扑朔迷离,信任正是利用确定的过去来保证安稳的现在,以减少未来的迷离^{[3]104}。通俗地说,信任可以通过过去推断未来,就像一个人过去值得信任,那么可以认为将来他也值得信任。虽然这种推断并没有十足的把握,但也确实能够提高效率,减少风险,相比一个完全未知的人,一个过去确定的人似乎更为安全。信任正是以上述形式,在复杂的合作关系中发挥简化功能,即通过多方信息,排除具有欺骗性的投机行为,选择合作愉快的伙伴继续合作,从而提高合作效率,降低未来损失的风险,这也是人们更愿意显示信任偏好以建立长期战略合作伙伴关系的原因。

(二) 信任偏好增进人际友善

社会和谐是美好生活的主要特征之一,但其诉诸社会关系的良性发展。一个社会的信任偏好显示越活跃,甚至出现信任偏好涌现,人际关系就越友善、和谐。这意味着信任偏好具有增进人际友善的价值。之所以如此,是由于信任与关心、利他、互惠、同理心等因素相关联。

神经哲学的开拓者帕特里夏·S.丘奇兰德曾追问:脑何以关心?关心自我福祉是动物的生存本能,但进化的压力使纯粹自我的关心并不能长久生存,于是动物必须将自我关心扩展至他人,如对后代的哺育。丘奇兰德从神经科学角度解释脑何以关心,并提出哺乳动物体内的一种强大的肽——催产素,是导致“自我关心扩展至婴幼儿,继而扩展至更广泛的关系圈”^{[17]50}的核心。催产素能够对人的行为产生显著影响,以致丘奇兰德将其视为关心得以扩展、良心得以塑造、道德直觉得以产生的重要生物基础,故催产素也有“道德分子”之称。

催产素与信任偏好显示紧密关联。丘奇兰德认为,两者是正相关关系,因为催产素“在提高对他者的容忍度时发挥了作用,以及对恐惧和回避反应进行向下调节”^{[17]50}。例如,当动物与家人或朋友处于一个较为安全的生存环境中体内的催产素水平较高时,它们通常十分放松,彼此相互关心并产生强烈的信任感,最突出的表现就是出现互相梳理和抚摸等行为。也就是说,高水平的催产素之下是高强度的关心和信任、低程度的冷漠和恐惧。另外,丘奇兰德也提出,生物存在行为循环,高水平的催产素能够增强信任感,导致梳理和抚摸动作;反之,梳理和抚摸动作也能够达到进一步的放松,维持高水平的催产素。也有学者曾将催产素作为变量因素,添加到信任博弈实验中,其实验证明,催产素与信任偏好显示度呈正相关关系,前者含量越高,后者越明显^{[17]56-57}。神经经济学家保罗·扎克也说:“在稳定且安全的环境下,催产素会让我们大多数人都表现出善良的一面。催产素引发了促使道德行为发生的移情作用,进而激发人们的信任感,继而促进更多催产素的释放,并产生更多的移情作用。”^[18]移情是一种能感知他人痛苦与快乐的共情能力,能够促使我们关心和帮助他人,在信任关系中表现为被信任方因对信任方在受到欺骗后的痛苦情绪感同身受,因而选择显示信任偏好,让信任方获得预期满足,减轻痛苦。因此,由催产素产生的同理心能够促使双边互相显示信任偏好,形成一种稳定的信任关系;这种关系又能反作用于催产素,使人们更具同理心。总之,在神经科学看来,信任偏好既是“果”,也是“因”。作为“果”,信任偏好来自催产素;作为“因”,信任偏好能够激发催产素,并导致更强烈的同理心,同时促进关心、互惠等亲社会行为的发生。

行为和实验经济学似乎也印证了信任偏好与同理心、关心及互惠有关的结论。有学者做了信任度、可信任度与利他偏好之关系的行为实验研究,通过计量回归分析,发现二者均受到利他偏好的显著影响,其结论为“更为利他的被试往往表现出更高的信任度和可信任度,而愿意信任他人的被试往往也更加值得信赖”^{[10]101}。综合以上成果可以发现,信任偏好与关心、利他、互惠、同理心等存在正相关关系,一个社会的信任度越高,意味着社会成员的“关心”越广泛,利他、互惠行为越明显,共情能力越显著。或者说,一个社会的信任偏好越涌现,人际关系就越友善。友善能够提升人际交往的暖度,化解生活的紧张,减少欺骗与暴戾,为美好生活的实现创造

和谐的社会环境。

三、自由平等、诚信提升、正当奖惩： 信任偏好显示以至涌现的基本条件

信任偏好显示不同于信任,前者是一种人性和个体行为,后者则是前者导致的结果。这种结果是显示信任偏好的双方之间建构的一种伦理结构、关系,存在于“信”与“任”的动态反馈机制中。信任建构的动态过程决定了信任本身具有时间上的滞后性、信息上的非对称性,这导致信任偏好显示不可避免地存在脆弱性。而这种脆弱性会随着陌生人社会的不断扩展,日益表现出人与人之间的道德模糊性,对人们追求美好生活构成妨碍。为了防止道德式微,顺利实现美好生活愿景,要从以下三个方面来推进信任偏好显示以至信任偏好涌现。

(一) 坚守自由平等原则为信任偏好显示以至涌现提供精神基础

要建构某种关系,首先必须确保该关系能够存在,关系存在又需要精神纽带发挥联结作用。因此,建构信任就需要夯实信任偏好显示以至涌现的精神基础。

从根本上说,信任偏好显示所建构的信任是以人为主体形成的伦理关系。从伦理学角度看,人之所以为人,其本质在于精神性,而“人的精神性的第一种体现是人的自由”“人因自由之能力而享有独特的尊严”^[19]。换言之,自由是人存在的先决条件。而自由之于人类的作用更在于“自由构成了人的全面发展的基础、道德产生的前提、社会建构的原则”^[20],自由使人类社会得以持续稳定地发展。信任作为社会发展的重要因素,同样以自由为前提,一旦“社会中只剩下指令、服从与执行,没有了交易与交换的自由空间,也就没有为兑现诺言与否留下经受考验的机会。当社会没有在其日常生活中为道德实践留下空间时,信任几成无本之木”^{[3]125}。所以,自由为信任创造可能,也为信任偏好显示创造可能,这种可能植根于与自由相伴随的不确定性。作为一种伦理价值,自由意味着行为主体自我做主、自主行动,而不受外在环境支配。一个社会是自由的社会,意味着这个社会既是开放性的,也是充满不确定性的。正是这种不确定性,使得信任偏好显示具有了价值。因为如果人们可以把握周遭环境、确信未来时,信任就没有必要存在,信任偏好显示也没有必要存在。

熟人社会是有确定性的社会,但随着自由市场的不断拓展,陌生人社会已然呈现。陌生人社会使人拥有了更多的自由,而自由又使人流动性增强。但流动性也使人更加碎片化,导致陌生人之间的交往更多地呈现为偶然性或一次性互动特征;同时,陌生人社会中人更具抽象性和隐匿性,以上种种预示着陌生人社会的不确定性显著化了。就当前社会发展状况而言,只有在不确定性中谈信任,信任偏好显示才能发挥其应有的价值,因为在“严格意义上,关于信任的合理伦理表现只发生在非信任即疑虑可能发生的伦理情境中”^[21]。也就是说,在一个完全确定的环境中,信任根本无从谈起。可能有人会疑惑,熟人社会中的信任不正是源于确定性吗?的确如此,但这仅仅是因为在熟人社会中,由于自由受限,相较于不确定性来说,确定性更为显著而已。其实熟人社会中的信任,究其根本,还是因为不确定性,因为在熟人社会依旧存在不确定性如“杀熟”(一种欺骗熟人的说法)。正因为这种不确定性,才彰显出信任偏好显示的价值。也就是用反义来体现词的本义,即用“欺骗”体现“信任”。当然,尽管自由之下的不确定给了信任生长的机会,但确定性是信任成长的基础,所以在不确定性中找确定,也是信任生长的关键。自由之于信任偏好显示的前提作用,与不确定性、确定性之间可谓“剪不断理还乱”的复杂关系^[22]。

综上所述,源自自由的不确定性,并不是一个社会消极的存在,而是人类进步的一种产物。以自由为前提,在不确定性中找确定,才能保障信任偏好显示的空间;反之,完全限制自由,消除一切不确定性,不仅信任关系的构建毫无可能,同时也将是社会的悲哀,更是历史的颠覆性倒退。与自由并驾齐驱的还有平等,二者共同构成信任偏好显示的沃土。应该说,平等是自由的产物,在一个毫无自由的地方追求平等是不可能的,这就犹如在奴隶社会中要求奴隶被平等地对待一样,实属天方夜谭。平等之于自由是必要的,但完全的自由可能会造成掠夺,甚至会引起暴动,也无益于信任的增长。自由面前人人平等,平等地享有“信”与“被信”的自由,不必因性别、种族冠以“不可信”的标签,也不必因身份、地位遵循“必须信”的命令。平等既是信任偏好显示的前提,同样“所有人都应该受到平等对待,也是我们不断扩大的信任圈的最终目标”^{[8]281},二者是一个双向奔赴的关系。

强调以自由和平等为原则,其实质是要尊重人

的主体性地位。信任偏好的主体是人,美好生活的追求者也是人。人是一切社会关系的总和,也是社会历史发展的主体。因此,信任偏好的显示必须以尊重人的自由平等权为前提,保障人的基本权益,以此为信任偏好显示以至涌现创造基本条件,进而助力美好生活的实现。

(二)提升个体诚信为信任偏好显示以至涌现提供内生驱动力量

自由和平等为信任偏好显示以至涌现创造条件,个体诚信则为信任偏好显示以至涌现提供内生动力。在马克思看来,“人对世界的关系是一种人的关系,那么你就只能用爱来交换爱,只能用信任来交换信任”^[23]。质言之,信任终归是人与人之间的交往活动,信任的产生始终是以人为根本。有学者认为,“所有信任的基础都是作为一种社会身份的个体自我的表现”^[24],那么信任偏好是基于个体的何种表现才能得以显示?对此,有学者认为,“当一个社群共同遵守一套道德价值观,并以此建立对彼此日常诚实行为的期许时,信任就产生了”^[25]。或者说,当个体日常表现出的“诚信”受到社会的普遍认同并上升为共同的道德价值观时,社会的普遍信任即信任偏好涌现便产生了。遵循上述逻辑,个体诚信与信任之间具有因果联系:个体诚信是信任偏好的原初状态,个体诚信的普遍化和常态化结果导致信任偏好的显示。因此,信任的构建,要以个体诚信的培养为核心,进而完成个人诚信到信任偏好显示再到信任偏好涌现的一体化建构。

信任偏好涉及主体“信任他人”及对象“为他人所信”问题,由此,个体诚信的培养也理应由角色的不同而有所区别。诚信以“诚”为始,释为“诚实”,寓意是信任偏好因主体之诚实而显示;以“信”为终,释为“守信”,寓意是信任偏好因对象之守信而完整。具体而言,作为信任偏好的主体,需以诚实之心对待信任对象的客观情况,避免两种极端的道德偏向:一是避免“明知不可信而信之”的“盲信”,即作为理性的个体要有正确的道德立场,在考察信任对象时需对客观对象的实际情况予以真实判断,对于扰乱秩序、危害社会等道德阙如之人不可盲目显示信任偏好;二是避免“明知其可信而不信之”的“异信”,即作为信任偏好显示的发起者需真实平等地对待每一位信任对象,避免明知其可信但却因其身份或性别而不显示信任偏好导致异化信任。而作为信任的对象,则要以守信为原则。守信是对信任偏好显示者的积极交代,是能够继续取信于人的前

提条件,若以失信回馈,将是信任危机的开始。当然,守信虽然重要,但在某些特殊场景不守信也不失为一种正确的价值立场。如你的朋友犯了罪,但出于对你的信任,他将事情告知于你并要求你保守秘密,而当你面对警察的询问时,就陷入了一种道德困境:守信意味着你欺骗警察,诚实则欺骗朋友。因此不管是信任偏好显示的主体抑或信任对象,在诚实和守信的抉择上必须拥有自己正确的道德判断。换言之,虽说信任偏好显示的主体与信任对象有着角色区别,但置身于信任关系中的人又何尝不是一种两面性的存在:在一组信任关系中可能扮演信任偏好的显示者,而在另一组信任关系中又可能充当信任对象。就此而言,信任关系中的主体既独立又统一,这也决定着诚实和守信的培养既有针对性,也有一致性。但作为一个完整的诚信个体,必须同时满足以上两个条件是毋庸置疑的。

然而,个体诚信到信任偏好显示再到信任偏好涌现,是以诚信上升为共同的价值规范作为中间环节的,“共同的规范构成了信任的坚实基础”^{[8]52}。相对来说,个体拥有的诚信、营建的信任带有范围的局限性,而美好生活的实现不是小范围信任所能及的,它必须依赖于整个社会的信任系统。要使信任偏好显示演进到信任偏好涌现以营建普遍的信任,则需将个人层面的诚信作为共同的价值规范,形成凝聚社会共识的诚信文化,并以此作为信任建构的价值引领,进而引领全社会信任体系的发展。

需要指出的是,诚信与信任偏好在字面意思上也能显示出最明显的关联,但仅靠个人诚信来引领信任偏好显示向信任偏好涌现似乎有些狭隘。正如前文所说,信任偏好与关心、利他、同理心相关,是一种友善情感和行。有学者认为,“信任是一种道德价值,它反映一种乐观的世界观,并且也有助于解释,为什么人们要向自己社群内的他人伸出双手,这些人不同于他们、而且比他们运气差”^[26]。因此,所谓个人诚信提升,实质上是人的道德品质的培养,使信任偏好显示者从动机上确立诚信、友善且不欺骗的信念,这样信任偏好显示以至涌现就具有了动力条件。

(三) 构建正当性奖惩机制为信任偏好显示以至涌现提供外在约束手段

如果说个体诚信的提升创造的是一种由内而外的信任偏好显示以至涌现的条件,那么依赖外在约束机制所建构的信任偏好显示以至涌现的环境所导致的信任,只能算作一种可操纵的信任。“真正的

信任固然是在无约束中建立的,但人类社会采取的恰恰是有约束的、不完整的、不纯粹的,充其量是有条件、有限制的信任。”^{[27]143}建立外在约束机制并非限制基于自由的信任偏好显示,也并非否定个人诚信之于信任偏好显示以至涌现的内生动力,而是在“向其成员宣告,虽然社会信任度很低,但由于这样的社会已经建立起了一整套监督和处罚机制,大家依然可以放心地生活”^{[27]143}。

外在约束机制之所以必要,一定程度上是源于信任偏好显示之内生动力的局限。我们把个体诚信的提升视为信任偏好显示以至涌现的内生动力,实际上就默认有一种自律性的信任偏好显示的存在。自律是一种自我约束,指按照自己内心的道德原则行事。因而可以把自律性的信任偏好显示理解为:在人际互动中,主体以诚信为动机显示信任偏好而构建的信任关系无关外在因素的影响。自律性的信任偏好显示的确不失为一种真正的信任偏好显示,但环境的差异性、主体的接受性决定了自律并不具有全覆盖性。就是说,在无法将生活环境、教育理念、个体接受能力等影响因素完全统一的情况下,自律性的信任偏好显示无法扩展至全社会。这样,作为一种外在约束的他律性的信任偏好显示就成为必要的补充机制。事实上,从信任偏好的概念上理解,约束机制也是必须的。信任偏好概念表达的是一种“不确定但想确定”的状态,环境的复杂、事物的多样及主体的多变都将信任偏好显示置于不确定之中,对比而言,约束机制下的他律性的信任偏好显示则显得更具确定性。当然,约束机制并不是构建信任偏好显示条件的独有的存在,“只要一个家庭、一个地方、一个社会、一个国家需要正常运转,就需要有约束机制”^{[27]148}。有学者通过做信任实验后说信任是好事,但信任并不总是比控制好,“在特别敏感的领域,不监控是行不通的,在工作环境中也是如此”^{[2]136}。目前的约束机制呈现多样化态势,主要有法律约束、制度约束、习俗约束、道德约束等,每一种约束机制都有其不同于他者的独特性,或以刚性支持,或以软性控制,又或刚柔并济。在此,笔者无法穷尽所有约束机制之于信任偏好显示的作用,但笔者认为不论何种约束机制,若要支持他律性的信任偏好显示以至涌现,都要冠以惩罚和激励两种约束手段。

第一种约束手段是建立惩罚制度。纵观整个人类社会,最能体现信任的行为其实是合作,但“合作究竟何以持续”的疑问也一直存在。抛开合作本

身,有学者提出使人类合作得以继续的是“强互惠行为”。强互惠行为是指在群体中存在部分愿意牺牲自身利益而去惩罚背叛和搭便车行为,以此来提高违背社会规范的成本,达到抑制失信行为的目的。强互惠理论是否是合作得以持续这一疑问的标准答案,还未见可知,但惩罚机制能够促进合作则是一个事实。就信任偏好与合作的关系而言,二者是相辅相成的:信任偏好能够产生合作,合作也能促推信任偏好显示。所以,惩罚机制既然能促进合作,也能促进信任偏好显示。惩罚不是目的,其最终指向的是形成惩罚知觉,构建信任偏好显示自觉。也就是让社会中的个体明晰惩罚机制,并在惩罚机制的压力下减少失信行为,树立诚信的行为意向,自觉显示信任偏好。当然,惩罚机制并不是随意制定的,它是一个严肃且有其判定标准的项目。有学者认为,惩罚源于对背叛者的负面情绪,“愤怒是标准道德行为(即惩罚做坏事的人)的强大驱动力”^[17]⁶⁷。那么进一步追问,为何背叛者会激发他人的愤怒情绪?其实是因为背叛者的行为违背了他们心中共同遵守的社会规范。由此可见,惩罚制度制定的依据是社会成员共同认定且具有正当性的价值规范,就信任偏好显示而言,就是以“诚信为荣,欺骗为耻”。“只有当人们根据广泛持有的社会规范认为其正当时,惩罚才是有效的”^[28];反之,非法、非正义的惩罚机制可能会破坏信任偏好显示环境。

第二种约束手段是建构奖励制度。对于一个社会而言,有惩罚就理应具备相应的奖励,只有惩罚没有奖励的结果可能是治理的失衡、成员的偏激、社会的病态。信任偏好显示既依赖正当性惩罚机制,也依靠正当性奖励机制。在广泛意义上,约束机制还是一种以惩罚为主要手段的消极控制,但奖励同样隶属约束机制,三者是“一体两翼”的存在,惩罚是对“不可做”的限制,奖励则是对“积极做”的鼓励。奖励之于信任偏好显示的作用,主要通过“已行动的肯定”和“未行动的激励”得以发挥,进而建构起从“巩固当前信任”到“鼓励未来信任”的持续的信任偏好显示环境。

无论是惩罚机制还是奖励机制,二者在本质上都是对公平伦理价值的践行。另外,在约束机制与信任偏好显示的问题上,可能有人认为约束机制会压制信任偏好的显示,但笔者认为不确定性是当前的社会现实,因此需要正当的约束机制来减少不确定,保障人们交往的安全性,并助力美好生活的实现。

结 语

当代行为和实验经济学提出的社会偏好理论的核心范畴即社会偏好包括利他偏好、公平偏好、互惠偏好三种典型类型,从这三种社会偏好中可以延伸出信任偏好。所谓信任偏好,意指个人相信他人、维护行为规则的亲社会情感、态度和行为。信任偏好由信任所定义,既是一种个体对信任对象的积极态度,也是一种积极行为。作为一种社会偏好,信任偏好对于人的行为具有极强的正向激励作用,是一种重要的人际伦理。

信任偏好作为人类的一种亲社会态度和行为,对于美好生活具有极为重要的价值。这主要体现在两点:一是由于美好生活以人际合作为重要规定,人际合作又离不开信任,借助于信任的黏合及简化作用,信任偏好就可以促进人际合作;二是由于美好生活以社会和谐为主要特征,社会和谐又建立在人际友善的基础上,借助于信任与关心、利他、互惠、同理心等因素的密切关联,信任偏好就可以增进人际友善。

在现实生活中,信任偏好显示具有不可避免的脆弱性,这种脆弱性会对人们追求美好生活构成妨碍。为顺利实现美好生活,需要从三个方面推进信任偏好显示以至信任偏好涌现:一是坚守自由平等原则,尊重人的主体性,为信任偏好显示以至涌现提供精神基础;二是提升个体诚信,为信任偏好显示以至涌现提供内生驱动力量;三是构建正当性奖惩机制,为信任偏好显示以至涌现提供外在约束手段。从信任偏好显示到信任偏好涌现再到普遍信任的构建,进而实现美好生活,是一个长期且艰辛的过程,需要全体社会成员共同努力。

参考文献

- [1]叶航,陈叶烽,贾拥民.超越经济人:人类的亲社会行为与社会偏好[M].北京:高等教育出版社,2013:35.
- [2]阿明·福尔克.人性的博弈:为什么做个好人这么难[M].杨亚庆,译.北京:中信出版社,2024.
- [3]郑也夫.信任论[M].北京:中信出版社,2015.
- [4]心理学大辞典:下[M].林崇德,杨治良,黄希庭,主编.上海:上海教育出版社,2003:1217.
- [5]Russell Hardin. Trust and Trustworthiness[M]. New York: Russell Sage Foundation, 2002:10.
- [6]科尔曼.社会理论的基础:上[M].邓方,译.北京:社会科学文献出版社,2008:96.
- [7]翟学伟,薛天山.社会信任:理论及其应用[M].北京:中国人民大

- 学出版社,2014.
- [8] 本杰明·何.信任:经济中的人际纽带[M].高李义,译.北京:中国原子能出版社,中国科学技术出版社,2024.
- [9] 伯纳德·巴伯.信任:信任的逻辑和局限[M].牟斌,李红,范瑞平,译.福州:福建人民出版社,1989:11.
- [10] 郑昊力.信任的本质:基于行为与神经实验的研究[M].杭州:浙江大学出版社,2022.
- [11] 马克·格兰诺维特.社会与经济:信任、权力与制度[M].罗家德,王水雄,译.北京:中信出版社,2019:95.
- [12] 常用汉字源流字典[M].魏励,编著.上海:上海辞书出版社,2018.
- [13] 格奥尔格·西梅尔.货币哲学[M].于沛沛,林毅,张琪,译.南昌:江西教育出版社,2014:138.
- [14] 马克思恩格斯文集:第4卷[M].北京:人民出版社,2009:193-198.
- [15] 亚当·斯密.国民财富的性质和原因的研究:上[M].郭大力,王亚南,译.北京:商务印书馆,1972:5.
- [16] 埃米尔·涂尔干.社会分工论[M].渠东,译.北京:生活·读书·新知三联书店,2000:24.
- [17] 帕特里夏·S.丘奇兰德.信任脑:来自神经科学的道德认识[M].袁莹,安晖,译.杭州:浙江大学出版社,2018.
- [18] 保罗·扎克.道德博弈:爱和繁荣究竟从何而来?[M].黄延峰,译.北京:中信出版社,2016:66.
- [19] 甘绍平.伦理学的当代建构[M].北京:中国发展出版社,2015:5.
- [20] 甘绍平.自由伦理学[M].贵阳:贵州大学出版社,2020:1.
- [21] 李建森.是否、应否、能否:信任的道德形上追问与经验省察[J].社会科学辑刊,2020(6):23-29.
- [22] 冯庆旭.自由与信任:“陌生人社会”的伦理视点[J].哲学动态,2022(5):101-107.
- [23] 马克思恩格斯文集:第1卷[M].北京:人民出版社,2009:247.
- [24] 尼克拉斯·卢曼.信任:一个社会复杂性的简化机制[M].瞿铁鹏,李强,译.上海:上海人民出版社,2005:81.
- [25] 弗朗西斯·福山.信任:社会美德与创造经济繁荣[M].郭华,译.桂林:广西师范大学出版社,2016:145.
- [26] 埃里克·尤斯拉纳.信任的道德基础[M].张敦敏,译.北京:中国社会科学出版社,2006:19.
- [27] 翟学伟.中国人的社会信任:关系向度上的考察[M].北京:商务印书馆,2022.
- [28] 塞缪尔·鲍尔斯,赫伯特·金迪斯.合作的物种:人类的互惠性及其演化[M].张弘,译.杭州:浙江大学出版社,2015:40.

On Trust Preference as the Interpersonal Ethics for a Better Life

Gong Tianping

Abstract: As a bond in interpersonal relationships, trust preference is a prosocial emotion, attitude, and behavior characterized by personally trusting others and acting in accordance with norms. It is both a positive attitude and a positive behavior. As a positive attitude, it refers to an individual's positive psychological expectation of the trusted party based on cognition and emotion. As a positive behavior, trust preference includes not only the externalized actions reflecting the intrinsic psychological inclination of those who show such preference, but also the positive responses from the trusted party to those who show trust preference. Trust preference can both facilitate interpersonal cooperation and enhance interpersonal harmony, thus being of great significance for achieving a better life. To promote the manifestation and even formation of trust preference for a better life, the following logic should be followed: uphold the principle of freedom and equality to lay a spiritual foundation for the manifestation and formation of trust preference; strengthen individual integrity to provide endogenous motivation for it; and construct a reward and punishment mechanism to offer external regulatory means for it.

Key words: interpersonal ethics; trust preference; a better life

责任编辑:思 齐