

收稿日期:2026-01-15

初始信用额度与信任的层级传导转化机制

——“信任的飞跃”现象的可能解释

孙正, 张晓东

(南京大学哲学学院, 江苏 南京 210023)

摘要:信任是陌生人社会中必要的社会资本,目前学界仍未就信任行为何以产生达成共识,主要存在认知主义与非认知主义两种解释路径,分别以理性计算期望与情感道德倾向作为信任行为的前提,但二者均难以合理说明对陌生人初始的“信任的飞跃”。在批判性考察两种路径局限的基础上,提出“信任层级传导与转化”模型,其核心是“初始信用额度”概念,即制度共同体通过有效回应失信行为而为成员提供的信用基础与风险兜底保障。进而尝试说明信任的生成是从宏观的制度共同体有效性,经中观的群体/系统信任提供附加,最终落实为微观个体信任的动态传导过程。此模型尝试为“信任的飞跃”等难题提供一个可能的统一解释框架,并对现代信任危机的诊断和化解作出说明。

关键词:信任;初始信用额度;认知主义;非认知主义

中图分类号:B82-02;C912.6

文献标识码:A

文章编号:1003-6873(2026)05-0030-13

作者简介:孙正(2003—),男,安徽阜阳人,南京大学哲学学院硕士研究生,主要从事当代道德哲学与应用伦理学研究;张晓东(1969—),男,安徽庐江人,南京大学哲学学院教授,博士生导师,主要从事马克思主义伦理学、中西伦理比较、伦理学与当代中国研究。

DOI:10.16401/j.cnki.ysxb.1003-6873.2026.05.004

我们在城市中向一位完全陌生的路人问路,在电商平台上确认收货,或是在公共医疗系统中选择信任一位素未谋面的医生……这些稀松平常的信任行为背后,隐藏着现代社会赖以运转的深层信任机制。这些行为既非完全基于理性计算的风险评估,也不单纯是情感驱动的道德选择,而是深刻依赖于一个超越具体人际关系的社会信任网络的支撑。正如社会学家尼克拉斯·卢曼(Niklas Luhmann)所认为的那样,信任是一种化约复杂性的机制,它使行动者能够超越对未来的无知,在不确定的情境中做出决策^{[1]1-2}。然而,这种简化机制正面临现代社会前所未有的挑战。

技术的发展改变着人的存在方式,而存在方式的改变也重新塑造着原先的交往方式以及交互策略。事实上,陌生人社会的演进与信息互联使得信任者正置身于一个深刻的信任困境之中。如社会学家安东尼·吉登斯(Anthony Giddens)所洞察的,高度复杂的现代社会本质上是一个以“脱域”机制为主导的系统,其运行极大地依赖于跨越时空的协同合作,“仅仅坐在家中,我

就已经被卷进了我所依赖的一种或一系列专家系统之中”^{[2]24}。从全球供应链的顺畅运转,到数字金融的交易安全,再到对专家知识(如医生诊断、科学家共识)的无条件依赖,现代社会系统的正常运转无疑需要广泛而深厚的信任作为不可或缺的社会资本。

但另一方面,现实的信任图景却布满裂痕。这种裂痕首先表现为宏观层面上对制度或权威信任的全球性下滑:2023年爱德曼信任度调查报告揭示了“信任与不信任的两极分化”现象,2024年报告则进一步指出对权威及公共权力的不信任感增强。与之相伴的是专家系统公信力的下滑,如特定公共事件中民众对专家的不信任。而在日常生活层面,从“扶不扶”困境到网络交往中因信息真伪、虚拟身份难以辨别而引发的普遍信任焦虑,人际信任的联结变得愈发艰难。数字时代的算法黑箱、数据滥用与数字深度伪造技术,又进一步加剧了对上述各个层面的信任焦虑。更为棘手的是,当人们试图通过加强透明度与建立问责制来重构信任环境时,这些手段本身也可能发生异化:若被简化为可量化的技术操作,被问责者往往凭借对程序知识的占有优势与制度本身的操作空间,在程序合规的表象下转移实质责任,致使公众面对看似透明的信息时,反而更难判断其是否属于自己理应知悉的内容;而且,过度依赖量化指标和程序合规,最终可能使信任被异化为一种可计算的“审计文化”,从而失去信任内在的道德与关系维度^[3]。

在此背景下,一个根本性的理论问题凸显出来:我们为何会对一位完全陌生的他人选择信任?这一问题不仅关乎个体日常交往的顺畅与安全,更关乎现代社会系统的整体运行效能。围绕这一核心之问,当代信任理论研究形成了两条主要的解释路径:一条是认知主义路径,将信任主要理解为基于证据和合理理由的理性推断,聚焦于理性计算期望;另一条是非认知主义路径,视信任为一种不可还原为纯粹计算的道德情感态度,聚焦于情感道德倾向。这两种路径各执一端,分别揭示了信任行为的不同面向,却在面对陌生人之间初始信任何以产生这一难题时,均暴露出解释效力不足的问题:既难以说明在决策信息匮乏时何以出现“信任的飞跃”,也无法为系统信任与个体信任的关联提供一个融贯的框架。因此,本文尝试在批判性考察既有信任理论局限的基础上,构建一个以“初始信用额度”为核心概念的“信任层级传导与转化”模型,为理解初始信任生成这一核心难题提供一种可能的理论阐释。

一、 认知主义与非认知主义的信任解释进路

当代信任理论研究陷入了一个看似无解的困境:认知主义与非认知主义这两种主导范式各执一词,分别揭示了引起信任行为的不同因素,却都无法对信任问题提供完整的解释图景。这一困境事实上削弱了哲学对现实信任危机的回应能力。只有深入剖析二者理论困境的形成过程与解释效力不足的根源,才能为走出这一困境指明方向。

(一) 认知主义路径:理性计算及其边界

认知主义路径将信任主要理解为一种基于证据和合理理由的理性行动,其核心论点是信任行为必然伴随着对被信任者可信度以及信任行为所带来的“风险—收益”的正面信念。无疑,这一路径在解释系统信任和专家信任时展现出了强大的理论力量。其中认知主义最具代表性的论者为拉塞尔·哈丁(Russell Hardin)与凯瑟琳·霍利(Katherine Hawley)。

1. 哈丁的利益内化模型

拉塞尔·哈丁提出的利益内化模型是认知主义解释路径的代表性观点。所谓“利益内化”,即“信任者对被信任者行为的预期,取决于对被信任者某些动机的评估”^{[4]10}。他认为,“主体A信任主体B”意味着“A相信B有动机在X事项上以符合A利益的方式行动,因为B将A的利益

内化到了其自身的利益之中”^{[4]3-4}。这一模型展现了信任的行动结构:信任者首先形成关于被信任者动机的信念,即相信对方已将其利益内化,并在此基础上产生对方将以符合其利益的方式行动的行为预期。因此,该模型通过将信任置于理性选择的框架内,使其成为一个可分析、可预测的行为。例如,信任者对司法系统以及司法人员的信任,可以被解释为基于“司法人员维护职业声誉的长期利益远大于徇私舞弊的短期收益”的信念,形成其将依法公正行事的行为预期。但这一模型的问题也很明显,其无法解释为何对信任的背叛会产生道德上的愤怒。对此,辛普森(Evan Simpson)批评道:“只要合作双方的关系被视为基于简单的约定而非情感纽带,那么理解为何退出行为会令人反感,就成了一个难题。”^[5]此外,哈丁在对信任关系分类时还区分了单向信任、重复信任与厚信任,在单向信任中,哈丁承认“在只玩一次的游戏里,几乎不会出现合作行为”^{[4]15},即在单次的“信任—回报”互动中,理性计算的结果当然是选择背叛。可问题是在生活中不乏单向信任的场景,信任者是出于什么原因甘愿冒险选择信任的呢?显然,仅凭利益内化模型不足以解释这种现象。

2. 霍利的承诺模型

继哈丁的利益内化模型之后,凯瑟琳·霍利以承诺模型进一步发展了认知主义路径。她首先区分了信任与依赖的区别,指出:信任者信任他人去保管花瓶却被辜负会感到背叛,而依赖架子支撑花瓶,花瓶却倾倒时,信任者只会对架子感到失望,却没有感受到背叛^{[6]5}。导致信任与依赖如此不同的原因被霍利归为“承诺”,即“信任某人去做某事,就是相信她有承诺去做,并且相信她会兑现承诺”^{[6]10}。因此,信任是对他人会履行其身份或角色所赋予的规范性承诺(如医生应当救治病人等)的信念,被信任者的背叛行为之所以会引发道德愤怒,原因在于,当信任者信任他人时,其内隐含着规范性期望,这种期望指向的是对方应当履行承诺,“值得信赖的人有时必须通过拒绝新承诺来预先避免失望,而非事后违背信任;这就是道德上的‘拒绝权’”^{[6]15}。也正是这种带有道德分量的期望,使得信任者被他人背叛时,会因这种期望落空而产生一种道德指向性,如指责对方不守信用。这一模型成功地将信任与现代社会角色和规范体系联系起来,能够解释为何信任者对陌生专家(如医生、飞行员)产生信任,这种信任本质上是对其专业角色所蕴含的承诺体系的信赖,但依然无法解释面对一个既不属于特殊群体又与信任者的社交圈毫无关联的陌生人时,信任者为什么依然有可能选择信任;它只能解释信任者选择信任带有一定的规范性期望,因为信任者希望对方履行特定的承诺,但无法准确解释这种期望从何而来。

总体而言,认知主义的有效性在于其理论的清晰性与解释力。它将信任置于理性计算的选择框架内,为理解现代社会的抽象系统信任提供了有力的分析工具。如吉登斯所指出的,现代社会的信任大量地寄托于专家系统和抽象系统之上,这种信任确实包含着显著的认知和计算成分^{[2]73}。

(二) 非认知主义路径:凸显道德情感的信任阐释

为了弥补认知主义的信任解释进路所忽视的情感与道德维度,非认知主义解释路径提出:信任行动在本质上是一种实践态度或情感立场,承载着不可还原为纯粹理性计算期望的信念与意图。

1. 琼斯的情感性态度理论

非认知主义者卡伦·琼斯(Karen Jones)将信任定义为一种情感性态度,即一种乐观地看待被信任者的倾向性方式。这种态度的核心可以分为乐观态度和期望:乐观态度指的是在互动领域中对他人的善意(goodwill)和能力(competence)的乐观,这种乐观不是通过理性计算他人可信

度而建立起的信念,而是一种“情感负载的看待方式”(affectively loaded way of seeing);期望所指向的则是被信任者会直接因他人的信任而感知到自己“被依赖”,进而采取有利于信任者的行动,而非仅因外部压力(如恐惧或习惯)^{[7]6}。在此,琼斯也对信任给出了一个不同于霍利承诺模型的解释:信任更像是一种主动采纳的乐观态度,包含着对他人的内在善意和主动回应依赖的期望^{[7]8}。这种信任的解释模型更进一步地回答了为什么信任者即便面对反证,也可能因为这种主动的倾向而选择继续对被信任者维持信任,“当我信任你时,我会相信你与指控的滔天罪行无关,并假设那些看似越来越多的罪证可以通过某种方式与你的清白相容”^{[7]16}。对于琼斯而言,构成一个正当的信任的前提条件不在于证据本身,而是由社会环境的氛围、信任行为发生的领域(如医疗领域因其更为重要所以需要更多证据)以及个体的倾向所决定^{[7]20}。但问题是,琼斯并没有解释为何主体会对被信任者产生如此的倾向性,而如果没有理由,信任就落入了盲目。琼斯也承认在医疗领域中的信任需要更多证据,这实则是对认知主义解释的妥协。此外,琼斯将信任的正当性条件归为社会整体环境等,但她并没有说明社会整体环境或信任行为所发生的领域本身是否正当,也没有回答基于一个扭曲的社会氛围而产生的信任倾向本身是否正当。最后,琼斯基于情感性态度的信任解释模型也无法解释系统信任的产生。

2. 希罗尼米的“完全信任”概念

在琼斯之后的帕梅拉·希罗尼米(Pamela Hieronymi)提出了更具哲学深度的分析,她提出了“完全信任”(full-fledged trust)的概念,完全信任是基于信任信念(trusting belief)的信任,即相信被信任者会做某事^{[8]229}。而所谓信任信念,其本质指向的是:选择信任的理由必须与信念内容的真实性相关(即被信任者的可信度),而非信任本身的价值^{[8]231-232}。但这一概念的解释并非指向信任就是单纯地将被信任者作为客体来进行工具性的预测,而同样承认对方的能动性。而事实上信任信念的核心所表达的是信任者对他人的自由意志所采取的特定立场。当信任者信任某人时,信任者不仅预测他会做某事,同时也期望他正是因为“认识到这是我的期望”^{[8]233}而选择如此行动。因此,背叛信任之所以构成无法化约的道德侵犯,是因为它是对信任者的意志,即一个自由主体对另一个自由主体的直接呼唤的拒绝与蔑视。尽管相较于琼斯的情感倾向模型,希罗尼米以基于信任信念的完全信任更好地解释了为何背叛信任会产生道德责难,但问题同样存在:其一,主体不可能期望系统能够认知到主体希望系统能够因“这是主体的期望”而产生行动,那么,系统信任如何产生?其二,信任者如何在初次信任中产生如此的道德期望?其三,完全信任所依靠的可信程度又如何产生?

事实上,非认知主义成功地捕捉了在信任现象中过于理性的认知主义进路无法触及的情感维度。类似于列维纳斯(Emmanuel Levinas)所强调的他者面孔的伦理召唤,“他的面容,是原始的表达,是开天辟地的第一句话:‘不可杀人’”^[9],他者的存在本身就已经蕴含着这种伦理表达,而无须通过理性计算的考量得出。由此,列维纳斯所构建的对他的前认知性回应结构为非认知主义提供了强有力的哲学支撑:它将信任从理性选择的狭窄框架中解放出来,揭示信任行动中那种不可被还原为利益计算的道德强制性,信任之所以可能“飞跃”,恰恰是因为他者的在场本身就构成伦理上的请求——他的求助、他的依赖、他作为与我共在的另一个自由主体的面孔,使回应成为我无法完全回避的道德处境,因此,借助列维纳斯,能够看到非认知主义所凸显的信任的伦理质性。

二、 认知主义与非认知主义解释进路的内在困境

尽管认知主义与非认知主义分别揭示了信任行为的理性维度与情感维度,但二者在面对

“初始信任何以可能”这一根本问题时,均陷入了难以摆脱的理论困境。以下将从各自的局限出发,逐步揭示这一困境的深层根源。

(一) 认知主义信任理论的三重局限

尽管认知主义的信任解释路径在多数场景下都具有解释效力,但当信任的场景从抽象的系统交互转换为具体的人际互动时,认知主义范式便暴露出其解释的乏力。

首先,认知主义对初始信任解释无力。当信任对象属于特定系统时,认知主义尚可借助系统信任形成预期;但面对初次信任,信任者却在证据严重不足的情况下仍可能授予信任。这种非理性的信任无法被利害计算容纳。霍布斯(Thomas Hobbes)的自然状态难题在此重现:无强制权力保障履约,首先践约者无异于“自弃于敌人”^[10]。在缺乏他人过往行为证据的条件下,认知主义的理性计算无法启动,单次互动也意味着遭遇背叛需要付出成本才能让对方得到惩罚。在这种收益与风险极其不对等、近乎霍布斯的自然状态中,信任从理性的考量上似乎无从发生。这构成了认知主义的起点悖论。

其次,正因为认知主义难以说明初始信任何以产生,因此它也无法解释面对单次互动的陌生人所出现的“信任的飞跃”现象。克尔凯郭尔(Soren Kierkegaard)曾以亚伯拉罕通过“信仰的飞跃”而非理性的皈依来建立个体直接面对上帝的生存关系为例,来描述宗教信仰的形成^[11]。这种基于非理性因素而达成行动结果的情况,类似地出现在建立起合作关系、亲密关系或应对危机事件时,达成初始信任的这一行为也往往要求信任者在信息并不完备的情况下做出主动走向信任的冒险。这种包含主观的意志决断和道德勇气的飞跃,显然不可能仅仅基于理性的“收益—风险”计算结论,抑或基于证据及其所导向的某种预期性结果的理性期望而产生。

最后,尽管霍利的承诺模型对信任被背叛所引发的道德伤害进行了一定说明,但基于理性计算的认知主义还是淡化了道德伤害的特殊性。当信任被背叛时,信任者所体验到的不仅仅是预测失误的懊恼,更有一种独特的、带有道德色彩的被背叛感。认知主义模型难以解释这种情感的深度和特殊性。如霍克海默(Max Horkheimer)与阿道尔诺(Theodor Wiesengrund Adorno)所言,工具理性的扩张可能导致对道德情感的忽视,“思维把自身客体化为一种不由自主的自我推动过程,客观化成一种机器的化身……这种将一般思想加以平均化的模仿形式,把现实的事物转变为特有的事物”^[12]。这恰可类比认知主义的信任理论,它同样将信任这一复杂的心理活动简化为“风险—利益”的理性计算,忽视了信任中固有的道德期待与情感投入,从而无法解释信任被背叛所带来的独特道德伤害。

(二) 非认知主义进路的理性局限与内在矛盾

尽管非认知主义凸显了信任的道德属性与主动性特征,弥补了认知主义路径所忽视的情感与道德维度,但具体来看,非认知主义的信任解释路径还是有着以下局限:

其一是理性缺失难题。如果非认知主义坚持初始信任无须任何证据,那么信任便沦为纯粹依靠情感与道德期望的盲目信任,进而无法区分合理信任与愚蠢轻信;而一旦承认需要理性证据,它又未能真正跳出认知主义的框架。威廉姆斯(Bernard Williams)在《真理与真诚》中强调信任应当有理性基础,他提到:“除了知道我们与某个人分享一个语言外,我们在其他方面对他一无所知,倘若如此,我们就应该从信任他入手。这不是一个理论假定,而是一项糟糕的建议。”^[13]因此必须承认,面对初次信任,信任者终究需要寻找支持其选择信任的理性理由。

其二,非认知主义对系统信任的解释乏力。很难说普通人对金融结算系统或日常生活中司法程序的信任是出于对系统的“自由意志”的道德期望,因为显然系统本身并不拥有自由意志。

我信赖司法程序,并非期待司法机构善意对待我,而是相信规则在类似情境中被一贯适用。卢曼所描述的“秩序被奉为规范并确立为可信准则”^{[1]18},精准刻画了这一现象,系统信任的本质是对规则的信赖,而非对他者意志的道德期待。因此,非认知主义在解释这类现代社会至关重要且体量巨大的信任形态时,无疑表现出了明显的局限性。

(三) 理论困境的根源:化约论与二元对立

认知主义与非认知主义之间的对峙及各自的困境本质上在于双方都陷入了一种化约论的追求,即试图将复杂的、多面向的信任现象归结为单一的本质。认知主义强调信任的理性计算,却忽略了信任被背叛带来的道德伤害;对系统信任和理性选择具有较强的解释能力,却对初始信任下的“信任的飞跃”解释乏力。非认知主义点明了信任的情绪与道德层面,却陷入了面对初始信任时难以完全摆脱理性计算的困境,并且难以解释对于系统的信任何以产生。

认知主义和非认知主义各自的困境在事实上表明,假定理性与情感、事实与价值可以清晰分开,进而将信任强行纳入某一单向度解释框架,这一思路并不可取。约翰·杜威(John Dewey)提出“当事实被认知时,它便进入了一个全新的环境。在不脱离现实环境的前提下,它还融入了人类活动的领域——欲望与厌恶、习惯与本能的交织”^[14]。这种实用主义认识论提示我们:信任是一种在具体社会情境中生成的,融合理性评估与情感托付、事实判断与价值承诺于一体的整体性伦理行动。因此,不能单一地以现实的理性计算或人类情感的倾向去理解信任这一伦理行动,而必须认识到信任本身是一个复杂心理活动过程的整体,人是一种社会性的存在,意味着信任行为本身也必须在复杂的社会关系中进行理解。

三、基于“初始信用额度”的“信任层级传导与转化”模型

如前所言,信任本身作为一种“伦理行动”^[15],如果要研究信任的本质,就不能将信任置于孤立的“二人交互关系”^①中进行理解,信任者对被信任者的信任,不是仅仅参考独属于被信任者的某些信息,正如马克思所言“人是社会关系的总和”^[16],处于同一个共同体^②中,就意味着个体之间共享着某些背景信息,至少作为同一生活世界的成员这一点是毋庸置疑的。因此,如果说认知主义与非认知主义分别刻画了主体间信任行为的理性与情感的维度,那么对信任的完整的解释框架还必须补上互动主体各自的社会维度因素。以下将提出一个以“初始信用额度”为核心的“信任层级传导与转化”模型,尝试为“信任的飞跃”现象提供一个可能的统一解释。

(一) 从系统信任与人际信任的二分到动态传导的视角

尼克拉斯·卢曼在其经典著作《信任与权力》中,区分了人际信任与系统信任。人际信任立足于个体间的熟悉与情感联系,是在面对他人自由所带来的不确定性时,所形成的对他人行为

① 此处并非仅指两个自然人之间的交往,而是泛指两个主体的交互,如个人与机构之间的互动等。

② “共同体”这一概念并非指称群体环境或是前现代的“温情集体”,而是指一个共享基本规则、具备基本制度效能的社会合作框架。它通过制度(法律、监管)、文化(诚信规范)、基础设施(信息透明)等为成员提供安全感。为避免概念混淆,对“共同体”一词的使用作如下界定:在讨论制度层面时,使用“制度共同体”一词,指代一个以法律、监管和争端解决机制为核心的规则体系;在讨论认同层面时,使用“文化共同体”一词,指代一个以共享价值观、历史记忆和生活方式为纽带的群体;在存在论层面讨论时,使用“生活世界”一词,指代个体与他人共在的基本情境。本文所构建的“信任层级传导与转化”模型,其宏观层面的信任基础主要依托于制度共同体的有效性,而非文化共同体的认同一致性。这一界定的好处在于:它使模型能够适用于价值多元的现代社会,也就是说,陌生人之间即使缺乏文化认同,仍可基于对制度共同体有效性(即双方均受同一制度约束且制度能够有效回应失信行为)的共识而形成信任预期。

的期望。而系统信任则指向对诸如货币、法律、政治制度等抽象系统的信赖,这种信任的根基在于系统功能的稳定与可靠,而非对系统内任何具体操作者的了解^{[1]48-49}。如前文所述,吉登斯对脱域机制与专家系统的分析同样深刻揭示了现代系统信任的基本形态。然而,这一区分本质上仍是一种静态的类型学。它描绘了信任的两种不同形态,却未能揭示二者之间动态的、微观的传导机制。当我走进一家陌生的医院,信任一位素未谋面的医生时,这种信任显然无法被简单归入人际信任或系统信任的单一类别:它既不是基于对该医生个人品行的熟悉,也并非仅仅抽象地信赖医疗系统而悬置具体个体。因此,问题的关键在于:对系统的信任,究竟通过何种机制传递到作为系统代表的个体身上?

尽管过往关于信任如何生成的理论中,这一传递过程并未得到详细解释,但可以分析出这一过程所涉及的不同社会层级。正如彼得·什托姆普卡(Piotr Sztompka)所指出的,“当信任文化存在时,交易成本会显著降低,合作机会也会增加”^{[17]105},这一判断隐含着一个重要洞见:社会背景层面的信任因素(对系统的信任、共同体的规则秩序)会向下渗透,影响具体情境中的人际信任。换言之,对陌生个体的信任,并非凭空产生,而是依托于一个由宏观至微观层层传导并转化的信任结构。当信任者走进一家陌生的医院,并对某位医生产生信任行为时,这一心理—社会事件可以大致分为以下的层级:

其一是对多个系统整合的制度共同体(一般意义上的政府)的监管体系与利益保证(包括法律体系建设与专业监管人员)的原初信任;

其二是对中观的医疗系统(包括其教育认证体系、行业监管规范、医疗科学技术)的基本信任;

其三是对医生所属的职业群体的整体能力和伦理水平的认知与评价;

其四,将这份基于系统和群体的信任,“授予”或“投射”到眼前作为其代表的个体医生身上,从而完成从系统信任到人际信任的现实转化。

因此,卢曼和吉登斯关于系统信任与人际信任的区分理论只是指出了信任的不同层级,但对于作为大背景的多系统的整合共同体的信任以及将宏观系统与微观个体联系起来的信任传导机制,却语焉不详。信任者如何从系统信任中提取信任并传递给具体的人际互动中的陌生人?这是需要回应的问题。因此,信任不应从分类学的角度孤立地各自看待,而应当被理解为一个多层次、动态的社会过程。微观的个体信任行动与宏观的社会信任结构并非相互独立,而是通过一系列中介机制相互构成。吉登斯在《社会的构成》指出:社会结构与个体行动之间存在相互构成的关系,“社会成员就是利用社会结构来使他们在特定情境中进行的行动为人理解,显得有条有理。在这一过程中,社会结构既是情境性互动的根本源泉,也是后者的产物”^[18]。因此,“利用社会结构来使社会成员在特定情境中进行的行动为人理解”,正是“信任层级传导与转化”模型的核心。

(二) 初始信用额度:作为信任起点的理论预设

在明确了信任本身存在着层级传导与转化的特性之后,需要重新回应最初的关于初始信任的问题:为什么行动者愿意对一个不知晓任何个人信息且只有单次互动的陌生人选择信任?对此,我们首先必须承认一个存在论的前提:信任者与被信任者共在于同一个生活世界之中。这种“共在”并非空间上的邻近,而是一种“共同生活形式”的共享,“他人的存在不但永远是无选择的给定条件,而且是每个人所必需的存在/生活条件”^[19]。这揭示了一个基本事实:个体的具体身份与行动可能性,总是也必须在特定的社会结构中得以界定。

正是依据这种共在的存在论前提,在此提出一个概念:初始信用额度。所谓“初始信用额

度”,是指制度共同体为其每一位成员预设的、非基于个人认知的信用资本。它是个体在面对完全陌生的他者时仍愿选择信任的基本理由,也是所有社会成员因共处于同一规则框架而被平等赋予的信用基础。但不能把初始信用额度等同于信任,它只是为初始信任提供一个必要的理性原因,正因为每个人都拥有一定的初始信用额度,所以,对陌生人的理性信任才是可能的。

这一概念的理论资源可追溯至朱克尔(Lynne G. Zucker)所分析的制度信任机制^[20]:当个体之间缺乏过往互动史,也无从辨识对方的人格特质时,信任仍可能发生,其依据在于共同遵守的制度规则为失信行为设置了可预期的代价。什托姆普卡提出的“相较于充满冲突、分裂对立的社会,信任者更倾向于选择和平、和谐、团结的社会”^{[17]123-124},也可支持这一概念。尽管信任者的这种倾向可能同时源于文化认同与制度保障,但对于价值日益多元的现代社会而言,后者的作用尤为关键。个体对制度共同体稳定性、公正性的信心,会转化为其在具体互动中愿意预支信任的心理倾向。之所以能够对陌生个体选择“信任的飞跃”,其前提不仅在于理性考量与行动者具有的一般性情感立场,更在于对方已经具备了可供信任的初始信用额度。这一额度为个体提供了理性的风险预期,使行动者敢于进入互动;而指向具体他者的情感倾向,则需在后续关系中逐渐形成。

需要强调的是,初始信用额度并非仅仅源于冷冰冰的制度规则,它更深层地扎根于一种为成员所共享的、最低限度的认同。这种认同可能是对法律规则的共同尊重,也可能是对某种道德直觉或对社会运行机制理解的共享。正如查尔斯·泰勒(Charles Taylor)所言,现代社会的个人本真性的确认与信任离不开“共同视野”的支撑,这种视野构成了我们与他人互动时能够“预判”对方行为的基础^{[21]47}。正是这种规则与认同的交织,构成了一个健全有效的制度共同体的基础,并以此生成了初始信用额度作为每位成员共有的基本信用,这种共有性意味着,虽然每个人都可能成为他人眼中的陌生人,但彼此陌生的双方对某一信任行为是否安全,拥有一个共同的判断依据。但初始信用额度在多元价值盛行的现代社会无法再直接源于某种抽象的共有性或单一的文化认同,而更多地源于共同体制度的客观有效性。当一个共同体的制度体系(包括法律、监管、争端解决机制等)能够以预期方式对失信行为作出有效回应时,该制度共同体便为其每一位成员提供了基本的风险兜底保障。这种保障独立于个体是否实际信任制度而存在,它构成了一种客观的信用基础:即便个体对具体他人一无所知,他也能基于对这一基础的认知,形成对互动风险的基本判断。因此,初始信用额度的本质,是共同体制度有效性所赋予的、可被个体理性认知的风险控制预期,而非某种先验的、预设的信任资本。如在现实中与一个陌生人产生交互关系时,即便信任者并不知晓其特殊的个人信息,也有以下的命题是必然成立的:

①对方必然与信任者处于同一个制度共同体中;

②这个制度共同体对这个被信任者而言至少是可忍受的(若不可忍受,则被信任者没有理由留在这个共同体中);

③被信任者必然也同样受这一制度共同体内部至少是部分规则的约束;

④即便被信任者与信任者的交互并非基于某些善意的企图,被信任者也仍然需要制度共同体最基础的功能,如基本的秩序稳定来保障自身的基本安全与行动的可能,这也是上述第三条信息成立的前提。

可以发现,前两个信息说明信任者与被信任者在背景信息上具有一定的相似性,后两者则说明决定一个完全陌生的个体拥有多少初始信用额度的根本条件在于制度共同体自身。制度共同体的秩序稳定性越高,那么,无疑每个人的初始信用额度也会更高,信任者就会对制度共同体弥补信任这一风险行为可能带来的损失拥有更高的信心——这种信心是基于秩序的稳定性、有效性以及规则的公正性与高惩戒速度。

初始信用额度的本质,是共同体通过其制度有效性与规则认同,为成员之间的互动提供一种风险兜底的承诺。它降低了达成信任这一冒险的潜在成本,使“信任的飞跃”不至于沦为盲目的赌博。初始信用额度并非以规则替代信任,而是以规则支撑信任。福山(Francis Fukuyama)则认为,正式规则与法律装备只是“信任的替代品”^[22],这一判断预设了规则与信任之间的零和关系。然而,慈继伟对正义结构的分析表明:正义的无条件性恰恰需要制度有效性为其奠基,规范与教育使正义从外在约束内化为心理认同^[23]。这一论证可迁移至信任问题,原因在于信任与正义在结构上共享一个关键特征:二者都内含对他人的规范性期望,而这种期望的维系高度依赖行动者对“他人也会同样守信”的相互性预期。因此,信任中的规范性期望同样需要制度有效性作为其外部支撑。休谟(David Hume)对财产规则的分析则从经验生成的角度提供了补充:稳定的财产规则虽在人类活动与交往中缓慢生成,却正是通过惩戒失信、稳定预期而获得效力^[24]。这一逻辑同样适用于信任所依托的规则约束,即制度通过有效回应失信行为,才能使信任者形成稳定的风险预期。信任的有效性也需要规则约束的支撑,只是必须区分如何做的规则与能否做的规则,初始信用额度所依托的正是后者——它为初始信任提供风险预期,使“信任的飞跃”不致沦为盲目的行动选择。

(三) 从初始信用额度到“信任层级传导与转化”模型

有了初始信用额度作为基础,信任的达成便可被理解为一个从宏观到微观的动态过程:层级传导,从对系统的信任转化为具体的人际信任。在共同体给予每个人的初始信用额度的基础上,每个人的信用是动态分层的,这种信任层级传导与转化使得因各自拥有不同群体身份的个体在不同场景中的信任资格被不断地评估与调整。

具体来看,这种“信任层级传导与转化”模型即为:

其一,在宏观上,一个健康、有序的制度共同体为每一位成员(包括陌生人)自动生成一个非零的初始信用额度。这个额度不是基于对个体的了解,而是基于一个普遍性的推定:既然你在这个我信赖的制度共同体中生存,那么,你受到该共同体规则的约束,因此你初步具备了一定的可信性。正如卢曼所言:“随着世界变得日益复杂,作为规范来源的那种旧有的自然统一体便瓦解了,从人们普遍共享的信念预设,跳到对个人的信任,这一跃再也无法逾越那不断扩大的鸿沟。”^[13]世界无法再用单一秩序进行解释,也无法确认他人共享着相同的世界观,作为共识的普遍信念也不复存在,亦如泰勒的论述,现代社会中的部分个体不再将“共同视野”(即价值共识)作为自我价值的锚点,而是追求自我实现的、琐碎的、自我中心式的生活方式^{[21]50-51}。也因此现代社会中,初始信用额度的高低不再依赖也无法依赖传统的那种不言而喻的普遍信念,而主要取决于制度执行的公正性、透明度与效率。在当今时代,价值多元与系统分化使得人们对“共同视野”的依赖日益减弱,因此,对制度有效回应失信行为提出了更高要求。

其二,在中观上,有了共同体产生的初始信用额度,具体情境下信任会迅速聚焦到个体所属的特定群体或扮演的社会角色上。例如乘警、医务人员或公共服务人员等会因其身份与共同体规则的紧密关联而增加信用额度,从而在初始信用额度基础上获得信任的叠加。当遇到一位陌生警察时,信任者不仅依托了共同体所赋予他的初始信用额度,更调动了对警察这个职业群体的特定信任。后者是在前者基础上的信用放大和聚焦。因为警察群体往往会受到更严格的职业规范规制、纪律约束和道德筛选,所以才会赋予陌生警察高于普通陌生人的信用额度,这与吉登斯所分析的专家系统的信任生成过程具有相似性。

其三,在微观上,当共同体以及系统的信用传递完成之后,一个特殊群体内的信用额度便已形成,毫无疑问,这一信用额度对处于这一群体的每一个成员都是通用的。信任者最终作出信

任行为,是在此基础上结合具体情境中的情感与道德倾向而达成的结果,从而既容纳了认知主义理性计算的考量,也结合了非认知主义的道德期望。

其四,在从宏观到微观的传导过程中,传导叠加的是信用,但信用额度的增加事实上不能等同于信任的达成,有更高的信用只意味着我们有更大的信心去选择信任。而在这个传导过程中,信任本身也发生着转化,即抽象的系统信任在具体的个体互动中转变为对另一个自由主体的道德期望。这个过程具体有三个部分:首先,当个体以特定社会角色(如医生、警察、教师)出现在信任者面前时,其承载的不仅仅是制度赋予的信用额度,更激活了信任者对该角色的道德期待,即该个体在互动中应该展现的社会角色。其次,正因为这种角色,个体成为系统在具体情境中的具身化载体,如警察背后代表的是公安系统。系统本身无法直接出场,必须通过其成员的实践而显现。正是经由这种具身化,对系统的信任转化为对具体他人的托付。最后是这种托付的保障,制度有效性所提供的初始信用额度为信任提供了安全网,使得信任者敢于将后果预期升级为道德期望,并进而做出信任的行为。

其五,这种传导与转化不是单向度的,而是带有反馈,从而形成自下而上的信任反向传导通道。当被信任者成功兑现信任时,这一积极的个体经验不仅会强化对被信任者个人的信用评价,还会通过反向传导巩固对该群体乃至整个系统的信赖,使未来的初始信用额度更加稳固。反之,若个体在互动中频繁遭遇背叛,则不仅会降低对失信者的信用评价,还可能侵蚀其对中观群体的信任,并最终动摇其对制度共同体的信心,如卢曼提到:“最重要的是(不仅在个体层面,在系统层面更是如此),信任取决于对风险倾向的可控性以及失望次数不会过度累积。”^{[1]86}这种双向互动机制,使得信任成为一个动态演化的社会过程。

由此可见,制度共同体提供的初始信用额度与个体的伦理信念并非相互替代或对立的关系,而是互动共建的动态组合。良好的制度环境(高初始信用额度)降低了个体选择信任的风险成本,进而鼓励其在日常交往中更加积极地展现信任、践行诚信;而广泛而持续的个体信任实践,又会通过正向反馈不断强化制度公信力,形成“制度保障—信任实践—制度强化”的良性循环。反之,若制度长期无法有效回应失信行为,个体的伦理信念将因缺乏外部支撑而趋于萎缩,信任文化亦无从生成。

四、“信任层级传导与转化”模型的理论优势与现实启示

相较于认知主义与非认知主义的传统进路,基于初始信用额度的“信任层级传导与转化”模型展现出显著的理论优势,为理解和应对现代信任危机提供了深刻的现实启示。

(一) 超越化约论:为信任的复杂性提供统一的解释框架

该模型成功弥合了认知主义与非认知主义的理论鸿沟。它在一个连贯的框架内同时容纳了认知主义的理性要素(体现在对系统和群体规则的认知评估上)与非认知主义的情感要素(体现在将系统信任主动寄托于个体时的道德期望与情感托付),为超越二者的理论局限提供了一条富有希望的道路。更重要的是,它为“信任的飞跃”这一经典难题提供了一个清晰的解释路径。所谓“飞跃”,并非盲目冒险,而是个体根据初始信用额度这一由制度所提供的安全网,所做出的一种兼具理性评估与情感勇气的伦理行动。类似于克尔凯郭尔所描述的“信仰的飞跃”建立在对更高秩序的托付之上,“信任的飞跃”同样建立在对共同体秩序的基本信任之上。这一解释既摆脱了认知主义将信任完全还原为理性计算的理论困境,也避免了非认知主义面对理性缺位的无力。

(二) 揭示传导与转化机制:澄清人际信任与系统信任的内在关联

“信任层级传导与转化”模型能够清晰地揭示初始信用额度如何作为人际信任的初始前提,以及信任从宏观系统经中观群体向微观个体传导、并通过个体实践进行反馈的双向运行机制。宏观层面的制度有效性为所有成员提供了基础的初始信用额度;中观层面的群体/系统(如职业群体、组织机构)因其与制度的紧密关联,在宏观制度有效性的基础上获得额外的信用叠加。在信用的传导叠加的基础上,系统或制度共同体下的个体间互动,就能够将制度共同体赋予的、对职业群体的信用,投射到具体的个体身上。而在这个过程中,信任的性质也发生了变化,它不再仅仅是对规则系统的信任,而是转化为对另一个自由主体的道德期望。借助这个模型,信任危机的扩散路径也得以完整呈现。当一个医生的失信行为发生时,如果缺乏中观层面的群体/系统反思和宏观层面的制度回应,那么,这种个体失信就可能侵蚀人们对整个医疗系统,甚至政府监管的信任。反之,当系统能够有效回应个体的失信行为时,信任的传导与转化机制就能够保持畅通,甚至得以强化。这一分析框架为诊断现代社会中各种信任危机的发生机制提供了新的理论视角。

(三) 重构信任建设的路径:从个体德性呼唤到制度与文化共建

本模型将信任问题的理论重心从初次信任中个体何以愿意冒险转移至制度共同体何以提供值得信任冒险的安全网与认同感。陌生人社会信任危机的深层根源,并非个体德性的普遍衰落,而是共同体规则有效性的减弱、回应能力的降低及文化认同感的弱化。由此,信任重建的根本着力点,便不再是抽象地呼唤个体德性回归,而是切实提升制度执行的公正性、透明度与效率,并在此基础上培育一种珍视诚信、尊重他者的公共文化。当制度能够有效降低信任的风险成本,个体的信任行动就会获得更坚实的支撑,并在持续的互动实践中形成广泛的信任文化氛围。以医患关系问题为例,从暴力伤医事件到对医疗系统性信任危机,医患信任的裂痕已从个别事件演变为结构性问题。运用本模型分析,医疗信任危机的深层原因可概括为传导链条的断裂与转化机制的失效。

首先,传导过程中,医疗信任的宏观基础是制度共同体的有效性,包括医疗监管体系、医疗纠纷解决机制、医生执业资格认证制度等。当这一制度有效性受到质疑(如医疗纠纷处理不公、监管缺位、违规执业成本过低)时,患者对医生所具备的初始信用额度的评价就会降低,信任的宏观基础也随之动摇。

其次,在中观群体层面,医生职业群体本应作为信用叠加的中介,患者对医生这一职业群体的整体能力和伦理水平的认知,会转化为对具体医生的信任预期。但当个别医生的失信行为未能得到职业群体的有效回应(如行业自律缺失、同行监督缺位),或媒体对医疗行业的负面报道持续累积时,中观层面的信用叠加效应便大打折扣,患者不再将医生视为一个值得信赖的职业群体,而是需要警惕的陌生人。

最后,这种传导叠加的信用会转化为具体情境中产生信任的理由。而即便宏观和中观层面的信用基础已经受损,在具体的医患互动中,信任仍可能通过医生的具身化行为(耐心沟通、专业解释、人文关怀)而部分实现。然而,当医患互动被简化为技术操作而缺乏情感交流,或者医疗体制迫使医生在有限时间内完成大量诊疗任务时,这种具身化的空间被严重压缩,患者难以从医生的面孔、语言和行为中感受到道德期望的可寄托性,从而系统信任向人际信任的转化机制失效。因此,医疗信任危机并非单纯的医德滑坡或患者焦虑,而是制度有效性不足(宏观)、职业群体信用弱化(中观)与医患互动异化(微观)三者相互作用的结果。信任重建需要在这三个

层面同时发力,即完善医疗纠纷解决机制、强化行业规范、重构医患互动的伦理空间。

由此,该模型与当代中国国家治理现代化的核心议题也形成了深刻的理论呼应。当“全面推进国家各方面工作法治化”成为改革共识,当政务诚信建设从道德号召转化为可量化、可追责的制度闭环,这一转型正通过提升制度执行的公正性、透明度与效率,为社会成员持续提供可感知的初始信用额度,进而提升公民选择信任的信心。在这个意义上,信任建设既是国家治理现代化的制度工程,更是一项在制度与人心之间持续传导、双向建构的意义工程,后者体现了“治理体系和治理能力现代化”这一时代命题所蕴含的人文温度与伦理厚度。当每一位公民都能在公正、透明、高效的共同体中生活,当每一次“信任的飞跃”都能获得制度有效保障,信任这一现代社会不可或缺的社会资本才能够得以持续积累,成为推动社会进步的内在动力。

综上所述,本文所构建的“信任层级传导与转化”模型,无意纠缠于“信任是什么”的定义问题,也不将信任单纯归结为理性计算或情感倾向,而是将信任置于具体的社会互动关系中,从宏观制度背景出发整体性地把握其生成条件。模型尤其强调,初始信任的理性基础不在于信任者对被信任者个体信息的预先判断,而在于双方所处制度共同体为被信任者提供的初始信用额度,取决于制度共同体能否有效回应失信行为并为成员提供风险兜底。正是从此制度条件出发,模型进一步说明初始信用额度如何经中观群体向微观个体传导,并最终转化为具体的信任行为。通过这个模型,既为“信任的飞跃”提供了理性基础,也为理解和化解现代社会的信任危机指明了根本着力点:提升制度有效性,而非仅仅呼唤个体德性。

参考文献

- [1] Luhmann N. Trust and Power [M]. Chichester, Toronto: John Wiley & Sons, 1979.
- [2] 吉登斯. 现代性的后果[M]. 田禾,译. 南京:译林出版社,2011.
- [3] Power M. The Audit Society: Rituals of Verification[M]. Oxford: Oxford University Press, 1997:96-97.
- [4] Hardin R. Trust and Trustworthiness[M]. New York: Russell Sage Foundation, 2002.
- [5] Simpson E. Reasonable Trust[J]. European Journal of Philosophy, 2013, 21 (3):402-423.
- [6] Hawley K. Trust, Distrust and Commitment[J]. Nous, 2014, 48(1).
- [7] Jones K. Trust as an Affective Attitude[J]. Ethics, 1996, 107 (1).
- [8] Hieronymi P. The Reasons of Trust[J]. Australasian Journal of Philosophy, 2008, 86(2):229-233.
- [9] Lévinas E. Totality and Infinity : An Essay on Exteriority [M]. Boston: Kluwer Academic Publishers, 1991:199.
- [10] 霍布斯. 利维坦[M]. 黎思复,黎廷弼,译. 北京:商务印书馆,2009:104.
- [11] Kierkegaard S. Kierkegaard's Writings: Volume 6: Fear and Trembling [M]. New Jersey: Princeton University Press, 1983:59-61.
- [12] 霍克海默,阿道尔诺. 启蒙辩证法[M]. 梁敬东,曹卫东,译. 上海:上海人民出版社,2006:19.
- [13] 威廉姆斯. 真理与真诚[M]. 徐向东,译. 上海:上海译文出版社,2013:143.
- [14] Dewey J. Human Nature and Conduct [M]. New York: Cosimo Classics, 2007:132.
- [15] 樊浩. 缺乏信用,信任是否可能[J]. 中国社会科学,2018(3):51-59.
- [16] 马克思,恩格斯. 马克思恩格斯选集:第1卷[M]. 北京:人民出版社,2012:135.
- [17] Sztompka P. Trust: A Sociological Theory [M]. Cambridge: Cambridge University Press, 1999.
- [18] 吉登斯. 社会的构成[M]. 李康,李猛,译. 北京:中国人民大学出版社,2016:313.
- [19] 赵汀阳. 论可能生活[M]. 北京:中国人民大学出版社,2004:125.
- [20] Zucker L G. Production of Trust: Institutional Sources of Economic Structure, 1840—1920 [J]. Research in

Organizational Behavior, 1986, 8(2):53-111.

[21] 泰勒. 本真性的伦理[M]. 程炼,译. 上海:上海三联书店,2012.

[22] 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,2016:30.

[23] 慈继伟. 正义的两面[M]. 上海:生活·读书·新知三联书店,2014:159-163.

[24] 休谟. 人性论[M]. 关文远,译. 北京:商务印书馆,2016:526-527.

Initial Credit Quota and the Hierarchical Transmission and Transformation of Trust: A Possible Explanation for the “Leap of Trust”

SUN Zheng, ZHANG Xiaodong

(School of Philosophy, Nanjing University, Nanjing, Jiangsu, 210023, China)

Abstract: Trust constitutes an essential form of social capital in a stranger-dominated society. However, no academic consensus has been reached on how trust-giving emerges. Two dominant interpretive approaches to trust-giving can be identified: cognitivism and non-cognitivism. The former regards “rational expectation calculations” as its premise, while the latter is built upon “affective-moral inclinations”. Nevertheless, neither can offer a rational explanation of the initial “leap of trust” among strangers. By examining the above-mentioned limitations, this study proposes a “hierarchical transmission and transformation of trust” model. The core concept is the “initial credit quota”, viz. the basis of credit and a risk-backstop provided by an institutional community to its members through effectively responding to breaches of trust. The researchers demonstrate that trust-giving is a dynamic transmission process: from the efficacy of the institutional community at the macro level, to the additional trust of the group-system at the medium level, and to the individual trust at the micro level. This model provides a unified explanation for theoretical puzzles such as the “leap of trust”, and facilitates the identification and resolution of trust crises in modern society.

Key words: trust; initial credit quota; cognitivism; non-cognitivism

[责任编辑:何敏敏]