

基于富兰克林效应的心理杠杆外交与国际信任生成机制研究

胡奕帆

中国政法大学, 北京 100088

摘要: 当代国际关系中的信任建构问题一直是国际政治理论探索中的一个容易被忽视但依然重要的命题。在无政府状态的国际体系下, 国家间的战略信任既非纯粹权力结构的附属品, 亦非制度约束的简单产物, 而是植根于复杂的行为互动与认知刷新之中。传统理论范式的局限性促使我们需要在微观心理机制与宏观政治实践之间架设一个沟通的桥梁, 进而从不同角度揭示出国际信任生成的底层逻辑。本文试图通过整合认知刷新、行为互动与技术中介的三重维度来系统阐释国际信任的动态生成机制, 并基于行为心理学、社会建构主义与经济学等跨学科理论的交叉提出“心理杠杆外交”的方法论框架。研究表明, 国际信任生成机制的本质在于: 国家之间的战略互信并非静态的外交成果, 而是在持续的互动中通过行为的实践不断重塑认知标签, 并在技术工具的催化下形成自我强化系统的一个动态过程。

关键词: 心理杠杆外交; 富兰克林效应; 国际信任; 沉没成本

引言

当代国际关系研究长期受制于结构现实主义与制度主义的二元分野, 将国际信任的生成机制简化为权力制衡的副产品或制度约束的被动结果^[1], 却在无形中忽视了行为互动与认知重构的深层动力。随着全球化进程的深化与数字技术的颠覆性变革, 国际信任的建构不再局限于物质资源的交换或条约文本的签署, 而日益呈现出复杂的社会心理属性——国家间战略互信的形成与维系, 既依赖于行为实践的累积效应, 又受到认知标签的动态调适与技术中介的加速催化。传统理论范式对微观心理机制的忽视使得对国际信任的研究陷入静态分析的窠臼: 既无法解释为何某些合作能在权力不对称条件下突破结构性桎梏, 也难以阐明技术革命如何重塑信任生成的底层逻辑。基于此, 本文试图突破学科边界的束缚, 以富兰克林效应为切入点, 构建一个整合认知刷新、行为互动与技术中介的三维咬合模型, 揭示国际信任从个体心

理向宏观政治场域转译的动态过程。通过融合行为心理学、社会建构主义与经济学的跨学科视角, 本文旨在阐明: 国际信任的本质并非外在制度安排的固化产物, 而是在持续互动中通过叙事重构、行为承诺与技术赋能的协同演化所形成的自组织系统。这一理论转向不仅为破解国际信任的“生成黑箱”提供了新的分析框架, 也为数字时代的外交实践开辟了方法论革新的可能路径。

1 理论框架: 从个体心理到国际互动的认知转译

国际信任的生成机制本质上是一种从微观心理向宏观政治场域的转译过程。这种转译并非简单的类比移植, 而是基于国家作为一定意义上“拟人化”行为体的认知特征^[2], 将个体层面的心理规律转化为集体决策逻辑的复杂适应过程。行为心理学中的认知失调理论、社会认同理论以及经济学中的沉没成本效应, 共同构成了理解这一转译机制的理论基石。其中, 富兰克林效应揭示的“行为

创造认同”的规律为破解国际信任困境提供了关键启示——当国家实施帮助行为后，其决策层与公众会经历一个“行为归因-认知重构-信任升级”的心理链条，这种内在的认知调整机制超越了传统外交策略的功利性计算，形成了信任建构的独特路径。

富兰克林效应的政治学转译在于其揭示了施助行为的自我说服功能。富兰克林效应的心理学本质是人类为消解“行为与态度不一致”的心理矛盾而进行的重新认知。富兰克林在自传中记载的“借书法则”——通过向政敌借阅一本书籍化解敌意——揭示了一个深刻规律：当个体为他人付出微小努力后，会不自觉地调整对对方的评价，以合理化自身行为^[3]。这种认知失调理论在国际关系场域的转译意味着国家的对外行为同样也受“心理一致性”的需求驱动。国家在实施援助或合作行为时，初始动机可能包含现实利益的考量，但行为本身的完成会触发认知系统的重构需求。根据认知失调理论，当行为与既有身份标签存在矛盾时（例如以现实主义自居的国家持续进行无偿援助），决策者会通过叙事重构将具体行为纳入更宏大的价值框架（如“全球责任”或“文明使命”），从而消解心理层面的认知张力。这种叙事重构并非简单的舆论宣传，而是深层次的认知调适过程——通过语言符号的重组，将工具性行为升华为价值观实践，进而形成对受助国的积极评价。在此过程中，国家的决策逻辑呈现出双重性：施助行为既是现实政治的计算结果，也是身份认同连贯性需求的产物。这种双重性使得合作行为能够超越短期功利性，成为信任生成的催化剂。

国家拟人化的社会建构属性则为心理杠杆理论提供了本体论基础。国际政治行为体并非完全意义上机械地遵循着权力逻辑行事，而是在互动过程中不断地刷新着符合彼此身份标签的认知。当国家 A 对国家 B 实施援助

时，其行为不仅仅局限于传递物质资源，更是通过叙事框架投射的特定的身份信号。若援助行为被解读为“伙伴责任”，则受助国在施助国的认知图谱中将被归类为“盟友”；若被解读为“战略性投资”，则可能强化“工具性合作”的标签。并且，这种身份建构具有自我强化的特性：一旦某种身份标签被确立，后续的合作将更易获得合法性支持，而冲突行为则需要付出更高的认知成本。值得注意的是，这种拟人化过程并非单向的标签投射，而是互动双方共同参与的意义协商。施助国通过行为传递身份信号，受助国则通过反馈（如公开致谢或反向回馈）不断地修正这一标签，使得信任的积累呈现出动态演进的螺旋特征，这种动态性便表明了国际信任的生成本质上是一种社会建构过程^[4]，其稳定性依赖于行为实践与叙事框架的持续匹配。

沉没成本效应则从行为经济学的视角揭示了信任维系的内在惯性。在个体决策中，前期投入的不可回收性会显著影响后续选择^[5]；而在国际互动中，这种心理机制则表现为“合作承诺的自我强化”。当国家投入了政治资本、经济资源或声誉资本后，退出合作的心理接受成本将远超物质损失。这种成本不仅包括前期投入的“打水漂”，更涉及其自身身份标签崩解的风险——例如施助国突然终止援助将导致“可靠伙伴”形象的质疑。沉没成本对信任的锁定作用体现在三个层面：物质投入的累积形成路径依赖，迫使国家持续追加资源以维持合作成果；认知标签的固化提高了政策转向的国内政治成本；国际声誉的积累具有不可逆性，因而一次信任的破裂可能导致长期声誉的折损。以上三个层面的相互作用共同形成了“认知-行为”的闭环反馈：初始合作触发认知刷新，认知刷新推动更大规模的行为投入，而新的投入进一步强化认知标签。这种动态循环使得信

任一旦建立，便倾向于更多地通过惯性维持而非纯理性计算，从而呈现出一种战略韧性的独特属性。

2 动态机制：认知、行为与技术的协同演化

国际信任的建构机制并非单一维度的线性过程，而是认知重构、行为互动与技术中介三者深度交织的一个动态系统。这一系统的核心特征在于其咬合性——认知调整驱动行为策略的优化，行为实践反哺认知标签的固化，而技术工具则通过重塑互动场景与信息传播路径来加速或扭曲这一循环的运转效率。理解上述三者间的协同与张力需要从微观心理机制到宏观技术生态逐层解析，从而揭示出国际信任生成的底层逻辑。

认知刷新作为信任生成的底层驱动力，其运作遵循“行为-叙事-认同”的转化逻辑。当国家实施施助行为时，内在的认知冲突源自行为动机与既有身份标签的不一致。为了消解这种张力，决策者必须通过叙事策略将具体行为纳入更宏大的价值框架。这一过程的关键在于“意义赋予”——通过语言符号的调整，将工具性合作升华为价值观实践。例如，将经济援助转化为“文明使命”的叙事，不仅能够有效降低公众的认同成本，还能为后续行为的升级提供认知基础和民意合法性^[6]。另外，叙事策略的连贯性与社会共鸣度也会直接影响认知刷新的深度：若施助国的叙事框架能与受助国的国内主流意识形态无缝衔接，则公众的认同成本显著降低；若存在冲突（如将军事干预包装为“民主输出”），则可能会触发认知的反弹。这种动态的平衡表明认知刷新既是信任生成的起点，也是其持续性的保障。国际信任的脆弱性往往源于叙事的断裂——当行为实践无法被既有认知框架容纳时，信任的根基随即面临崩塌风险。

行为层面的信任建构则遵循“互惠-强化”的螺旋动力学。初始合作行为的完成（如技术援助协议的签署）触发施助国的认知刷新，进而推动更大规模的行为投入（如建立长期合作机制）；新投入的履行则进一步巩固认知标签，形成“行为强化认知，认知驱动行为”的正反馈循环^[7]。而这一螺旋的核心在于“行为承诺的可信度”——只有当合作行为具备可验证性与连续性时，认知标签的稳定性才能得到保障。在此过程中，行为策略的设计需要兼顾“低门槛入口”与“高扩展潜力”。借鉴心理学中的登门槛效应，初始合作的微小性可以有效降低心理阻力，从而为后续的深度合作铺平道路。例如通过文化合作项目建立初步互信，再逐步扩展至经济或安全领域。这种从“低级政治”到“高级政治”的渐进策略的关键在于行为序列的连贯性：每一步合作都需要强化前一阶段的认知标签，同时为下一阶段的行为升级预留接口。若行为序列出现断层（如援助突然中断），则螺旋动力可能逆转为信任消解的恶性循环。因此，行为互动的节奏控制与风险缓冲机制成为了维持信任螺旋上升的关键要素。

技术中介作为信任生态的重塑力量，其作用实质上远超普遍认为是工具性范畴。如今数字技术的广泛应用已经通过重构国际互动的信息环境而对认知标签的生成与传播带来深刻影响。恰如社交媒体平台通过“叙事裂变”加速认知标签的扩散——当施助国的合作行为被转化为短视频、推文或虚拟场景中的符号互动时，其传播效率与情感动员能力呈指数级增长。然而，这种技术赋能的传播模式也可能导致认知浅表化——复杂的合作逻辑被简化为情感符号，从而削弱公众对合作本质的理性认知。算法技术的风险则在于“认知操纵”与“信任熵增”——基于用户偏好的内容推送可能会将合作叙事局限在

特定受众圈层，从而加剧国际社会的认知割裂与极化。施助国的善意行为在受援国获得算法加持的正面传播，却很可能在第三方国家因政治偏见被曲解为“地缘扩张”。更危险的是，如今 AI 技术的滥用可能会被制造出越来越多的虚假合作场景，彻底瓦解信任的信息基础。因此，技术工具在提升信任建构效率的同时也带来了认知真实性与系统稳定性的双重挑战。

3 系统整合：信任生成的三维咬合模型

国际信任的生成机制最终呈现为认知、行为与技术三维度的深度咬合。这种咬合性体现在三个层面：首先，认知刷新为行为互动提供意义框架，行为实践为认知标签提供验证基础，技术工具则为两者的转化提供加速器；其次，三者间的反馈循环形成动态平衡——认知偏差可能通过行为调整得以修正，技术异化可能触发新的认知刷新；最后，系统的稳定性依赖于三者的协同进化速度。当技术变革速度超过认知适应能力时（如当下人工智能技术爆炸所引发的信息过载），或着行为实践滞后于认知期待时（如承诺因各种原因而未能兑现），系统将面临失衡风险。

在此模型中，信任的生成不再是单向的因果关系，而是多要素之间非线性互动的涌现现象。认知刷新通过叙事策略将离散行为整合为连贯的身份叙事，行为互动通过可验证的承诺强化叙事可信度，技术中介则通过信息传播放大两者的协同效应。这种咬合模型的最大理论价值在于揭示了国际信任的复杂系统属性——它既非完全由物质实力决定，亦非单纯依赖观念塑造，而是在认知、行为与技术的持续互动中动态演化的自发形成的组织系统。这种系统视角为理解信任的脆弱性与韧性提供了新的分析框架：信任危机本质上是系统要素的协同断裂，而信任修复则需要三者的重新校准。

4 心理杠杆外交：国际信任建构的方法论革新

本文的核心观点在于提出“心理杠杆外交”的理论范式。心理杠杆外交理论的突破在于将信任建构从结构决定论中解放出来，赋予行为实践以本体论地位。传统理论往往将信任视为权力结构或制度环境的衍生品，忽视了个体心理机制向集体决策逻辑转译的可能性。本文提出的三维咬合模型证明，即便在权力不对称的体系中，弱国通过精心设计的行为策略（正如前文所提到的低门槛的合作请求）也能激活强国的认知刷新机制，进而在其身份叙事中植入平等伙伴的认知标签。这种机制的有效性源自人类认知系统的普遍规律——行为对态度的塑造作用具有跨层级的稳定性。

同时，该理论对数字时代的外交实践具有重要启示。技术工具不再是中立化的传播渠道，而是深度嵌入信任生成过程的认知转化器。信任建构的成功不仅取决于行为策略的设计，更依赖于对技术中介效应的精准把控——如何在提升叙事传播效率的同时有效防范认知操纵风险成为“数字外交”所面临的的核心挑战^[8]。

最终，心理杠杆外交理论将指向一种新的国际关系认识论：信任不再是外交行为所追求的一个目标，而是外交实践本身的存在方式。国家在与他者的持续互动中，通过行为实践不断重塑着自我认知与他者形象，这种双重建构过程使得信任既是合作的条件，也是合作的结果。这种认识论转向要求我们超越传统的主客二分思维，将国际关系视为认知与实践共同演化的动态场域。从这个意义上讲，国际互信的建构并非一蹴而就的制度安排，而是人类在无政府世界以互动为桥梁、以实践为路径，不断追索共同意义的漫长历程。

结论

本文通过构建认知、行为与技术的三维咬合模型系统揭示了国际信任生成机制的内在逻辑：战略互信的建立既非单纯依赖物质利益的理性计算，亦非被动接受制度规范的约束，而是一个在行为实践触发认知刷新、技术工具加速叙事传播的循环反馈中动态演化的复杂过程。研究表明，富兰克林效应所揭示的“行为塑造认同”规律在国际政治场域具有显著的转译价值——无论是施助行为的自我说服功能，还是沉没成本对信任锁定的惯性效应，均表明心理机制能够穿透权力结构的表层逻辑，在互动中重构国家的身份叙事。心理杠杆外交理论的提出，不仅将信

任建构从结构决定论的桎梏中解放出来，更赋予弱势行为体通过策略性合作撬动强国认知标签的能动性空间。在技术深度介入国际互动的当下，数字平台既可能通过信息裂变放大信任螺旋的正向效应，也可能因算法操纵与虚假叙事导致认知熵增，这要求外交实践必须兼顾技术赋能的效率和认知真实性的把控。未来研究可进一步探索三维模型的实证边界，例如数字外交中技术中介的阈值效应，或者危机情境下信任修复的动态机制，从而深化对国际信任的复杂属性的理解。归根结底，国际信任的生成与维系是一场认知与实践共舞的永恒博弈，唯有在动态平衡中把握行为策略的精确刻度，方能在无政府状态的国际体系中编织出韧性共生的合作之网。

参考文献

- [1] 葛腾飞. 世界动荡变革期的国际信任建设：国际体系的视角[J]. 人民论坛·学术前沿, 2024, (16): 66-75.
- [2] 鲍高志. 国家间信任研究的问题之争与理论构建 [N]. 中国社会科学报(理论版), 2024-10-29.
- [3] 陈允皓著. 富兰克林效应[M]. 湖南文艺出版社, 2021. 08.
- [4] (美) 温特(Wendt, A.) 著, 秦亚青. 国际政治的社会理论[M]. 上海人民出版社, 2014.
- [5] 郑先著. 投资心理学：投资其实是一场心理游戏[M]. 中国法制出版社, 2017.
- [6] 霍淑红. 政策扩散视角下美国私人基金会的对外援助[J]. 山西大学学报(哲学社会科学版), 2018, 41(05): 87-93.
- [7] (美) 罗伯特·西奥迪尼著, 闫佳. 影响力：全新升级版：the psychology of persuasion[M]. 北京联合出版公司, 2021.
- [8] 陈小鼎, 刘洋. 数字外交对国际话语权博弈的影响及中国应对[J]. 吉林大学社会科学学报, 2024, 64(05): 177-189+239.

作者简介：胡奕帆（2002--），男，汉族，河南周口人，中国政法大学政治与公共管理学院国际政治系外交学专业硕士研究生，学生，研究方向：国际战略与大国外交