

# 互惠策略与国际关系

黄真<sup>1</sup>

(1. 上海行政学院, 上海 200233)

**摘要:** 互惠策略及其制度化被新自由主义者认为能够实现合作进化, 本文是对这个基本认识的思考。互惠策略在理论上的意义, 在于一定程度上突出了国家的能动性、适应性和“进程建构”的特征, 使人对国际关系的进化持有审慎的乐观; 但理论意义又是有限的, 突出表现在互惠策略预定国家的理性利己本性, 忽视了若干互惠是运作于非自利的动机和关系背景之下。互惠策略在政策上的意义, 在于鼓励国家超越“霍布斯逻辑”, 敢于给予、敢于接受和敢于回报; 不过其政策意义也是有限的, 表现在使领导人和民众对重复博弈和确定性产生具有危险性的偏好。未来的互惠策略研究应该聚焦于三个方面: 互惠的政策研究, 即国家应该怎么给予、怎么接受和怎么回报; 多元互惠研究, 即多国的战略互惠研究; 互惠实践与国际关系和国家动机的再造, 明确互惠策略与合作进化或退化之间的机制。

**关键词:** 互惠策略; 条件; 悖论; 未来研究

中图分类号: D81

文献标识码: A

文章编号: 1004 - 8049(2012)03 - 0048 - 10

著名的社会学家格奥尔格·西美尔 (Georg Simmel) 曾经提出社会科学的本性问题——社会何以可能? (How is society possible?)<sup>①</sup>他给出的答案是人际之间的互惠。其实, 在诸多文化人类学、社会交换理论和社会生物学著作中,<sup>②</sup>很多著名学者也把互惠作为社会或自然界的基本运行规则。因此, 倘若果真存在一个“国际社会”的话, 这个国际社会何以可能便成为国际问题研究的根本性问题。英国学派认为, 正是国家在主权等根本性价值

问题上的互惠, 使得一个弱质并有可能不断增强的国际社会存在。<sup>③</sup>

国际关系理论中存在两种形式的互惠, 一种是作为道德规范的互惠, 一种是作为国家行动策略的互惠。<sup>④</sup>但相比而言, 在国际关系理论和理论图谱中, 互惠更多的是指代一种国家策略。之所以说更多地指代策略, 是因为在西方国际关系理论史上的第三次新新对垒中, “互惠”基本上构成新自由制度主义的一个核心词汇, 互惠策略被新自由主义者认为能够极为有

收稿日期: 2011-10-03; 修订日期: 2012-02-20。

基金项目: 国家社会科学基金项目“金融危机后新兴经济体参与全球经济治理的挑战及我国对策研究”(11CGJ007)的阶段性成果。

作者简介: 黄真(1983—), 男, 江西广昌人, 上海行政学院政治学教研部讲师, 法学博士, 主要研究方向: 国际关系理论。

① Georg Simmel, “How is society possible?”, *American Journal of Sociology*, 1910, Vol. 16, No. 3, pp. 372 - 391.

② 如马林诺夫斯基 (Bronislaw Malinowski)、马塞尔·莫斯 (Marcel Mauss)、列维-斯特劳斯 (Claude Levi-Strauss)、彼得·布劳 (Peter Blau)、霍曼斯 (George C. Homans)、汉密尔顿 (William Hamilton)、克鲁泡特金 (Peter Kropotkin) 等。

③ Hedley Bull, *The Anarchical Society: A Study of Order in World Politics*, London: Macmillan, 1977.

④ 黄真“国际关系中的互惠: 一项研究评估”, 《当代亚太》, 2010年第5期。

效地促成合作。为此,米歇尔·泰勒(Michael Taylor)等人通过理论和实践验证,<sup>①</sup>广泛运用互惠理论分析国家间合作的形成与发展。

下文将重点讨论国际关系中的互惠策略。第一,介绍互惠策略本身及其效度所必须严格满足的条件;第二,互惠策略本身蕴含的四大悖论;第三,互惠策略在理论和政策上的意义及其有限性;第四,互惠策略的未来研究方向。

### 一、互惠策略及其有效性的条件

在国际关系研究中,互惠作为一种国家策略存在历时已有半个世纪,早在20世纪60年代,以查理斯·奥斯古德(Charles Osgood)、阿米塔·埃兹奥尼(Amitai Etzioni)等为代表的学者探究国家应该如何给予才能较为容易地引发对方的积极反应并回报善意。他们认为,GRIT(gradual reciprocation in tension-reduction)可以有效缓解安全困境,施惠方要持续地、渐进地给予(而不必太在乎对方是否即刻回报),这些给予需具有一致性、连贯性和代价性,以让对方逐步确定自身的善意。<sup>②</sup>之后拉波波特(Anatol Rapoport)、谢林(Thomas Schelling)都意识到互惠策略在解决囚徒困境问题上的重要意义。<sup>③</sup>不过,互惠策略研究的最著名学者是阿克塞尔罗德(Axelrod)。这一方面得益于他个人丰富的学科背景(生物学、经济学、政治学和国际关系学),另一方面在于国际关系理论史上第三次争论的高峰期恰逢他的名著《合作的进化》(1984)出版。此外,与基欧汉的个人私交使得他主动或“被动”地成为新自由主义的重要人物之一。

阿克塞尔罗德的互惠策略很简单。不过在论述他的互惠策略之前,有必要澄清“互惠”的英文翻译问题,互惠基本上被翻译成reciprocity,而reciprocity是正向的互惠,即合作对合作,有礼尚往来、投桃报李之意。但实际上,国际关系意义上的互惠不仅包括正向互惠,也包括消极互惠,如世仇、相互威慑、相互确保摧毁、僵局等。实际上,互惠的最为准确的英文翻译应当是reciprocation,这个词才囊括了两种互惠类型。

阿克塞尔罗德的互惠策略遵循以下逻辑:一方首先以合作的姿态对待他者,之后便根据对方的反应而做出选择,如果对方合作,则合作,如果对方背叛或欺骗,则惩罚或威慑。阿克塞尔罗德认为,这个策略尽管是计算机模拟的策略,但是在国际关系中的效用却非常明显。国际关系的最重要特征便是无政府状态,如果每个国家都实行“互惠”策略,理性利己的国家通过博弈会逐步认识到,合作比不合作好,最终将共同超越“囚徒困境”,实现合作的进化,而且进化是不可逆转的。<sup>④</sup>互惠策略由于具备“投桃报李”、“以牙还牙”的特征,在国际关系学中又被称为“一报还一报”(tit-for-tat)策略或“针锋相对”策略。自互惠策略提出以来,它便和收益结构(payload-structure)、未来阴影/未来的重要性(the shadow of future)以及博弈人数等因素一道,几乎构成新自由制度主义合作理论的分析范式,众多学者又通过各种案例试图验证互惠策略的有效性和重要性,并探究互惠策略与其他三个因素的相互关系。<sup>⑤</sup>

一般而言,互惠策略作为一种计算机模拟,必须满足苛刻的条件方能产生预期的合作结果,这些条件至少包括:

① 代表人物还有:拉波波特(Anatol Rapoport)、阿克塞尔罗德(Robert Axelrod)、基欧汉(Robert Keohane)、斯坦(Arthur Stein)、米尔纳(Helen Milner)、奥伊(Kenneth Oye)等。

② Charles Osgood, *An Alternative to War or Surrender*, Urbana: University of Illinois Press, 1962. Amitai Etzioni, *The Hard Way to Peace: A new Strategy*, New York: Crowell-Collier, 1962.

③ Anatol Rapoport and Chammah Albert, *Prisoner's Dilemma*, Ann Arbor: University of Michigan Press, 1965. Anatol Rapoport, "Escape from Paradox", *Scientific American*, Vol. 217, 1967, pp. 50-57. Thomas Schelling, *The Strategy of Conflict*, Massachusetts: Harvard University Press, 1963.

④ [美]罗伯特·阿克塞尔罗德著,吴坚忠译《合作的进化》,上海人民出版社,1995年版,第40、135页。

⑤ Kenneth Oye, eds., *Cooperation Under Anarchy*, N. J: Princeton University Press, 1986; Joanne Gowa "Anarchy, Egoism and Third Images: The Evolution of Cooperation and International Relations", *International Organization*, 1986, pp. 167-186; Joshua Goldstein and John Freeman, *Three-Way Street: Strategic Reciprocity in World Politics*, Chicago: Chicago University Press, 1990; Robert Axelrod "On Six Advances in Cooperation Theory", Prepared for a Special Issue of *Analyse & Kritik on The Evolution of Cooperation*, 2000, pp. 1-31, available at: <http://www.fordschool.umich.edu/research/papers/PDFfiles/00-003.pdf>.

第一,国家必须处于无限博弈链中。按照经典博弈论的基本常识,如果博弈是有限的,那么理性的行为体根据回溯推理将使有限博弈等同于单次博弈,因而相互背叛成为首选,互惠的合作功能便会消失。<sup>①</sup>当然,正如霍布斯所说,与人类的高死亡率截然不同,国家的消亡率是如此的低,以至于国际关系学者基本上认为,国际关系注定是无限博弈的。

第二,国家必须处于决策相互依赖中。每一方都没有主导战略(dominant strategy)而必须处于共同决策(joint-decision)之中,即每一方的选择都在相当程度上取决于对方的选择。<sup>②</sup>如果一方或两方具有主导战略(如永远合作或永远背叛),那么任何策略,包括互惠策略都不起作用。

第三,博弈结构不发生急剧式的重大改变,即每次博弈的收益值大体稳定,收益不稳无法保证行为模式的稳定性。<sup>③</sup>例如,当一国认为当下背叛的收益是如此之大以至于可以一劳永逸地解决所有难题时,那么即使对方实施互惠策略进行威慑和惩罚,该国也不在乎。与此紧密相关的另一个条件是,未来足够重要且行为体关注未来。在阿克塞尔罗德和基欧汉等人看来,“未来重要”实际上有两个含义:一是未来相遇的可能性很大;二是未来收益与当下背叛的收益相比足够重要。<sup>④</sup>行为体如果认为未来不重要,互惠策略将失去价值。

第四,国家必须具备实施互惠策略的意愿和能力。就意愿而言,除了愿意以合作回报对方合作外,还愿意以威慑或惩罚回应对方的欺骗和背叛。但当需要惩罚时,又面临三大难题:一是如果在集体中惩罚背叛者,则很有可能存在集体行动的困境,集体惩罚往往因搭便车者、漠不关心者而受阻。<sup>⑤</sup>二是有时候人们会允许或容忍一定数量的背叛者存在,这样有利于集体收益的优化或最大化,这种情况在共同规避困境(common aversion dilemma)或协调型博弈(coordination game)中最为明显。<sup>⑥</sup>三是惩罚背叛者往往会同时惩罚到其他合作者。就能力而言,当对方背叛后,国家是否具有相应的能力实

施惩罚。国家应该具备问题领域的大致对等的合作或报复能力,或具备不同问题领域的议题联系能力,或者具备跨越国内障碍的能力。<sup>⑦</sup>

第五,信息透明和共同理解。尽管信息充足并非总是好事,但确定性,尤其是确定对方是否是合作者的问题相当关键,这被认为是进行互惠的前提条件,也是互惠策略发挥效用的前提条件。共同理解意味着双方尽管都有合作的意图,但必须对什么构成合作、什么构成背叛具有共同的理解,否则没有互惠只有讨价还价。制度在这方面发挥了重要作用。<sup>⑧</sup>

## 二、互惠策略的四大悖论

互惠策略的合作效应除了要满足以上五大条件外,还存在四大悖论,这四大悖论决定了互

<sup>①</sup> Arthur Stein, "Coordination and Collaboration: Regimes in An Anarchical World", in Stephen Krasner, eds., *International Regimes*, Ithaca: Cornell University University Press, 1983, pp. 115 - 140. [美]道格拉斯·诺思著,杭行译《制度、制度变迁与经济绩效》,上海人民出版社,2008年版,第79页。

<sup>②</sup> [美]托马斯·谢林著,谢静等译《微观动机与宏观行为》,中国人民大学出版社,2005年版,第4页。

<sup>③</sup> Arthur Stein, *Why Nations Cooperate: Circumstance and Choice in International Relations*, Ithaca: Cornell University Press, 1990, pp. 100 - 101.

<sup>④</sup> Robert Axelrod, "On Six Advances in Cooperation Theory", p. 18.

<sup>⑤</sup> [美]曼库尔·奥尔森著,陈郁等译《集体行动的逻辑》,上海人民出版社,1995年版。Garrett Hardin, "The Tragedy of the Commons", *Science*, Vol. 168, 1968, pp. 1243 - 1248.

<sup>⑥</sup> 同②,第102 - 110页。关于共同规避困境和协调型博弈,可见莉萨·马丁:“利益、权力和多边主义”, [美]利萨·马丁、贝思·西蒙斯编,黄仁伟等译《国际制度》,上海人民出版社,2006年版,第35 - 64页。Duncan Snidal "Coordination Versus Prisoners' Dilemma: Implications for International Cooperation and Regimes", *The American Political Science Review*, Vol. 79, 1985, pp. 923 - 942.

<sup>⑦</sup> Kenneth Oye, "Explaining Cooperation Under Anarchy: Hypotheses and Strategies", in Kenneth Oye, eds., *Cooperation Under Anarchy*, p. 16. Helen Milner, "International Theories of Cooperation among Nations: Strengths and Weakness", *World Politics*, Vol. 44, 1992, pp. 466 - 496.

<sup>⑧</sup> Geogre Downs, David Rocke and Randolph Siverson "Arms Race and Cooperation", in Kenneth Oye, eds., *Cooperation Under Anarchy*, pp. 133 - 137. Robert Keohane, "The Demand for International Regimes", *International Organization*, Vol. 36, 1982, pp. 325 - 355. 但关于制度的混合(积极与消极)效应,见 [美]道格拉斯·诺思著《制度、制度变迁与经济绩效》,第108,137页。

惠策略不仅在理论上,而且在实践操作上都面临诸多的困难,也因此遭到不少的诟病。

### 第一,向前看与向后看的悖论。

在新自由主义者看来,互惠策略的顺序是这样的:首先合作,对方反应,然后根据反应再做出反应。也就是说,除了第一轮是由施惠方首先合作外,从第二轮开始双方都是以对方上一轮或者说历史的行为为基础。但互惠策略又是以国家意识到“未来的重要性”(即以未来对方会回报或再次给予)为基础的,也就是国家往往通过前瞻未来而决定当下的选择,因而这便产生了一个最为显要的矛盾,国家到底是历史反应型的还是未来前瞻型。<sup>①</sup> 比如说,当对方背叛后,我是报复呢,还是意识到未来相当重要从而容忍呢? 如果报复,便很容易形成僵局,如果重视未来收益,则必须宽容。阿克塞尔罗德也意识到,博弈中容易出现“噪音”(noise,即实施选择过程中的随机错误),如果奉行严格的针锋相对策略,将导致无休止的背叛,结果是国家理性导致绝对的无效率。鉴于此,他进行了一定程度的“策略改进”。改善后的互惠策略包括:宽容(generosity)的“一报还一报”,即允许一定比例的实施背叛的博弈者不受惩罚,尤其是在对方可识别的非故意性犯错的情况下给予原谅;悔悟(contrition)的“一报还一报”,如果自己不小心犯错,立马向对方表明悔悟并立即改正,不再背叛下去。阿克塞尔罗德通过计算机模拟得出结论,宽容和悔悟的“一报还一报”策略在实际运用中极为有效,博弈者合作的几率十分高。<sup>②</sup> 不过总体看来,尽管阿克塞尔罗德试图超越向前看向后看的困境,但似乎成效不大。

### 第二,信息透明的悖论。

新自由主义的一个核心认识便是,信息充足和透明是国际合作的必需。就互惠策略而言,信息透明有利于将合作者与背叛者区分,也将导致误解、误算和错误知觉的消失或大幅度减弱,而正是这些错觉导致冲突。<sup>③</sup> 但是,信息透明可能产生不良后果,例如当一个机会主义者(对方合作我背叛,对方惩罚我合作)具备充足信息得知对方是一个合作者(总是合作)的话,那么机会主

义者便倾向于剥削,合作失败;反过来,有时信息不明和错觉反而有利于互惠策略发挥效用。如果一个机会主义者把一个合作者误认为是互惠者(对方合作我合作,对方背叛我背叛)的话,互惠策略的威慑力量可能导向双方合作,对此,斯坦做了有趣而又深刻的分析。<sup>④</sup>

### 第三,未来重要性的悖论。

尽管未来重要性表明了“行为体未来相遇的可能性很大和未来收益相当有吸引力”,但它还内含一种对互惠策略极为不利的含义,即未来重要意味着一方灭绝(extinction)的可能,毫无疑问,这仍然符合未来重要的含义。这说明不能忽视利益分配问题,即相对收益问题。换句话说,如果我能在相互背叛中把对方拖垮,我将采取背叛策略(谁失去更多),或者如果我能在相互合作中获取更多收益(谁得到更多),从而在未来存在把这种收益转化为毁灭对方的能力的话,那么对方一旦预期到这个结果,将肯定不会合作,这两种情况都大大削弱了互惠策略的有效性。这也从另一个侧面说明,未来收益的确定性将损害合作,而某种不可知或模糊性反而有利于合作。<sup>⑤</sup>

### 第四,未来与当下的悖论。

互惠策略十分看重“未来的阴影”,但正如著名的互惠理论研究者科尔姆(Serge - Christophe Kolm)所言:“(人们的)顺序推理实际上难以超过

<sup>①</sup> Arthur Stein, *Why Nations Cooperate*, p. 100.

<sup>②</sup> [美]罗伯特·阿克塞尔罗德著,梁捷等译《合作的复杂性》,上海人民出版社,2008年版。黄真“从‘互惠利他’到‘强互惠’——国际合作理论的发展与反思”,《国际关系学院学报》,2009年第4期。

<sup>③</sup> [美]斯蒂芬·范·埃弗拉著,何曜译《战争的原因》,上海人民出版社,2007年版。Robert Jervis, *Perception and Misperception in International Politics*, Princeton: Princeton University Press, 1976.

<sup>④</sup> Arthur Stein, *Why Nations Cooperate*, chapter 3.

<sup>⑤</sup> Joseph Grieco “The Relative - Gains Problem for International Cooperation”, *American Political Science Review*, Vol. 87, No. 3, 1993, pp. 729 - 743. Joseph Grieco, *Cooperation among Nations: Europe, America and Non - tariff Barriers to Trade*, Cornell University Press, 1990. David Baldwin, eds., *Neorealism and Neoliberalism: The Contemporary Debate*, New York: Columbia University Press, 1993. Arthur Stein, *Why Nations Cooperate*, chapter 4. D. Kreps and R. Wilson, “Rational Cooperation in the Finitely Repeated Prisoner’s Dilemma”, *Journal of Economic Theory*, Vol. 27, 1982, pp. 245 - 252.

两或三步——有限理性的原因……至少一个行为体对于进程终于何时或对方将如何行为的不可知已经足以让彼此的关系维系。”<sup>①</sup>换句话说,由于未来收益的不确定性与当下收益的吸引力和相对确定,大多数国家倾向于关注当下,国家总体上是保守理性,而非富有远见的。<sup>②</sup>更为重要的是,基于未来回报的合作在合作最需要的时候却恰恰最容易崩溃,<sup>③</sup>例如当一国面临生死存亡或巨大灾难的危急时刻,如果仍固守对方未来回报的原则将不可能存在任何合作,因为该国很可能没有未来,很难想象汶川地震世界大援助的行为是为了指望受益国未来的回报。所以,互惠策略实施可能更多的是依据当下。

### 三、互惠策略的理论和政策意义

尽管互惠策略需要满足苛刻的条件才能促进合作演化,也尽管互惠策略内含四大悖论,但并不因此掩盖互惠策略的理论意义和政策意义。

#### 3.1 理论意义

就理论意义而言,互惠策略一定程度上突出了国家的能动性、适应性和某种“进程建构”<sup>④</sup>的特征,使人对国际关系的进化持有审慎的乐观。研习国际关系理论,总会有这么一种判断,国际关系的性质是由先验设定的人性、国家性和体系性推导出来的,这种研究路径不重视国家间合作与冲突的动力学,忽视对“关系”的研究,<sup>⑤</sup>因而是静止与僵化的。战略互动,按照谢林的话,就是行为体的“决策相互依赖”,<sup>⑥</sup>即A的选择构成B的环境,B的反应反过来又构成A的环境,正是这种选择与反应的“进程”决定了行为体间的关系及其性质以及由此决定的行为体彼此的角色定位,也正是在这个意义上,我们认可了温特(Alexander Wendt)“无政府状态是由国家造成的”这句话的深刻含义。简单说,国际关系不是必然依靠先验设定和外部强制的,而是可以在国家的相互作用中内生而成的。外部强制以促成合作或互惠是国际关系和政治思想史中的一大特点。按照霍布斯的理论思路,

没有一个利维坦(大致可理解为“作为国家政府的集权机构”),人类合作是不可能产生的。而按照卢梭的理论路径,一些自由人之间相互交往会产生“公意”,在公意存在的情况下,通过某种“社会契约”,就可以达到维持人类社会或社群之间的某种合作。但通观卢梭著作,不难发现,他的社会理论是以共和制为形式的专制独裁理论为导向的。概言之,从霍布斯到卢梭,西方古典社会契约理论的一个基本假设是:一个具有共同利益的群体会在某种外在强制力量和社会安排下为实现共同利益而采取集体行动,并且这要么导致君主专制,要么导致共和独裁。今天我们也许会发现,在20世纪由美国著名经济学家奥尔森所诘问的人类社会的“集体行动的逻辑”,实际上是把“人类合作如何可能”这类霍布斯和卢梭式的老问题,在现代社会科学的话语语境中重新提了出来,而奥尔森给出的答案是“除非存在强制或其他某些特殊手段以使个人按照他们的共同利益行事,有理性的、寻求自我利益的个人不会采取行动以实现他们共同的或集团的利益。”很显然,奥尔森只不过在20世纪复述了一遍霍布斯和卢梭的答案。如果简单概括一下西方社会政治和经济思想史中的合作问题,不难看出,思想家们对合作的态度总体上是悲观的,这种悲观来自于其对人性的看法,人虽然有理性,但整体上是自利的,霍布斯的自然状态、卢梭的猎鹿游戏、奥尔森的经济人逻辑等无不体现了这种先验设定,因而外部强制力

① Serge-Christophe Kolm, *Reciprocity: An Economics of Social Relations*, Cambridge: Cambridge University Press, 2008, p. 72.

② Arthur Stein, *Why Nations Cooperate*, p. 89, 189.

③ Samuel Bowles and Herbert Gintis, “The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Population”, *Theoretical Population Biology*, Vol. 65, No. 1, 2004, pp. 17-28.

④ 秦亚青、魏玲“结构、进程与权力的社会化”,《世界经济与政治》,2007年第3期。

⑤ Arthur Stein, *Why Nations Cooperate*, p. 13, 20-21, 176-184, 192. 秦亚青“关系本位与过程建构:将中国理念植入国际关系理论”,《中国社会科学》,2009年第3期。

⑥ [美]托马斯·谢林著,谢静等译《微观动机与宏观行为》,中国人民大学出版社,2005年版,第4页。

量是合作必须的。<sup>①</sup>

互惠策略实际上就隐含了战略互动和进程建构的要义。一国做什么取决于他国做了什么,正是在这“一报还一报”的博弈进程中,国家实现了合作或有较大的合作可能。尤其需要特别指出的是,阿克塞尔罗德本不是学国际关系的,他具有深厚的生物学知识背景,也就是说,他是在进化博弈的基础上理解互惠的。<sup>②</sup>进化过程中行为体的适应能力、学习能力、创造和超越能力,给所有国家以及全人类一个反思自身的机会,那就是,我们每一个人都要克服任何虚假的决定论思想,意识到我们自身实践具有的巨大能量,同时还要开拓视野,有意识地思考国际社会应朝哪个方向发展。

不过,互惠策略的理论意义又是有限的,它忽视了国家的非理性特质(如道德、情感、文化等)。阿克塞尔罗德认为,除了自然选择<sup>③</sup>外,人们可以通过试错(try and error)学会针锋相对,并最终发现合作而非占有头筹是有益的。阿克塞尔罗德的理论无比接近低级学习理论和行为主义的斯金纳法则,即假定人们寻求报偿规避痛苦,而对信念、认知、情感等其他心理要素完全忽视,这也是为什么阿克塞尔罗德的理论适用于从细菌聚集到国际关系所有领域的原因。<sup>④</sup>由此,互惠策略关于国家本性的先验假定(国家的理性利己)必然导致它在“进程建构”方面的贡献大打折扣。国家之间形成何种“关系”、形成何种身份定位、彼此具有何种动机,是由彼此的给予、接受与回报的实践或进程建构的。<sup>⑤</sup>不可否认,互惠策略大部分是在国家间的功利性关系和动机下实施的,但也不能遮掩有那么一部分国际关系事实,其中国家的给予、接受和回报并非出自理性的利己主义,国家间的关系除了功利性成分,还有道德的、心理的(尤其是情感的)、社会的成分。<sup>⑥</sup>实际上,基欧汉在论述扩散互惠(diffuse reciprocity)时,动机已经放在了自身所属群体的福祉而非狭隘的个人利益上,颇具有利他主义的特征,可是由于秉持理性思维,他最终陷入了“为群体也是为了自我利益”的逻辑。<sup>⑦</sup>阿克塞尔罗德的互惠策略充满了亲

社会情感(pro-social emotions),善良、宽容、悔悟、不要嫉妒、伦理仪式、自己活也让他人活……可是这所有的利他情感都建立在未来的收益(尤其是物质收益)的基础上,阿克塞尔罗德自己也说“这是指未来重要时,如果似乎不再见面,马上背叛比善良要好。”<sup>⑧</sup>

实际上,今天最为前沿的互惠策略——强互惠(strong reciprocity)已经意识到这个问题。强互惠是指一个行为体以合作姿态对待群体内任何一个成员,当某个成员违背规范时,即使违规者对该行为体没有造成利益损害,也即使在可预见的未来没有利益回报,这个行为体仍然实施惩罚行为。<sup>⑨</sup>强互惠体现出的利他主义、道德主义、集体主义使得我们思考互惠策略的关系性环境和动机时,要突破狭隘的理性利己视角,充分意识到国家的非理性关系和非理性动机。非理性关系和动机以及由此导致的给予、接受与回报可能得自道德规范的内化、和平主

① 以上分析可见:Thomas Hobbes, *Leviathan*, Oxford: Oxford University Press, 1943; Jean-Jacques Rousseau, *The Social Contract*, tran. by Maurice Cranston, New York: Penguin Books, 1968, p. 83. [美]奥尔森著,陈郁等译《集体行动的逻辑》,上海人民出版社,1995年版,第2页;关于猎鹿游戏,见[美]卡尔·多伊奇著,周启明等译《国际关系分析》,世界知识出版社,1992年版,第200-202页。

② Robert Axelrod, "An Evolutionary Approach to Norms," *American Political Science Review*, Vol. 80, No. 4, 1986, pp. 1095-1111. [美]罗伯特·阿克塞尔罗德《合作的进化》,上海人民出版社,1995年版,第5章。

③ 实际上,阿克塞尔罗德的进化原理同样不适合人际和国际交往,因为没有自然选择机制以消除无效的个体或国家。

④ Robert Axelrod, *The Evolution of Cooperation*, New York: Basic Books, 1984, pp. 12, 18, 125-126, 173-174, 182.

⑤ 黄真“国际互惠的基本类型及其伦理取向——兼论利他主义的动力根源”,《教学与研究》,2010年第6期。

⑥ 参见《外交评论》2011年第4期的主题文章。

⑦ Robert Keohane, "Reciprocity in International Relations", *International Organization*, Vol. 40, No. 1, 1986, p. 21, 注释63. Stephen Krasner, eds., *International Regimes*, Ithaca: Cornell University Press, 1983, pp. 11-12.

⑧ [美]罗伯特·阿克塞尔罗德著,吴坚忠译《合作的进化》,上海人民出版社,1995年版,第89页。

⑨ Herbert Gintis "Strong Reciprocity and Human Sociality", *Journal of Theoretical Biology*, Vol. 206, No. 2, 2000, pp. 169-179. Samuel Bowles and Herbert Gintis "The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Population", *Theoretical Population Biology*, Vol. 65, No. 1, 2004, pp. 17-28.

义传统或民族性格,但不管出自何种原因,非理性关系和动机将使我们正视虽不算主流但仍十分重要的国际事实。<sup>①</sup>

### 3.2 政策意义

互惠中的给予、接受与回报至少对国际关系实践具有以下启示。第一,敢于给予。国家十分注重对他国动机和意图的判断。现实主义表面上说只关注能力而非意图,其实不然,现实主义其实最为关注意图,只不过考虑到意图的不确定性便退而求其次聚焦于能力。国际关系的实务操作最需要预防的一点便是预定他者的恶意,如此便没有了任何改善关系的可能。最为可行的方法是“试试看”,在自身能够承担的成本代价范围内,敢于给予,不要过于敏感对方的欺骗和剥削,尽可能昭示自身善意,寻求对方回报。第二,敢于接受。接受方不要过于夸大施惠方的虚情假意(如炫耀、羞辱、欺骗等),要敢于接受对方的好意,当然接受也有一定的成本和代价,这同样需要领导人审慎的分析方可做出决策。第三,敢于回报。对施惠方的合作行为,要敢于以合作对合作,甚至有时更多的回报也是相当必要的,但不可忘记,回报的资源在量和质上同样要在本国可以接受的成本和代价范围内。之所以鼓励国家主动行动,是为了把握住拉森(Deborah Welch Larson)所说的由于不信任而导致的太多的“失去的机会”。<sup>②</sup>当然,以上分析均是在国家能大体判断对方动机和意图非恶的前提之下,一旦意识到对方意图不轨(如修正主义动机),则互惠策略的政策意义便在于以威慑、惩罚对抗对方的欺骗或者对方某种“虚情假意”的恩惠。

不过,与大打折扣的理论意义一样,互惠策略的政策意义同样是有局限的,其中最大的弊病就在于对重复博弈的偏好,若沉醉于诸如交流、沟通、学习、理解等政策,结果可能陷国家于不利境地。不知为何,重复博弈这个本来中性的路径却基本上被认为是合作的天然路径,这是自由主义的影响所致。一般认为,国家如能保证重复的博弈便保证了持续的沟通、学习甚

至理解,由之不断增强的确定性、熟悉和充足信息必将生成“合作对合作”的互惠模式。这里存在两大误解:第一,重复博弈同样可以是重复性冲突,国家的关系既可以通过重复博弈得以实现进化,但也可能退化;第二,即使重复博弈确实增强了熟悉程度和确定性,但正如前文所言,确定性可能更为危险。“黔驴技穷”的成语深刻地说明,双方通过持续的博弈所获的熟悉是如何导致悲剧性的结果,所以,一定程度的不可知可能反而有利于维系“合作对合作”的互惠模式。

## 四、互惠策略的未来研究

针对互惠策略的已有成果和不足,未来的互惠策略研究应该集中在以下几个方面。

### 4.1 互惠策略的政策研究

互惠的政策研究是指互惠策略应该如何在实践中操作。由于互惠是一个给予、接受与回报的过程,因此这里蕴含三个小问题,国家怎么给予、怎么接受和怎么回报?毫无疑问,这三个小问题又是密切相关的。以“给予”为例,我们至少面临以下几个实际的政策操作问题:时间(timing)、包容度(degree of accommodation)、条件性(conditionality)和资源类型(type of gift)。<sup>③</sup>就时间而言,一国是一开始就给予以示好,还是在表示出对抗决心之后再给予以威慑对方不敢剥削,还是并非一开始就给予而是在对方示好之后才给予从而遵守严格的“一报还一报”策

<sup>①</sup> 不过,这里存在的一个问题是,强互惠表现出的某种单方面付出何以能称为“互惠”?互惠的根本性要义便是资源的大体对等的交换。恩斯特·费尔(Enst Fehr)指出,在没有外部补偿和激励的条件下,利他行为的自激励机制使行为主体从利他行为本身获得满足感,从而无需依赖有形的物质报偿。Fehr et al. "The Natural Basis of Altruistic Punishment", *Science*, Vol. 8, 2004, pp. 1254 - 1258.

<sup>②</sup> Deborah Welch Larson "Trust and Missed Opportunities in International Relations", *Political Psychology*, Vol. 18, No. 3, 1997, pp. 701 - 734.

<sup>③</sup> Russell J. Leng, *Interstate Crisis Behavior, 1816 - 1980: Realism Versus Reciprocity*, New York: Cambridge University Press, 1993, pp. 166 - 171.



略?就包容度而言,一国是单边给予而不附带任何强制性行动,还是胡萝卜加大棒,还是就对对方的某种强制性行为做出更轻而非对等的回应?就条件性而言,一国的给予是单边无条件的,还是以对方的回报为条件?如果是以对方的回报为条件,这回报是要求同等还是要求更多?就资源类型而言,一国是给予物质商品还是地位承认,是给予某种情感还是某种行动(含话语和实际行动)?另外,给予本身持续的时间和给予的质量也是相当重要的研究课题,这方面奥古德( GRIT)、埃兹奥尼( Progressive/ Gradualist GRIT)做了开创性的研究,歌德斯坦和弗里曼( Extended GRIT)作了补充性的研究,<sup>①</sup>概括起来无非是这么几个问题,给予要持续多长时间(即使在对方不即刻反应的情况下要坚持多长时间)?给予的资源成本要多大?给予是否要呈现出一定规律性(即给予是连续的还是零星分散的)?给予的质量是从小到大、从少到多、从象征到具体、从一般到珍贵还是相反?尽管以上学者作了一定的分析,但经验研究还较为不足。

对接受和回报的政策问题同样缺乏研究。<sup>②</sup>就“接受”而言,接受国是全盘接受,部分接受,还是完全拒绝?就“回报”而言,接受国是同等回报,超额回报还是低额回报,还是占有剥削“礼物”?不得不说,关于给予、接受和回报的所有政策研究都基于具体的情境,因此与国际互惠的政策研究紧密相关的一个研究课题是国际互惠的环境判断。环境很多时候不是外在客观的东西,它是行为体自己“创造”的。一般而言,一个互动的环境包含诸多变量,但其中三个尤为显著:(1)行为体对于利益结构(或博弈结构)的认知;(2)行为体的角色是现状的挑战者还是维护者;(3)每一方的互动策略。<sup>③</sup>所以,未来的国际互惠政策研究可以概括为:在大致掌握环境的情况下,国家应该怎么给予、怎么接受和怎么回报?

#### 4.2 多元互惠研究

由于在现实世界里,多国互惠、集体行动和

全球治理都是较为常见的现象,因而研究多国互惠就十分必要。

首先,尽管时至今日,N人博弈(大致等同于三人博弈)的理论和模型已经较为成熟,但它并没有有效捕捉到国际关系的基本特质。N人博弈理论倾向于把多人博弈简单转化为双人博弈。研究者认为,多个行为体通过联盟可以划归为两大阵营,每个阵营的行为体享有同一的偏好顺序和收益结构,一个阵营中的任何一个行为体的行动具有集体针对性(即对另一个阵营的任何一个行为体而言是一样的),因此多人博弈等同于经典囚徒困境,导致相互背叛(DD)和次优的收益。<sup>④</sup>但是,国际关系绝非如此简单。多国交往中,联盟现象并不普遍或清晰,很难说今天中印巴(基斯坦)、美俄欧、中朝美等三边关系等同于双边关系。另外,一国的行动并非如N人博弈理论宣称的那样对其他两国具有同样的效应。实际上,每一国的行动可能招致另两国不同的反应。再者,现实中A国对B国的行动可能出自C国本身,或出自C国之前对A国或B国的行动,或者出自A国之前对C国或B国之前对C国的行动。<sup>⑤</sup>总之,在诸如三国关系的互动中,存在着相互连接和相互影响的三对关系,每一组的互动及其结果受另一组或另两组互动及其结果的影响。多元互惠的高度复杂性和现实性要求我们进行扎实的理论和经验研究。

其次,多元互惠研究不仅包含行动上的复

<sup>①</sup> Charles Osgood, *An Alternative to War or Surrender*. Amitai Etzioni, *The Hard Way to Peace: A new Strategy*. Joshua Goldstein and John Freeman, *Three - Way Street: Strategic Reciprocity in World Politics*.

<sup>②</sup> 最近的文献来自 Tang Shipin, *A Theory Of Security For Our Time: Defensive Realism*, New York: Palgrave Macmillan, 2010, pp. 137 - 139. 但唐世平也只是提出了在接受和回报问题上存在的几种理论可能而没有经验研究。

<sup>③</sup> Russell J. Leng, *Interstate Crisis Behavior, 1816 - 1980: Realism Versus Reciprocity*, pp. 179 - 187.

<sup>④</sup> Anatol Rapoport, *N - Person Game Theory*, Ann Arbor: University of Michigan Press, 1970, p. 79.

<sup>⑤</sup> 典型的案例研究是冷战期间中美苏三角关系,见 Joshua Goldstein and John Freeman, *Three - Way Street: Strategic Reciprocity in World Politics*.



杂关系,也包含心理学上的复杂关系,已有的国际互惠心理学仅仅关注双边的心理互动,或者把集体成员的心理互动简化为双边互动,心理互动的三边性(triangularity)并没有受到重视。<sup>①</sup>现实世界里,一国很难对国际社会其他成员持有同一的认知和情感,而是怀有不同的态度,且这不同的态度往往又是相互依赖的。例如,在三边关系中,A国对B国是温和的(soft-line),而对C国却是强硬的(hard-line),而A国对B国的温和可能正是因为A国对C国是强硬的,反过来A国对C国的强硬也可能正是因为A国对B国是温和的。<sup>②</sup>同样,A国对B国的认同可能源自AB两国对C国的共有的“恨”,或者源自A国对C国的认同,而C国也认同B国(认同的连锁反应),或者A国认同C国,而B国也认同C国。

最后,多元互惠研究必须正视政策惯性(policy inertia)之于多边互动的影 响。政策惯性指的是A国对B的政策是A国对C国的一贯性政策(既可以是主动的政策也可是反应性政策)的复制,即A国把它的C国政策外溢到B国。国际关系中,出于各种原因,一国在一定时间内持有对他国的稳定的政策,这种政策惯性会影响到该国对第三国的政策取向。因而一个新的课题便是:政策惯性之于多边互惠的影响。

以上三点说明,未来互惠研究的重心应该聚焦于多元行为体的真正的战略性互动。

#### 4.3 互惠策略与关系的生产与再生产

前文已经表明,当前互惠策略基本基于国家的“经济人/理性人”假定。没有人否认“理性人/经济人”假定的强大功效,即使在当下,它也是极为有用和说服力极强的逻辑起点。但时至今日,在经济学内部已经越来越关注社会性因素(如正义、利他主义、情感等)并日益显现出“经济科学社会化”的时候,包括国际关系学在内的社会科学就要尽力避免深陷“社会科学经济化”的窠臼。其实,国家通过实践所形成的关系不仅仅只是功利性的,还有可能是情感性的、

道德性的。在不同的关系属性(relational property)中,国家具有不同的动机(功利性 utilitarian 和表达性 expressive<sup>③</sup>),甚至在同一个关系中,有情感也有算计,有自愿也有强迫,有为公益也有为私利,所以关系从来不是纯粹的,而是信任、情感、义务和算计等要素的混合,这些要素的强弱决定了关系的性质和类型。<sup>④</sup>科尔姆也说“互惠……由三种更为根本的要素——社会平衡义务(duty of social balance)、爱的相互作用(interaction of liking)、利益的相互性(mutuality of interest)——推动,而这三个要素本身又是由一系列更为基本的心理要素推动的。”互惠就是“建立在(行为体)心理动机相互联结(motivationally interrelated)基础上的礼物和恩惠的交换”。<sup>⑤</sup>

由此,我们要探究一些机制,这些机制能够较好地反映国家何以通过互惠实践和进程(如合作对合作,惩罚对背叛)实现关系和国家动机的再造。我们一旦发现了机制,就不仅更加明白合作何以进化或退化,也更有利于国家政策操作。

编辑 薛迎春 罗凤灵

<sup>①</sup> Deborah Welch Larson, “The Psychology of Reciprocity in International Relations”, *Negotiation Journal*, Vol. 4, No. 3, 1988, pp. 281-301. 尹继武著《社会认知与联盟信任形成》,上海人民出版社,2009年版。

<sup>②</sup> Glenn Snyder and Paul Diesing, *Conflict among Nations: Bargaining, Decision Making, and System Structure in International Crises*, Princeton: Princeton University Press, 1977, p. 298. Robert Jervis “Realism, Game Theory and Cooperation”, *World Politics*, Vol. 40, 1988, p. 326.

<sup>③</sup> 关于两种动机的划分,笔者参考了台湾著名的社会科学本土化运动领袖、社会心理学教授黄光国的研究。黄光国著《儒家关系主义:文化反思与典范重建》,北京大学出版社,2006年版,第3-30页。另外,功利性动机也叫做工具性动机,表达性动机也称为情感性动机,所以表达性和功利性等同于情感性(emotional)和工具性(instrumental)。

<sup>④</sup> [美]杨美惠著,赵旭东等译《礼物、关系学与国家——中国人际关系与主体性建构》,江苏人民出版社,2009年版。[美]阎云翔著,李放春等译《礼物的流动》,上海人民出版社,2000年版。

<sup>⑤</sup> Serge-Christophe Kolm, *Reciprocity: An Economics of Social Relations*, p. 11, 13.

## Study on the Strategy of Reciprocity and International Relations

HUANG Zhen<sup>1</sup>

(1. *Shanghai Administration Institute, Shanghai 200233, China*)

**Abstract:** Reciprocity with its institutionalization is considered to be able to make cooperation evolve by neo-liberalists. This paper is meant to be a reflection on this view. The theoretical significance lies in highlighting states' initiative, adaptation and the traits of "process constructing", making us prudently optimistic about the evolution of international relations. But this significance is quite limited, because reciprocity presumes the nature of state as rational and selfish, and ignores some reciprocities are operated by non-selfish motive and in non-selfish relational background. The practical significance is to encourage states to transcend the "Hobbes logic", and dare to give, accept and return. But this significance is also limited, because reciprocity leads the leaders and the people to have risky preference for the repeated game and certainty. The future researches of the strategy of reciprocity should focus on three aspects: The policy researches of reciprocity, namely how to give, accept and return; pluralistic reciprocity researches namely the multi-nation strategic reciprocity researches; and the reciprocity practice, reproduction of international relations and national motive, in order to make clear the mechanisms between reciprocity strategy and the evolution and degradation of cooperation.

**Key words:** strategy of reciprocity; contingency; dilemma; future study