

社会困境问题的理论架构与实验研究*

刘长江 郝 芳

(南京师范大学心理学院, 南京 210097)

摘 要 社会困境是指个体理性导致集体非理性这样的情境。社会困境研究的推进要求一个新的框架来理解已有理论, 这一框架将这些理论区分为宏观、中观和微观三个层面。实验研究依据这些不同层面的理论, 以实验游戏为工具来探讨人们在面临社会困境时做出选择的心理机制及其影响因素, 从而对社会困境问题发展出相应的问题解决路径。理论的发展、实验任务本身的改进以及研究的外在效度预示了通过实验方法来研究社会困境问题有着广阔的前景。

关键词 决策; 社会困境; 合作; 实验研究

分类号 B849:C91

人类的很多决策发生在自己决策结果与他人决策结果相互依赖的情境中。人口控制、公共物品的供给、社区公园、小组作业等都体现出个人的决定既影响自己的结果, 也影响到他人的结果。这样的互依情境具有一个共同的特征, 即个人由于追求即时的自我利益(即个人理性的结果), 进而损害了他们所在群体的利益(即集体理性的结果)。这便是“社会困境(social dilemmas)” (Dawes, 1980)。

社会困境问题引起了经济学家、社会学家、心理学家、政治学家、生物学家、数学家以及人类学家等的广泛关注(Messick & Brewer, 1983; Van Lange, Joireman, Parks, & Van Dijk, 2013)。时至今日, 社会困境依然是个严峻的问题。当前, 资源短缺、环境污染、道德水平下降、人际信任滑坡等问题严重困扰着社会和谐与发展。本文阐述了社会困境的本质及其表现形式, 提出从 3 个层面来概括相关理论。本文还总结了实验社会科学在开展社会困境研究中的共性与差异, 指出存在的问题, 探讨未来的研究方向。

1 社会困境的实质与分类

社会困境是指个体理性导致集体非理性这样

的情境(Kollock, 1998)。社会困境通常涉及两人或更多人, 它具有两个属性(Dawes, 1980): 1) 不管社会中其他个体怎么做, 相比于做出社会性合作选择, 每个个体做出社会性背叛选择要获得更高的收益; 2) 相比于所有人都合作的情况, 如果所有人都背叛, 那么所有个体得到更低的收益。社会困境中的合作选择是指个体做出有利于集体或他人利益的决策, 而背叛选择是指做出有利于自己个人的决策。比如选择为所在小区公共服务提供捐助是合作选择; 从公共资源中索取更多资源便是背叛选择。

总体来讲, 社会困境分为社会陷阱和社会藩篱这两类(Messick & Brewer, 1983) (表 1)。所谓社会陷阱(social traps)是指奖赏与惩罚的特定组合, 在这种组合中, 那些短期内对个体是奖赏的行为, 长期而言, 无论是对于个体还是其他人则都意味着惩罚。也就是说, 人们为了眼前的好处, 掉进了从长远角度来看需要付出代价的陷阱中。资源困境或索取困境属于社会陷阱。

社会藩篱(social fence)则是一种反向陷阱, 某种行为从短期角度来看会产生令人厌恶的结果, 因而个体可能不从事这样的行为, 但是, 假如个体这样行动, 那么从长远角度来看对自己与他人都会产生积极的收益。也就是说, 即时的个人负性行为阻止人们提供长期的正性行为。公共物品困境或给予困境属于社会藩篱。

收稿日期: 2013-12-13

* 南京师范大学百名青年领军人才培养计划项目资助。

通讯作者: 刘长江, E-mail: chjliu@njnu.edu.cn

表1 社会困境的分类

	集体结果	
	即时或短期	延迟或长期
社会陷阱 索取困境或资源困境	企业盲目追求利益导致每日空气污染加重。	如果每个家庭都生育更多子女,社会和自然资源最终将无法满足人口膨胀的需要。
社会藩篱 给予困境或公共物品困境	来自于社会的自愿捐助挽救一个垂危的生命。	如果社会中每个公民积极选择参与到公益事业中,那么长期以后,社会道德水平终将提高。

参照: Messick & Brewer, 1983; Van Lange et al., 2013

2 社会困境的理论

2.1 宏观-中观-微观框架

我们用宏观、中观和微观三个层面来概括社会困境研究中所涉及到的理论。这一区分是基于理论出发点所欲解释的对象的组织层次(即单个-总体, Liljenström & Svedin, 2005)及行为解释的距离差异(即近端-远端, Tinbergen, 1968)。图1描绘了宏观、中观和微观三个层面在由对象的组织层次和行为解释的距离所组成的二维空间中的关系。理论上,社会困境研究关注决策行为的机制或者动机及其影响因素(Messick & Brewer, 1983)。这样,微观层面的理论更多关注个体实际做出的行为选择,研究个体行为的近端机制,强调个体内在的动机因素;而宏观理论关注个体行为总和的结果,通常强调人类或普通人行为的远端机制(包括策略、进化机制等)。中观层面则在行为解释的距离方面更接近于近端机制,而在对象的组织层次方面通常用来解释来自于总体

中的群体,或者关注于特定群体的行为特征。情境在中观层面上尤为重要,诸如不同的群体构成、互动特征等不仅提供了行为决策的特定情境,而且也强调了人们决策的交互过程。更为凸显且重要的是,中观层面强调情境认知在决策行为中的作用(Halevy & Katz, 2013)。

更多的理论或观点是微观层面的个体动机研究以及中观层面的特定情境下的过程研究,而宏观的理论包括博弈论与进化论等为实际的选择行为提供了规范性指导,经常成为社会困境中行为的标准或参照,而且宏观层面的理论也往往落脚于更低层次的理论来得以支持和解释。总的来说,微观之和就是宏观,而宏观分解则为微观。中观则介于宏观与微观两者之间。中观层面关注情境因素和结构性机制,以及宏观与微观两个层面之间的交互。不同层面的分析水平是相互补充的,而非相互排斥的,因为任何现象的完整解释要求在每个层面上都有所回答(Barclay, 2012)。

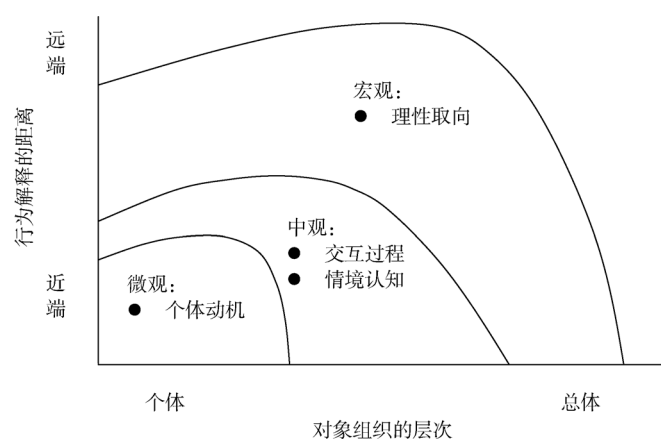


图1 社会困境理论的宏观-中观-微观框架

2.2 宏观理论

进化论和博弈论是分析人类或一般意义上人在社会困境中理性行为的典范, 两者均将理性取向作为人在社会困境中行为决策解释的基础, 指出人应该如何行为使得自己利益最大化。进化论提出自然选择是生物进化的动力, 从而使得那些强者更能够适应环境并且生存下来, 强调人类的利己动机在其进化过程中的源动力; 而博弈论则着重分析最优的策略结构, 认为经济理性驱动人的行为, 因此, 也可以说博弈论提供了一种供物种(如人类)适应的最佳环境。博弈论和进化论已广泛应用于社会科学中, 并且, 研究者结合博弈论来解释生物进化从而为进化博弈论(Evolutionary Game Theory)奠定基础(Vincent & Brown, 2005)。

2.2.1 博弈论

博弈论是关于有理性但彼此互不信任的人之间的冲突的科学, 它通过数学形式来研究游戏中个体的行为及优化策略。博弈论源于应用数学, 但是它广泛用于分析几乎所有社会科学里的基本问题, 尤其是在对个体而言有着不同目标或偏好的经济、政治或社会情境中(Myerson, 1997)。因而, 博弈论能够对社会情境中行为进行建模, 从而对人们最优的行为选择做出预测。

博弈论通过建构游戏来考察理性人的选择策略。这样的游戏通常要求几名参与者(称为玩家)一起参与, 他们做出一些行为或策略, 并因此而得到不同的酬劳(payoffs)。为了简约起见, 博弈论会分析最简单的游戏形式。囚徒困境(prisoner's dilemma)就是博弈论中最具代表性的例子, 它反映了个人的理性选择并非集体的理性选择(表2呈现一个一般形式的囚徒困境矩阵结构)。研究已经提出并考察了不同的选择策略。例如, 人们会使用针锋相对策略(tit-for-tat)、宽容的针锋相对策略(generous-tit-for-tat)、赢定输移策略(win-stay, lose-shift)等(见 Nowak, 2006)。

博弈论不仅阐释人们的选择策略, 也能够更低层面上解释微观理论所关注的动机因素。例如, 在囚徒困境中, 合作动机的基础在于两个成分, 即“担心”和“贪婪”(Komorita & Parks, 1995)。“担心”表示人们避免自己的利益被那些不合作者所剥夺, 它说明的问题是, “假设对方背叛, 我该如何反应?”例如, 那些不为集体做贡献的人是

因为他们相信没有人会做出贡献。贪婪表示诱惑人们搭他人合作这一便车的各种刺激。贪婪要回答这样的问题: “假设对方合作, 我该如何反应?”这样就为搭便车的存在提供了合理的理由。在社会困境中, 担心成分用相互背叛的收益和单方合作的收益之差表示, 而贪婪成分用数学表达为单方背叛的收益和相互合作的收益之差。就表2中的例子而言, 担心收益为1, 贪婪收益为2, 因而, 个体更多受贪婪所诱惑。

表2 一般形式囚徒困境的矩阵

玩家甲	玩家乙	
	合作	背叛
合作	(R, R) (3, 3)	(S, T) (0, 5)
背叛	(T, S) (5, 0)	(P, P) (1, 1)

注: 单元格内包含符号的括弧构成一般形式的囚徒困境, 包含数字的括弧为一符合囚徒困境本质特征的实例。其中, R 表示相互合作的收益; S 表示单方合作的收益; T 表示单方背叛的收益; P 表示相互背叛的收益。

2.2.2 进化论

进化论是当代生物学的主导思想, 不仅人类生物特征一代一代地遗传进化, 人类的社会行为同样具有进化基础。在人类社会早期, 人类很多关键活动构成一种公共物品, 比如一起狩猎、维持稀有资源等。自然选择使得个体偏爱合作, 即使参与这样的活动需要个人付出成本, 因为这样做有助于提高群体成员生存和繁衍的机率。例如, 那些基因上相近的个体之间更愿意选择合作, 这样自己的基因可以得到更好的繁衍(Hamilton, 1963)。

然而, 人类(也包括其他物种)的合作不仅发生在基因相似的个体身上, 还发生在没有基因关系的个体之间。假如自己与他人相遇的概率超出了利他行为的成本-收益比例, 那么自然选择通过直接互惠机制促进合作, 即人们与那些直接给予自己帮助的人合作(Nowak, 2006)。的确, 基于重复性囚徒困境的研究发现, 针锋相对策略是最好的合作策略(Axelrod, 1984)。在社会中, 人们为了获得名声而帮助那些与自己基因无关且不会得到直接回报的他人, 这一间接互惠机制也能有效地促进合作(Nowak & Sigmund, 2005)。不过, 即使在基因无关的人群中、单次互动中以及没有声

誉情况下合作依然会产生,这是因为利他惩罚机制在发挥作用:尽管惩罚搭便车者对个人是付出成本的,但是如果足够多的人执行利他惩罚,或者通过惩罚他人能够得到足够的回报,那么合作就会维持下去(Fehr & Gächter, 2002)。可见,基于进化论的不同机制能够为广泛的社会情境中人的行为做出合理解释。

2.3 中观理论

在社会困境研究中,能够体现中观层面、典型的理论是社会互依理论和适当性模型。适当性模型的提出是基于作者对社会困境研究中决策行为的影响因素的总结与提炼,而社会互依理论则根植于博弈论,强调群体中人际依赖与互动。两个理论的共同点在于两者均关注决策的过程,强调情境因素,以及决策者对情境因素的主观解释。因此,中观理论论述的是,决策者的决策结果取决于个体本身及其与情境之间的互动。因而,中观理论构成了实验社会科学开展社会困境研究的结构框架。理论之间不同层次的交互渗透在中观层面上尤为突出。对于微观理论而言,这些中观理论整合诸多不同方面的动机理论或观点于一个结构框架中;而对于宏观理论而言,中观理论可以看作是向上一级层次的理论整合的基础,也是宏观理论进行实证检验的过渡。

2.3.1 社会互依理论

社会互依理论(Social Interdependence Theory)(Rusbult & Van Lange, 1996)假设,互依情形取决于三个要素,即一种互依的结构(如囚徒困境)、彼此影响的各方(如玩家A和B)以及互动的动态性(如使用针锋相对策略)。在互依情境中,人们不止关注自己,还关注互依情境中的他人,而且,人们的这种对自己的结果和他人结果的相对偏好具有稳定性。例如,亲社会取向的人偏好最大化自己和他人的结果;个人取向的人偏好最大化自己的结果而不在于他人结果怎样;竞争取向的人偏好拉大自己与他人结果之间的差异等(Messick &

McClintock, 1968)。在互动过程中,决策者会对互依情境做出主观解释,也就是决策者对自己和他人结果的偏好进行某种“结果转换”。这样,当决策者面临一个用一种结构或矩阵来表示的客观结果时,他们会根据内在的规范将这一给定的结构转换成一个主观的、有效的矩阵结果,正是这一有效矩阵影响了他们的行为。

我们假定在表2中玩家甲是一名亲社会取向的个体。亲社会者要最大化自己与同伴的利益,这样,每个单元格里首先计算双方的结果之和;同时他们尽可能结果均等,这样,每个单元格里再减去最初单元格数值之差(Van Lange, 1999)。经过这样的转换,亲社会者甲将最初的囚徒困境矩阵转换成了一个有效矩阵(图2):如果甲乙均合作,那么甲收益为6;如果甲合作而乙背叛或者如果甲背叛而乙合作,那么甲收益为0;如果甲乙均背叛,那么甲收益为2。因而,如果乙合作,那么甲最好是合作,因为这是收益是6;如果乙背叛,那么甲背叛的收益(2)要比他合作的收益更大(0)。可见,亲社会者未必总是一贯地表现出合作,当对方出现背叛时,亲社会者也将表现出背叛行为。

2.3.2 适当性模型

在社会困境领域里,同样关注情境因素及其主观解释的理论还有适当性模型(Appropriateness Framework)(Weber, Kopelman, & Messick, 2004)。根据这一模型,人们在做决策时需要回答这样一个基本问题,“像我这样的人在这样的情境中会做什么呢?”概括来说,这一模型认为决策取决于三个基本要素,即决策者对情境的定义(或认知)、决策者的身份以及决策规则的应用。人们通过认知过程(或原型分类)界定了他们所面临的情境,也就是将情境的特征与他们所理解的其他情境加以匹配。因此,一种情境越是与既有的事件原型相匹配(越典型),那么人们越有可能相信所面临的情境也具有这一事件原型所具有的特征。另一

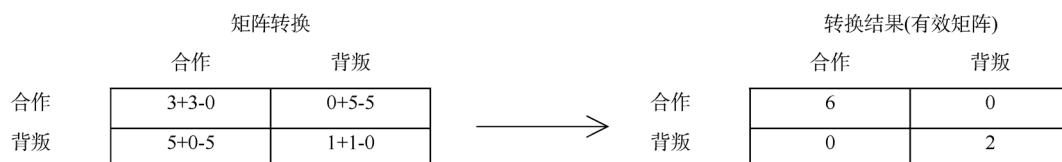


图2 亲社会者甲的结果转换过程

个重要的问题是, 决策者预期他人对这一情境的认识, 即这一情境的规范是什么, 其他人期望我在这种情境下做什么。

适当性模型强调不同的个体在知觉情境过程以及对情境的反应可能不同。该模型用身份一词来泛指个体的所有气质性因素, 它是一个复杂的、多层面的、重要的调节变量, 包括人格因素、社会价值、个人历史和体验等等。此外, 适当性模型接受“人是有限理性”这一假设。有限理性的人 would 采用各种规则来应对他们所面临的刺激, 减少选项, 做出行为选择。人们不仅会使用外在的、成型的规则来指导行为, 也可能应用一些更内在的社会启发式与习惯方式。人们使用何种规则通常取决于他们所知觉到的社会规范。

2.4 微观理论

微观层面的理论从个体行为的动机角度来解释人们为什么会做出有利于他人和/或集体的决策。社会心理学认为人的动机是多样化的, 人不仅有利己动机, 还关注他人利益以及集体利益。在社会困境中, 一些行为源于个体内在的动机, 一些则源于与他人之间的社会比较过程。由于社会困境本身是涉及多人所构成的群体, 因而, 群体特征也可能构成一种动机从而促进个体的合作决策。这些不同类型的动机因应了复杂的任务特征和情境因素, 也使得个体的行为不再符合宏观理论的理性预期, 事实上, 多样化的动机理论在一定程度上促进宏观理论的发展。

2.4.1 自我服务: 个体的内在动机

从个体层面来讲, 个体在面临社会困境时的心理表现与行动策略可能受到自我服务(self-serving)动机所驱动。对自我服务动机的探讨不仅有着很长的历史, 而且在归因、社会判断等领域里也得到广泛证实。目标-期望理论(Goal-Expectation Theory) (Pruitt & Kimmel, 1977)提出特定情境下的合作依赖于两个条件, 即人们必须具有合作性目标和人们必须期望群体中的其他成员也会合作。根据社会困境的本质特征, 对个人最有利的结果是个人背叛而他人合作, 其次是所有成员都合作。如果个人合作, 同时(预期到)他人合作, 那么他可能赢得物质上的利益, 还可能赢得社会声誉, 因而, 这种合作的动机本质上是利己的。

解释自我服务动机的另一个有影响的理论是

前景理论(Prospect Theory) (Kahneman & Tversky, 1979)。该理论指出, 人们的风险知觉和偏好在收益领域和损失领域里是不同的。在收益领域里个体是风险规避, 即个体偏好确定量, 而不偏好风险或不确定量; 而在损失领域里却表现为风险寻求。在公共物品困境中, 来自人们的捐献可能会实现公共物品, 而这会使得群体中每个人收益更多。这种潜在的收益使得人们规避风险, 这样人们更不愿意做出捐献。相反, 在资源困境中, 公共资源业已存在, 但有可能被消耗殆尽。这种潜在损失使得人们更倾向于寻求风险, 这样人们不愿意从公共资源中索取更多, 以期保留或维持公共资源的存在。研究证实, 在同等成本和收益条件下, 资源困境要比公共物品困境可能更能促进合作(Son & Sell, 1997)。不过, 无论是个体寻求风险还是规避风险, 其行为的最终目的是自我服务取向的(如最大化自己的利益, 尽管结果未必这样)。

2.4.2 社会比较: 自我-他人利益权衡

基于相似他人的社会比较是社会互动的基本现象(Festinger, 1954), 不对称社会困境为社会比较提供丰富的情境因素。不过, 由于社会困境问题本身强调个人利益与他人利益之间的权衡, 因而已有的理论更多关注于个体互动后的结果差异, 即人们不仅关注自己的结果, 也关注他人的结果以及两者结果的关系。例如, 根据社会价值取向研究, 亲自我者的动机就是要最大化个人结果, 无论是以相对的方式(竞争者)还是绝对的方式(个人主义者); 然而, 亲社会者则偏好最大化彼此共同的结果, 最小化彼此结果的差异或者是仅仅为了均等考虑(Eek & Gärling, 2008; Van Lange, 1999)。

贪婪-效率-公平假设(Greed-Efficiency-Fairness) (Wilke, 1991)也关注社会比较过程。该假设提出三个影响人们在社会困境中合作的动机因素, 即贪婪、效率和公平。贪婪是指最大化自己的结果, 人们在社会困境中都尽可能为自己获得最多的资源。贪婪动机还受到效率和公平两种动机的约束。效率是指人们期望有效地保留资源。公平是指人们期望实现公平。前两个动机可以认为是更多地关注自我服务, 而公平动机则关注个人结果与他人结果之间的差异, 它通常指所有参与者最终获得的结果要平等。这样, 如果群体资源获取量不平衡, 即便是能有效率地利用资源,

人们也会不满意,并且寻求某种改变。

2.4.3 共同命运:个体与群体之间的联结

群体成员之间的互依是社会困境的一个重要特征。在社会困境中,不同的成员共同完成贡献或限制自我利益,但是个人的结果不仅依赖于个人的决定,也依赖于群体其他成员的选择。个体追求自我利益及权衡个人利益与他人利益均能够激发个体在社会困境中的合作行为。同时,从社会性角度来看,个体与群体之间联结也可能激发其合作行为。个体如果能将自己看成是社会性成员,或者从社会性来定义自我,那么合作水平会提高(De Cremer & Tyler, 2005)。

我们用格式塔心理学创始人威特海默提出的共同命运(common fate)(引自 Campbell, 1958)来概括上述效应的动机机制。根据共同命运原理,朝相同方向运动的物体更可能被感知为属于一组或者彼此相关。在社会互动中,如果某个个体与其他个体都朝集体目标努力,那么他们更可能是一类人群(如亲社会者);如果个体对自己在社会生活中所担任的角色融入到自我概念(如社会自我)中,那么他们更可能表现出与集体目标一致的行为。社会身份理论指出如果人们感到自己归属于这一社会类群,那么他们会根据这一类群的定义特征来为自己界定了他们的身份,正是这种身份动机促进了个体朝向集体目标努力(Hogg, Terry, & White, 1995)。此外,对社会结构因素(例如群体规模、权力、地位、领导等)、信任、归属感等在社会困境中的作用探讨也能够支持这种共同命运原理。

3 社会困境的实验研究

在过去半个世纪的研究中,实验社会科学家广泛地考察社会困境中个体选择行为的机制与影响因素。近期,一本著作《社会困境:理解人类合作》(Van Lange, Balliet, Parks, & Van Vugt, 2014)的问世,标志着社会困境问题已经得到系统而全面的认识。在本部分,再总结这些相关研究已是多余。理论对于社会困境研究是至关重要的,前面已经展现了不同的理论取向及其要素。下面,我们概括实验研究在其他方面所表现出共同之处以及差异。

3.1 研究问题

尽管社会困境有着不尽相同的表述形式,但

是其本质含义已经得到一致认可。实验室研究是社会科学研究社会困境最常见的研究手段,而且研究问题通常包含两种类型:第一,研究基本的心理过程;第二,研究影响行为的结构或环境因素(Messick & Brewer, 1983)。社会困境问题研究也不脱离这两个基本类型。更细致的研究是考察情境与个体因素之间的交互影响,即在特定情境中个体行为的心理机制。因而,社会困境研究在研究问题的本质上具有共同性。但是,在解决问题的过程中相关研究表现出一定的差异,这具体表现为下面两点。

第一,由于理论假设的不同,同样的问题可能产生不同的解决路径。例如,社会价值取向对社会困境中合作选择产生影响(Balliet, Parks, & Joireman, 2009)。如前所述,社会互依理论强调不同类型价值取向对结果偏好的有效转换过程(Rusbult & Van Lange, 1996)。目标-期望理论(Pruitt & Kimmel, 1977)则与情境线索交互影响了个体的合作选择,也就是说,合作意愿(个体要具备合作愿望)是在外在的激励因素的作用下促进合作的实现,而对他人的预期(个体预期到他人行为)需要建立在信任的基础上(Bogaert, Boone, & Declerck, 2008)。

第二,同一现象可以依据不同的理论,而基于这些不同理论所开展的研究彼此能够推动我们对现象本质的认识。例如,研究发现,在人们经历过大自然的威胁(如地震)后,随着居住地灾情越来越严重,所在地居民的亲社会行为水平平均随之增加(Rao et al., 2011)。基于进化论的观点,这些研究者认为,自然灾害使人们处于劣势情境,通过个体间(包括无血缘关系,甚至是敌对的双方)相互帮助与合作,个体生存的概率会提高,基因得到更好的遗传。内隐社会认知的观点认为,与劣势有关的信息与亲社会性评价之间的联系是自动化加工的(乐国安,李文姣, 2010)。然而,从个体差异的角度,只有亲社会者在处于劣势时更可能表现出合作行为(刘长江,郝芳, 2011)。

3.2 研究工具

在社会冲突和人际关系研究中,研究者以实验游戏作为主要的研究工具。事实上,社会困境问题的实验室研究在很大程度上依赖于也得益于这些实验游戏的开发。在实验游戏中,每个参与者需要做出一些影响自身及他人收益的决定;决

策的结果是用数字形式来表示,而且这些数字是由实验者事先设计好的(Pruitt & Kimmel, 1977)。

实验游戏能够抓住社会困境的本质特征,然而,其表现形式却是多样化的。研究者通过任务参数来设定个人利益与他人或集体利益之间的冲突形式,从而使得实验中的社会困境与现实中的社会困境更为贴近。最常见的一些游戏形式包括:改变困境中涉及到的行为者的数量可以设定游戏为两人困境或多人困境;根据游戏之初每个行动者所拥有的资源或禀赋是否相同确定游戏为对称游戏还是非对称性游戏;根据行动者互动的次数可以确定重复与非重复困境等。再如,嵌套社会困境(nested social dilemmas)模拟了同属于一个更高位群体的两个亚群体之间的社会困境问题,这反映了现实中人通常具有多重群体身份这一特征(Wit & Kerr, 2002)。

3.3 研究目的

科学就是要寻求真理、解释常识,是要进步与创新,还要能够将知识应用到实践(Van Lange, 2013)。尽管现实生活中象囚徒困境这样纯粹的困境游戏很少遇见,但是这一游戏所揭示的社会现象却普遍存在,而且具有说明性。因而,通过社会困境研究可以揭示出个体在面临个人利益与他人利益冲突情境时的策略选择,更深一步讲是探讨人性的本质,促进人类利他行为和社会发展。同时,社会困境研究绝非仅仅只是停留在对经验和现象的科学解释。从前面所提及的多篇综述文献来看,研究者都期望能够为解决现实社会困境问题提供思路,这是众多研究所追求的目标之一。

然而,由于理论假设各不相同,实验研究的最终目的在实践指导的层面上有所差异。从宏观层面上讲,理性假设是稳健的、整体性的。例如,从进化的角度来看,理性的决策习惯才被完整保留下来。在解决诸如国家政策制定、集体行为管理等方面的社会困境问题时,理性假设提供了一个准则,显现出很强的效度。但是在具体问题处理上,人性化理念则需要关注个体本身及其与情境的交互,做到有差别对待。因而,研究在追求这一终极目标过程中表现出的差异更多体现它们所面临问题的层面是不同的。可以说,对于社会困境问题的解决,“我们既需要自上而下的规范性方法,也需要自下而上的描述性方法”(Hastie & Dawes, 2010/2013, p.21)。

4 总结与展望

社会困境研究已经经历了近半个世纪的发展,然而,如今相应的社会问题依然凸显和严峻,这足以说明社会困境研究的重要性。在研究层面上,社会困境依然是社会科学乃至自然科学的一个热门课题。

4.1 不同层次的理论交融促进理论的完善与发展

我们总结了社会困境研究中经常被引用的理论,并且提出从微观、中观和宏观三个层面来加以划分。尽管这些具体的理论观点在社会困境研究中并不新鲜,但是这一划分是必要的,因为这让我们清晰地看到具体理论的科学问题解决能力及其实践价值。宏观层面的理论在确定关键要素以及整体指导方面起着重要的作用;中观层面强调情境及其主观解释的重要性,因而对于解决社会困境问题的系统设计起重要作用;而微观层面更强调个人内在的心理动机与过程,因而在解决社会困境问题过程中有助于认识如何有效地调动个体积极性。例如,博弈论和进化论经过分析(如计算、模拟等)指出惩罚在社会困境中的重要性,但是诸如人们行为的原因,如何实施惩罚,惩罚的效果及其反作用等问题则需要从中观层面和微观层面提供更为细致的执行策略。

尽管社会困境理论尚未有更大突破,而且即使在同一层面中不同观点之间的关系也鲜有比较和论述,但是不同理论之间相互碰撞促进了理论本身的发展。近期有关社会投射(social projection)的工作就是一个很好的例子(Krueger, DiDonato, & Freestone, 2012)。这一理论不再视理性假设为唯一假设,因而不单是关注于宏观层面,相反,它着眼于微观层面,认识到社会投射(即相信别人像自己一样行为)是合作的一个重要条件,并将其作为假设前提,建构行为模型。只是,即使做出修改,新的理论解释也是有一定的适用范围的(例如不清楚是否适合多人困境)。

在中观层面上,社会科学假设情境对人行为的影响至关重要,其目标之一就是要发展出完善的情境理论(Reis, 2008)。这一理论的一个重要基础是对情境的主观解释(Rusbult & Van Lange, 1996; Weber et al., 2004)。近期,研究者从微观层面上提出用冲突模板(conflict templates)来表示对结果互依情境(包含社会困境情境)的主观知觉,

并且证实了不同的冲突模板能够预测冲突行为、社会关系以及社会政治态度等(Halevy & Katz, 2013)。这样,未来研究需要在较低理论层面上阐释这种认知因素(如表征、理解等)在社会困境中的作用机制,在较高理论层面上解释认知因素的理性与进化基础。

4.2 分析实验任务本身有助于丰富社会困境研究

尽管不同类型的实验游戏都声称反映社会困境的本质特征,然而,人们在不同类型的社会困境中的心理与行为仍可能存在差异。例如,人们在具有同等成本和收益条件的资源困境和公共物品困境中有着不同的合作水平(Sell & Son, 1997)。即使是同一种类型的社会困境,不同形式的困境仍有可能存在差异。例如,跨阶层公共物品(step-level public good)与连续型公共物品(continuous public good)在博弈结构和社会互依方面都存在差异(Abele, Stasser, & Chartier, 2010)。犹如社会困境本身就体现了混合动机一样,人们在不同社会困境中的心理过程也可能是复杂的。例如,人们知觉资源困境和公共物品困境这两者的竞争程度是不同的(Apesteguia & Maier-Rigaud, 2006)。研究发现,当人们直觉状态下做决策时,他们在给予困境中要比在具有相同损益矩阵的索取困境中花费更多时间来思考,体验更强的心理冲突(Liu & Hao, 2011)。

我们认为,用数字形式来表示任何价值、信念等可能简化对事物本质的认识。例如,在研究中一个很少被提及的问题是,公共资源或公共物品本身的差异。实验室研究用数量化形式来表达公共资源或公共物品,泛化地表达了小组作业、自然资源(如牧场、水资源等),甚至道德、民主等。不同于个人物品,这些公共物品和公共资源具有非排他性,而且,贡献(或者限制索取)与个人获取之间不存在直接的关系。比如,公共资源私有化是解决社会困境的一种策略(Van Vugt, 1998),但是,不同资源甚至是物品私有化的形式以及对群体成员所产生的心理效应是否一致,这种差异是否影响随后的行为,尚未见探讨。

因而,未来研究需要在社会困境这一大主题下挖掘不同情境下数字的特殊表达形式及其心理意义。例如,在考察领导在社会困境中的作用时,研究者不再单纯强调领导是外在的安排,还体现了领导的功能,比如领导者也需要有利己的一面,

担当领导者是需要付出成本,甚至不能通过为集体贡献来做出弥补(章哲明,金盛华,吴嵩,周翔, 2013; Arbak & Villeval, 2013; Van der Heijden & Moxnes, 2013)。这样,用数字的形式(包括得与失)来表达更为现实的领导功能,便能更有效地揭示出领导在社会困境中的作用机制。

4.3 基于现实的理论整合有助于提高研究的外在效度

研究中的实验任务抓住社会困境的本质特征(见前面对社会困境属性的说明),这样能更好地理解社会困境中决策行为的基本过程与内在机制(Messick & Brewer, 1983)。然而社会现实是复杂的,社会困境问题的解决脱离不开现实特征。因而,在不同层次理论的指导下,一些因素融入到具体的理论中,从而有助于提高研究的外在效度。尽管实验社会科学已经广泛地探讨社会困境问题,但是很多因素诸如社会规范、角色原型、奖赏与惩罚、互惠、公平、情绪、解析过程、表情信息、群体间关系等在社会困境中的作用依然没有得到解决(Guala, 2012; Van Lange et al., 2013; Weber et al., 2004)。我们注意到,资源不对称性、社会困境的时间性、噪音情境、社会角色和功能的内置化等思想也逐渐进入研究者的视野,这些更接近现实社会困境的考虑将成为未来实验研究的重要趋向。

从实践的角度来看,上述这些考虑源于现实社会中可能影响公共物品供给和公共资源索取的因素。社会科学家将这些因素概括为两大类,一是改变个人行为(即个体策略);二是对群体进行有组织的集体行为(即集体策略或结构性策略)(Van Vugt, 1998)。可以肯定的是,集体策略要比个体策略在解决社会困境问题方面效果更强。例如,通过惩罚机制完全可以抵制一些司机私自停车。然而,社会科学研究还提示我们,个体策略在解决社会困境中的作用是不容忽视的。集体策略或结构性策略“依赖于群体成员对那些决定他们社会互动模式的规则或结构的选择……[这些]变化能够对群体成员的期望和对他人的行为有着间接的影响,也对收益结构有着直接的影响。”(Samuelson & Messick, 1995)。从理论区分的视角来看,不同层面的理论在社会困境问题解决的思想、组织和行动上互为补充,这既能拓展这种基于因素的解决策略,也能为寻求新的解决策略提

供思路。

致谢: 感谢外审专家和编委会专家对本文提出批判性以及建设性意见; 感谢 Brandon Bufe 校对英文摘要, 并启迪改进摘要; 感谢编辑部细致入微的工作。

参考文献

- 海斯蒂, R., 道斯, R. M. (2013). *不确定世界的理性选择——判断与决策心理学* (谢晓非, 李纾 译). 北京: 人民邮电出版社.
- 刘长江, 郝芳. (2011). 不对称社会困境中社会价值取向对合作的影响. *心理学报*, 43, 432–441.
- 乐国安, 李文姣. (2010). 弱势引发亲社会行为——来自贫困大学生的实证研究. *南开学报: 哲学社会科学版*, 60, 63–68.
- 章哲明, 金盛华, 吴嵩, 周翔. (2013). 公共资源困境中领导对群体成员合作行为的影响. *心理学报*, 45, 453–465.
- Abele, S., Stasser, G., & Chartier, C. (2010). Conflict and coordination in the provision of public goods: A conceptual analysis of continuous and step-level games. *Personality and Social Psychology Review*, 14, 385–401.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Apesteguia, J., & Maier-Rigaud, F. P. (2006). The role of rivalry public goods versus common-pool resources. *Journal of Conflict Resolution*, 50, 646–663.
- Arbak, E., & Villevall, M.-C. (2013). Endogenous leadership: Selection and influence. *Social Choice and Welfare*, 40, 635–662.
- Balliet, D., Parks, C. D., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes and Intergroup Relations*, 12, 533–547.
- Barclay, P. (2012). Proximate and ultimate causes of punishment and strong reciprocity. *Behavioral and Brain Sciences*, 35, 16–17.
- Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology*, 47, 453–480.
- Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, 3, 14–25.
- Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, 31, 169–193.
- De Cremer, D., & Tyler, T. R. (2005). Managing group behavior: The interplay between procedural justice, sense of self, and cooperation. *Advances in Experimental Social Psychology*, 37, 151–218.
- Eek, D., & Gärling, T. (2008). A new look at the theory of social value orientations: Prosocials neither maximize joint outcome nor minimize outcome differences but prefer equal outcomes. In A. Biel, D. Eek, T. Gärling, & M. Gustafsson (Eds.), *New issues and paradigms in research on social dilemmas* (pp. 10–26). New York: Springer.
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415, 137–140.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.
- Guala, F. (2012). Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, 35, 1–59.
- Halevy, N., & Katz, J. J. (2013). Conflict templates thinking through interdependence. *Current Directions in Psychological Science*, 22, 217–224.
- Hamilton, W. D. (1963). The evolution of altruistic behavior. *American Naturalist*, 97, 354–356.
- Hastie, R., & Dawes, R. M. (2013). *Rational choice in an uncertain world: The psychology of judgment and decision making* (X. F. Xie, & S. Li, Trans.). USA: Sage. (Original work published 2010).
- Hogg, M. A., Terry, D. J., & White, K. M. (1995). A tale of two theories: A critical comparison of identity theory with social identity theory. *Social Psychology Quarterly*, 58, 255–269.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.
- Komorita, S. S., & Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology*, 46, 183–207.
- Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology*, 24, 183–214.
- Krueger, J. I., DiDonato, T. E., & Freestone, D. (2012). Social projection can solve social dilemmas. *Psychological Inquiry*, 23, 1–27.
- Liljenström, H., & Svedin, U. (2005). *Micro, meso, macro: Addressing complex systems couplings*. World Scientific.
- Liu, C.-J., & Hao, F. (2011). An application of a dual-process approach to decision making in social dilemmas. *American Journal of Psychology*, 124, 203–212.
- Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology*, 4, 1–25.
- Messick, D. M., & Brewer, M. B. (1983). Solving social dilemmas: A review. In L. Wheeler & P. Shaver (Eds.), *Review of personality and social psychology* (Vol. 4, pp. 11–44). Beverly Hills, CA: Sage.

- Myerson, R. B. (1997). *Game theory: Analysis of conflict*. Cambridge: Harvard University Press.
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314, 1560–1563.
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, 437, 1291–1298.
- Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, 28, 363–392.
- Rao, L. L., Han, R., Ren, X. P., Bai, X. W., Zheng, R., Liu, H., ... Li, S. (2011). Disadvantage and prosocial behavior: The effects of the Wenchuan earthquake. *Evolution and Human Behavior*, 32, 63–69.
- Reis, H. T. (2008). Reinvigorating the concept of situation in social psychology. *Personality and Social Psychology Review*, 12, 311–329.
- Rusbult, C. E., & Van Lange, P. A. M. (1996). Interdependence processes. In: E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 564–596). New York: Guilford Press.
- Samuelson, C. D., & Messick, D. M. (1995). When do people want to change the rules for allocating shared resources? In D. A. Schroeder (Ed.), *Social dilemmas: Perspectives on individuals and groups* (pp. 143–162). Westport, CT: Praeger.
- Sell, J., & Son, Y. (1997). Comparing public goods with common pool resources: Three experiments. *Social Psychology Quarterly*, 60, 118–137.
- Tinbergen, N. (1968). On war and peace in animals and man: An ethologist's approach to the biology of aggression. *Science*, 160, 1411–1418.
- Van der Heijden, E., & Moxnes, E. (2013). Leading by example to protect the environment: Do the costs of leading matter? *Journal of Conflict Resolution*, 57, 307–326.
- Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology*, 77, 337–349.
- Van Lange, P. A. M. (2013). What we should expect from theories in social psychology: Truth, abstraction, progress, and applicability as standards (TAPAS). *Personality and Social Psychology Review*, 17, 40–55.
- Van Lange, P. A., Balliet, D. P., Parks, C. D., & Van Vugt, M. (2014). *Social dilemmas: Understanding human cooperation*. Oxford: Oxford University Press.
- Van Lange, P. A., Joireman, J., Parks, C. D., & Van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, 120, 125–141.
- Van Vugt, M. (1998). The psychology of social dilemmas. *The Psychologist*, 6, 289–292.
- Vincent T. L., & Brown, J. S. (2005). *Evolutionary game theory, natural selection, and Darwinian dynamics*. Cambridge: Cambridge University Press.
- Weber, J. M., Kopelman, S., & Messick, D. M. (2004). A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality and Social Psychology Review*, 8, 281–307.
- Wilke, H. A. (1991). Greed, efficiency and fairness in resource management situations. *European Review of Social Psychology*, 2, 165–187.
- Wit, A. P., & Kerr, N. L. (2002). “Me versus just us versus us all” categorization and cooperation in nested social dilemmas. *Journal of Personality and Social Psychology*, 83, 616–637.

Social Dilemmas: Theoretical Framework and Experimental Research

LIU Changjiang; HAO Fang

(School of Psychology, Nanjing Normal University, Nanjing 210097, China)

Abstract: Social dilemmas are situations in which individual rationality leads to collective irrationality. Advances in research for social dilemmas require a new kind of micro-meso-macro framework for understanding already existing theories. Based on these different theories, experimental research uses games as tasks to investigate basic psychological processes or to study the structural or environmental factors that influence individuals' behavioral choice in social dilemmas, and thus develops corresponding solutions to social dilemmas. Development of theory, improvement of experimental tasks, and experimental external validity all indicate prospects for future social dilemmas research via experimental methods.

Key words: Decision-making; social dilemmas; cooperation; experimental research