

社会合作秩序何以可能：社会科学的基本问题*

韦倩 姜树广

内容提要：无亲缘关系的社会成员之间的大规模合作构成了人类社会的主要特征。一个显而易见的问题是：社会合作秩序得以维持的原因是什么呢？或者说，与其它物种相比，人类具有什么样的独一无二的特征才导致大规模群体中陌生人之间的合作呢？这个问题几乎是所有社会科学都关心的问题，但是，这个问题目前为止仍然存在一些未解之谜，仅仅依靠单一学科很难回答这个问题。本文对这个问题进行了详细的探索和深层次的追问，试图以跨学科的视角对这一社会科学的基本问题给出合理的解释。

关键词：社会秩序 合作 一阶搭便车 二阶搭便车 神经元经济学

与其他物种不同，即使在规模很大的群体中，人类也可以与非亲属成员进行合作(Boyd et al., 2003; Gintis, 2000)。这种无亲缘关系的社会成员之间的广泛合作构成了人类社会的主要特征。那么，社会合作秩序得以维持的原因是什么呢？或者说，与其它物种相比，人类具有什么样的独一无二的特征才导致大规模群体中陌生人之间的合作呢？这些问题几乎是所有社会科学都关心的问题，经济学、社会学、生物学、人类学等学科都对其给予了大量关注，并提出许多理论解释。但是，这些解释目前为止仍然存在一些问题和未解之谜，仅仅依靠单一学科很难回答这个问题。《科学》杂志在创刊 125 周年之际曾组织全世界最优秀的科学家提出了 125 个“驱动基础科学研究以及决定未来科学研究方向”的科学难题（其中 25 个“大问题”，100 个“小问题”），25 个“大问题”中有 23 个是纯粹的自然科学问题，只有两个问题同时还是社会科学问题，一个是“马尔萨斯的人口论观点会继续错下去吗”，一个是“人类合作行为如何演进”（黄少安和张苏，2013）。Pennisi（2005）在同期的《科学》杂志上指出，人类合作演进的问题是一个需要演化生物学家、动物学家、神经科学家、经济学家共同努力来研究的课题，需要最终探索出“到底是什么因素在促成我们的合作精神”。英国皇家学会主席 May 在 2005 年的主席致辞中也曾说道：“人类社会和其他动物种群中的合作行为如何演化和维持是演化生物学和更广泛的社会科学中最重要的未解之谜”（Colman, 2006）。在本文中，我们将对这个问题进行详细的探索和深层次的追问，试图以跨学科的视角对这些问题和谜团进行探索和分析并给出合理的解释。

一、“一阶搭便车”之谜

按照进化论的观点，人类可以在大规模群体中与非亲属成员进行合作的现象令人困惑(Boyd et al., 2003; Gintis, 2000)，因为合作具有利他性，实施合作行为需要合作者自己负担成本，但

*韦倩，山东大学经济研究院副教授、清华大学社科学院博士后，Email: weiqian1979@163.com；姜树广，山东大学经济研究院博士研究生，Email: jsg123000@sina.com。作者感谢国家社科基金一般项目（项目编号：12BJL014）、教育部人文社科青年基金项目（项目编号：10YJC790270）、中国博士后科学基金特别资助项目（项目编号：2013T60092）、中国博士后科学基金面上资助项目（项目编号：2012M510389）、山东大学自主创新基金青年团队项目（项目编号：IFYT12096）、山东大学人文社会科学重大项目（项目编号：13RWZD07）、山东大学自主创新基金（项目编号：IFW10058）对本研究的资助；感谢浙江大学叶航教授慷慨分享仿真程序；感谢匿名审稿人的宝贵意见，但文责自负。原文 18000 字，因篇幅所限，缩减至 14000 字。

却可以使群体内的非亲属成员获益。由于背叛者无需付出任何成本就可以从合作者身上获益，因此，理性的社会成员会采取背叛行为，对合作者进行“搭便车”。这种现象被称为“一阶搭便车”行为。如果没有其他机制的约束，“一阶搭便车”现象会逐渐蔓延并最终瓦解整个社会。那么，人类是如何避免“一阶搭便车”现象从而保持社会的合作秩序呢？除了同样适用于动物界的亲缘选择理论（Hamilton, 1964）外，学者们还提出了重复互动、声誉、群体选择与空间选择等理论。

一些学者认识到，很多社会交往是重复发生的，这或许能够解释存在于亲缘关系之外的合作行为。其代表性理论为互惠利他理论（reciprocal altruism theory）。该理论认为，如果未来重复交易的机会足够高，那么，在双边交易中，自然选择会偏爱合作行为（Trivers, 1971; Axelrod and Hamilton, 1981）。合作者在未来交易中收回合作的威胁会使得对手基于长远利益而抑制其背叛行为。与此相似，博弈论的研究成果也已表明，如果博弈双方再次相遇的概率比较大，折扣率也比较低，那么，即使参与者具有短期欺骗的动机，他们也会达成一个完全合作的均衡（Friedman, 1971; Fudenberg and Maskin, 1986）。

除此之外，社会秩序还经常表现为大量多边交易中的合作行为。基于此，一些学者提出声誉考虑的理论，比如间接互惠理论（indirect reciprocity）、有成本的信号理论（costly signaling theory）等。与互惠利他理论不同，间接互惠理论（Nowak and Sigmund, 1998; Wedekind and Milinski, 2000; Leimar and Hammerstein, 2001）的理念是：个体之所以选择合作行为是期望以此获得良好声誉，从而今后能得到第三方的奖励，而不是仅来自受惠个体的回报。Nowak 和 Sigmund (1998) 考察了间接互惠的演化稳定条件，他们发现，如果合作行为的成本收益率小于行为主体获知受惠者类型的概率，那么，建立在间接互惠基础上的合作均衡将会非常稳固。与间接互惠理论不同，有成本的信号理论认为，通过合作行为，人们可以传递其作为潜在伙伴的信号，这使他们具有结盟优势，从而导致合作在无亲缘关系的社会群体中成功演化（Zahavi, 1995; Wright, 1999; Gintis et al., 2001）。比如 Gintis et al. (2001) 通过构建一个以多人公共物品博弈作为分析框架（即多个发信号者—多个观察者）的理论模型发现，由于合作行为可以显示出发信号者的潜在品质，从而引出观察者的一个响应，这会使双方都受益，从而该行为是演化稳定的，在满足一定的条件下，尽管这种发信号行为起初非常稀少，但却可以在人口中增殖。

由于生物的进化不但可以在个体层次上实现，也可以在种群层次上实现，因此，一些学者（Henrich, 2004; Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981）提出了文化群体选择理论（cultural group selection theory）来解释社会合作秩序的维持。他们认为，当个体的合作行为有利于整个种群时，合作程度更高的群体会因适应性强而生存下来，从而合作的特征就可能随种群的保存而得以繁衍和保留。

另外一些学者注意到了社会交往的网络结构特征对群体合作秩序的影响，比如空间选择（spatial selection）或网络互惠（network reciprocity）理论认为，社会交往存在空间结构或网络差异，一些人的交往显然要比其他人更为紧密，这种结构差异会使得合作者通过空间聚集来避免被背叛者剥夺，从而空间选择效应可以促进合作的演化（Nowak and May, 1992; Ohtsuki et al., 2006）。比如 Ohtsuki et al. (2006) 发现，如果合作或利他性行为的成本收益率超过相邻个体的平均数目，空间选择就会偏好合作行为。

令人遗憾的是，重复互动、声誉、群体选择与空间选择等任何一种机制都不允许利他性合作在规模很大的群体中演化。比如 Joshi (1987)、Boyd and Richerson (1988) 等已经论证，有条件合作的针锋相对之类的策略只能在 4-8 人的小规模群体中才能够成功演化，即使再次相遇的概率足够大，仅仅依靠针锋相对之类的策略或其他互惠行为也很难维持大规模群体的合作秩序。可见，重复交往只能支持一定规模内的二元关系之间的合作。而声誉、群体选择与空间选择机制只

有在群体规模比较小且迁移很少发生的条件下才能导致合作行为的稳定演化 (Leimar and Hammerstein, 2001; Rogers, 1990), 比如 Leimar and Hammerstein (2001) 发现, 当群体发生混合时, 合作率会急剧下降并逐渐趋近于零。

可见, 上述理论均不能完美解释社会合作秩序何以可能的问题。一些学者相信, 依靠惩罚威胁也许可以解决这个问题。由于背叛可能会遭受严厉的惩罚, 使得背叛行为的成本大于合作的成本, 这使得合作而不是背叛成为理性选择策略。大量实验 (Fehr and Gächter, 2002; Fehr and Fischbacher, 2003; Camerer and Fehr, 2006) 和人类学 (Boehm, 1993; Henrich, 2004) 的证据表明, 许多人愿意自己承担成本去惩罚背叛者, 甚至在一性交往情况下也是如此。因此, 惩罚可能是人类维持社会合作秩序的主要机制, 惩罚行为的存在可以有效确立社会合作秩序 (Gächter, et al., 2008; Henrich et al., 2006; Henrich and Boyd, 2001; Boyd et al., 2003)。这似乎表明, 援引惩罚机制可以解决上述困惑, 然而, 非常遗憾, 它却带来了一个新的演化之谜, 下面我们将论述并尝试解释这个二阶搭便车之谜。

二、“二阶搭便车”之谜

虽然对背叛进行惩罚可以解决“一阶搭便车”问题, 可是, 惩罚需要付出成本, 使得惩罚者的适应性降低。于是, 纯粹合作者相对于惩罚性合作者来说就是一个“二阶搭便车者(second-order free riders)” (Panchanathan and Boyd, 2004; Fowler, 2005), 这样, 就产生了一个新的演化之谜: 由于惩罚者相对于二阶搭便车者来说需要负担额外成本, 这必然使其在进化中丧失优势而消亡, 那么, 减少个人适应性的惩罚行为如何能通过自然选择得以进化呢?

Gintis (2000) 认为, 人类群体会以一定正的概率面临灭绝的威胁, 比如战争、饥荒、自然灾害等, 当人类面对灭绝的威胁时, 由于未来太过渺茫, 互惠利他和间接互惠都不能维持可以帮助群体生存所必需的合作, 并且大部分人类群体不仅仅局限于亲缘关系, 使得亲缘选择在这里也不适合, 但是, 具有一定利他性惩罚者(也称强互惠者)的群体则可以在这种环境下幸存。利他性惩罚者自己承担成本而使整个群体受益, 这虽然使得群内选择的演化压力不利于利他性惩罚行为, 但却使得群内选择的压力更偏爱利他性惩罚行为, 含有一定利他惩罚者的群体在这种情况下可以更好的生存。可见, 利他性惩罚是一种演化上的适应反应, 它可以给拥有它的族群带来可观的利益, 并且该利益会超出个体为利他性惩罚行为所付出的成本。

Bowles and Gintis (2004) 也从群体选择的角度去探讨二阶搭便车问题, 但他们探究了一个不同的机制。在他们的模型中, 发信号或群体灭绝都不起作用, 惩罚采取驱逐(ostracism)和避开(shunning)的方式。在这种情况下, 虽然利他性惩罚者承担了惩罚成本从而比其他人适存度降低, 但是当群体成员都从共同遵守社会规范中获益时, 他们也许会服从规范并惩罚违反规范者。他们使用系统动力学方法模拟了距今10—20万年以前更新世(pleistocene)晚期以狩猎——采集为生的人类社会环境, 仿真结果表明, 由突变产生较小数量的利他性惩罚者可以侵入自私者人群, 使族群内的合作行为与适应性维持在一个较高水平。他们还运用Price方程实证检验了人口层次的力量会阻止合作者把利他性惩罚者驱逐出去。合作者由于是“消极利他主义者”, 虽然在群体内的适存度相对利他性惩罚者占优, 但这个优势被大量合作者导致的群体较低平均适存度所抵消。如果群体效应足够大, 那么合作者就不会把总人口中的利他性惩罚者驱逐出去, 即便他们可以把利他性惩罚者从与他们共存的群体里赶走。

Boyd et al. (2003) 给出了另外一种解释: 尽管利他性合作和利他性惩罚的逻辑基础相似, 但它们的演化逻辑却不相同, 利他性合作者相对于背叛者的支付劣势独立于人口中背叛者的频率, 但利他性惩罚行为的成本劣势却会随背叛者人数的下降而下降, 当惩罚行为普遍时, 群内反

对惩罚的选择性压力相当微弱。正是这种非对称性的机制可以支持利他性惩罚在大规模群体中演化。他们的模拟结果显示，当没有惩罚时，群体选择只有在小规模情况时才能维持高频率的合作，而加上惩罚条件后，便可以在更大规模的群体中维持高频率的合作，且提高灭绝率可以提高长期平均的合作水平；另外，长期平均合作水平对惩罚成本非常敏感，当惩罚成本较大时，群间选择便不能补偿群内惩罚者的下降，最终使得惩罚者和捐献者都会减少；当背叛者减少时，只有在惩罚者的支付成本可以随之降低的情况下，惩罚才可以导致合作的提高，而如果惩罚者承担的成本固定，即独立于群体中背叛者的人数，则惩罚不会对合作的演化有所帮助。他们的模型揭示了不能维持利他性合作的群间选择却可以产生利他性惩罚演化的可能。

与上述研究不同，Henrich and Boyd (2001) 的社会顺从 (conformism) 理论认为，合作和惩罚的演化是人类在同化过程中采用流行行为的一个副效应，文化演化模型可以解释一次交往情境中的惩罚行为。他们认为，人类的社会学习并不是随机复制别人的行为，而是复制成功者和大多数人的行为。这种学习机制虽然不能总是产生最优行为，但却可以快速地获得高适应行为，正是这种收益偏好传播和从众传播机制使得社会可以建立合作和惩罚的规范。他们的文化演化模型显示，在有限次惩罚阶段的条件下，仅有少数尊奉习俗者的传播就可以在演化均衡中稳定群体的合作和惩罚行为。一旦一个群体的合作稳定后，便可以通过群体选择机制使得群体受益的文化特征在一个多群体的人口口扩散。文化群体选择进程主要以两种方式扩散一些亲社会行为，一是合作的群体具有更高的生产能力，从而可以支持更快的人口增长；二是模仿成功人士的行为也会使得合作行为被模仿。最后，一旦合作均衡变得普遍时，自然选择也会偏爱那些喜欢合作和惩罚背叛者的基因，因为这些基因降低了人们遭受惩罚的成本。

另有一些学者注意到，人类社会中存在各种社会排斥 (social exclusion) 机制，比如驱逐、躲避和忽视等，这会使群体中并非所有成员都能享受合作或惩罚行为的益处，从而有利于群体合作秩序的演化。这种观点被称为社会排斥理论。Panchanathan and Boyd (2004) 构建了一个包含间接互惠机制的互助游戏环境，其中，回避者 (Shunner) 只帮助具有好名声的弱者，而不会为坏名声者提供帮助。他们发现，这种基于声誉的社会排斥机制具有演化上的稳定性，可以提供对有成本合作行为的激励。同样，利他性惩罚者也可以通过排斥来约束搭便车者参与利益分享。Sasaki and Uchida (2013) 的博弈理论模型表明，排斥可以减少合作受益者的数目，从而增加其利益分享，由于驱逐者的收益超过被驱逐的背叛者，只要驱逐的成本比较低，即使在背叛者较多的群体驱逐行为也会占优而得到扩散。这表明，社会排斥可以有效克服二阶搭便车问题。

还有一些学者注意到，社会交往的空间结构特征也会对惩罚的扩散有影响，当它可以促进惩罚行为在群体中的扩散时，就会淘汰二阶搭便车者。Helbing et al. (2010) 认为，与利他性惩罚者不同，人群中还存在另一类惩罚者，虽然他们自身也是背叛者，但却会对背叛者进行惩罚，这类人被称为伪善者或背德者。在可与邻近个体进行空间交往的情形下，利他性惩罚者可以通过空间分割完全驱除搭便车者，虽然他们起初数量会严重下降，但最终会扩散到整个人群；利他性惩罚者还可以从与伪善者的罪恶同盟 (unholy cooperation) 中获益。不同于 Helbing et al. (2010) 的伪善者，Szolnoki and Perc (2013) 在空间交往情景中引入了有条件的惩罚者以区别于一般模型中的无条件惩罚者，有条件惩罚者的惩罚程度与其他惩罚者的数量成比例。他们发现，当惩罚成本较高时 (更符合现实的情况)，有条件惩罚比无条件惩罚更能有效维持合作效率。

还有其他一些非常有价值的研究也对二阶搭便车难题提供了部分解决方案，如多重惩罚 (Boyd and Richerson, 1992)、变异 (Sethi and Somanathan, 1996)、公共品的回报额度 (叶航, 2012) 等，不再一一而述。

虽然上述研究从不同角度对“二阶搭便车问题”进行了解释，提出了非常有价值的洞见，但

却仍然存在一个明显的未解之谜。它们均假设惩罚者能够低成本实施严厉惩罚，即惩罚者付出的成本要小于惩罚对背叛者造成的伤害，只有这样的条件下，惩罚和合作行为才能够成功演化。这个假设在一定程度上是合理的，因为人类具有高超的制造工具和狩猎的能力，比如投掷抛射物等，从而对背叛者可以低成本施加严厉处罚的唯一物种（Bingham, 1999）；或者少数目的惩罚者也可以通过协商、秘密行动和致命武器等手段来降低适应性成本从而击败最强大的单个敌人（Gintis, 2000）。但是，这些能力显然不只惩罚者独有，这使得惩罚者有时不是背叛者的对手，因此，低成本惩罚的假设并不总令人信服。另外，人们对惩罚的反应并非总是容忍，报复非常常见，这也会加大惩罚的实施成本，从而降低了利他性惩罚理论的可信性与适用度。比如：Janssen and Bushman（2008）经过模拟发现，如果报复行为被允许，合作和利他性惩罚不会在大规模群体中演化。很明显，如果惩罚者不能低成本实施严厉惩罚，那么，依靠惩罚来维持社会合作秩序的信念将很难再被坚持。可见，惩罚强度和惩罚成本这两个变量对合作的演化之谜起到至关重要的作用。下面我们通过仿真分析来表明这点。

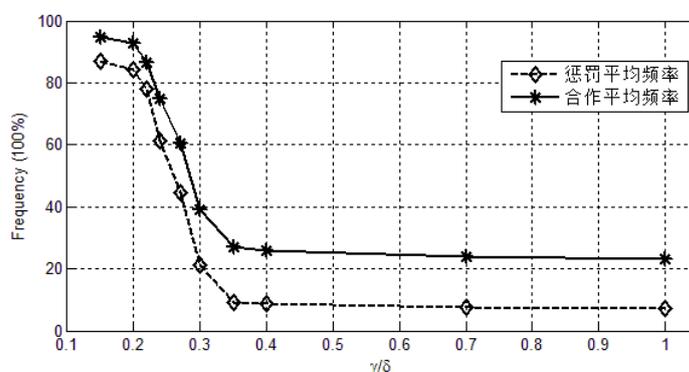


图 1 惩罚强度 δ 和惩罚成本 γ 对群体合作行为的影响

考虑一个规模恒为 m 的人群，每次随机抽取 n 个成员参加一个公共物品博弈，假定每个成员都有三种行为选择——只捐献不惩罚（合作者）、不捐献（背叛者）、既捐献又惩罚（惩罚者）。假设该群体中合作者的数量为 x ，背叛者的数量为 y ，惩罚者的数量为 z ，那么一定有 $x + y + z = n$ 。捐献需要个人支付成本 c ，但会产生一个可以被全体成员均分的总收益 $b = rc$ ，其中 $r > 1$ 为合作回报乘数；背叛虽然不需支付任何成本，但也不会给群体带来任何收益；惩罚会使群体内的每一个背叛者减少 δ 的收益，但也会给惩罚者带来 γ 的成本。不考虑同情以及监视成本等其他因素，可知：每轮博弈中捐献者的收益为 $\frac{x+z}{n}b - c$ ，背叛者的收益为 $\frac{x+z}{n}b - \delta z$ ，惩罚者的收益为 $\frac{x+z}{n}b - c - \gamma y$ 。可以看出，惩罚相对于合作来说损失了 γy 的成本，理性的个体是不愿意付出惩罚的。但是，群间选择的力量却偏爱惩罚。假设个体的行为可以进行模仿，并且假定个体更倾向于模仿更高收益的行为。此外，还假定在每一轮博弈开始之前， m 中的所有成员都将以一个很小的概率 μ 随机突变成其他类型。在完成这些步骤后，再从 m 中随机抽取 n 个成员进行新一轮博弈。假定这个公共物品博弈的随机演化过程是一个频率依赖性的 Moran 过程。我们使用浙江大学跨学科社会科学研究中心 (ICSS) 开发的仿真软件 PGG-SP C++2.0 对这个模型进行了代数为 1×10^5 的仿真^①（仿真具体结果参见附录），仿真支持了我们的观点。如图 1 所示，

^① 该软件的详细介绍参见浙江大学叶航教授的个人主页：<http://mypage.zju.edu.cn/yehang>。

横轴表示惩罚成本与惩罚强度之比 γ/δ ，纵轴表示各行为在人群中的平均频率，随着 γ/δ 的增大，惩罚者在人群中的平均比例逐渐下降，从而群体的合作程度也逐渐下降。主要参数的取值状况^①为： $\gamma/\delta = (0.15, 0.20, 0.22, 0.24, 0.27, 0.30, 0.35, 0.40, 0.70, 1.00)$ 、 $m=100$ 、 $n=5$ 、 $x=30$ 、 $y=40$ 、 $z=30$ 、 $c=1$ 、 $r=4$ 、 $\mu=0.001$ 。除此之外，我们还分别对初始状态 (x, y, z) 以及公共物品博弈的人数 n 进行了稳健性检验，仿真结果与图 1 完全相似，限于篇幅略去。

因此，一个值得进一步分析和探索的问题将是：族群或社会如何保证惩罚者能以低成本对背叛者进行严厉的惩罚？我们认为，族群中必然存在一些可以提高人类惩罚能力的机制，这有助于利他性惩罚行为在人群中稳定演化，并最终维持了群体的合作秩序。比如：一些学者认为惩罚行为并非总是无条件的或者是没有协调的个人自发行为，而是常常通过小道传闻或与其他惩罚者的沟通而得以协调进行 (Wiessner, 2005; Boehm, 1993)，因此，惩罚可以作为一种集体行为向背叛者传递群体的谴责信息。Boyd et al. (2010) 的模型就认为，协调的惩罚更符合人类祖先的现实条件，这会使得随着惩罚者人数增加，惩罚单个搭便车者的平均成本会降低，从而即使在惩罚者稀少的情况下，惩罚也会得到繁衍从而增加群体的收益；而自发的同辈惩罚 (peer punishment) 只在群体的规模足够小或个体有不合作的选择权利等限制条件下才有效。还有一些学者认为，群体具有趋向形成自我管理的机制，惩罚背叛者的权力一般被赋予特定的权威集中者，这可以提高惩罚的效力 (Henrich et al., 2010; Greif et al., 1994)。比如传统社会的村民会请求他们的首领来调节争端 (Gibson and Marks, 1995)、中世纪的欧洲商人创立了工会来维持商业秩序 (Greif, 1993) 等。在人类社会中，惩罚性的制度安排实际担当了维持社会秩序的主要责任。惩罚基金 (punishment fund) 可被看作维护公共利益的制度雏形，许多小规模的社会都使用这种方式 (如通过雇佣一个执法者)，如 Ostrom (1990) 描述了许多自筹资金保证合约执行的例子。由于集权化的制度可以更有效的克服协调问题和二次搭便车问题，因此，从演化的角度来看比同辈惩罚更有效率 (Güth et al., 2007)，可见，人类社会的层级结构和权威在合作秩序维护中扮演了重要的角色，甚至角色分化和等级统治也是大多数动物群体的组织原则。特别是在复杂的大规模社会，集权化惩罚 (centralized sanctioning) 和法定权威对维持社会的合作秩序更有意义，他们的权威来自所处社会阶层的位置 (Baldassarri and Grossman, 2011)。韦倩 (2009) 思考了人类社会中存在的提高惩罚能力的三种社会机制——规范的内化、缔结同盟与第三方介入，分析了它们在人类社会中的作用机理和相关影响因素，并构建了一个动态演化模型探讨了这三种机制对于维持利他性惩罚行为的演化和社会合作秩序的作用。

可见，人类族群或社会中存在一些可以提高惩罚能力的社会机制，即使在单个惩罚者的惩罚能力不能够满足惩罚行为稳定演化的条件下，这些机制也会自动调整其惩罚能力，使得利他性惩罚行为能够在人群中保持稳定演化，从而维持了社会的合作秩序。因此，我们的研究可以大大增强利他性惩罚理论的解释力度和适用范围。

三、合作行为的神经学基础

虽然利他性惩罚是维持社会合作秩序的主要机制，并且族群或社会中存在许多机制使利他性惩罚得以实施，从而避免了二阶搭便车问题，但是这相对于人类丰富多彩的合作行为来讲仍显不足，因为它并不一定可以保证惩罚行为对背叛者造成的伤害一定大于惩罚者付出的成本，即使高级形式的制度安排仍然面临执法者犯法的危险。

对不惩罚背叛者的人进行惩罚 (二阶惩罚) 也许可以完美解决“二阶搭便车”的问题 (Boyd and Richerson, 1992; Colman, 2006)，但是，这会带来更高阶的“三阶搭便车”，依次类推，还

^① 各参数的取值理由及含义参见浙江大学叶航教授的个人主页：<http://mypage.zju.edu.cn/yehang>。

可能存在四阶、五阶乃至无穷阶的搭便车问题。比如，集权化制度中的权威者也可能会监守自盗，以及即使在社会制度已经发育完善的现代社会中仍然存在无法被完全监督和制约的个体（如国家元首）等。现有文献对高阶搭便车现象的分析非常少见。运用逆推法可知，当搭便车的阶数 N 趋向于无穷时，由于不存在更高价的惩罚，此时再也不能依靠物质利益来约束搭便车者，这样，搭便车者会破坏这一层级内的合作；再推到前一阶，当成员预期到上一层级内的合作秩序不能维持时，他们也不会愿意自己承担成本去维持这一层级内的合作秩序，如此类推，最终整个社会体系都会崩溃，无论 N 的链条延长到何种程度都不管用。因此，当考虑到无穷阶搭便车问题时，再也不能从物质利益的角度来解释搭便车问题了，这时候社会秩序的维持和运转只能依靠社会成员内心的道德约束。那么，人类内心是否天生就具有愿意与别人合作、愿意自己付出成本惩罚别人、具有关心他人的亲社会情感呢？神经元经济学^①的研究表明，人是被情感、情绪、道德等因素支配的综合社会人，人类心智结构的发展也体现出对合作秩序遵从的固有倾向。

首先，人类具有奖励合作行为的心智结构。人类之所以可以牺牲自身的物质利益参与合作，可能源自个体在相互合作中可以获得更高的快乐体验。比如 Rilling et al. (2002) 发现，被试在与人类对手相互合作时可以比他与电脑对手合作时获得更高层次的纹状体激活，由于在这两种情形下其货币收益是相同的，因此，这表明他从合作行为中获得了额外收益。但这个建立在重复博弈之上的实验可能会混同其他影响因素，从而影响结果的可信性，因此，Rilling et al. (2004) 进行了关于一次序贯社会困境博弈的实验，实验结果再一次表明，与计算机对手情况相比，被试与人类对手的合作可以产生更高的纹状体激活。Singer et al. (2004) 完成了一个更深层次的实验，他们发现，仅仅看到以前合作的个体的脸就可以激活激励相关脑区，这揭示了合作行为的特殊效用品质。这个结果意味着人们可以从与合作的人们的相互交往中获得更大的效用，这不仅是因为他们在这些互动中可以获得更多的金钱，而是因为这些互动本身就具有奖励作用。

其次，人类具有奖励惩罚行为的心智结构。神经与认知科学的证据表明，观察到背叛行为没有得到惩罚是一件痛苦的事情，同时人们从惩罚背叛行为本身中也可以获得效用的满足。De Quervain et al. (2004) 利用正电子发射成像技术对二人序贯社会困境博弈进行的研究对此提供了充分的神经学证据。由于大脑的背侧纹状体包括尾核与壳核的神经回路是人类及灵长类动物的一个与激励相关的关键部位，因此，当对背叛进行惩罚时如果能观察到这一脑区的激活，那么，就可以证实人们能从惩罚本身中获得收益。在被试 A 将自己的货币全部交给 B 而后 B 选择什么也不给 A 的情况下，实验者为 A 设计了 4 个不同场景：(1) IC，即惩罚对于 A 和 B 都是有代价的，此时 A 施加给 B 的每一个惩罚点将会给 A 带来一单位代币的成本，但却可以让 B 损失两单位代币；(2) IF，即惩罚对于 B 有代价但是 A 没有代价，此时 A 花费一个惩罚点来惩罚 B 时自己没有任何成本但是会让 B 损失两单位货币；(3) IS，即 A 对 B 的惩罚是象征性的，此时，惩罚既不能减少 A 的货币也不能减少 B 的货币；(4) NC，即 B 的行为是随机的且 A 事先知道，但是惩罚对于 A 和 B 来说仍然是有代价的。实验结果显示，A 在 IC、IF 和 IS 三种情况下都表现出强烈的惩罚愿望：在 IF 条件下，全体被试都对 B 实施了惩罚；在 IC 条件下，14 个被试中有 12 个对 B 实施了惩罚；在 IS 条件下，14 个被试中有 6 个对 B 实施了惩罚。与此形成鲜明对比的是，在 NC 条件下，A 几乎没有惩罚愿望，14 个被试只有 3 个惩罚了 B，而且惩罚强度相当低。这个实验揭示了人们宁愿舍弃私利也要惩罚背叛者的动力所在，即惩罚者可以从惩罚行为本身中获得欲望满足。在另一项研究中，Singer et al. (2006) 在被试或其搭档被电击（轻微疼痛）时使用 fMRI 扫描其大脑发现，当合作的搭档被电击时，被试的前扣带区和前脑岛都观察到移情反应（empathic responses），而伏核（NAcc）和 OFC 区域则在不合作者被电击时表现活跃，被试报

^① 关于神经元经济学的研究参见叶航、汪丁丁和贾拥民（2007）的综述。

复动机的强烈程度也与 NAcc 的活跃程度相关，这表明对不合作者惩罚可以得到奖赏。

再次，人类具有关爱他人的心智结构。这种心智结构可能起源于哺育后代的挑战(Hrdy, 2011)或来自种群内的冲突(Bowles and Gintis, 2011)。最近有关大脑神经系统的一些研究显示，人们生来就具备关心他人的神经系统——迷走神经(vagus nerve)。迷走神经位于胸部，从脊髓顶端开始，然后在全身漫游，分布于面部肌肉组织、与发音相关的肌肉、心脏、肺、肾脏、肝脏以及各种消化器官等，主管人的发声、心脏和呼吸节律、消化和其他功能。克特纳(2009)认为，当人产生同情时，迷走神经会处于最兴奋的状态，这会让人作出牺牲。Porges(2011)认为，迷走神经通过以下三种方式产生怜悯之情：首先，迷走神经可以调动参与关爱行为的各个肌肉群，包括面部肌肉组织和发音器官等。比如当一个人听到受苦受难的描绘时，往往会发出有节奏地叹息。这声叹息是关切与理解的情感表达，它会使叹息者平静下来，并可以缓解描绘者的痛苦，增加其信任。其次，迷走神经是控制人类心率的首要制动装置。如果迷走神经没有被激活，人的心脏平均要跳动 115 次，而不是一般情况下的 72 次。当人们愤怒或害怕时，心跳就会加快，把血液输送到各肌肉群，让身体做好战斗或逃跑准备。迷走神经的作用正好相反，它可以降低心率，让人心平气和，增加了与他人近距离温柔接触的可能性。再次，当迷走神经启动时，会引发后叶催产素的释放，而催产素是引致亲社会行为的一种激素，它会在大脑和全身各处发出了温暖、信任、热爱的信号，刺激出现了亲切友好的说话腔调以及更为平静的心血管生理现象，这些信号最终还会传达到别人身上。可见，迷走神经经常是各种亲社会行为的来源。值得一提的是，迷走神经是哺乳动物所特有的神经，或许正是它支撑了哺乳动物特殊的照管后代行为。

最后，人类具有同情共感的心智结构。迷走神经只可以激发人的“同情”，但是个体如何感知别人的痛苦、快乐呢？这就要依靠镜像神经元。神经与认知科学的研究发现，包括人在内的灵长类生物的大脑中央运动前皮层有一个 F5 区域，该区域的功能与其他绝大多数神经元不同，大多数神经元的活跃程度是与特定的动作联系在一起的，而这些神经元的活动并不与任何简单的、构成一个完整动作的“元动作”相关联，而是与一个完整的动作相关联(Rustichini, 2005)。镜像神经元的激活具有以下特征：首先，当它们观察到一个特定动作时，就会作出反应，且与这个动作主体相距的距离无关；其次，镜像神经元是对动作，而不是对动作的对象作出反应；再次，镜像神经元是对动作在内心的呈现产生反应，而不是针对动作的物理现实，比如 Umiltá et al. (2001)通过实验证实，激活神经元需要完整的动作，而且这与动作是否有一部分被隐藏而不能被观察到无关；最后，内在心理状态也存在镜像，Wickers et al. (2003)通过实验证实，被试观察到的情绪也会激活其相关的“镜像神经元”，因此，这表明人们亲身感受某种情绪时所激活的脑区，与他观察到其他人感受同一种情绪时所激活的脑区确实存在着重叠部分(Rustichini, 2005)。可见，同情具有大脑活动的基础，“同情”只不过是做某件事和观察到同一件事时基于“镜像神经元”实现的“神经网络共享”。在人类具有了同情共感的能力后，才可以将制度内化于心。

可见，正是人类天生具有愿意与别人合作、愿意自己付出成本惩罚别人、具有关心他人的亲社会情感的禀赋，社会秩序才得以维持。最近的一项研究也表明，合作在很大程度上由直觉神经系统实现而不是由慎思神经系统来完成的。由于直觉神经系统的决策时间一般较短，而慎思神经系统的决策时间相对较长，因此，Rand et al. (2012)通过公共物品实验研究被试的决策时间与最终捐献率之间的相关性，他们发现决策时间短的被试捐献得更多，从而更倾向于合作；而且当要求被试加快决策速度时可以提高其捐献率，当提示他们放慢决策速度时会降低其捐献率。可见，合作在很大程度上不是算计的结果，而是人类天生的心智结构自动完成的。对照我们日常生活中的行为决策，我们每个人都应该对此有所感知。当然，这里还存在一个问题：人类是如何演化出这样的心智结构呢？虽然这需要对人类原始社会的生活形态进行更深入的探索 and 更严谨的分析，

但道理却可能非常简单：人类偏爱合作的心智结构是人类在长期演化过程中通过自然选择固化的产物，即自然选择会使得人类进化出偏好合作行为的神经机制，并把它编码到自动或情感的神经过程中。我们认为，一种很可能的途径是：原始人最初只是小规模的人群生活，日常交往对象都是亲戚和近邻，这时候亲缘选择和直接互惠的动机发挥主要作用，在这种情景下，人类出于哺育后代的压力等演化出同情共感、关爱他人、惩恶扬善的能力以及丰富的情感情绪，从而具有了在更大范围合作的神经基础。在此基础上，人类的交往与合作范围开始扩大，因为此时间接互惠以及利他性惩罚开始扮演维持合作秩序的主要角色，再加上人类社会已经演化出增强惩罚能力的一些机制，人类的生产水平也能够支持微弱劣势的个体不会立刻被演化淘汰，从而利他性惩罚者可以始终在人群中维持一定的比例，构成社会秩序的终极守护者，使人类成为自然界中唯一一个可以在大规模群体中与非亲属成员进行合作的物种。

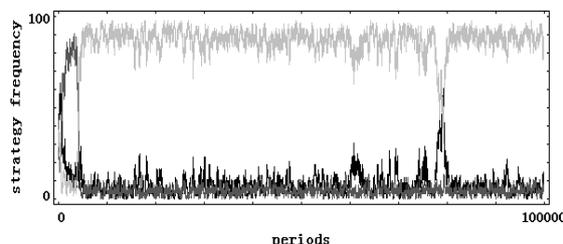
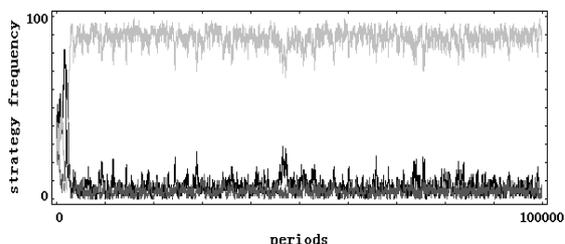
四、结论

与其他物种不同，即使在规模很大的群体中，人类也可以与非亲属成员进行合作。这种无亲缘关系的社会成员之间的广泛合作构成了人类社会的主要特征。那么，社会合作秩序得以维持的原因是什么呢？对人类合作难题的探究引发了学术界极大的兴趣。首先，学者们先注意到了“一阶搭便车”之谜，即背叛者对合作者的搭便车现象，并基于重复互动、声誉与群体选择等机制进行解释，但是，它们中的任何一种机制都不支持利他性合作行为在大规模群体中演化。于是，许多研究人员相信，依靠惩罚威胁可以维持人类大规模的合作。但它却带来了另外一个新的演化之谜，与惩罚性合作者相比，只合作而不惩罚者就成为一种“二次搭便车者”。然后，本文对二阶搭便车问题的相关文献进行总结，对比分析了它们的不足和缺陷。还通过仿真讨论了惩罚强度和惩罚成本这两个变量对二次搭便车问题的影响，并讨论了人类社会上存在的一些可以改变它们的社会机制，发现即使单个惩罚者的惩罚能力不能够满足惩罚行为稳定演化的条件，一些社会机制也会自动调整其惩罚能力，使得利他性惩罚行为能够在人群中稳定演化，维持了群体合作秩序，这大大增强了利他性惩罚理论的解释力度和适用范围。或许对不惩罚背叛者的人进行惩罚可以完美解决“二阶搭便车”的问题，但是，这会带来更高阶的“三阶搭便车”问题，依次类推，还可能存在四阶、五阶乃至无穷阶的搭便车问题。“社会何以可能”的问题最终需要向人的内心去追问。神经经济学的最新研究成果支持了我们的观点，神经经济学的研究表明，人类内心天生就具有愿意与别人合作、愿意自己付出成本惩罚别人、具有关心他人的亲社会情感的心智结构。

附录：仿真图（参数的取值状况为： $m=100$ 、 $n=5$ 、 $x=30$ 、 $y=40$ 、 $z=30$ 、 $c=1$ 、 $r=4$ 、 $\alpha_1=1$ 、 $\delta=1$ 、 $\omega=1$ 、 $\mu=0.001$ ），其中黑色线条表示合作者、褐色线条表示背叛者、浅色线条表示惩罚者。

(1) $\gamma/\delta=0.15$

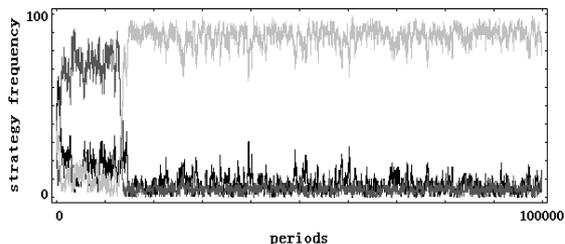
(2) $\gamma/\delta=0.20$



平均比例：合作者 7.87%、背叛者 5.16%、惩罚者 86.97%；合作者 8.42%、背叛者 7.31%、惩罚者 84.27%

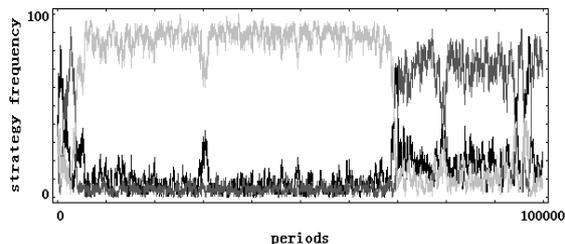
(3) $\gamma/\delta=0.22$

(4) $\gamma/\delta=0.24$

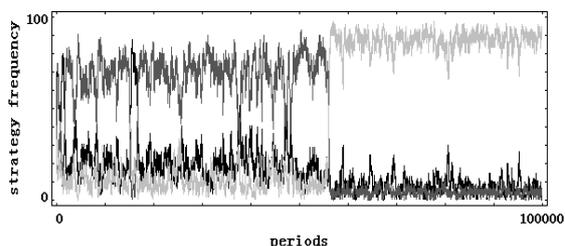


平均比例：合作者 8.84%、背叛者 13.33%、惩罚者 77.83%；合作者 13.54%、背叛者 25.11%、惩罚者 61.35%

(5) $\gamma/\delta = 0.27$

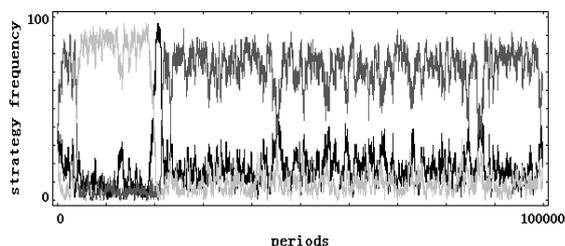


(6) $\gamma/\delta = 0.30$

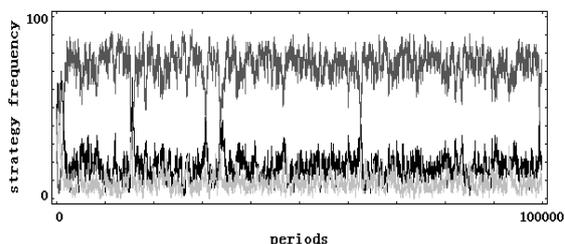


平均比例：合作者 15.70%、背叛者 39.64%、惩罚者 44.56%；合作者 71.82%、背叛者 60.81%、惩罚者 21.37%

(7) $\gamma/\delta = 0.35$

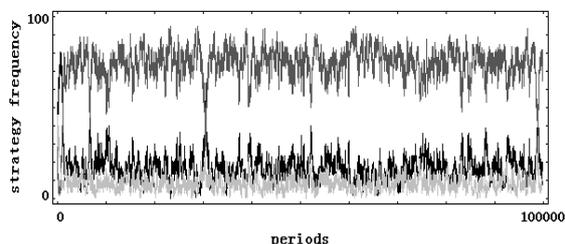


(8) $\gamma/\delta = 0.40$

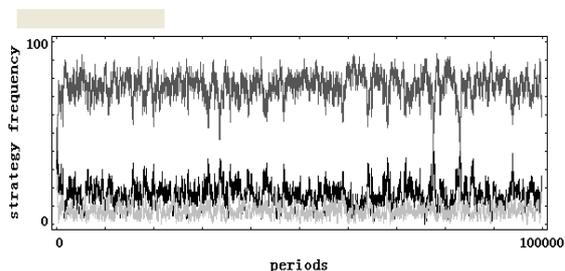


平均比例：合作者 17.95%、背叛者 72.89%、惩罚者 9.16%；合作者 17.24%、背叛者 74.06%、惩罚者 8.70%；

(9) $\gamma/\delta = 0.70$



(10) $\gamma/\delta = 1$



平均比例：合作者 16.28%、背叛者 75.99%、惩罚者 7.73%；合作者 16.08%、背叛者 76.67%、惩罚者 7.25%

参考文献

达契尔·克特纳，2009：《生而向善》，中国人民大学出版社。

黄少安和张苏，2013：《人类的合作及其演进：研究综述和评论》，《中国社会科学》第7期。

- 韦倩, 2009: 《增强惩罚能力的若干社会机制与群体合作秩序的维持》, 《经济研究》第10期。
- 叶航, 2012: 《公共合作中的社会困境与社会正义》, 《经济研究》第8期。
- 叶航、汪丁丁和贾拥民, 2007: 《科学与实证: 一个基于“神经元经济学”的综述》, 《经济研究》第1期。
- Axelrod, R., and W. D. Hamilton, 1981, “The Evolution of Cooperation”, *Science*, 211: 1390-1396.
- Baldassarri, D., and G. Grossman, 2011, “Centralized Sanctioning and Legitimate Authority Promote Cooperation in Humans”, *Proceedings of the National Academy of Sciences of the United States of America*, 108: 11023-11027.
- Bingham, P. M., 1999, “Human Uniqueness: A General Theory”, *Quarterly Review of Biology*, 74(2): 133-169.
- Boehm, C. H., 1993, “Egalitarian Behavior and Reverse Dominance Hierarchy”, *Current Anthropology*, 34: 227-254.
- Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson, 2003, “The Evolution of Altruistic Punishment”, *Proceedings of the National Academy of Sciences of the United States of America*, 100(6): 3531-3535.
- Boyd R., and P. J. Richerson, 1985, *Culture and the Evolutionary Process*, Chicago: University of Chicago Press.
- Boyd, R., and P. J. Richerson, 1988, “The Evolution of Reciprocity in Sizable Groups”, *Journal of Theoretical Biology*, 132: 337-356.
- Boyd, R., and P. J. Richerson, 1992, “Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups”, *Ethology and Sociobiology*, 13: 171-195.
- Boyd, R., H. Gintis, and S. Bowles, 2010, “Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare”, *Science*, 328: 617-620.
- Bowles, S., and H. Gintis, 2004, “The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Populations”, *Theoretical Population Biology*, 65: 17-28.
- Bowles, S., Gintis, H., 2011. *A Cooperative Species: Human Reciprocity and its Evolution*, Princeton, NJ: Princeton University Press.
- Camerer, C. F., and E. Fehr, 2006, “When Does Economic Man Dominate Social Behavior?”, *Science*, 311: 47-52.
- Cavalli-Sforza, L. L. and M. W. Feldman, 1981, *Culture Transmission and Evolution*, Princeton, NJ: Princeton University Press.
- Colman, A. M., 2006, “The Puzzle of Cooperation”, *Nature*, 440: 744-745.
- De Quervain Dominique J.-F., U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck and E. Fehr, 2004, “The Neural Basis of Altruistic Punishment”, *Science*, 305: 1254-1258.
- Fehr, E., and S. Gächter, 2002, “Altruistic Punishment in Humans”, *Nature*, 415: 137-140.
- Fehr, E., and U. Fischbacher, 2003, “The Nature of Human Altruism”, *Nature*, 425: 785-791.
- Fowler, J.H., 2005, “Second-order Free-riding Problem Solved?”, *Nature*, 437(7058): E8-E8.
- Friedman, J., 1971, “A Noncooperative Equilibrium for Supergames”, *Review of Economic Studies*, 38: 1-12.
- Fudenberg, D., and E. Maskin, 1986, “The Folk Theorem in Repeated Games with Discounting or with Incomplete Information”, *Econometrica*, 54: 533-556.
- Gächter, S., E. Renner, and M. Sefton, 2008, “The Long-Run Benefits of Punishment”, *Science*, 322(5907): 1510-1512.
- Gibson, C. C., S. A. Marks, 1995, “Transforming Rural Hunters into Conservationists: An Assessment of Community-Based Wildlife Management Programs in Africa”, *World Development*, 23: 941-957.
- Gintis, H., 2000, “Strong Reciprocity and Human Sociality”, *Journal of Theoretical Biology*, 206: 169-179.
- Gintis H., E. A. Smith and S. Bowles, 2001, “Costly Signaling and Cooperation”, *Journal of Theoretical Biology*, 213: 103-119.

- Greif, A., 1993, "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition", *American Economic Review*, 83: 525–548.
- Greif, A., P. Milgrom, and B. R. Weingast, 1994, "Coordination, Commitment, and Enforcement: The Case of The Merchant Guild", *Journal of Political Economy*, 102: 745–776.
- Güth, W., M.V. Levati, M. Sutter, E. van der Heijden, 2007, "Leading by Example with and without Exclusion Power in Voluntary Contribution Experiments", *Journal of Public Economics*, 91: 1023–1042.
- Hamilton, W. D., 1964, "The Genetical Evolution of Social Behaviour, I and II", *Journal of Theoretical Biology*, 7(1): 1–52.
- Helbing, D., A. Szolnoki, M. Perc, and G. Szabó, 2010, "Evolutionary Establishment of Moral and Double Moral Standards through Spatial Interactions", *PLoS Computational Biology*, 6(4): e1000758.
- Henrich, J. P., 2004, *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence From Fifteen Small-Scale Societies*, Oxford: Oxford University Press.
- Henrich, J., and R. Boyd, 2001, "Why People Punish Defectors: Weak Conformist Transmission can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas", *Journal of Theoretical Biology*, 208: 79–89.
- Henrich, J., R. McElreath, A. Barr, J. Ensminger, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. Tracer, and J. Ziker, 2006, "Costly Punishment across Human Societies", *Science*, 312(5781): 1767–1770.
- Henrich, J., J. Ensminger, R. McElreath, A. Barr, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. Tracer, and J. Ziker, 2010, "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment", *Science*, 327(5972): 1480–1484.
- Hrdy, S.B., 2011, *Mothers and Others: The Evolutionary Origins of Mutual Understanding*, Cambridge, MA: Harvard University Press.
- Janssen, M. A., and C. Bushman, 2008, "Evolution of Cooperation and Altruistic Punishment When Retaliation Is Possible", *Journal of Theoretical Biology*, 254: 541–545.
- Joshi, N.V., 1987, "Evolution of Cooperation by Reciprocation within Structured Demes", *Journal of Genetics*, 66(1): 69–84.
- Leimar, O., and P. Hammerstein, 2001, "Evolution of Cooperation through Indirect Reciprocity", *Royal Society*, 268: 745–753.
- Nowak, M.A., and K. Sigmund, 1998, "Evolution of Indirect Reciprocity by Image Scoring", *Nature*, 393: 573–577.
- Nowak, M. A., and R. M. May, 1992, "Evolutionary Games and Spatial Chaos", *Nature*, 359(6398): 826–829.
- Ohtsuki H., C. Hauert, E. Lieberman, and M. A. Nowak, 2006, "A Simple Rule for the Evolution of Cooperation on Graphs and Social Networks", *Nature*, 441: 502–505.
- Ostrom, Elinor, 1990, *Governing the Commons*, New York: Cambridge University Press.
- Panchanathan, K., and R. Boyd, 2004, "Indirect Reciprocity Can Stabilize Cooperation without the Second-Order Free-Rider Problem", *Nature*, 432: 499–502.
- Pennisi, E., 2005, "How Did Cooperative Behavior Evolve", *Science*, 309(5731): 93.
- Porges, S. W., 2011, *The Polyvagal Theory: Neurophysiological Foundations of Emotions, Attachment, Communication, and Self-Regulation*, New York: W. W. Norton.
- Rand, D. G., J. D. Greene, and M. A. Nowak, 2012, "Spontaneous Giving and Calculated Greed", *Nature*, 489: 427–430.
- Rilling, J. K., D. A. Gutman, T. R. Zeh, G. Pagnoni, G. S. Berns, and C. D. Kilts, 2002, "A Neural Basis for Social

Cooperation”, *Neuron*, 35: 395-405.

Rilling, J. K., A. G. Sanfey, J. A. Aronson, L. E. Nystrom, and J. D. Cohen, 2004, “Opposing Bold Responses to Reciprocated and Unreciprocated Altruism in Putative Reward Pathways”, *Neuroreport*, 15(16): 2539-2543.

Rogers, A. R., 1990, “Group selection by selective emigration: The effects of migration and kin structure”, *American Naturalist*, 135: 398-413.

Rustichini, A., 2005, “Neuroeconomics: Present and Future”, *Games and Economic Behavior*, 52: 201-212.

Sasaki, T. and S. Uchida, 2013, “The Evolution of Cooperation by Social Exclusion”, *Proceedings of the Royal Society B*, 280: rspb.20122498.

Sethi, R., and E. Somanathan, 1996, “The Evolution of Social Norms in Common Property Resource Use”, *American Economic Review*, 86(4): 766-788.

Singer, T., S. J. Kiebel, J. S. Winston, H. Kaube, R. J. Dolan, and C. D. Frith, 2004, “Brain Responses to the Acquired Moral Status of Faces”, *Neuron*, 41(4): 653-662.

Singer, T., B. Seymour, J. P. O’Doherty, K.E. Stephan, R. J. Dolan, and C.D. Frith, 2006, “Empathic Neural Responses Are Modulated by the Perceived Fairness of Others”, *Nature*, 439: 466-469.

Trivers, R. L., 1971, “The Evolution of Reciprocal Altruism”, *Quarterly Review of Biology*, 46: 35-57.

Umiltà, M. A., E. Kohler, V. Gallese, L. Fogassi, L. Fadiga, C. Keysers, and G. Rizzolatti, 2001, “I Know What You Are Doing: A Neurophysiological Study”, *Neuron*, 31: 155-165.

Wedekind, C., and M. Milinski, 2000, “Cooperation through Image Scoring in Humans”, *Science*, 289: 850-852.

Wickers, B., C. Keysers, J. Plailly, J. Royet, V. Gallese, and G. Rizzolatti, 2003, “Both of Us Disgusted in My Insula: The Common Neural Basis of Seeing and Feeling Disgust”, *Neuron*, 40: 655-664.

Wiessner, P., 2005, “Norm Enforcement among the Ju/’hoansi Bushmen”, *Human Nature*, 16(2): 115-145.

Wright, J., 1999, “Altruism as Signal: Zahavi’s Alternative to Kin Selection and Reciprocity”, *Journal of Avian Biology*, 30(1): 108-115.

Zahavi, A., 1995, “Altruism as a Handicap: The Limitations of Kin Selection and Reciprocity”, *Journal of Avian Biology*, 26: 1-3.

How Is Social Cooperative Order Possible: Exploring Mysteries

Wei Qian & Jiang Shuguang

(Center for Economic Research, Shandong University; School of Social Science, Tsinghua University)

Abstract: Cooperation between genetically unrelated individuals in large groups constitutes the main feature of human society. So we want to know how social order is possible. Or, which kind of the unique features of human can lead to cooperation between strangers in large groups? This problem attracts almost all the researchers in social sciences. However, there are still many mysteries about this issue. Maybe only relying on a single discipline is very difficult to answer the question. So we try to explore this basic problem of social science in an inter-disciplinary perspective in this paper.

Keywords: social order, cooperation, first-order free rider, second-order free rider, neuroeconomics

JEL Classification: D710