

集体风险困境的研究进展与应用综述

陈鸿窈

浙江理工大学心理学系, 浙江 杭州

收稿日期: 2023年5月15日; 录用日期: 2023年6月19日; 发布日期: 2023年6月28日

摘要

集体风险困境作为社会困境的一种变体, 可以更确切地描述许多当代社会挑战, 比如气候变化和传染病的传播等问题, 人们在该情境下需要共同合作以避免集体性的灾难。集体风险困境的研究与全球人民的福祉息息相关。研究得到的理论成果和实证证据, 有利于人们全面理解集体风险下的行为决策并提出解决措施。目前该领域的研究结论较为分散, 因此本文对集体风险困境进行了概念梳理, 并对已有研究结论进行回顾和归纳, 在此基础上展望未来的研究, 为后续的研究提供参考方向。

关键词

集体风险, 社会困境, 合作, 行为决策

A Review on the Research Progress and Application of Collective Risk Dilemma

Hongyao Chen

Department of Psychology, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: May 15th, 2023; accepted: Jun. 19th, 2023; published: Jun. 28th, 2023

Abstract

The collective risk dilemma, as a variant of social dilemma, can accurately describe many contemporary social challenges such as climate change and the spread of infectious diseases. In this context, people need to cooperate to avoid collective disasters. The study of collective risk dilemma is closely related to the well-being of global citizens. The theoretical achievements and empirical evidence obtained from this research are conducive to a comprehensive understanding of the actions under collective risk and proposing solutions. Currently, the conclusions in this field are relatively scattered. Therefore, this article summarizes the concept of collective risk dilemma, reviews and summarizes research conclusions, and looks forward to future research to provide ref-

erence for subsequent studies.

Keywords

Collective Risk, Social Dilemma, Cooperation, Decision-Making

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

从气候变化、生态环境破坏到传染病的传播,许多当代社会挑战引发了集体行动问题。在面临这些问题时,人们可以选择共同合作来避免集体损失,但也可能为了个体利益选择搭便车,希望从别人的努力中获益,这种做法对集体合作有害,最终会形成社会困境(Dreber & Nowak, 2008)。这些问题让研究者认识到开发一个新的综合框架来研究人类合作行为与集体风险的重要性,同时也引发了心理学、组织行为学、环境学等多学科的广泛关注(Chen & Szolnoki, 2018; Tilman, Plotkin, & Akçay, 2020)。集体风险困境(collective risk dilemma, CRD)就是针对类似的全球性问题的一种概括框架,与过去常见的社会困境不同的是,在这种情况下,参与者的目标是避免损失,而不是获得利益,而且只有当所有参与者的共同贡献超过集体目标时,才能避免损失(Milinski et al., 2008)。

集体风险困境的研究虽然近几年才得到关注,但是其与全球人民的福祉息息相关。研究得到的理论成果和实证证据,有利于人们科学全面地理解集体风险下的行动并提出解决措施。目前该领域的研究结论较为分散,故本文将系统性地对集体风险困境进行概念梳理,并对已有研究结论进行回顾和归纳,在此基础上对未来研究进行展望,为后续的研究提供参考方向。

2. 社会困境的概念基础

2.1. 基本概念

Hardin (1968)认为,当一种公共的有限资源被想从中谋取私人利益的个体所管理时,会不可避免地导致公地悲剧。在 Hardin 看来,个人利益和集体利益之间的两难困境最终会破坏公共资源,除非介入“道德”或者是实施“相互牵制”的制度,这就是社会困境的由来。Dawes (1980)则对社会困境进行了更加明确的表述,社会困境具有两个属性:不管其他社会成员做什么,个人背叛行为的社会回报都高于合作行为的回报;然而如果所有人都不合作会比社会中所有人都合作的回报更低。Ostrom (1990)指出,个体理性决策与集体最优结果不一致的社会困境无处不在,但并非完全依靠强制性措施才能解决。

2.2. 研究范式

当前学界有多种社会困境研究范式,对于双边博弈,可以采用囚徒困境、最后通牒博弈、小鸡博弈、独裁者博弈、信任博弈等(Güth, Schmittberger, & Schwarze, 1982; Van Lange et al., 2013; Sigmund et al., 2010)。而在生活中还存在着多人的合作互动,针对这种情况,更常见的是使用公共物品博弈(Public Goods Game, PGG)(Nowak & May, 1992)。公共物品存在着非竞争性和非排他性,也即不论个体有无贡献,都可以享受到集体利益,这也导致团队中可能存在搭便车的问题,因此团队合作中的利益冲突能够用公共物品困境来模拟(Balliet & Lange, 2013)。

在典型的公共物品博弈中，参与者会获得初始的游戏货币，通常被称为禀赋。参与者可以选择是否将其中的一部分投入到公共资源库中。公共资源库中收集到的集体投资金额会乘以增强因子，这被称为人均边际收益，并将最终得到的总金额平均重新分配给所有参与者。没有投入公共资源池的部分则始终为个人所有，也不会乘以增强因子。因此，在公共物品博弈的情境下，理性或自私的选择是不向公共资源库贡献，这种做法可能会损害其他贡献人的利益。相反，当每个个体将其禀赋全部贡献到公共资源库时，可以达到集体最优结果。在公共物品博弈中，贡献的大小可以用来衡量群体合作水平(Fehr & Gächter, 1999, 2002)。

2.3. 影响因素

目前学界对促进社会困境中合作水平的机制进行了广泛的研究，将社会困境的解决方案大致分为两大类：动机性解决方案和结构性解决方案(Kollock, 1998)。动机性的解决方案假设行动者不是完全以自我为中心的，人们在做决定时会考虑伙伴的结果。与之相关的是社会价值取向以及沟通交流(Bochet, Page, & Putterman, 2006; Murphy & Ackermann, 2014)、间接互惠机制和声誉监控(Nowak & Sigmund, 2005; Okada, 2020)，此外还可以从心理机制出发，增加同伴之间的信任、群体身份认同以及培养对长远利益的考虑(Chaudhuri, 2011; Parks, Joireman, & Van Lange, 2013)。

结构性解决方案则涉及到改变博弈规则，如重复博弈、设置阈值、改变群体规模、设置奖励与惩罚、加入领导者等外部机构(Merhej, Santos, Melo, & Santos, 2021; Nockur, Pfattheicher, & Keller, 2021; Rondeau, Poe, & Schulze, 2005)。这些发现为社会困境下新的研究工作提供了坚实的基础。

3. 集体风险困境研究现状

世界经济论坛的全球风险报告指出，当前最令人担忧的风险包括气候变化、生物多样性丧失、极端天气以及社会分裂和经济脆弱性(Merhej et al., 2022)。值得注意的是，通过调整个人和集体的行为，仍然可以避免灾难性的后果，这需要从普通公民到上层决策者等众多主体之间的合作与协调。无论是在地区、国家还是国际层面，环境、社会、经济和其他可预见的集体风险都需要集体努力，才能成功地解决问题，进而避免危机。在这种背景下的合作会带来一个全新的社会困境：如果其他人对集体利益做出贡献，并且自身不做出贡献，通过搭便车避免风险，则会出现最佳的个人结果。这种自私的完全理性的决策，以及将责任转移给他人的做法，会导致所谓的社会困境(Dawes, 1980)。

3.1. 集体风险的概念

全球变暖和其他有灾难性后果的社会困境不能被简单地框定为标准的公共物品博弈。针对这一情况，Milinski 等(2008)引入了“集体风险困境”的概念，以描述气候变化问题所带来的广泛性和灾难性的经济损失威胁(Milinski et al., 2008)，并在此基础上对参与者的决策行为进行分析研究。虽然最初集体风险困境的概念是针对气候危机而提出的，但是它所描述的情境也同样适用于其他需要集体努力来避免共同损失的情况，比如近年来发生的 COVID-19 大流行，需要个体主动做好防疫措施，自觉接种疫苗(Osama, Pankhania, & Majeed, 2020)。也就是说，集体风险困境也属于社会困境，但在集体风险困境中，合作努力的目的是防止公共灾难而不是获得集体利益。

随着集体风险困境研究的发展，该情境下的群体行为模式以及心理机制可以采取多种探究工具，目前的研究主要集中在行为实验、进化博弈论模型和强化学习模型这几个方面，后两者都属于仿真模型模拟的范畴。

3.2. 行为研究范式

集体风险困境是公共物品博弈的一种变式，具体表现为在集体风险困境中集体贡献必须达到一个阈

值才能产生公共利益。此外，集体风险困境还假设在没有达到阈值的情况下，每个参与者都有损失大部分剩余禀赋的风险。

在最初的也是目前使用最为广泛的集体风险困境范式中，参与者以 6 人为一组进行互动，每个成员在每一轮中都需要从他们的私人账户(40 个实验货币单位)中选择向公共账户贡献 0、2 或 4 个实验货币单位。实验一共进行 10 轮，最后，如果所有参与者的贡献总和高于或等于 120 个实验货币的集体目标，那么所有参与者将保留其贡献外的剩余部分。否则，所有参与者将以 90% 的概率失去剩余禀赋，面临一无所获的结果，这代表了该困境下的集体风险(Milinski, Röhl, & Marotzke, 2011)。

也有研究者依据不同的实验需求，对上述范式中的设定进行修改，如更改参与者的初始禀赋用以探究贫富差距对合作水平的影响(Milinski et al., 2011)；更改风险发生的概率以及损失的大小来探究不同的风险对合作水平的影响(Tavoni et al., 2011)；不告知参与者所需要达到的最终贡献或者实验结束的时间以探究未来不确定性对于决策的影响(Domingos et al., 2020)。

3.3. 仿真模拟研究

集体风险困境的理论仿真模拟采用计算机编程的方法建立模型，通常会涉及到一些基本参数，比如：群体规模，群体中每个个体的初始禀赋，需要达到的最终总贡献(阈值)，合作者向公共资源库中贡献的货币数量，以及搭便车者向公共资源库中贡献的货币数量等。在设置了这些基本参数后，即可采用不同的理论模型对人们的行为进行预测。

3.3.1. 进化博弈论模型

早期研究者大多使用进化博弈论(evolutionary game theory, EGT)对集体风险博弈进行理论研究，这种方法用博弈论的思维模式设计群体行为，模仿定义个体行为随时间的变化。进化博弈理论针对一个由博弈参与者组成的群体，在博弈的过程中，每个参与者随机相遇，并采取各自的固定策略。每一次随机相遇，参与者都会进行两两博弈，博弈结束后得到各自的收益。在此基础上，比较自己与对手的收益，收益越大代表着个体能够成功繁衍的概率就越大，个体即可依据收益比较的结果进行策略更新。

在进化博弈论的框架下，较为简单的设计是在这些模拟中人们只有有限的行为策略：合作或背叛。博弈参与者每个人都有初始禀赋。合作者贡献一部分禀赋，而搭便车者则不贡献(Santos & Pacheco, 2011)。也有一些研究者通过更为复杂的参数设定模拟个体的行为策略，比如以上一轮总贡献是否达到阈值为标准，决定本轮贡献者的贡献(Abou Chakra & Traulsen, 2012)；在模型中加入惩罚者这一角色并定义惩罚者每轮的惩罚花费，以此来评估外部制裁机构在风险存在时的作用(Vasconcelos, Santos, & Pacheco, 2013)。这些理论模型不仅证实了行为实验的结果，还允许研究者们将分析扩展到任意大小的群体规模、风险感知和群体网络结构之中。

3.3.2. 强化学习模型

对于集体风险困境，既可以从静态视角来研究，也可以从动态视角来研究。静态视角通常假设参与者是理性的，并且知道所有可能的策略以及各自的贡献和收益结果。然而，实验研究表明，人类决策过程中并不总是做出理性的选择(Erev & Roth, 2014)，而是会根据博弈中获得的经验调整他们的策略。强化学习(Reinforcement learning, RL)提出了分析决策动态的新角度，已被证明可以准确地模拟几种社会困境中的人类行为。在过去的几年里，强化学习模型迅速发展，并且专门开发了几种变体来研究社会困境中的合作(Jacq et al., 2020)。有研究者受此启发，采用独立的强化学习(个体只能观察到自己的贡献行为和收益)研究在风险多样性和异质性博弈中出现的集体风险困境(Merhej et al., 2022)。与此相关的，还有基于人口的强化学习模型(Population-based Learning, PBL)，该模型主要参照 Roth-Erev 和 Q-learning 的简单个体

学习方案来更新行为,从而可以得到更广泛的混合策略,并允许更大群体规模的模拟参与者学习最适合的行为策略来解决集体风险困境问题(Domingos et al., 2021)。

3.4. 相关影响因素

3.4.1. 风险损失比例和发生概率

在集体风险困境实验中,常常设定一旦风险发生,参与者将会损失所有财富, Milinski 等(2008)在此基础上,评估了不同的风险发生概率如何影响人类的决策。他们发现,大多数面临高风险的小组(风险发生概率为 90%)成功了,而大多数面临低风险的小组(风险发生概率为 10%)失败了。Domingos 等(2020)以 90%的失败风险概率重复了该处理,发现了类似的结果。还有研究者发现,当风险损失比例越大时,群体成功达成目标的概率更高(Dannenberget al., 2015)。使用模型模拟可以选取更广泛的风险概率变化,并结合不同的风险损失比例,模型模拟的研究结果也表明,群体只能在高风险发生概率时进行协调合作(Abou Chakra & Traulsen, 2012);风险损失概率越大,比例越高,越有利于合作(Wang et al., 2009)。

3.4.2. 群体异质性

在集体风险困境的研究中,群体异质性主要可分为风险感知异质性、禀赋异质性和损失分配异质性。个体对于相同的风险会产生不同的主观感知,这会导致合作水平的差异,因此研究者认为风险感知异质性是重要的研究因素。实验结果表明,高风险感知异质性会导致群体成功率显著降低,而且随着这种异质性的增加,高风险感知人群的合作水平会增加,低风险感知人群的合作水平会降低(Merhej et al., 2022)。禀赋不平等的提出主要是为了研究贫富差异对于合作的影响,因此在相关实验中,初始禀赋高的代表富有的人或国家,初始禀赋低的代表贫穷的人或国家。研究结果表明,当穷人和富人各占一半时,成功的可能性比全部是穷人的更高,但低于全部都是富人的情况,这可能是因为三种处理条件下群体总资产不同,因而面对同样的贡献总目标会产生不同的难度(Milinski et al., 2011)。而当不同条件下群体总资产相同时,研究发现禀赋异质性不利于群体合作,但是这种不利影响可以通过沟通机制减弱,因为沟通使得富人更愿意贡献较多的资产(Tavoni et al., 2011)。也有研究者认为应当将禀赋异质性和损失分配异质性结合起来考虑,因为不同贫富情况的个体的抗风险能力往往是不同的。在此基础上,研究发现在禀赋不平等的情况下,穷人面临更大风险损失的时候往往合作更容易崩溃(Burton-Chellew, May, & West, 2013)。

3.4.3. 不确定性

在集体风险困境的研究中,不确定性主要可分为时间不确定性和阈值不确定性。时间不确定性指的是最终发生风险的轮次是不确定的,研究发现这种不确定性会增加早期和极端的贡献,高贡献行为会在博弈前半部分随着不确定性的增高而增加,这说明了在时间不确定的情况下尽早做出贡献和不拖延的重要性(Domingos et al., 2020)。有研究者采用模型模拟的方法,建立了可以连续贡献的集体风险困境,研究了时间不确定性对于合作成功率的影响,并得出结论,在这种不确定性下,最佳策略是立即做出贡献行动,即尽早贡献可以提高合作成功率(Abou Chakra et al., 2018)。因此,面对高时间不确定性,人们需要尽早做出高贡献才能提高成功率。

阈值不确定性是指最终需要收集的总金额是不确定的,这种类型的不确定性存在于许多集体风险情景中,例如气候谈判,其中温室气体减排的确切值是未知的(Barrett, 2013; Dannenberg et al., 2015)。研究表明当阈值存在不确定性时,群体成就也会显著降低(Dannenberget al., 2015)。此外,研究者也表明,高阈值不确定性也将这种博弈从避免灾难的合作困境转变为经典囚徒困境(Barrett & Dannenberg, 2012)。因此,在集体风险困境中,阈值不确定性往往会对群体合作产生负向影响。

3.4.4. 群体规模

在集体风险困境中，参与者被组织成一定规模的小组，这就涉及到了群体规模的影响。在一项计算机模拟集体风险困境的研究中，研究者发现较小的群体规模能够增加合作成功的可能性(Santos & Pacheco, 2011)。同样的，有研究者在行为实验中设置了三种不同的群体规模(3人，7人，11人)，也发现在11人为一组的条件下群体成功率最低(Santos & Pacheco, 2011)。总的来说，研究表明，增大群体规模不利于合作，因为更大的团队更难协调，所以大规模的群体整体成功率比小规模群体更低。

3.4.5. 人为干预措施

目前集体风险困境研究涉及到的人为干预措施主要有金钱激励和沟通机制。研究表明，在具有共同风险的多人博弈中，以惩罚或奖励形式出现的金钱激励是增加合作和目标达成的有效机制(Dong et al., 2021)。而且当奖励和惩罚联合使用时，奖励有利于一开始形成较高贡献趋势，而惩罚有利于在后续阶段维持合作(Góis et al., 2019)。对于沟通机制，有研究发现当参与者可以针对自己的贡献与成员进行沟通时，团队成功达成目标的可能性更大(Askari, Gordji, & Park, 2019)；当参与者可以就团队贡献应该达到的数目进行沟通时，也可以促进协调，更有利于合作目标达成(Dannenberget al., 2011)。

4. 总结与展望

本文通过文献梳理，归纳了社会困境的概念以及常见范式，进而引出集体风险困境的概念，并对集体风险困境常用的研究方法和相关影响因素进行了整理。

集体风险困境作为公共物品博弈的一种变式，能够更准确的描述人类当下面临的气候危机、疾病流行等问题。综上所述，虽然目前已经对一些影响因素进行了研究。但在社会困境合作领域中的许多其他机制是否有利于克服集体风险困境还有待考察，比如：亲缘关系、直接互惠和间接互惠机制、增加信任和群体认同等(Bochet et al., 2006; Murphy & Ackermann, 2014; Nowak & Sigmund, 2005; Okada, 2020)。此外，在解决 COVID-19 大流行或气候变化等全球危机时，有许多研究者主张可以使用人工智能来帮助解决协调合作问题(Oliveira et al., 2021)。未来的研究可以在集体风险困境中引入不同行为策略的人工智能，与真实的个体进行交互，探究包含人工智能和预定义决策过程的混合群体能否取得更大的集体成功。同时，也可关注真实个体在面临集体风险困境时决策的心理动机以及神经机制，这将有助于更全面地理解人类与人工智能的差异性和互补性。

参考文献

- Abou Chakra, M., & Traulsen, A. (2012). Evolutionary Dynamics of Strategic Behavior in a Collective-Risk Dilemma. *PLoS Computational Biology*, 8, e1002652. <https://doi.org/10.1371/journal.pcbi.1002652>
- Abou Chakra, M., Bumann, S., Schenk, H., Oshlies, A., & Traulsen, A. (2018). Immediate Action Is the Best Strategy When Facing Uncertain Climate Change. *Nature Communications*, 9, Article No. 2566. <https://doi.org/10.1038/s41467-018-04968-1>
- Askari, G., Gordji, M. E., & Park, C. (2019). The Behavioral Model and Game Theory. *Palgrave Communications*, 5, Article No. 57. <https://doi.org/10.1057/s41599-019-0265-2>
- Balliet, D., & Lange, P. (2013). Trust, Punishment, and Cooperation across 18 Societies A Meta-Analysis. *Perspectives on Psychological Science*, 8, 363-379. <https://doi.org/10.1177/1745691613488533>
- Barrett, S. (2013). Climate Treaties and Approaching Catastrophes. *Journal of Environmental Economics and Management*, 66, 235-250. <https://doi.org/10.1016/j.jeem.2012.12.004>
- Barrett, S., & Dannenberg, A. (2012). Climate Negotiations under Scientific Uncertainty. *Proceedings of the National Academy of Sciences*, 109, 17372-17376. <https://doi.org/10.1073/pnas.1208417109>
- Bochet, O., Page, T., & Putterman, L. (2006). Communication and Punishment in Voluntary Contribution Experiments. *Journal of Economic Behavior & Organization*, 60, 11-26. <https://doi.org/10.1016/j.jebo.2003.06.006>

- Burton-Chellew, M. N., May, R. M., & West, S. A. (2013). Combined Inequality in Wealth and Risk Leads to Disaster in the Climate Change Game. *Climatic Change*, *120*, 815-830. <https://doi.org/10.1007/s10584-013-0856-7>
- Chaudhuri, A. (2011). Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature. *Experimental Economics*, *14*, 47-83. <https://doi.org/10.1007/s10683-010-9257-1>
- Chen, X., & Szolnoki, A. (2018). Punishment and Inspection for Governing the Commons in a Feedback-Evolving Game. *PLoS Computational Biology*, *14*, e1006347. <https://doi.org/10.1371/journal.pcbi.1006347>
- Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., & Tavoni, A. (2011). *Coordination under Threshold Uncertainty in a Public Goods Game*. ZEW-Centre for European Economic Research Discussion Paper (11-065). <https://doi.org/10.2139/ssrn.1983266>
- Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., & Tavoni, A. (2015). On the Provision of Public Goods with Probabilistic and Ambiguous Thresholds. *Environmental and Resource Economics*, *61*, 365-383. <https://doi.org/10.1007/s10640-014-9796-6>
- Dawes, R. M. (1980). Social Dilemmas. *Annual Review of Psychology*, *31*, 169-193. <https://doi.org/10.1146/annurev.ps.31.020180.001125>
- Domingos, E. F., Grujić, J., Burguillo, J. C., Kirchsteiger, G., Santos, F. C., & Lenaerts, T. (2020). Timing Uncertainty in Collective Risk Dilemmas Encourages Group Reciprocation and Polarization. *Iscience*, *23*, Article ID: 101752. <https://doi.org/10.1016/j.isci.2020.101752>
- Domingos, E. F., Grujić, J., Burguillo, J. C., Santos, F. C., & Lenaerts, T. (2021). Modeling Behavioral Experiments on Uncertainty and Cooperation with Population-Based Reinforcement Learning. *Simulation Modelling Practice and Theory*, *109*, Article ID: 102299. <https://doi.org/10.1016/j.simpat.2021.102299>
- Dong, Y., Ma, S., Zhang, B., Wang, W.-X., & Pacheco, J. M. (2021). Financial Incentives to Poor Countries Promote Net Emissions Reductions in Multilateral Climate Agreements. *One Earth*, *4*, 1141-1149. <https://doi.org/10.1016/j.oneear.2021.07.006>
- Dreber, A., & Nowak, M. A. (2008). Gambling for Global Goods. *Proceedings of the National Academy of Sciences*, *105*, 2261-2262. <https://doi.org/10.1073/pnas.0800033105>
- Erev, I., & Roth, A. E. (2014). Maximization, Learning, and Economic Behavior. *Proceedings of the National Academy of Sciences*, *111*, 10818-10825. <https://doi.org/10.1073/pnas.1402846111>
- Fehr, E., & Gächter, S. (1999). Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, *90*, 980-994. <https://doi.org/10.1257/aer.90.4.980>
- Fehr, E., & Gächter, S. (2002). Altruistic Punishment in Humans. *Nature*, *415*, 137-140. <https://doi.org/10.1038/415137a>
- Góis, A. R., Santos, F. P., Pacheco, J. M., & Santos, F. C. (2019). Reward and Punishment in Climate Change Dilemmas. *Scientific Reports*, *9*, Article No. 16193. <https://doi.org/10.1038/s41598-019-52524-8>
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior & Organization*, *3*, 367-388. [https://doi.org/10.1016/0167-2681\(82\)90011-7](https://doi.org/10.1016/0167-2681(82)90011-7)
- Hardin, G. (1968). The Tragedy of the Commons: The Population Problem Has No Technical Solution; It Requires a Fundamental Extension in Morality. *Science*, *162*, 1243-1248. <https://doi.org/10.1126/science.162.3859.1243>
- Jacq, A., Perolat, J., Geist, M., & Pietquin, O. (2020). Foolproof Cooperative Learning. In *Proceedings of the 12th Asian Conference on Machine Learning* (pp. 401-416). <https://proceedings.mlr.press/v129/jacq20a.html>
- Kollock, P. (1998). Social Dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology*, *24*, 183-214. <https://doi.org/10.1146/annurev.soc.24.1.183>
- Merhej, R., Santos, F. P., Melo, F. S., & Santos, F. C. (2021). Cooperation between Independent Reinforcement Learners under Wealth Inequality and Collective Risks. In *The Proceedings of the 20th International Conference on Autonomous Agents and Multi-Agent Systems* (pp. 898-906). http://web.ist.utl.pt/franciscocsantos/MyArticles/Merhejetal_AAMAS_2021.pdf
- Merhej, R., Santos, F. P., Melo, F. S., Chetouani, M., & Santos, F. C. (2022). Cooperation and Learning Dynamics under Risk Diversity and Financial Incentives. In *The Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems* (pp. 908-916). <https://ifaamas.org/Proceedings/aamas2022/pdfs/p908.pdf>
- Milinski, M., Röhl, T., & Marotzke, J. (2011). Cooperative Interaction of Rich and Poor Can Be Catalyzed by Intermediate Climate Targets. *Climatic Change*, *109*, 807-814. <https://doi.org/10.1007/s10584-011-0319-y>
- Milinski, M., Sommerfeld, R. D., Krambeck, H. J., Reed, F. A., & Marotzke, J. (2008). The Collective-Risk Social Dilemma and the Prevention of Simulated Dangerous Climate Change. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 2291-2294. <https://doi.org/10.1073/pnas.0709546105>
- Murphy, R. O., & Ackermann, K. A. (2014). Social Value Orientation: Theoretical and Measurement Issues in the Study of Social Preferences. *Personality and Social Psychology Review*, *18*, 13-41. <https://doi.org/10.1177/1088868313501745>

- Nockur, L., Pfattheicher, S., & Keller, J. (2021). Different Punishment Systems in a Public Goods Game with Asymmetric Endowments. *Journal of Experimental Social Psychology*, *93*, Article ID: 104096. <https://doi.org/10.1016/j.jesp.2020.104096>
- Nowak, M. A., & May, R. M. (1992). Evolutionary Games and Spatial Chaos. *Nature*, *359*, 826-829. <https://doi.org/10.1038/359826a0>
- Nowak, M. A., & Sigmund, K. (2005). Evolution of Indirect Reciprocity. *Nature*, *437*, 1291-1298. <https://doi.org/10.1038/nature04131>
- Okada, I. (2020). A Review of Theoretical Studies on Indirect Reciprocity. *Games*, *11*, Article No. 27. <https://doi.org/10.3390/g11030027>
- Oliveira, R., Arriaga, P., Santos, F. P., Mascarenhas, S., & Paiva, A. (2021). Towards Prosocial Design: A Scoping Review of the Use of Robots and Virtual Agents to Trigger Prosocial Behaviour. *Computers in Human Behavior*, *114*, Article ID: 106547. <https://doi.org/10.1016/j.chb.2020.106547>
- Osama, T., Pankhania, B., & Majeed, A. (2020). Protecting Older People from COVID-19: Should the United Kingdom Start at Age 60? *Journal of the Royal Society of Medicine*, *113*, 169-170. <https://doi.org/10.1177/0141076820921107>
- Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511807763>
- Parks, C. D., Joireman, J., & Van Lange, P. A. M. (2013). Cooperation, Trust, and Antagonism: How Public Goods Are Promoted. *Psychological Science in the Public Interest*, *14*, 119-165. <https://doi.org/10.1177/1529100612474436>
- Rondeau, D., Poe, G. L., & Schulze, W. D. (2005). VCM or PPM? A Comparison of the Performance of Two Voluntary Public Goods Mechanisms. *Journal of Public Economics*, *89*, 1581-1592. <https://doi.org/10.1016/j.jpubeco.2004.06.014>
- Santos, F. C., & Pacheco, J. M. (2011). Risk of Collective Failure Provides an Escape from the Tragedy of the Commons. *Proceedings of the National Academy of Sciences*, *108*, 10421-10425. <https://doi.org/10.1073/pnas.1015648108>
- Sigmund, K., De Silva, H., Traulsen, A., & Hauert, C. (2010). Social Learning Promotes Institutions for Governing the Commons. *Nature*, *466*, 861-863. <https://doi.org/10.1038/nature09203>
- Tavoni, A., Dannenberg, A., Kallis, G., & Löschel, A. (2011). Inequality, Communication, and the Avoidance of Disastrous Climate Change in a Public Goods Game. *Proceedings of the National Academy of Sciences*, *108*, 11825-11829. <https://doi.org/10.1073/pnas.1102493108>
- Tilman, A. R., Plotkin, J. B., & Akçay, E. (2020). Evolutionary Games with Environmental Feedbacks. *Nature Communications*, *11*, Article No. 915. <https://doi.org/10.1038/s41467-020-14531-6>
- 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. <https://doi.org/10.1016/j.obhdp.2012.11.003>
- Vasconcelos, V. V., Santos, F. C., & Pacheco, J. M. (2013). A Bottom-Up Institutional Approach to Cooperative Governance of Risky Commons. *Nature Climate Change*, *3*, 797-801. <https://doi.org/10.1038/nclimate1927>
- Wang, J., Fu, F., Wu, T., & Wang, L. (2009). Emergence of Social Cooperation in Threshold Public Goods Games with Collective Risk. *Physical Review E*, *80*, Article ID: 016101. <https://doi.org/10.1103/PhysRevE.80.016101>