

社会规范对合作影响机制的研究¹

——基于内化效应和社会认同效应的分析

崔驰² 安露露³ 戴明⁴

摘 要 人们的合作行为受到社会规范的驱使,这一观点近期得到了学术界的证实。本文设计了一个在遵守社会规范和个人利益之间权衡的社会规范实验,借此将被试者分为遵守社会规范组和不遵守社会规范组,进而分解内化效应和社会认同效应。研究发现:(1)内化效应与社会认同效应都可以从整体上有效提高社会合作水平,但是在维持高水平合作方面的作用有限;(2)随着社会规范遵守程度的增强,社会认同效应在促进社会合作方面的主导作用逐渐被内化效应取代;(3)内化效应在维持高水平合作方面的作用随社会规范遵守程度的增强逐渐凸显,社会认同效应则在降低平均合作率衰减速度方面表现得更好,但短期内仍无法扭转平均合作率高速下降的趋势。本文利用实验室实验为社会规范促进社会合作的研究提供了一种新的视角。

关键词 社会规范;合作;社会认同效应;内化效应;公共品博弈

0 引言

经济社会的正常运转和社会合作的有效进行都离不开社会规范(social norms)的驱使和约束。在人类社会的演化发展过程中,社会规范作为一种非正式制度对人们的反社会行为——如欺骗(Bicchieri et al., 2019)、歧视(Barr et al., 2018)、贪污腐败(Banerjee, 2016; 张新超等, 2016; Muthukrishna et al.,

1 本研究获得国家社会科学基金年度项目“社会经济地位对儿童成长的综合影响及其破解的实地实验研究”(项目编号:22BJL082);教育部人文社会科学青年基金项目“社会规范影响合作水平的机制研究”(项目编号:20YJCZH013)的资助。作者非常感谢两位审稿专家宝贵的修改意见,同样感谢荷兰马斯特里赫特大学经济系 Alexander Vostroknutov 教授、瑞士苏黎世大学经济系 Roberto Weber 教授、澳大利亚昆士兰大学经济系 Daniel Zizzo 教授、新加坡南洋理工大学经济系包特教授对本文提出的宝贵意见,同时非常感谢我的妻子和我相亲时的言行为本研究提供了最初的灵感。

2 崔驰,东北师范大学经济与管理学院副教授,E-mail:cuic354@nenu.edu.cn。

3 安露露,北京师范大学经济与工商管理学院硕士研究生,E-mail:202021030002@mail.bnu.edu.cn。

4 戴明,海德堡大学经济与社会科学学院博士研究生,E-mail:daiming9999@hotmail.com。

2017)以及逃税行为(Hallsworth et al., 2017)——起到了较强的约束作用,进而有利于经济社会的稳定运行。与此同时,社会规范也时刻影响人们的亲社会行为,例如利他偏好(Krupka and Weber, 2013; Gächter et al., 2017a; 崔驰等, 2019, 2022)、公平偏好(Bicchieri and Chavez, 2010; 崔驰和路智雯, 2018; 戴明, 2020)、互惠(Gächter et al., 2013)以及人际交往中的信任(Bicchieri et al., 2019)等,进而促进了社会合作效率的提升。

从现有的研究来看,有学者认为社会规范可以直接作用于人们心理层面的认知活动(Ohtsuki et al., 2009),使人们在作出行为决策之前,考虑其行为是否符合社会规范并将这一过程内化,在这里我们将其称为“内化社会规范”。内化社会规范是人类社会长期演变发展的产物,它可以通过两种渠道影响人们的亲社会行为,如合作等。一方面,社会规范对个人行为的约束力可能源于人们对其社会形象的普遍关注,即为了达到某个荣誉性目标或者避免遭受污名惩罚(Bénabou and Tirol, 2006)。例如,Andreoni and Bernheim(2009)发现完全均等分配(五五分)的规范并不能解释大部分实验结果,人们可能会为了迎合他人的公平偏好而决定自己的行为模式。另一方面,社会规范也可能通过影响个人的认知水平,使人们形成一种内在遵守社会规范的本能(Frey et al., 2010; Gavrillets and Richerson, 2017),进而潜移默化地影响人们的行为决策,尤其是合作行为。与社会规范的内化相关,内化效应则是指以内化社会规范程度高低为标准分组后,社会合作效率的相应变化。近些年来,越来越多的实验研究涉及社会规范的内化对人们合作行为的影响。Kimbrough and Vostroknutov(2016)在Krupka and Weber(2009)的研究基础上通过交通灯实验(traffic lights game)在被试者未知的情况下将其分为遵守社会规范组和不遵守社会规范组,研究了内化社会规范对人们合作行为的影响,并证明了规范敏感度较强(内化社会规范较强)的个体更加倾向于选择合作。随后的多数研究都得到了类似的实验结论,即社会规范内化的存在确实极大地促进了社会合作,内化效应也是显著正向的(Gächter et al., 2017b; Barr et al., 2018; Chang et al., 2019; Rössler et al., 2019)。

此外,社会认同(Social Identity)可能是社会规范促进合作的另外一种途径。Tajfel等人最早于20世纪70年代就提出了社会认同理论,该理论逐渐成为解释群体间行为的最具有影响力的理论之一。该理论认为个体对群体的认同是群体行为的基础,其主要作用原理是通过增强个体对所处群体的社会偏好,引起群体偏见,进而使其作出更加有利于群体的行为,促进群体合作水平的提升(Tajfel, 1974; Jetten et al., 1996; Eckel and Grossman, 2005; Goette et al., 2006; Chen and Li, 2009; 周业安和王一子, 2016)。现实生活中,人们往往将他人行为是否合乎自身认同的规范作为判别可否与其展开合作的标准,即合作团队内成员之间存在一定程度的社会认同,而这种认同的产生显然源于社会规

范。由此,理论上可以通过激发特定的社会认同影响人们的亲社会行为,进而提升社会合作效率。

随着通信技术的革新发展,人们的交际圈子快速扩大,由社会认同推动的传统熟人合作模型也逐渐转向现代化的由社会规范推动的大规模陌生人合作模式。此外,随着全民受教育水平和道德素质的提高,个人价值观逐渐向社会规范靠拢(Tabellini,2008),遵守社会规范的意愿也越加强烈。那么,随之而来的问题是随着个人规范意识的增强,社会合作对社会认同和社会规范内化的相对依赖度是否会发生变化?然而,现有研究大多是单一地探究其中一项对合作的促进效果,缺乏将两种效应直接分离的研究,其原因可能是:①传统新古典研究范式的“惯性”,将社会规范视为既定的社会条件而忽视了对它的研究;②过去的文献由于研究方法和研究工具的缺失(Manski,2000),导致研究者很难从定量的角度对社会规范促进合作的内化效应和社会认同效应进行直接分离。而随着20世纪80年代行为与实验经济学的兴起,逐渐打破了研究方法和工具的瓶颈,故本文将采用实验方法研究社会规范和社会认同对合作的影响效应,即社会认同效应和内化效应,并探究在不同的社会规范遵守程度中,社会合作对两种效应的依赖度差异。

本文利用社会规范实验和公共品博弈实验的组合简化现实情境研究社会规范①对社会合作的整体影响机制,分离社会认同效应和内化效应。本文后续内容安排如下:第1部分是文献综述,第2部分是实验设计,第3部分是实验结果和数据分析,最后是结论与讨论。

1 文献综述

近年来,由于行为与实验经济学的兴起,越来越多的研究者采用公共品博弈实验研究人们的合作行为(Santos et al.,2008;Szolnoki and Perc,2010;Brekke et al.,2011;陈叶烽和何浩然,2012;连洪泉等,2013;周业安等,2014;Chen et al.,2015)。根据传统经济学的理性人假定,个体在公共品博弈中的最优选择是“搭便车”,但是大量的实验研究结果对这一理论预测提出了挑战,即发现大多数人都具有很强的社会偏好,愿意为了实现合作而支付私人成本(Andreoni,

①从严格的定义上看,不同学科(从经济学到社会学再到心理学)对社会规范界定描述各不相同,但是其核心内容大同小异。规范通常被理解为在群体内占主导地位的行为范式,被共同理解和支持,并且通过群体中的社交互动得以维持(Ostrom,2000)。在这里我们将社会规范解释为一种特殊的行为规则,指其在特定情况下向特定人群规定或禁止某种行为(Jackson,1965;Cialdini,2003;Lapinski and Rimal,2005)。如果个人在以下条件更偏向于遵循规则,则该规则满足社会规范的性质:①人们认为有足够多的其他人也遵守它(经验期望);②人们认为有足够多的其他人也应该遵守该规则,并可能愿意制裁他人的违反行为(规范性期望)(Bicchieri,2005)。

1988,1995;Keser and van Winden,2000;Fischbacher et al.,2001;Hauert et al.,2002;Rege and Telle,2004)。

近期,有研究者发现人们的亲社会行为,尤其是合作行为,不仅受到社会偏好的影响,也受到社会规范的驱使(Krupka and Weber,2009,2013;Kessler and Leider,2012;Kimbrough et al.,2014;关宏宇等,2015)。例如, Ellingsen and Johannesson(2008)认为大多数人对自身的社会形象非常敏感,而与社会形象相关的自豪感也是个体亲社会行为的源泉。社会规范作为维持社会形象的基础,可以通过间接提升个人行为的自豪感强化人们的亲社会行为。Krupka and Weber(2009)直接研究了社会规范对人们亲社会行为的影响,发现社会规范会强化人们的亲社会行为。但他们仅仅是比较了第三方遵守社会规范的信息(information treatment^①)和引起被试者注意社会规范的环境或线索(focusing treatment^②)对个体行为的影响,并没有对形成机制进行深入探究。Kimbrough and Vostroknutov(2016)根据被试者遵守社会规范的程度分组来研究社会规范的内化效应对合作的促进效果。他们通过交通灯实验直接分离了个体遵守社会规范的程度,然后依据实验结果将被试者分成遵守社会规范组和不遵守社会规范组,并让其参加有限期^③固定搭配的公共品博弈实验,结果发现遵守社会规范组的合作率显著高于不遵守社会规范组,且遵守社会规范组的合作率保持了较强的稳定性,在一定程度上证实了社会规范内化对合作的积极影响。随后,越来越多的学者开始对社会规范内化进行研究,但无论是实验室实验(Gächter et al.,2017b;Barr et al.,2018;Chang et al.,2019)还是田野实验(Rössler et al.,2019),研究结果都支持 Kimbrough and Vostroknutov(2016)的研究结论。

除内化效应之外,社会规范也可能通过在群组内部形成特定的社会认同,进而影响群组成员的合作行为。中国自古以来就是礼仪之邦,人们在社会交往过程中非常重视社会规范,常常将社会规范遵守程度视为一种与人相处的认知聚点,也就是俗语中的“人以类聚”。Calvert(1995)认为“识别或者创造聚点是保证参与者成功合作的一种方式”,而产生聚点的同时也伴随着社会认同的出现。但是,现有研究关于此机制的探讨相对缺乏,仅有的研究主要是集中在发现社会认同对社会合作具有促进作用上。他们认为社会认同(social identity)或身份认同感可以增强成员对群体的社会偏好,导致群内偏见,从而使人们作出

① information treatment 是指被试者可以看到或了解其他人是否遵守社会规范,这里强调了被试者可以直接获取他人如何遵守社会规范的相关信息。

② focusing treatment 是指通过某些社会规范的环境线索引起被试者产生相应的社会规范行为,这里被试者没有直接获取相关社会规范的信息,而是被启发出社会规范行为。

③ 在 Kimbrough and Vostroknutov(2016)公共品博弈实验中,实验员没有明确告诉被试者参加公共品博弈实验的具体轮数,因此,对于被试者来说无法准确预期博弈进行的轮数,但实际上该博弈进行了10轮。

更加利于群体的行为,促进群体合作水平的提升(Tajfel et al., 1971; Tajfel, 1974; Jetten et al., 1996; Terry et al., 1999; Akerlof and Kranton, 2000; Eckel and Grossman, 2005; Chen et al., 2010; Benjamin et al., 2010; Chen and Li, 2009)。如 Eckel and Grossman(2005)通过构造群内身份,研究社会认同强度对公共品博弈实验中被试者合作行为的影响,发现社会认同可以很好地提升组内成员的合作水平,同时还发现社会认同显著提高了组员的捐献水平。Guala et al. (2013)利用囚徒困境实验改变参与者的信念,让他们产生群体归属感和身份认同感,发现群内成员比群外成员捐献的更多并且合作率更高。这些研究证据都表明了社会认同确实可以有效提高社会合作效率(Solow and Kirkwood, 2002; Goette et al., 2012)。

基于以上两方面,本文借鉴了 Kimbrough and Vostroknutov(2016)的实验设计,通过引入社会规范实验将被试者按照其遵守社会规范的程度进行分组探究不同组被试者在公共品博弈实验中的合作行为,引入不同规则使实验情境更加贴近现实生活,进而试图分析比较内化效应和社会认同效应在不同社会规范遵守程度下促进社会合作的相对作用。文章的主要贡献在于:①采用实验经济学方法研究社会规范影响合作的效应,丰富了国内对社会规范作用的研究;②改进了 Kimbrough and Vostroknutov(2016)的研究设计,首次尝试将社会规范和社会认同相结合以探究这两种效应在不同的社会规范遵守程度下促进社会合作的相对作用,在一定程度上丰富了 Krupka and Weber(2009)以及 Kimbrough and Vostroknutov(2016)的研究成果,并且实验设计更加贴近现实,提高了以往研究的外部效应。

2 实验设计

本实验在东北师范大学经济管理实验室和数学与统计学院教育部重点实验室进行。我们在长春市多所大学张贴招募海报,并从招募的实验志愿者中随机抽取 272 名被试者,这些被试者分别来自社会科学、理学、工学等各个专业。本次实验采用 Z-tree 实验软件(Fischbacher, 2007)进行实验程序编程,整个实验过程都是在计算机上完成的,每个实验局的被试者数量为 16 人,耗时约 60 分钟。

实验由社会规范实验^①和公共品博弈实验两个环节组成。除此之外,为了控制条件性合作及其他因素对公共品博弈的影响,被试者在公共品博弈实验结

^① 本文的社会规范实验沿袭 Kimbrough and Vostroknutov(2016)的实验设计并对原设计的社会情境进行了改进。在 Kimbrough and Vostroknutov(2016)的研究设计中,采用了较为复杂的交通灯情境(简称“RF 实验”),但交通灯实验的现实情境可能会引起被试者的侥幸心理,即当马路上没有行人和车辆时,即使不遵守交通规则也不会对他人造成影响。为了排除这种侥幸心理的影响和便于被试者理解,我们将此情境简化为丢垃圾情境。而且,绿色环保也是近年来广受关注的热点话题,人们对此也更加敏感。

束后还完成了条件合作量表 (Keser and van Winden, 2000; Fischbacher et al., 2001) 和人口统计量数据的填写。社会规范实验的目的是通过构建遵守社会规范和个人利益之间冲突的决策情景, 使被试者面临两难的权衡, 进而根据实验结果将被试者分为遵守社会规范组和不遵守社会规范组。与此同时, 根据被试者在公共品博弈实验之前是否被告知具体的四人组分规则信息, 形成两个实验局即内化组和社会认同组。内化组的被试者不知晓分组信息, 社会认同组则被告知^①。然后, 探究不同实验局以及它们各自的遵守社会规范组和不遵守社会规范组在公共品博弈实验中的合作情况。最后, 通过与没有施加分组干预的基础组进行对比, 分析比较社会规范影响合作的两种效应, 即内化效应和社会认同效应。具体实验内容如下。

被试者首先参加社会规范实验, 然后参与固定搭配的 10 轮公共品博弈实验。在社会规范实验中, 每个被试者有 100 个小球, 并被明确告知小球代表“垃圾”, A 盒子代表“垃圾箱”, B 盒子代表“随地乱扔”^②, 而被试者的任务就是决定如何在 A 盒子(“垃圾箱”)和 B 盒子(“随地乱扔”)之间分配这 100 个小球。如果被试者将小球放进 A 盒子里, 那么其收益为每个小球 0.25 个筹码; 而放进 B 盒子里的收益为每个小球 0.4 个筹码, 社会规范实验的收益为两个盒子的收益之和。宣读完实验规则后, 实验员会明确地用指令性规范^③告诉被试者: “社会规范是人们应该将小球(“垃圾”)全部放进 A 盒子(“垃圾箱”)中”^④。社会规范实验结束后, 每个被试者都将看到自己在 A 盒子中的收益、B 盒子中的收益以及加总后的总收益(A 盒子和 B 盒子的收益之和)。之后进行第二环节, 被试者参加每四人一组的公共品博弈实验, 每个被试者的初始禀赋是 20 个筹码。然后, 告知被试者公共品博弈实验重复进行 10 轮, 四人组的人员构成在整个实

① 按照心理学定义, 社会认同就是“个体知晓他/她归属于特定的社会群体, 而且他(她)所获得的群体资格会赋予其某种情感和价值意义”, 因此我们通过告知被试者其所在小组的成员在遵守社会规范方面与自己表现相似, 进而激发被试者的社会认同感。

② Zizzo (2010) 提出实验者需求效应 (experimenter demand effects) 主要是由于恰当行为信息提示所引起的被试者的行为变化。因此, 在社会规范实验中, 我们并未明确提及将小球扔进 B 盒子(“随地乱扔”)里会造成环境污染以及其他可能带来的负效应, 其目的是尽可能减少需求效应的影响并显示出被试者真实的遵守社会规范的倾向性。

③ Cialdini 及其同事 (Cialdini et al., 1990; Reno et al., 1993) 区分了两种类型的规范, 即描述性规范和指令性规范。其中, 指令性规范涉及对个人行为的道德价值和社会标准判断, 即明确规定“什么是对的或错的”“人们应该做什么”或“什么行为是社会可接受的和有价值”。多数研究发现指令性规范对人们的行为具有更显著的影响 (Brauer and Chaurand, 2010), 且现实生活中的社会规范带有指令性规范的特征, 故本实验选用指令性规范。

④ 在社会规范实验中, 其实验结果可能会被解释为实验者需求效应 (experimenter demand effects), 即被试者因对实验者的假设进行推论而相应地改变其决策行为 (De Quidt et al., 2019)。Levitt and List (2007) 认为需求效应也可以解释为在亲社会行为方面许多明显的异质性, 且 Kimbrough and Vostroknutov (2016) 认为任何实验者的需求效应实际上都是我们试图衡量的规范依赖性的体现。因此, 我们认为社会规范实验可以反映出被试者对规范的遵守情况, 且依此分组存在一定的合理性。

验过程中不会发生改变。被试者可以选择向公共项目中投资0到20个筹码中的任何一个数量,如果被试者选择投资 X ,小组的其他三个成员分别投资 Y 、 A 和 B ,那么该被试者此轮获得的收益为 $20-X+0.4 \times (X+Y+A+B)$ 。在每一轮公共品博弈实验结束后,所有被试者都可以看到自己对公共项目的投资额(X),小组所有成员对公共项目的总投资额($X+Y+A+B$)、保留在自己私人账户中的筹码数($20-X$)、从公共项目中获得收益($0.4 \times (X+Y+A+B)$)和自己在该轮次中的总收益($20-X+0.4 \times (X+Y+A+B)$)。被试者在公共品博弈实验中获得最终收益为这10轮实验所挣得的筹码之和乘以0.1倍。实验结束后,被试者的最终收益为两实验收益之和再加上10元出场费。所有实验组都经过这两阶段的实验,但是分组方式和告知被试者的信息略有不同。根据这些不同分为了三个实验组,它们分别是:基础组(T1)、内化组(T2)和社会认同组(T3),其整体实验流程如图1所示。下面是对每个实验组的详细描述。

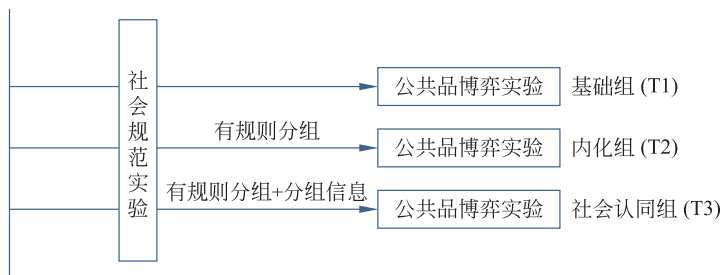


图1 实验流程图

基础组(T1):实验开始前,实验员告诉被试者将进行社会规范实验,向被试者发放实验说明书并解释实验流程,然后开始社会规范实验。社会规范实验完成后,向被试者发放公共品博弈实验的说明书并解释实验流程。在该环节电脑将被试者分为每4人一组,然后开始10轮人员构成固定化的公共品博弈实验。

内化组(T2):除了公共品博弈实验的分组方式与基础组不同外,其他实验设置均与基础组相同。即与基础组(T1)的唯一区别为在社会规范实验完成后,根据被试者在社会规范实验中的选择结果,将其分为遵守社会规范组和不遵守社会规范组^①,然后再进行公共品博弈实验。但是被试者并没有被告知这种分

^① 在本文的实验中,每个实验局共有16名被试者,我们根据被试者在社会规范实验中扔进A盒子(“垃圾箱”)的小球数量进行从大到小排序,扔进A盒子(“垃圾箱”)中数量最多的为序号1,最少的为序号16,并将被试按照该排列序号依次分为每四人一组,即1~4号为一组,5~8为一组等。其中,排名靠前的两组为遵守社会规范组,排名靠后的两组分为不遵守社会规范组。而另外一种分组方式是首先将被试者分为遵守社会规范组(即前8名被试),然后按照1,3,5,7为一组,2,4,6,8为一组的方式,同理不遵守社会规范组则是9,11,13,15为一组,10,12,14,16为一组。由于我们度量的是平均效应,故这两种方式相差不大。本文则采用前一种方式。

组方式,即他(她)们并不知晓自己所在四人组的成员特征信息(特指遵守社会规范方面的特征信息)。此外,在内化组中,被试者被分入遵守社会规范组还是不遵守社会规范组是由其自身是否遵守社会规范的特点所决定的,他(她)们并不知晓自己所在四人组属于遵守规范组还是不遵守规范组。我们将被试者在本组的合作行为与基础组的合作行为的差别界定为内化效应。

社会认同组(T3):此实验组的实验设置与内化组的不同点在于,被试者知道自己在公共品博弈中的分组是依据社会规范实验测度出的特征进行的。具体来说,被试者知晓自己所在四人组的成员特征信息(特指遵守社会规范方面的特征信息)。即,当完成社会规范实验后,实验员向被试者发放公共品博弈实验的说明书,向其解释实验流程,并强调公共品博弈实验的分组原则,即告知“在本实验中,和您同组的三位组员是和您同类型的实验者即在实验三^①中,他们放入A盒子中的小球个数和您放入的个数相近”,这样做的目的是激发组内成员之间的社会认同感。同时,为了检验分组规则是否有效地激发了组员之间的社会认同感,我们采用问卷形式测度受访者在知晓分组规则后(公共品博弈实验之前)其对同组成员和不同组成员的态度差异^②。同理,在社会认同组中,被试者被分入遵守社会规范组还是不遵守社会规范组是由其自身是否遵守社会规范的特点所决定的,他(她)们并不知晓自己所在小组属于遵守规范组还是不遵守规范组。因此,该组被试者的合作行为与基础组的合作行为的差别可以被解释为内化效应和社会认同效应的叠加。

3 实验结果分析

3.1 描述性统计

本实验于2018年5月、6月和10月、2019年4月、5月,以及2021年12月在东北师范大学经济管理实验模拟中心和数学与统计学院教育部重点实验室

① 第一个实验是采用量表测度被试者的风险偏好(Holt and Laury, 2002),第二个实验是借鉴Fischbacher and Föllmi-Heusi(2013)和Gächter and Schulz(2016)中的方法测度被试者的诚实水平,由于这两个实验与后续实验之间不存在直接的联系,所以我们认为它们对后续实验的影响是相当小的。

② 在问卷设计上,我们借鉴了Chen and Li(2009)的实验设计,即在受访者线上完成社会规范实验后告知分组规则(与他们分为一组的其他3位受访者是与他们投入A盒子(“垃圾箱”)中小球(“垃圾”)个数相近的受访者),之后参与4种分配情境下的第三方独裁者博弈实验,即作为第三方分别在两名其他受访者之间分配100个筹码。其中,两位其他受访者的身份组合分别是:(同组参与者1,同组参与者2)、(不同组参与者1,不同组参与者2)、(同组参与者1,不同组参与者2)以及(不同组参与者1,同组参与者2)。此外,为了避免单纯“组”概念的影响,我们还设计了没有社会认同激发的对照组,即受访者在完成社会规范实验后仅被告知他们将会与其他3位受访者分为一组。具体的问卷设计和结果详见附录1。

进行。实验被试者总数为 272 人,每场中每个实验局的被试者数量为 16 人,耗时约 60 分钟/场,被试者的平均收益为 35.5 元。

表 1 描述了三个实验局的基本统计量,其中包括每个实验局的被试者人数、社会规范实验中被试者投入 A 盒子中的小球数量及其标准差、公共品博弈实验中被试者的平均投资比例及其标准差。三个实验局的被试者人数分别为:基础组(T1)有 80 名被试者,内化组(T2)有 96 名被试者,社会认同组(T3)有 96 名被试者。基础组(T1)、内化组(T2)和社会认同组(T3)分别向 A 盒子中投入的小球数平均为 58.8、51.3 和 58.4,三个实验局的标准差基本上比较接近。基础组在公共品博弈中的平均投资比例为 31.3%,内化组为 34.1%,社会认同组为 43.5%。

表 1 三个实验组的基本统计量

实验组	被试数量	A 盒子中的平均小球数	公共品博弈实验的平均投资比例
基础组(T1)	80	58.8 (35.7)	0.313 (0.337)
内化组(T2)	96	51.3 (34.2)	0.341 (0.332)
社会认同组(T3)	96	58.4 (35.6)	0.435 (0.336)

注:三个实验局的平均投资比例的最小值均为 0,最大值均为 1;三个实验局 A 盒子的小球数最小值均为 0,最大值均为 100;括号内为标准差。

3.2 社会规范实验的数据分析

图 2 描述的是被试者在社会规范实验中投入 A 盒子里的小球数量的分布状况。我们发现基础组大约有 17.5%的被试者选择向 A 盒子(垃圾箱)投入 0 个小球,27.5%的被试者投入 100 个;内化组大约有 18.75%的被试者向 A 盒子投入 0 个小球,16.67%的被试者投入 100 个小球;社会认同组大约有 14.58%的被试者选择向 A 盒子投入 0 个小球,21.88%的被试者投入 100 个小球。表 2 是对基础组(T1)、内化组(T2)和社会认同组(T3)中 A 盒子里小球数进行 Wilcoxon 秩和检验(即 Wilcoxon Rank-sum Test 非参数检验)得到的结果。根据表 2 可知,在基础组(T1)、内化组(T2)和社会认同组(T3)中,任何两组之间的小球分布比例都无显著性差别,说明这三个实验局的被试者类型的分布是没有区别的。图 3 是内化组(T2)和社会认同组(T3)的遵守社会规范组和不遵守社会规范组在 A 盒子中的投球情况,可以发现遵守社会规范组的小球分布主要集中在 50~100 区间内,绝大部分人会向 A 盒子投入至少 50 个小球,而不遵守社会规范组的小球分布主要集中在 0~50 区间内,并且投入 A 盒子的小球数普遍都小于 50。这说明我们划分的遵守社会规范组的被试者倾向于牺牲自己的部

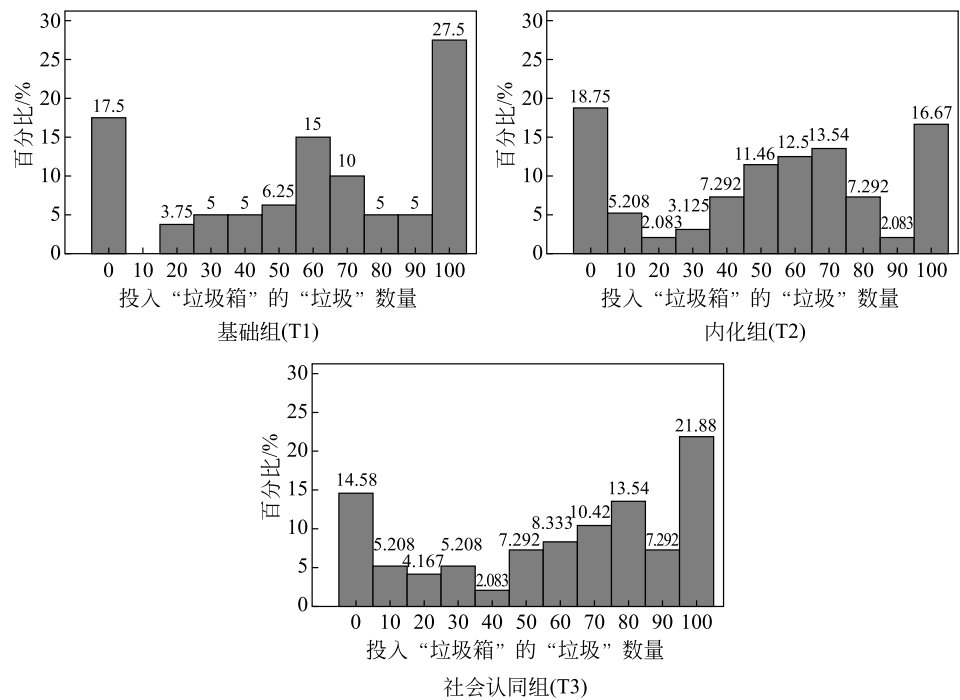


图 2 社会规范实验的小球分布(总体)

分收益从而将更多的小球放入 A 盒子(垃圾箱)中,而不遵守社会规范组的被试者则为了追求更多的私人利益将更多的小球放入 B 盒子(随地乱扔)中,这与现实生活中观察到的情况基本一致。

表 2 Wilcoxon Rank-sum Test 检验结果

实验局	原假设	z 值	p 值
基础组 vs 内化组	两组分布比例无显著性差异	-1.217	0.2237
基础组 vs 社会认同组	两组分布比例无显著性差异	-0.030	0.9761
内化组 vs 社会认同组	两组分布比例无显著性差异	-1.572	0.1160

3.3 公共品博弈实验的数据分析

本小节将分析社会规范和社会认同对合作的影响效应。由于内化组和社会认同组的分组方式是根据社会规范实验的结果将被试者分成了遵守社会规范组和不遵守社会规范组,因此本节先从总体上分析不同实验局在公共品博弈实验中的投资水平,然后再细分为遵守社会规范组和不遵守社会规范组进行讨论。

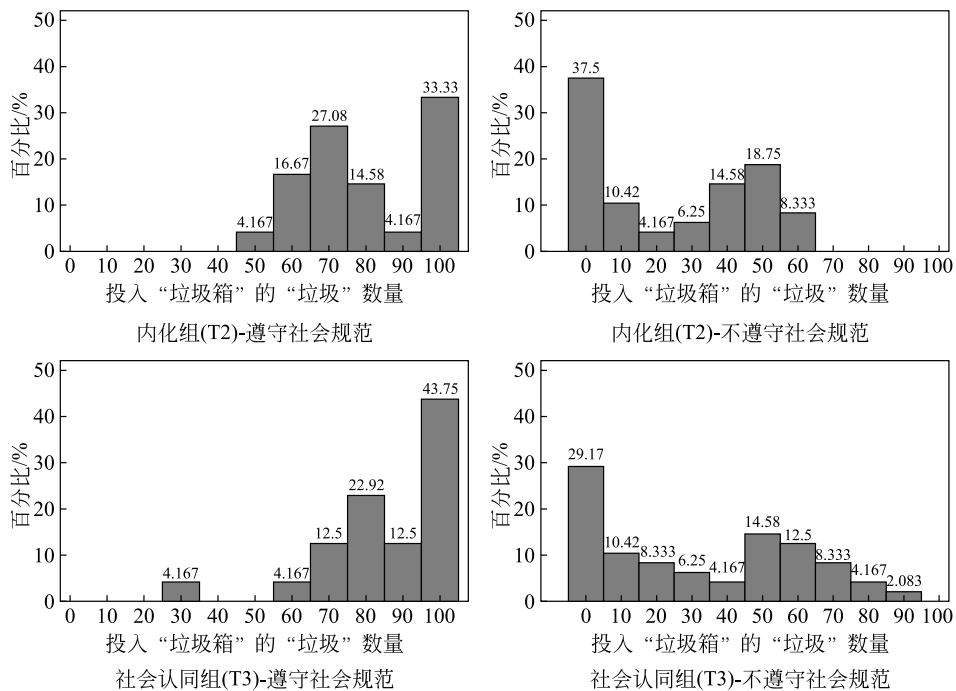


图3 社会规范实验的小球分布(分组)

3.3.1 公共品博弈实验的总体数据分析

图4描述了不同实验局的公共品博弈实验投资比例的分布情况,即基础组(T1)、内化组(T2)和社会认同组(T3)。总体上,多数被试者的投资比例集中在0.1及以下。具体来说,基础组投资比例小于0.1的被试者占比达到40.88%,为三个实验局之最,而投资比例大于0.5的仅占34.75%;内化组投资比例小于0.1的被试者占比为35.42%,投资比例大于0.5的占比为36.873%;社会认同组投资比例小于0.1的被试者占比为22.08%,而投资比例大于0.5的占比为49.17%。此外,在社会认同组和内化组中投资比例达到100%的被试者占比也明显高于基础组,依次是14.38%、10.52%和9.875%。显然,从各实验局投资比例分布的情况来看,社会认同组的合作率最高,内化组次之,基础组最差。

尽管在社会规范实验中已经证明了基础组、内化组和社会认同组中被试者并不存在明显差异,但为了保证实验结果的可比性和有效性,我们利用三组被试者在公共品博弈第一轮实验结果进行检验。正如前面实验设计中所提及的,在内化组中,被试者并不知晓公共品博弈实验的分组规则,因此如果不同组被试者之间不存在差异,内化组和基础组在第一轮中的投资水平应该不存在显著性差异,检验结果也证明了这一点,如表3所示: $z = -1.033$, $p = 0.3017$,并不

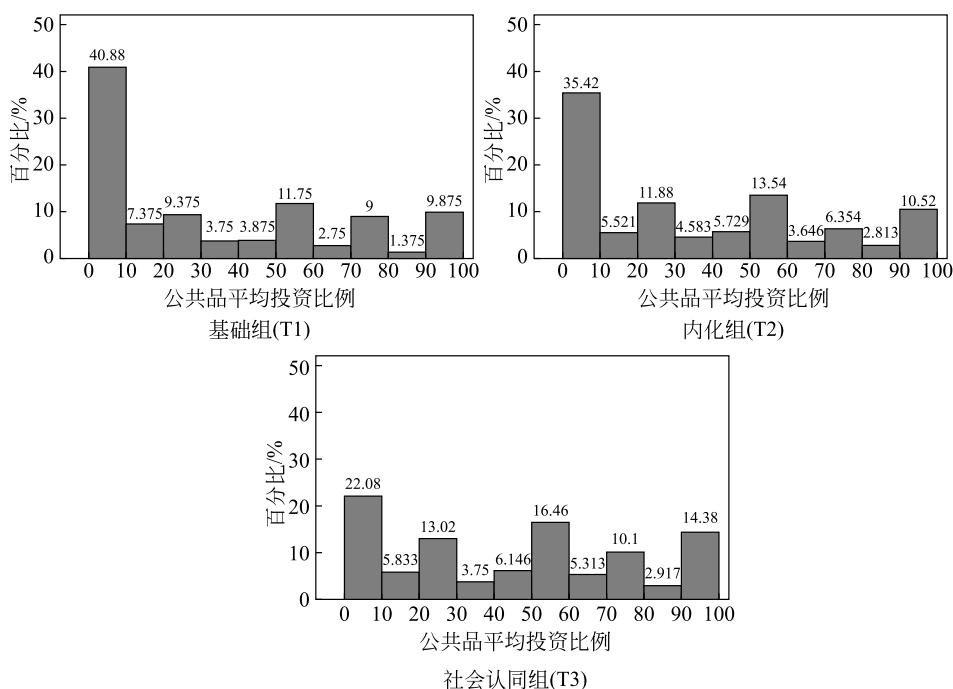


图4 公共品博弈实验的投资比例分布

能通过显著水平为1%的统计检验,即内化组和基础组在第一轮中的投资水平不存在显著性差异。

在正式分析内化效应^①和社会认同效应能否显著提升合作率之前,我们首先需要检验社会认同组中的分组规则是否有效激发了人们的社会认同感。我们对问卷结果进行了比较分析:发现相对于没有明确分组规则的受访者,知晓分组规则的受访者对组内成员表现得更加慷慨,即在第三方独裁者博弈中他们分配给组内成员显著更多的筹码(组内偏好),而分配给组外成员显著更少的筹码(组外歧视)。由此说明,社会认同中的分组规则设置确实能够有效激发被试的社会认同感(具体结果详见附录1)。

为探究内化效应和社会认同效应能否显著地提升合作率,我们利用Wilcoxon秩和检验对基础组(T1)、内化组(T2)和社会认同组(T3)总体平均投资水平进行两两检验,结果如表3所示。基础组(T1)和内化组(T2): $z = -2.112, p = 0.0347$,在5%的显著性水平下拒绝原假设,这说明基础组的平均投资水平显著小于内化组,即内化效应能够显著地提高合作水平。内化组(T2)和

^① 内化效应是指内化组与基础组的平均合作率差异;同理,社会认同效应是指社会认同与内化组的平均合作率差异。

社会认同组 (T3): $z = -6.455, p = 0.000$, 在 1% 的显著性水平下拒绝原假设, 这说明社会认同组 (T3) 的平均投资水平显著高于内化组 (T2) 的投资水平, 即证明了由社会规范产生的特定社会认同确实极大地促进了社会合作。基础组 (T1) 和社会认同组 (T3): $z = -8.051, p = 0.0000$, 在 1% 的显著性水平下拒绝原假设, 表明基础组的平均投资水平显著小于社会认同组, 这说明由社会规范引致的社会认同效应和内化效应显著提升合作水平。

表 3 Wilcoxon Rank-sum Test 检验结果

实验组	z 值 (第一轮)	p 值 (第一轮)	z 值 (总体)	p 值 (总体)
基础组 vs 内化组	-1.033	0.3017	-2.112	0.0347
基础组 vs 社会认同组	-2.372	0.0177	-8.051	0.0000
内化组 vs 社会认同组	-1.395	0.1631	-6.455	0.0000

接下来, 我们对内化效应和社会认同效应在维持社会合作方面的作用进行分析。图 5 展示的是基础组 (T1)、内化组 (T2) 和社会认同组 (T3) 合作率的变化趋势。从图 5 可以直观地看出, 基础组、内化组和社会认同组的投资比例都随着实验轮数的增加有明显的下降趋势, 这与以往的研究结果相一致。对各实验局进行期数效应检验发现, 基础组、内化组和社会认同组的投资比例确实随着实验轮数的增加而显著下降。其中, 社会认同组的下降趋势最为明显, 下降速度最快 ($b = -3.375, p = 0.000$), 内化组次之 ($b = -2.621, p = 0.000$), 基础组的下降速度最慢 ($b = -1.816, p = 0.003$), 这主要是由于基础组的整体投资率都很低, 下降空间有限。但我们也可以明显看到在内化效应和社会认同效应作用下, 内化组和社会认同组的平均合作率在大多数轮次中均高于基础组。

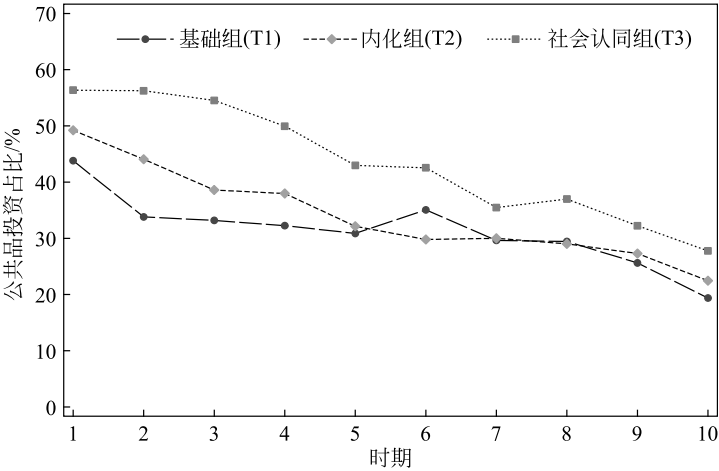


图 5 公共品博弈实验的投资趋势

为了进一步探究三组平均合作率随轮次衰减的速度变化,我们利用 Matlab 的 curve fitting tool 分别对三组公共品博弈实验的数据进行非线性拟合(具体拟合曲线和方程形式详见附录 2)并根据拟合效果选择最佳的最高次项和具体方程形式。根据拟合方程,我们可以求出各组在公共品博弈实验中的平均合作率随轮次变化的下降速度变化(加速度),进而探究内化效应和社会认同效应在改变平均合作率变化趋势惯性方面的角色,结果如表 4 所示。首先,对基础组而言,在第 1~5 期,其平均合作率的衰减加速度一直处于大于 0 的状态但不断减小,到第 5 期时的衰减加速度已经接近于 0,即公共品投资已经处于或接近于匀速下降。在第 5~10 期,其平均合作率的衰减加速度开始小于 0 且绝对值不断增大,即合作率水平的下降速度越来越快,这可能是受到公共品博弈的实验结束效应影响。其次,对内化组而言,其在前 6 期的合作率衰减速度不断减缓,后 4 期则逐渐加速,但是相较于基础组,其平均合作率的加速度绝对值都要更小,即内化效应可以在一定程度上降低衰减速度的变化。最后,对社会认同组而言,不同于基础组和内化组,其平均合作率在前 5 期加速衰减但加速度越来越小,而在后 5 期开始逐渐降低衰减速度且降低幅度越来越大,即加速度越来越大。但是,相较于基础组,其平均合作率的加速度绝对值都要更小,即内化效应和社会认同效应能够在一定程度上维持甚至降低平均合作率的衰减速度且社会认同效应的作用更强。

表 4 三组平均合作率随轮次衰减的速度变化 %

轮次	基础组	内化组	社会认同组
1	3.8244	1.8688	-1.12874
2	2.9508	1.5436	-0.89168
3	2.0772	1.2184	-0.65462
4	1.2036	0.8932	-0.41756
5	0.33	0.568	-0.1805
6	-0.5436	0.2428	0.05656
7	-1.4172	-0.0824	0.29362
8	-2.2908	-0.4076	0.53068
9	-3.1644	-0.7328	0.76774
10	-4.038	-1.058	1.0048

注:衰减加速度大于 0 表示下降速度越来越小,衰减速度等于 0 或接近 0,表示下降速度处于匀速下降或接近匀速下降,而衰减速度小于 0 表示下降速度越来越大。

综上,从三组衰减速度变化的角度上看,内化效应和社会认同效应也能够在一定程度上降低平均合作率的衰减速度,但是相对于基础组,内化组和社会

认同组的下降速度仍然更大,即内化效应和社会认同效应维持高水平合作方面的作用有限。

结论1: 内化效应与社会认同效应都可以从整体上有效促进社会合作水平的提升,但是在维持高水平社会合作方面的作用有限。

3.3.2 遵守和不遵守社会规范组的公共品博弈实验数据分析

图6展示的是遵守社会规范组和不遵守社会规范组在公共品博弈实验中的投资率分布。总体上,遵守社会规范组的投资率分布更加均匀,而不遵守社会规范组的投资率分布偏左,这说明不遵守社会规范组的整体合作率较低。具体来说,在遵守社会规范组中,内化-遵守社会规范组的投资比例小于0.1的占比为28.75%,而投资比例大于0.5的占比为41.875%;社会认同-遵守社会规范组的投资比例小于0.1的占比为14.38%,这明显小于内化组,而其投资比例大于0.5的占比为57.712%,大于内化组。在不遵守社会规范组中,内化-不遵守社会规范组的投资比例小于0.1的占比为42.08%,这一比例远高于内化-遵守社会规范组,而其投资比例大于0.5的占比仅为31.878%;社会认同-不遵守社会规范组的投资比例小于0.1的占比为29.79%,而其投资比例大于0.5的占比为40.625%。

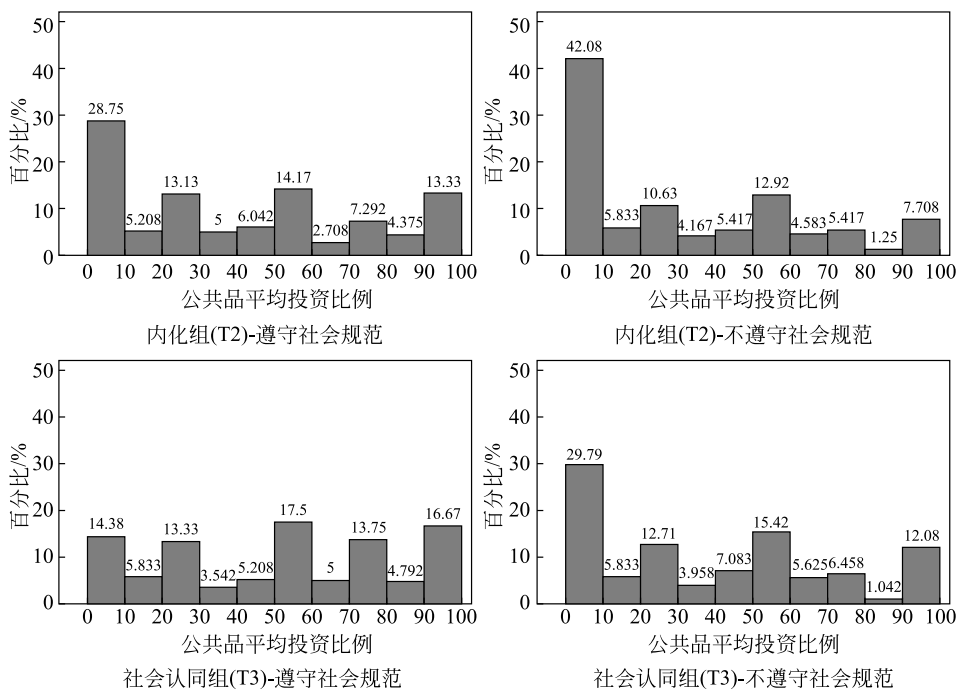


图6 遵守社会规范和不遵守社会规范的投资比例分布

然后,我们对四组数据进行 Wilcoxon 秩和检验(如表 5 所示)。在内化组中,遵守社会规范组和不遵守社会规范组之间存在显著性差异: $z=4.697, p=0.0000$ 。在社会认同组中,遵守社会规范组和不遵守社会规范组之间也存在显著差异: $z=6.365, p=0.0000$;这说明无论是在内化组还是社会认同组中,遵守社会规范组的投资比例都显著大于不遵守社会规范组的投资比例,即遵守社会规范者比不遵守社会规范者更加倾向于合作,进一步证明了内化效应在促进社会合作方面有效应。此外,我们将内化组和社会认同组中的遵守社会规范组和不遵守社会规范组分别进行比较,结果也再次证明了社会认同效应在促进社会合作方面的有效性。具体来说,在遵守社会规范组中,内化组与社会认同组之间存在显著性差异($z=-5.273, p=0.0000$),即社会认同组-遵守社会规范组的合作率显著更高。在不遵守社会规范组中,内化组和社会认同组也存在显著性差别($z=-3.877, p=0.0001$),且社会认同组-不遵守社会规范组的合作率显著更高。

表 5 Wilcoxon Rank-sum Test 检验结果

实验组	原假设	z 值	p 值
T2-遵守规范组 vs T2-不遵守规范组	两组投资比例无显著性差异	4.697	0.0000
T3-遵守规范组 vs T3-不遵守规范组	两组投资比例无显著性差异	6.365	0.0000
T2-遵守规范组 vs T3-遵守规范组	两组投资比例无显著性差异	-5.273	0.0000
T2-不遵守规范组 vs T3-不遵守规范组	两组投资比例无显著性差异	-3.877	0.0001

注:T2 表示内化组,T3 表示社会认同组。

为了进一步探究社会合作对内化效应和社会认同效应的依赖程度是否会随人们对社会规范的遵守程度变化,即内化效应和社会认同效应在遵守社会规范组和不遵守社会规范组中推动社会合作的主导角色或相对强度是否发生变化,我们将公共品博弈实验的投资率对两组(内化组和认同组)的遵守社会规范组和不遵守社会规范组进行回归分析,结果如表 6 所示。首先,当以基础组为对照组时,内化-遵守社会规范组、社会认同-遵守社会规范组、内化-不遵守社会规范组和社会认同-不遵守社会规范组与其合作率都具有显著性差别。其中,如表 6 第(5)和第(6)列的结果所示,在不遵守社会规范组中,内化效应^①并没有带来合作率的显著提升,而由社会认同效应所带来的合作率提升为 7.101%。在遵守社会规范组中,由内化效应所带来的合作率提升为 8.145%,即内化组-遵守社会规范组的合作率比基础组高 8.145%,而由社会认同效应所带来的

① 内化效应是内化组-遵守社会规范组(内化组-不遵守社会规范组)与基础组的平均合作率差异;同理,社会认同效应是社会认同组-遵守社会规范组(社会认同-不遵守社会规范组)与内化组-遵守社会规范组(内化组-不遵守社会规范组)的平均合作率差异。值得注意的是,在我们的计算定义里内化效应可能是负向的,例如由于内化组的最不遵守社会规范组在社会规范遵守方面的平均水平低于基础组,即内化组的最不遵守社会规范组的内化效应显然低于基础组,使得计算结果是负向的。

表 6 遵守和不遵守社会规范组在公共品博弈实验中平均合作率的差异比较

模型	被解释变量：被试者在公共物品博弈实验中的投资百分比					
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)
T2-遵守规范组	7.725 *** (1.964)	8.758 *** (1.950)			7.725 *** (1.964)	8.145 *** (1.956)
T2-不遵守规范组	-2.233 (1.194)	-0.941 (1.846)			-2.233 (1.870)	-1.368 (1.862)
T3-遵守规范组			18.881 *** (1.916)	19.235 *** (1.905)	18.881 *** (1.917)	19.257 *** (1.910)
T3-不遵守规范组			5.516 *** (1.919)	5.288 *** (1.922)	5.516 *** (1.194)	5.733 *** (1.918)
性别		-1.555 (1.996)		5.488 *** (1.874)		-2.424 (1.540)
家庭兄弟姐妹数		3.721 *** (1.007)		2.620 *** (0.868)		2.328 *** (0.738)
其他控制变量	NO	YES	NO	YES	NO	YES
常数项	31.306 *** (1.194)	25.056 *** (3.719)	31.306 *** (1.193)	38.591 *** (3.578)	31.306 *** (1.194)	33.426 *** (2.990)
F 值	12.09	10.72	49.03	31.23	32.87	24.63
R ²	0.0137	0.0220	0.0525	0.0604	0.0452	0.0488
观测值	1760	1760	1760	1760	2720	2720

注：括号里的数值为标准误，以分组后的不同实验组别为虚拟变量。如模型(1)中，以内化-遵守社会规范组为虚拟变量，将其设为 1，其他取值为 0，以内化-不遵守社会规范组为虚拟变量，将其设为 1，其他为 0；其他控制变量为：民族、户口、常住地。* $p<0.1$ ，** $p<0.05$ ，*** $p<0.01$ 。

合作率提升为 11.112%。因此，在人们的社会规范意识相对较为薄弱的情况下，相较于内化效应，社会认同效应是助推高效率合作的主导方式。而随着社会规范意识的提高，内化效应对合作的促进效果逐渐凸显并逐渐取代社会认同效应成为推动高效率合作的主导方式。此外，在表 6 的第(6)列中，可以看到家庭中兄弟姐妹的数量越多，个人的合作倾向也越高。此结论和探究独生子女和非独生子女偏好差异的研究结论相一致，即有兄弟姐妹的青少年在信任、值得信任和合作倾向等方面表现得更好 (Dunn, 1988; Cameron et al., 2013; 陈刚, 2020)，其可能的机制是与兄弟姐妹间的交流互动有效提高了个人的社交技能并促进了心理层面的发展，使他们形成了更信任他人和更愿意与他人合作的特性。为了更清晰地看出内化效应和社会认同效应在不同社会规范强度（即人们的社会规范意识）下相对主导角色的变化，我们将内化组和社会认同组被试者

按照遵守社会规范的程度高低分组^①进行分析,结果如表 7 所示。总的来看,随着社会规范遵守程度的提高,内化效应在促进社会合作方面的作用不断增强,逐渐取代社会认同效应成为推动高水平合作的主导力量,社会认同效应则维持在一个相对平稳的水平。

结论 2: 随着社会规范遵守程度的增强,社会认同效应在促进社会合作方面的主导作用逐渐被内化效应取代,即对于不遵守社会规范组来说,社会认同效应对合作的促进强度相对更大,对于遵守社会规范组来说,内化效应在促进社会合作方面起到了主导作用。

表 7 不同社会规范遵守程度组的内化效应和社会认同效应

遵守社会规范程度	内化效应	认同效应
最不遵守社会规范组	-11.743 ***	11.729 ***
不遵守社会规范组	-2.233	7.75 ***
遵守社会规范组	7.725 ***	11.156 ***
最遵守社会规范组	13.881 ***	11.979 ***

注:从上到下表示遵守社会规范的程度逐渐增强,其中遵守规范组是内化-遵守规范组和社会认同-遵守规范组的组合,而不遵守规范组是内化-不遵守规范组和社会认同-不遵守规范组的组合。^{*} $p<0.1$,^{**} $p<0.05$,^{***} $p<0.01$ 。

最后,我们对内化效应和社会认同效应在遵守社会规范组和不遵守社会规范组维持社会合作方面的作用进行分析。图 7 展示的是各个实验组中遵守社会规范组和不遵守社会规范组的合作率变化趋势,我们对四组数据分别进行了期数效应检验。首先,在不遵守社会规范组中,不管是内化-不遵守社会规范组还是社会认同-不遵守社会规范组,其合作率均随着轮次增加呈现明显的下降趋势,其中内化-不遵守社会规范组的期数效应为: $b=-3.839, p=0.000$,结果与 Kimbrough and Vostroknutov (2016) 的结论一致;而社会认同-不遵守社会规范组的期数效应为: $b=-3.149, p=0.000$,显然,这两组的合作率下降速度显著快于基础组($b=-1.816, p=0.003$)。其次,在遵守社会规范组中,内化-遵守社会规范组和社会认同-遵守社会规范组的合作率水平都存在明显的下降趋势,社会认同-遵守社会规范组的下降速度最快($b=-3.602, p=0.000$),基础组次之($b=-1.816, p=0.003$),内化-遵守社会规范组最慢($b=-1.405, p=0.003$)。其中,内化-遵守社会规范组的结果与 Kimbrough and Vostroknutov (2016) 的结论显著

^① 在前文实验设计中的分组规则下,我们分别将社会规范遵守程度最强(排序为 1~4)和最弱(排序为 13~16)的两组抽离出来进行分析,并将社会规范遵守程度最强的组定义为最遵守规范组,最弱的组定义为最不遵守规范组。值得注意的是,我们并没有采用排序 5~8 组和 9~12 组的数据,其主要原因是公共物品博弈实验中的分组规则是按照被试者在社会规范实验中的表现从大到小进行排序分组,因此处在中间的两组受到随机因素的干扰较大,在社会规范遵守程度方面的区分度较小。

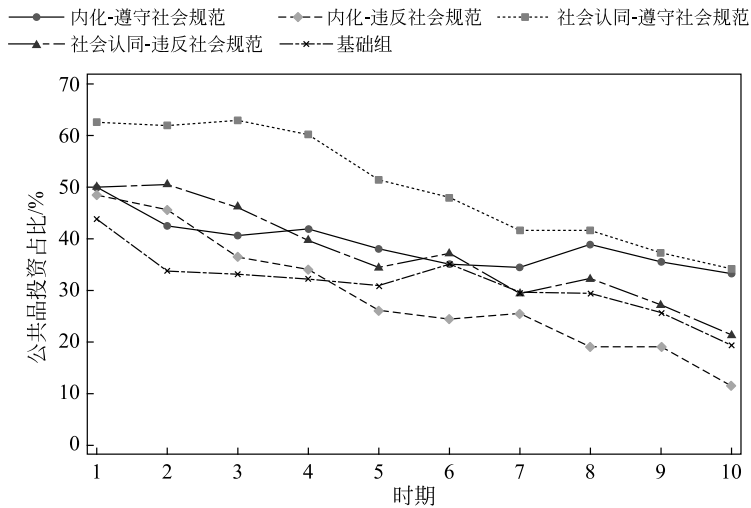


图7 公共品博弈实验的投资趋势

不同,他们认为遵守社会规范组的期数效应不显著,即合作率较高且稳定,这可能与他们的实验设计没有明确告诉被试者进行几轮公共品博弈实验有关,而本实验明确告诉了被试者会进行10轮公共品博弈实验。从结果来看,社会认同-遵守规范组具有最明显的下降趋势(下降速度最快),可能是由于其初始合作率水平最高,下降空间更大。但有趣的是,随着社会规范意识的增强(即社会规范遵守程度的提高),内化效应在维持高水平合作方面的作用逐渐凸显,即内化-遵守社会规范组的合作率下降速度远低于基础组和内化-不遵守规范组。

同理,为了进一步探究各遵守规范组和不遵守规范组平均合作率随轮次衰减的速度变化,我们利用Matlab的curve fitting tool分别对内化组和社会认同组的遵守规范组和不遵守规范组的公共品博弈实验的结果进行非线性拟合(具体拟合曲线和方程形式详见附录2)。各组平均合作率随轮次变化的加速度结果如表8所示,可以看到不管是内化-遵守规范组、内化不遵守规范组还是社会认同-不遵守规范组,其平均合作率的衰减加速度变化趋势都与基础组相似,即在前6期,平均合作率的衰减速度在减缓并逐渐接近于均速下降,后4期则逐渐加速。其中,相对于内化-不遵守规范组,内化-遵守规范组的平均合作率的加速度绝对值在各轮次中相对更小,即随社会规范遵守程度的增强,内化效应在维持甚至降低平均合作率衰减速度方面的作用也有所增强。而对社会认同-遵守社会规范组,其平均合作率在前5期加速衰减,但是加速度越来越小,在后5期则加速降低其衰减速度,且平均合作率的加速度绝对值在各轮次中都比基础组更小。显然,相对于基础组,社会认同组-遵守规范组结合了内化效应和社会认同效应,结合以上分析结果可知内化效应和社会认同效应都能够在一定程度上

维持甚至降低平均合作率的衰减速度,但从长期来看,社会认同效应在降低平均合作率衰减速度方面表现更为突出。

表 8 不同组平均合作率随轮次衰减的速度变化 %

轮次	T1	T2_遵守规范组	T2_不遵守规范组	T3_遵守规范组	T3_不遵守规范组
1	3.8244	1.82682	1.91078	-2.9146	0.65692
2	2.9508	1.51164	1.57556	-2.3152	0.53164
3	2.0772	1.19646	1.24034	-1.7158	0.40636
4	1.2036	0.88128	0.90512	-1.1164	0.28108
5	0.33	0.5661	0.5699	-0.517	0.1558
6	-0.5436	0.25092	0.23468	0.0824	0.03052
7	-1.4172	-0.06426	-0.10054	0.6818	-0.09476
8	-2.2908	-0.37944	-0.43576	1.2812	-0.22004
9	-3.1644	-0.69462	-0.77098	1.8806	-0.34532
10	-4.038	-1.0098	-1.1062	2.48	-0.4706

注:T1表示基础组,T2表示内化组,T3表示社会认同组;衰减加速度大于0表示下降速度越来越小,衰减速度等于0或接近0,表示下降速度处于匀速下降或接近匀速下降,而衰减速度小于0则表示下降速度越来越大。

综上,我们可以发现内化效应在维持高水平合作方面的作用随社会规范遵守程度的增强逐渐凸显,社会认同效应则没有发挥出维持高水平合作的作用。内化效应和社会认同效应在维持甚至降低平均合作率的衰减速度方面都具有重要的作用,其中社会认同效应在降低平均合作率衰减速度方面表现更为突出,但短期内并没有扭转社会认同组(包括遵守规范组和不遵守规范组)高速下降的趋势。

结论 3: 内化效应在维持高水平合作方面的作用随社会规范遵守程度的增强逐渐凸显,社会认同效应则在维持甚至降低平均合作率衰减速度方面表现更好,但短期内仍无法扭转平均合作率高速下降的趋势。

3.3.3 稳健性检验

1) 条件合作者的影响

不同组之间的合作率差异可能会受到各组条件合作者比例的影响,从而对内化效应和社会认同效应的估计造成混杂。为检验这一点,我们利用实验结束后通过问卷测度的条件合作者数据^①对此进行比较分析,结果如表 9 所示。实际上,我们可以看到基础组、内化组和社会认同组及其细分的遵守规范组和不

^①但值得注意的是,我们对于条件合作者的度量是在公共品博弈实验结束后进行的,可能存在很多混杂因素的干扰,例如顺序效应的影响,因此该结果仅能作为一些补充,我们对于该结果也是保持谨慎的态度。

遵守规范组在条件合作者比例上并没有明显的差异,比例均在 54%~63% 范围内,由此可以在一定程度上排除各组条件合作者比例差异对合作结果所造成的干扰。此外,可以看到认同组(T3)-不遵守规范组相对于基础组有较大比例的搭便车者(20.7% vs 16.2%),但是其合作行为仍然显著优于基础组,这说明社会认同可能起着比较大的作用。

表 9 不同组中被试类型分布 %

实验局	搭便车者	条件合作者	U 型	完全合作	无法分类者
T1	16.2	60.0	15.0	1.3	7.5
T2	18.7	56.3	12.5	1.0	11.5
T3	13.5	58.3	14.6	1.0	12.6
T2-遵守规范组	14.5	58.4	10.4	2.0	19.0
T2-不遵守规范组	22.9	54.2	14.6	0	13.0
T3-遵守规范组	6.3	62.5	17.6	0	13.6
T3-不遵守规范组	20.7	54.1	11.6	2.0	11.6

注:T1 表示基础组,T2 表示内化组,T3 表示社会认同组。

2) 策略性行为影响

由于在社会认同组中,被试者被明确告知其在公共物品博弈实验中的分组规则,由此可能会引发被试者的策略性行为,例如积极互惠和消极互惠行为,进而对内化效应和社会认同效应的估计造成偏差。为此,我们利用受分组时的随机因素影响最小的最遵守社会规范组和最不遵守社会规范组的数据对此进行分析。

假设存在策略性行为的影响,社会认同-最遵守规范组和内化-最遵守规范组的差异将是积极互惠和社会认同效应的叠加,而社会认同-最不遵守规范组和内化-最不遵守规范组的差异将是消极互惠和社会认同效应的叠加。(a)假设社会认同效应在社会认同-最遵守组和社会认同-最不遵守组之间不存在差异,且可以看到前者的差异显著大于后者,那就可能说明存在策略性行为的干扰,但是我们发现社会认同-最遵守和内化-最遵守之间的合作率差异为 11.979%,而社会认同-最不遵守和内化-最不遵守之间的合作率差异为 11.729%,两者之间并不存在显著性的差异,这在一定程度上证明了策略性行为的干扰可能是微不足道的。(b)如果假设社会认同效应在社会认同-最遵守组更大,我们应该期望可以看到更显著的差距,但并没有得到数据支撑,所以(a)中结论同样适用。(c)如果假设社会认同效应在社会认同-最不遵守组更大,策略性行为可能是一个相对较大的干扰,但是由于文章所研究的社会认同恰是以规范认同为聚点引发对同组个体成员的认同,所以我们认为(c)的情形可能是不符合现实的。综上,我们认为被试者的策略性动机对主要实验结果的影响可能是有限的。

4 结论与讨论

本文通过社会规范实验和公共品博弈实验研究了社会规范对社会合作的影响,以进一步分解依托于社会规范促进合作的两种常见的机制效应:内化效应和社会认同效应。研究发现:(1)内化效应与社会认同效应都可以从整体上有效促进社会合作水平的提升,但是在维持高水平社会合作方面的作用有限;(2)随着社会规范遵守程度的增强,社会认同效应在促进社会合作方面的主导作用逐渐被内化效应取代;(3)内化效应在维持高水平合作方面的作用随社会规范遵守程度的增强逐渐凸显,社会认同效应则在维持甚至降低平均合作率衰减速度方面表现得更好,但短期内仍无法扭转平均合作率高速下降的趋势。

社会规范是社会生活中一个普遍的特征,几乎遍及人类社会交往的各个方面,对促进合作也有着重要的作用。然而,有关社会规范的实证研究相对较少,尤其是关于社会规范如何促进合作的机制研究,而本文利用实验室实验研究了社会规范对合作的影响效应,试图为社会规范研究提供一种新的思路。本文与多数关于社会规范如何影响合作的研究不同:首先,以往研究大多是对公共品博弈实验的合作规范进行研究,比如惩罚搭便车者的规范和条件合作者规范(Fehr and Fischbacher, 2004a, 2004b),本文则强调可以在合作之前进行规划,通过社会规范激发特定的社会认同进而促进社会合作,这种方式很好地避免了前者在维持合作方面的高成本执行传导问题。其次,以往研究大多是单一地探究内化效应或社会认同效应对社会合作的促进作用,较少涉及现实生活中社会规范促进社会合作的两种效应的结合,即内化效应和社会认同效应的叠加,本文则对社会规范的这两种效应进行研究。最后,本研究通过变化社会规范的遵守程度模拟了现实世界中社会规范发展的不同阶段,探究了依托于社会规范发挥作用的内化效应和社会认同效应在推动社会合作方面的主导角色变化,为推动我国社会规范意识建设提供证据支撑。

当然,本文还有诸多不完善的方面:其一,本文利用实验室实验研究了社会认同效应和内化效应对社会合作的促进强度,但是我们是通过社会规范这个认知聚点引发人们的社会认同感,不同于以往自然的身份认同。有研究将社会认同划分为两类,即固有身份(Hoff and Pandey, 2006; de Benedictis-Kessner et al., 2013)和最简群体范式(Eckel and Grossman, 2005),本研究则关注于后者。之后的研究可以进一步考虑将个体自然身份纳入研究框架。此外,还有更复杂的社会情境可以被考虑进来,比如外部环境的不确定性如何影响这两种机制的作用、个体的抗风险能力是否会影响人们对社会规范行为的需求和公共品博弈合作水平,这也是我们实验进一步扩展的方向。其二,研究结论的外部有效性有待检验。一般来说,实验室实验相较于实地实验具有更好的内部有效性,但由

于实验室实验本身的局限性,我们需要进一步通过实地实验方法检验其外部有效性。其三,本文在检验分组规则是否有效激发了被试者的社会认同上即在论证群体身份上的证据相对较弱。尽管我们通过补充问卷的形式弥补了这一缺陷,但由于问卷缺乏真实的货币激励,可能存在一定的偏差。此外,本实验采用的环保类的社会规范尽可能有效激发实验参与者的社会认同(Dono et al., 2010; Fritsche et al., 2018),可以进一步加强实验结果的有效性,但即使如此,这也是我们在后续研究中需要注意的问题和完善的地方。以上的不足和缺陷也是我们未来进一步完善扩展现有研究的方向,希望为丰富社会规范与合作的研究尽一点力量。

附录

附录 1

我们采用问卷方法检验社会认同组的分组规则能否有效激发受访者对组内成员的社会认同感。具体来说,我们设计了两个实验局:

(1) 社会认同组:首先,受访者需要完成社会规范实验,完成后被告知其将与另外3位受访者分为一组,且这三位受访者投入A盒子(“垃圾箱”)的小球个数与其投入A盒子中的小球个数相近;其次,该受访者需要完成四种分配情形下的第三方独裁者博弈实验,即受访者被告知其需要在两位受访者之间分配100个筹码,分配决策不会影响受访者本人的收益。四种分配情形下的两位受访者组合分别是:独裁者游戏1(同组参与者1,同组参与者2)、独裁者游戏2(不同组参与者1,不同组参与者2)、独裁者游戏3(同组参与者1,不同组参与者2)和独裁者游戏4^①(不同组参与者1,同组参与者2);

(2) 对照组:受访者的任务与社会认同组相同,唯一的区别在于受访者在完成社会规范实验后仅被告知其将与另外3位受访者分为一组,即没有特殊的分组规则。

有效收集到的对照组和社会认同组的问卷数量分别为43和44,具体分析结果展示如下。

首先,我们检验了问卷受访者和真实实验被试者在社会规范遵守程度上是否存在显著性差异^②,我们将问卷中的对照组和社会认同组分别与真实实验中

① 可以看出独裁者游戏3和独裁者游戏4本质上是一样的,而设计这两个游戏的目的是另一方面可以检验受访者是否认真填写问卷,另一方面可以排除标签对受访者的影响。结果发现受访者在独裁者游戏3和独裁者游戏4中的分配结果并没有显著性差异。

② 一个有关调查问卷的担忧是被试者缺乏真实的货币激励(调查问卷无法进行真实支付),但由于线上问卷的匿名性更高,受访者出于社会形象顾虑进行伪装的动机更弱,可以进一步抵消缺乏货币激励所带来的问题。

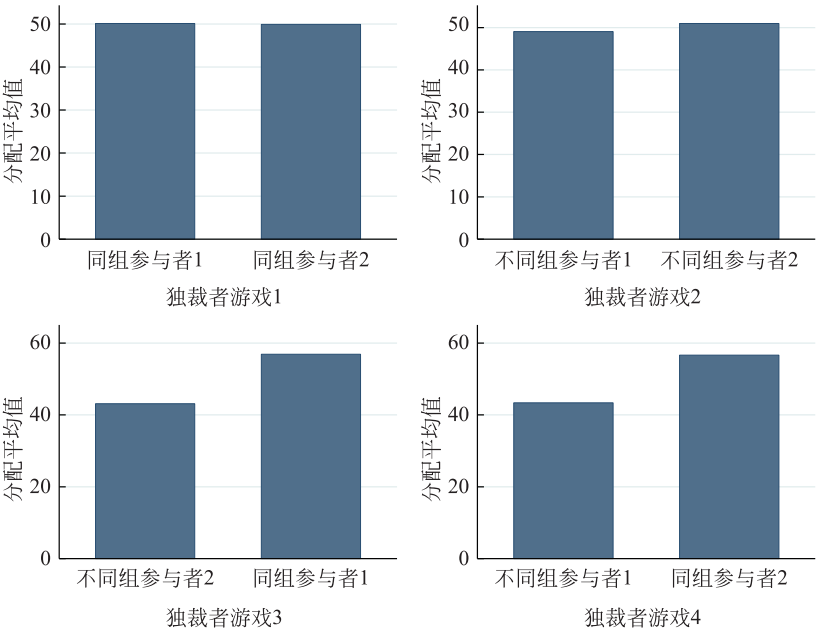
的基础组和社会认同组在社会规范实验中的结果进行比较,结果如附表 1 所示,即真实实验被试者和问卷受访者在社会规范实验中的表现并不存在显著性差异。

附表 1 Wilcoxon Ranksum Test 检验结果

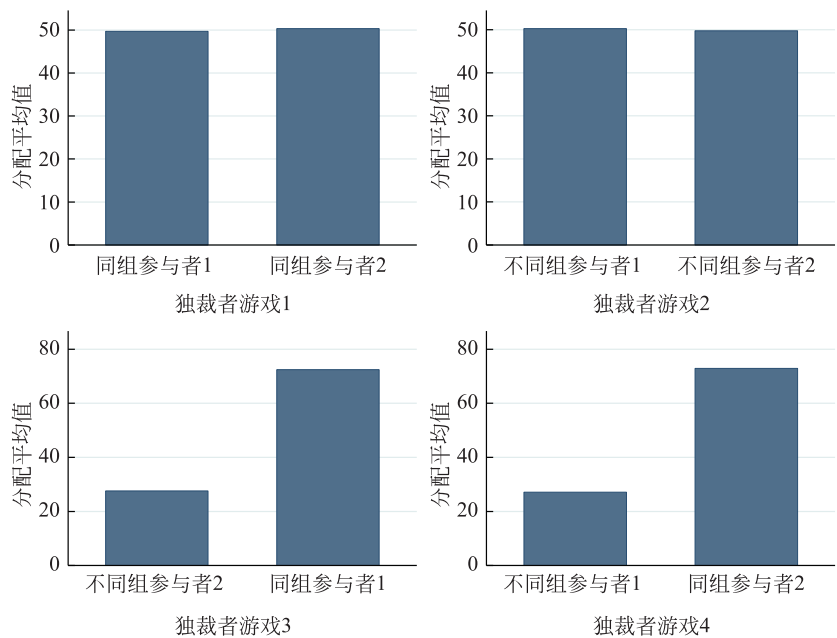
实验组	H ₀	z 值	p 值
基础组 vs 对照组	两组分布无显著性差异	-0.752	0.4523
社会认同组 I vs 社会认同组 II	两组分布无显著性差异	-0.689	0.4907

注:原假设是两组被试者投入 A 盒子中的小球个数相同。为便于区分,我们将真实实验中的社会认同组称为“社会认同组 I”,将问卷中的社会认同组称为“社会认同组 II”。

其次,我们对问卷受访者在不同独裁者博弈实验中的结果进行分析。由附图 1、2 和附表 2 可知,对照组和社会认同组的受访者在独裁者游戏 1 和独裁者游戏 2 中对参与者 1 和参与者 2 的分配都遵循“50-50”的公平分配。此外,我们发现无论是对照组还是社会认同组,受访者在独裁者游戏 3 和独裁者游戏 4 中都表现出一定的组内偏好,即对同组成员分配更多的筹码。这说明即使没有特殊的分组规则,单纯“同组”的概念就可能增加被试者对组内成员的偏好。附表 3 是分别对四种分配情形下的分配结果进行 T 检验的结果,可以发现在游戏 1 和游戏 2 中,对照组和社会认同组的受访者在不同受访者之间的分配值并不存在显著差异,但是其在游戏 3 和游戏 4 的分配值存在显著性差异,即受访者均显著偏好于组内成员。



附图 1 第三方独裁者游戏(对照组)



附图 2 第三方独裁者游戏(社会认同组)

附表 2 描述性统计

	独裁者游戏 1 (同组 1, 同组 2)	独裁者游戏 2 (不同 1, 不同 2)	独裁者游戏 3 (同组 1, 不同 2)	独裁者游戏 4 (不同 1, 同组 2)
对照组	(50.1, 49.9)	(49, 51)	(56.9, 43.1)	(43.3, 56.7)
社会认同组	(49.7, 50.3)	(50.25, 49.75)	(72.4, 27.6)	(27.1, 72.9)

注：括号内的数字是不同组受访者在不同独裁者博弈中分配给对应同组或不同组成员的平均筹码数。

附表 3 T 检验

	H ₀	t 值	p 值
对照组			
独裁者游戏 1	同组 1=同组 2	-0.4330	0.6672
独裁者游戏 2	不同组 1=不同组 2	0.1179	0.9067
独裁者游戏 3	同组 1=不同组 2	8.7282	0.0000
独裁者游戏 4	不同组 1=同组 2	-8.8023	0.0000
社会认同组			
独裁者游戏 1	同组 1=同组 2	0.7809	0.4392
独裁者游戏 2	不同组 1=不同组 2	-1.4680	0.1496
独裁者游戏 3	同组 1=不同组 2	4.3187	0.0001
独裁者游戏 4	不同组 1=同组 2	-4.1568	0.0002

最后,为检验社会认同组的分组规则是否有效激发了受访者的社会认同感,我们将对照组和社会认同组受访者在独裁者博弈 3 和独裁者博弈 4 中的分配结果进行对比。如果社会认同组受访者在独裁者博弈 3 和 4 中分配给组内成员的筹码数显著高于对照组,而分配给组外成员的筹码数显著低于对照组,即可以表明社会认同组的分组规则确实激发或强化了受访者对组内成员的社会认同感。结果如回归附表 4 所示,我们发现相对于对照组,社会认同组的受访者在独裁者游戏 3 和 4 中分配给同组成员显著更多的筹码,而分配给不同组成员显著更少的筹码。这说明社会认同组中的分组规则确实有效激发或强化了受访者的社会认同感。

附表 4 对照组和社会认同组在不同独裁者博弈中分配结果比较

	独裁者游戏 1-同组	独裁者游戏 2-不同组	独裁者游戏 3-同组	独裁者游戏 3-不同组	独裁者游戏 4-同组	独裁者游戏 4-不同组
社会认同组	-0.698 (0.932)	0.861 (2.207)	14.251 *** (3.135)	-14.251 *** (3.135)	15.249 *** (3.234)	-15.249 *** (3.234)
年龄	0.204 (0.827)	2.333 (1.937)	-3.515 (2.581)	3.515 (2.580)	-2.967 (2.828)	2.967 (2.828)
性别	-1.206 (1.188)	-3.599 * (2.076)	1.919 (3.181)	-1.919 (3.181)	1.798 (3.194)	-1.798 (3.194)
教育	-0.367 (0.742)	-1.462 (1.789)	-0.991 (2.292)	0.991 (2.292)	-0.625 (2.234)	0.625 (2.234)
常数项	50.76 *** (0.745)	47.51 *** (2.853)	66.242 *** (5.919)	33.758 *** (5.919)	64.109 *** (6.506)	35.891 *** (6.506)
R ²	0.0344	0.0385	0.2516	0.2561	0.2582	0.2582
观测值	87	87	87	87	87	87

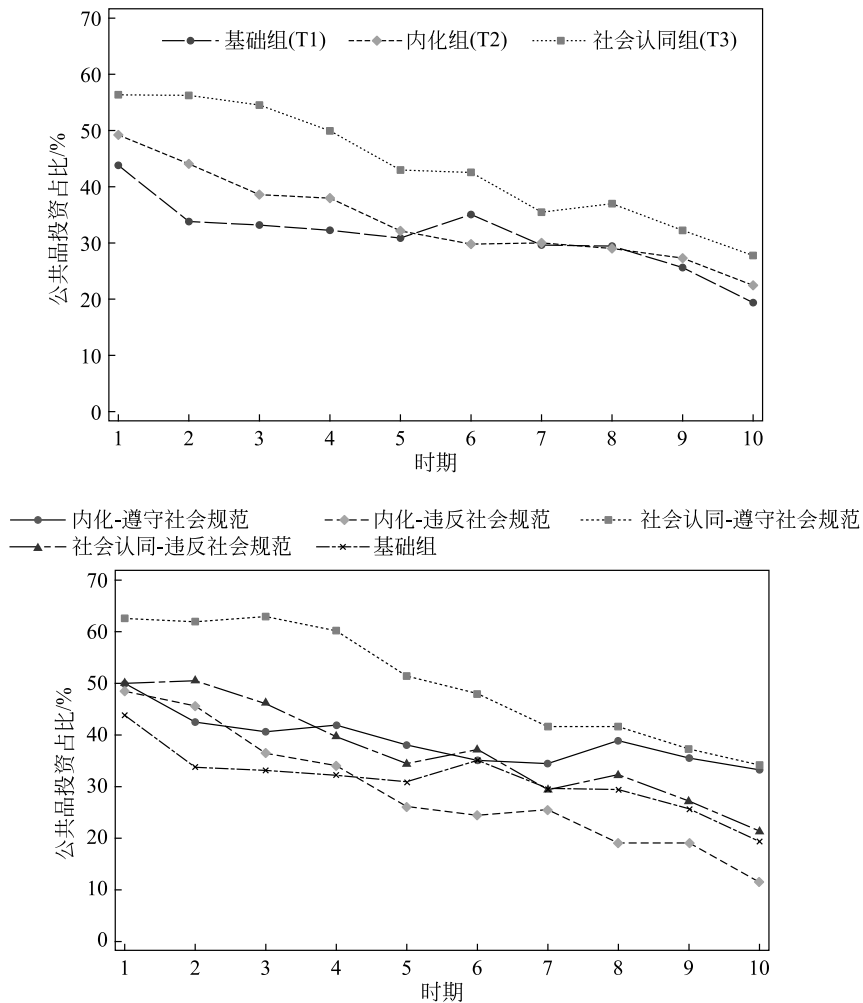
注:括号里的数值为标准误。不同实验组别为虚拟变量,社会认同组设为 1,对照组设为 0。被解释变量分别表示在不同独裁者游戏中,受访者分配给同组或者不同组的筹码数。* $p<0.1$, ** $p<0.05$, *** $p<0.01$ 。

附录 2

我们采用 Matlab 软件的 curve fitting tool 对各组的公共品博弈投资水平数据进行非线性拟合。具体步骤如下。

第一步:筛选非线性拟合方程

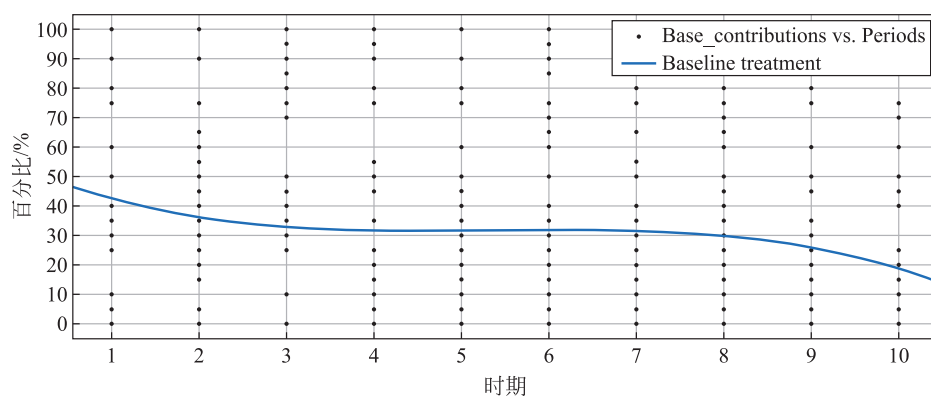
首先,通过各组在公共品博弈实验中的平均合作率下降趋势图确定非线性拟合方程。如附图 3 所示,可以看到曲线下降趋势比较符合一元多项式方程。



附图3 各组的平均合作率趋势

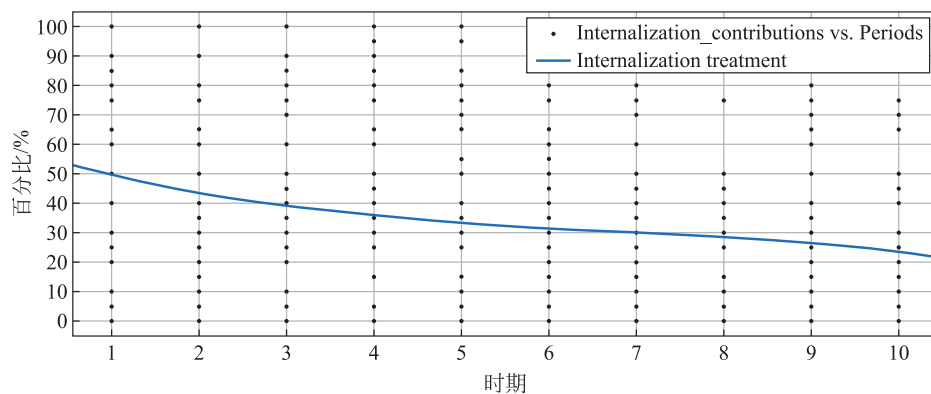
第二步: 确定最高次项和方程形式

在确定一元多项式方程形式后, 需要进一步寻找多项式方程的最高次项, 即我们利用 Matlab 拟合工具确定最高次项以及方程的具体形式。由于 curve fitting tool 中的最高次项的选择范围为 1~9, 我们对 9 种最高次项进行逐个测试并根据拟合效果选择最佳的最高次项和具体方程形式。结果显示, 不管是基础组、内化组、社会认同组, 还是内化组和社会认同组的遵守规范组和不遵守规范组, 其拟合优度达到最大时的最高次项均为 3, 具体拟合曲线图和方程形式如附图 4~10 所示。



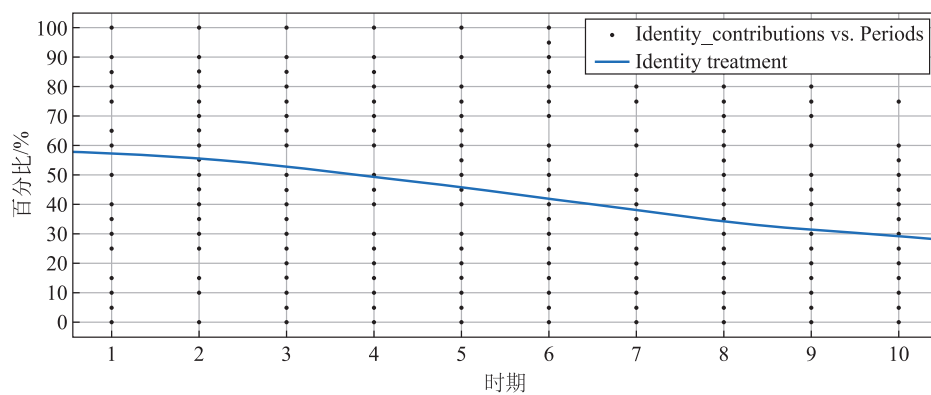
附图4 基础组拟合曲线

基础组的拟合方程: $\hat{y} = -0.1456 \cdot x^3 + 2.349 \cdot x^2 - 12.31 \cdot x + 52.64$



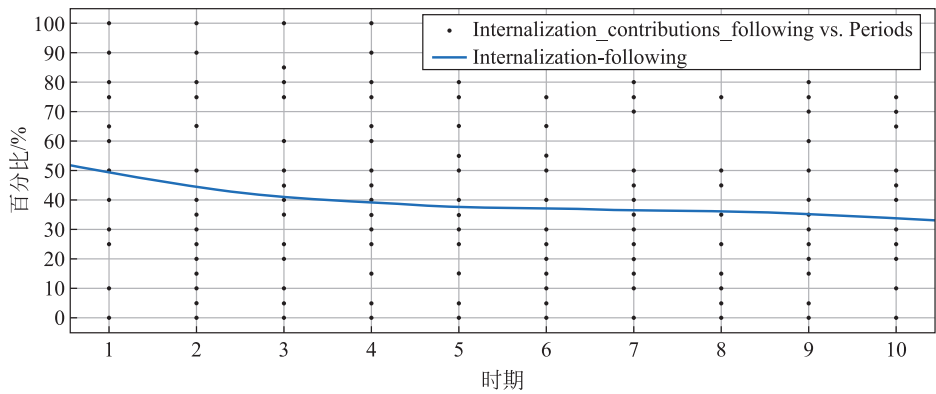
附图5 内化组拟合曲线

内化组的拟合方程: $\hat{y} = -0.0542 \cdot x^3 + 1.097 \cdot x^2 - 8.974 \cdot x + 57.58$



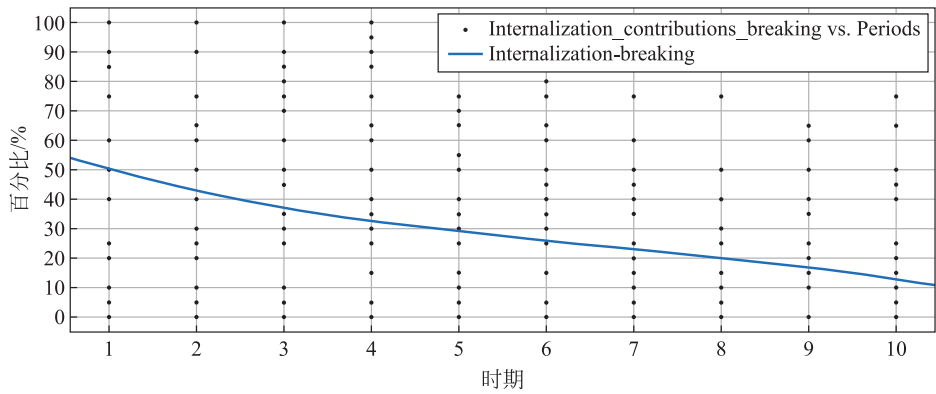
附图6 社会认同组拟合曲线

社会认同组的拟合方程: $\hat{y} = 0.03951 \cdot x^3 - 0.6829 \cdot x^2 - 0.0287 \cdot x + 58$



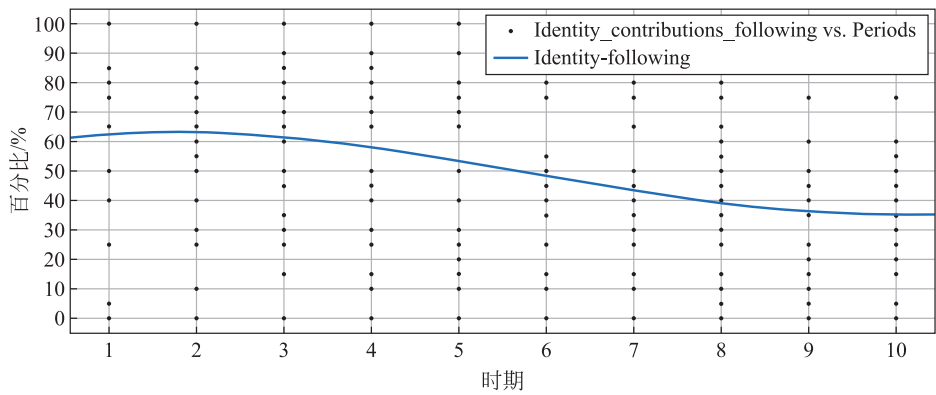
附图 7 内化-遵守规范组拟合曲线

内化-遵守规范组的拟合方程： $\hat{y} = -0.05253 \cdot x^3 + 1.071 \cdot x^2 - 7.652 \cdot x + 55.77$



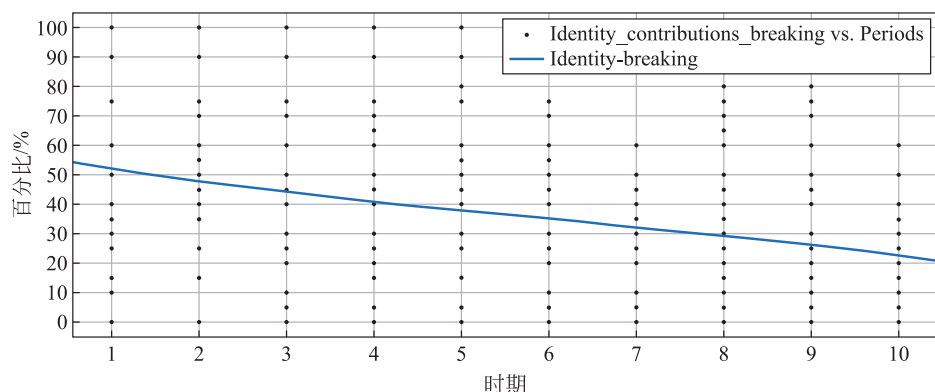
附图 8 内化-不遵守规范组拟合曲线

内化-不遵守规范组的拟合方程： $\hat{y} = -0.05587 \cdot x^3 + 1.123 \cdot x^2 - 10.3 \cdot x + 59.39$



附图 9 社会认同-遵守规范组拟合曲线

社会认同-遵守规范组的拟合方程： $\hat{y} = 0.0999 \cdot x^3 - 1.757 \cdot x^2 + 5.193 \cdot x + 59.04$



附图 10 社会认同-不遵守规范组拟合曲线

社会认同-不遵守规范组的拟合方程: $\hat{y} = -0.02088 \cdot x^3 + 0.3911 \cdot x^2 - 5.251 \cdot x + 56.96$

参考文献

陈刚. 2020. 独生子女政策与消失的企业家精神[J]. 经济学动态, (7): 84-98.

Chen G. 2020. China's one-child policy and the missing entrepreneurship [J]. *Economic Perspectives*, (7): 84-98. (in Chinese)

陈叶烽, 何浩然. 2012. 是什么影响了人们的自愿合作水平——基于公共品博弈实验数据的分析[J]. 经济学家, (5): 18-26.

Chen Y F, He H R. 2012. What is the impact of people's voluntary cooperation—Analysis based on experimental data of public goods game[J]. *Economist*, (5): 18-26. (in Chinese)

崔驰, 戴明, 孙可可. 2022. 中国大学生是不是精致利己主义者? ——基于显示偏好原理的独断博弈实验研究[J]. 东北师大学报(哲学社会科学版), (1): 88-99.

Cui C, Dai M, Sun K K. 2022. Are Chinese college students refindly self-interested? —An experimental study of dictator game based on GARP[J]. *Journal of Northeast Normal University: Philosophy and Social Sciences*, (1): 88-99. (in Chinese)

崔驰, 路智雯, 戴明. 2019. 简单、复杂劳动获得禀赋是否影响人们的亲社会行为[J]. 税务与经济, (3): 25-31.

Cui C, Lu Z W, Dai M. 2019. Does simple and complex labor gain endowment affect people's prosocial behavior? [J]. *Taxation and Economy*, (3): 25-31. (in Chinese)

崔驰, 路智雯. 2018. 禀赋来源和框架影响人们的分配行为吗——基于 Dictator

- game 的实验研究[J]. 南方经济, (9): 107-122.
- Cui C, Lu Z W. 2018. Does the source of endowment and framework affect people's distribution behavior: Experimental research based on dictator game [J]. *South China Journal of Economics*, (9): 107-22. (in Chinese)
- 戴明. 2020. 不同年龄维度的青少年公平观异质性的研究[D]. 东北师范大学.
- Dai M. 2020. A study on the heterogeneity of teenagers of fairness views across ages [D]. *Northeast Normal University*. (in Chinese)
- 关宏宇, 朱宪辰, 周彩霞. 2015. 规则公平偏好对个体公共物品自愿供给的影响: 一项实验研究[J]. 财贸经济, (5): 35-47.
- Guan H Y, Zhu X C, Zhou C X. 2015. Fair-rule preference and the voluntary provision in public goods: An experimental investigation [J]. *Finance and Trade Economics*, (5): 35-47. (in Chinese)
- 连洪泉, 周业安, 左聪颖, 陈叶烽, 宋紫峰. 2013. 惩罚机制真能解决搭便车难题吗? ——基于动态公共品实验的证据[J]. 管理世界, (4): 69-81.
- Lian H Q, Zhou Y A, Zuo C Y, Chen Y F, Song Z F. 2013. Can the punishment mechanisms really solve the free-rider problems? —A study based on the evidence from the test on the dynamic public goods game[J]. *Journal of Management World*, (4): 69-81. (in Chinese)
- 张新超, 范良聪, 刘璐. 2016. 腐败何以持续: 基于社会规范视角的实验研究[J]. 南方经济, (5): 97-114.
- Zhang X C, Fan L C, Liu L. 2016. Corruption as a social norm: An experimental investigation[J]. *South China Journal of Economics*, (5): 97-114. (in Chinese)
- 周业安, 黄国宾, 何浩然, 等. 2014. 领导者真能起到榜样作用吗? ——一项基于公共品博弈实验的研究[J]. 管理世界, (10): 75-90.
- Zhou Y A, Huang G B, He H R, et al. 2014. Can leaders really play a role of model? An investigation based on public goods game[J]. *Journal of Management World*, (10): 75-90. (in Chinese)
- 周业安, 王一子. 2016. 社会认同、偏好和经济行为——基于行为和实验经济学研究成果的讨论[J]. 南方经济, (10): 95-105.
- Zhou Y A, Wang Y Z. 2016. Social identity, preferences and economic behaviors: Based on the reviews of the research findings of behavior and experimental economics [J]. *South China Journal of Economics*, (10): 95-105. (in Chinese)
- Akerlof G A, Kranton R E. 2000. Economics and identity[J]. *The Quarterly Journal of Economics*, 115(3): 715-753.
- Andreoni J. 1988. Why free ride?: Strategies and learning in public goods experiments [J]. *Journal of public Economics*, 37(3): 291-304.
- Andreoni J. 1995. Cooperation in public-goods experiments: Kindness or confusion?

- [J]. *The American Economic Review*, 891-904.
- Andreoni, J., Bernheim, B. D. 2009. Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects[J]. *Econometrica*, 77(5), 1607-1636.
- Banerjee R. 2016. On the interpretation of bribery in a laboratory corruption game: Moral frames and social norms[J]. *Experimental Economics*, 19(1): 240-267.
- Barr A, Lane T, Nosenzo D. 2018. On the social inappropriateness of discrimination [J]. *Journal of Public Economics*, 164: 153-164.
- Bénabou R, Tirole J. 2006. Incentives and Prosocial Behavior[J]. *American Economic Review*, 96 (5): 1652-1678.
- Benjamin D J, Choi J J, Strickland A J. 2010. Social identity and preferences [J]. *American Economic Review*, 100(4): 1913-28.
- Bicchieri C, Chavez A. 2010. Behaving as expected: Public information and fairness norms[J]. *Journal of Behavioral Decision Making*, 23(2): 161-178.
- Bicchieri C, Dimant E, Sonderegger S. 2019. It's not a lie if you believe it: Lying and belief distortion under norm-uncertainty[J]. Available at SSRN.
- Bicchieri C. 2005. *The grammar of society: The nature and dynamics of social norms* [M]. Cambridge: Cambridge University Press.
- Brauer M, Chaurand N. 2010. Descriptive norms, prescriptive norms, and social control: An intercultural comparison of people's reactions to uncivil behaviors [J]. *European Journal of Social Psychology*, 40(3): 490-499.
- Brekke K A, Hauge K E, Lind J T, et al. 2011. Playing with the good guys. A public good game with endogenous group formation [J]. *Journal of Public Economics*, 95(9-10): 1111-1118.
- Calvert S L. 1995. Pictorial discourse and learning[J]. *Discursive psychology in practice*, 194-204.
- Cameron, L, Erkal, N, Gangadharan, L, Meng, X. 2013. Little emperors: Behavioral impacts of China's one-child policy[J]. *Science*, 339(6122), 953-957.
- Chang D, Chen R, Krupka E. 2019. Rhetoric matters: A social norms explanation for the anomaly of framing[J]. *Games and Economic Behavior*, 116: 158-178.
- Chen X, Sasaki T, Brännström Å, et al. 2015. First carrot, then stick: How the adaptive hybridization of incentives promotes cooperation [J]. *Journal of the royal society interface*, 12(102): 20140935.
- Chen Y, Harper F M, Konstan J, et al. 2010. Social comparisons and contributions to online communities: A field experiment on movielens [J]. *American Economic Review*, 100(4): 1358-98.
- Chen Y, Li S X. 2009. Group identity and social preferences[J]. *American Economic Review*, 99(1): 431-57.

- Cialdini R B, Reno R R, Kallgren C A. 1990. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places[J]. *Journal of personality and social psychology*, 58(6): 1015.
- Cialdini R B. 2003. Crafting normative messages to protect the environment[J]. *Current directions in psychological science*, 12(4): 105-109.
- de Benedictis-Kessner J, de Kadt D, Wise T. 2013. Playing with race: How racial diversity affects redistributive behavior in simulated public goods games[J]. *SSRN Electronic Journal*.
- De Quidt J, Vesterlund L, Wilson A J. 2019. Experimenter demand effects [M]// *Handbook of research methods and applications in experimental economics*. Cheltenham: Edward Elgar Publishing.
- Dono J, Webb J, Richardson B. 2010. The relationship between environmental activism, pro-environmental behaviour and social identity [J]. *Journal of environmental psychology*, 30(2): 178-186.
- Dunn, J. 1988. Sibling influences on childhood development [J]. *Journal of child psychology and psychiatry*, 29(2), 119-127.
- Eckel C C, Grossman P J. 2005. Managing diversity by creating team identity [J]. *Journal of Economic Behavior & Organization*, 58(3): 371-392.
- Ellingsen T, Johannesson M. 2008. Pride and prejudice: The human side of incentive theory[J]. *American Economic Review*, 98(3): 990-1008.
- Fehr E, Fischbacher U. 2004a. Social norms and human cooperation [J]. *Trends in cognitive sciences*, 8(4): 185-190.
- Fehr E, Fischbacher U. 2004b. Third-party punishment and social norms [J]. *Evolution and human behavior*, 25(2): 63-87.
- Fischbacher U, Föllmi-Heusi F. 2013. Lies in disguise—An experimental study on cheating [J]. *Journal of the European Economic Association*, 11(3): 525-547.
- Fischbacher U, Gächter S, Fehr E. 2001. Are people conditionally cooperative? Evidence from a public goods experiment [J]. *Economics letters*, 71(3): 397-404.
- Fischbacher U. 2007. z-Tree: Zurich toolbox for ready-made economic experiments [J]. *Experimental economics*, 10(2): 171-178.
- Frey B S, Savage D A, Torgler B. 2010. Interaction of natural survival instincts and internalized social norms exploring the Titanic and Lusitania disasters [J]. *Proceedings of the National Academy of Sciences*, 107(11): 4862-4865.
- Fritsche I, Barth M, Jugert P, et al. 2018. A social identity model of pro-environmental action (SIMPEA) [J]. *Psychological Review*, 125(2): 245.
- Gächter S, Kölle F, Quercia S. 2017a. Reciprocity and the tragedies of maintaining and providing the commons [J]. *Nature human behaviour*, 1(9): 650-656.

- Gächter S, Gerhards L, Nosenzo D. 2017b. The importance of peers for compliance with norms of fair sharing[J]. *European Economic Review*, 97: 72-86.
- Gächter S, Nosenzo D, Sefton M. 2013. Peer effects in pro-social behavior: Social norms or social preferences? [J]. *Journal of the European Economic Association*, 11(3): 548-573.
- Gächter S, Schulz J F. 2016. Intrinsic honesty and the prevalence of rule violations across societies[J]. *Nature*, 531(7595): 496-499.
- Gavrilets S, Richerson P J. 2017. Collective action and the evolution of social norm internalization[J]. *Proceedings of the National Academy of Sciences*, 114(23): 6068-6073.
- Goette L, Huffman D, Meier S. 2006. The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups[J]. *American Economic Review*, 96(2): 212-216.
- Goette L, Huffman D, Meier S. 2012. The impact of social ties on group interactions: Evidence from minimal groups and randomly assigned real groups[J]. *American Economic Journal: Microeconomics*, 4(1): 101-15.
- Guala F, Mittone L, Ploner M. 2013. Group membership, team preferences, and expectations[J]. *Journal of Economic Behavior & Organization*, 86: 183-190.
- Hallsworth M, List J A, Metcalfe R D, et al. 2017. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance[J]. *Journal of Public Economics*, 148: 14-31.
- Hauert C, De Monte S, Hofbauer J, et al. 2002. Replicator dynamics for optional public good games[J]. *Journal of Theoretical Biology*, 218(2): 187-194.
- Hoff K, Pandey P. 2006. Discrimination, social identity, and durable inequalities[J]. *American Economic Review*, 96(2): 206-211.
- Holt C A, Laury S K. 2002. Risk aversion and incentive effects[J]. *American Economic Review*, 92(5): 1644-1655.
- Jackson J M. 1965. Variability in visual thresholds[J]. *The Journal of Psychology*, 59(1): 17-28.
- Jetten J, Spears R, Manstead A S R. 1996. Intergroup norms and intergroup discrimination: Distinctive self-categorization and social identity effects[J]. *Journal of Personality and Social Psychology*, 71(6): 1222.
- Keser C, Van Winden F. 2000. Conditional cooperation and voluntary contributions to public goods[J]. *Scandinavian Journal of Economics*, 102(1): 23-39.
- Kessler J B, Leider S. 2012. Norms and contracting[J]. *Management Science*, 58(1): 62-77.
- Kimbrough E O, Miller J B, Vostroknutov A. 2014. Norms, frames and prosocial

- behavior in games [R]. Working paper, Simon Fraser University, Bocconi University, Maastricht University.
- Kimbrough E O, Vostroknutov A. 2016. Norms make preferences social[J]. *Journal of the European Economic Association*, 14(3): 608-638.
- Krupka E L, Weber R A. 2009. The focusing and informational effects of norms on pro-social behavior[J]. *Journal of Economic Psychology*, 30(3): 307-320.
- Krupka E L, Weber R A. 2013. Identifying social norms using coordination games: Why does dictator game sharing vary? [J]. *Journal of the European Economic Association*, 11(3): 495-524.
- Lapinski M K, Rimal R N. 2005. An explication of social norms[J]. *Communication theory*, 15(2): 127-147.
- Levitt S D, List J A. 2007. What do laboratory experiments measuring social preferences reveal about the real world? [J]. *Journal of Economic Perspectives*, 21(2): 153-174.
- Manski C F. 2000. Economic analysis of social interactions[J]. *Journal of Economic Perspectives*, 14(3): 115-136.
- Muthukrishna M, Francois P, Pourahmadi S, et al. 2017. Corrupting cooperation and how anti-corruption strategies may backfire[J]. *Nature Human Behaviour*, 1(7): 0138.
- Ohtsuki H, Iwasa Y, Nowak M A. 2009. Indirect reciprocity provides only a narrow margin of efficiency for costly punishment[J]. *Nature*, 457(7225): 79-82.
- Ostrom E. 2000. Collective action and the evolution of social norms [J]. *Journal of Economic Perspectives*, 14(3): 137-158.
- Rege M, Telle K. 2004. The impact of social approval and framing on cooperation in public good situations[J]. *Journal of Public Economics*, 88(7-8): 1625-1644.
- Reno R R, Cialdini R B, Kallgren C A. 1993. The transsituational influence of social norms[J]. *Journal of Personality and Social Psychology*, 64(1): 104.
- Rössler C, Rusch H, Friehe T. 2019. Do norms make preferences social? Supporting evidence from the field[J]. *Economics Letters*, 183: 108569.
- Santos F C, Santos M D, Pacheco J M. 2008. Social diversity promotes the emergence of cooperation in public goods games[J]. *Nature*, 454(7201): 213-216.
- Solow J L, Kirkwood N. 2002. Group identity and gender in public goods experiments [J]. *Journal of Economic Behavior & Organization*, 48(4): 403-412.
- Szolnoki A, Perc M. 2010. Reward and cooperation in the spatial public goods game [J]. *EPL (Europhysics Letters)*, 92(3): 38003.
- Tabellini G. 2008. Presidential Address Institutions and Culture [J]. *Journal of the European Economic Association*. 6(2-3): 255-94.
- Tajfel H, Billig M G, Bundy R P, et al. 1971. Social categorization and intergroup

- behaviour[J]. *European Journal of Social Psychology*, 1(2): 149-178.
- Tajfel H. 1974. Social identity and intergroup behaviour[J]. *Social Science Information*, 13(2): 65-93.
- Terry D J, Hogg M A, White K M. 1999. The theory of planned behavior: Self-identity, social identity and group norms[J]. *British Journal of Social Psychology*, 38(3): 225-244.
- Zizzo D J. 2010. Experimenter demand effects in economic experiments[J]. *Experimental Economics*, 13(1): 75-98.

The Influence Mechanism of Social Norms on Cooperation: An Elicitation of Internalization Effect and Identity Effect

Chi Cui¹ Lulu An² Ming Dai³

(1. School of Economics and Management, Northeast Normal University;

2. Business School, Beijing Normal University;

3. Faculty of Economics and Social Sciences, University of Heidelberg)

Abstract The idea that people's cooperative behavior in games is driven by social norms has recently gained academic consensus. In this study, we constructed a social norm experiment in which participants were divided into a social norm compliance group and a non-social norm compliance group to further decompose the two common mechanism effects that promote cooperation based on social norms: Internalization effect and social identity effect. The results show that: (1) Both the internalization effect and the social identity effect can effectively improve the level of social cooperation on the whole, but the role of maintaining high level of social cooperation is limited; (2) With the enhancement of compliance with social norms, the dominant role of social identity effect in promoting social cooperation is gradually replaced by internalization effect; (3) The internalization effect plays an important role in maintaining high level of cooperation with the increase of compliance with social norms, while the social identity effect plays a better role in maintaining or even reducing the decline rate of average cooperation rate, but it still cannot reverse the rapid decline trend of average cooperation rate in the short term. This paper uses laboratory experiments to provide a new perspective for the study of social norms promoting social cooperation.

JEL Classification C92, H41, Z13