

共识铸就合作^{*}

——基于和平制造者博弈的实验研究

姜树广¹ 张正平²

摘要 个体间或国家间的合作往往受到第三方干预的影响。其中,观点的共识对于合作达成以及第三方介入至关重要。本研究在基于和平制造者博弈的框架下,引入了基于观点的分组,以探究共识对合作的影响。研究发现,在不同实验局及不同对手情境下,参与者的行为存在显著差异,这与理性经济人假设下的预测不符。进一步的分析表明,共识的形成主要通过产生共同的群体身份,影响个体与他人的社会亲近程度,从而影响个体的合作行为和第三方的干预行为。此外,通过增加共同的群体身份,当事方的合作行为得到促进,同时第三方的干预也得到增加,从而提高了整体社会效率。

关键词 共识;合作;第三方干预;和平制造者博弈;社会亲近程度

0 引言

近年来,国际竞争与合作问题备受关注。在2020年联合国成立75周年纪念峰会上,多国元首和政府首脑强调了多边主义和国际合作的重要性。习近平主席在2022年6月的全球发展高层对话中也强调:“合作才能办成大事,办成好事,办成长久之事。”随着中国国际地位的不断提升,中国在国际事务中越来越多地扮演合作发起方和冲突调解方的角色,例如2013年秋季作为“一带一路”倡议的发起方,以及2023年3月作为沙特阿拉伯和伊朗恢复外交关系的见证方。在沙伊复交的过程中,中方推动三方形成对未来发展的共识起着重要作用。

尽管传统国际关系理论和经典经济学逻辑将国际关系视为基于理性利益算计的博弈,在当代国际合作与竞争中各方观念差异的影响日益明显。特别地,在多方博弈的情况下,各方间是否存在某种程度的观念共识对于各方的策略选择尤为重要。因此,深入研究共识对合作的影响,有助于拓展我们对于国际合作机制的认识,为构建更加和谐、稳定的国际关系提供理论支撑和实践指导。

为了刻画存在第三方影响的合作场景,本研究借鉴了和平制造者博弈(Halevy and Halali, 2015)的基本框架。在和平制造者博弈中,两个对称的当事方与一个第三方需同时

^{*} 本研究受国家社科基金一般项目“合作与冲突:基于演化博弈与实验经济学的研究”(21BJL126)的资助。作者感谢编辑部老师和两位匿名审稿专家对本文提出的富有建设性的评论和修改意见,文责自负。

1 姜树广(通信作者),山东大学经济研究院教授, E-mail: sgjiang@sdu.edu.cn。

2 张正平,山东大学经济研究院博士研究生, E-mail: zhangzhengping@mail.sdu.edu.cn。

做出决策,当事方的选择为合作或冲突,第三方的选择为干预或旁观,第三方不同的选择决定着当事方不同的博弈支付结构。在该博弈中,第三方选择干预(相较于旁观)具有一定的风险,但会奖励选择合作的当事方并惩罚选择冲突的当事方,因此第三方的干预也可以被视为一种广义的合作行为。

为了研究共识对合作行为的影响,我们在和平制造者博弈的基础上引入了不同形式的分组。共识意味着不同主体对某一话题持有相同或相近观点。在本研究中,我们通过让被试针对某一社会话题选择观点并进行讨论,人为地在实验室环境中制造观点的共识与分歧。如果对这一简单话题的观念共识对合作行为产生影响,那么现实世界不同人群间巨大的信仰、价值观念和意识形态的差异则具有更大的影响。在“观点分歧”实验局中,被试依据对话题的观点倾向被划分为对立的两组。通过策略方法,我们构造了博弈对手方来自相同或不同观点分组的多种独立场景,被试需在各场景中分别做出和平制造者博弈的相关决策。

为了识别由观念共识形成的群体身份的作用大小,本研究进一步通过被试间设置引入仅由随机分组产生群体身份的“最小分组”实验局^①,以及不进行分组的“控制”实验局。我们希望探究,这种观念共识在多大程度上加强了个体的身份认同,及对合作与干预行为产生多大程度的影响。基于 Goette and Tripodi(2021)的一体性量表(Oneness Scale),测度了在“观点分歧”实验局与“最小分组”实验局中,被试对同组、异组成员的社会亲近程度(对自己与他人之间社会距离的感知),以深入理解共识对合作的影响机制。

本研究发现以下结论:首先,与众多研究群体身份导致群体内偏爱及群体外偏见的文献结论相一致(Chen and Li, 2009; Goette et al., 2012),个体更容易对同组成员表现出亲社会性(当事方的合作及第三方的干预)。其次,不同形式的分组影响了个体作为当事方的合作行为,作为第三方的干预行为,以及个体对同组、异组成员的社会亲近程度。具体来说,在基于共识的分组中,个体的合作与干预倾向更强且对异组成员有着更低的社会亲近程度(相较于随机的分组)。此外,基于观点共识所构建的群体身份能够更好地促使当事方合作,并吸引第三方的干预,从而提高社会效率。然而,当两名当事方在观点上存在分歧时,这种分歧则会破坏合作,进而损害社会效率。

本研究对现有文献的可能贡献主要体现在以下几个方面。第一,本研究在观念共识的视角下拓展了合作影响因素的研究。人类合作难题作为重大的社会科学问题吸引了国内外学者的广泛关注(见黄少安和张苏, 2013; 韦倩和姜树广, 2013 综述)。实验经济学文献在微观层面对人类合作的决定因素与影响机制做出了大量探索(周晔馨等, 2014; 姜树广和陈叶烽, 2016; 姚澜和朱迅, 2020; Chaudhuri, 2011; Cooper and Weber, 2020)。文献中考察的重要合作影响因素包括惩罚和奖励、交流、群体规模、领导角色等,少有研究从观点共识与分歧的视角来分析认知观念对合作行为的影响。

第二,在以往关于合作的文献中尽管对第三方角色的作用进行了大量研究,这些研究主要聚焦于作为独立第三方的利他惩罚者在解决社会困境中的作用(Fehr and Gächter, 2002; Fehr and Fischbacher, 2004; Kamei, 2020)。考虑可能利益相关第三方的角色, Evdokimov and Garfagnini(2018)和 Hörner et al. (2015)构建的和平制造者博弈刻画了自

^① 在实验研究中,这一概念主要包括被试随机分为无重叠的组、被试之间无互动、分组是匿名等方面(Tajfel and Turner, 1986)。

利且利益相关的第三方如何通过改变博弈结构影响合作的问题。在此基础上,本研究进一步考察了第三方与当事方在身份或观念上的异同如何影响第三方行为及当事双方合作的实现,拓展了第三方研究的视角,有助于更好地认识现实世界中的合作与干预行为。

第三,本研究还在第三方干预和合作背景下对当前学术界高度活跃的身份认同(见 Charness and Chen, 2020 综述)和极化(Waller and Anderson, 2021; Dimant, 2024)的研究领域做出边际贡献。近年来,情感极化问题日益引起政治学、心理学和经济学的关注。现有研究主要基于西方的政治背景,发现由于党派身份的差异导致对同党派的积极情感和对对立派别的消极情感(McConnell et al., 2018; Iyengar et al., 2019)。本研究发现,即使没有政治身份的差异,仅仅由于对单一社会话题的看法差异也可以导致对两派成员间的情感差异。这对在更广泛的人群中认识情感极化问题具有重要意义。

本文的结构安排如下:第1节回顾相关的文献,第2节详细介绍了实验设计并提出研究假设,第3节对实验结果进行分析,第4节为总结与展望。

1 文献回顾

本研究主要与以下三个领域的文献密切相关:一是群体身份与合作的关系,二是有关情感极化影响个体行为的研究,三是第三方在合作中的作用。

1.1 群体身份与合作

持有相同的观点本身会形成基于观点的群体身份,因此本研究与众多对群体身份的文献相关。相关文献广泛探讨了群体身份与合作之间的关系,发现群体身份主要通过使个体形成对内群体成员的一般互惠预期来影响合作(Yamagishi and Kiyonari, 2000),因此相较于外群体成员,人们更愿意与内群体成员进行合作(Chen and Li, 2009; Goette et al., 2012),并会报复那些不与内群体成员合作的人(Goette et al., 2012)。本研究使用的和平制造者博弈存在两个纯策略纳什均衡,且其中一个帕累托优于另一个,因此在一定程度上具有协调博弈的性质,而 Chen and Chen(2011)以及 Chen et al. (2014)通过最小努力博弈发现,共同的群体身份有助于在协调问题中达成更好的均衡结果^①。然而,以往的研究主要集中在共同的群体身份如何影响个体行为上,而本研究的重点在于研究共同的群体身份以及不同群体之间的群体身份对立如何影响个体的行为。

1.2 情感极化如何影响个体行为

极化问题是当前社会科学领域高度关注的热点问题,也是当代西方国家面临的重大现实问题。大量研究考察了不同身份群体的观念极化如何影响个体行为(Waller and Anderson, 2021; Flamino et al., 2023; Dimant, 2024)。例如,Dimant(2024)从民众对时任美国总统特朗普的好恶出发,研究了政治极化对社会偏好的影响,发现个体不愿意同与自己政治倾向对立的人合作。这种对不同群体身份人不同的态度的极端化体现为情感极化,即对同群体的人产生好感而对不同群体身份的人产生敌意。Iyengar et al. (2019)对情

^① 更多有关群体身份与合作的文章,可参见 Charness and Chen(2020)关于群体身份的综述。

感极化的起源以及后果进行了全面梳理,认为情感极化起源于党派成员将党派身份作为一种社会认同,而这种极化会影响人们的社会关系和经济行为。政治宗派主义(Finkel et al., 2020)的概念进一步概括了身份认同与情感极化的现象,其包括三个核心要素:他者化(Othering),即将对立党派视为与自己本质不同或格格不入的倾向;厌恶(Aversion),即讨厌和不信任对立党派的倾向;道德化(Moralization),即将对立党派视为邪恶的倾向。现有研究主要关注不同群体(例如美国的共和党 and 民主党)之间的行为差异,而我们的研究强调观点的共识与对立本身通过引发情感差异而影响个体在合作方面的行为。

1.3 第三方在合作中的作用

第三方在合作和冲突中也发挥着关键作用。Hörner et al. (2015)通过理论证明,即使第三方提出的建议不具有强制执行力,这些建议也会促进合作的达成。Evdokimov and Garfagnini(2018)在实验中发现,第三方可以通过向其他玩家发送公开信号来创建“焦点”,从而改变其他博弈方的预期并影响博弈结果。Bland and Nikiforakis(2015)的实验结果表明,其他博弈方的行为带给第三方的外部性大小与方向会显著影响博弈结果。此外,第三方惩罚也在促进合作方面发挥着积极作用(Carpenter and Matthews, 2012)。一方面,第三方愿意惩罚违反合作规范的行为,即使这些惩罚需要付出成本(Krasnow et al., 2016)或者承担风险(Nakashima et al., 2017)。另一方面,由于其他博弈方预期到第三方将惩罚不合作的行为,因此第三方的存在增加了他们合作的倾向(Halevy and Halali, 2015)。而第三方的惩罚行为会受到群体身份的影响,群体身份使得个体更关心内群体成员的福祉,这导致第三方更愿意惩罚那些损害内群体成员利益的不合作者(Bernhard et al., 2006; Goette et al., 2012)。

鉴于第三方在冲突调解和促进合作中的重要作用,本研究主要采用 Halevy and Halali(2015)提出的和平制造者博弈作为实验框架,该博弈包含当事方的合作、冲突以及第三方的干预。此外,本研究还通过引入不同形式的分组产生的群体身份,探究观点共识、群体身份等因素对合作和干预行为的影响以及相应的作用机制。

2 实验设计与研究假设

2.1 基本框架:和平制造者博弈

本实验的基本框架基于 Halevy and Halali(2015)提出的和平制造者博弈,并参考了 Nakashima et al. (2017)的参数设置。在该博弈中,两个对称的当事方与一个第三方需要同时行动。每个当事方都需要在合作和冲突中选择一项,第三方需要在干预和旁观中选择一项。图1展示了博弈各方在不同博弈结果下的收益情况,图1(a)、图1(b)分别展示了第三方旁观或干预情况下两名当事方不同行动组合对应的实验币收益(每个数字组合中的两个数字分别表示当事方1、当事方2在该行动组合下的收益),图1(c)展示了不同行动组合下第三方的实验币收益。为了消减实验者需求效应,实验过程中使用了中性语言对相关词语进行了替代。在该博弈中,若第三方选择旁观,则冲突双方实际上在进行囚徒博弈。若第三方选择干预,则当事双方选择合作均成为最优策略。第三方的选择干预通过奖励选择合作的一方,惩罚选择冲突的一方改变博弈的结构,同时将自身利益与冲突双方的行为进行绑定。

当事方2的选择				当事方2的选择			
		合作	冲突			合作	冲突
当事方1 的选择	合作	300,300	100,400	当事方1 的选择	合作	400,400	200,300
	冲突	400,100	200,200		冲突	300,200	100,100

(a) 第三方选择旁观时，当事方的选择与收益

(b) 第三方选择干预时，当事方的选择与收益

第三方的选择			
		旁观	干预
当事方1 与当事方 2的选择	合作 合作	250	400
	冲突 合作	250	250
	合作 冲突	250	250
	冲突 冲突	250	100

(c) 当事方不同选择组合时，第三方的选择与收益

图1 和平制造者博弈的收益结构

与其他包含第三方的合作博弈框架相比(如 Kamei, 2020), 和平制造者博弈具有以下两个特点。首先, 第三方与其他博弈方的选择是同时进行的, 它需要选择是否介入其他博弈方的博弈过程, 而非事前制定规则或事后惩罚或奖励其他博弈方。其次, 第三方如果选择干预, 那么它的收益将会受到当事方选择的影响。因此, 第三方的决策不仅涉及社会偏好, 还包括对自身利益的考虑, 以及对两名当事方所做选择的信念。这一特点更能反映现实世界中的第三方关于声誉的考虑。例如, 在人际关系中, 中间人在为他人合作牵线搭桥时, 会考虑每次合作的成败对自身口碑的影响; 在国际关系中, 一些国家或国际组织在为他国间的合作达成或冲突化解进行背书时, 也会重点考虑这种支持将如何影响自身的国际影响力。

2.2 实验局设置

为探究观点共识所形成的分组如何影响个体的合作及干预行为, 本研究采用了被试间和被试内设计相结合的方式进行实验局的设置。在被试间的设计中, 不同的实验参与者样本被随机分配到“观点分歧”实验局、“最小分组”实验局以及“控制”实验局三种不同的实验处理中。在被试内的设计中, 每一个实验局中的同一个参与者在不同的决策场景中就以上的和平制造者博弈做出选择, 而其中场景的差异依赖于实验局设置形成的分组。

在“控制”实验局中, 每位被试需按照策略方法, 分别按照作为当事方(场景 A)或第三方的角色(场景 B)做出和平制造者博弈的选择。通过在实验局的选择, 可以得到在没有任何群体身份的影响下, 个体在和平制造者中的合作和干预行为, 并作为比较的基准。

在“最小分组”实验局中, 被试首先被计算机程序随机分配到两组, 并需要按照策略方法在 7 种不同的分组场景中分别进行和平制造者博弈的决策。换句话说, 每位参与者在假定个人和对手方不同分组身份的 7 种不同场景中分别做出 7 次决策, 具体的场景设置如表 1。在场景 1 至 4 中, 被试作为当事方进行决策, 根据前期随机分组的相同或不同, 所面对的另一当事方以及第三方与自己所在分组之间的关系分别为(相同, 相同)、(不同, 不同)、(不同, 相同)、(相同, 不同)。在场景 5 至 7 中被试作为第三方进行决策, 根据前期随

机分组的相同或不同,所面对的两名当事方与自己所在分组之间的关系分别为均相同、均不同、一方相同一方不同。与“控制”实验局不同,“最小分组”实验局的参与者可能因博弈对手方是来自与自己同一分组或不同分组的身份不同而产生情感的差异,从而影响合作制造者博弈中的选择。

表 1 实验局与场景设置

实验局	场景	角色	场景描述(决策者关系)
“观点分歧”实验局	1	当事方	另一当事方观点相同;第三方观点相同
	2	当事方	另一当事方观点不同;第三方观点不同
	3	当事方	另一当事方观点不同;第三方观点相同
	4	当事方	另一当事方观点相同;第三方观点不同
	5	第三方	当事双方观点都相同
	6	第三方	当事双方观点都不同
	7	第三方	当事一方观点相同,一方观点不同
“最小分组”实验局	1	当事方	另一当事方同组;第三方同组
	2	当事方	另一当事方不同组;第三方不同组
	3	当事方	另一当事方不同组;第三方同组
	4	当事方	另一当事方同组;第三方不同组
	5	第三方	当事双方都同组
	6	第三方	当事双方都不同组
	7	第三方	当事一方同组,另一方不同组
“控制”实验局	A	当事方	无情境
	B	第三方	无情境

在“观点分歧”实验局中,被试首先需要针对一个有争议的话题“当下是否应该在公共场合佩戴口罩”选择观点倾向并展开讨论。被试需要针对该话题在公共聊天窗口中进行 400 秒的讨论,在讨论开始前和结束后,每名被试都要在 1 至 6 的程度上选择观点倾向(1 表示完全不用戴,6 表示必须强制戴),在讨论过程中屏幕上还会轮番展示两方网友的观点。我们根据被试在讨论后最终报告的观点倾向将其分为观点倾向对立的两组,其中选择程度为 1、2、3 的被试为一组,而选择程度为 4、5、6 的被试为另一组^①。

基于该观点分歧,每位实验参与者需要按照策略方法在 7 种不同的分组场景中进行和平制造者博弈的决策,具体的场景设置如表 1。与“最小分组”实验局类似,在场景 1 至 4 中被试作为当事方进行决策,根据前期观点的相同或不同,所面对的另一当事方以及第三方与自己所在分组之间的关系分别为(相同,相同)、(不同,不同)、(不同,相同)、(相同,不同)。在场景 5 至 7 中被试作为第三方进行决策,根据前期观点的相同或不同,所面对的两名当事方与自己所在分组之间的关系分别为均相同、均不同、一方相同一方不同。

与“最小分组”实验局中的分组身份仅仅由随机分配产生不同,“观点分歧”实验局的设置使我们得以诱导出更强的对内外群体的情感差异。我们允许被试在初次报告态度后进行 400 秒的充分讨论和信息交流。通过在该环节展示正反两方面的不同观点,可以激发被试对持有不同观点的其他人的情感反应。我们以最终报告的观点作为分组的依据,

① “观点分歧”实验局中的 54 名被试在讨论结束后,观点倾向于支持不戴口罩与支持戴口罩的人数分别为 26 人与 28 人。

从而被试感知到的分组差异为最终不可调和的观点分歧。也就是说,最终的分歧是他们在全面地了解支持或反对某一观点的信息和进行充分的讨论后仍然坚持的分歧,而不是由于获取信息的差异导致的认知差异,这更可能引发强烈的对内外群体情感差异。由于“观点分歧”实验局的设计目的是探究观点分歧所带来的情感如何影响合作与第三方干预行为,我们并不关心讨论的过程以及这个过程是否会导致观点的改变,而只需通过这个过程的最最终分歧进行分组。此外,为了尽可能地控制外部信息对被试的差异化扰动,我们所选取的两方网友观点的文本长度与信息量基本一致,正反方观点交替展示。并且在同一时刻,所有被试在屏幕上所展示的信息完全一致。

2.3 实验实施

本研究在预实验后通过 AsPredicted 网站对实验进行了预注册(预注册号:126835)。实验使用 Z-Tree 软件(Fischbacher,2007),于2023年3月至4月在山东大学进行,所有实验局的样本均为山东大学的大学生,他们来自不同的年级与专业。在每场实验前,我们在控制男女性别平衡的情况下,通过计算机随机抽取未参加过此场实验(或与本研究主题相近的实验)的学生,并向他们发送该场的实验邀请。因此,这些样本可以很好地代表山东大学的大学生群体。此外,由于我们研究的问题本身不涉及被试在现实中的特定身份背景,并且大学生样本较少受到社会经验等复杂因素的影响,因此更加符合我们的研究问题。

实验总共进行了8场:“观点分歧”实验局3场,“最小分组”实验局3场,“控制”实验局2场。每场被试人数均为18人,共招募144人。在每个实验局中,当全部阶段完成后,计算机将随机将每场实验的18名被试分成6个三人组,每组随机抽取2人作为当事方、1人作为第三方,并根据三人的分组情况匹配相应场景。被试到达实验室后,通过抽签的方式被随机分配座位,当一场实验的全部被试均到场后,实验员发放实验说明并大声朗读。为了确保每名被试理解实验规则,被试们必须独立回答3道控制性测试题。只有全部被试都通过3道控制性问题后,实验才正式开始。

为了确保被试可以注意到不同决策场景中其他博弈方分组情况的变化,我们在实验说明与决策环节开始前均向被试做了提醒。同时,在决策界面中,我们还特别用红色文字醒目标注了其他博弈方的分组情况。为了控制由于决策顺序差异可能导致的学习效应,在“最小分组”和“观点分歧”实验局中,我们为每名被试随机生成一个独立的7个场景呈现顺序。为了防止前期的结果对后续实验造成影响,在所有实验局的实验过程中被试均不会收到任何有关他人决策的反馈信息,从而被试不可能在前期场景的决策中学习到任何其他人的策略。为了控制可能的财富效应,我们使用了随机支付的方式使得被试在7个场景中的决策最终仅有一个实际得到执行。具体而言,每场实验的被试在全部实验决策完成后将被试随机配为3人一组,每组随机选取2人作为当事方、1人作为第三方,并根据三人的分组情况匹配相应场景,之后根据三人在相应场景中的决策核算三人的收益情况。

我们在实验后的问卷环节收集了被试的部分人口统计学特征以及使用 Falk et al.(2023)的主观态度问卷方法测度被试的社会偏好。我们在“观点分歧”与“最小分组”实验局的调查问卷中还使用了一体性分数(Goette and Tripodi,2021)分别测量了被试对同组与异组成员的社会亲近程度。各组被试人口统计学特征与社会偏好的描述性统计信息以及样本的平衡性检验如表2所示,可以看出人口统计学特征在实验局之间基本平衡。

参与者在“观点分歧”实验局的平均收益为 41.57 元人民币,“最小分组”实验局的平均收益为 42.87 元人民币,“控制”实验局的平均收益为 40.42 元人民币。每场实验时间约 40 分钟。

表 2 各实验局被试的人口统计学特征

变量名	“观点分歧” 实验局	“最小分组” 实验局	“对照” 实验局	Kruskal-Wallis 检验
性别(女性=0,男性=1)	0.444 (0.502)	0.407 (0.496)	0.528 (0.506)	1.272 (0.530)
专业(非经管=0,经管类=1)	0.315 (0.469)	0.278 (0.452)	0.444 (0.504)	2.814 (0.245)
年龄	21.056 (2.430)	21.000 (2.510)	21.000 (2.597)	0.027 (0.987)
样本量	54	54	36	

注:二至四列报告了每个人口统计学特征在各实验局中的均值,下方括号内的数值为对应标准差。最后一列报告了 Kruskal-Wallis 检验的调整 H 统计量($\chi^2_{d.f=2}$),下方括号内的数值为对应 p 值。

2.4 研究假设

在和平制造者博弈中,如果记当事方 1 选择合作的概率为 p (选择冲突的概率为 $1-p$),当事方 2 选择合作的概率为 q (选择冲突的概率为 $1-q$),第三方选择旁观的概率为 r (选择干预的概率为 $1-r$),那么有序数对 (p, q, r) 便可以表示该博弈的一个混合策略组合。

对于一个混合策略组合 (p, q, r) ,如果每个博弈方都预期他人会使用该组合中的策略,那么当事方选择合作严格占优于(严格劣于,无差异于)冲突,当且仅当 $r < 0.5 (r > 0.5, r = 0.5)$;第三方选择干预严格优于(严格劣于,无差异于)旁观,当且仅当 $p + q > 1 (p + q < 1, p + q = 1)$ 。进一步的分析可知,该博弈存在(合作,合作,干预)与(冲突,冲突,旁观)两个纯策略纳什均衡,以及一系列混合策略纳什均衡 $(\tilde{p}, \tilde{q}, 0.5)$,其中 $\tilde{p}$ 与 $\tilde{q}$ 满足 $\tilde{p} + \tilde{q} = 1$ 。

尽管博弈理论无法预测哪一个均衡会被实现,但每个博弈方的占优策略取决于他对其余博弈方所做选择的预期。由上面的分析可知,当事方的占优策略仅取决于他预期第三方选择旁观的概率 $\hat{r}$,而不取决于他预期另一当事方所做选择;第三方的占优策略取决于他预期两名当事方选择合作的概率之和 $\hat{p} + \hat{q}$ 。

尽管在“观点分歧”实验局与“最小分组”实验局中,对于社会争议话题所持观点或随机分组并不包含他人合作或干预概率的信息,但个体可能会预期同组成员的选择更加具有亲社会性。因此,当事方会预期与自己同组的另一当事方更有可能合作,与自己同组的第三方更有可能干预,第三方会预期与自己同组的当事方更有可能合作。在假定个体理性自利的条件下(理性经济人假设),结合上述预期与前文对博弈占优策略的讨论,我们有如下假设:

假设 1(经典假设): 在“观点分歧”实验局与“最小分组”实验局中,当事方面对同组第三方时的合作率高于面对异组第三方,但不受另一当事方所在分组的影响;第三方的干预率随着同组当事方的增多而增加。

在经典的理性经济人假设下,个体只关注自身物质利益。但行为经济学认为个体在

做出决策时并不仅追求自身的物质利益,同时也考虑他人的利益。因此,互惠、利他等因素也可能会影响个体的选择。从群体身份相关的理论来看,任何分组形式都会产生群体身份,这可能导致个体更为关心同组成员的福利,而对异组成员的福利更加冷漠。对当事方而言,他的选择必然会影响另一当事方的收益,但可能不影响第三方的收益(因为第三方有可能选择旁观),因此他的选择更多地受到另一当事方所在分组的影响。对第三方而言,他更愿意在面对与自己同组的当事方时选择干预。除此之外,“观点分歧”实验局相较于“最小分组”实验局,可能会产生更强的群体身份感。因此我们做出如下假设:

假设 2(行为假设):在“观点分歧”实验局与“最小分组”实验局中,当事方面对同组另一当事方时的合作率高于面对异组当事方,但几乎不受第三方所在分组的影响。第三方的干预率随着同组当事方的增多而增加。并且这一现象在“观点分歧”实验局中要比在“最小分组”实验局中更为明显。

尽管基于观点的分组与随机分组都会产生群体身份,但前者可能会使个体对同组成员产生更多的共鸣,从而产生更强的群体身份感。因此,观点分组使个体对同组、异组成员的社会亲近程度差异更大。为检验上述说法,我们通过一体性分数来度量同组与异组成员的社会亲近程度,并做出如下假设:

假设 3(行为假设):相较于“最小分组”实验局,“观点分歧”实验局中的个体对同组成员的一体性分数更高,对异组成员的一体性分数更低。

3 实验结果

3.1 分组形式对个体行为的影响

表 3 报告了作为当事方的合作率在各种场景中的平均值,图 2 和图 3 进一步展示了被试内与被试间行为差异的检验结果。在没有任何身份场景的“控制”实验局中,被试在 69%的情况下选择了合作。只要配对的另一当事方为来自同一组的成员(场景 1 和场景 4),“观点分歧”实验局的合作率都高于“控制”实验局且该差异在统计上是显著的,“最小分组”实验局的合作率也高于“控制”实验局但差异不显著。当配对的另一当事方是非同组成员时(场景 2 和场景 3),“观点分歧”实验局的合作率均低于“控制”实验局,且“最小分组”实验局在场景 3 的合作率也低于“控制”实验局,但差异均不显著。

表 3 各实验局与场景中的当事方合作行为

实验局	场景	另一当事方、第三方 所在的分组情况	合作率	标准差	样本量
“观点分歧”实验局	1	(相同,相同)	0.870	0.339	54
	2	(不同,不同)	0.648	0.482	
	3	(不同,相同)	0.611	0.492	
	4	(相同,不同)	0.870	0.339	
“最小分组”实验局	1	(相同,相同)	0.815	0.392	54
	2	(不同,不同)	0.759	0.432	
	3	(不同,相同)	0.667	0.476	
	4	(相同,不同)	0.722	0.452	
“控制”实验局	A	—	0.694	0.467	36

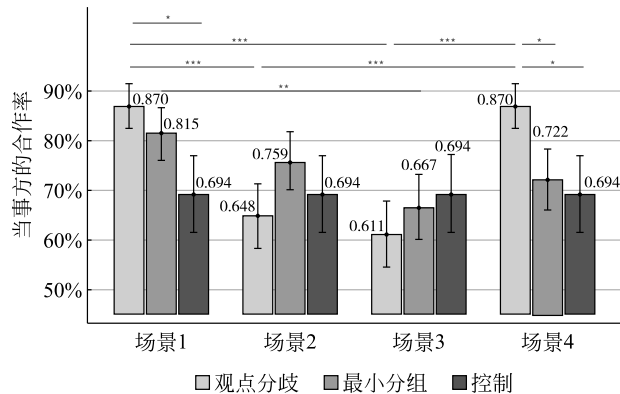


图 2 不同实验局与场景下的当事方合作率

注:数字标注了均值,误差条表示平均值的标准误。***、**、* 分别表示 1%、5% 和 10% 的水平上显著,同一实验局在不同场景间的比较使用配对样本双侧 t 检验,同一场景在不同实验局间的比较使用双总体双侧 t 检验。

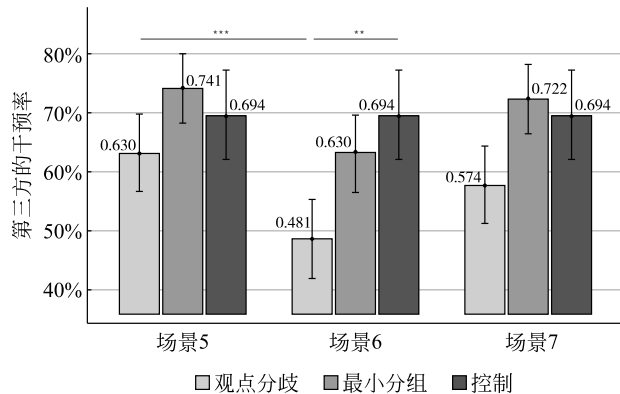


图 3 不同实验局与场景下的第三方干预率

注:数字标注了均值,误差条表示平均值的标准误。***、**、* 分别表示相应检验在 1%、5% 和 10% 的水平上显著,同一实验局在不同场景间的比较使用配对样本双侧 t 检验,同一场景在不同实验局间的比较使用双总体双侧 t 检验。

在“观点分歧”实验局中,不论第三方的身份如何,当事方面对同组另一当事方的合作率均显著高于面对异组另一当事方(场景 1 vs. 场景 2、场景 1 vs. 场景 3、场景 2 vs. 场景 4、场景 3 vs. 场景 4 均存在显著差异)。在“最小分组”实验局中,仅当第三方来自同组的情况下,当事方在面对同组另一当事方时的合作率才会显著高于面对异组另一当事方(场景 1 vs. 场景 3)。

表 4 反映了作为第三方的干预率在各实验局的平均值。图 2 和图 3 进一步展示了被试内与被试间行为差异的检验结果。在没有任何身份场景的“控制”实验局中,第三方在 69% 的情况下选择了进行干预。“观点分歧”实验局在各种场景下的干预率均低于“控制”实验局,且在配对的两名当事方均来自不同观点组时呈现显著更低的干预率(48%, 场景 6)。“最小分组”实验局干预率在配对的两名当事方均来自不同分组时的干预率也低于“控制”实验局,而在其他场景的干预率则略高于“控制”实验局,但这些组间的差距均不显著。

在“观点分歧”实验局中,第三方的干预率会随着同组当事方增多而增加,且在面对两名

同组当事方时的干预率显著高于面对两名异组当事方(场景 5 vs. 场景 6)。在“最小分组”实验局中,尽管第三方的干预率会随着同组当事方增多而增加,但这一结果并不显著。

表 4 各实验局与场景中的第三方干预行为

实验局	场景	同组当事方的人数	干预率	标准差	样本量
“观点分歧”实验局	5	2	0.630	0.487	54
	6	0	0.481	0.504	
	7	1	0.574	0.499	
“最小分组”实验局	5	2	0.741	0.442	54
	6	0	0.630	0.487	
	7	1	0.722	0.452	
“控制”实验局	B	—	0.694	0.467	36

进一步地,我们对“观点分歧”与“最小分组”实验局中各场景间合作率或干预率的差值进行了比较。我们将每个实验局中的被试作为当事方的四种场景进行两两组合,得到六组差值;将被试作为第三方的三种场景进行两两组合,得到三组差值。九组差值在“观点分歧”实验局与“最小分组”实验局之间的对比与统计检验如图 4 所示。作为当事方时,通过“2-1”“4-2”“4-3”的比较可以看出,“观点分歧”实验局在面对同组当事方和面对异组当事方时合作率的差异显著大于“最小分组”实验局。作为第三方时,随着同组当事方人数的减少,第三方的干预率在“观点分歧”实验局中比“最小分组”实验局下降得多,尽管这一差异并不显著。

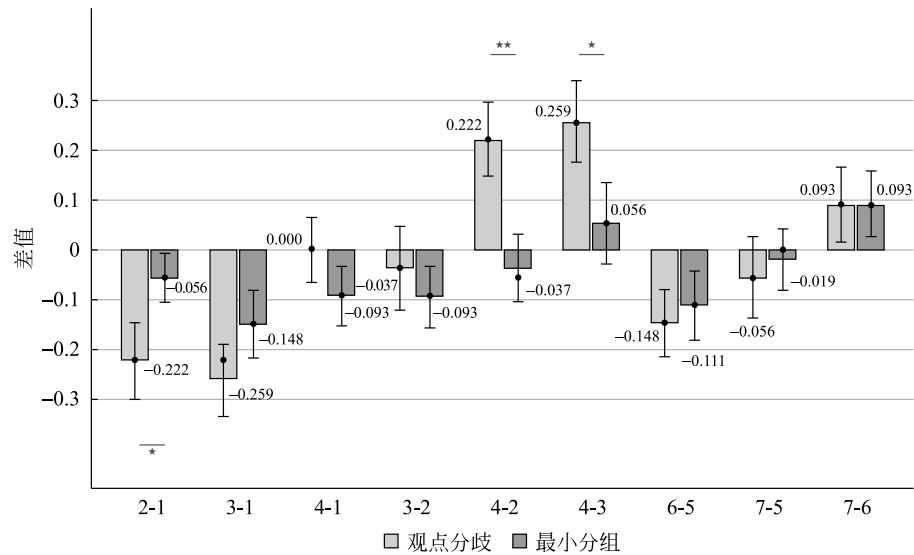


图 4 “观点分歧”与“最小分组”实验局各场景间合作率或干预率的差值

注:横轴的数字标注了相互比较的场景间的两两组合,图中数字表示在不同场景间当事方合作率(第三方干预率)差值(前一场景减后一场景)的均值,误差条表示差值均值的标准误。***、**、* 分别表示相应检验在 1%、5%和 10%的水平上显著,使用的检验为双总体双侧 *t* 检验。

接下来,本研究运用线性概率模型(LPM),在控制顺序效应、个人特征(性别、年龄、是

否为经管背景、是否具有策略性思考能力^①、社会偏好(信任、利他、正互惠、负互惠)的情况下,使用被试个体在一种场景下的决策作为被解释变量进行回归。在针对 7 个场景的每个回归中,我们将“控制”实验局作为参照组,“观点分歧”和“最小分组”实验局的虚拟变量作为主要解释变量。表 5 报告的回归结果与以上统计分析的初步结果一致,在场景 1 和场景 4 中,“观点分歧”实验局与“控制”实验局相比,个体的合作率显著更高。在场景 6 中,与“控制”实验局相比,“观点分歧”实验局的干预率显著更低。我们进一步检验了“观点分歧”实验局和“最小分组”实验局的差异,表 5 报告的结果显示当面对同组当事方时,“观点分歧”实验局的合作率比“最小分组”实验局更高(模型(1)和(4)),而面对异组当事方时,则比“最小分组”实验局更低(模型(2)和(3)),且该差异在模型(4)中是显著的。“观点分歧”实验局在所有场景下的第三方干预率均低于“最小分组”,但差异不显著,这表明在存在观点分歧的情况下第三方更不愿意干预的普遍倾向。

表 5 不同的实验局对个体行为的影响

场景	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	场景 1	场景 2	场景 3	场景 4	场景 5	场景 6	场景 7
被解释变量	当事方的行为(合作=1,冲突=0)				第三方的行为(干预=1,旁观=0)		
最小分组	0.156 (0.113)	0.098 (0.121)	-0.051 (0.128)	0.034 (0.118)	0.077 (0.138)	-0.163 (0.144)	0.045 (0.126)
观点分歧	0.212* (0.108)	-0.024 (0.117)	-0.079 (0.137)	0.214* (0.112)	-0.001 (0.143)	-0.297** (0.141)	-0.093 (0.129)
观点分歧-最小分组	0.055 (0.077)	-0.122 (0.097)	-0.027 (0.097)	0.180** (0.087)	-0.078 (0.097)	-0.134 (0.101)	-0.138 (0.101)
样本量	144	144	144	144	144	144	144
顺序效应	控制	控制	控制	控制	控制	控制	控制
个人特征	控制	控制	控制	控制	控制	控制	控制
社会偏好	控制	控制	控制	控制	控制	控制	控制

注:括号内为聚类到个体层面的稳健标准误。***、**和*分别表示在 1%、5%和 10%水平上显著。在场景 1~4 中,我们的分析使用了“最小分组”实验局、“观点分歧”实验局中对应场景下的行为数据,以及“控制”实验局中个体作为当事方的行为数据;在场景 5~7 中,我们的分析使用了“最小分组”实验局、“观点分歧”实验局中对应场景下的行为数据,以及“控制”实验局中个体作为第三方的行为数据。在每个场景中,基准组均为“控制”实验局。顺序效应位表示相关决策为被试第几个决策的虚拟变量。

值得说明的是,对于“观点分歧”实验局中所有场次的 54 名被试,在讨论前仅有 2 人持有极端的观点(2 人选择 1,无人选择 6),讨论后仅有 1 人持有极端的观点(1 人选择 1,无人选择 6)。这说明尽管观点的分歧并不严重,但是由分歧形成的两派共识仍会影响个体行为。此外,为了确保结论的稳健性,我们还运用 Logit 模型和 Probit 模型对上述回归分析进行了重新验证,结果与线性模型一致。

我们进一步运用线性概率模型,在控制顺序效应以及个体固定效应的情况下,使用被

① 被试在所有场景的和平制造者博弈实验结束后,还进行了如下的 2/3 选美博弈:每名被试需要独立地选择一个 0 至 100 之间的整数,最接近全部被试所选数字平均数 2/3 的参与者将获得 300 实验币的奖励。所选数字≤33 的被试具有策略性思考能力。

试内的数据分别分析了“观点分歧”实验局和“最小分组”实验局中不同场景对个体行为的影响。表 6 报告的回归结果显示,在两个实验局中,个体面对各种分组组合下的其他博弈方时,其行为存在显著差异,并且在“观点分歧”实验局中,这种差异更大。

表 6 不同的场景对个体行为的影响

实验局	(1)	(2)	(3)	(4)
	最小分组	观点分歧	最小分组	观点分歧
被解释变量	当事方的行为(合作=1,冲突=0)		第三方的行为(干预=1,旁观=0)	
场景 2	-0.052 (0.052)	-0.242 *** (0.079)		
场景 3	-0.146 ** (0.072)	-0.246 *** (0.074)		
场景 4	-0.088 (0.063)	0.010 (0.067)		
场景 6			-0.125 * (0.073)	-0.150 ** (0.071)
场景 7			-0.034 (0.071)	-0.035 (0.081)
样本量	216	216	162	162
个体固定效应	控制	控制	控制	控制

注:括号内为聚类到个体层面的稳健标准误。***、**和*分别表示在1%、5%和10%水平上显著。在模型(1)、(2)中,基准组为场景1,在模型(3)、(4)中,基准组为场景5。

接下来,我们进一步分析“观点分歧”实验局在剥离“最小分组”的影响后的净效应,即仅仅由于观点的异同而不是分组本身带来的影响。在此仅使用“观点分歧”与“最小分组”的数据并运用线性概率模型,在控制个人特征和社会偏好的情况下^①,以场景间的行为差异为被解释变量,以“最小分组”作为参照组,“观点分歧”实验局作为主要的解释变量进行回归,结果如表 7 所示。这一结果与图 4 的分析基本一致,表明“观点分歧”带来的影响并非仅仅由于分组本身,观点的异同造成的情感加强了分组身份的效应。

表 7 剔除最小分组影响的场景间差异分析

被解释变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	当事方合作率在场景间的差值						第三方干预率在场景间的差值		
对应两场景	2-1	3-1	4-1	3-2	4-2	4-3	6-5	7-5	7-6
观点分歧	-0.174 * (0.095)	-0.080 (0.099)	0.081 (0.096)	0.093 (0.113)	0.254 ** (0.105)	0.161 (0.118)	-0.049 (0.100)	-0.053 (0.106)	-0.005 (0.104)
样本量	108	108	108	108	108	108	108	108	108
个人特征	控制	控制	控制	控制	控制	控制	控制	控制	控制
社会偏好	控制	控制	控制	控制	控制	控制	控制	控制	控制

注:括号内为异方差稳健标准误。***、**和*分别表示在1%、5%和10%水平上显著。回归的基准组均为“最小分组”实验局。

① 由于被解释变量涉及 2 个场景,对应于 2 个不同的轮次,因此对顺序效应的控制不适用。事实上,我们通过随机顺序、策略方法等实验技术,已经消减了每一场景出现的顺序位置对被试行为的影响。

3.2 社会亲近程度对个体行为的影响

由上部分的分析可以看出,在不同形式的分组下以及在不同分组组合的场景中,个体的行为存在差异。本小节旨在探究个体行为差异是否由与其他博弈方的社会亲近程度不同所导致。根据 Goette and Tripodi(2021),我们用以度量社会亲近程度的一体性分数是被试对如下两个问题回应的平均分数:1)We 问题:在多大程度上愿意使用“我们”来描述自己与某人的关系(1 至 7,数字越大代表越愿意);2)IOS(Inclusion of Other in the Self)问题:在 7 组两两重叠的圆圈中选择最能代表他们与某人的关系的一组(1 至 7,数字越大重叠部分越多)。

图 5 展示了不同实验局中,被试对同组、异组成员的一体性分数,并使用 t 检验对一体性分数的差异进行检验。可以看出,无论在哪个实验局下,被试对同组成员的一体性分数都显著高于异组成员 1 分以上。被试对同组成员的一体性分数在“最小分组”与“观点分歧”实验局下几乎没有差异;被试在“观点分歧”实验局下对异组成员的一体性分数显著低于“最小分组”实验局,这一结果支持了假设 3 对异组成员社会亲近程度的预测,但却反驳了假设 3 对同组成员社会亲近程度的预测,说明不同的个体对他人的社会亲近程度会受到分组方式的影响,这可能是不同实验局中的个体行为存在差异的原因。

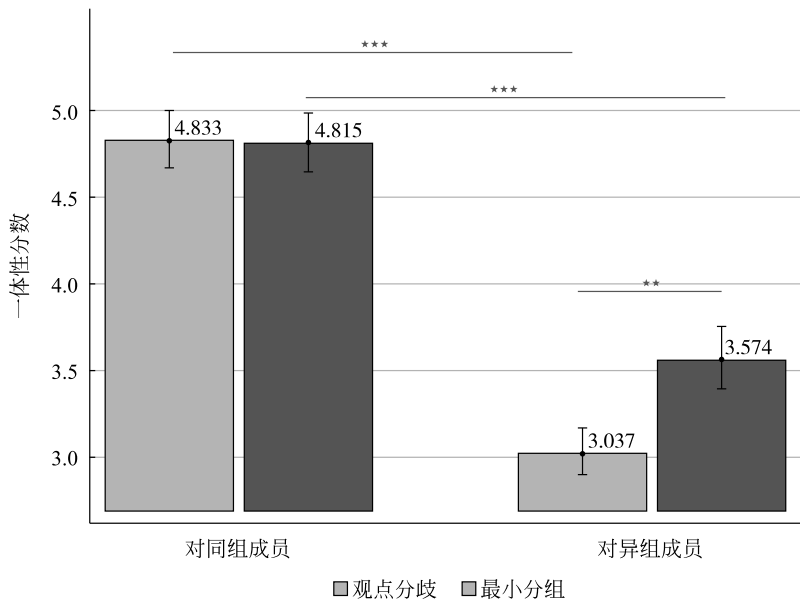


图 5 不同实验局中个体对同组、异组成员的一体性分数

注:数字标注了一体性分数的均值,误差条表示平均值的标准误。***、**、* 分别表示相应检验在 1%、5% 和 10% 的水平上显著。同一实验局对同组与异组成员一体性分数的比较使用配对样本 t 检验,不同实验局间对于同组(异组)成员一体性分数的比较使用双总体双侧 t 检验。

为具体分析社会亲近程度对个体行为所起的作用,本研究使用“观点分歧”实验局被试与“最小分组”实验局中 108 名被试的数据进行分析。首先,我们根据每名被试对同组、异组成员的一体性分数,为其作为当事方的 4 个决策场景生成了对应场景的“另一当事方

一体性分数”以及“第三方一体性分数”两个解释变量；同时，也为其作为第三方的3个决策场景生成了对应场景的“两名当事方一体性分数之和”这一解释变量。表8中的示例展示了一名对同组成员一体性分数为7，对异组成员一体性分数为1的被试在7个场景中对应的上述三个解释变量的赋值情况。

表8 不同场景下三个解释变量的赋值情况示例

场景	角色	另一当事方、第三方 所在的分组情况	另一当事方 一体性分数	第三方 一体性分数	两名当事方 一体性分数之和
场景1	当事方	(相同, 相同)	7	7	
场景2	当事方	(不同, 不同)	1	1	
场景3	当事方	(不同, 相同)	1	7	
场景4	当事方	(相同, 不同)	7	1	
场景5	第三方				14
场景6	第三方				2
场景7	第三方				8

本研究通过LPM分析,在控制了顺序效应、个人特征、社会偏好的情况后,使用当事方的行为对“另一当事方一体性分数”以及“第三方一体性分数”进行了回归,使用第三方的行为对“两名当事方一体性分数之和”进行了回归,结果如表9所示。从模型(1)可以看出,个体作为当事方时,他对另一当事方的一体性分数每增加1分,平均而言,其选择合作的可能性便增加5.63%;而他对第三方的一体性分数并不影响他合作的可能性。从模型(3)可以看出,个体作为第三方时,他干预的可能性随着他对两名当事方一体性分数之和的增加而上升。考虑社会亲近程度对个体行为所起的作用可能在不同形式的分组中存在异质性,模型(2)与模型(4)通过加入一体性分数与“观点分歧”实验局虚拟变量的交互项,进行了异质性分析。可以看出,加入交互项并不改变模型(1)与模型(3)中系数的符号与显著性,而且交互项的系数也不显著。说明给定个体对其他博弈方的一体性分数后,不管是哪一种实验处理导致的该一体性分数水平,一体性分数对个体行为的影响都是相同的,即社会亲近程度对个体合作行为所起的作用在各种分组形式之间不存在异质性。

表9 一体性分数对冲突与合作行为的影响

被解释变量	(1)	(2)	(3)	(4)
	当事方的行为 (合作=1, 冲突=0)		第三方的行为 (干预=1, 旁观=0)	
另一当事方一体性分数	0.056*** (0.018)	0.052*** (0.019)		
第三方一体性分数	-0.009 (0.018)	-0.008 (0.019)		
另一当事方一体性分数×“观点分歧”实验局		0.010 (0.013)		
两名当事方一体性分数之和			0.021* (0.013)	0.025* (0.013)

续表				
被解释变量	(1)	(2)	(3)	(4)
	当事方的行为 (合作=1,冲突=0)		第三方的行为 (干预=1,旁观=0)	
另一当事方一体性分数×“观点分歧”实验局				-0.012 (0.009)
常数项	0.389 (0.283)	0.410 (0.296)	0.351 (0.418)	0.293 (0.423)
样本量	432	432	324	324
顺序效应	控制	控制	控制	控制
个人特征	控制	控制	控制	控制
社会偏好	控制	控制	控制	控制

注:括号内为聚类到个体层面的稳健标准误。***、**和*分别表示在1%、5%和10%水平上显著。

3.3 分组对社会效率的影响

为了分析不同形式的分组对社会效率的影响,我们构建了一个效率指数,用于反映各种情况下的社会总效率。

首先,在“观点分歧”实验局被试与“最小分组”实验局中,如果我们用1、2表示不同的分组,并且考虑到当事方1与当事方2的对称性以及组1和组2的对称性,那么当事方1、当事方2、第三方的分组组合情况可以分成以下三类:(a)111、222,(b)112、221,(c)121、122、212、211。分别表示(a)博弈中的三方来自同组;(b)两名当事方来自同组,且均与第三方不同组;(c)第三方与一名当事方同组,且与另一名当事方不同组。不失一般性,我们将上述三类组合中当事方1所在的组规定为组1,那么三种分组组合可被记为111、112、121。然后,我们定义某个实验局与分组组合的效率指数为三个博弈方在该情况下的收益之和,即

Efficiency_Index = \sum_{j \in J} P_j \left(\sum_{i=1}^3 \text{payoff}_{ij} \right)

其中, J 表示所有可能的博弈结果, P_j 为某个博弈结果 j 发生的概率, $\sum_{i=1}^3 \text{payoff}_{ij}$ 表示在某个博弈结果 j 下三个博弈方的收益之和。例如,根据表3与表4中的数据可知,在最小分组111的情况下, $P_{(\text{合作}, \text{合作}, \text{干预})} = 0.82 \times 0.82 \times 0.74$, $\sum_{i=1}^3 \text{payoff}_{i(\text{合作}, \text{合作}, \text{干预})} = 400 + 400 + 400$ 。

各种情况下的效率指数如表10所示。可以看出,无论在“观点分歧”实验局还是在“最小分组”实验局中,三个博弈方均来自同一组时的效率指数最高,且高于“控制”实验局100个实验币左右。而在“观点分歧”实验局中,两名当事方来自不同组的效率指数最低,低于“控制”实验局超过50个实验币。而其他情况下的效率指数与“控制”实验局的差异均在20以内。

表 10 各种情况下的效率指数

实验局与分组组合	效率指数
观点分歧 111	987.311
观点分歧 112	879.287
观点分歧 121	828.018
最小分组 111	976.200
最小分组 112	874.588
最小分组 121	900.257
控制	883.410

综上,在和平制造者博弈中,个体对同组、异组成员的社会亲近程度不同,且会受到分组形式的影响。这种社会亲近程度的差异会导致个体在不同场景下的行为存在差异,具体表现为对同组成员更多的合作与干预行为,但社会亲近程度对个体合作行为所起的作用在各种分组形式之间不存在异质性。此外,构建共同的群体身份通常会提高社会效率,尤其是基于观点共识所构建的群体身份,对社会效率的提升效果最为明显。

4 结论及展望

本研究在和平制造者博弈的基础上引入分组概念,分析不同形式的分组以及与分组情况相关的不同场景对个体行为的影响。通过被试内对比,可以发现在和平制造者博弈中,当事方的合作意愿在另一当事方来自同组时高于来自异组,但却不受第三方来自哪一组的影响,第三方的干预意愿随着异组当事方数量的增加而减弱。通过被试间对比,可以发现上述的现象在“最小分组”实验局中更为显著。进一步的分析表明,不同形式的分组通过影响个体对他人的社会亲近程度的感知,进而影响个体在博弈中的合作与干预行为。此外,不同形式的分组与分组组合也对社会效率有影响,基于共识所形成的共同群体身份对提高社会效率具有积极意义。

本研究发现对个体层面的合作及冲突调解具有重要启示。首先,个体与同组成员互动时的行为更具有亲社会性,表明找到共同的群体身份是确保合作和干预冲突的重要措施。其次,随机分组形成的松散群体在促成合作及冲突调解中的作用可能十分有限,而努力挖掘各方的共识并塑造基于共识的群体身份可使彼此更加亲近,这有助于各方开展后续合作与实施第三方有效干预。

本研究发现对群体间合作以及冲突调解也具有一定的启发性。尽管群体身份在合作行为中的作用被广泛地探究(Chen and Li,2009),但与之相关的内群体偏爱、外群体歧视以及跨群体偏见等概念主要基于个体层面,因此相关理论可能无法为群体间的合作及冲突调解提供指导。而共识却在个体层面与群体层面普遍存在,尽管现实世界中的共识往往存在于群体内部,如宗教的教义信仰、民族的文化认同、党派的政治主张等,但不同的群体在面对同一问题时也可能存在相同或相似的观点,从而形成不同群体在同一问题上的共识。这些共识或许可以作为群体间进行合作或化解冲突的契机,但本研究基于个体间共识所得的结论是否适用于群体间冲突,需要进一步研究讨论。

参考文献

- 黄少安, 张苏. 2013. 人类的合作及其演进研究[J]. 中国社会科学, (7): 77-89.
- Huang S A, Zhang S. 2013. Study of the evolution of cooperative behavior[J]. *Social Sciences in China*, (7): 77-89. (in Chinese)
- 姜树广, 陈叶烽. 2016. 腐败的困境: 腐败本质的一项实验研究[J]. 经济研究, 51(1): 127-140.
- Jiang S G, Chen Y F. 2016. The dilemma of corruption: A lab experiment on the nature of corruption[J]. *Economic Research Journal*, 51(1): 127-140. (in Chinese)
- 韦倩, 姜树广. 2013. 社会合作秩序何以可能: 社会科学的基本问题[J]. 经济研究, 48(11): 140-151.
- Wei Q, Jiang S G. 2013. How is social cooperative order possible: Exploring mysteries[J]. *Economic Research Journal*, 48(11): 140-151. (in Chinese)
- 姚澜, 朱迅. 2020. 通过额外兑付改善协调困境: 一项实验研究[J]. 经济研究, 55(3): 163-178.
- Yao L, Zhu X. 2020. Pregame resource exchange as a mechanism in overcoming the coordination dilemma: An experimental study[J]. *Economic Research Journal*, 55(3): 163-178. (in Chinese)
- 周晔馨, 涂勤, 胡必亮. 2014. 惩罚、社会资本与条件合作——基于传统实验和人为田野实验的对比研究[J]. 经济研究, 49(10): 125-138.
- Zhou Y X, Tu Q, Hu B L. 2014. Social capital, institutional environment and conditional cooperation: An experimental study of VCM of public goods based on “Core + Shell” framework[J]. *Economic Research Journal*, 49(10): 125-138. (in Chinese)
- Bernhard H, Fehr E, Fischbacher U. 2006. Group affiliation and altruistic norm enforcement[J]. *American Economic Review*, 96(2): 217-221.
- Bland J, Nikiforakis N. 2015. Coordination with third-party externalities[J]. *European Economic Review*, 80: 1-15.
- Carpenter J P, Matthews P H. 2012. Norm enforcement: Anger, indignation, or reciprocity? [J]. *Journal of the European Economic Association*, 10(3): 555-572.
- Charness G, Chen Y. 2020. Social identity, group behavior, and teams[J]. *Annual Review of Economics*, 12: 691-713.
- Chaudhuri A. 2011. Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature [J]. *Experimental Economics*, 14(1): 47-83.
- Chen R, Chen Y. 2011. The potential of social identity for equilibrium selection[J]. *American Economic Review*, 101(6): 2562-2589.
- Chen Y, Li S X. 2009. Group identity and social preferences[J]. *American Economic Review*, 99(1): 431-457.
- Chen Y, Li S X, Liu T X, et al. 2014. Which hat to wear? Impact of natural identities on coordination and cooperation [J]. *Games and Economic Behavior*, 84: 58-86.
- Cooper D J, Weber R A. 2020. Recent advances in experimental coordination games[M]//Capra C M, Croson R T A, Rigdon Mary L, et al. *Handbook of Experimental Game Theory*. Cheltenham: Edward Elgar Publishing, 149-183.
- Dimant E. 2024. Hate trumps love: The impact of political polarization on social preferences[J]. *Management Science*, 70(1): 1-31.
- Evdokimov P, Garfagnini U. 2018. Third-party manipulation of conflict: An experiment [J]. *Experimental Economics*, 21(1): 27-49.
- Falk A, Becker A, Dohmen T, et al. 2023. The preference survey module: A validated instrument for measuring risk, time, and social preferences[J]. *Management Science*, 69(4): 1935-1950.

- Fehr E, Gächter S. 2002. Altruistic punishment in humans[J]. *Nature*, 415(6868): 137-140.
- Fehr E, Fischbacher U. 2004. Third-party punishment and social norms[J]. *Evolution and Human Behavior*, 25(2): 63-87.
- Finkel E J, Bail C A, Cikara M, et al. 2020. Political sectarianism in America: A poisonous cocktail of othering, aversion, and moralization poses a threat to democracy[J]. *Science*, 370(6516): 533-536.
- Fischbacher U. 2007. z-tree: Zurich toolbox for ready-made economic experiments[J]. *Experimental Economics*, 10(2): 171-178.
- Flamino J, Galeazzi A, Feldman S, et al. 2023. Political polarization of news media and influencers on Twitter in the 2016 and 2020 US presidential elections[J]. *Nature Human Behaviour*, 7(6): 904-916.
- 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): 101-115.
- Goette L, Tripodi E. 2021. Social influence in prosocial behavior: Evidence from a large-scale experiment[J]. *Journal of the European Economic Association*, 19(4): 2373-2398.
- Halevy N, Halali E. 2015. Selfish third parties act as peacemakers by transforming conflicts and promoting cooperation [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 112(22): 6937-6942.
- Hörner J, Morelli M, Squintani F. 2015. Mediation and peace [J]. *The Review of Economic Studies*, 82(4): 1483-1501.
- Iyengar S, Lelkes Y, Levendusky M, et al. 2019. The origins and consequences of affective polarization in the United States[J]. *Annual Review of Political Science*, 22: 129-146.
- Kamei K. 2020. Group size effect and over-punishment in the case of third party enforcement of social norms[J]. *Journal of Economic Behavior & Organization*, 175: 395-412.
- Krasnow M M, Delton A W, Cosmides L, et al. 2016. Looking under the hood of third-party punishment reveals design for personal benefit[J]. *Psychological Science*, 27(3): 405-418.
- McConnell C, Margalit Y, Malhotra N, et al. 2018. The economic consequences of partisanship in a polarized era[J]. *American Journal of Political Science*, 62(1): 5-18.
- Nakashima N A, Halali E, Halevy N. 2017. Third parties promote cooperative norms in repeated interactions[J]. *Journal of Experimental Social Psychology*, 68: 212-223.
- Tajfel H, Turner J C. 1986. The social identity theory of intergroup behavior[M]//Stephen W, Austin W. *Psychology of Intergroup Relations*. Chicago: Nelson Hall, 7-24.
- Waller I, Anderson A. 2021. Quantifying social organization and political polarization in online platforms[J]. *Nature*, 600(7888): 264-268.
- Yamagishi T, Kiyonari T. 2000. The group as the container of generalized reciprocity[J]. *Social Psychology Quarterly*, 63(2): 116-132.

Consensus Forges Cooperation: An Experimental Study Based on Peacemaker Game

Shuguang Jiang Zhengping Zhang

(The Center for Economic Research, Shandong University)

Abstract We run an experiment to investigate how ideology, exemplified by differing opinions on a

specific social topic, shapes group identity and influences behavior in conflict scenarios. Participants engaged in the Peacemaker Game, where disputants decide to cooperate or defect, and a third party chooses whether to intervene. The experimental design includes an opinion-divergence treatment, a minimal grouping treatment and a control condition. Results reveal that “opinion-based group identity” significantly influences in-group favoritism, out-group bias, and the likelihood of third-party intervention. The findings underscore the substantial impact of ideological alignment on social interactions and provide valuable implications for understanding and mitigating conflicts in real-world scenarios marked by increasing polarization.

JEL Classification C91, C92, D74, D91