

惩恶、扶弱与扬善：欺凌防治视角的 第三方惩罚、补偿与奖励*

王帆 陈光辉 陈培坤 智诺 常源 贤业敏 张文新

(山东师范大学心理学部; 山东省脑科学与心理健康重点实验室, 济南 250014)

摘要 中小学生会为了维护反欺凌群体规范而做出间接的第三方行为。本研究采用经济博弈范式与假想情境范式, 以 1011 名小学、初中与高中学生为对象, 考察欺凌现象中具有防治作用的第三方惩罚、补偿、奖励及其性别与学段差异。结果表明: 欺凌假想情境下, 中小学生均愿意对欺凌者实施第三方惩罚、对受欺凌者实施第三方补偿(二元情境中社会性第三方补偿除外)、对保护者实施第三方奖励; 相比于没有保护者的欺凌情境, 中小学生在有保护者的欺凌情境中会对欺凌者实施显著更强的经济性第三方惩罚。男生的经济性第三方惩罚与补偿水平显著高于女生; 高中生的社会性第三方惩罚水平显著高于初中生和小学生。研究发现证实了中小学生在欺凌假想情境中会表现出“惩恶、扶弱、扬善”的群体内互动方式, 为后续探索非冲突性反欺凌行为及干预方案开发提供了方向性依据。

关键词 中小学生, 欺凌, 第三方惩罚, 第三方补偿, 第三方奖励

分类号 B844

1 问题提出

欺凌是力量较强的一方蓄意伤害较弱一方并对其造成伤害的行为(陈光辉, 2014; 张文新, 2023; Olweus, 1993)。学生群体中欺凌行为的频发, 往往伴随反欺凌群体规范的缺失。形成和维护反欺凌的群体规范, 既需要群体成员敢于向欺凌说“不”, 也需要他们做出实际的反欺凌行动(Gaffney et al., 2021; Salmivalli & Voeten, 2004)。然而, 直接与欺凌者进行对抗的反欺凌行为通常会遭到欺凌者的报复或引发其他风险(Ma et al., 2019)。因此, 对于发生在中小学生群体中的欺凌现象, 间接的、非直接对抗性的反欺凌行为, 可能是群体成员有效维护反欺凌群体规范、降低欺凌发生率的可行路径。

1.1 群体规范维护与第三方惩罚

惩罚是一种通过施加厌恶刺激或撤销愉悦刺激以降低目标行为发生频率的方式, 可有效约束群

体成员的行为以维护群体规范(Molho et al., 2020)。根据惩罚实施者与违规事件的利益关联, 惩罚可分为第二方惩罚与第三方惩罚。第二方惩罚(又称直接惩罚)是指直接利益关联者对违规者实施的惩罚, 实施过程既能够维护群体规范, 又可同步提升自身收益(Molho et al., 2020; Turillo et al., 2002)。与此不同的是第三方惩罚, 它是指与违规事件无直接利益关联的个体, 通过事后或间接的方式(如取消帮助、减少资源分配)对违规者实施的惩罚行为。有时需要惩罚者主动付出成本, 其核心目的是维护群体规范而非获取个人即时收益(Balafoutas et al., 2014; Fehr & Fischbacher, 2004)。

第三方惩罚可进一步区分为社会性第三方惩罚与经济性第三方惩罚。社会性第三方惩罚侧重通过降低对违规者的帮助意愿、社会接纳等关系性联结来表达反对态度和实施非冲突性惩罚(Balafoutas et al., 2014), 其更加注重立场与意愿的间接表达,

收稿日期: 2026-02-08

* 2025 年度教育部哲学社会科学研究重大专项项目(2025JZDZ025)。

通信作者: 陈光辉, E-mail: chengh2100@sdnu.edu.cn

且隐蔽性高(Zhao et al., 2025)。例如, Balafoutas 等人(2014)在真实生活场景中的研究发现, 车站广场流动的路人更不愿意给乱扔垃圾者(违规者)提供帮助, 且当存在间接惩罚(不提供帮助)机会时, 路人选择直接惩罚(如当面指责)的比例显著下降。而经济性第三方惩罚则侧重通过减少金钱、游戏币等有形物品的给予来表达反对态度和非冲突性惩罚(Jordan et al., 2016)。由于物品分配具有明确的数量单位, 所以经济性第三方惩罚能够较直接地反映实施者为惩罚违规者所承担的成本及其惩罚强度。例如, 在公平分配情境(独裁者游戏)与合作情境(囚徒困境)中, 第三方参与者即便付出代价也要减少给违规者的代币数量以惩罚违规者, 且惩罚力度随着违规程度的增加而增强(Fehr & Fischbacher, 2004)。

进一步而言, 社会性第三方惩罚与经济性第三方惩罚分别体现了两种不同的非冲突性规范维护方式。社会性第三方惩罚更多表现为基于态度或意愿的关系性或情感性回应, 其实施门槛相对较低; 经济性第三方惩罚则需要实施者以出让一定的自身客观利益为代价, 通常在实施前需要进行必要的成本收益分析。二者均具备间接对抗的特征, 其相对较低的实施风险与广泛适用性使其在复杂社会互动(如欺凌事件)中更具潜在应用价值。鉴于已有研究尚未对两类惩罚进行比较研究, 本研究将考察中小学生在欺凌假想情境诱发的两类第三方惩罚上存在的可能差异。

1.2 欺凌防治视角的第三方惩罚、补偿与奖励

学生群体中的欺凌行为是一种违反群体规范的不道德行为, 且超过 80% 的 10 岁儿童已能辨识欺凌, 并将其判定为“不道德”或“错误”行为(陈光辉, 2014)。然而, 儿童青少年在面对不道德行为时, 需要谨慎考虑惩罚欺凌者可能引发的报复或其他潜在风险。因此, 探索非直接对抗的第三方行为对于学生欺凌防治可能更具可行性和潜在的积极意义。

从欺凌防治的视角而言, 除去对欺凌者进行间接惩罚(第三方惩罚), 群体成员对受欺凌者进行帮助(第三方补偿)和对反对欺凌者(或保护者)进行奖励(第三方奖励)可能也会有助于反欺凌规范的建立与维护。第三方补偿是指事后或通过间接的方式向弱者(如受欺凌者)提供支持、帮助或增加资源分配的行为, 有时为了达成此目的自身会主动付出一定的成本(Darley & Pittman, 2003; Dhaliwal et al., 2021)。第三方奖励则是指事后或通过间接的方式对主动维护群体规范的成员(如保护者)给予支持、

赞许或增加资源分配的行为, 有时为了达成此目的自身会主动付出一定的成本(de Kwaadsteniet et al., 2013; Geraci & Surian, 2021)。需要明确的是, 本研究聚焦的第三方惩罚、补偿与奖励, 特指欺凌事件结束后未直接参与事件的知情者, 在后续人际互动中对相关角色实施的非冲突性回应。这一事后性和非冲突性特征, 使其不同于事件发生过程中旁观者的即时干预, 也不同于与事件有直接利益关联者所采取的应对行为。

需要注意的是, 以往研究大多重点考察不公平事件中对违规者的第三方惩罚或对受害者的第三方补偿(马睿 等, 2023; Fehr & Fischbacher, 2004), 其情境互动结构表现为“违规者-受害者”的二元互动。与不公平情境不同, 有些欺凌事件是限于“欺凌者-受欺凌者”的二元互动, 有些欺凌事件中则会出现对抗欺凌者、求助成人或安慰受欺凌者的保护者(Chen et al., 2020; Salmivalli et al., 1996), 从而形成“欺凌者-受欺凌者-保护者”的三元互动。在“三元互动”情境中, 保护者作为主动维护反欺凌规范的一类群体成员, 为群体成员的第三方奖励提供了明确对象。群体成员对保护者的正向反馈, 不仅传达了对保护行为的认可, 也可能通过提升其同伴地位与被接纳感, 进一步增强其持续实施保护行为和维护反欺凌规范的动力(Laninga-Wijnen et al., 2023; van der Ploeg et al., 2017)。由此, 第三方惩罚、补偿与奖励共同构成了欺凌防治中的三种非冲突性同伴互动方式: 惩恶、扶弱与扬善。

鉴于对受欺凌者的第三方补偿和对保护者的第三方奖励均具有非冲突性特征, 本研究推测, 受欺凌者可能会在互动中因获得他人帮助或关爱而增强对抗欺凌者的力量, 保护者也可能因获得他人奖励性的支持与拥护而增强后续反欺凌的力量。更为重要的是, 这种隐藏于日常群体互动的事后行为较不易被欺凌者直接察觉, 从而减少了对欺凌者的激惹和可能引发的报复行为。与第三方惩罚类似, 第三方补偿和奖励也应有社会性和经济性之分。基于此, 本研究假设, 学生欺凌情境能够诱发中小学生对欺凌者实施第三方惩罚、对受欺凌者实施第三方补偿、对保护者实施第三方奖励, 且这三种行为在社会性和经济性维度上可能存在差异。

1.3 “二元互动”与“三元互动”欺凌情境下的第三方惩罚与补偿

学生欺凌事件中, 欺凌者与受欺凌者是核心互动主体, 而 80% 以上的欺凌事件会有旁观者在场

(Salmivalli et al., 1996)。旁观者的不同参与行为(如袖手旁观或出手相助)会重塑事件的互动结构,且相当比例的旁观者(19% ~ 43%)愿意主动制止欺凌或陪伴受欺凌者(Hawkins et al., 2001; Ma et al., 2019)。因此,从同伴干预的欺凌防治视角而言,欺凌情境可分为两类:一类是仅包含“欺凌者-受欺凌者”的二元互动情境(无保护者介入);另一类是包含“欺凌者-受欺凌者-保护者”的三元互动情境(有保护者介入)。不同情境可能对知情者的第三方惩罚、补偿、奖励产生重要影响。

在“二元互动”情境中,知情的第三方需独立完成对事件性质的判断,既包括确认行为事件是否属于“强者故意伤害弱者”的欺凌(Olweus, 1993),也包括对受欺凌者是否存在过错、惩罚欺凌者是否存在风险的权衡(Ng et al., 2020)。这种信息不确定性与风险顾虑,可能削弱第三方实施惩罚或补偿的意愿(Toribio - Flórez et al., 2023)。同时,由于情境中缺乏主动维护规范的保护者角色,仅可能诱发第三方惩罚与补偿两类行为。

在“三元互动”情境中,保护者的主动干预(如制止欺凌、陪伴受欺凌者、告知教师)会影响第三方的决策过程。一是保护者的出现为群体成员实施第三方奖励提供了奖励对象;二是保护者的保护弱者行为传递了反对欺凌的清晰信号(van der Ploeg et al., 2017),为第三方的反欺凌行为决策提供了指引;三是保护者作为可直接模仿的行为榜样,为第三方提供了安全的同伴干预参照,且无需独自承担反对欺凌的潜在风险(Ng et al., 2020),从而可能提升个体对欺凌者实施第三方惩罚、对受欺凌者实施第三方补偿的意愿。鉴于此,本研究假设,相较于“二元互动”情境,存在保护者的“三元互动”情境能够诱发中小學生更高水平的第三方惩罚和第三方补偿。

1.4 第三方惩罚、补偿与奖励的性别与年龄段差异

学生欺凌情境诱发的第三方惩罚、补偿与奖励可能存在性别差异。根据社会角色理论,女性更注重人际联结与关系维护,会偏好采取非对抗性的社会互动行为(Eagly & Wood, 1999; Xiao et al., 2019)。然而在欺凌事件中,一方面,男生更多采取以强凌弱的行为方式解决人际冲突或达成目标(Feijóo et al., 2021);另一方面,女生扮演保护者的数量显著高于男生,且保护行为水平也显著高于男生(Chen et al., 2020; Salmivalli et al., 1996)。这提示女生在欺凌事件后的第三方惩罚、补偿、奖励等维

护群体规范的行为上可能显著高于男生。

学生欺凌情境诱发的第三方惩罚、补偿与奖励可能存在年龄段差异。作为欺凌事件后的非冲突性互动,第三方行为的发生与强度会受到个体发展阶段的影响。随着年龄增长,个体的认知控制系统日益完善(Andrews et al., 2021),对干预行为可能引发的报复或社交风险的预判和权衡能力逐步提升(Romer et al., 2017),从而影响其采取干预行为的意愿和强度。同时,进入高年龄段(如初中至高中),个体的道德认知水平显著提升,对欺凌这类违背群体规范的行为容忍度降低,维护群体规范的动机随之增强(Decrop et al., 2025)。因此,高年龄段的个体在欺凌事件后的第三方惩罚、补偿、奖励等维护群体规范的行为可能显著高于低年龄段个体。

1.5 研究概览

本研究立足非冲突性人际互动视角,系统考察中小學生由欺凌假想情境引发的“惩恶、扶弱、扬善”的反欺凌群体互动行为。本研究拟针对无保护者介入的“二元互动”欺凌假想情境与有保护者介入的“三元互动”欺凌假想情境,考察第三方惩罚、第三方补偿与第三方奖励,并进一步比较保护者是否在场对第三方惩罚与补偿的影响。具体而言,研究1在无保护者介入的“欺凌者-受欺凌者”二元互动情境中考察中小學生对欺凌者的第三方惩罚和对受欺凌者的第三方补偿;研究2在有保护者介入的“欺凌者-受欺凌者-保护者”三元互动情境中进一步考察中小學生的第三方惩罚、第三方补偿以及对保护者的第三方奖励;研究3则基于研究1和研究2的数据,进一步分析有无保护者的欺凌情境对中小學生实施第三方惩罚与补偿的差异影响。三个研究均同时从社会性和经济性两个维度考察第三方行为,并分析其性别差异与年龄段差异。

2 研究1:“二元互动”欺凌假想情境下的第三方惩罚与补偿

研究1基于“二元互动”模式构建“欺凌情境”与“日常情境”两种互动情境,旨在考察中小學生对欺凌者实施的第三方惩罚和对受欺凌者实施的第三方补偿,及其性别与年龄段差异。

2.1 被试

研究采用2(性别:男 vs. 女)×3(年龄段:小学 vs. 初中 vs. 高中)的两因素组间实验设计,考察中小學生的第三方惩罚与第三方补偿及其性别与年龄段差异。参考社会心理学的平均效应量 $f =$

0.2148 (Richard et al., 2003), 使用 G*Power 3.1 估计满足本研究所需样本量(Faul et al., 2009), 达到 0.95 的统计检验力($\alpha = 0.05$)需要 338 人。本研究实际招募有效被试 508 人($M_{\text{age}} = 13.37$ 岁, $SD = 2.54$ 岁; 其中男生 255 人), 包括小学生 169 人($M_{\text{age}} = 10.41$ 岁, $SD = 0.75$ 岁; 其中男生 87 人)、初中生 167 人($M_{\text{age}} = 13.39$ 岁, $SD = 0.66$ 岁; 其中男生 83 人)、高中生 172 人($M_{\text{age}} = 16.21$ 岁, $SD = 1.25$ 岁; 其中男生 85 人)。研究征得了学生监护人、学校管理者知情同意以及学生的自愿参与, 学生在参与过程中随时可以选择退出; 研究由山东师范大学心理学院伦理委员会审查通过(伦理审查编号: SDNU2023083, 以下同)。

2.2 实验材料

2.2.1 “二元互动”情境

参考 Fox 等人(2014)编制的欺凌情境设计了两类“二元互动”情境: 欺凌情境、日常情境, 具体描述见表 1。欺凌情境描述了 A (欺凌者)故意绊倒 B (受欺凌者)并加以嘲笑的典型欺凌场景, 通过“故意伸脚绊倒并加以嘲笑”体现欺凌的故意伤害性特征; 通过“周围同学起哄附和、受欺凌者处于被动弱势位置”体现欺凌的力量不均衡性特征; 通过“这已经不是 A 第一次这样对待 B, 之前就有过类似行为”体现欺凌的重复发生性特征。日常情境则描述了甲、乙两位普通同学关于水杯的日常交谈场景。日常情境在本研究中主要作为欺凌情境的基线对照条件。

为检验实验材料在区分情境类别(欺凌 vs. 日常)与参与者角色(欺凌者 vs. 受欺凌者 vs. 普通同学)上的有效性, 选取 38 名学生($M_{\text{age}} = 13.09$ 岁, $SD = 0.44$ 岁; 其中男生 21 人)对情境类别与参与者角色进行判断。McNemar 检验表明, 被试能准确识别欺凌情境, 即被试判断欺凌情境中存在欺凌行为的比例(97.37%)显著高于日常情境(2.63%), $p < 0.001$ 。进一步地, Mann-Whitney U 检验表明, 被试将欺凌情境中 A 判断为欺凌者的比例(94.74%)显著

高于将日常情境中甲判断为欺凌者的比例(2.63%), $U = 57.00$, $Z = -7.98$, $p < 0.001$; 被试将欺凌情境中 B 判断为受欺凌者的比例(94.74%)显著高于将日常情境中乙判断为受欺凌者的比例(2.63%), $U = 57.00$, $Z = -7.78$, $p < 0.001$ 。以上结果表明, “二元互动”情境材料能够有效区分欺凌情境与日常情境, 且能够有效区分欺凌者、受欺凌者与普通同学。

2.2.2 第三方惩罚与第三方补偿

基于帮助意愿的测量。借鉴已有研究(Artavia-Mora et al., 2017; Balafoutas et al., 2014), 改编形成两个亲社会情境材料用于测量社会性第三方惩罚与第三方补偿。具体描述如下。

1. 你正准备要去上体育课, 【角色】(如甲、乙、A、B)过来找你借字典, 你的字典就在书包里。这时候你愿意把字典借给他吗?

2. 在大课间的时候, 【角色】(如甲、乙、A、B)去操场打球了, 他的水杯不小心被路过的同学碰倒。你注意到杯子里的水马上就要流到他的课本上了, 这时候你会愿意帮他把水杯扶起来吗?

被试需要针对上文“二元互动”情境中的角色(如甲、乙、A、B)来回答对目标对象的帮助意愿。采用 5 点计分, 1 表示“非常不愿意”, 5 表示“非常愿意”, 总得分越高, 则帮助意愿水平越高。若被试对欺凌者 A 的帮助意愿显著低于对普通同学甲的帮助意愿, 则表明存在社会性第三方惩罚。二者差值(帮助甲的意愿减去帮助 A 的意愿)越大, 表明社会性第三方惩罚水平越高。若被试对受欺凌者 B 的帮助意愿显著高于对普通同学乙的帮助意愿, 则表明存在社会性第三方补偿。二者差值(帮助 B 的意愿减去帮助乙的意愿)越大, 表明社会性第三方补偿水平越高。

基于游戏币分配的测量。采用独裁者博弈范式(Turillo et al., 2002)测量经济性第三方惩罚与第三方补偿。被试假想与情境中的角色(如甲、乙、A、B)进行游戏币分配游戏, 并作为“独裁者”拥有对 10 枚游戏币的完全分配权。若被试分配给欺凌者 A 的

表 1 “二元互动”情境

情境类型	内容
欺凌情境	在课间, B 出去接水回来往自己的座位上走。A 故意伸脚绊倒了 B, B 杯子里的水洒了一地。A 嘲笑地说: “你真够笨的, 走个路都能摔倒! 哈哈!” 周围有几个同学也跟着起哄, 哈哈大笑。 这已经不是 A 第一次这样对待 B 了, 之前就有过类似的行为。B 忍着疼痛扶着凳子站了起来, 拿拖把拖干地后出去重新接水。
日常情境	在课间, 乙出去接水回来往自己的座位上走。甲叫住他说: “这是你新换的水杯吗? 真好看!” 两人开始聊起关于新水杯的事情。

游戏币数量低于分配给普通同学甲的游戏币数量, 则表明存在经济性第三方惩罚。二者差值(分配给甲的游戏币数量减去分配给A的游戏币数量)越大, 表明经济性第三方惩罚水平越高。若被试分配给受欺凌者B的游戏币数量高于分配给普通同学乙的游戏币数量, 则表明存在经济性第三方补偿。二者差值(分配给B的游戏币数量减去分配给乙的游戏币数量)越大, 表明经济性第三方补偿水平越高。

2.3 研究结果

2.3.1 实验条件操纵性检验

采用配对样本 t 检验, 比较被试对欺凌情境中欺凌者A与日常情境中普通同学甲的行为合适程度的评估差异, 以检验被试能否区分二者行为所代表的欺凌情境与日常情境。结果显示, 被试认为欺凌者A行为的合适程度($M = 1.06$, $SD = 0.24$)显著低于普通同学甲行为的合适程度($M = 4.04$, $SD = 0.82$), $t(507) = -74.76$, $p < 0.001$, Cohen's $d_z = 3.32$, 表明情境类别的实验操纵有效。

2.3.2 欺凌情境与第三方惩罚、第三方补偿

采用配对样本 t 检验对被试的帮助意愿和分配游戏币数量进行差异分析, 结果如图1所示。在帮助意愿上, 被试对欺凌者A的帮助意愿($M = 6.17$, $SD = 2.84$)显著低于对普通同学甲的帮助意愿($M = 8.66$, $SD = 1.55$), $t(507) = -20.50$, $p < 0.001$, Cohen's $d_z = 0.91$, 由此表明存在社会性第三方惩罚; 然而, 被试对受欺凌者B的帮助意愿($M = 8.78$, $SD = 1.70$)与对普通同学乙的帮助意愿($M = 8.69$, $SD = 1.55$)不存在显著差异, $t(507) = 1.64$, $p = 0.102$, 由此表明不存在社会性第三方补偿。

在分配游戏币数量上, 被试给欺凌者A分配的游戏币数量($M = 3.76$, $SD = 2.01$)显著低于给普通

同学甲分配的游戏币数量($M = 4.89$, $SD = 0.99$), $t(507) = -12.77$, $p < 0.001$, Cohen's $d_z = 0.57$, 由此表明存在经济性第三方惩罚; 被试给受欺凌者B分配的游戏币数量($M = 5.29$, $SD = 1.28$)显著高于给普通同学乙分配的游戏币数量($M = 4.89$, $SD = 0.98$), $t(507) = 7.02$, $p < 0.001$, Cohen's $d_z = 0.31$, 由此表明存在经济性第三方补偿。

分别以社会性、经济性第三方惩罚与第三方补偿为因变量, 性别(男、女)和学龄段(小学、初中、高中)为自变量, 进行 2×3 的两因素方差分析。结果显示, 在社会性第三方惩罚上, 性别的主效应不显著, $F(1, 502) = 2.25$, $p = 0.134$; 学龄段的主效应显著, $F(2, 502) = 6.91$, $p = 0.001$, $\eta_p^2 = 0.03$, 表现为高中生的社会性第三方惩罚($M = 3.08$, $SD = 2.65$)显著高于初中生($M = 2.40$, $SD = 2.67$)与小学生($M = 2.00$, $SD = 2.81$); 性别和学龄段的交互作用不显著, $F(2, 502) = 0.51$, $p = 0.601$ 。在经济性第三方惩罚上, 性别的主效应显著, $F(1, 502) = 5.19$, $p = 0.023$, $\eta_p^2 = 0.01$, 表现为男生的经济性第三方惩罚($M = 1.33$, $SD = 2.17$)显著高于女生($M = 0.92$, $SD = 1.75$); 学龄段的主效应、性别和学龄段的交互作用均不显著, $ps > 0.238$ 。在社会性与经济性第三方补偿上, 性别和学龄段的主效应及交互作用均不显著, $ps > 0.125$ 。

2.4 讨论

本研究结果显示, 欺凌假想情境能够显著诱发中小学生的第三方惩罚(包括社会性与经济性)以及第三方补偿(仅限于经济性), 表明中小學生普遍持有明确的反欺凌态度, 并会做出非对抗性的惩恶与扶弱。具体而言, 第三方会通过事后非冲突性方式参与反欺凌, 即对欺凌者实施社会疏离(社会性第

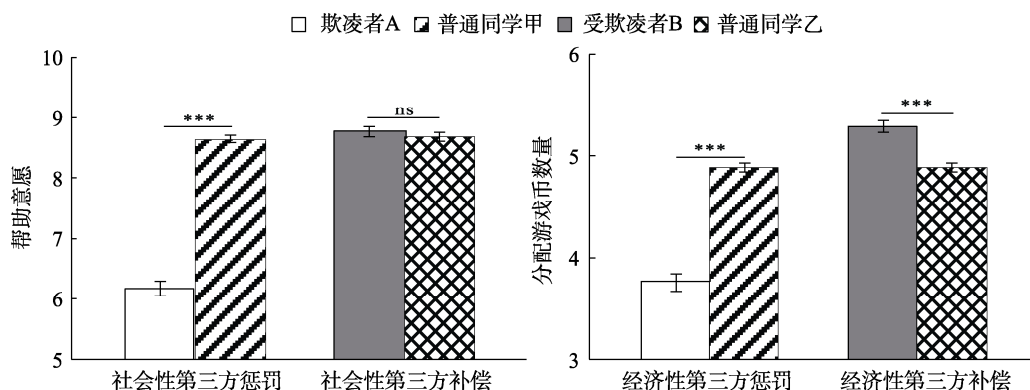


图1 中小学生的社会性与经济性第三方惩罚和补偿

注: *** $p < 0.001$

三方惩罚)或资源剥夺(经济性第三方惩罚),并对受欺凌者进行资源倾斜(经济性第三方补偿)。第三方惩罚(社会性与经济性)的一致性结果表明,惩戒欺凌者的反欺凌规范在学生群体中具有较高的内在一致性(Thornberg et al., 2022),即中小学生无论是以社会疏离的方式,还是以牺牲个人利益的方式,都愿意让欺凌者付出代价。而第三方补偿仅发现了经济性维度上的显著差异,这可能是由于经济性第三方补偿(如资源倾斜、代币分配)能修复受欺凌者的物质损失与第三方的不公平感,且能为干预者带来潜在的声誉收益(Dhaliwal et al., 2021)。相比之下,社会性第三方补偿(如简单帮助、情感支持)可能难以有效缓解欺凌给受欺凌者带来的社交孤立、情感自卑等深层心理创伤(Al - Alawi et al., 2025),第三方可能主观上判断这类干预的效果有限,从而未形成稳定的补偿倾向。

此外,本研究发现男生的经济性第三方惩罚显著高于女生。这与 Yang 等人(2023)的研究结果一致,男生在第三方惩罚中更倾向于对违规者(如欺凌者)实施更严厉的惩罚。这可能与社会角色期待有关,男性常与帮助弱者、维护秩序等工具性角色相联系(Alsawalqa et al., 2021; Gutierrez et al., 2020),使男生更倾向于采用直接、量化的惩罚手段来应对不公平情境。同时,本研究发现高中生的社会性第三方惩罚显著高于初中生与小学生,这与高中阶段前额叶皮层功能成熟、规范认知能力提升的生理发展特征一致(Baum et al., 2020),也支持了第三方惩罚随年龄增长而增强的研究假设。

需要注意的是,有些欺凌事件发生时会有保护者出现,保护者的介入会改变欺凌事件中亲欺凌者与反欺凌者的力量对比关系,并可能会对知情者后续的第三方惩罚、第三方补偿带来影响。欺凌情境中是否会出现对保护者的第三方奖励也尚未可知。因此,研究 2 将通过构建“欺凌者-受欺凌者-保护者”的三元互动假想情境进一步探究第三方行为。

3 研究 2:“三元互动”欺凌假想情境下的第三方惩罚、补偿与奖励

在研究 1 的两类“二元互动”情境中分别加入保护者与新普通同学,形成“三元互动”欺凌情境与日常情境,旨在探究个体对欺凌者的第三方惩罚、对受欺凌者的第三方补偿、对保护者的第三方奖励及其性别与年龄段差异特征。

3.1 被试与实验设计

研究采用 2 (性别:男 vs. 女) × 3 (年龄段:小学 vs. 初中 vs. 高中)的两因素组间实验设计,考察个体的第三方惩罚、第三方补偿与第三方奖励及其性别与年龄段差异。参考社会心理学的平均效应量 $f = 0.2148$ (Richard et al., 2003),使用 G*Power 3.1 估计满足本研究所需样本量(Faul et al., 2009),达到 0.95 的统计检验力($\alpha = 0.05$)需要 338 人。本研究实际招募有效被试 503 人($M_{\text{age}} = 13.28$ 岁, $SD = 2.44$ 岁;其中男生 248 人),包括小学生 173 人($M_{\text{age}} = 10.45$ 岁, $SD = 0.82$ 岁;其中男生 85 人)、初中生 170 人($M_{\text{age}} = 13.31$ 岁, $SD = 0.61$ 岁;其中男生 86 人)、高中生 160 人($M_{\text{age}} = 16.21$ 岁, $SD = 0.62$ 岁;其中男生 77 人)。

3.2 实验材料

3.2.1 “三元互动”情境

在“二元互动”的欺凌情境与日常情境中分别增加保护者 C 与普通同学丙,构成“三元互动”情境(见表 2)。

为检验实验材料在区分情境类别(欺凌 vs. 日常)与参与者角色(欺凌者 vs. 受欺凌者 vs. 保护者 vs. 普通同学)上的有效性,选取 40 名学生($M_{\text{age}} = 13.00$ 岁, $SD = 0.28$ 岁;其中男生 20 人)对情境类别与参与者角色进行判断。McNemar 检验表明,被试能准确识别欺凌情境,即被试判断欺凌情境中存在欺凌行为比例(95.00%)显著高于日常情境(7.50%), $p < 0.001$ 。进一步地, Mann-Whitney U 检验结果显示,被试将 A 判断为欺凌者的比例(92.50%)显著高于将甲判断为欺凌者的比例(7.50%), $U = 120.00$, $Z = -7.56$, $p < 0.001$;被试将 B 判断为受欺凌者的比例(92.50%)显著高于将乙判断为受欺凌者的比例(10.00%), $U = 140.00$, $Z = -7.34$, $p < 0.001$;被试将 C 判断为保护者的比例(92.50%)显著高于将丙判断为保护者的比例(7.50%), $U = 120.00$, $Z = -7.56$, $p < 0.001$ 。以上结果表明,“三元互动”情境材料能够有效区分欺凌情境与日常情境,且能够有效区分欺凌者、受欺凌者、保护者与普通同学。

3.2.2 第三方惩罚、补偿与奖励

基于帮助意愿的测量。社会性第三方惩罚与社会性第三方补偿的测量方式同研究 1。在此基础上,如果被试对保护者 C 的帮助意愿显著高于对普通同学丙的帮助意愿,则表明存在社会性第三方奖励。二者差值(帮助 C 的意愿减去帮助丙的意愿)越大,表明社会性第三方奖励水平越高。

表 2 “三元互动”情境

情境类型	内容
欺凌情境	在课间, B 出去接水回来往自己的座位上走。A 故意伸脚绊倒了 B, B 杯子里的水洒了一地。A 嘲笑地说: “你真够笨的, 走个路都能摔倒! 哈哈!” 周围有几个同学也跟着起哄, 哈哈大笑。 这已经不是 A 第一次这样对待 B 了, 之前就有过类似的行为。这时, 一旁的 C 看不下去了, 过去将 B 扶起来并和他一起去重新接水。放学后 C 将此事告诉了班主任老师。
日常情境	在课间, 乙出去接水回来往自己的座位上走。甲叫住他说: “这是你新换的水杯吗? 真好看!” 两人开始聊起关于新水杯的事情。一旁的丙也加入了聊天。

基于游戏币分配的测量。经济性第三方惩罚与经济性第三方补偿的测量方式同研究 1。在此基础上, 若被试分配给保护者 C 的游戏币数量高于分配给普通同学丙的游戏币数量, 则表明存在经济性第三方奖励。二者差值(分配给 C 的游戏币数量减去分配给丙的游戏币数量)越大, 表明经济性第三方奖励水平越高。

3.3 研究结果

3.3.1 实验条件操纵性检验

采用配对样本 t 检验, 比较被试对欺凌情境中欺凌者 A 与日常情境中普通同学甲的行为合适程度的评估差异, 以检验被试能否区分二者行为所代表的欺凌情境与日常情境。结果显示, 被试认为欺凌者 A 所做行为的合适程度($M = 1.06, SD = 0.24$)显著低于普通同学甲所做行为的合适程度($M = 3.98, SD = 0.88$), $t(502) = -69.78, p < 0.001$, Cohen's $d_z = 3.11$, 表明情境类别的实验操纵有效。

3.3.2 欺凌情境与第三方惩罚、补偿、奖励

采用配对样本 t 检验对被试的帮助意愿和分配游戏币数量进行差异分析, 结果如图 2 所示。在帮助意愿上, 被试对欺凌者 A 的帮助意愿($M = 5.91, SD = 2.93$)显著低于对普通同学甲的帮助意愿($M = 8.61, SD = 1.61$), $t(502) = -21.19, p < 0.001$, Cohen's

$d_z = 0.94$, 由此表明存在社会性第三方惩罚; 被试对受欺凌者 B 的帮助意愿($M = 8.85, SD = 1.65$)显著高于对普通同学乙的帮助意愿($M = 8.67, SD = 1.63$), $t(502) = 3.12, p = 0.002$, Cohen's $d_z = 0.14$, 由此表明存在社会性第三方补偿; 被试对保护者 C 的帮助意愿($M = 9.00, SD = 1.68$)显著高于对普通同学丙的帮助意愿($M = 8.57, SD = 1.75$), $t(502) = 6.15, p < 0.001$, Cohen's $d_z = 0.27$, 由此表明存在社会性第三方奖励。在分配游戏币数量上, 被试给欺凌者 A 分配的游戏币数量($M = 3.39, SD = 2.17$)显著低于给普通同学甲分配的游戏币数量($M = 4.83, SD = 1.01$), $t(502) = -15.02, p < 0.001$, Cohen's $d_z = 0.67$, 由此表明存在经济性第三方惩罚; 被试给受欺凌者 B 分配的游戏币数量($M = 5.19, SD = 1.16$)显著高于给普通同学乙分配的游戏币数量($M = 4.86, SD = 0.96$), $t(502) = 6.20, p < 0.001$, Cohen's $d_z = 0.28$, 由此表明存在经济性第三方补偿; 被试给保护者 C 分配的游戏币数量($M = 5.22, SD = 1.14$)显著高于给普通同学丙分配的游戏币数量($M = 4.86, SD = 0.97$), $t(502) = 7.35, p < 0.001$, Cohen's $d_z = 0.33$, 由此表明存在经济性第三方奖励。

分别以社会性、经济性第三方惩罚、第三方补偿与第三方奖励为因变量, 性别(男、女)和年龄段

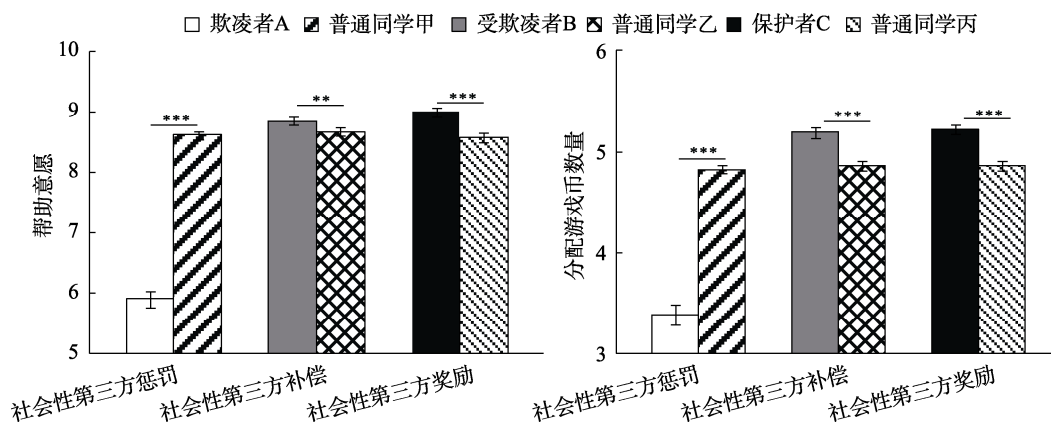


图 2 中小学生的第三方惩罚、补偿与奖励

注: ** $p < 0.01$, *** $p < 0.001$

(小学、初中、高中)为自变量,进行 2×3 的两因素方差分析。结果显示,在社会性第三方惩罚上,性别的主效应不显著, $F(1, 497) = 0.47, p = 0.494$; 学龄段的主效应显著, $F(2, 497) = 10.03, p < 0.001, \eta_p^2 = 0.04$, 表现为高中生的第三方惩罚($M = 3.52, SD = 2.74$)显著高于初中生($M = 2.31, SD = 2.67$)与小学生($M = 2.34, SD = 3.00$)。性别与学龄段的交互效应不显著, $F(2, 497) = 2.54, p = 0.080$ 。在经济性第三方惩罚上,性别和学龄段的主效应及交互作用均不显著, $ps > 0.132$ 。

在社会性第三方补偿上,性别和学龄段的主效应及交互作用均不显著, $ps > 0.399$ 。在经济性第三方补偿上,性别的主效应显著, $F(1, 497) = 8.34, p = 0.004, \eta_p^2 = 0.02$, 表现为男生的第三方补偿($M = 0.50, SD = 1.39$)高于女生($M = 0.18, SD = 1.01$); 学龄段的主效应以及性别与学龄段的交互效应均不显著, $ps > 0.843$ 。

在社会性第三方奖励上,性别与学龄段的主效应均不显著, $ps > 0.065$; 性别和学龄段的交互作用显著, $F(2, 497) = 5.51, p = 0.004, \eta_p^2 = 0.02$ 。简单效应分析发现,初中女生的社会性第三方奖励($M = 0.98, SD = 1.74$)显著高于初中男生($M = 0.34, SD = 1.26$), $F(1, 497) = 7.31, p = 0.007, \eta_p^2 = 0.02$; 高中男生的社会性第三方奖励($M = 0.57, SD = 1.84$)显著高于高中女生($M = 0.08, SD = 1.54$), $F(1, 497) = 4.00, p = 0.046, \eta_p^2 = 0.01$ 。在经济性第三方奖励上,性别与学龄段的主效应均不显著, $ps > 0.191$; 性别和学龄段的交互作用显著, $F(2, 497) = 3.27, p = 0.039, \eta_p^2 = 0.01$ 。简单效应分析发现,高中男生的经济性第三方奖励($M = 0.62, SD = 1.30$)高于高中女生($M = 0.18, SD = 1.07$), $F(1, 497) = 6.55, p = 0.011, \eta_p^2 = 0.01$ 。

3.4 讨论

本研究首次发现,在欺凌假想情境中,保护者的反欺凌行为会引发其他人的社会性支持(提升帮助意愿)和经济性资源倾斜(分配更多游戏币),证实了“三元互动”情境下群体成员会对保护者进行第三方奖励。保护者主动制止欺凌、支持受欺凌者的行为,本质上是维护群体反欺凌规范时做出的亲社会行为,而同伴的社会认可与资源倾斜,正是对这种亲社会行为的间接回报。进一步而言,这一发现连同此前证实的第三方惩罚与补偿,共同构成“惩恶、扶弱、扬善”的全景性反欺凌群体互动方式,呈现出“欺凌者失道寡助、受欺凌者因弱受助、保护

者得道多助”的群体互动行为倾向。

研究发现“三元互动”情境下第三方行为的性别差异并非单一不变。与“二元互动”中男生的经济性第三方惩罚显著高于女生不同,在“三元互动”情境中,男生的经济性第三方补偿显著高于女生。这一差异可能源于保护者的即时援助缓解了受欺凌者所受到的伤害(Ma & Chen, 2019),降低了即时对抗不公的行为需求,使男生的干预重心转向弥补受欺凌者伤害,进而表现为经济性第三方补偿的显著提升。由此可见,保护者的在场可能改变第三方行为的性别差异表现,使其呈现出一定的情境适应性。

本研究还发现,第三方奖励的性别差异具有明显的发展性特征。小学阶段男女生在第三方奖励上的差异并不显著,这可能表明在较早发展阶段,学生对“保护弱者”这一行为总体持有较为一致的积极评价。进入初中阶段后,随着青春期性别角色社会化压力增强,男女生在社会行为表达上的分化逐渐增加(Priess et al., 2009)。研究表明,女生在青春期的共情关怀、关系维系与亲社会行为上总体更为突出(Michalska et al., 2013; Van der Graaff et al., 2018),且在欺凌情境中更常担任保护者、其保护行为水平也更高(Chen et al., 2020)。因此,初中女生更高的社会性第三方奖励,可能更多体现了其基于关怀、支持与关系维系的反应倾向。

随着学龄段提升至高中,第三方奖励的性别差异出现反转,高中男生的社会性与经济性第三方奖励水平均显著高于女生。这一结果可能与男生的亲社会行为在青春期中后期出现更明显的上升趋势有关(Van der Graaff et al., 2018)。同时,第三方奖励不仅体现一般性的亲社会支持,也包含对规范维护者的积极评价与正向反馈。研究指出,男性对高显著性激励线索的行为、生理和神经反应可能强于女性(Warthen et al., 2020),因此男生可能更容易对欺凌情境中出现的规范性线索做出积极回应,从而表现出更高水平的第三方奖励。需要说明的是,本文并未直接测量性别角色认同、规范维护动机或奖赏敏感性,相关解释仍需后续研究检验。

综上,本研究发现中小学生在欺凌假想情境中会表现出与“惩恶、扶弱、扬善”相一致的反欺凌群体互动行为,证实了同伴群体中可能存在对亲社会干预者的奖励行为;同时还发现保护者在场的欺凌假想情境下,第三方惩罚与补偿可能会受到影响。为此,研究 3 将直接对比有无保护者对第三方惩罚与补偿的影响,并分析可能存在性别与学龄段差异。

4 研究 3: 欺凌假想情境中有无保护者对第三方惩罚、补偿的影响

4.1 研究方法

在综合研究 1 和研究 2 所收集数据的基础上, 研究 3 将“保护者是否在场”(或“有无保护者”)纳入研究范畴, 考察有无保护者及其与性别、学龄段的交互作用对第三方惩罚、补偿的影响。研究 3 实际纳入有效被试 1011 人($M_{\text{age}} = 13.32$ 岁, $SD = 2.49$ 岁; 其中男生 503 人), 包括小学生 342 人($M_{\text{age}} = 10.43$ 岁, $SD = 0.78$ 岁; 其中男生 172 人)、初中生 337 人($M_{\text{age}} = 13.35$ 岁, $SD = 0.63$ 岁; 其中男生 169 人)、高中生 332 人($M_{\text{age}} = 16.21$ 岁, $SD = 1.00$ 岁; 其中男生 162 人)。研究采用 2 (有无保护者: 无 vs. 有) $\times$ 2 (性别: 男 vs. 女) $\times$ 3 (学龄段: 小学 vs. 初中 vs. 高中) 的三因素组间设计。

研究 3 所分析的社会性与经济性第三方惩罚、补偿, 均沿用研究 1 和研究 2 的实验材料与操作测量方式。其中, 社会性第三方惩罚与补偿分别通过帮助意愿差值表示, 经济性第三方惩罚与补偿分别通过游戏币分配差值表示, 差值越大表明相应第三方行为水平越高。

4.2 研究结果

4.2.1 有无保护者对第三方惩罚的影响及其性别与学龄段差异

分别以社会性、经济性第三方惩罚为因变量, 有无保护者、性别和学龄段为自变量, 进行 $2 \times 2 \times 3$ 的三因素方差分析, 结果如图 3 所示。在社会性第三方惩罚上, 学龄段的主效应显著, $F(2, 999) = 16.30, p < 0.001, \eta_p^2 = 0.03$, 表现为高中生的第三方惩罚($M = 3.29, SD = 2.70$)显著高于初中生($M =$

$2.35, SD = 2.67$)与小学生($M = 2.17, SD = 2.91$); 其余主效应与交互效应均不显著, $ps > 0.098$ 。在经济性第三方惩罚上, 有无保护者的主效应显著, $F(1, 999) = 6.04, p = 0.014, \eta_p^2 = 0.01$, 有保护者时个体的第三方惩罚($M = 1.44, SD = 2.15$)显著高于无保护者($M = 1.13, SD = 1.95$)。其余主效应与交互效应均不显著, $ps > 0.107$ 。

4.2.2 有无保护者对第三方补偿的影响及其性别与学龄段差异

分别以社会性、经济性第三方补偿为因变量, 有无保护者、学龄段和性别为自变量, 进行 $2 \times 3 \times 2$ 的三因素方差分析。结果显示, 在社会性第三方补偿上, 有无保护者、性别与学龄段的主效应与交互效应均不显著, $ps > 0.119$ 。在经济性第三方补偿上, 性别的主效应显著, $F(1, 999) = 8.62, p = 0.003, \eta_p^2 = 0.01$, 男生的经济性第三方补偿($M = 0.48, SD = 1.41$)高于女生($M = 0.25, SD = 1.06$); 其余主效应与交互效应均不显著, $ps > 0.291$ 。

4.3 讨论

分析比较发现, 有保护者的欺凌假想情境仅在经济性第三方惩罚上显著高于无保护者的欺凌情境。相比之下, 社会性第三方惩罚、社会性第三方补偿与经济性第三方补偿均未因保护者是否存在而表现出显著差异。这表明, 对于中小學生而言, 社会性第三方惩罚(如降低对欺凌者的帮助意愿)、社会性第三方补偿(如增加对受欺凌者的日常帮助)以及经济性第三方补偿(如向受欺凌者分配更多资源)的反应强度在有无保护者的情境下表现相对稳定。一种可能的解释是, 对弱者的同情与支持是一种较为稳定的亲社会倾向(Grueneisen & Warneken, 2022), 且受欺凌者所遭受的伤害具有持续性(Chen

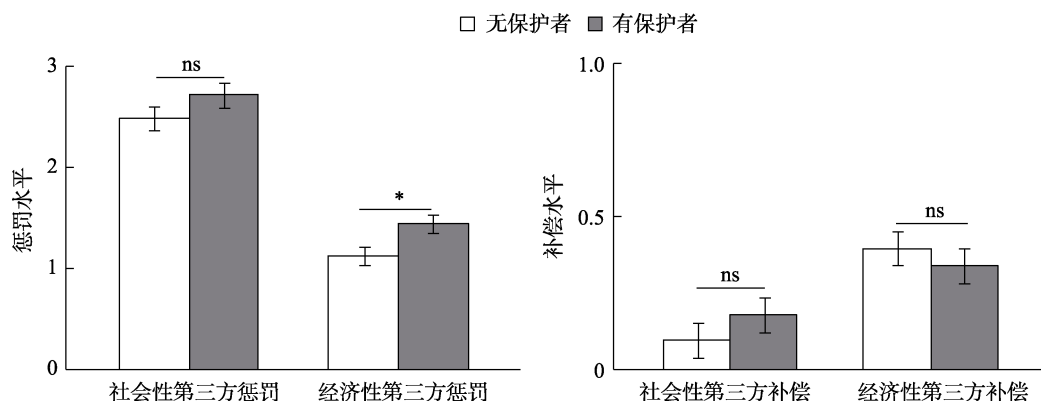


图 3 有无保护者下的社会性与经济性第三方惩罚、补偿

注: * $p < 0.05$

et al., 2018), 从而使第三方实施补偿的行为独立于保护者是否存在的情境性影响。对于惩罚而言, 社会性第三方惩罚更多体现为基于态度或意愿的关系性回应, 其诱发可能主要来自个体对欺凌行为本身的否定, 因此在有无保护者的情境下差异不明显。相比之下, 经济性第三方惩罚需要实施者付出一定的自身资源代价, 因而更依赖保护者在场所传递的规范支持信号。保护者的出现不仅有助于他人进一步确认欺凌行为的违规性质, 也会强化群体中的反欺凌规范(Camacho et al., 2025)。因此, 保护者在场会促使个体提升实施经济性第三方惩罚的水平。

5 总讨论

学生欺凌的有效防治离不开同伴群体对反欺凌规范的自发维护(Gaffney et al., 2021)。本研究将非冲突性第三方行为的研究范式应用于学生欺凌假想情境, 首次在中小學生群体中发现了与“惩恶、扶弱、扬善”相一致的反欺凌群体互动行为倾向。这一发现不仅拓展了经典研究中以第三方惩罚为主的公平维护行为, 还提供了一种启示: 反欺凌群体规范的维护中可能存在一种集约束(惩戒违规)、修复(弥补伤害)与强化(激励善行)功能于一体的第三方行为互动。本研究为理解反欺凌群体规范的维护行为提供了有益启示, 并为后续探索非冲突性反欺凌行为及干预方案开发提供了方向性依据。

首先, 本研究首次在欺凌假想情境中证实了中小学生会表现出“惩恶、扶弱、扬善”的非对抗性反欺凌行为, 突破了传统反欺凌研究局限于直接对抗行为的单一研究视角。研究将关注点转向事件结束后的后续群体互动, 发现在不直接对抗欺凌者的情况下, 事后知情的中小學生可能会通过减少帮助、资源分配倾斜等非冲突性方式来维护反欺凌规范。由此提示, 学生群体中的欺凌防治不应只关注直接制止欺凌、保护受害者、支持保护者的显性干预行为, 还应重视日常同伴互动中非对抗性的规范维护行为。相较于容易引发后续冲突的对抗性干预行为, 第三方惩罚、补偿与奖励无需直面欺凌者、受欺凌者与保护者, 而是通过关系调整与资源分配来削弱欺凌者的社交优势、弥补受欺凌者的处境劣势、强化保护者的积极行为, 从而为校园欺凌防治提供了一种相对安全且可行的同伴干预方式。这种规范维护方式也与中国传统文化的价值导向相契合。中国传统文化以“仁”为核心, 倡导“薄责于人”“为善者赏, 为不善者罚”, 强调帮助他人、避免直接冲突

(马睿 等, 2023), 这种文化导向为中国文化背景中的学生采用非对抗性方式维护反欺凌群体规范提供了文化适应性支撑。

其次, 本研究发现保护者在场会增强中小學生对欺凌者的经济性第三方惩罚。这一研究发现预示着, 欺凌假想情境中保护者的干预行为既可能有助于明确群体反欺凌共识, 降低其他旁观者的干预顾虑(Chen et al., 2025); 也可能有助于提供亲社会示范(Knauf & Eschenbeck, 2025), 激活旁观者的规范维护动机; 甚至使其从原本侧重于对欺凌者的否定(惩罚)与对受欺凌者的支持(补偿), 延伸至对保护者亲社会行为的积极反馈(奖励)。这种从“惩恶”到“扬善”的行为维度扩展, 不仅呼应了中国传统文化中“善不可失, 恶不可长”的基本价值取向, 也契合“隐恶而扬善”以及“道人善, 即是善; 人知之, 愈思勉”等强调劝善、扬善与促善的文化期许, 更让原本属于个体层面的反欺凌行为在同伴互动中得到群体层面的正向反馈。本研究发现还表明, 保护者在场对第三方惩罚与补偿的影响并非普遍性的增强, 而是具有一定特异性。由此在理解欺凌情境中的第三方行为时, 有必要进一步区分社会性与经济性维度, 并关注不同类型行为对具体情境线索的差异化反应。

最后, 本研究发现第三方行为在性别与学龄段上呈现出差异化模式, 反映了性别角色分化与社会规范内化的动态发展特征。相较于女生, 男生实施的经济性第三方惩罚与补偿水平更高, 这与社会对男性作为规则执行者的角色期待相契合(Alsawalqa et al., 2021; Eagly & Wood, 1999)。相较于社会性第三方行为, 经济性第三方行为具有更高的可见性和可量化性, 其惩戒与补偿意图更容易被群体清晰识别, 因此也更符合男性在群体互动中被期待承担的直接、明确的规则维护角色。随着学龄段增长, 个体的认知控制与情绪调节系统日趋成熟, 社会观点采择能力与规范内化水平显著提升(Hall et al., 2021; Silvers, 2022)。这意味着, 性别角色倾向的行为表达并非固定不变, 而会随着发展逐渐表现出更强的策略性, 即个体会结合具体情境和规范维护需求, 选择更安全、有效的行为方式。本研究中第三方奖励所呈现的性别与学龄段交互作用, 正体现了这种发展中的差异化选择。鉴于此, 欺凌防治实践中需要转向性别优势与发展阶段适配的整合策略, 既要关注不同性别在规范维护方式上的相对优势, 又要根据不同学龄段的认知与情绪特点进行差异

化引导。

本研究仍存在以下局限。首先, 本研究采用的欺凌假想情境(如绊倒、嘲笑)属于欺凌现象中较为常见、程度偏轻的类型。在更严重的身体欺凌或网络欺凌中, 第三方是否会因恐惧报复或感知无力而改变其行为决策, 有待未来研究探讨。其次, 本研究未对研究对象的过往欺凌参与经历(如曾是欺凌者、受欺凌者或保护者)、未来被报复预期以及性别角色认同等变量进行测量与控制。这些因素可能影响第三方行为的强度及其性别与年龄段差异。未来研究可将其纳入分析模型, 以更全面地理解行为背后的成因。最后, 本研究采用的社会性与经济性第三方惩罚、补偿与奖励范式, 主要以假想情境中的帮助意愿与游戏币分配为指标, 只能为后续探索非冲突性反欺凌行为及干预方案开发提供方向性依据, 不能直接外推至真实的校园欺凌互动情境中。未来研究可基于追踪设计的日记法或自然观察法, 探究真实校园互动情境中个体实施第三方行为的方式与强度。

6 结论

在欺凌假想情境中, 中小学生均愿意对欺凌者实施第三方惩罚、对受欺凌者实施第三方补偿(二元情境中社会性第三方补偿除外)、对保护者实施第三方奖励; 相比于没有保护者的欺凌情境, 中小学生在有保护者的欺凌情境中会对欺凌者实施显著更强的经济性第三方惩罚; 第三方惩罚、补偿与奖励上存在显著的性别差异与年龄段差异。

参 考 文 献

- Al-Alawi, K. S., Eltayib, R. A. A., Al Saadoon, M., Alhaj, A. H., Aldhafri, S., Al-Adawi, S., & Chan, M. F. (2025). Efficacy of psycho-educational interventions on school bullying in social isolation, loneliness and PTSD among schoolchildren: A systematic review and meta-analysis. *Child: Care, Health and Development*, 51(2), e70062.
- Alsawalqa, R., Alrawashdeh, M., & Hasan, S. (2021). Understanding the Man Box: The link between gender socialization and domestic violence in Jordan. *Heliyon*, 7(10), e08264.
- Andrews, J. L., Ahmed, S. P., & Blakemore, S. J. (2021). Navigating the social environment in adolescence: The role of social brain development. *Biological Psychiatry*, 89(2), 109–118.
- Artavia-Mora, L., Bedi, A. S., & Rieger, M. (2017). Intuitive help and punishment in the field. *European Economic Review*, 92, 133–145.
- Balafoutas, L., Nikiforakis, N., & Rockenbach, B. (2014). Direct and indirect punishment among strangers in the field. *Proceedings of the National Academy of Sciences of the United States of America*, 111(45), 15924–15927.
- Baum, G., Cui, Z., Roalf, D., Ciric, R., Betzel, R., Larsen, B., ... Satterthwaite, T. (2020). Development of structure-function coupling in human brain networks during youth. *Proceedings of the National Academy of Sciences of the United States of America*, 117(1), 771–778.
- Camacho, A., Runions, K., Ortega-Ruiz, R., & Romera, E. (2025). Norm misperception and witnessing bullying: The role of individual and contextual characteristics. *Journal of Adolescence*, 97(7), 1807–1817.
- Chen, G. H. (2014). The perception to definition of Qifu phenomenon from primary and junior middle students. *Studies of Psychology and Behavior*, 12(5), 639–644.
- [陈光辉. (2014). 中小学生对欺负现象本质内涵的感知. *心理与行为研究*, 12(5), 639–644.]
- Chen, G., Kong, Y., Deater-Deckard, K., & Zhang, W. (2018). Bullying victimization heightens cortisol response to psychosocial stress in Chinese children. *Journal of Abnormal Child Psychology*, 46(5), 1051–1059.
- Chen, G., Zhang, W., Zhang, W., & Deater-Deckard, K. (2020). A “defender protective effect” in multiple-role combinations of bullying among Chinese adolescents. *Journal of Interpersonal Violence*, 35(7–8), 1587–1609.
- Chen, Q., Lin, W., Wu, Q., & Chan, K. L. (2025). The effectiveness of interventions on bullying and cyberbullying bystander: A meta-analysis. *Trauma, Violence, & Abuse*, 26(5), 938–954.
- Darley, J. M., & Pittman, T. S. (2003). The psychology of compensatory and retributive justice. *Personality and Social Psychology Review*, 7(4), 324–336.
- de Kwaadsteniet, E. W., Rijkhoff, S. A., & van Dijk, E. (2013). Equality as a benchmark for third-party punishment and reward: The moderating role of uncertainty in social dilemmas. *Organizational Behavior and Human Decision Processes*, 120(2), 251–259.
- Decrop, R., McManamon, B., Williams, K., Houlihan, K., Bok, H. C., Knestrick, K., ... Cauffman, E. (2025). Cognitive loopholes of crime: Mapping the codevelopment of moral disengagement within perceptions of risks and rewards. *Journal of Research on Adolescence*, 35(3), e70056.
- Dhaliwal, N. A., Patil, I., & Cushman, F. (2021). Reputational and cooperative benefits of third-party compensation. *Organizational Behavior and Human Decision Processes*, 164(7), 27–51.
- Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist*, 54(6), 408–423.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160.
- Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2), 63–87.
- Feijóo, S. S., O’Higgins-Norman, J., Foody, M., Pichel, R., Braña, T., Varela, J., & Rial, A. (2021). Sex differences in adolescent bullying behaviours. *Psychosocial Intervention*, 30(2), 95–100.
- Fox, C. L., Jones, S. E., Stiff, C. E., & Sayers, J. (2014). Does the gender of the bully/victim dyad and the type of bullying influence children’s responses to a bullying incident? *Aggressive Behavior*, 40(4), 359–368.
- Gaffney, H., Ttofi, M., & Farrington, D. (2021). What works in anti-bullying programs? Analysis of effective intervention components. *Journal of School Psychology*, 85, 37–56.

- Geraci, A., & Surian, L. (2021). Toddlers' expectations of third-party punishments and rewards following an act of aggression. *Aggressive Behavior*, 47(5), 521–529.
- Grueneisen, S., & Warneken, F. (2022). The development of prosocial behavior—from sympathy to strategy. *Current Opinion in Psychology*, 43, 323–328.
- Gutierrez, B. C., Halim, M. L. D., Martinez, M. A., & Arredondo, M. (2020). The heroes and the helpless: The development of benevolent sexism in children. *Sex Roles*, 82, 558–569.
- Hall, H. K., Millea, P. M., Summers, M. J., & Isbel, B. (2021). Longitudinal research on perspective taking in adolescence: A systematic review. *Adolescent Research Review*, 6(2), 125–150.
- Hawkins, D. L., Pepler, D. J., & Craig, W. M. (2001). Naturalistic observations of peer interventions in bullying. *Social Development*, 10(4), 512–527.
- Jordan, J., Hoffman, M., Bloom, P., & Rand, D. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature*, 530, 473–476.
- Knauf, R. K., & Eschenbeck, H. (2025). School bullying and cyberbullying—associations of student's roles with social-cognitive and affective reactions. *Social Psychology of Education*, 28(1), 51.
- Laniga - Wijnen, L., Malamut, S. T., Garandeau, C. F., & Salmivalli, C. (2023). Does defending affect adolescents' peer status, or vice versa? Testing the moderating effects of empathy, gender, and anti-bullying norms. *Journal of Research on Adolescence*, 33(3), 913–930.
- Ma, R., Wu, N., Tian, M. Q., & Wu, Z. (2023). Punishing perpetrators or compensating victims: The third-party punishment motivation of children aged 7–10. *Chinese Science Bulletin*, 68(17), 2258–2268.
- [马睿, 吴南, 田莫千, 伍珍. (2023). 惩恶还是扶弱: 7~10 岁儿童第三方惩罚的动机. *科学通报*, 68(17), 2258–2268.]
- Ma, T. L., & Chen, W. T. (2019). The benefits of being defended: Perceived bystander participant roles and victims' emotional and psychosocial adjustment. *Journal of School Violence*, 18(1), 77–91.
- Ma, T. L., Meter, D. J., Chen, W. T., & Lee, Y. (2019). Defending behavior of peer victimization in school and cyber context during childhood and adolescence: A meta-analytic review of individual and peer-relational characteristics. *Psychological Bulletin*, 145(9), 891–928.
- Michalska, K. J., Kinzler, K. D., & Decety, J. (2013). Age-related sex differences in explicit measures of empathy do not predict brain responses across childhood and adolescence. *Developmental Cognitive Neuroscience*, 3, 22–32.
- Molho, C., Tybur, J. M., Van Lange, P. A., & Balliet, D. (2020). Direct and indirect punishment of norm violations in daily life. *Nature Communications*, 11(1), 3432.
- Ng, K., Niven, K., & Hoel, H. (2020). 'I could help, but...': A dynamic sensemaking model of workplace bullying bystanders. *Human Relations*, 73(12), 1718–1746.
- Olweus, D. (1993). *Bullying at school: What we know and what we can do*. Oxford, UK: Blackwell.
- Priess, H. A., Lindberg, S. M., & Hyde, J. S. (2009). Adolescent gender-role identity and mental health: Gender intensification revisited. *Child Development*, 80(5), 1531–1544.
- Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7(4), 331–363.
- Romer, D., Reyna, V. F., & Satterthwaite, T. D. (2017). Beyond stereotypes of adolescent risk taking: Placing the adolescent brain in developmental context. *Developmental Cognitive Neuroscience*, 27, 19–34.
- Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, 22(1), 1–15.
- Salmivalli, C., & Voeten, M. (2004). Connections between attitudes, group norms, and behaviour in bullying situations. *International Journal of Behavioral Development*, 28(3), 246–258.
- Silvers, J. (2022). Adolescence as a pivotal period for emotion regulation development. *Current Opinion in Psychology*, 44, 258–263.
- Thornberg, R., Pozzoli, T., & Gini, G. (2022). Defending or remaining passive as a bystander of school bullying in Sweden: The role of moral disengagement and antibullying class norms. *Journal of Interpersonal Violence*, 37(19–20), NP18666–NP18689.
- Toribio-Flórez, D., Saße, J., & Baumert, A. (2023). "Proof under reasonable doubt": Ambiguity of the norm violation as boundary condition of third-party punishment. *Personality and Social Psychology Bulletin*, 49(3), 429–446.
- Turillo, C. J., Folger, R., Lavelle, J. J., Umphress, E. E., & Gee, J. O. (2002). Is virtue its own reward? Self-sacrificial decisions for the sake of fairness. *Organizational Behavior and Human Decision Processes*, 89(1), 839–865.
- Van der Graaff, J., Carlo, G., Crocetti, E., Koot, H. M., & Branje, S. (2018). Prosocial behavior in adolescence: Gender differences in development and links with empathy. *Journal of Youth and Adolescence*, 47(5), 1086–1099.
- van der Ploeg, R., Kretschmer, T., Salmivalli, C., & Veenstra, R. (2017). Defending victims: What does it take to intervene in bullying and how is it rewarded by peers? *Journal of School Psychology*, 65, 1–10.
- Warthen, K. G., Boyse-Peacor, A., Jones, K. G., Sanford, B., Love, T. M., & Mickey, B. J. (2020). Sex differences in the human reward system: Convergent behavioral, autonomic and neural evidence. *Social Cognitive and Affective Neuroscience*, 15(7), 789–801.
- Xiao, S. X., Hashi, E. C., Korous, K. M., & Eisenberg, N. (2019). Gender differences across multiple types of prosocial behavior in adolescence: A meta-analysis of the prosocial tendency measure-revised (PTM-R). *Journal of Adolescence*, 77, 41–58.
- Yang, H., Zhang, Y., Lyu, Y., & Tang, C. (2023). Group bias under uncertain environment: A perspective of third-party punishment. *Acta Psychologica*, 237, 103957.
- Zhang, W. X. (2023). The definition of bullying: From cross-cultural and developmental perspectives. *Psychological Development and Education*, 39(4), 590–598.
- [张文新. (2023). 欺凌的界定: 文化和发展的视角. *心理发展与教育*, 39(4), 590–598.]
- Zhao, K., Yan, Y., Zhou, Z. E., Xiao, S., Shan, D., & Iqbal, M. (2025). Punish or help? Third parties' constructive responses to witnessed workplace incivility: The role of political skill. *Humanities and Social Sciences Communications*, 12, 516.

Punishing bullies, compensating victims, and rewarding defenders: Third-party behaviors from the perspective of bullying prevention and intervention

WANG Fan, CHEN Guanghui, CHEN Peikun, ZHI Nuo, CHANG Yuan, XIAN Yemin, ZHANG Wenxin

(Faculty of Psychology, Shandong Normal University; Shandong Provincial Key Laboratory of

Brain Science and Mental Health, Jinan 250014, China)

Abstract

Bullying is a widespread aggressive behavior in school environments in which stronger individuals intentionally harm weaker peers who are unable to defend themselves effectively. Direct confrontation with bullies may incur retaliation or social risks, making indirect and non-conflictive third-party behaviors feasible paths to uphold anti-bullying norms. On the basis of the framework of punishing bullies, compensating victims, and rewarding defenders, this study systematically explored third-party punishment, compensation, and reward behaviors among primary, junior high, and senior high school students, as well as their differences across gender and school stages in bullying contexts.

A total of 1011 students from primary, junior high, and senior high schools participated in three sequential studies using economic game paradigms and hypothetical scenario paradigms. Study 1 adopted a 2 (gender: male vs. female) $\times$ 3 (school stage: primary vs. junior high vs. senior high) between-subjects design to examine third-party punishment and compensation in the “bully-victim” dyadic interaction context. Study 2 extended to the “bully-victim-defender” triadic interaction context, adding third-party rewards as an outcome variable while maintaining the same design as Study 1. Study 3 integrated data from Studies 1 and 2, using a 2 (defender presence: absent vs. present) $\times$ 2 (gender) $\times$ 3 (school stage) between-subjects design to compare differences in third-party punishment and compensation across the two contexts.

The results showed that, in hypothetical bullying contexts, students showed both social and economic third-party punishment toward bullies, third-party compensation toward victims (except for social third-party compensation in the dyadic context without a defender), and both social and economic third-party rewards toward defenders when a defender was present, with both social (e.g., helping willingness) and economic (e.g., token allocation) dimensions examined. Compared with the dyadic context (without defenders), the triadic context (with defenders) significantly enhanced students' economic third-party punishment toward bullies. Senior high school students exhibited significantly higher levels of social third-party punishment toward bullies than junior high school and primary school students. Males showed significantly higher levels of economic third-party punishment and compensation than females. Junior high school females demonstrated higher social third-party rewards, whereas senior high school males showed higher levels of both social and economic third-party rewards.

These findings suggest that primary and secondary school students show third-party behavioral tendencies consistent with punishing bullies, compensating victims, and rewarding defenders in hypothetical bullying contexts. The presence of defenders strengthens norm enforcement by enhancing economic punitive behaviors, and third-party behaviors exhibit distinct developmental and gender-specific characteristics. Given that this study relied on hypothetical scenarios, helping willingness, and token allocation, the findings should not be interpreted as direct evidence for the effectiveness of real-world school-based interventions. Rather, they provide a directional basis for future exploration of non-conflictive anti-bullying behaviors and intervention design.

Keywords primary and secondary school students, school bullying, third-party punishment, third-party compensation, third-party reward