

校园霸凌行为问卷的编制*

赵福棠¹, 何 壮¹, 袁淑莉¹, 黄希庭²

(1. 贵阳学院教育科学学院, 贵阳 550005; 2. 西南大学心理学与社会发展研究中心, 心理学部, 重庆 400715)

摘 要:目的:从贵州省青少年校园霸凌行为的实际出发,以大样本为基础,遵循心理测量学的方法编制《校园霸凌行为问卷》。方法:经理论建模、题库建设编制初测问卷,以3923名学生为被试进行项目分析和探索性因子分析;以3899名被试进行信度检验及验证性因子分析,形成最终问卷。结果:问卷由7个一阶因子和2个二阶因子组成,共40个题项。问卷及各维度的内部一致性系数在0.7~0.927之间,重测信度在0.291~0.49之间,结构效度良好,与OBVQ-V和MBVS的相关分别为0.646和0.726。结论:本研究探索、发现并整合创新了反映当代校园霸凌治理难点的“恶整霸凌”;问卷信效度良好,可作为国内青少年校园霸凌行为的有效测量工具。

关键词:校园霸凌;问卷构建;信度;效度

中图分类号:B841.2

文献标识码:A

文章编号:1003-5184(2021)01-0064-05

校园霸凌是中小学校学生个体或群体受到力量较强一方蓄意或恶意、持续多次的身心攻击行为。校园霸凌有恃强凌弱、蓄意伤害、重复发生的本质特征,对霸凌者、受害者及旁观者的身心都存在不同程度的影响和危害,扭曲其对教育、社会的看法进而改变正常的行为方式,甚至导致受害者长期抑郁、失去或放弃生命,霸凌者也会遭到报复。

校园霸凌是世界各国致力研究和防治的重大现实问题,2014年中国基础教育质量监测协同创新中心“区域教育质量健康体检”项目的调查发现,遭受过身体霸凌、言语霸凌和关系霸凌学生占抽样总体的比例分别为30.85%、58.28%、43.34%(胡咏梅,李佳哲,2018)。国内各项研究的调查结果无疑有重要的研究价值和应用价值,但就测量工具的适切性值得商榷。国内相关研究常用的量表较新的有谢家树等2015年引入中国的特拉华欺凌受害量表(Delaware Bullying Victimization Scale-Student, DBVS-S)(谢家树,魏宇民, Bear, 2018), Harbin等2019编制的多维霸凌受害量表(Multidimensional Bullying Victimization Scale, MBVS)。其中影响力最大、使用最多的问卷为张文新等1999年修订的Olweus欺负问卷中文版(Olweus Bully/Victim Questionnaire, OBVQ)(张文新,武建芬,1999)。但上述工具在不少环节上存在缺陷(赵福棠,何壮,袁淑莉,黄希庭,2020)。首先,这些工具的测量内容不

统一,影响了研究结论可比性。OBVQ、DBVS-S将霸凌行为划分为身体霸凌、言语霸凌和关系霸凌, MBVS则划分为直接霸凌、间接霸凌和评价霸凌。其次,题项长度不一,测量的精度差异较大。第三,题项内容没有包含信息技术的普及和发展情况下明显呈逐年增长趋势的网络霸凌(Smith, Mahdavi, Carvalho, Fisher, & Tippett, 2008);肢体霸凌也不再只是显性的身体伤害,转而多用对身体的恶整以及胁迫器官功能的倒置等不易发现的、非人性的、教育与社会难以规制的羞辱方式。最高检向教育部发出史上首份检察建议书(一号检察建议)直指对中小校园性霸凌的关注,遗憾的是,上述研究工具均未纳入相应内容。最关键的是,这些翻译过来的测量工具都存在跨文化适切性问题,甚至同一工具的不同译本在措辞上仍存在语义差异(Cheng, Chen, Liu, & Chen, 2011)。有鉴于此,此前多用的测量工具因其内容上的涵括性不够、文化的适切性欠缺和题项设计的局限性,其科学性也就存疑。

因此,以我国中小学实际为基础,编制一份适用于我国青少年学生的校园霸凌问卷很有必要。

1 校园霸凌初测问卷编制

1.1 方法

1.1.1 被试

样本1:采取分层抽样的方法,在贵州省抽取6所学校,其中小学、高中、完全中学、十二年一贯制学

* 基金项目:国家社会科学基金教育学一般课题“青少年校园霸凌的特点和机制研究:以贵州省为例”(BBA170070)。

通讯作者:黄希庭, E-mail: xthuang@swu.edu.cn。

校各 1 所,初中 2 所;涵盖小学四年级至高三 9 个年级。共收集有效问卷 3923 份,年龄分布在 9~18 岁之间,平均年龄 13.6 ± 1.7 岁,小学、初中、高中被试比例 1: 6: 3,男女比例 5: 5,城乡比例 6: 4,汉族与少数民族比例 6: 4。

1.1.2 初测问卷编制

(一)理论建模

根据心理测量学的要求,基于我国校园霸凌行为的现实,在编制测量校园霸凌行为工具时的理论构建上,突出了以下几点。

首先,突出中国文化背景下校园霸凌的特点。我国的传统文化反对霸凌,但另一方面,黑恶崇拜、码头文化、拉帮结派、江湖义气等不良文化侵蚀校园。

其次,突出多元特点。通过案例分析发现国内校园霸凌的行为多种多样,可以从现实/虚拟、个体/群体等多个角度描述。在分类上承继学界认同的“言语霸凌”、“关系霸凌”,整合“肢体霸凌”及当下对身体的恶搞与整蛊、器官功能的倒置等为“恶整霸凌”。恶整霸凌是指缺乏人性、残害身体的攻击行为。还增加了“性霸凌”、“网络霸凌”。

第三,突出题项测量指标的具象化。研究过程中采取了“动机→行为指标→题项设计”的思路,总结出不同类型校园霸凌从动机到行为表现的特点,为编制题项提供参考。

(二)题库建设

在建模的基础上,结合案例提取信息总结出的霸凌行为,决定题项的编制从校园霸凌发生的时空(现实空间/虚拟世界)和类型(言语霸凌/关系霸凌/性霸凌/恶整霸凌/网络言语霸凌/网络关系霸凌/网络性霸凌)角度展开。题项编写过程:先由中小学教师、管理干部和校长,心理学专业研究生,大学心理学、教育学、社会学、法学专业教师组成课题组编写出初稿。再请中小学生学习阅读题项,根据其反馈修改内容、删除内容重复题项。最后经多轮小范围试测得到初测问卷(共 63 题),并将其命名为校

园霸凌行为问卷(School Bullying Victims Questionnaire,SBVQ)。

(三)初测问卷

初测问卷包括人口学变量 5 题、校园霸凌行为 63 题;除人口学变量外,校园霸凌行为问题均为李克特 5 级自评量表,1 代表从未发生、5 代表频繁发生。

1.1.3 问卷施测与数据处理

采取纸笔填写方式施测,要求学生独立作答,并给予充足的时间。对收回的问卷进行筛选,将作答不完整、填写不规范的问卷剔除。数据分析使用软件为 SPSS22。

1.2 结果与分析

1.2.1 项目分析

采用高低分组(前 27%、后 27%)独立样本 t 检验和决断值指标检验题项的鉴别力。结果显示,所有题项高低分组之间均值都存在显著差异($p < 0.01$);且临界值在[4.58,60.8]之间,均大于相关研究建议的标准 3。采用修正后的题总相关和内部一致性系数检验题项的同质性。结果显示第 50 题(被他人强吻)修正后的题总相关为 0.375,低于 0.4,应予删除;其他题项相关系数在[0.4,0.67]之间。内部一致性统计结果显示,各题项删除后的内部一致性系数均低于全部题项的内部一致性系数 0.957。

1.2.2 探索性因子分析

对剩余 62 题进行探索性因子分析,KMO = 0.963,Bartlett 球形检验卡方值 = 106935.394, $p < 0.01$ 。采用主成分分析、正交最大旋转法,共获得 7 个特征根大于 1 的主成分,首成分方差解释率占 33.57%,累计方差解释率 65.44%。综合特征根、碎石图拐点、因子载荷(大于 0.5)、共同度(大于 0.2)、MSA(大于 0.5)等指标,经三次因子分析后:MSA 在[0.869,0.978]之间,共同度在[0.428,0.84]之间,载荷在[0.547,0.867]之间。最终得到 7 个因子、40 个题项,因子命名及题项内容如表 1 所示。

表 1 初测问卷一阶因子命名

因子命名及编号	题项数量	霸凌行为
言语霸凌(F1)	8	嘲笑谩骂、浑名戏谑、流言蜚语、揭短痛斥、造谣污蔑
恶整霸凌(F2)	8	踢打推毆、器械伤害、物品羞辱、恶搞整蛊、抓扯撕逼
性霸凌(F3)	3	下流动作、胁迫非礼、语言挑逗
关系霸凌(F4)	6	社群排挤、仗势隔离、挑拨离间
网络言语霸凌(F5)	5	恐吓威胁、匿名栽赃、盗名蛊惑、水军谩骂、造谣污蔑
网络关系霸凌(F6)	4	社群排挤、散播隐私、挑拨离间
网络性霸凌(F7)	6	散播隐私、推送淫秽、偷拍威胁

对一阶模型 7 个因子分析后发现,除必须面对面实施的恶整霸凌外,言语霸凌、关系霸凌和性霸凌都存在一个与之对应的网络霸凌因子。因此,结合已有研究和理论,进一步探索高阶因子结构。高阶因子分析 $KMO = 0.806$, Bartlett 球形检验卡方值 = 10882.56, $p < 0.01$, 共获得 2 个特征根大于 1 的主

成分,首成分方差解释率 36.49%,累计方差解释率 66.48%。一阶因子中所有在物理空间发生的霸凌行为共同组成了“现实世界”因子、所有在互联网发生的霸凌行为共同组成了“虚拟空间”因子,各因子载荷、MSA 和共同度如表 2 所示。

表 2 初测问卷二阶因子命名 ($n = 3923$)

一阶因子及编号	高阶因子及编号		共同度	MSA
	现实世界(H1)	虚拟空间(H2)		
言语霸凌(F1)	0.869		0.765	0.777
恶整霸凌(F2)	0.81		0.686	0.793
性霸凌(F3)	0.751		0.592	0.794
关系霸凌(F4)	0.772		0.671	0.875
网络言语霸凌(F5)		0.63	0.606	0.805
网络关系霸凌(F6)		0.54	0.596	0.814
网络性霸凌(F7)		0.856	0.738	0.773

2 正式问卷测试及信效度检验

2.1 方法

2.1.1 被试

样本 2:在样本 1 所选学校抽样,共获得有效问卷 3899 份,其中 2409 人曾参加前测,将其作为重测样本,另有 1490 人为初次参加测试。年龄分布在 9~18 岁之间,平均年龄 13.15 ± 2.5 岁,小学、初中、高中被试比例 2:5:3,男女比例 5:5,城乡比例 7:3,汉族与少数民族比例 7:3。

2.1.2 工具

正式问卷包含初测问卷的 40 个题项和人口学变量 5 题,共 45 题。效标问卷 1 为 OBVQ 的被欺负部分,共 9 题,测量言语、身体两种霸凌行为,采用 5 级计分,1 代表从未发生、5 代表一周好几次;在此次研究中的内部一致性系数为 0.834。效标问卷 2 为 MBVS,共 26 题,分为直接、间接、评价霸凌 3 个因

子,采用 4 级计分,1 代表从未发生、4 代表频繁发生,在此次研究中的内部一致性系数为 0.916。

2.1.3 施测与数据处理

采取纸笔填写与涂卡相结合的方式施测,前测、后测间隔 4~6 周;无效问卷剔除标准同初测;验证性因子分析使用 AMOS 22。

2.2 结果与分析

2.2.1 信度

信度检验结果显示,各一阶因子内部一致性系数在[0.763,0.907]之间,二阶两个因子的内部一致性系数在[0.7,0.703]之间,问卷总体的内部一致性系数为 0.927。

4~6 周后 2409 人参加重测,采用前、后测相关法评估重测信度,各一阶因子重测信度在[0.291,0.476]之间,二阶两个因子的重测信度在[0.439,0.473]之间,问卷总体的重测信度为 0.49。

表 3 正式问卷信度

因子编号	F1	F2	F3	F4	F5	F6	F7	H1	H2	总体
$\alpha (n = 3899)$	0.895**	0.885**	0.793**	0.873**	0.856**	0.763**	0.907**	0.703**	0.7**	0.927**
重测信度($n = 2409$)	0.476**	0.348**	0.44**	0.395**	0.442**	0.291**	0.393**	0.473**	0.439**	0.49**

注:*代表 $p < 0.05$, **代表 $p < 0.01$;下同

2.2.2 效度

(一)结构效度

各因子与总分间的相关系数如表 4 所示。一阶各因子之间相关系数在[0.11,0.637]区间,二阶两个因子之间相关系数为 0.4,均为中低程度正相关;表明各因子在确保测量内容一致的情况下又有差

异,互相不可替代。一阶因子与总分的相关在[0.298,0.895]区间,二阶两个因子与总分相关在[0.566,0.982]区间;因子与总分的相关大于因子间相关;各相关系数均达到了相关研究建议的标准(Tucker & Lewis,1973),表明各因子与问卷总体测量的内容一致性较高。

表 4 因子与总分相关系数矩阵($n = 3899$)

因子编号	F1	F2	F3	F4	F5	F6	F7	H1	H2
F2	0.637**								
F3	0.349**	0.376**							
F4	0.552**	0.471**	0.275**						
F5	0.301**	0.233**	0.302**	0.346**					
F6	0.346**	0.263**	0.302**	0.405**	0.611**				
F7	0.11**	0.16**	0.324**	0.132**	0.402**	0.285**			
H1	0.92**	0.832**	0.468**	0.733**	0.353**	0.403**	0.174**		
H2	0.327**	0.28**	0.390**	0.381**	0.857**	0.822**	0.684**	0.4**	
总分	0.895**	0.806**	0.51**	0.738**	0.495**	0.532**	0.298**	0.982**	0.566**

采用结构方程模型对问卷正式版进行验证性因子分析以评估其结构效度,考察模型的合理性。为检验假设模型是否为最优模型,还将假设模型与不区分因子模型(竞争模型 M1)和一阶 7 因子模型

(竞争模型 M2)的拟合进行了比较。三个模型的拟合指数如表 5 所示,M1 不符合相关理论假设,M2 和假设模型均达到了相关研究建议的标准,但假设模型拟合更好、且模型的简洁性更好。

表 5 不同模型拟合指数($n = 3899$)

模型	χ^2/df	IFI	TLI	CFI	RMSEA(90% CI)
假设模型	12.358	0.913	0.906	0.913	0.054(0.053,0.055)
M1	73.188	0.408	0.377	0.408	0.136(0.135,0.137)
M2	12.839	0.911	0.902	0.911	0.055(0.054,0.056)

进一步检验高阶模型的聚合效度,拟合指标如表 6 所示。一阶因子题项载荷在[0.514,0.984]区间,均在相关研究建议的 0.5 以上;二阶因子下的七个一阶因子载荷在[0.331,0.898]区间,性霸凌因子载荷为 0.482,接近 0.5 的建议;网络性霸凌因子载荷 0.331,虽然载荷较低,但未对高阶因子的聚合效度产生较大影响,因此予以保留。一阶因子的组合信度在[0.773,0.907]之间,二阶因子的组合信度在[0.718,0.811]之间,均在相关研究建议的 0.7 以上;一阶因子的平均变异抽取量(AVE)在[0.46,0.625]之间,二阶因子的平均变异抽取量在[0.45,0.528]之间,均接近或达到相关研究建议的 0.5 以上标准(吴明隆,2013)。

表 6 高阶模型聚合效度($n = 3899$)

因子名称及编号	组合信度	AVE
言语霸凌(F1)	0.907	0.531
恶整霸凌(F2)	0.887	0.497
性霸凌(F3)	0.811	0.595
关系霸凌(F4)	0.877	0.548
网络言语霸凌(F5)	0.851	0.546
网络关系霸凌(F6)	0.773	0.46
网络性霸凌(F7)	0.907	0.625
现实世界(H1)	0.811	0.528
虚拟空间(H2)	0.718	0.45

(二)效标效度

正式问卷的效标相关分析结果显示:正式问卷总分与 OBVQ-V 相关为 0.646,与 MBVS 的相关为

0.726,具有较高的效标效度。

3 讨论

3.1 理论建模反映了我国校园霸凌的现状

理论建模过程中正视了当代中国负面世俗文化对青少年成长的影响。这种文化与中华民族优秀传统文化和社会主义核心价值观格格不入,背离党的教育方针、立德树人的根本任务,反映出西方不良文化、传统文化的阴暗残留、媒体渲染误导、社会的不良风气等在中国社会快速变迁中对社会主义教育价值观、学校主流文化的冲击与侵蚀,集中体现出与学校主流文化不相容的亚文化形态。经过分类剖析,使模型建构紧贴中国校园霸凌实际,为铲除不良思想的污染,正本清源,实施校园治理、教育预防、心理干预提供依据,进而为弘扬正气、探究培养青少年学生养成正确价值观、积极自我观和追求理想的优秀品质打下良好基础(黄希庭,2014)。

模型反映了当下学校亚文化形态中校园霸凌的特点与现状,并体现在模型的建构中,使其客观的反映校园霸凌的全貌。在建模与问卷编制过程中,遵循心理学测量学要求,借鉴和吸收国外已有校园霸凌行为测量工具研究的合理成分,在其基础上增加了新时代背景下网络霸凌表现的不同样态,将社会关切的、困扰教育的性霸凌、恶整霸凌纳入到模型中,并本土化测量指标,提出了现实/虚拟的高阶心理结构假设。

贵州的校园霸凌的现状能代表全国的一般状

况。比较同一时期的几项调查发现,在对校园霸凌的类型划分方面,此次研究的结论在相关研究中均得以印证。研究结论中,现实世界二阶因子下包含言语霸凌、恶整霸凌、关系霸凌和性霸凌四种类型,这与谢家树等人的系列研究相一致(谢家树,谢璐, Yang, 2016)。研究结论中,虚拟空间二级因子下包含网络言语霸凌、网络关系霸凌和网络性霸凌三种类型,典型的行为包括:恐吓威胁、匿名诽谤、冒名顶替、水军谩骂等十余种行为,这些典型行为的选择与陈启玉等人的研究一致(陈启玉,唐汉瑛,张露,周宗奎, 2016)。

3.2 类型划分及测量指标的选择

在测量指标的选择上,基于对学校师生的访谈、监狱在押罪犯的回溯式访谈,汇集了来自学校、法院、网上收集的个案,分类梳理、剖析特点、提取信息、反复比对,发现校园霸凌发生、发展的原因和机制等复杂情况。一种动机可能会引起多种霸凌行为,例如中学生群体中常见的因情感问题而产生的霸凌问题,因卷入者身份不同可能会导致多种霸凌行为;比如因歧见、嫉妒或挫败等而产生的排挤(关系霸凌)、因情感竞争而产生的争斗(恶整霸凌)。一种霸凌行为也可能是多种动机共同引起的,多种霸凌行为之间还有可能互相转换。鉴于心理测量学的要求,以行为指标设定题项,区分出一个指标的多个题项,再邀请有关专家、中小学校长与教师评议、补充和完善,最后,通过问卷的初测、正式施测,基于结构方程建模得到 7 个一阶因子和 2 个二阶因子组成的问卷。可以看出,该问卷与以往问卷相比,指标更全面,题项更客观更具可操作性,同时体现了一般性和代表性。

分析结果表明,问卷所测内容与两个效标问卷有较高的一致性,具有较高的效标效度。现实世界因子和虚拟空间因子与 MBVS 的直接霸凌因子相关分别为 0.707 和 0.248,这一结果进一步证明了问卷的内容效度。

3.3 研究的突破与发现

虽然研究借鉴并吸收了国内外学者已有的相关研究成果,但这不是重复他人的研究,不是修订别人的问卷,而是有一个重要突破:这就是研究路径的突破。突破了既往学界认同并使用的校园霸凌问卷基本上源自于对国外学者问卷的修订,而本问卷是从我国校园霸凌的现实出发,用心理测量学的方法而编制出来的,有其独特的特点。研究有三个发现:一是在问卷研制的过程中及时捕捉并发现了有关恶整

霸凌的新证据。恶整霸凌是对肢体伤害、胁迫生理器官功能倒置、暴虐体表与恶意整蛊的整合,进而提出了一个概念,即关于恶整霸凌的概念。这类霸凌没有外伤,难以及时发现、处罚,伤害程度难以评估。长久以来,这种攻击行为被校规、校纪及相关法律所忽视,心身严重危害性急需加以关注。二是及时发现了网络霸凌的新形式。在网络信息空间背景下,现实世界中言语霸凌、关系霸凌、性霸凌表现出新的形态,而且决定霸凌的强弱关系也可能会发生改变;现实世界中生理或心理处于弱势的群体可能转变为网络空间的强者,并实施霸凌。三是探索性因子分析发现,言语霸凌、关系霸凌、性霸凌均存在现实和虚拟两种表现形式,因此进一步提出校园霸凌的高阶因子。虚拟空间的霸凌是现实世界霸凌的新型表现形式,具有匿名性、持续伤害性等特点。当今信息时代对虚拟空间霸凌尤其值得关注。本问卷不仅全面涵盖各种网络霸凌行为,且进一步提出的高阶模型假设为后续问卷调查和数据分析提供了更宽广的视角。总之,编制的测量工具基本上包含了当下中国校园霸凌行为的不同形态,有利于开展全方位的调查研究以及后续的治理。

参考文献

- 陈启玉,唐汉瑛,张露,等.(2016).青少年社交网站使用中的网络欺负现状及风险因素——基于 1103 名 7-11 年级学生的调查研究.《中国特殊教育》,(3),89-96.
- 黄希庭.(2014).《探究人格奥秘》.北京:商务印书馆.
- 胡咏梅,李佳哲.(2018).谁在受欺凌?——中学生校园欺凌影响因素研究.《首都师范大学学报(社会科学版)》,(6),171-185.
- 吴明隆.(2013).《结构方程模型:Amos 实务进阶》.重庆:重庆大学出版社.
- 谢家树,魏宇民, Bear, G. (2018). 特拉华欺凌受害量表(学生卷)中文版再修订及初步应用.《中国临床心理学杂志》,26(2),259-263.
- 谢家树,谢璐, Yang, C. Y., 等.(2016).中美青少年校园欺负受害问题的比较研究.《中国临床心理学杂志》,24(4),706-709.
- 张文新,武建芬.(1999).Olweus 儿童欺负问卷中文版的修订.《心理发展与教育》,(2),8-12.
- 赵福巢,何壮,袁淑莉,黄希庭.(2020).奥尔维斯欺负量表的 Rasch 模型分析.《西南大学学报(社会科学版)》,46(5),115-121.

(下转第 90 页)