

第十四届全运会女子体操E组裁判评分的客观性评析

王丰彩, 李明

(集美大学诚毅学院, 福建 厦门 361021)

摘要:采用文献资料法、视频观察法和数理统计法对第十四届全运会女子体操单项决赛E组裁判评分数据进行了客观性分析,结果表明:四个单项决赛项目的E组裁判员评分的客观性系数都比较高,裁判员对本届参赛运动员评分的尺度把握总体较为一致,评分的随机误差总体上控制较好;四个项目的评分差异不显著,E组裁判员评分的标准比较一致。虽然有个别裁判员评分的个体间差异及随机误差稍大,极差值和无效分的出现机率稍大,但是对运动员最后得分的可信度影响不大。建议:合理控制和降低随机误差,提高裁判评分的区分度,减少无效分的出现机率,进一步缩小极差评分值的波动范围,确保E组裁判评分的客观性。

关键词:全国运动会;体操比赛;裁判员;E组;客观性

中图分类号:G 832

文献标识码:A

文章编号:1007-7413(2023)02-0062-05

Objective Evaluation on the Judging Score of Group E of Women's Gymnastics in the 14th National Games

WANG Feng-cai, LI Ming

(Chengyi College, Jimei University, Xiamen 361021, China)

Abstract: By using the methods of literature, video observation and mathematical statistics, the objective analysis of the scoring data of group E of women's individual gymnastics finals in the 14th National Games was carried out. The results showed that the objectivity coefficient of the scores of Group E judges in the finals of four single event final events was relatively high, the judges had a consistent grasp of the scale of the athletes' scoring, and the random error of the scoring was generally well controlled. There was no significant difference in the scores of the four items, and the scores of the judges in group E were consistent. Although there were some individual differences and random errors of individual judges' scores, and the occurrence probability of range value and invalid score was slightly greater, it had little influence on the credibility of the final score of athletes. It is suggested that all judges should rationally control and reduce the random error, improve the differentiation degree of scoring, reduce the occurrence probability of invalid scores, further reduce the fluctuation of range score value to ensure the scoring objectivity of the judges of group E.

Key words: National Games; gymnastics competition; judge; group E; objectivity

全运会作为规模最大的全国性体育赛事,运动员的参赛情况可以准确地反映出各个地区的体育发展水平。体操是我国的传统优势项目,在发展全民健身,增强全民体质的大众体育推广中,有着不可替代的助推作用。体操运动员的竞赛成绩显示了其代表地方的体操技术发展水平和综合实力。在体操比赛过程中,运动员的比赛成绩一方面受到自身动作完成情况的影响,另一方面也受到裁判员主观评分因素的影响。

随着2021年第十四届全运会体操项目比赛在陕西奥体中心体育馆成功举办,运动员的精彩表现为陕西奥体中心的赛事拉开了帷幕。此届女子体操比赛采用《2017—2020国际女子竞技体操评分规则》的评分标准,按照竞赛规程,进入决赛的运动员是预赛中各个单项的前八名。女子竞技体操评分规则中,有两组评委给参赛队员评分。D组裁判评难度分,由动作难度、连接难度、特定要求构成,不设上限。E组裁判评完成分,对参赛队员成套动作中出现的错误进行扣

收稿日期:2021-08-16

第一作者简介:王丰彩(1981—),女,河南汝州人,副教授。研究方向:学校体育学、体操、排舞。

分,E组裁判评定E分,是从10分开始,以0.1分为单位进行扣分,小失误扣0.10分,中等失误扣0.30分,大失误扣0.50分,重大失误扣1.00分。两组裁判评分都是独立的,去掉最高分和最低分,最后的平均分即为参赛队员的最后得分。^[1]因同年龄段的参赛队员难度得分即D分的差异性不是很大,因此,E组裁判评分的客观性对比赛结果起着决定性的作用。代表国内最高执裁水平的全运会比赛,尤其是最能体现裁判员执裁水平的主观评分类项目比赛中的体操E组裁判员,其评分的客观性如何?为研究女子体操单项决赛E组裁判员评分的客观性,本研究以第十四届全运会女子体操项目比赛单项决赛成绩中E组裁判的评分成绩作为研究对象,对裁判员评分的客观性进行多方面的综合分析与评价,研究结果将有助于我们认识和了解全运会体操比赛裁判员执裁水平,优化我国体操裁判员的业务培训及遴选工作,为进一步科学化监督和评价裁判员提供参考依据。

1 研究对象与方法

1.1 研究对象

以2021年第十四届全运会体操项目比赛女子项目中的高低杠、平衡木、跳马和自由体操四个单项决赛前八名运动员完成情况即E分作为研究对象。

1.2 研究方法

1.2.1 文献资料法

通过中国知网和国际体操联合会官网等查阅相关文献资料,并以《2017—2020国际女子竞技体操评分规则》和第十四届全运会女子体操成绩册作为参考依据进行分析。

1.2.2 视频观察法

通过观看第十四届全运会的现场直播以及赛后视频回放,对女子体操各单项决赛的赛事过程进行记录,对运动员的动作完成情况进行技术观察。

1.2.3 数理统计法

为更客观地分析E组裁判评分,运用SPSS23.0和Excel 2010软件对第十四届全运会女子体操成绩册数据进行统计与分析,并制作相关图表。

2 结果与分析

2.1 客观性检验

采用SPSS23.0统计软件中肯德尔和谐系数(Kendall's W)即一致性系数的检验,将第十四届全运会女子体操单项决赛中E组裁判员对前八名运动员动作完成质量评分的一致性进行处理,数据如表1所示。

表1 女子单项决赛E组裁判员评分的一致性系数(Kendall's W)检验分析

项目	<i>n</i>	Kendall's Wa	Chi-Square	<i>df</i>	<i>P</i>
高低杠	9	0.820	51.646	7	0.000
平衡木	9	0.898	56.563	7	0.000
跳马	9	0.797	107.580	15	0.000
自由体操	9	0.848	53.395	7	0.000

注: $P < 0.01$ 差异非常显著; $P < 0.05$ 差异显著; $P > 0.05$ 差异不显著。

从表1中女子单项决赛四个项目检验统计结果可以看出,女子体操单项决赛项目的高低杠、平衡木和自由体操三个项目的E组裁判员评分一致性系数都在0.8以上,跳马项目的E组裁判员评分一致性系数为0.797,稍低于0.8;渐进显著性 P 值均为0.000,小于0.01,从统计学角度来看属于高度相关,则可以认为进入各单项决赛的8名运动员的得分存在显著差异,E组裁判员的整体评分标准一致性较高。从个案样本一致性结果来看,女子平衡木中E

组裁判评分的一致性系数相对最高,即平衡木裁判评分客观性系数最高,自由体操裁判评分客观性次之,跳马裁判评分客观性相对稍低,这说明裁判员的评分误差相对于运动员个体之间现场水平发挥的差异较小,也就是评分结果对运动员的变差小,进而区分度比较高(跳马稍低)。从视频观察及成绩册评分数据来看,参加女子跳马决赛的8名运动员均需完成两跳动作(共16套),对两跳动作的搭配类型和完成情况进行列表统计与对比分析发现:女子跳马决赛中的难

度得分即 D 分在 4.6~6.0 之间,完成分即 E 分的差异对运动员的决赛成绩有着至关重要的影响,运动员的跳马完成情况与两跳动作的选择类型及难度使用情况密切相关,这对裁判员主观评价的客观水平影响较大。相比女子高低杠、平衡木和自由体操三个决赛项目的一致性结果来看,女子跳马项目 E 组裁判评分的一致性系数反映了该项目中 E 组裁判业务水平及对评分尺度的把握有待进一步加强,评分客观性仍有进一步提高的空间。

2.2 单因素方差分析

为进一步分析 E 组裁判员对运动员的评判标准是否一致,采用 SPSS23.0 统计软件中单因素 ANOVA 即方差分析,以运动员决赛成绩中的 E 分作为观测变量,裁判员作为控制变量,通过单因素方差分析方法,将第十四届全运会女子体操四个单项决赛中 E 组裁判员对各单项前八名运动员 E 分的影响进行了方差分析,采用多重比较检验法,结果如表 2 和表 3 所示。

表 2 女子单项决赛四个项目的方差齐性检验结果

项目	裁判评分数据			
	Levene 统计量	df1	df2	P
高低杠	0.405	8	63	0.913
平衡木	0.392	8	63	0.921
跳马	0.149	8	135	0.996
自由体操	0.359	8	63	0.938

表 2 的方差齐性检验结果表明,四个项目的 E 组裁判员评分方差齐性检验统计量的观测值分别为:0.405、0.392、0.149、0.359,对应的概率 P 值分别为:0.913、0.921、0.996、0.938,9 位裁判员在四个项目中的渐进显著性均大于显著性水平 0.05,显著性较强。因此接受原假设,认为它们的方差是齐性的,即不同项目的 E 组裁判员评分数据的总体方差不存在显著差异。

表 3 女子单项决赛四个项目的单因素方差分析结果

项目		SS	df	MS	F	P
高低杠	组间	0.399	8	0.050	0.314	0.958
	组内	10.011	63	0.159		
	总和	10.410	71			
平衡木	组间	0.277	8	0.035	0.070	1.000
	组内	31.303	63	0.497		
	总和	31.579	71			
跳马	组间	0.151	8	0.019	0.131	0.998
	组内	19.433	135	0.144		
	总和	19.584	143			
自由体操	组间	0.097	8	0.012	0.055	1.000
	组内	13.922	63	0.221		
	总和	14.019	71			

注: $P < 0.01$ 差异非常显著; $P < 0.05$ 差异显著; $P > 0.05$ 差异不显著。

从表 3 中的单项决赛四个项目的方差分析结果可以看出,如果仅考虑裁判员单个因素的影响,则运动员决赛成绩中的 E 分总变差中,四个项目中不同裁判的评分标准可解释的变差分别为 0.399、0.277、0.151、0.097,随机误差引起的变差分别为 10.011、31.303、19.433、13.922,他们的 F 统计量的观测值分

别为 0.314、0.070、0.131、0.055,所对应的概率 P 值分别为 0.958、1.000、0.998、1.000,在显著性水平 α 为 0.05 时,女子单项决赛中四个项目的 E 组裁判员评分标准均没有显著差异 ($P > 0.05$)。认为各项目的 E 组裁判员评分标准对运动员决赛成绩均没有产生显著影响,即裁判员对第十四届参赛运动员评分的

尺度把握总体较为一致,没有出现倾向性的偏高或偏低情况^[2]。

对 SPSS 方差分析中输出的分析数据,结合 Excel 2010 软件进行再次分析处理,其结果如图 1、图 2 所示。从图 1、图 2 四个项目的均值图和标准误差值图可以看出,个别裁判的 E 分平均值出现过高或过低的情况,相对于运动员的个体差异来说,个别裁判员的评分出现了较大的随机性,随机误差较大。

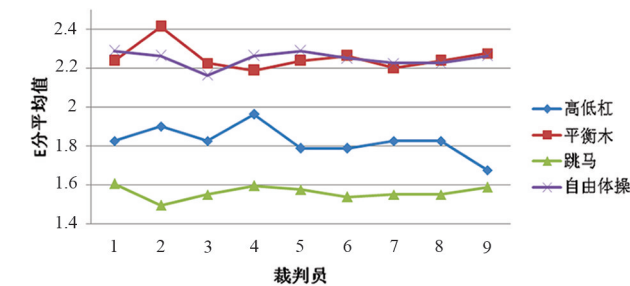


图 1 四个项目 E 分均值折线图

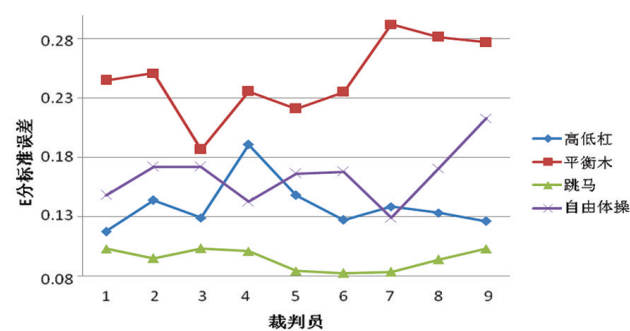


图 2 四个项目 E 分标准误差值折线图

造成这种结果的原因可能有两方面:一方面从客观角度看,运动员的运动水平较为接近,裁判员在评分时难以评判出高低;另一方面从主观角度来看,裁判员在评分过程中相对保守,为让自己的评分成为有效分,评分的波动性较小,可能会降低区分度。从整体来看,虽然裁判员对所有运动员的总体评分尺度一致,没有倾向性的偏高或偏低,但对于水平相近的运动员来讲,个别 E 组裁判员的评分随机误差较大,会出现偏高或偏低,成为无效分。综合多方面分析表明:个别裁判员评分的区分度有待进一步提高^[3]。

2.3 极差分析

采用 SPSS23.0 统计软件中的描述分析(单因素 ANOVA 即方差分析),并结合 Excel 2010 软件对第十四届全运会女子体操前八名单项决赛中的 E 组裁判员评分出现极差的频率,进行分析统计,结果如表 4

和图 3 所示:

表 4 女子单项决赛 E 组裁判员评分出现极差的频率

极差/分	高低杠	平衡木	跳马	自由体操	合计
0.0	0	0	0	0	0
0.1	1	0	2	0	3
0.2	0	0	4	0	4
0.3	3	1	4	4	12
0.4	2	1	5	2	10
0.5	0	2	1	1	4
0.6	1	0	0	1	2
0.7	0	3	0	0	3
0.8	0	1	0	0	1
1.3	1	0	0	0	1

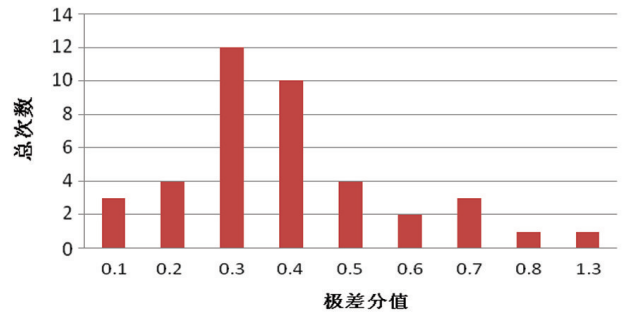


图 3 E 组裁判评分极差分值与出现总次数分布图

通过表 4 和图 3 对女子体操单项决赛 E 分的分析和统计可以看出,E 组裁判对运动员的评分极差值在 0.1~1.3 分之间波动,最小极差评分值为 0.1 分在高低杠和跳马两个项目中出现了 3 次,最大极差评分值为 1.3 分仅在高低杠决赛中出现了 1 次,极差值为 0.3 分和 0.4 分的这两个分值均在四个单项决赛中出现过,评分极差值中出现最多的极差值为 0.3 分,在女子单项决赛中共出现了 12 次。所有极差值的出现机率在裁判员的评分数据中大致成偏正态分布,符合统计学的要求。极差值的出现机率说明了裁判员个体间评分尺度存在差异,且存在一定的评分误差随机性。对于个别运动员的 E 分,裁判员的评分极差较大,例如在女子高低杠决赛项目的评分中,裁判的评分极差值为 1.3 分,说明裁判员对该运动员的评分存在争议,评判的尺度有一定的差异,从而使评分数据的客观性受到质疑。虽然裁判员的评判结果

出现显著的个体间差异,但是依据体操竞赛规则去掉最高分和最低分,仅出现一次差异显著的评分并不会影响 E 组整体裁判员的评分结果,这与数理统计中的一致性系数分析评价是一致的^[4-6]。

3 结论与建议

第十四届全运会体操比赛裁判绝大部分是由国家级及以上级别的裁判担任,并且 E 组裁判在赛前是通过抽签来决定裁判员的分工。综合以上客观性统计分析的结果说明:(1)从整体情况来看:第十四届全运会女子体操单项决赛中四个项目 E 组裁判员的评分客观一致性较高,裁判员对第十四届参赛运动员评分的尺度把握总体较为一致,裁判员评分的随机误差总体上控制较好,没有出现倾向性的偏高或偏低情况。一方面说明了第十四届全运会中女子体操 E 组裁判员的执裁水平相对较高,另一方面也说明了我国的体操裁判员培养和遴选工作达到了较高水平。(2)从四个项目的局部统计学分析来看:女子体操 E 组裁判员评分的客观一致性系数虽然较高,但通过评析 E 组裁判员评分客观性的一系列指标数据后发现,仍有个别裁判员评分的随机误差稍大,评分的区分度有待提高,需要进一步加强控制和降低随机误差;通过统计学分析结果可以看出:极差值和无效分的出现机率和出现位置,说明个别裁判员对评分尺度的把握不准确,使评分个体间差异及随机误差较大,评分数据客观性受到质疑,但是,依据体操竞赛规则去掉最高分和最低分之后,最终对运动员最后得分的可信度影响不大。

E 组裁判员作为竞技体操比赛中的重要组成部

分,不仅应严格遵守竞赛规则,确保比赛公平顺利进行,而且应加强个人业务能力,提高对竞赛规则理解的深度和执行力度。^[7]为了提高 E 组裁判员执法水平,进一步提高裁判员评分客观性,一方面,建议合理控制和处理评分误差,减少无效分的出现机率,缩小极差评分值的波动范围,进一步提高裁判员评分的客观一致性,推进体操裁判队伍建设^[8];另一方面,通过加强监控机制,提高业务培训,使裁判员的评分更加的公正、合理和科学,确保体操比赛结果更具客观性,进一步推动体操项目的健康持续发展。

参考文献

- [1] 国际体操联合会. 2017—2020 国际女子竞技体操评分规则[S]. 利斯:国际体操联合会,2016.
- [2] 赵晨琼,赵玉华. 索契冬奥会单板滑雪 U 型场地技巧决赛裁判员评分的客观性研究[J]. 哈尔滨体育学院学报, 2014,32(5):21-25.
- [3] 崔园园,郑幸红. 第 29 届奥运会女子体操单项完成分的客观性[J]. 体育科学研究,2010,14(2):71-73.
- [4] 徐亚楠. 女子竞技体操评分规则演变的奥运周期特征及发展趋势研究[D]. 天津:天津体育学院,2020.
- [5] 李明,王丰彩. 福建省青少年体操锦标赛 E 组评分客观性分析——以 2021 年男子体操单项决赛为例[J]. 体育科学研究,2022,26(5):52-56,87.
- [6] 谢杰. 东京奥运会男子双杠决赛成套动作编排分析[J]. 体育科技文献通报,2022,30(4):28-31.
- [7] 何俊. 东京奥运周期中国女子体操发展路径探索[J]. 体育科学,2017,37(12):14-22.
- [8] 徐瑞芳. 北京奥运会男子体操单杠决赛 E 组评分研究[J]. 体育科学研究,2009,13(3):50-51.

[责任编辑 江国平]