

三种综合评价(评分)方法在 体育评价中的比较研究^①

李 毅¹, 沈建国²

1. 中山大学 体育部, 广州 510275; 2. 浙江工商大学 体育部, 杭州 310118

摘要: 为将科学合理、简便易行的先进评价方法引入体育评价, 对 3 种体育综合评价方法进行了比较研究. 研究显示: 3 种综合评分方法均科学合理(信度高)、简便易行(效度好), 且对研究样本的计算结果基本趋向一致, 可以为体育实践中的教学、训练、竞赛、管理、健身、科研评价、对比、衡量、判断、分类、预测和决策的量化工作提供标准和调控依据.

关 键 词: 经验定性; 科学定量; 综合评价; 比较研究

中图分类号: G807

文献标志码: A

文章编号: 1000-5471(2017)08-0112-05

在学校体育的教学、训练、竞赛、管理、健身及科研实践中, 经常需要对所获取的相关数据(信息)进行对照比较、综合分析, 但因数据(信息)类别、项目、性质和计量单位的不同而不能直接进行对比和统计分析, 必须将这些数据(信息)转变成可以计量的分数形式方来完成, 为更好地贯彻“科学发展观”指导思想, 更快地实施“阳光体育”教育理念, 执行“健康第一”教学模式, 彻底摒除陈旧的“传习式”和“大锅饭式”体育教学, 启用科学的“以人为本”即以学生为主体的培养现代人(KAQ—知识、能力、素质相结合的复合型人才)的体育教学与训练方法, 学校体育改革已被提到重要日程. 为适应学校体育教育的改革, 迫切需一个全面、客观与合理量化的赋予综合评价(评分)体系^[1]. 多年来, 我国各级各类学校仍旧停留在经验定性评估基础上, 采用听、看、访、议的简单比较后, 以举手投票的方式进行综合评价, 难以精确定量判断其评价对象的优劣, 故科学的综合评价势在必行.

目前, 国内外通常使用的评价(评分)方法有十多种, 各具其优、缺点, 其中权重综合评分法(I)、模糊综合评分法(II)、多项综合均分法(III)等 3 种方法比较常用. 权重综合评分一般先运用主成分分析和层次分析法对相关因素进行实证分析, 其程序分 3 步: ①结合专家的经验, 运用主成分分析法筛选出相对重要的指标及客观权重; ②通过建立层次分析模型确定关键指标, 构造判断矩阵、确定指标主观权重; ③将主客观权重采用“加法”集成法计算出最终的集合权重. 模糊综合评价法(Fuzzy Comprehensive Evaluation, FCE)是一种基于模糊数学的综合评价方法^[1]. 它是根据模糊数学的隶属度理论把定性转化为定量评价, 用模糊数学对受到多种因素限制的事物或对象做出一个整体的评价, 具有结果清晰与系统性强的特点, 适合各种多层次、多因素不确定性问题的评判效果, 如方案的优选、重要项目的影响因素筛选等. 在实际应用中, 常常需要对受多种因素影响的事物进行评价, 由于同一事物具有多种属性, 受多种不确定因素的影响, 因此在评价的过程中, 必须对多个相关的因素作综合性考虑及全部评价, 这便是所谓的评价, 若评判过程涉及模糊因素, 称为模糊综合评价法. 多项综合均分不计权重, 只依据单个评委的评分进行汇总综合评分, 为了剔除评委之间出现的分歧, 基于习惯做法, 去掉一个最低分、去掉一个最高分再求得平均分^[2].

① 收稿日期: 2017-03-23

作者简介: 李 毅(1972-), 男, 湖南湘潭人, 讲师, 硕士研究生, 主要从事体育教学与运动训练.

从科学定量评价理论得知,评价即是对事物(研究对象)价值的界定,它必须用评分方法来完成.而综合评分又是一种能全面、客观与定量反映评价对象真实全貌的主要手段,其遵循的原则为:①综合评价要如实反映客观事实,准确计量、明晰分辨出被评对象个体间的本质差异,避免评分偏高和偏低的现象;②评价给分幅度与水平难度相适应,也就是说难度越大,付出的劳动越艰苦则评分越高,反之亦然;③采用的评分方法应科学合理、简便易行,可以为各级各类学校广泛运用.本文以同一例题对 3 种综合评价评分方法进行比较研究,拟为体育科研方法的采用提供参考.

2 应 用

2.1 权重综合评分法(简称方法 I)

例:某高校对体育老师教学质量水平进行量化评价,以作为绩效工资、职称评定、奖金定级等的依据.这是一个典型的定量分析的综合评价,用一个特定的统一评价模式对在校每位体育老师进行综合打分,这无疑是一个科学量化的好措施.

2.1.1 步骤操作

设计一份体育教学质量综合评价表,见表 1.

表 1 体育教学质量综合评价细则表

因 素		等 级 与 细 则					分项 权重	总体 权重	
		1 栏	2 栏	3 栏	4 栏	5 栏			
		优 95 分	良 85 分	中 75 分	差 65 分	劣 55 分			
		(90~100)	(80~90)	(70~80)	(60~70)	(50~60)			
A	课 前 准 备	1. 目的任务	很明确、严谨	明确、严谨	较明确	不明确	很不明确	0.3	0.2
	2. 备课教案	很认真、有条理	认真、条理	较认真	不认真	很不认真	0.4		
	3. 场地器材	布局很合理	合理	较合理	不合理	很不合理	0.3		
B	课 的 组 织	4. 讲解	很生动、形象	生动、形象	较生动	不生动	很不生动	0.4	0.4
	5. 示范	很准确、熟练	准确、熟练	较准确	不准确	很不准确	0.3		
	6. 教法、教态	很新颖、自若	新颖、自若	较新颖	不新颖	很不新颖	0.3		
C	课 后 效 应	7. 任务完成情况	很好	好	较好	不好	很不好	0.4	0.25
	8. 密度、运动量	很科学、合理	科学、合理	较科学	不科学	很不科学	0.3		
	9. 学生反映	很好、欢迎	好	较好	不好	很不好	0.3		
D	教 育 功 能	10. 品德教育	收效很好	好	较好	不好	很不好	0.3	0.15
	11. 体质、安全教育	很到位	到位	较到位	不到位	很不到位	0.3		
	12. 理论知识教育	很结合实际	结合实际	较结合实际	不实际	很不实际	0.4		
分值与权重来源		1. 表中优等 95 分为 $(90+100) \div 2$; 良为 $(80+90) \div 2$; ……劣为 $(50+60) \div 2$ 求得. 2. 分项与总体权重由专家投票通过“特尔菲”法求得.							

2.1.2 步骤操作

请 30 名评审人员(由评价专家、主管教学领导、同行与学生代表组成)观摩被评教师上课,并在表 1 各栏划记“√”(每一因素每一等级栏只能划记一次),如 A 组因素之 1 为本课目的任务栏,若认为该教师很明确严谨则在等级与细则栏优 95 分处打“√”即可,依此类推.

2.1.3 步骤操作

将各评审人员的评价表(表 1)进行统计,按各因素、各栏各等级划记“√”汇总填入体育教学质量综合评价统计表(表 2).

权重综合评分法(简称方法 I)计算过程:

A 因素
(课前准备)

1. $\frac{5 \times 95 + 15 \times 85 + 10 \times 75}{30} \times 0.3 = 25(\text{分})$

2. $\frac{2 \times 95 + 18 \times 85 + 10 \times 75}{30} \times 0.4 = 32.93(\text{分})$

3. $\frac{20 \times 85 + 10 \times 75}{30} \times 0.3 = 24.5(\text{分})$

注:未予权重分配的纯得分

$\frac{5 \times 95 + 15 \times 85 + 10 \times 75}{30} = 83.33(\text{分})$

A 的总体得分 $(25+32.93+24.5)\times 0.2=16.49$ (分)

同理,B 因素(课中组织)总体得分: $(32.67+25+25.3)\times 0.4=33.19$ (分)

C 因素(课后效果)总体得分: $(34+24+24.5)\times 0.25=20.63$ (分)

D 因素(教育功能)总体得分: $(24.5+23.7+32.13)\times 0.15=12.10$ (分)

则 I 法最后总得分:I 总分 $=16.49+33.19+20.63+12.10=82.41$ (分)

表 2 体育教学质量综合评价统计细则表

因 素		等级划记数量统计					分项 纯得分	分项 权重	分项 得分	总体 权重	总体 得分
		优 95 分	良 85 分	中 75 分	差 65 分	劣 55 分					
A	1. 目的任务	5	15	10	0	0	83.33	0.3	25	0.2	16.49
	2. 备课教案	2	18	10	0	0	82.33	0.4	32.93		
	3. 场地器材	0	20	10	0	0	81.67	0.3	24.5		
B	4. 讲解	4	12	14	0	0	81.67	0.4	32.67	0.4	33.19
	5. 示范	4	17	9	0	0	83.33	0.3	25		
	6. 教法、教态	8	12	10	0	0	84.33	0.3	25.3		
C	7. 任务完成情况	10	10	10	0	0	85	0.4	34	0.25	20.63
	8. 密度、运动量	0	15	15	0	0	80	0.3	24		
	9. 学生反映	6	12	8	4	0	81.67	0.3	24.5		
D	10. 品德教育	3	17	10	0	0	82.67	0.32	4.8	0.15	12.10
	11. 体质、安全教育	0	14	14	2	0	79	0.3	23.7		
	12. 理论知识教育	2	15	10	3	0	80.33	0.4	32.13		
总 分										82.41	

2.2 模糊综合评分法(简称方法 II)

模糊综合评分法的等级评判,与方法 I 相同.

2.2.1 确定评价因素评语等级

同方法 I,确定不同级别的评语等级.

2.2.2 划分等级和权重

同方法 I,使用同样原班评审专家组成员.

2.2.3 制订模糊综合评分计算表(表 3).

模糊综合评分法(II)法:
采用模糊矩阵计算,求出 A,B,C,D 各因素评判集.
条件:

A. $(0.3 \quad 0.4 \quad 0.3) \begin{pmatrix} 0.17 & 0.5 & 0.33 & 0 & 0 \\ 0.07 & 0.6 & 0.33 & 0 & 0 \\ 0 & 0.67 & 0.33 & 0 & 0 \end{pmatrix}$

按矩阵计算,横行与纵列相圈乘,依先取小、再取大模式处理:

$(0.3 \wedge 0.17) \vee (0.4 \wedge 0.07) \vee (0.3 \wedge 0)$

$(0.3 \wedge 0.5) \vee (0.4 \wedge 0.6) \vee (0.3 \wedge 0.67)$

$(0.3 \wedge 0.33) \vee (0.4 \wedge 0.33) \vee (0.3 \wedge 0.33)$

$(0.3 \wedge 0) \vee (0.4 \wedge 0) \vee (0.3 \wedge 0)$

$(0.3 \wedge 0) \vee (0.4 \wedge 0) \vee (0.3 \wedge 0)$

}

$(0.17 \quad 0.4 \quad 0.33 \quad 0 \quad 0)$

将 $(0.17 \quad 0.4 \quad 0.33 \quad 0 \quad 0)$ 归一化后得 $(0.19 \quad 0.44 \quad 0.37 \quad 0 \quad 0)$

表 3 模糊综合评分计算细则表

因 素		等级划记数量统计					总分
		优 95 分	良 85 分	中 75 分	差 65 分	劣 55 分	
A	1. 目的任务	5(0.17)	15(0.5)	10(0.33)	0	0	81.8
	2. 备课教案	2(0.7)	18(0.6)	10(0.33)	0	0	
	3. 场地器材	0	20(0.67)	10(0.33)	0	0	
B	4. 讲解	4(0.13)	12(0.4)	14(0.47)	0	0	
	5. 示范	4(0.13)	17(0.57)	9(0.3)	0	0	
	6. 教法、教态	8(0.27)	12(0.4)	10(0.33)	0	0	
C	7. 任务完成情况	10(0.33)	10(0.33)	10(0.33)	0	0	
	8. 密度、运动量	0	15(0.5)	15(0.5)	0	0	
	9. 学生反映	6(0.2)	12(0.4)	80(0.27)	4(0.13)	0	
D	10. 品德教育	3(0.1)	17(0.57)	10(0.5)	0	0	
	11. 体质、安全教育	0	14(0.4)	14(0.47)	2(0.07)	0	
	12. 理论知识教育	2(0.07)	15(0.5)	10(0.5)	3(0.1)	0	

注：表 3 中，A 因素 1 优 95 分栏因有 5 名专家投票赞成，共有 30 名专家，则 $5\div30=0.17$ 记入 5(0.17)，其余类推。同理：

$$B. \begin{pmatrix} 0.13 & 0.4 & 0.47 & 0 & 0 \\ 0.13 & 0.57 & 0.3 & 0 & 0 \\ 0.27 & 0.4 & 0.33 & 0 & 0 \end{pmatrix} = (0.27 \quad 0.4 \quad 0.4 \quad 0 \quad 0) \text{ 归一化后为 } (0.25 \quad 0.37 \quad 0.37 \quad 0 \quad 0)$$

$$C. \begin{pmatrix} 0.33 & 0.33 & 0.33 & 0 & 0 \\ 0 & 0.15 & 0.15 & 0 & 0 \\ 0.2 & 0.4 & 0.27 & 0.13 & 0 \end{pmatrix} = (0.33 \quad 0.33 \quad 0.33 \quad 0.13 \quad 0) \text{ 归一化后为 } (0.29 \quad 0.29 \quad 0.29 \quad 0.12 \quad 0)$$

$$D. \begin{pmatrix} 0.1 & 0.57 & 0.33 & 0 & 0 \\ 0 & 0.47 & 0.47 & 0.07 & 0 \\ 0.07 & 0.5 & 0.33 & 0.1 & 0 \end{pmatrix} = (0.1 \quad 0.4 \quad 0.33 \quad 0.1 \quad 0) \text{ 归一化后为 } (0.11 \quad 0.43 \quad 0.35 \quad 0.11 \quad 0)$$

接着求综合决断集：

$$(0.2 \quad 0.4 \quad 0.25 \quad 0.15) \begin{pmatrix} 0.19 & 0.44 & 0.37 & 0 & 0 \\ 0.25 & 0.37 & 0.37 & 0 & 0 \\ 0.29 & 0.29 & 0.29 & 0.12 & 0 \\ 0.11 & 0.43 & 0.35 & 0.11 & 0 \end{pmatrix} = (0.25 \quad 0.37 \quad 0.37 \quad 0.12 \quad 0)$$

归一化后为(0.23 0.33 0.33 0.11 0)

II 总分=0.23×95+0.33×85+0.33×75+0.11×65=81.8(分)

2.3 多项综合均分法(简称Ⅲ法)

本方法不计权重，也不必作矩阵计算，只要将 A,B,C,D 因素的 12 项指标得分相加，取其平均值。

即：Ⅲ 总分 =
$$\frac{83.33+82.33+81.67+81.67+83.33+84.33+85+80+81.67+82.67+79+80.33}{12}$$

82.11(分)

3 分 析

由上可见，Ⅰ，Ⅱ，Ⅲ 3 种综合评分法的总分分别为Ⅰ 总分=82.41 分，Ⅱ 总分=81.8 分，Ⅲ 总分=82.11 分。其中最大差值 0.61 分，最小差值 0.3 分，3 种方法结果基本一致。若按实际分值比较，方法Ⅰ评

分值最高(82.41 分)。如何科学、客观地评价体育教师的教学工作,为体育教师提供具体、准确的反馈信息,帮助教师改进教学工作,提高自身专业素质具有非常重要的现实意义^[3]。对任何复杂的评价体系而言,各个指标的权重分配是其评价过程重要程度的反映,是指标相对重要性的一种综合度量,权重确定的合理与否,直接关系到整个体系的科学性。近 10 年关于我国体育教师评价的研究文献,主要使用 4 类指标权重确定方法,即层次分析法、特尔菲法、自拟权重法和模糊数学法。本研究中,方法Ⅱ评分值最低(81.8 分),虽采用了模糊数学方法对难以精确量化的指标进行合理定量,但在计算处理时丧失了部分信息(只取大、取小值计算中间值被忽略);方法Ⅲ评分值居中,各数据的信息,向平均值靠拢,且该方法较之方法Ⅰ、方法Ⅱ更简便易操作,其结果也与之相接近^[4],但它未涉及因素所应具有的重要性不同的权重,属各因素平权处理。故 3 种方法各有优缺点。

4 结 论

1) 方法Ⅰ、方法Ⅱ、方法Ⅲ 3 种综合评分方法均科学合理(信度高)、简便易行(效度好),且对研究样本的计算结果基本趋于一致。

2) 此 3 种方法可为体育实践中教学、训练、竞赛、管理、健身、科研的评价、对比、衡量、判断、分类、预测和决策的量化工作提供标准和调控依据,具有推广和应用价值,但具体应用时要注意详加考察,综合实施,以求准确。

参考文献:

- [1] 黄明教,陈志强,丁红翎,等. 体育试验设计与科学量化方法 [M]. 北京: 高等教育出版社, 2002.
- [2] 陈 嵘. 体育研究中多因素综合分析方法与应用 [M]. 北京: 人民日报出版社, 2005.
- [3] 靳玉乐,倪 哲. 中小学综合素质评价相关制度的建立及其运用 [J]. 西南师范大学学报(自然科学版), 2014, 39(2): 147—153.
- [4] 刘 成,刘兰娟,司虎克. 体育竞争情报分析要素和分析框架的理论与实证研究 [J]. 西南师范大学学报(自然科学版), 2014, 39(2): 120—126.

Comparative Research on Three Comprehensive Evaluation (Grading) Methods in Evaluation of Physical Education

LI Yi¹, SHEN Jian-guo²

1. The P. E. Department of SUN YAT-SEN University, Guangzhou 510275, China;

2. P. E. Department of Zhejiang Gongshang University, Hangzhou 310118, China

Abstract: Various methods, scientific and reasonable, simple and feasible, have been introduced into the evaluation of sports, with three sports comprehensive evaluation method for physical education stated in this article to lead to the discard of the disadvantage of assessments that rely on the daily experience, and thus rise to the height of scientific quantitative evaluation, moreover, accomplishing the purpose of macro-control and micro-analysis. They also have great profound strategic significance for better guiding the evaluation of physical education practice, deepening the education reform, and cultivating KAQ (knowledge, ability, quality) talents who can participate in the international competition. The results show that three comprehensive evaluation methods are scientific and reasonable, high reliability, easy (validity), and the basic trend is consistent, the calculation results of the sample provides the basis of the standard and regulation for quantitative work in physical education, mass sports, training, competition, management, health, scientific research evaluation, contrast, measure, judgment, classification, prediction and decision.

Key words: experience factors in qualitative assessment; scientific quantitative analysis; comprehensive evaluation; comparative research

责任编辑 胡 杨