



体育测量 与评价

人民体育出版社

体育测量与评价

编译 陈骏良

审校 薛济英

人民体育出版社

此书得以出版，除了组织上的支持与鼓励，编辑们的帮助之外，在此我要特别感谢薛济英教授为此而付出的心血。广州体院杨迺军同志替我编译了第二十五章和第二十六章的内容；暨南大学宋兆鸿同志审阅了第四章；华南师大陈斌同志审阅了第二、第三章，在此一并致以衷心的感谢。

体 育 测 量 与 评 价

陈骏良 编译 薛济英 审校

人 民 体 育 出 版 社

新华书店总店北京发行所发行 北京印刷二厂印刷

850×1168毫米 1/32 400千字 印张22

1991年3月第1版 1991年3月第1次印刷

印数：1—2000册

ISBN 7-5009-0670-6/G·640 定价：10.00元

责任编辑：卢 锋

目 录

前言	(1)
第一章 绪论	(5)
测量	(6)
评价	(7)
测量和评价的职能	(11)
体育任务	(12)
教学体系的模式	(14)
教育体制对测量与评价的影响	(16)
第二章 测量与评价中的基本统计手段	(18)
分析观测数据的要素	(18)
观测数据的分类	(18)
统计方法的分类	(22)
正态分布曲线	(22)
正态曲线在测量与评价中的主要用途	(24)
频数分布	(27)
集中趋势量数	(31)
变异性量数	(35)
标准分	(37)
相关	(42)
积差相关系数	(43)
预测公式	(49)
第三章 测量与评价中的高级数理方法	(52)
方差分析	(52)

差异观测数的可靠性·····	(54)
成绩提高的评价问题·····	(55)
其它相关法·····	(59)
等级相关(又称序相关)·····	(59)
点双列相关系数(Paint biserial)·····	(61)
ϕ 系数(FHI)·····	(63)
联列系数(Contingency)·····	(65)
偏相关系数·····	(67)
偏标准差·····	(71)
复相关系数·····	(72)
识别运动基本因素的程序·····	(77)
第四章 常模参照性测验的有效性 ·····	(83)
测验有效性定义·····	(83)
内容有效性·····	(84)
效标有效性(又称相合有效性)·····	(92)
预测有效性·····	(94)
结构有效性·····	(103)
体育中使用结构有效性的范例·····	(109)
有效性系数的大小·····	(111)
运动技术测量中的无效源泉·····	(111)
第五章 常模参照性测验的可靠性 ·····	(113)
可靠性定义·····	(114)
可靠性和有效性的关系·····	(114)
可靠性系数的分类·····	(115)
测量变异性的根源·····	(115)
运动成绩测验中估价可靠性的统计方法·····	(122)
影响测验可靠性系数的因素·····	(138)
可靠性系数的大小·····	(140)
估价可靠性的非参数方法·····	(141)
第六章 效标参照性测量 ·····	(144)

评价学习的掌握程度·····	(145)
效标参照性测验的定义·····	(147)
效标参照性测验的用途·····	(148)
有效性·····	(148)
可靠性·····	(154)
第七章 体育评价中的量表和常模·····	(158)
原始观测数和导出观测数·····	(158)
常模·····	(159)
如何选择量表·····	(174)
选用常模的准则·····	(174)
地区性常模和全国性常模·····	(175)
第八章 运动成绩测验的编制·····	(177)
创编单项技术测验·····	(178)
测验的设计·····	(179)
有效性和可靠性·····	(184)
制定常模·····	(185)
成套技术的测验·····	(185)
竞技能力的测验·····	(186)
第九章 笔试的编制·····	(188)
测验说明·····	(188)
测验的实施·····	(189)
选择性试题·····	(190)
替换选择题·····	(190)
选答题·····	(193)
配对选择题·····	(196)
填空测验题·····	(197)
短答测验题·····	(197)
论文式问题·····	(198)
权重各个测验成分·····	(201)
猜题倒扣·····	(202)

知识测验的评价·····	(203)
测验题目的分析·····	(204)
知识测验中存在的问题·····	(209)
知识测验和研究中的成果和结论·····	(209)
第十章 测验的选择、编制、组织与实施·····	(212)
测验的形式·····	(212)
测验的属性与要求·····	(212)
测验实施的三个步骤·····	(218)
第十一章 体育评价的形式和方法·····	(220)
终结期评价·····	(222)
形成期评价·····	(224)
形成期评价的应用·····	(226)
教学大纲的评价·····	(231)
课程的评价·····	(235)
教学效果的评价·····	(238)
评价方式的应用与发展·····	(247)
体育活动课中的教师教学效果评价·····	(249)
第十二章 体育成绩的评价和定级·····	(251)
评价·····	(252)
定级·····	(254)
定级方法·····	(258)
最后定级·····	(266)
未完成全部测验项目学生的定级问题·····	(272)
评价中的计算机运用·····	(272)
第十三章 发育、营养、体型及身体成分的测量和评价·····	(274)
分级指数·····	(274)
生长发育和营养状况测量·····	(276)
梅雷迪斯身高一体重量表·····	(280)
威策尔网格图表·····	(280)

韦尔斯利体重预测法.....	(281)
彼尔迪西营养指数.....	(285)
身体成分与体型测量.....	(285)
肥胖问题.....	(285)
体型.....	(295)
第十四章 身体姿势及结构的测量与评价.....	(306)
姿势测验的用途.....	(306)
使用评价量表的姿势测验.....	(307)
使用工具的客观性姿势测验.....	(311)
足部结构测量.....	(317)
姿势测量中存在的问题.....	(320)
姿势测量和研究中的成果和结论.....	(322)
第十五章 心肺机能测量和评价.....	(324)
需氧工作能力.....	(325)
实验室方法.....	(325)
非实验室方法.....	(331)
适用于群体测量的其它心肺机能测量.....	(359)
心血管机能测验中存在的问题.....	(365)
心血管机能测验的研究成果概述.....	(366)
第十六章 柔韧性测量与评价.....	(369)
柔韧性测验的用途.....	(370)
实用柔韧性测验.....	(370)
柔韧性测量中存在的问题.....	(385)
柔韧性测量和研究中的成果和结论.....	(385)
第十七章 力量测量与评价.....	(387)
力量测验的用途.....	(388)
相对力量的实用测验.....	(388)
等张性力量测验.....	(388)
弹簧秤力量测验.....	(399)
等长性力量测验.....	(404)

等动性力量测验·····	(405)
力量指数和身体素质指数的评价·····	(405)
身体素质指数的测量与评价·····	(406)
力量测量中存在的问题·····	(429)
力量测量和研究中的成果与结论·····	(430)
第十八章 肌肉耐力测量与评价 ·····	(433)
肌肉耐力的三种测验方式·····	(433)
肌肉耐力测验的用途·····	(434)
肌肉耐力的实用测验·····	(434)
肌肉耐力测量中存在的问题·····	(453)
肌肉耐力测量和研究中的成果和结论·····	(454)
第十九章 功率(爆发力)的测量和评价 ·····	(457)
功率测验的用途·····	(458)
运动性功率的实用测验·····	(458)
做功性功率测验·····	(465)
功率测验中存在的问题·····	(470)
功率测量和研究中的结论和成果·····	(470)
第二十章 灵敏性测量与评价 ·····	(472)
灵敏性测验的用途·····	(472)
灵敏性的实用测验方法·····	(472)
灵敏性测验中存在的问题·····	(483)
灵敏性测量中的研究成果与结论·····	(484)
第二十一章 速度和反应的测量与评价 ·····	(485)
速度和反应测验的用途·····	(487)
手的反应测验·····	(487)
足的反应测验·····	(489)
动作速度测验·····	(490)
四秒钟冲刺跑测验·····	(490)
六秒钟冲刺跑测验·····	(492)
50码跑测验·····	(492)

“选择—反应—动作”测验·····	(495)
反应和速度测量中存在的问题·····	(498)
速度和反应测验及研究中的成果与结论·····	(499)
第二十二章 平衡及韵律测量与评价 ·····	(502)
平衡测验的用途·····	(502)
静力性平衡的实用测验·····	(502)
动力性平衡的实用测验·····	(508)
平衡测量中存在的问题·····	(517)
平衡测量和研究中的一些成果与结论·····	(518)
韵律测量·····	(519)
韵律测量中存在的问题·····	(521)
韵律测量中的研究成果与结论·····	(521)
第二十三章 知觉性运动与动觉测量及评价 ·····	(523)
动觉性知觉测量·····	(525)
成套知觉性运动测验·····	(533)
知觉性运动与动觉测验中存在的问题·····	(539)
知觉性运动测量中的研究成果与结论·····	(541)
第二十四章 社交、态度测量与评价 ·····	(543)
社交测验的用途·····	(543)
社交因素的实用测量·····	(544)
社交行为和顺应的量表·····	(544)
人类关系的认可测量·····	(545)
与众人相处中领导能力和才干的测量·····	(547)
社交测量中存在的问题·····	(549)
社交测量和研究中的成果与结论·····	(550)
态度测量和评价·····	(550)
态度测验的用途·····	(551)
态度评分量表·····	(551)
体育态度的普通量表·····	(553)
对身体活动的态度量表·····	(558)

运动道德的测量.....	(561)
态度测量中存在的问题.....	(562)
态度测量和研究的成果与结论.....	(563)
第二十五章 运动技术测量与评价.....	(564)
运动技术测验的用途.....	(565)
实用运动技术测验.....	(565)
羽毛球.....	(566)
篮球.....	(574)
排球.....	(596)
足球.....	(602)
体操.....	(603)
游泳.....	(612)
射箭.....	(613)
墙球.....	(614)
网球.....	(617)
棒、垒球.....	(620)
运动技术测验中存在的问题.....	(622)
运动技术测验的研究成果和结论.....	(626)
第二十六章 运动成绩成套测验的设计及评比.....	(628)
成套运动成绩测验的用途.....	(629)
巴罗运动能力测验.....	(629)
斯科特运动能力测验.....	(635)
拉奇华四、五、六年级男女生运动成绩测验.....	(644)
美国青少年素质测验.....	(648)
约翰逊室内运动成绩测验.....	(649)
跳、攀、跑测验.....	(655)
女子(竞技运动)体能测验.....	(661)
运动成绩测验中存在的问题.....	(664)
运动成绩测验和研究的结果或结论.....	(665)
第二十七章 其它测量.....	(639)

麦克乐运动潜力测验.....	(669)
运动可塑性测量.....	(670)
测量结果再现性的估价.....	(678)
克劳斯-韦伯测验.....	(680)
疲劳的测量与判断.....	(682)
附录一	(685)
附表1 F值表(单侧检验, 方差分析用) $\alpha = 0.05$	(685)
附表2 F值表 $p = 0.01$	(688)
附录二 测量单位的英制、公制转换表	(691)
参考文献	(693)

前 言

体育测量与评价是研究体育的教学、训练、科研、管理等范畴中如何做到“有数”的一门学科。我国有句成语，叫做“心中有数”。这个“数”除能表明事物的量外，还能推知其质。

体育测量学起源于美国的教育测量学，它是教育测量学的一个分支。十八世纪中叶，教育领域首先应用了测验与测量。从1860年起，克伦姆威尔（Cromwell）和希契科克（Hitchcock）分别由简单的人体测量着手，研究了学龄儿童的年龄与生长发育方面的问题，首先将测验与测量引入了美国体育。嗣后菲利普斯（Phillips）等人从事人体测量学的研究长达五十余年之久。与此同时，又有萨金特（Sargent）等人先后将力量、机能、形态等研究引进了体育测验与测量的范畴。此外，1885年在美国“健康、体育与娱乐协会”的首届会议上，将测验的一致性和使用常模等问题引入了议程。这就是体育测量学的开拓阶段。

到了本世纪二十年代，可以认为体育测量学开始逐渐由萌芽期进入了现代发展阶段。根据体育对学生的全面发展所作的贡献，把它作为教育的一个组成部分得到了教育界的承认。体育教师在教学中重视启发学生的专业兴趣，引导学生不仅仅局限于关心对教材和技能的掌握。美国体育界许多著名学者，如布雷斯（D. Brace）、麦克乐（C. H. McCloy）、寇瑞顿（T. C. Cureton）、克拉克（H. H. Clarke）及其后继者不仅针对人体素质（包括力量、爆发力、肌肉耐力、心肺耐力、速度及反应时、灵敏性、协调性、柔韧性、平衡以及身体适应性），心肺机能，形态姿势，一

般运动能力与专项运动能力等等属性作了如何测验、如何评价的大量研究，而且在情感、行为及体育知识等测验和评价方面也作了相当深入的研究。其中包括将描述性和推断性数理统计方法应用于体育测量的研究，从而为使用复杂的教育测量技术奠定了基础。

在此，值得一提的是对体育测量学作出很大贡献的具有权威性的体育学术刊物《研究季刊》(Research Quarterly)，它于1930年创刊以来，从未间断地刊登世界各国有关体育及运动方面的测量及评价的学术论文。

在教育测量中，效标参照性测量和常模参照性测量两个新概念的产生是教育测量学发展的一个里程碑。从此测量学逐步结合评价而不断发展，使教育测量学更趋完善。作为教育的一部分的体育测量学也逐渐发展成测量与评价的两个主要内容。尤其是近七、八年来，体育测量学的理论体系进入了更为完善的飞跃发展阶段。具体表现为测量与评价两个内容要自始至终紧密结合，对测量效标或常模的要求以及测量有效性、可靠性的检验等均更加严格、更需符合实际，对测量结果的评价方法也更趋精确、合理。

体育测量与评价从开始研究人体测量及机体各种具体属性发展成为今天的研究体育教师、教练员如何带着持续改进教学与训练的观点，为达到预定目标去主导教学与训练。

我国解放前曾设此课程，但内容是十分肤浅的。尤其正值此学科飞跃发展的近三十多年间，我国却中断了对它的教学和研 究。其严重后果，绝非仅仅使我国广大体育工作者对其内容越加变得茫然所知，更重要的是对于我国体育教学、训练中如何克服盲目性，进行有效改革；对于科学研究中如何提高设计上的合理性，增强方法上的严谨性和分析论证上的说服力；对行政管理工作中如何根据可靠的客观依据，对所管辖工作给予合理评价、提出改进目标等等方面，都已造成极为不利的影响。今天，历史和时代都要求我们，必须抓紧时间，奋起直追，迎头赶上。

本书以介绍测量与评价相结合这一崭新概念为主导的体育测量学完整理论体系为主，以有选择的范例形式介绍其实践内容为辅。针对我国当前情况，这个编译本的指导思想既有现实意义，又具深远的战略观点。理由是，不论从自然或人为条件看，我国与其它国家之间都存在着种种差异，因而国外的各种具体测量内容与方法，尤其是具体的评价标准，多数是不适于我们直接引用的。此书范例能从思路上给我们以借鉴和启发。至于其理论的完整体系则不然，它是放之四海而皆准的，是能指导实践的武器。

编者以体育测量学理论体系为主导的编写方针，同当代教育测量学专家布朗（Brown）教授的观点是不谋而合的。布朗教授在他1983年出版的《教育与心理测量学》的论著中这样写道：“妨碍现代测验得以普遍应用和健康发展的主要原因在于对测验的误用，从而造成混乱，而误用的原因在于对测量与评价的逻辑缺乏理解……所谓逻辑乃指当前教育与心理测量各个体系据以建立的各项基本原理。为此，必须使读者回答如下问题，例如：为什么要测量各种心理（或教育）特征？有哪些特征能加以测量？为什么测验必须可靠、有效和标准化？在测量成就、各种能力和性格特征时，作出了哪些假设？这些假设对测验的性质和得分的解释与使用将会带来什么意义？考虑这些‘为什么’势必会带来各种‘怎么样’的问题——也就是那些涉及编制测验和评价测验结果的方法问题。

“本书还着重阐明测验的应用方法，着重说明各种测验如何应用于各项教育决策和就业决策的制订，以及如何促进个人的发展。着重于这些方面，不仅对测验使用者而且对测验编制者和测验研究人员都同样是需要，因为他们都同样必须充分理解测量过程的性质，以及这一过程所依据的推理和各种假设。”

布朗教授的这一权威性论点，对目前我国体育研究的实践，应该说具有普遍的指导意义。如在我国全国范围内的“体质”调研工作中，教学、训练现行的考核内容、方法以及科学研究工作的设计、论证，资料搜集等程序，确实存在一些测验误用问

题，对测验的有效性、可靠性和客观性缺乏检验，对测验结果的评价方法的科学依据不足等现象说明，实践迫切需要这样的理论来指导。

总之，这本书为在我国重建体育测量学这一课程提供了良好的物质基础。让我们群策群力，在学习和借鉴国外先进经验基础上，从体育实践的实际需要出发，通过认真、艰苦的努力，尽快创建成适应于当前国情及未来需要的体育测量学的科学体系，为促进我国四化建设进程，为有效地进一步推动我国体育事业的普及和提高，作出应有贡献。

北京体院薛济英教授

第一章 绪 论

体育测量和评价，是为制订体育教学大纲提供客观依据的科学基础。不言而喻，它们应对教学大纲的制订和检验发挥重要作用。体育测量和评价着重于研究青少年从事体育运动过程中的智力、能力、潜力，以及他们所应具备的条件；它是为考察、检验教学与训练效果服务的。我们知道，每个学生的个性特征和属性是通过他们的智力、能力、潜力、欲望以及环境影响而形成的兴趣等所表现出来的。而所有这些因素在教学过程中必将带来引人注目的差异，通过测量也可以将它们反映出来。因此，教师要从了解学生当前可以做什么，并熟悉其具体条件。因而也就可以展望学生将来能获得一些怎样的成就。

由此可见，测量学的根本目的是为了改善教学与训练过程。测量是教学与训练的重要组成部分，两者不仅相互联系在一起，而且必须使教学与训练在体育测量学的科学监督下有计划地进行。换句话说，体育测量学是研究教师如何抱着持续改进教育的观点，为达到自己的预定目标去从事教学与训练的一门学科。

在通常情况下，人们往往把测验结果当作判断事物的基础。这可以说是带有普遍性的一种倾向。而这样一些判断，往往又影响乃至决定着其中每个人的命运，如学生的升级、大学的录取、征兵入伍是否够格等等。然而遗憾的是，有些人在做出这种判断的同时，却往往忽视了其测验结果的质量以及对它的评价。

正是在这样的情况下，他们尽管可以广泛使用测量和评价这两个术语，但实际上他们并非真正了解它们的含义。测量包括对