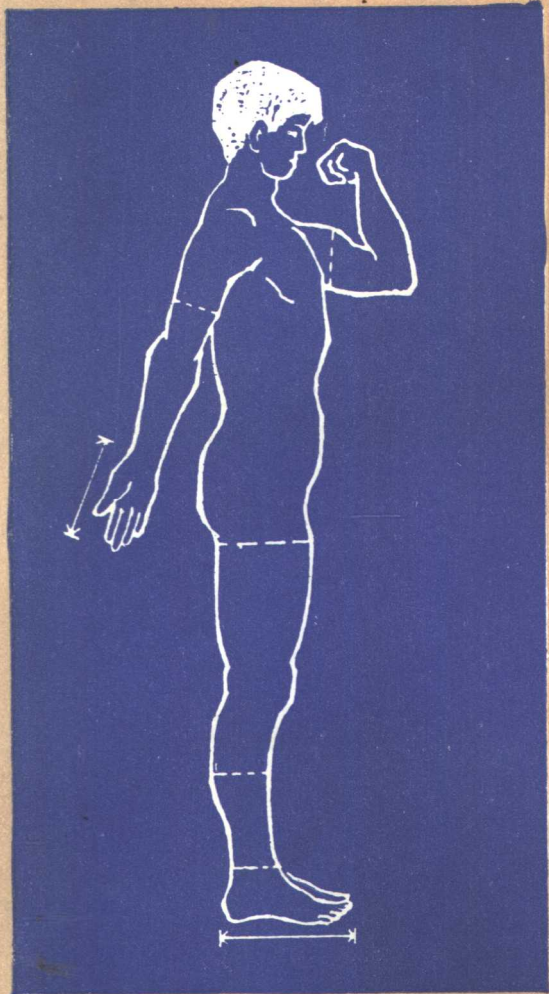


《体育测量学基础》编写组

TIYU
CELIANGXUE
JICHU



体育
测量学
基础

高等教育出版社

体育测量学基础

体育测量学基础编写组

高等教育出版社

内 容 提 要

本书介绍了体育测量学中基本的测量理论和评价原理,并结合我国体育教学和锻炼的实践,介绍了对各种能力、素质、机能、形态、情感行为、动觉、知觉 and 知识等常用的测量方法及评价标准。此书除了作为体育专业通用教材外,还可供体育教师、教练员和其他体育工作者参考。

(京) 112 号

体育测量学基础

《体育测量学基础》编写组

高等教育出版社出版
新华书店北京发行所发行
文字六〇三厂印装

开本 850×1168 1/32 印张 7.75 字数 190 000

1988年10月第1版 1991年9月第2次印刷

印数 7 671—10 182

ISBN 7-04-001088-7/G·89

定价 2.50 元

前 言

《体育测量学基础》是高等师范院校体育专业学生必修的专业基础理论课,此教材通用于高师体育系本、专科。

按照原教育部(84)教体厅字 009 号文件《关于高师体育专业教材工作安排的通知》的意见,把教学计划中的《人体测定》课程改为《体育测量学基础》。1985 年 1 月,国家教委委托华南师范大学作为编写召集单位组织编材小组,参加的院校有北京体育学院、苏州大学、北京师范学院。在教材定稿前,曾先后召集了有关师范院校和师范专科学校,讨论了教学大纲、编写细目和初稿。教材的编写细目经 1985 年 7 月全国高师体育专业教材编审会议审订通过,教材初稿亦于同年 10 月经部分体育学院、师范院校及师范专科学校等教师讨论修改,最后于 1986 年 6 月经全国高师体育专业教材编审组审定通过。

体育测量学是起源于教育测量学的一门应用学科。在国外,此学科体系也只是在近几年才逐渐完善起来;在国内则刚刚起步。编写组参考了近年来国外出版的多种测量学教材,结合国情,在着重介绍学科基础理论的同时,还注重于介绍常用的素质、能力、机能、形态、情感行为、知觉、动觉及知识等范畴的测量与评价方法及评价标准。

根据目前国内的实际情况,考虑到学科体系的完整性,我们介绍了一些国外的评价标准。鉴于国情,标准不一定适用,仅供参考。在使用标准进行评价时,应依据当时的具体情况,尽量选用国内最新发表的评价标准。本书分前七章基本内容和后三章参考内容。全书的编写分工为(以章目顺序排列):第一、三、四、九、十章

由华南师范大学陈骏良执笔,第二、五章由北京师范学院晏嘉玲执笔;第六、七章由苏州大学李佑文执笔;第八章由北京体育师范学院张玉青执笔。最后统稿人陈骏良。

把“体育测量学”作为高师体育专业必修的基础理论教材是个初步尝试。限于编者的学识水平和实践经验,书中难免有缺点和错误,诚恳地希望试用单位和同志们提出意见,批评指正。

《体育测量学基础》编写组

1987年5月

目 录

前言

第一章 绪论	1
第一节 体育测量学的发展概况	1
第二节 体育测量学的研究对象及任务	2
第三节 体育测量学的几点预备知识	3
第二章 测量的有效性、可靠性和客观性	10
第一节 测量的有效性	10
第二节 测量的可靠性	17
第三节 测量的客观性	29
第三章 评价中的量表与常模	32
第一节 评价量表的概念与分类	32
第二节 常模的定义和种类	47
第三节 如何在评价中选择量表与常模	52
第四章 测验的编制、组织和实施	55
第一节 测验的选择和编制	55
第二节 测验的组织和实施	59
第五章 体育成绩的评价和定级	63
第一节 评价的意义、参照标准及形式	63
第二节 定级的准则和方法	66
第三节 成绩评价与定级的特殊方法	77
第四节 评级量表(rating scale) 在主观评价中的运用	83
第六章 身体素质和运动能力的测量与评价	88
第一节 力量测量与评价	89
第二节 肌肉耐力测量与评价	98

第三节	心血管耐力(心血管功能) 测量与评价	106
第四节	柔韧性测量与评价	113
第五节	爆发力测量与评价	121
第六节	灵敏性测量与评价	128
第七节	速度和反应测量与评价	134
第八节	平衡和韵律测量与评价	141
第七章	人体形态和结构测量与评价	150
第一节	人体测量	150
第二节	身体结构与姿势测量与评价	157
第三节	身体发育和营养测量与评价	164
第四节	身体成分和体型测量与评价	170
△第八章	动觉运动和知觉性运动测量与评价	183
第一节	动觉运动测量与评价	183
第二节	知觉性运动测量与评价	188
△第九章	情感行为测量与评价	193
第一节	获取情感行为领域信息的方法	193
第二节	情感行为量表的编制方法	200
第三节	各种情感行为测量实例	205
△第十章	知识测量与评价	218
第一节	知识测量的意义	218
第二节	知识测验的型式与命题准则	219
第三节	试卷的有效性和可靠性	227
第四节	试题分析及若干评分法则	231
附表 1	标准正态概率分布和纵标高度	241

△：为参考内容

第一章 绪 论

第一节 体育测量学的发展概况

体育测量学起源于美国的教育测量学。1861年,美国阿默斯特(Amherst)学院的首任体育主席爱德华·希契科克(E.Hitchock)第一个把测验引入美国的体育运动中,他与菲列普斯(Phillips)从事人体测量学研究工作整整50年,这是体育测量学的开拓性工作。随后,各方面专家先后在力量、形态、功能等范畴中使用了各种测验和测量。至1885年,美国召开了第一次“保健、体育与娱乐协会”的成立大会,讨论了测验的一致性和评价中使用常模(norm)的问题。从本世纪20年代起,逐步把体育测验应用于体育研究中,并结合统计手段进行,从而为在体育中使用更复杂的教育测量手段奠定了基础。

值得一提的是,1930年在美国创刊的《研究季刊》(Research Quartely),对体育测量的发展作出了很大贡献。它至今还是以刊登世界各国有关体育、运动及娱乐等测量与评价的学术论文为主。1936年编辑《研究季刊》的“保健、体育和娱乐协会”正式成立了体育测量的专门机构,专职从事测验设计与编制的研究工作,从而更加巩固了体育在教育中的重要地位。

随着教育和心理测量学的深入发展,教育测量中产生了两个新的概念:目标参照性测量(criterion-referenced measurement)和常模参照性测量(norm-referenced measurement),这是教育测量学发展史上的一个里程碑。从此测量学的发展开始与评价紧

密结合起来，使教育测量学更趋完善。不言而喻，以教育测量学为基础的体育测量学，同样也是由测量与评价两部分组成。

在我国，解放前曾一度开设过“体育测验与统计”课程。解放后在某些学院曾开设过“人体测量学”，这些课程无论从体系或内容来看，都与今天的体育测量学有很大的不同。近年来在某些学院开始引进美国的《体育测量与评价》教材。在1982年举办了全国第一次《体育测量与评价》学习班；1984年原教育部把“人体测定”正式更名为“体育测量学基础”。这说明体育测量学在国内只是在近年来才逐渐被重视。

在体育测量与评价过程中，均要使用统计知识。从测量学的研究对象来看，测量是前提，是基础；统计是手段，是方法；评价是目的，是结果。为了获取信息，必须有正确的方法和原则，才能取得真实可靠、有效而客观的数据，并通过统计方法的处理，得出有价值的结果，最终用于评价。否则，不可靠、不客观而无效的数据，即使使用最先进的统计手段，也无法得出有价值的结果。

目前，我国各级学校的体育教师在其测量与评价的实践过程中，尚缺少使用统计程序。原因是统计学知识的普及工作还做得不够，资料的统计分析缺乏电子计算机的帮助。随着电子计算机的广泛应用，体育测量学近年来发展非常迅速，各种先进和复杂的计算方法已渗透到学科中来。那些过去认为是既费时、效率又低的计算，如用麦克乐(McCloy)设计的因素分析算法，对20—30个测验进行相关系数的因素分析需要花费半年的时间才能完成。使用电子计算机后可在不到1分钟(min)的时间内，即可运算完毕，而且所得的结果却是非常有效、非常可靠的。

第二节 体育测量学的研究对象及任务

初期的体育测量学单纯是对人体各部比例的测定。后来才逐

步发展到对人体各属性的测量与评价。随着当代三大科学理论——控制论、系统论、信息论的创立与发展,随着教育测量学理论体系的日趋成熟体育测量学的研究也逐渐由对人体诸属性的测定,转变成对学校体育范畴各种属性的测定。体育测量学不仅是研究学校体育中各种信息的产生、反馈、收集、处理和评价等一系列理论及实践的一门应用学科,而且已成为在体育教学过程中,获得信息(测量学)、处理信息(统计学)和评价信息(评价学)的综合性学科。

体育测量与评价,为制定体育教学大纲提供了客观依据。在学校体育中,它着重于研究学生在从事体育运动过程中的智力、潜力、能力以及形态、功能等的变化和发展,体育测量学是为考察、检验教学、训练及锻炼效果服务的。测量可反映每个学生的属性,教师从中能了解到学生当前的训练情况效果,从而可预测学生的未来,这就是体育测量与评价的具体任务。

第三节 体育测量学的几点预备知识

一、关于测量与评价的基本概念

测量是收集资料的过程,评价则是对所测得的各种资料进行科学分析、并作出判断的过程。测量和评价是两个互相依存的概念,评价成功与否,直接取决于所收集资料的价值。倘若各次测量的结果互不一致,那么这些资料就是不真实、不可靠的,因而也是无效的,故也不能据此作出准确无误的评价。由此可知测量过程最关键,改善测量过程,就能提高评价的精确性。

(一) 测量(measurement)

在体育实践中,测量是一种手段,通过测量,可以获取客体(对象)的各种属性及其与特征有关的信息。然而客体所具有的属性

及功能都是在不断变化着的,因此,取得的信息必须以变量来说明,才具有科学意义。

体育测量的广义概念是指在体育运动中以仪器或者其它测量手段,对人体各种属性特征进行定量化的过程。就其测定的具体内容来讲,它可分为狭义的体育测量和体育测验。狭义的体育测量是指以仪器或其他测量手段,直接对受试者测定其固有的各种指标,如身高、体重等。而体育测验则是指在受试者进行某种特定身心活动的条件下,以仪器或其它测量手段,测定其从事各种活动的的能力(包括强度、时间、次数、得分等方面的指标)。

鉴于人的个体特征不同,在形态机能、素质及能力等方面也就各不相同,为此,测量前要首先确定测量的具体内容(属性或特征),选择相应的测量手段。但还有某些特征或属性,在现有的测量手段和条件下,仍很难测定,甚至是无法测定的,比如“良好的教学”即属此类。因此,只有对某种特征能明确地加以限定,并对拟测个体作出区分时,才具有测量意义。

测量过程通常使用标准化程序,其目的在于减少误差。但判断一个测量是否成功,则要涉及到用可靠性(信度)、有效性(效度)和客观性的问题。现仅就测量的客观性加以说明,因为测量既可以是客观的,也可以是主观的,比如,两名以上的教师对同一名学生进行同一测验时,若彼此所给的评分或记数极为接近,这样的测验是比较客观的。只要有可能要尽量使用客观性高的测验。

(二) 评价 (evaluation)

根据教育和心理测量专家们的最新研究,澄清了评价的基本方式和内涵。目前普遍认可的评价定义,可以从三方面来阐述,

1. 评价是对学生行为变化的信息进行收集和整理的过程,以便依据教育大纲的要求,对此作出判断;
2. 评价是指系统地收集必要的证据,借以判断学生变化的过程,并对个体变化的特点、程度和质量作出分析与判断;
3. 评价还泛指对所研究的事物拟定决策,

选择适宜材料进行分析,按照具体需要作出书面报告。

总的来说,在当今教育范畴中,评价已从单一的价值判断发展到对学生行为变化动态的一种决策过程。一般是:1. 收集相应的资料,亦即进行测量;2. 根据判断标准去鉴别资料,判明它们的真正价值;3. 依据其中有价值的资料,对受测对象的行为变化作出判断。评价的最终目的即是作出科学的合理的判断,以便教师能不断地改进教学过程。

二、测量量表的定义与分类

量表即量具,它是应用于测量或评价中的工具,故有测量量表和评价量表之分。

通过测量,我们获得各种各样的数据,要解释这些数据,我们必须知道这些数据是依据何种量表测量的。所谓测量量表(measurement scale)就是指被测事物属性的数量表述规则。这种规则是由实数数列的诸特征来决定的。

实数数列有三个特征:有序、距离、原点。量表中包括的实数数列特征越多,其反映的信息量也越多。一般以含有信息量的多寡来区分量表的级别。含信息量最多的(包含实数数列的所有特征)是最高级的量表。依此类推,不包含实数数列任何特征的量表(如称名型)属最低级量表。

实数数列的第一特征是有序。即任意两个实数都有次序。第二特征是距离,即两实数之间的差距,它是指任何一对数之间的差异是大于、等于或小于另一对数之间的差异。第三个特征是原点,即指一列数中被零位数代表的那一点,原点(即绝对零点)的重要特征是,含有零位数在内的任何一对数之间的距离,即等于该对数的另一数之值。比如0和5这对数的距离即是5。

当赋予数字以各种性质(事物、人或其它)时,便使数字之间反映了被测属性之间的关系。因此,测量量表的分类,是决定于所得

的观测数所反映实数列特征的刻度形式。这些量表不仅被实数数列的诸特征所描述,而且还被具有不变性的排列所限制。所谓量表的不变性(invariance)可定义为不改变量表结构性质的各种数字变换。下面对各种测量量表加以举例分析:

(一) 称名量表: 这种量表不含有实数列的任何一种特征, 当具有各种具有相同属性的事物集合成类时, 称名量表即形成。这些类别具有互相排他性, 即无任何事物可同属于两个范畴或类别, 例如, 足球运动员的编号即是称名量表, 9号运动员未必比8号运动员更好, 因序数不是此量表的一要素, 因为不论属任何一个类别(或编号)的成员, 并不会比其他类别的人员有更大或更小的量值意义。运动员甲不可能既是9号, 同时又是8号, 因称名的类别(编号)是互不相容的。称名测量中所取的数字属数码, 它仅是数目的符号, 不是数目, 而数目具有量值的意义。这种量表内的数字置换, 不会改变量表不变性的结构, 任何一对一的数字置换是允许的。例如, 人们可以用序数0, 1, 2, 3……10来表示足球运动员的位置, 以0号代表守门员, 1号代表右后卫……10号代表左边锋。但如改用另一套号码来表示运动员的位置, 也不会改变原位置的属性。称名量表一般不含有序数的意义。但是在实践应用中, 有些分类具有基本的序数意义。例如, 可根据素质水平把学生分成高、中、低之类, 这些类别是互相排他性的, 在高级组的学生表示比中、低级组的学生有更高的素质水平, 这一类量表可称准顺序量表。由上可知, 称名量表不含有实数数列的任何特征, 所含信息量最少, 它属最低级量表。

(二) 顺序量表: 顾名思义顺序量表是有数序的, 数序是实数数列的重要特征。比如, 有10名运动员参加比赛, 冠军为第一名, 最后为第十名, 第一名比第二名好, 第二名比第三名好, 余类推, 故数序是有意义的。但第一名可能是出类拔萃的, 而第二、三、四名的能力也许差别不大, 很可能是近似的。故能力上的差异并不能

由名次之间的差异来确切反映。因此,实数数列的第二个特征——距离不是顺序量表的一个要素。

当事物属性被观测而非排列时,也可使用顺序量表,但观测数之间的差异是无意义的。比如,评选最佳运动员,要求记者投票。假如运动员甲得 232 票,运动员乙得 93 票,运动员丙得 190 票。显然甲是得票最多的运动员,但票数尚不能说明甲比乙好多少,因运动员得票数的差距,尚不具有水平高低的意义。

数序的任何变换将偏离顺序量表的不变性结构,故顺序量表的最重要特征是数序。上面选最佳运动员的例子,只要甲、丙、乙的名次保持不变,运动员实际得票数的多少是无多大意义的。不管甲比丙是多 42 票,还是多 2 票,甲总是被定为第一名。

实数数列的第三特征原点,也不是顺序量表的一个特征。比如,评选最佳足球运动员,得零票数的运动员并不能说明他完全缺乏足球能力。故顺序量表称为低级量表。

(三) 等距量表:此量表具有实数数列的两个特征——次序与距离。原点不是它的特征,但它具有相对零点。比如,在引体向上测验中,得零分的学生(未能拉引一次),并非说明他完全无力,故零分是相对的。

此量表除具有序数意义之外,还具有等距离的特征。 70° 与 80° 之差等于 40° 与 30° 之差。运动成绩测验中的 93 分和 90 分之差等于 53 分和 50 分之差。这种情况从测量上说是对的,但从评价上说是不真实的,因为得 90 分的人增加 3 分远比得 50 分的人增加 3 分要困难得多。这是属于成绩评价问题,应用其他方法来解决。

在保持量表结构不变性的情况下,对等距量表作任何线性变换是可行的。量表中的每个值被一具体数除或乘,只改变其测量单位;在每个值加上或减去一个数,也仅改变数列的相对零点。由于此量表的零点是一个相对的数值,它非绝对零点,故不能说一个

值是另一个值的两倍或说 $1/3$ 倍。在等距量表中如果取任何两个数的比率,将会改变其量表的性质。只有以零代表不存在的拟测属性时,作上述的比较才是有意义的。因等距量表所含的信息量较多故称为高级量表。

(四) 比率量表:这种量表含实数数列的全部特征。由于它的绝对零点已定,只测量单位可自由变动,如果从某个变量加上或减去一个数就会改变其原点,即失去比率量表的性质之一。量表中的任何值乘以或除以某一值也会偏离量表的不变性结构。

比率性变量包括长度、重量、时间等的测量,这类量表是体育中最常用的。这种量表除表明量的大小和同等测量单位外,还有绝对零点,以至测量的变量可作比率性陈述。比如,可以说:170 cm 的成人身高是 85 cm 的小孩身高的两倍。在运动成绩的评价中,比率数据的某些基本假定,存在一些有待解决的问题。能否说跳 2 m 的成绩是跳 1 m 的两倍呢?当您意识到这是在对成绩作评价时,会发现这种等值差异的观念是不成立的,实践中的这类观测数是普遍作为等距资料来处理的。这说明了成绩评价的复杂性。

比率量表含最多的信息量,故称最高级的量表。

三、数的精确限(real limits)

连续性变量(由比率量表所测定)是指量尺上任意两点之间具有无限个可能取值的变量,而实测所得的数据,取其整数,是一个近似值。这就出现了数的界限问题。例如,时间这一变量是用年(a)、小时(h)、分(min)、秒(s),以至毫秒(ms)为单位来测量的。如果测量到 s,即使用最精密的计时表,也只能得到 s 的较好近似值。比如,小孩跑 60 m,成绩是 8 s,并不意味着是精确的 8 s,而只能理解成“近似的 8 s”,意味着在 7.5 s 和 8.5 s 之间(不含 8.5 s),于是 7.5 s 和 8.5 s 就成为 8 s 这个数的精确限。这个数的下精确限是 7.5,因为小于 7.5 并大于 6.5(包含 6.5)的任何数(如 7.3),都将

四舍五入取其约整数记为7而不记8。其上精确限是8.5,因大于8.5并小于9.5的任何数,也将四舍五入取其约整数记为9而不是8。一个数的精确限,是指位于这个数的上下各半个测量单位的那些点。如果测量的时间要精确至 s 时,其测量单位是 $1s$,因此 $8s$ 的精确限是8.5和9.5,在 $8s$ 的上下分别占半个测量单位($0.5s$)。如果测量要求精确到 $1/10s$,其记录值应为8.0,那么该值的精确限便是7.95 s 和8.05 s 。如果测量要精确至 $1/100s$,其记录值应为8.00,于是8.00的精确限是7.995和8.005 s 。

第二章 测量的有效性、可靠性和客观性

我们在选择或设计一个测验或测量手段时，必须考虑到测量的有效性、可靠性和客观性。如果缺乏这些特性，则测量的结果就不能充分置信。

在测量的这些特性中，首先要考虑到有效性，然而离开了对可靠性的分析来讨论有效性是极不全面的。一个有效的测验必须是可靠的，但一个可靠的测验并不一定有效。如50 m跑是可靠的测验，也是测量速度的有效手段，但它不是测量灵敏性的有效手段。当然，如果一个测量或测量手段缺乏客观性，则减少了其可靠性和有效性。

第一节 测量的有效性(validity)

一、有效性的概念和种类

测量的有效性是指一种测验或测量手段能够测出它所要测量的因素(内容、性质)的准确程度。例如：单杠引体向上是测臂部肌肉力量和耐力的有效手段，因为肱二头肌的力量和耐力对引体向上是关键因素，但该测验不是上肢运动速度的有效测量，因引体向上次数不受臂部运动速度的影响。

根据测验的性质和目的，有效性可分为四种：内容有效性、编制有效性，协同有效性和预测有效性。