

学生体育学业成绩评定真相的实证研究

程传银,孟凡花

(南京师范大学体育科学学院,江苏 南京 210023)

[摘要] 体育教师评价学生的体育成绩时,最重要的评价标准是学生的努力程度,其次是行为表现,运动能力居第 3 位.然而相反的是,从教师们给学生的最终分数来看,运动能力却是最重要的评价标准,其次分别为努力程度和行为表现.研究表明:教师们在教学计划中呈现的评价方法与他们实际评分间存在差异.因此,我们在处理仅来自教师报告中的信息时,应小心谨慎.

[关键词] 评价,评价标准,评价实施

[中图分类号] G806 [文献标志码] A [文章编号] 1001-4616(2013)01-0148-05

Empirical Research on the Latent Facts in Grading the Students' Performance in PE Classes

Cheng Chuanyin, Meng Fanhua

(School of Sport Sciences, Nanjing Normal University, Nanjing 210023, China)

Abstract: According to teachers' reports, the most important grading criterion was effort. Behaviour was ranked in the second place and performance in the third. On the contrary, according to the grades they assigned to their pupils, performance was the most important criterion followed by effort and behaviour. The results of this study indicated a discrepancy between teachers' written reports and how they actually assigned grades. Therefore, information from studies based solely on teachers' reports should be handled with caution.

Key words: grading, grading criterion, grading complement

有关评价标准的讨论通常集中于 2 个问题:1. 我们应该采用哪些评价标准? 2. 评定学生最终成绩所用各标准的比重是多少^[1]?

关于评价标准,可以把它们分成 3 类:学生的运动能力、努力程度和行为表现. 学生的运动能力包括学生达到某标准或规定数据的程度,如技能测试的分数、游戏比赛的评估等^[2]. 学生的努力程度包括学生对任务或游戏的积极参与程度. 由于进步被看作是努力的结果,因而它也被列入努力程度这一类中. 而且,学生的进步也说明他参与其中,并付出了努力. 因此,学生的努力、参与及进步被认为是紧密相关的 3 要素,并被列入同一类别中^[3]. 评价标准的第 3 类即学生的行为表现,包括学生服装得体、按时上课、认真听讲以及学校教育相关的其他行为,如诚实、尊重他人等^[4].

关于教师的评价方法的研究揭示,在评价学生时,学生的努力程度和行为表现比他们的运动能力更为重要. Morrow 再次研究表明,1956 ~ 1973 年,学生的努力程度和行为表现在评价中占据主要位置^[5]. Mathews 的研究表明,80% 的教学情境中,评定成绩只与学生的参与程度相关^[6]. Hensley 等的报告中指出,在调查的 3 000 多位教师中,他们最常用的评价标准是努力程度和参与程度. 时代在变化,实践却止步不前^[7]. Matanin 和 Tannehill 的调查表明,教师在评定学生最终成绩时,努力程度和行为表现约占评价标准的 2/3^[8]. 最终,Duchane 和 French^[9]以及 Chatzopoulos 和 Derri^[10]的研究表明,最常用的评价标准为努力程度和行为表现.

然而,这些有关评价标准的信息依据的仅仅是教师们的书面或口头报告. 在相关研究中使用的研究方

法是问卷调查、专家访谈和观察法(即观察教师们在日常上课时对学生强调的评价标准是什么)^[11]。然而,关于教师评价标准的报告的有效性,却还未见报道。

对于人们言与行之间的差异,这已经是一个古老的话题。然而,这项研究的调查分析却远远超出了我们的共识:语言与行为之间常不一致。最初研究这个问题的是 Argyris 和 Schon,他们认为人们在任何情况下采取何种行动,这在他们思想上早已做好规划,这涉及到他们如何计划、如何实施以及如何检查他们的行为^[12]。此外,Argyris 声称:正是这些所谓的思想规划图指导着人们的行动,而不是那些他们所信奉的口头理论。更重要的是,很少有人能意识到他们所用的这些规划图或理论。我们可以用这种方式来解释此现象:理论与实践间存在差距^[13]。然而,Argyris 认为这涉及到两种行为理论,即一种理论用于实践;而另一种理论只是口头说给别人听而已^[14]。

当一个人被问到在一定情况下他将如何表现时,他的回答往往是在特定情况下的“口头理论”。这是一种表现他忠诚并用来与别人交流的行为理论。然而,真正控制一个人行动的理论是用于实践的理论,这种理论不一定与他的“口头理论”相一致。此外,人们也不一定能意识到这两种理论间的不一致性^[15]。

把这两者相区别,我们会问到这样的问题:哪些行为适合“口头理论”? Argyris 等指出,教师们的行为(把理论用于实践)取决于使得这些行为必须的背景因素,尽管这些行为与教师们所说所想不一致,但是却与教师们所信奉的日常因素相一致^[14]。

然而,关于教师们的“口头理论”和“用于实践的理论”是否一致的研究十分有限,且结果不一致。一些研究者报告中指出,教师的“口头理论”与他们的实际行动是相关的^[15]。然而另一些研究者却认为教师们的“口头理论”与实际行动不一致。可是,在体育教育中,教师们是否按他们教学计划中所写的方法去评定学生成绩?对于这个问题,尚无人研究。

本研究的目的是为了调查关于评定学生成绩时,教师们的书面报告是否与他们实际评价相一致。具体说来,第1个问题强调的是:教师们的“口头理论”中,判定学生最终成绩时所用的评价标准的各个比重是多少?第2个问题是:在学期末评定学生成绩时,教师们所用的评价方法是否与他们“口头理论”中所讲的方法相一致?

1 方法

用2个步骤来说明以下方法。

1.1 参与者

在华东地区选择51位自愿参与这项研究的体育教师,这些教师教5~6年级学生(年龄在11~12岁),且教龄都在5~23年。在这项研究中,对教师采取机密和匿名的方式,最后获得一份正式书面报告。对5~6年级学生的等级评价尺度用数字1~10分来衡量,10分表示最好,4分表示不及格。对体育教育的评价结果是学生们的最终总成绩,因为评价结果对学生的奖励有非常重要的影响,所以教师的评价行为也会因此而改变。

1.2 收集资料

评价标准可分为3类:学生的运动能力、努力程度和行为表现。学生的运动能力包括:根据标准数据(考试分数)来评定学生所处级别。学生的努力程度是指根据学生如何努力地去提高自己来评定,这是教师的一种主观判断。努力、进步和参与被认为是密切相关的3要素,因而被归于同一类。第3类,即学生的行为举止,包括服装得体、按时上课、认真听讲等。为了调查教师们的“口头理论”,他们被要求写一份书面报告,这是一份关于组成评价体系各标准的百分比的报告。体育教师们把评价按照运动能力、努力程度、行为表现这3个维度来处理。

试验可靠性的建立来源于对18位体育教师的调查问卷,这些教师是在2个不同的时期(间隔3个月)完成调查问卷的。学生的运动能力、努力程度和行为表现的组间系数分别为0.89、0.83和0.92。为了调查教师们实际上是如何评定学生的,根据以上对评价标准的分类,我们在表1中运用了新方法。其涵盖了8个学生的3个评价标准的各个分数。我们要求教师对表1中提及到的8个学生进行评价,然后在学期末再次对抽样的8个学生进行评分。

关于表 1 中的运动能力,100 分表示满分,50 分表示均值. 例如第 1 位学生,运动能力优秀(80 分),非常努力以及行为表现好. 相比,第 8 位学生运动能力差(20 分),努力程度一般以及行为表现差. 为了计算出用于总成绩评定的各个评价标准的比重,应进行以下操作:

运动能力:我们把所有运动能力优秀者的分数相加($P1+P2+P3+P4=PG$,运动能力优秀者),然后减去运动能力差者分数和($P5+P6+P7+P8=PB$,运动能力差者),用运动能力优秀者分数总和减去运动能力差者分数总和等于成绩占的比重($PW=PG-PB$). 其他标准的比重用类似的方法来计算.

努力程度:我们把非常努力的学生的分数相加($P1+P2+P5+P6=EG$,非常努力者),然后我们把努力程度一般的学生的分数相加($P3+P4+P7+P8=EL$,努力一般者),用 $EG-EL$ 所得结果即为努力程度的比重(EW).

行为表现:我们把行为表现好的学生的分数相加($P1+P3+P5+P7=BG$,行为表现好者),然后我们把行为表现差的学生的分数相加($P2+P4+P6+P8=BB$,行为表现差者),用 $BG-BB$ 所得结果即为行为表现的比重(BW).

因此,我们可以根据各标准所占比重来划分评价标准级别. 运动能力比重(PW)、努力程度比重(EW)和行为表现比重(BW),这 3 者中哪个数字最大表明这项标准在评价体系中是最重要的.

试验可靠性的建立来源于对 17 位体育教师调查时所运用的方法,是通过对 17 位教师在 2 个不同的时期(间隔 1 个学期)进行调查,最后统计出数据(教师每一学期都要评定学生成绩). 运动能力比重、努力程度比重和行为表现比重的组间可靠系数分别为 0.90、0.91 和 0.88. 由于这种方法的有效性,且对学生分数精确计算的结果表明:对评价标准各比重的计算是一个逻辑过程.

2 结果

教师们的书面报告 and 实际评价中所用到的 3 个评价标准的比重(运动能力、努力程度和行为表现)在表 2 中分别做出了描述统计.

在书面报告中,教师们把努力程度看作是最重要的评价标准,其次是行为表现. 与努力程度比重(45.49)和行为表现比重(29.41)相比,运动能力(25.10)列居第 3 位. 具体来说,29 位教师(56.8%) 在报告中指出,努力程度在评价标准中占最大比重;10 位教师(19.6%)认为行为表现占最大比重;9 位教师(17.64%)认为努力程度和行为表现占同等比重,运动能力则占较小比重;剩余 3 位教师(5.88%)则认为努力程度和运动能力占同等比重,行为表现则占较小比重.

根据实践中教师评定 8 位学生的分数来看,运动能力却占最大比重(4.72),其次为努力程度(3.31),行为表现列居第 3 位(3.19). 具体说来,有 31 位教师(60.78%)依据学生的运动能力来评定学生;5 位教师(9.8%)认为 3 个评价标准占同等比重;4 位教师(7.84%)依据学生的努力程度来评定学生;2 位教师(3.92%)认为运动能力和努力程度占同等比重,行为表现占较小比重;4 位教师(7.84%)认为运动能力和行为表现占同等比重,努力程度占较小比重;3 位教师(5.88%)认为努力程度和行为表现占同等比重,运动能力占较小比重.

从书面报告中可以得知:38 位教师声称努力程度和行为表现是最重要的评价标准,然而在实际评价学生时,他们中的大部分(20 位)却把运动能力作为最重要的评价标准. Spearman 相关系数表明,教师们的言与行之间没有显著相关性.

表 1 8 位学生的运动能力、努力程度和行为表现

Table 1 Athletical skills,level of efforts and class performance of 8 students

学生	评价标准		
	运动能力	努力程度	行为表现
1	80	非常	好
2	80	非常	差
3	80	一般	好
4	80	一般	差
5	20	非常	好
6	20	非常	差
7	20	一般	好
8	20	一般	差

表 2 评价标准的各个比重、教师们的书面报告 and 实际评价

Table 2 Ratio of the criterion,teachers' written reports and actual grading

评价标准	书面报告 M±SD	实际评价 M±SD
运动能力	25.10±16.98	4.72±1.89
努力程度	45.49±14.04	3.31±1.44
行为表现	29.41±10.98	3.19±1.09

3 讨论

这项研究结果表明,教师的书面报告和他们的实际行动是不一致的.在教师们的书面报告中,努力程度在评价标准中占最大比重,行为表现列居第2,而运动能力在评价标准中占最小比重.教师们在报告中的这些回答是与以往研究的结果相一致的^[9].然而,根据教师们给予学生的分数来看,运动能力却成为最重要的评价标准.这表明,教师们言与行之间存在差距.这种差距可以这样解释:我们假定教师们实际评价行为基于一定的环境因素,这些环境因素使得他们必须那样去做^[14].尽管这些行为可能与他们所说的不一致,但是这些行为却可能与学校方针政策和功能相一致(如奖励、优秀学生的选择等).

具体来讲,在以往的研究中,教师们在他们书面的教学计划中指出:在评定学生时,努力程度和行为表现比运动能力重要得多^[16].这个研究结果和我们在研究中教师们的书面回答是一致的.然而,根据我们的研究结果,教师们在学期末给学生评定分数时,最重要的评价标准却是运动能力.导致教师们言与行之间差异的可能原因是教师们对处理纪律问题的忧虑.为了避免学生对纪律漠不关心和处理一些纪律方面的问题,一些体育教师强调行为表现和努力程度在评价标准中的重要性.然而,在学期末,教师们必须给学生打分,以决定哪些学生可以获得奖励.分数的一个基本功能是告知学生、家长及管理者:和其他学生相比,这个学生处于哪个等级^[17].根据学生们的行为表现或努力程度来给学生评分,以此来实现分数的基本功能.或许,一些教师希望根据学生的行为表现或努力程度来评定分数,然而他们的这种希望却与学校的方针政策不一致.因此,在将来的研究中,查明教师们言与行不一致的原因是十分重要的.为了提高教师自身的教学知识,教师们需要明确“口头理论”与实际行动,并且找出这两者不一致的原因.“反思行为”在教师们是否有能力使得他们的理论与行动更加协调一致中是一个至关重要的因素.

这项研究结果对有关评价标准的书面报告提出了强烈的质疑,并同时支持了 Argyri 的理论与实践不一致的观点.如果关于评价标准的书面报告与教师们的实际评价间确实存在差异的话,那么基于教师们书面报告的研究结果应被小心应用.对于我们而言,关于教师们对于评价标准各比重的报告与他们实践中的应用是否一致性的研究,没有任何经验可借鉴.因此,我们需要对教师们的理论与实践的关系的研究有更深的洞察力,并做进一步的研究.

此外,在这项研究中,教师们并没有意识到他们言行间的不一致性,直到他们获得研究结果时才意识到这种差异.因此,教师培养与发展的课程,不仅要为教师获得新知识创造良好的机会,而且还要培养他们把理论用于实践的能力.教育者在评价时有必要认同在学习中的真实经历的价值.我们不仅期望教师们能够明确他们的“口头理论”,而且还要提供给他们把理论用于实践的机会,这两者在教师职前培训和在职培训中都是十分重要的.尽管我们并不打算改变教师们在实际中的评价行为,然而有关行为和信念的反思能够帮助教师们对行为以及潜在的变化取得一致性的观点.

最后,在这项调查中我们使用了两种手段:1.关于评价标准各比重的调查问卷;2.对8位学生的评分.这些应该会对职前和在职的教师在实际上如何评价学生有帮助.如果我们热衷于体育教学实践,我们就必须对教师们的“口头理论”和“实际行动”格外关注.此外,这项研究是对教师们在实际中是如何评价学生的第一次尝试.因此,为了拓展我们对有关教师们“口头理论”和“实际行动”间关系的认知,我们需要更进一步地研究这层关系.

[参考文献]

- [1] Safrit M, Wood T. Introduction to Measurement in Physical Education and Exercise Science [M]. 3th ed. St. Louis: Mosby, 1995.
- [2] Veal M. School-based theories of pupil assessment: A case study [J]. Research Quarterly of Exercise and Sports, 1992, 63: 48-59.
- [3] Rink J. Teaching Physical Education for Learning [M]. 4th ed. Boston: McGraw-Hill, 2001.
- [4] Chatzopoulos D, Derri V. Grading profiles of high school physical educators: A cluster analysis [J]. Journal of Human Movement Studies, 2004, 47: 61-73.

- [5] Desrosiers P, Genet-Volet Y, Godbout P. Teachers' assessment practices viewed through the instruments used in physical education classes[J]. *Journal of Teaching in Physical Education*, 1997, 16: 211–228.
- [6] Mathews D. *Measurement in Physical Education*[M]. 5th ed. Philadelphia: Saunders, 1978.
- [7] Hensley L, Lambert L, Baumgartner T, et al. Is evaluation worth the effort? [J]. *Journal of Physical Education, Recreation and Dance*, 1987, 58(6): 59–62.
- [8] Matanin M, Tannehill D. Assessment and grading in physical education[J]. *Journal of Teaching in Physical Education*, 1994, 13: 395–405.
- [9] Duchane K, French R. Attitudes and grading practices of secondary physical educators in regular education settings[J]. *Adapted Physical Activity Quarterly*, 1998, 15: 370–380.
- [10] Chatzopoulos D, Mouratidou K. Grading criteria of students in physical education of primary education[J]. *Hellenic Journal of Physical Education & Sport*, 2004, 52: 33–44.
- [11] Veal M. Pupil assessment issues: A teacher educators perspective[J]. *Quest*, 1988, 40: 151–161.
- [12] Argyris C. *On Organizational Learning*[M]. Oxford, UK: Blackwell Publishers, 1992.
- [13] Argyris C, Schon D. *Theory in Practice: Increasing Professional Effectiveness*[M]. London: Jossey-Bass, 1974.
- [14] Argyris C, Putnam R, Smith D. *Action Science*[M]. San Francisco: Jossey-Bass, 1985.
- [15] Chen A, Ennis C. Teaching value-laden curricular in physical education[J]. *Journal of Teaching in Physical Education*, 1996, 15: 338–354.
- [16] Tsangaridou N, O'Sullivan M. Physical education teachers' theories of action and theories-in-use[J]. *Journal of Teaching in Physical Education*, 2003, 22: 132–152.
- [17] Miller D, Keith D. *Measurement by the Physical Educator: Why and How*[M]. Boston: WCB McGraw-Hill, 2002.

[责任编辑:黄 敏]