

编制教育评价方案的一个关键问题

——加权及其方法研究

霍力岩

教育评价工作不仅要求对指标体系中的各项指标作出单项的评价,而且要求对评价对象的整体作出综合的、统一的评价,而综合的、统一的评价比对各项指标的单项评价在教育评价中的意义更为重大。例如,我们在对学生的发展状况进行评价时,不仅要对学生德、智、体、美各个方面的发展状况作出分别的评价,更要对包括学生德、智、体、美各个方面的发展状况在内的全面发展状况进行综合的、统一的评价。对评价对象作出综合的、统一的评价就需要我们通过一定的方法把对各项指标的评价值进行统计分析,得出总的评价值。这样一来,编制一份教育评价方案就不可避免地涉及到各级、各项指标在整个指标体系中的地位 and 相互关系问题,也可以说是涉及到各级、各项指标对整个指标体系贡献大小的问题。由此,加权就成了编制教育评价方案时的一个关键问题。事实上,编制一份科学、合理的教育评价方案来评价学校教育工作和学校中学生的状况是教育评价工作者需要认真面对的问题,而编制教育评价方案中的加权问题又是教育评价方案编制中的一个关键问题和令教育评价工作者感到困惑、困难的问题。

一、加权及其意义

编制一份教育评价方案必须解决好加权的问题,而要解决好加权的问题,必须对加权

及其意义有清晰的认识。

(一)加权的概念

加权就是根据各级、各项指标在整个指标体系中的地位和重要性程度,给有关指标分派代表其地位和重要性程度的不同的数值(这个数值一般被称为权重值)。

通过加权对教育评价方案中的各项指标进行综合统计时,常用的表示方法是:

$$S = W_1 X_1 + W_2 X_2 + \cdots + W_i X_i + W_n X_n$$

S 为总评价价值;

W_i 为指标 i 在指标体系中的权重值;

X_i 为某一指标 i 的评价值;

上式也可以表示为:

$$S = \sum W_i X_i \quad (\sum \text{为求和符号})$$

教育评价中的加权主要是指对教育评价指标体系中的各个指标进行加权。加权要求我们先将权重值分派到一级指标,然后再随着一级指标的分解而将权重值逐层分解下去,直至最后一级指标(通常是指经过三级分解后所得的指标)。这种加权分配形式是:

总目标为 1

一级指标权重值为 $W_1, W_2, \cdots, W_i, \cdots, \sum W_i = 1$; 二级指标权重值为 $W_{i-1}, W_{i-2}, \cdots, W_{i-j}, \cdots, \sum W_{i-j} = 1$

简言之,假定总目标的权重值为 1,其后每一级指标的权重值之和则均为 1,但下一级指标随着指标级别的降低,在总权重值(一级指标 1)中所表现出的贡献量就越越来越小。

表 1 是一个指标权重值分配的实例:

表 1 学校工作评价指标权重值分析表

一级指标	权重	二级指标	权重	折 合 总权重	三级指标	权重	折 合 总权重
A-1 办 学 条 件	0.13	B-1 规模、设备	0.2	0.026	C-1 校舍规模 教学设备	0.6	0.022
					C-2 校舍利用 设备维修	0.4	0.010
		B-2 领导班子	0.3	0.039	C-3 结构合理	0.6	0.023
					C-4 配合协作	0.4	0.015
		B-3 教师队伍	0.5	0.075	C-5 配备齐全 合 理	0.4	0.030
					C-6 素质高 水平好	0.6	0.045
A-2 管理水平	0.25	略					
A-3 工作质量	0.29						
A-4 学生发展水平	0.33						

(二)加权的意义

我们可以这样认为,教育评价方案中的指标体系构成了评价因素集(即评价哪些项目),加权就是这些因素的关系集(即所评价的这些项目各自的重要性程度如何)。各评价因素的权重值既表示它们在因素集中的地位,又表示它们与其它诸因素的关系。加权的意义也正是由此引申而来:

1. 加权影响评价结果。

不同的加权观念会导致不同的评价结果。由于权重值反映的是指标体系中各项指标在整体中的地位和作用,是每一项指标的评价值对总的评价值结果的影响程度,如果人们对某一指标的重要性认识不一,给予同一指标不同的权重值,就会导致不同的评价结果。例如表 2 中的甲、乙、丙三种加权观念,就得出了三种不同的评价值,显示了人们评价学生发展状况的三种价值观念。

表 2 权重值影响学生发展水平评价结果表

指 标	权重值			评价等级				评价值		
	甲	乙	丙	A(95)	B(80)	C(65)	D(50)	甲	乙	丙
体育成绩	0.25	0.4	0.1	✓				23.8	38	9.5
智育成绩	0.25	0.1	0.4				✓	12.5	5	20
德育成绩	0.25	0.4	0.1	✓				23.8	38	9.5
美育成绩	0.25	0.1	0.4				✓	12.5	5	20
合 计	1.00	1.00	1.00					72.6	86	59

不同的加权观念会导致评价结果的不同,这一点告诉我们在给各项指标分派权重值时一定要慎之又慎。

2. 加权指引人们重视“重要”的指标。

各项指标权重值的大小,从数量上向人们传达了各项指标重要程度的信息,必然会产生人们对不同指标的不同重视程度。如果我们在学生发展评价指标体系中,给智育成绩分派了较大的权重值,那么,人们就会给予智育较大的重视;相反,如果我们给智育成绩分派了较小的权重值,就会引导人们不重视智育,而是去重视其它权重值较高的指标了。从这个意义上来讲,当前的教育改革要求我们的加权能够引导人们重视素质教育的一些关键指标,当然这种引导必须通过科学、正确地分派权重值来实现。

二、加权的两个有效方法

教育评价方案中的加权是人们对各级、各项指标之间关系认识的凝聚物,同时也是人们的价值认识的凝聚物,因此,进行加权的过​​程既是反映人们价值认识的过程,又是统一人们价值认识的过程,进行加权必须考虑权集的这一根本特性并从这一根本特性出发。根据我们长期的理论研究和实践研究,我们认为加权有两个有效的方法,即特尔斐法、层次分析法。

(一) 特尔斐法(Delphoi Technique)

特尔斐法是一个使专家集体在各个成员互不见面的情况下对某一项指标的重要性程度达成一致看法的方法,它是编制教育评价方案中进行加权时经常使用的一种方法。

特尔斐法因出自古希腊特尔斐地区的预言家而得名,它最早是古希腊特尔斐地区的预言家预测未来时经常使用的方法。特尔斐法在现代的运用是从 50 年代开始的,主要被作为预测未来的工具使用在未来学的研究中。60 年代中期以来,特尔斐法的应用范围

迅速扩大,在未来学以外的其它领域特别是评价领域得到了广泛的运用,为在解决各种形式的复杂问题,如某一指标的重要性程度时促进信息交流并取得一致意见发挥了重要作用。现代教育评价中的特尔斐法实际上是一种在编制教育评价方案时为了取得对某一指标或某些指标重要性程度的一致认识而进行的专家意见征询法,它以分发问题表的形式,征求、汇集并统计一些资深人员对某一项指标重要性程度的意见或判断,以便在这一问题的分析上使大家取得一致的意见。

1. 步骤

应用特尔斐法分配权重值的步骤是:

(1) 设计意见征询表。

根据所要咨询的问题设计好问卷表格,意见征询表的一般形式如表 3:

表 3 学校工作评价权重值分配意见征询表

指标项目	很重要	重要	一般	不重要
办学条件				
工作质量				
学生发展水平				

设计意见征询表时,需要特别注意两个问题:其一,表中所列的重要性等级如上表中所列“很重要”、“重要”、“一般”、“不重要”等必须有明确定义,即需要明确说明在何种情况下才能算得上很重要,在何种情况下才算是重要等,以免由于对这些词语的误解造成误判,从而影响意见征询的科学性;其二,为了使专家容易将上面的重要性等级换算成权重值,事先应对这些重要性等级赋值。

(2) 选择专家并请他们填写问卷表格。

选择参加咨询的专家并要求他们不署名地根据要求将对某一指标或某些指标重要性程度的看法写在问卷表格中。选择专家时应注意专家既要权威性又要代表性,即所选择的专家应是对所要咨询的问题有深入了

解和研究的人士,所持的观点具有权威性;同时所选择的专家来源应涉及到和要咨询问题有关的各个方面,即所选择的专家应是各个方面如行政管理人员、科研人员、实际工作者等的代表。请专家填写意见征询表时应注意以书面或口头的形式(最好以书面的形式)提醒他们完全按规范和要求填写,不应随着展开或以其它不被允许的方式回答咨询。

(3)整理和反馈专家意见。

所有专家将意见征询表格填好交回后,组织者要整理专家们的意见,求出某一项指标或某些指标的权重值平均数,同时求出每一专家给出的权重值与权重值平均数的偏差,然后将求出的权重值平均数反馈给各位专家,接着开始第二轮意见征询,以便确定专家们对这个权重值平均数同意和不同意的程度。

在这一步骤中,需要对下面两项内容作统计处理:

a. 每一指标权重值的平均数:

$$M(W_i) = 1/n \sum W_{ij}$$

其中:

M 为权重值的平均数;

i 为第 i 项指标;

W_i 为第 i 项指标的权重值;

$M(W_i)$ 为第 i 项指标权重值的平均数。

W_{ij} 为第 j 项专家对第 i 条指标给出的权重值。

b. 每一专家给出的权重值与权重值平均数的偏差:

$$\Delta_{ij} = W_{ij} - M(W_i)$$

其中:

Δ_{ij} 表示与权重值平均数的偏差。

(4)不断整理和反馈专家意见。

再一次将权重值平均数反馈给各位专家并给出某些专家不同意这个平均数的理由,让各位专家在得知少数人不同意这个平均数的理由后再一次作出反应。重复进行上述整理和反馈专家意见的步骤,直至再重复下去

观点集中程度或认识统一程度不能增加多少时停止。这样重复几次以后,各位专家对某一指标或某些指标的权重值的看法就会趋向一致,组织者也就可以由此得到比较可靠的权重值分配结果。

2. 特点

特尔斐法作为编制教育评价方案时经常使用的一种专家意见征询法,它与一般性的意见征询不同,它具有以下几个方面的典型特点:

(1)参加意见征询的人士对所咨询问题的回答具有权威性。

参加意见征询的人士一般是对所咨询的问题有比较深入研究的专家和权威,他们中某一人士对所咨询问题的回答就已具有某种意义上的权威价值,众多这些人士对所咨询问题的一致回答就更具有权威价值。如果专家都认为某一指标重要,要给它较高的权重值,那就说明该项指标确实重要。反之亦然。

(2)参加意见征询的人士对所咨询问题的回答具有独立性。

参加意见征询的专家和权威在整个征询意见的过程中并不见面,不是“面对面”就某一指标的重要性程度进行讨论,而是“背靠背”地回答组织者的咨询。专家之间没有面对面的相互影响和相互对抗,从而有效地减少了专家中资历、口才、人数优势等方面因素对他们回答问题的影响。可以说,参加意见征询的人员不受权威左右,不受口才好坏、人数多少的影响,对所咨询问题的回答均是自己独到的见解,具有较强的独立性。

(3)参加意见征询的人士能够逐步取得价值认识和判断的一致。

在意见征询的一轮又一轮反复中,有关专家可以通过反馈回来的经过整理的各轮应答情况,了解并认真考虑他人的思想和意见,在此基础上,决定是否修正和如何修正自己原来的想法。一般来说,整个意见征询过程中专家的意见一轮比一轮相对集中,呈逐步

收敛的趋势。这就保证了根据大多数人的价值认识去统一所有人员的人价值认识,保证了参加意见征询专家的价值认识能够逐步地取得一致。

(4)意见征询是一个有组织、有控制的过程。

专家只能按照意见征询表中所列非常明确、具体的问题依照指定的回答方式简单明了地表示自己的意见。

(二)矩阵对偶法

矩阵对偶法是通过指标体系中的各项指标进行两两比较,在一系列可供比较的指

标中,找出各项指标的优先顺序,从而确定各项指标权重值的方法,它也是编制教育评价方案中进行加权的常用方法。

矩阵对偶法的数学基础是矩阵理论。由于一般矩阵的求解比较复杂、繁琐,在教育评价指标的加权中一般采用一种改进型的矩阵对偶法来求解权重值。这种矩阵对偶法规定判断的等级为三个即同等重要、稍微重要、明显重要,判断值一律是整数。

假设是甲、乙、丙、丁四项指标,要使用矩阵对偶法确定它们各自的权重值,可以先制好如下表格,以确定各项指标比较的顺序:

表 4 矩阵对偶法比较顺序表

比较前项 比较后项	甲	乙	丙	丁
甲		①	④	⑤
乙	1		②	⑤
丙	4	2		③
丁	6	5	3	

表中数字指各对指标比较的顺序,其中加圈的数字是比较的前项,不加圈的数字是比较的后项。有了上述比较的顺序,我们就可以进行比较了。如第一次比较乙和甲,第二次比较丙和乙……。

上表中比较的顺序应该是:①乙和甲;②丙和乙;③丁和丙;④丙和甲;⑤丁和乙;⑥丁和甲

需要指出的是,使用矩阵对偶法进行比较时,比较的顺序是固定的,不能随意改变。每次将两个指标进行比较后,按照上表中比较前项和比较后项的标记给各指标记分,并将分数记在类似于上面表格的空白表格中,同时记上配对比较时另一项指标的相应评分。记分方法如表 5:

表 5 矩阵对偶法记分表

计 分		含 义
甲	乙	
2	2	两项指标相比,甲与乙同等重要
3	1	两项指标相比,甲比乙稍微重要
4	0	两项指标相比,甲比乙明显重要

现以学校工作评价方案中四项指标权重值的确定为例说明如何运用矩阵对偶法确定各项指标的权重值:

现假定学校工作评价方案的一级指标有 4 个,即办学条件、管理水平、工作质量和学生发展水平。根据矩阵对偶法比较顺序表,我们可以得到上述 4 个指标的比较顺序如下:

①管理水平——办学条件

- ②工作质量——管理水平
- ③学生发展水平——工作质量
- ④工作质量——办学条件
- ⑤学生发展水平——管理水平
- ⑥学生发展水平——办学条件

经过以上对各项指标依次进行比较,我们认为:

- ①管理水平比办学条件稍微重要;

- ②工作质量和管理水平同等重要;
- ③学生发展水平和工作质量同等重要;
- ④工作质量比办学条件稍微重要;
- ⑤学生发展水平比管理水平稍微重要;
- ⑥学生发展水平比办学条件稍微重要。

根据矩阵对偶法计分表,将对以上各项指标比较后的结论转化成分数,计入“学校工作评价一级指标权重值分配表”,由此得到表6。

表6 学校工作评价一级指标权重值分配表

比较前项 比较后项	办学条件	管理水平	工作质量	学生发展水平
办学条件		3	3	3
管理水平	1		2	3
工作质量	1	2		2
学生发展水平	1	1	2	
得分 X_i	3	6	7	8
权重值 W_i	0.13	0.25	0.29	0.33

由上表可见,当管理水平与办学条件相比时,我们认为管理水平比办学条件稍微重要,故在①处记3,1处记1;当工作质量与办学条件相比时,我们认为工作质量比办学条件稍微重要,则在④处记3,4处记1。依次类推。将每对指标都比较过之后,可以看到四个指标的得分分别是3、6、7、8,首先将这些得分相加得到总分是24,然后用每项指标的得分去除以总分,就可以得到每项指标的权重值(W_i),各项指标的权重值分别是:办学条件0.13;管理水平0.25;工作质量0.29;学生发展水平0.33。4项指标的权重值之和为1。

总而言之,教育评价方案编制中的加权问题是教育评价中的一个非常重要的问题。加强对加权问题的研究并掌握加权的方法对于我们提高编制教育评价方案的水平,进而促进我们教育评价质量的提高是非常有益

的。

参考文献:

- ①[德]特尔威斯特(Thwaite N.P.)主编,许建斌等编译,《简明国际教育百科全书——教育测量与评价》,教育科学出版社,1992年版
- ②Mcafee, Dralie. Assessing and Guiding Young children's Learning and Development(1997) Akktb&Bacib A Viacom Company
- ③Airasian, P. W. (1994). Classroom Assessment. New York: McGraw - Hill
- ④Marzano, R. J. Pickering, D., & McTighe, J. (1993). Assessing Student Outcomes: Performance Assessment Using the Dimensions of Learning Model. Alexandria, VA: Association for Supervision and Curriculum Development.

(作者单位 北京师范大学国际与比较教育研究所)
(责任编辑 赵福江)