

社会指标运动中法治评价的演进*

屈茂辉 匡 凯

内容提要:法治评价的演进与社会指标运动的四个阶段紧密相连,呈现出主观性和知识多元性不断增强的趋势。理论阶段,法治评价借鉴社会指标理论研究成果,被部分地涵盖在社会指标中或以政策评估的方式出现;其哲学基础是实证主义和科学主义,实践技术是测量,且几乎没有主观性色彩。初期应用阶段,法治评价理论与社会指标研究难以区分,在评价标准和方法模型运用上,也深受后者的影响;其哲学基础是后实证主义,实践技术是描述,且显现出知识的多元性特征。生活质量指标阶段,两类指标体系都扩展了公民满意度调查的视野;法治评价的哲学基础是解释主义,实践技术是判断,且增加了科学的主观指标。社会指标浪潮复兴阶段,法治评价指标体系逐渐成熟;哲学基础是批判主义,实践技术是回应式评估,且通过交流理论丰富了知识的多元性色彩。中国的法治评价可以分为早期的定性评价与后期的科学评价两个阶段,大体上经过了一个由宏观到微观,从整体评价到区域评价的过程。当前,我国在理论上面临着法治理论和评价指标的问题,在实践中面临资料和数据难以达到研究要求的困境。

关键词:法治评价 社会指标运动 主观性 知识多元性

屈茂辉,湖南大学法学院教授;匡凯,湖南大学法学院博士研究生。

一 引 言

社会指标运动自 20 世纪 60 年代在美国兴起后,历经理论研究、初期应用研究和生活质量指标三个阶段。^[1] 20 世纪 70 年代短暂沉寂之后,在 80 年代又掀起了一股研究和应用的新热潮,我们称之为社会指标浪潮复兴阶段。^[2] 政府和研究单位都参与到了这个运动中,利用社会指标来测量社会民主程度、社会福利程度、人们的主观生活态度以及描述社会变迁趋势。社会指标运动作为理念、理论和方法传播的手段,以指标设定、权重设置和指标体系

* 本文为湖南省软科学重点项目“法律制度绩效评价理论与技术研究(2009ZK2004)”的中期成果。

[1] 秦麟征:《关于美国的社会指标运动》,《国外社会科学》1983 年第 3 期。为了区别生活质量指标阶段和复兴后浪潮中的应用,本文特称第二阶段为初期应用阶段。

[2] Wouter van Dooren & Maria P. Aristigueta, “The Rediscovery of Social Indicators in Europe and the USA: an international Comparison”, Paper for the EGPA Annual Conference (in Berne, Switzerland), 30/08 - 03/09 (2005), p. 2.

构建为中心,在不同的学科中得到承继和发展。在法学领域内,基于科学方法的法治评价研究就源于社会指标运动。通过对法治评价演进历程的梳理,我们发现其发展几乎与社会指标运动的几个阶段相重合,所以可以说法治评价深受社会指标运动的影响。例如在立法评价、司法评价、法律实效评价和法治环境评价中的失业率、通货率、居住率、出生率等都是社会指标所关注的内容。^[3]

当然,在充分运用社会指标运动成果的同时,法治评价又反过来促进了这个运动的发展。它们背后的哲学理念、理论设计和方法创新共同促进了这两者的演进。目前,社会指标和法治评价标准已经从单纯的科学性逐渐转向了科学性与合理性相结合。而且,我们通过文献也可以看出,法治评价在社会指标运动中呈现出两个趋势:一是评价和指标设定时主观性不断增强;二是评价理论本身的知识多元性特征也逐渐突出。

二 社会指标运动理论阶段中的法治评价

20世纪60年代中期到60年代末是社会指标运动的理论阶段,同时也是法治评价兴起的阶段,主要表现是法治评价指标被涵盖在社会指标中,并表现为政策评估的形式。在反思经济指标被滥用的基础上,学者发起了社会指标运动。因为用前者度量社会福利和幸福时,容易造成对生活质量、阶层平等和财富积累等方面的负面影响。^[4]单纯的GNP测量无法得出市场价值对社会福利的贡献度、无法得出非市场行为对社会福利的影响、无法得出社会规则体系和经济之间的依存度等重要信息,^[5]所以一类新的指标理论应运而生。

(一) 理论阶段法治评价的主要成果

1966年鲍尔(R. A. Bauer)主编《社会指标》^[6]的出版,标志着社会指标运动的兴起。在书中他首次提出了“社会指标”这一概念,并随后指出应当将多元目标纳入项目考核,^[7]这种系统思维方式迅速被学界和政界所认可。20世纪60年代后,政府部门和国际组织也纷纷参与到社会指标运动中来,其中较为活跃的有美国、奥地利、比利时、丹麦、联邦德国、法国、美国、荷兰、挪威、瑞典、联合国和经济合作与发展组织。

美国作为社会指标运动起源地,其研究具有示范作用。例如在“人口流动和工作更换”的项目中,社会服务研究委员会(Social Service Research Council简称“SSRC项目”)提出了开发“社会指标体系”的要求,进而指出收集各类可得统计数据的必要性。一方面,这些工作初步确定了法治评价指标和社会指标的运行规则,^[8]另一方面,这些研究也辨析了社会

[3] Frank. M. Andrews, “Social Indicators and Socioeconomic Development”, 8 *The Journal of Developing Areas*. 3 - 12 (1973).

[4] Carley, M., *Social Measurement and Social Indicators: Issues of Policy and Theory* (London: George Allen & Unwin, 1981), p. 195.

[5] Gross. B. M., “The state of the nation: social systems accounting”, in R. A. Bauer (ed.), *Social Indicators* (Cambridge, MA: MIT, 1966), pp. 154 - 271.

[6] 从实质意义上来说,社会指标最早出现在1929年由胡佛总统(President Hoover)发起,总统委员会(President's Committee)发表的《当前社会潮流》报告中,在书中他们运用了统计方法来描述美国在1890年-1930年间社会阶层的变迁情况。但是这并没有引起其研究的热潮。

[7] Raymond A. Bauer(ed.). *Social Indicators*, p. 6.

[8] 在20世纪60年代中期,有一系列数据和文章记录了社会科学家发展的想法,这些想法是关于对一套社会指标潜在的运用和过度运用,例如什么指标可得和什么指标需要增加等。其中较为重要的是Bauer的 *Social Indicators* (1966); *Annals of the American Academy of Political and Social Science* (1967, 1970, 1971); HEW's previously cited *Toward a Social Report* (1969) 和 *Foundation The Human Meaning of Social Change and Indicators of Social Change*.

指标、社会统计、社会科学、社会预测、社会报告和社会账簿等相关概念。

在社会指标运动的理论阶段,法治评价的特征包括:第一,法治状况作为社会指标研究的内容之一,只是部分地被涉及,而缺乏单独的研究成果。例如,1971 年 SSRC 项目中,研究者重新收集了 50 年代“底特律项目”中 8 个指标的数据,随后描绘出了底特律地区这 8 个指标所代表现象的变迁路线。通过社会报告的形式,社会测量有效地反映了社会变迁的情况。^[9]而测量社区情况,则实际上从规则变迁的纵向层面上暗合了法治评价的内在需求。第二,法治评价主要体现为政策评估的形式,^[10]而法治评价的最初阶段就是政策评估。现代政策评估从 20 世纪 50 年代出现,持续到 60 年代末,期间的研究重点也发生了转向。最初政策评估是预决算分析,运用投入—产出分析法,主要研究行政投入、行政过程、行政产出和部门效率,而没有涉及到法律和政策效果的评价。此时,政策评估等同于政府绩效评估。但是政策评估因为专注于行政机构解决问题的速度而非有效性上,所以其反而增加了政府运行成本,降低了效率。意识到这个问题后,研究便发生了转向:在美国彼得·罗西、马克·李普西和霍华德·弗里曼合著的《评估:方法与技术》一书中,他们开启了从系统社会科学角度研究评估的历史。书中包括对评估历史、评估内容和评估方法的梳理,重点论述了需求评估、项目理论评估、过程评估、结果评估和效率评估等内容。^[11]此后,美国学者对“政策评估”的内涵和外延有了更深层次的理解。托马斯·戴伊提出了“政策效用”的概念。他指出,政策评估是探寻法律政策实施后会发生什么,着眼点是规则的后果。^[12]在学者努力下,成本—收益、投入—效用理论也被有效地运用到了政策评估中。^[13]

法治评价结合各指标内容与社会指标的相关概念,形成了政策评估的核心概念。例如,社会指标和政策评估几乎可以相互借鉴对“政策”的定义。政策评估中指标设定标准同样也受到了社会指标的影响。例如,政策评价指标的“有限性”、“综合性”、“相符性”、“协调性”和“层次性”标准^[14]都与社会指标运动中理论研究的成果密切相关。所以,利用社会指标运动理论阶段的成果,法治评价(主要是政策评估)在本体理论和方法论上都有了较大的发展。

(二)理论阶段法治评价的哲学基础和实践技术

受到孔德实证主义哲学和自然科学的影响,理论阶段的社会指标和法治评价研究哲学上都保持着实证的精神。政策评估理论认为应当采取价值中立的态度,利用自然科学的方法来科学地认知社会现象和法律现象。而实证主义注重归纳逻辑对“经验层面”进行研究和测量。当然,测量的目的不仅局限于了解现象,更多的是为了探寻现象背后的规律。另外,科学主义中“科学”被狭义的界定为是通过现象可以直接获取规律的自然科学。实证研究则强调通过经验方法揭示“客观”规律,而且,这种揭示是建立在扩展已有经验基础之上,

[9] 这 8 个指标分别是婚姻、性别角色、小孩抚养、社会参与、宗教行为和信仰、价值、政治定向和种族态度等。参见 Otis Dudley Duncan, “Developing Social Indicators”, 71 *Prov. Nat. Acad. Sci.* 5096-5102 (1974)。

[10] 一般而言,评估与评价两者在理论上还是有所差别。前者侧重于对未来情况的预测,而后者则是侧重于对已有现实的判断。虽然如此,但是这两者在方法、技术和模型构建等问题上却是可以互相借鉴的,甚至已经达到了非常高的融合度,以至于在翻译过程中对于“evaluation”可以依据其语境的需要来译作“评估”或者“评价”。

[11] 参见[美]彼得·罗西、马克·李普西、霍华德·弗里曼:《评估:方法与技术》,邱泽奇等译,重庆大学出版社 2007 年版,第 15 页。

[12] [美]托马斯·R.戴伊等:《理解公共政策》(第十版),彭勃等译,华夏出版社 2004 年版,第 280-281 页。

[13] 李瑛等:《政策评估理论与实践研究综述》,《公共管理评论》(第四卷)2006 年第 1 期。

[14] Kenneth C. Land, “On the Definition of Social Indicators”, 6 *American Sociologist.* 25-322 (1971)。

那么归纳与演绎两种逻辑推理方式也就成为其必然的选择。

政策评估主要受实证主义和科学主义的影响。研究者定位于实验者角色,将社会视作是实验室,然后运用自然科学方法寻求和验证各因素间的因果关系。学者主要采用定量分析方法,因为他们认为这能够保证结果的客观化和科学化,显示其中立的价值。随着自然科学、数学的发展,法治评价客观性特征更为明显,它已经由最初单纯的数据列举,发展到计量分析预测和数理模型建构。数理化技术保障了评价的精确性,研究者中立的政治立场,实现了形式的科学性。基于认知的理念,此时的评价方法主要是测量,也即运用问卷设计的方法实现主客观相结合。总结过往经验并设计出问卷,分析问卷中的回答,最终得出对现象的(新)观点。前者是归纳逻辑,后者是演绎推理。测量的是第一代评估的方法,这也被视为是法治评价的实践技术。依据西方政策评估理论的研究,测量主要是评估者利用掌握的工具,对指定的任何调查变量予以测量的一种评价(评估)方式。^[15] 评估者几乎等同于技术人员,他们不断测量法律现象与主观设计间的吻合度,这样评估的过程和结果实际上成为了对假设的检验。例如任职法官知识考核就是一种最常见的法治评价方法。

此阶段社会指标与法治评价在本体理论和方法论上重合程度较大,而且社会指标对后者具有指导作用。因为深受古典实证主义的影响,采用了科学主义的研究进路和测量技术,法治评价客观应用成果也主要表现为测量法治和政策存在。另外,由于孔德古典实证思想将心理学视为是形而上学,所以在指标设定中基本上排除了主观指标和主观因素。

三 社会指标运动初期应用阶段中的法治评价

在社会指标运动理论阶段,研究者在指标理论、概念辨析等方面做出了大量基础性工作。但是,基于科学主义和测量技术的第一代评估也出现了诸多问题。例如它无法稳定社会运动造成的动荡状态;在约翰逊总统的“大社会”计划与“对贫困宣战计划”失败后,研究理论无法解决其带来的消极后果,同样也难以解释其失败的原因。^[16] 这些都促使了初期应用阶段社会指标和法治评价研究的转型。20世纪60末到70年代初,社会指标从理论走向了应用,法治评价也从“实验室”走向了“社会实践”。

(一) 初期应用阶段法治评价的主要成果

在初期应用阶段,美国社会科学研究委员会通过会议、创办刊物等方式积极地活跃在社会指标运动中。其中最具代表性的成果是联邦统计出版物——《社会指标》。^[17] 研究成果主要是改变测量的方法:通过控制和分析社会变迁过程,收集整理现有数据,描述先前空白指标的内容。描述性指标实现了需求标准、模型标准的分类,为最终构建描述型社会指标体系打下了基础。^[18] 另外,社会报告开始逐渐受到重视。虽然奥格本(Ogburn)时期,在描述性指标中就有社会报告的身影,但是现代社会报告是以《社会变迁的指标》(1969)和《迈向

[15] [美]埃贡·G.古贝、伊冯娜·S.林肯:《第四代评估》,秦霖、蒋燕玲等译,中国人民大学出版社2003年版,第5页。

[16] 社会运动中人们通过组建行业协会、公共集会、媒体声明、游行示威等方式表达他们反对越战的态度。这就使得社会处于不稳定状态。参见PAUL MEADOWS,“Theses on Social Movement”,24 *Social Force*. 408-412 (1946)。

[17] Executive Office of the President, *Office of Management and Budget: 1973, Social Indicators* (Washington: US Government Printing Office, 1973)。

[18] Organization for Economic Co-operation and Development, *List of Social Concerns Common to Most OECD Countries* (Paris: Organization for Economic Co-operation and Development, 1973)。

社会报告》(1969)两本书的出版为标志的。自此以后,联邦政府就开始定期公布美国主要社会趋势的统计纲要。^[19]

虽然社会报告是用统计信息描述社会潮流和历史变迁,但是实际上依据不同的主体,报告的侧重点会有所不同。例如,在社会报告目标的解释上,美国政府报告只解释数据的特征和不足;而各种实体行业的特殊报告则偏向于描述指标之间的关系,报告主体甚至还会谈到数据的来源。^[20] 研究者通过大规模的实证调查,分析模型、定义概念和阐明理论,其成果体现了他们对社会指标理解的多样性。指标体系研究对象则主要集中在人口、迁徙、教育、人力资源、家庭等传统的“社会关注问题”上,缺乏人口统计学和社会会计的研究。^[21]

初期应用阶段的社会指标与法治评价(政策评估)研究尚未完全分离。例如,政策会计要求社会指标应当与公共项目相联系,运用输入—输出表格描述政策运行实际状况。“合理的政策使得社会指标与项目的投入成为必要,这意味着社会指标并没有为决策提供所有的定量信息。最终,我们必须综合运用社会指标与政策会计,从而获得所需社会指标来衡量变革,而指标来源则是政策期望的输出标准。”^[22] 当然,这也可以认为社会指标受到了非公共政策因素的影响:“社会状况的优劣并非仅仅取决于特定公共项目。例如,健康状况和寿命长短,除了受公共健康项目影响外,还受到个人的医疗支出、生活水平、健康状况,抵抗力等因素的制约。因此,我们必须辨别出社会指标中的这些因素,从而调整公共项目支出。”^[23]

在研究对象和指标选择上,社会指标与法治评价难以区分,在评价标准和方法模型运用上,后者也深受前者的影响。此阶段中,法治评价已从单纯的依靠经济指标评价法治状况过渡到利用经济理论和社会理念构建评价指标。这意味着,经济状况已经不再是衡量法治优劣的唯一标准,但经济学方法则在政府绩效评价、司法绩效评价和司法管理中流行起来。此外,“法治”概念多元化的理解直接影响到了“法治评价”对象的范围:从单纯的政府绩效评价发展到了立法、司法和政府行为评价,这将法律运行的各个静态阶段都纳入到了法治评价的领域。通过借鉴社会指标的模型和获取指标的方法,法治指标自身也得到了应用和发展。例如,借鉴“目标获取模型”、“侧面影响模型”、“自由目标评估模型”和“相关利益人模型”就丰富了法治评价的方法和技术。另一方面,模型也被视为是分析问题的进路和组织数据的工具。法治指标和法治指数在广义上可视为是一种社会指标和社会指数,这也丰富了社会指标运动的研究对象。于是,法治评价成为了社会指标运动中不可缺少的部分。

在初期应用阶段后期,法治评价逐渐与社会指标运动相分离,出现了独立的“法治政府指标体系”,这可以运用权力制约理论实证地描述和监督法治状况。1968年美国学者伊万(W. M. Evan)构建了一个法治指标体系,他利用70个具体指标描述了美国当时“法治政府”

[19] Zapf, W., “International, Public, and Private Actors in Social Reporting”, 10 *Social Indicators Newsletter*. 4-5 (1976).

[20] 参见 Tauber, C., “America in the Seventies: Some Social Indicators”, 435 *Ann. Am. Acad. Pol. Soc. Sci.* 355 (1978); Tauber, C., “America Enters the Eighties: Some Social Indicators”, 435 *Ann. Am. Acad. Pol. Soc. Sci.* 308 (1981).

[21] 参见 Stone, R (ed.), *Demographic Accounting and Model-Building* (Paris: Organization for Economic Co-operation and Development, 1971); R Stone, Statistical Commission, *An Integrated System of Demographic, Manpower and Social Statistics and Its Links with the System of National Economic Accounts* (New York: United Nations, 1970).

[22] U. S. Department of Health, Education and Welfare, *Toward a Social Report*. (Washington DC: USGPO, 1969), p. 101.

[23] U. S. Department of Health, Education and Welfare, *Toward a Social Report*, p. 5

的运行状况。^[24] 社会指标只是作为观测社会变迁和预测社会趋势的一个手段,而法治评价与社会指标运动在研究目的和研究对象上还是存在一定的差异,所以社会的法治状况不可能完全依赖于社会指标。社会实验方法虽在理论上可以对法治状况进行探讨,但是考虑到社会系统的复杂性,纯粹运用社会指标来研究指标数量增减与法律规章的因果关系,并对少数重要法条之间的因果关系进行科学探究并不可行。因此,研究对象上的专业性缺失是法治评价要求独立的内生条件。

(二)初期应用阶段法治评价的哲学基础和实践技术

在理论阶段的社会指标和法治评价初步实践中,学者已经意识到了实证主义和科学主义方法容易导致唯科学主义的独断性,并且完全采用经济学方法会遭遇到诸多的质疑,所以他们已经开始自觉的实现古典实证主义和科学主义进路向后实证主义的转型。

社会指标初期应用阶段的哲学基础主要是后实证主义。它在形式上保留了实验方法、定量方法以及遵循因果律的特征。其进步之处在于修正了哲学上真理唯一性和同一性理念,转向了柏拉图哲学中关于理念世界和现象世界的划分。也即,研究者认识到理念世界高于现象世界,后者只能无限的接近前者。理解理念世界就必须要从理解多元的现象世界中获得。在法治评价理论中,通过综合各类学科的知识,社会指标和法治评价成为了交叉学科研究的领域。并且,由于无法获得完整的“终极知识”,所以各学科的目标、进路甚至价值都是可以共存的,理论视角和价值框架的多元化得到了方法上的支撑。特别是在统计学中,社会指标与法治评价已经摆脱了简单要素的影响分析,走向了复杂的多变量描述,甚至是在数理上描述和展现复杂变量之间的互相影响情况。在多元哲学思维、新理论和新方法的引导下,社会指标和法治评价指标构建也已经开始走向系统化,并以形成“指标体系”为最终目标。

从学科分类上来看,法治评价的理论来源包括了法学、管理学、经济学、统计学和计量学等学科。从理论构建上来看,法学和管理学是法治理论和评估理论的基础。例如,学者对“法治”和“评估”概念的理解,直接关系到他们研究的范围和对现象的挖掘程度。统计学和计量学侧重于方法支撑,例如,处理资料和数据时,实践者需要依赖于这两个学科的方法和技术。而经济学的作用则是提供连接评价理论和方法的理念,无论是古典经济学的“成本—收益”分析,还是新制度经济学中交易成本概念,它们的目的是使抽象的理论和现象能够得到度量。简而言之,经济学提供了度量转换的工具。

在研究技术上,法治评价已经从单纯的测量走向了描述型评价,后者的目标是描述某些特定对象的优劣模式。在初期应用阶段中,测量被视作是评价和评估的一种工具。^[25] 单纯从理解“实证”出发,描述才是其对应的根本性方法,但是理论阶段的研究受到科学主义影响,已经将现象和规律的理解绝对化了,加之不同学科局限在各自理论视角下,无视互相的矛盾。这些消极因素都促使了研究方法从单纯测量转向描述。在描述方法阶段,评价理论和方法建立在多学科支撑的基础之上,力图展现社会和法律现象的多维度特征。在传统学者看来,这种方式没有提供研究的结论。但是实践者自称,他们秉持的是价值中立的精神,而任务只是为规则制定者提供现象全貌。应当说,法治评价和法治评价指标体系在社会上真实地扮演了启蒙的角色——它们让公众能够了解社会法治现状,理解法律、规则以及监督

[24] 耿浩飞:《法治政府指标评估体系研究》,《行政法研究》2012年第1期。

[25] [美]埃贡·G.古贝、伊冯娜·S.林肯:《第四代评估》,第7页。

法治主体的行为,这客观上起到了将“集团”或者“团体”运作透明化的作用。而且,研究者已经开始运用定期社会报告的方法及时公布法治运行状况。例如,规则制定者借鉴社会报告中描述的法治发展状况,推理得出社会综合发展趋势,把握各类重点关系,进而制定科学的规则。在此可以看出,初期应用阶段的指标和评价已经开始注重综合运用交叉学科知识,开始在实践中融入多元化思路,并打破了经济学帝国主义的约束。

四 社会指标运动生活质量指标阶段中的法治评价

进入 20 世纪 70 年代后,国际社会开始反思政策与法律的制定目标和运行准则问题。社会运动加速了福利国家的出现,指标也从效率标准转向了公平标准。在世界各国关注公民福利的时候,社会指标与法治评价也开始研究与公民日常生活相关的问题。社会科学研究团体发起的“学术性的改革运动”将传统社会“物质—经济视角”扩展为“幸福观视角”,进而以此驳斥物质主义者的物质福利观。社会指标运动则顺应时代潮流,进入了生活质量指标阶段。

(一) 生活质量指标阶段法治评价的主要成果

本阶段的研究领域从观测社会变迁转至评价公民福利。早期测量社会变迁中关注的是文化变迁趋势,指标包括:主流文化、传统主流文化的载体、主流文化的支持者、主流文化的参与者、主要团体等。后来,受社会学“失范”理论影响,美国学者利用当时最先进的时间序列指标对“社会健康”问题进行研究,指标体系中包括了政治、选民、经济环境、生态、劳工趋势、工作满意度、宗教、人口预测等具体指标。通过借鉴和回归社会健康理论,他们以社群为研究单位,将视野扩展到单位人的生活质量领域上,例如雇佣、收入、健康、住宅和教育等都是当时的热门主题。而且,随着社会指标理论的成熟,学者针对不同的主题开发了相应的指标类型。例如信息型指标就适用于描述健康地位、休闲、公共安全、社会福利主题;预测型指标适用于雇佣工人的工作分配、家庭构成、家庭需求、学校招生主题;问题导向型指标则对特定区域群中婴儿死亡率、犯罪率、贫困率的比较分析较为适用。^[26]

生活质量阶段的指标体系特征是增加了大量的主观指标。《人类社会变革的意义》文集主要讨论了理解社会变迁过程中对态度、期望、感觉、激情和价值等社会心理密切关注的必要性。^[27]“基于主客观联系不完备理念,我们确定了这本书的研究范围。为了解对象的生活质量,必然需要其直接描述自身的感受。虽然并没有精细地定义‘生活质量’这个概念,但是其理应当包括工作成就、美的感受、群体归属感以及个人潜力的被认同感。”^[28]于是,社会指标通过测量满意度、幸福度和生活满足感来直接确定“主观现实”状况。

实际上,法治评价中主客观联系不但不完备,甚至是达到了难以明确的地步。埃斯特林(Easterlin)^[29]和邓肯(O. D. Duncan)^[30]的研究指出,就单个研究对象而言,收入与满意度相

[26] Frank. M. Andrews, “Social Indicators and Socioeconomic Development”, 3 - 12.

[27] Campbell A & Converse. P E (eds), *The Human Meaning of Social Change* (NY: Russell Sage Found, 1972), p. 547.

[28] Campbell, A., Converse, P. E., Rodgers, W. L., *The Quality of American Life: Perceptions, Evaluations, and Satisfaction* (NY: Russell Sage Found, 1976), p. 1.

[29] Easterlin R. A., “Does money buy happiness?”, 30 *Publ. Int.* 3 - 10 (1973).

[30] Duncan. O. D., “Does money buy satisfaction?”, 2 *Soc. Ind. Res.* 74 - 267 (1975a).

关。但这个命题无法类推适用到群体研究对象上。因为 GNP 每提高一个单位,国民的平均满意度并没有相应地提升。所以从社会理论的观点来看,满意度是否与期望值、兴奋度、态度或者价值等指标相关还难以确定。特别是在法治评价的语境中,个人和集体的满意度同时最大化的状况是无法达到的。^[31] 因为个人的满意度会随着融入到群体中而被降低。与个人生活满意度最相关的领域是法治环境中的基本行为,如警察的执法行为,法官在审判时的表情和语调等,这些都无法通过累加的方式类推得出群体满意度。

(二)生活质量指标阶段法治评价的哲学基础和实践技术

在生活质量指标阶段中,法治评价指标体系构建方式、法治评价方法以及观念都有所转变。前两代法治评价侧重于事实领域,但是对实证和“事实”地过渡推崇使得科学主义反而掩盖了客观实在。学者将科学理性理解成了工具理性,所以价值取向和价值判断才变得多余。这种缺乏了内在目标的法治评价已经失去了“评价”本身的含义。在对工具理性下的现实与目的相分离、科学主义中过度崇尚技术进行反思之后,研究者开始意识到了价值的作用。

本阶段受解释主义哲学的影响,研究者认为社会现象和法律现象都不具备普遍性。社会现象是由诸多不确定因素的合力形成的,而这些因素很多都是源于人的非理性行为。另外,各人对客观世界的理解也不同,社会和法治运行的状况是否符合规律已经显得不再那么重要,重要的是社会规则和法律规定的效果。这导致了政府角色由社会控制者转向了社会服务者,公权力机构在运行中也开始注重公民的感受。

解释主义主要体现在对现象进行测量和描述之后,运用社会科学、法学等学科理论来进行评价和判断。因为多元理念下的法治评价需要以学科理论的合理性解释作为支撑。这里没有唯一正确的真理,有的只是解释力强弱的评价。研究者运用各自理论解释指标所代表的现象状况、发展历程及其意义。虽然他们最后难以得出统一的结论,而且在形式上也丧失了科学唯一性的特征,但是从更为广阔的角度上来看,纳入了相对合理性的解释更加符合实际状况。指标设计和法治评价的过程中,增加了被调查者和利益相关者变量的输入机会,这客观上使得最终颁布出来的法律和政策成为多种力量平衡的结果。

从评估的发展来看,此时已经进入了以“判断”为特征的第三代评估。实证主义被实证方法所取代,研究者抛开了价值祛除的理念,开始将价值分析与规范分析相结合。测量和描述在此已经成为了“判断”中的技术和工具。而价值分析则是主观指标的评价标准,以此来对现实中的情况进行优劣排序。受到罗尔斯观点的影响,此时的评价标准已经从效率转向了公平。

在方法上,研究者引进了主观指标。因为:第一,主观指标能有效地测量满意度;第二,主观指标与地区政府服务评价密切相关,它们基本上能够反映行政计划和区域行政服务实效;第三,主观指标能够防止“文化霸权”的产生;^[32] 第四,主观指标能够辨别广义上的社会问题,它们作为社会基本定位的工具之一,可以用来重新分配资源,这也就为政策制定者提供了更广的视野和更新的服务对象。

[31] Land, K. C & Juster, F. T., “Social Accounting Systems: An Overview”, in F. T. Juster & K. C. Land (eds.), *Social Accounting Systems: Essays on the State of the Art* (USA: Academic Press, 1981), pp. 1 - 22.

[32] 这种文化霸权则是被解释为:行政主体通过“家长式的作风”将他们的理念和推理方法灌输到社会其他成员的思想与行为中。

当然,运用主观指标也有几个问题:第一,研究者还不习惯用心理学技术来保证数据的信度,特别是他们经常引用的感知数据很容易受到被调查者即时环境的暗示,这将影响到资料收集的稳定性。例如,突发性非重要性事件可能会改变最终的决策。而且主观态度也取决于一系列潜在的事件,例如焦虑性或者骚扰性事件。^[33] 第二,被调查者可能会给出误导性的答案。面对问卷中的价值性问题,被调查者可能会选择正面答案以显示他们为人正直。第三,人们对问卷理解的角度不同。例如对司法效果的主观感受,有人关注于法官是否遵守程序,有人则关注于法官在庭上对双方的态度。

本阶段的特征是将主观指标纳入到指标体系中,这实质上加重了主观要素在整个体系中的比重,通过引入外部多元观点弱化指标设计者的前见或偏见。在最终成果的报告上,研究者和实践者通过对指标体系及指标的研究,向立法者和政策制定者提出自己的建议,这具有了第四代评估“交流”和“沟通”性质的雏形。

五 社会指标浪潮复兴阶段中的法治评价

20 世纪 70 年代至 80 年代,由于经济危机导致大量管理人员被裁减,社会指标运动也失去了其经济基础和人力资源,所以一段时间内社会指标运动和法治评价活动都步入了低潮。20 世纪 90 年代,随着经济的复兴和法治评价在社会服务等方面重要性增强,社会指标和法治评价也得到了复兴。

(一) 社会指标浪潮复兴阶段法治评价的主要成果

社会指标复兴阶段的研究不能称为运动,而必须称为“浪潮”。因为此时社会指标和法治评价虽与先前各阶段的研究在对象上有重合性,但是其发展程度和速度却是其他阶段难以企及的,此时的研究包括了概念的重新定义,模型和方法的运用,以及在生活质量指标基础之上进行的其它研究。

在基本理论方面,学者对社会指标进行了重新定义,“在描述社会体系的理念核心上,社会指标只是个技术工具。”^[34]一方面,这意味着为了操作的方便,社会指标可以运用技术将抽象概念转为具体概念。另一方面,社会指标概念涉及到了社会体系的理念核心——政治领域,所以其必然要有政治上的目的。

在具体运用上,对“浪潮”一词最为有力的解释是在欧洲开展的系列社会指标研究和应用的活动。2000 春,里斯本欧洲委员会制定了一个十年计划,项目包括对 35 个主要指标进行监控和考量,其内容与就业、创新、经济改革、社会和谐度紧密相关。2001 年 3 月,斯德哥尔摩欧洲委员会年度综合报告扩充了上述 35 个指标内容,在 2002 年向巴塞罗那欧洲委员会进行汇报时,报告已经构建了包括了就业、创新和研究、经济改革、社会和谐度、环境以及一般经济背景在内的 42 个指标。2003 年,布鲁塞尔欧洲委员会春季报告的指标内容虽然与以前一样,但资料收集的范围却扩展到了欧盟候选国家中。2004 年,欧洲委员会发布了 100 个二级指标,它们都作为指标子集分配到了 14 个构建性指标中。^[35]

[33] A. McKennell, “Surveying Subjective Welfare”, 74 *Strategies and Methodological Considerations*. 135 - 159 (1982).

[34] Vogel, J. “The Future Direction of Social Indicators Research”, 42 *Social Indicator Research*. 103 - 116 (1997).

[35] Wouter van Dooren & Maria P. Aristigueta “The Rediscovery of Social Indicators in Europe and the USA: an international Comparison”, pp. 8 - 9.

在生活质量指标的研究中,研究者已经设计了“问卷模板”来解决主观指标构建中存在的问题。例如,自我健康报告主要是通过 SF-36 表的 8 个维度来测量对象的健康状况。对同一组对象自测健康的解释中,如果采用整体福利理论进行预先判断,能够解释 56% 的变量;而如果使用多重差异理论,则能够解释关于满意度中 51% 的变量。^[36]

政府机构、社会组织对部分地区的法治发展状况也进行了有益探索。值得一提的是“世界正义工程”(the World Justice Project),这是由美国律师协会(ABA)、联合国国际律师协会、泛美律师协会、泛太平洋律师协会等组织发起的一个项目。他们于 2008 年 7 月出台了旨在促成全球统一规范适用,衡量国家法治程度的“法治指数”。该指标体系通过将行政、立法与执法过程浓缩为联合国公约的 13 个核心原则,再对每项原则制定 3—7 个具体指标,形成了一套包含 60 个指标的评价体系。特别是该套指标力求通过反映法治过程中的共性,以确保其在不同经济、文化、社会、政治背景的国家或者地区都能够得到适用。^[37] 还有些组织力图通过定量研究方法,从治理指数、腐败指数和民主状况等领域来反映区域法治水平。如:世界银行的“善治(good governance)”指标体系、美国国际开发署支持的“拉美舆论项目”、世界银行发布的“世界治理指数”(Worldwide Governance Indicators)、联合国奥斯陆治理研究中心的“民主治理测评体系”(Measuring Democratic Governance)、由非洲公共调查机构 Afrobarometer 发布的“Afrobarometer 指数”等。此外,诸如非洲治理指数、经合组织(OECD)的“人权与民主治理测评”指标体系、透明国际的腐败认知指数、全球清廉指数等,都把法治指数作为其中一项重要内容纳入了指标体系之中。

应该说,法治评价在指标研究的层面上是属于社会指标运动(浪潮)的一个部分,而且社会指标运动与法治评价的发展阶段也基本重合。但是这两者还是不尽相同,其中最明显的就是多元观念接受的程度。社会和社会理论的多元性使得社会指标运动在很大程度上是方法论与理念的扩展,通俗的说,这是一种思维进路和问题意识的延伸。所以,如果国际组织利用社会指标进行区域排名的话,并不会遭到太多的非议。而法治评价在一定意义上是具体研究细节的扩展,是文化多元与绝对正义之间在理念上的冲突,所以在利用正义指标、法治指标进行区域排名的时候则会引起非议。^[38]

(二) 社会指标浪潮复兴阶段法治评价的哲学基础和实践技术

从法治评价的方法角度而言,其主要经历了从实证主义导向到批判主义导向的转变,也即,从实证导向到规范与实证的综合导向。^[39] 批判主义理论实际上带有怀疑论的影子,它要求从经验,现象的认识,以及研究对象主客体三者之间的互动得出评价结果,这需要建立在对科学和历史经验认识的基础上。并且在公正性评价的过程中,需要不断的在正反双方立场、语言逻辑、指标及其目的之间进行沟通和检验。这种建立在哈贝马斯交流理论之上的公正确保了法治评价过程中程序的民主性。另外,此阶段对评价也有更为科学的理解。由先前的自然科学理解,到现在科学多维面向的理解,后者在普遍性和统一性基础上扩展了解

[36] Najman, J. M and S. Levin, "Evaluating the Impact of Medical Care and Technologies on the Quality of Life: A Review and Critique", 15 *F Social Science and Medicine*. 107 - 115 (1981).

[37] 参见世界正义工程: *The WJP Rule of Law Index* (2010), <http://worldjusticeproject.org/rule-of-law-index/>, 最后访问时间: 2012 年 3 月 15 日。当然,报告的具体内容肯定会存在争议。

[38] 例如中美双方针锋相对发表的“人权报告”。

[39] 一般认为实证导向向批判实证发展到批判实证实现的是实证向规范(或者价值)的转移,但是从方法上而言,其是从完全的实证(价值中立)走向了混合实证与价值结合(价值作为目标,而实证则是证明的手段)。

释性与相对性。这在哲学上基本排除了科学主义的负面影响,在方法上彻底地打破了“经济学帝国主义”的僵局,保障了知识的多元性。并且解释性和相对性赋予了“法治”区域性的特色,这为国家法治程度和法治评价指标的选择提供了理论支持,同时也为区域法治建设评价提供了哲学指导。对于解释而言,一方面是在现象和理论之间构建沟通的桥梁,另一方面,也是构建理论自身的一种方法。

在方法论上,受到多面向理解科学的哲学影响,此阶段其方法则已经由单纯定量描述和评价方法走向了混合方法。混合方法可以粗略的视为是实证研究方法和诠释学等方法的结合。其三个步骤包括:一是确定理论结构。二是确定如何实施数据收集。三是确定切入和整合数据材料的方式。“混合方法”的运用掀起了被视为继定性和定量方法之后的“第三次方法论运动”。^[40]

在方法和技术运用的层面上,第四代评估理论深刻的影响了社会指标浪潮和法治评价。它最明显的特征是交流与沟通,包括了研究者与研究对象之间,利益关系者之间,法治评价者与法律政策执行者之间的协调和互动。在不断的输出和反馈中,最终得出评价的结果。第四代评估方法中的法治评价形成了具有自身特色的“回应式法治评价”。这种评估采用构建方式,运用诠释、分析、定量等方法来进行评价。而且,回应式法治评价吸收了前几代评估理论的成果和认识。测量与描述成为了评价中的步骤,评价已经从绝对的科学主义转型到了科学性和合理性相结合。在理论解释中承认了社会和法律现象的多元存在,而指标也从单纯的客观设定转变为主客观相结合,这实际上意味着研究者已经实现了观察者到参与者的角色转变。

此阶段中,哲学基础中包含了罗尔斯“相对正义”思想,而在具体的方法论运用层面上则是采用的哈贝马斯的交流理论。社会指标设定与法治评价两者都在很大程度上注入了主观色彩,特别是在问卷设计、对象访谈上,混合方法运用得较为广泛。在具体技术使用上则体现出了各个层面上的沟通,促进了评价和指标理论本体知识构成的多元性。

六 我国社会指标运动中的法治评价

中国的社会指标运动起步较晚,由于历史的原因,改革开放后学界才开始着手引进和建立社会统计学。1982年,联合国教科文组织继在亚、非、拉一些国家和地区举办了各种有关社会经济指标的讲习班之后,在中国社会科学院主持下于北京举办了一期“社会经济指标研究班”。在这个研究班上,第一次对国内研究者进行“社会指标”系统化的介绍和阐述。^[41]随后,社会指标运动在中国得到快速的发展,目前社会指标体系已经成为包含各类满意度指标、法治指标、社会经济类指标等多种类型指标在内的综合性指标体系了。

中国的社会指标运动承继了美国“理论研究阶段”传统后,直接进入了“实践运用阶段”。在指标运用过程中,我国借鉴了美国的社会指标理论、技术操作理论和社会理论。所以在指标的选取、数据调查、模型构建、理论解释,甚至刚兴起时的研究内容都与美国几近相同。社会指标运动第二个和第三个阶段在中国具有一定的重合性。初期应用阶段与生活质

[40] 蒋逸民:《混合方法研究:方法论研究的新取向》,《中国社会科学报》2010年8月5日,第A-05版。

[41] 陈立新:《社会指标与社会协调发展》,湖南大学出版社2004年版,第85页。

量指标阶段区别不明显,所以法治评价的应用和法治环境评价两个阶段(主要是通过引用主观指标)区分也不明显。目前关于法治评价理论介绍和实际运用的文献几乎都强调了主客观评价相结合、主客观指标并存等内容。受到混合研究方法的影响,中国的法治评价也正向着“回应式法治评价”的方向发展。

国内的法治评价理论来源也不外乎于社会指标运动的促进,法学、管理学、经济学、统计和计量学知识的综合借鉴等。需要注意的是,法治评价舶来品的性质使得它可以与国外理论迅速对接,这大大提高了其发展的速度。所以国内法治建设评价的理论发展历程在很大程度上是国外法治评价理论的移植过程。

中国的法治建设评价大体上经过了一个由宏观到微观,从整体评价到区域评价的过程。早期的评价受到美国政府绩效评价的影响,主要是针对我国政府绩效进行评价,几乎完全采用了美国预决算方法对政府运行和政府行为进行监督和预测。整个过程中,经济因素占有主导地位,因为预测(也可以被称为评估)是通过政府对政府运行透明度的监督来达到合理制定预算的目的;从社会指标运动的角度来看,其尚处于运用经济指标的初级阶段。到了第二阶段,随着立法评价、司法评价、司法管理评价逐步纳入到法治评价的视野中,法治环境评价才渐渐成型。此时的法治评价偏向于宏观层面的研究,也即以大范围(例如全国)数据作为研究基础。本阶段最具代表性的成果是北京大学实证法研究中心“北大法意”数据库的推广和朱景文主编的“中国法律发展报告数据库和指标体系”的成形。这两者共有的特征就是“全面”,无论是在研究对象还是在区域上,都涵括了较广的范围。第三个阶段是区域法治评价。上一阶段中,范围广、摊子大,这在某种程度上使得法治评价成为了轻实质的形式研究。例如为了形成立法、执法、司法和守法的数据结构,他们不得不借助官方的统计数据,甚至在某些数据库没有及时更新时只能发布历史数据。这极大地影响了资料的信度,对结果的效度也产生了负面作用。因此,通过对地区数据的有效收集,虽然某种程度上舍弃了统一性的形式,但是通过实地调查等方式确保了数据和资料的真实性,这为后续的评估工作打下了坚实的基础。

第一个阶段最重要的成果就是将“实效”或者“绩效”的概念引入到了国内,逐渐形成了系统的法治评价理论,这大大扩展了研究和执政者的视野,为中国法治建设提供了一个新的理念。当然,从知识学角度来说,新学科、新知识初期阶段的任务主要在于传播理论和理念,所以应用上并没有自己的创建。第二阶段宏观层面的评价研究,意图以一种高屋建瓴的方式来对国家层面的法治状况进行概览和评价,它立意虽高,但是缺乏相应的方法和技术,导致有效数据空白面较大,反而降低了科学性和可信度。第三阶段从小角度切入,保障了数据的信度和效度;此时的问题在于数据收集的口径难以统一,研究成果与外界难以沟通。学者虽然也会参与行政机构组织的法治评价,但受到政府立场的影响,难以保障其以民间学者身份评价时的中立性。

上述三个阶段并没有出现非此即彼式的断裂性研究,各个阶段既有侧重,又互相融合,中国法治评价已经在研究群体意义上形成了自己的模式。虽然我国没有如同国外那样形成法治评价方法的三个明显时期,但是从总体上而言,也可以分为早期的定性评价与后期的科学评价两个阶段。

受传统法学以及评价科学本身特性的影响,我国早期法治评价研究以定性方法居多,忽视了对定量方法的运用。例如当时的法学类文章在批判和构建制度时多采用价值分析、规

范分析和逻辑分析等定性方法,较少地从交叉学科、综合性研究角度运用计算机、统计学、运筹学、计量经济学等学科的定量技术来分析。而且当时的研究只注重考虑经济、技术因素,较少考虑政治、社会因素。法治评价局限在了对法律制度和法律条文的影响估计上,缺乏了整体性的评价。

随着中国法治进程的不断进步和政治经济体制改革的深入,法治评价开始运用经济审计,目标责任考核和公民评议等方法。国外先进经验的介绍、自然科学的发展、信息技术的进步以及计算机应用的推广都为法治评价提供了科学、客观、可信的技术基础,从而使得政策评估摆脱了过去以定性分析为主、作用范围有限的局面,逐渐呈现出多样化、系统化和数理化的趋势。从评价方法的发展阶段而言,我国未经过测量阶段而进入了描述阶段,并且直接适用国外先进模型,用数学方法,计量方法等进行政策评估和法治评价。较有影响力的研究包括了:中国法学会“法治建设状况评价指标体系”、法治昆明综合评价指标体系、北京市哲学与社会科学“十五”规划研究项目“城市法治环境评价体系与方法研究”、蒋立山主编“现代化法治城市评价”、浙江法治环境评估体系、香港法治指数、张志铭等编写的“世界城市的法治化治理”著作以及钱弘道组织的余杭法治评价。此外,为了进一步贯彻落实国务院《全面推进依法行政实施纲要》,加快我国法治政府进程,据不完全统计,全国内地 31 个省(自治区、直辖市)中,已经有 22 个出台了相关文件,开始了推行法治政府和法治社会的工作。

虽然目前我国在理念上已经接受了法治评价,但是在理论和实践中,还有诸多问题,具体包括了理论与实践两个方面。第一是法治评价理论上的问题。这又包括了法治理论和评价指标两个方面。其一,在法治理论上,难以处理“法治”与“自治”的关系,以及对“法治”的过度依赖。例如,如果将“法”限定在法律和法规的层面上的话,村一级更加倾向于“自治”,因为其规范和运行更多的要遵循乡俗民约。一旦盲目的按照规范法进行评价的话,那么村级的法治指数将会很低。另外,考虑到法治的统一性,“地方法治”概念本身就具有争议。其二,在评价指标理论中,由于指标设定理论论述不明确,这会导致出现重复测量和空白测量的情况。例如,区级政府无财政权,如果要对法治状况、政府运行透明度进行评价时,便会遇到困难。权力机关运行的透明度状况在深层次意义上远远比其研究中所倡导的公民“法治感受程度”重要。因为前者与法治运行的底线密切相关,而满意度调查则倾向于是表面层次的法治服务状况。

第二是实践过程中存在的问题,主要体现为资料和数据难以达到研究要求。其一,在法治评价中,人员构成是研究的内容,而且可以将性别、年龄、民族、党派和学历等作为研究法律职业构成的变量。但是在实际中,各单位人员构成的公开程度不一。例如全国人大代表和人大常委会委员的资料和法官、检察官、律师的资料完整度就不可能一样,前者能够通过正式的渠道获取,而后三者只是零散的和系统的资料,散见于报纸、杂志和书籍中,难以从官方的渠道完整获得。其二,经费的管理和支出监督应该是法治评价内容中一个非常重要的部分,但是研究中却缺乏相关领域的调研材料。在提出法治建设后,国家对立法、执法、司法和法治环境建设等方面的投入都很大,但这一部分经费公开程度却不高,以至于无法从财政的角度评价法治水平和法治进程。其三,由于缺乏资料,经费计算方式也是研究的空白领域。例如要对法院的审判绩效进行分析的话,就缺乏科学计算方法来获取法院的投入成本。

现代法治评价研究发展的趋势之一就是研究方法的综合化,它作为一门介于自然科学

和社会科学之间的交叉性学科,其研究方法也必然体现出系统性和综合性的特点,无论是单独的定量研究还是定性分析显然都不能满足法治评价科学性的发展要求。因此,引入客观分析,渗透主观分析,采用混合研究方法才能更好地促进法治评价的科学发展,从而增强法治评价的科学性和有效性。

[**Abstract**] There is a closely connection between the progress of assessment on rule of law and the four steps of social indicators movement. The assessment on rule of law utilized the researching achievement on the theory of social indicators during the theoretical steps. The indicators of rule of law was partly covered by social indicators or appeared as policy evaluation. The philosophical foundation of the assessment in that time was positivism and scientism and the technique was measurement and there was rarely subjectivity in this step. It was difficult to tell the difference between the research on social indicators and the theory of assessment on rule of law in the early practical step. Besides, the latter was deeply influenced by the former in the areas of the standards and the models. The philosophical foundation of assessment in that time was post-positivism and the technique was description. Further more, there was plurality of disciplinary in the theory. Both of the indicators system of assessment on rule of law and the social indicators system were expanded the range of vision by adding the scientifically subjective indicators during the step of life quality indicators. The philosophical foundation of assessment in that time was hermeneutics, the technique was judgment and the scientific subjective indicators were added. The indicators system of assessment on rule of law was mature in the recovery step of social indicators movement. The philosophical foundation of assessment in that time was criticism; the technique was responding evaluation and the plurality of knowledge is developed by communicating theory. The assessment on rule of law in China can be divided into two steps which were early quality evaluation and latter scientific assessment from the perspective of methodology. Moreover, it roughly experiences the macroscopic view to the microscopic one, which means the overall assessment to the regional assessment. Nowadays, the Chinese assessment is exposed to the problems of theory of rule of law and the assessing indicators in the theoretical aspect and disqualification of material and data in the practical aspect.

(责任编辑:支振锋)