

法治评估专题研究 (栏目主持人: 李瑜青)

在全面推进依法治国的背景下,本刊以“法治评估”为主题刊发四篇有理论价值和实践意义的文章。通过对余杭法治指数的表达与余杭法治实践相互关系的分析,《法治评估实践的效度》一文认为,余杭法治评估在效度存在事实与宣传的背离,评估技术不完善、评估缺乏有效接纳是导致问题产生的原因。针对实践中各级法院在创新社会管理方面呈现的模仿性与随意性趋势,《我国人民法院创新社会管理的司法法治逻辑考量》一文认为,有必要明确将司法职业责任伦理作为我国人民法院创新社会管理的理性考量,通过遵循法治建构的逻辑来保障创新社会管理功能的实现。对宪法基本概况、宪法权利及宪法义务等知识的掌握情况开展的调查显示,东部地区和华南地区虽然在经济上具有优势但并没有显现出应有的法治优势。《地域差异与公民宪法意识》一文认为,有必要对先发地区法治优先的判断加以检验。《论法治评估的立场与类型》一文对法治评估的类型及特点展开了研究,将法治评估区分为国家视角、社会立场及专家视阈三种类型,三者的侧重点分别为法治目标实现状况、社会利益表达及法治理念的落实。

法治评估实践的效度

——基于“余杭法治指数”的文本研究

李瑜青,张 玲

摘 要:“余杭中泰垃圾焚烧厂事件”的发生撕裂了法治评估与法治现实,令法治评估效度成为重要问题,引起人们关注。所谓法治评估的效度,即法治评估的有效性,包括内容效度和结果效度两个层面。通过对指标体系构建、评估信息收集、评估数据处理、评估报告结论等余杭法治指数报告核心内容的文本研究发现,余杭法治评估实践存在评估的效度上事实与宣传的背离。而这一情况是评估技术不完善、评估缺乏有效接纳等内外因共同作用下结果。

关键词:法治评估效度;余杭法治指数;余杭事件

作者简介:李瑜青,华东理工大学法学院教授,博士生导师,中国法律社会学专业委员会副会长;张玲,华东理工大学法学院博士研究生。

基金项目:中国法学会重大研究课题“法律实施的保障机制研究”(CLS2011A04)。

中图分类号:D902 **文献标识码:**A **Doi:** 10. 3969/j. issn. 2095-042X. 2016. 05. 005

一、“余杭中泰垃圾焚烧厂事件”与余杭法治指数

2014年5月10日,浙江余杭爆发了一次震惊中外的群体性事件——“余杭中泰垃圾焚烧厂事件”(以下简称“余杭事件”),即“民众为反对杭州市余杭区中泰乡九峰村生活垃圾焚烧发电厂项目建设,

封堵高速公路省道、打砸车辆等违法事件”^①。事件起源于2014年4月,杭州市政府面对“垃圾围城”的难题做出了在余杭九峰村建造垃圾焚烧发电厂的项目规划。该工程项目规划一经公示,民众便因存在身体健康、环境质量、资产价值等多重负面影响而采取集会等方式进行反对。随后,政社双方经由听证会、信息发布会等方式维持一定频率的沟通,但是总体来说政府回应缓慢滞后,民众的担忧与质疑一直未得消解,最后在一根导火索——“政府未经民众认可暗自开建垃圾焚烧厂”的谣传——的刺激下发展成了一件规模性聚集(5 000+人)、性质恶劣的群体性事件。而在事件当场更有部分不法分子制造传播谣言,采取打砸车辆甚至围殴民警等暴力方式来宣泄情绪,导致现场秩序一度混乱,直到2014年5月11日凌晨秩序才基本得以恢复。在事件的后续处理中公检法司联合发文惩治不法分子,官方通过大众媒体加强政社沟通并作出“项目在没有履行完法定程序和征得大家理解支持的情况下,一定不开工”的声明,由此该群体性事件的危机警报才逐渐得以解除。

“余杭事件”的群体聚集性渐次消去,但随后媒介平台对该事件的反思报道继续,一时间普通网民、评论员、学者纷纷参与该事件的讨论。如《凤凰网评论·自由谈》栏目,以“余杭事件:邻避思维难免邻避冲突可免”为题展开了专题讨论。大众网也有“余杭‘5·10’事件分析”这样的全视角冷静分析。可以说,讨论热度很高、信息量很大,有的认为是民众问题,认为民众的暴力行为体现了民众利益表达方式不理性,公民维权意识的扭曲^②;有的认为问题在政府,地方政府未及时公开信息是该事件的导火索^③;也有认为问题是混合的,民众与政府之间沟通不畅,尤其是民众缺乏对政府的信任^④。

以上三种问题分析所下的结论,其功能主要停留在提高关注度兼普及知识的层面。随着新一轮热点事件的发生,媒体对该事件的探讨便渐入终结,但是该事件的深度反思显然尚未完成。而对“余杭事件”的反思态度归根到底是取决于如何定性该事件,即到底是将该事件当成一个简单的偶发性事件,还是余杭法治建设中的必然事件?笔者认为“余杭事件”是余杭法治的一个缩影,民众、政府、媒体在此交锋,是余杭法治条件下的必然产物,无论是发生原因还是处理过程都可用法治尺度去衡量或评价。从政府透明度、民众守法、执法程序规范化等法治角度予以考察不难发现,“余杭事件”发生前后呈现的确是政府信息低度透明、民众守法意识薄弱、行政程序不规范等法治基本要求不达标的事实,不由得令人联想到“余杭法治指数”。

作为中国地区法治指数的首发地,余杭自2008年公布第一份余杭法治指数至2014年“余杭事件”发生,余杭已连续进行了6次法治评估。虽然在评估启动初期余杭法治评估也曾遭受过一些质疑,但更多是以一种发挥积极作用的形象展示给外人,并引领法治评估实践在其他省市地区蓬勃兴起。2007年至2012年的余杭法治指数分别为71.60、71.84、72.12、72.48、72.56、73.66。这一系列总体指数能反映当年度法治基本状况,给人们一个整体印象,令人一目了然。尽管起初还有人质疑分值不高法治不理想,但课题组给出的解释是在香港75分已是高分,所以余杭法治指数的分值充分说明余杭良好的法治状况,从而从总体上肯定了余杭政府的法治绩效。而法治指数也并非简单的绝对值的概念,而具有相对值的意义,其突出表现是为进行历年法治建设状况的对比提供参考。逐年递增的法治指数向公众展示了余杭的法治

① 百度百科:余杭中泰垃圾焚烧厂事件, http://baike.baidu.com/link?url=w9JDSdlR1yGGif2-yi0EdLkk6L-roULiLUjSe3YNoxNoA7yJJsx7kwDq86nmB8EykatqXbs0C-_sxIyxKyJqLa, 最后访问日期2016年7月17日。

② 晏佳伟:余杭事件,请理性表达利益诉求, <http://focus.cnhubei.com/original/201405/t2924365.shtml>, 最后访问日期2016年7月17日。

③ 新京报社论:“垃圾焚烧厂争议”需要冷静和理性 http://news.ifeng.com/a/20140512/40251250_0.shtml, 最后访问日期2016年7月17日。

④ 矫志欢:“余杭事件”背后的互信与沟通, http://www.cnepaper.com/lrb/html/2014-05/12/content_2_2.htm, 最后访问日期2016年7月17日。

状况一直保持着上升趋势,余杭政府在法治绩效方面始终追求更好的有力证明。

法治指数所塑造的良好政府形象,为余杭获得了上级的认可,以2010年为例,余杭司法系统收获了一系列国家级荣誉^{[1]255},如“全国优秀法院”“全国先进基层检察院”“全国司法所建设先进单位”,可见法治指数能够助力余杭地区在全国的评比,已经成为展现余杭司法水平的一项重要象征性标志。可见,法治指数的积极效应是非常明显的。但是,对于法治指数明确的绩效式定位,也出现过质疑之声,质疑政府是在花钱买数字^[2],质疑法治指数的有效性^[3],质疑法治指数缺乏含金量及评估的形式化^[4]。对于这些质疑,余杭法治指数的课题组也做过总结,诸如“没有给出得出这一结论的证据”“对余杭法治评估的质疑多发生在2008年上半年,此后质疑声减弱”,并给出了质疑声减弱的三点原因,包括公众的理解、评估分值与研究文献的公开,以及余杭法治指数具有对其他省市的导向意义和示范价值,从而否定了余杭的法治指数是一场“学术盛装和政治作秀”^{[1]264}。

但是,“余杭事件”所暴露的法治问题,与蒸蒸日上的“法治指数”形成了鲜明的对照,不免令人再次质疑指数的价值,以及法治评估在余杭发挥的实际作用到底有多大。经由该事件暴露的法治问题显然不是突然发生的,而是长久存在于体制内的,这些问题单靠法治指数的数值是无法反映的,更不是其能够解释说明的。因此,应当通过法治指数报告这一详细的评估结果文本来理性探讨地区法治状况。但是以发现问题解决问题为旨归的法治评估,为何在连续6~7年的实施后,对一些基本的法治问题依旧束手无策。而且,类似性质的事件在其他大力推动法治评估的地区也曾出现,比如湖南的贿选事件,闵行史上最贵拆迁案等案件均发生在法治评估的先行区域。可见,尽管法治评估在其启动之时便自诩为“倒逼地区法治建设的手段”,然而一些突发的规模性事件却表明评估理想与法治现实间是撕裂状态。这足以警示我们在高热度推行法治评估的同时,应当对法治评估的效度,进行反省并形成清醒认识。所谓法治评估的效度,即法治评估的有效性,这里囊括两个层面:其一是内容效度,即所测量结果与所要考察内容的吻合程度。其二是结果效度,即法治评估对推进地区法治建设的作用效果。而这两个层面的效度既独立又相互影响,前者是后者的基础,后者是对前者的验证。

我们的研究依据余杭法治评估积累的文本材料,尤其是《法治评估的实验——余杭案例》和《中国法治指数报告(2007—2011年——余杭的实验)》两本专著。以余杭法治指数报告为基础形成的报告是值得重视的,它是当时专业团队所做工作的反映。但具体研究路径上我们则是遵循社会学文献研究方法的准则,提出法治评估在其初期阶段效度有限的研究假设,通过指标体系构建、信息采集和资料处理三个关键的评估操作过程的文本分析验证法治评估的内容效度,通过历年法治指数报告结论部分的对比分析验证法治评估的结果效度,最后就两类效度的本文分析获得结论并进行原因解析。

二、从评估操作过程看法治评估的内容效度

尽管,我们通过法治指数评估报告结论部分的对比分析摆明了法治评估的结果效度有限的事实,但我们并非旨在否定法治评估,而是以改进法治评估为研究指向。因此,进一步的研究任务在于找到评估低效的原因。现有研究更多将问题归因于评估结果缺乏接纳,但笔者认为这只是结果导向的解析,要完整认识该问题,还需要注重过程分析。由于缺少入场的机会,本文仅借助文本材料对余杭法治指数对操作过程进行分析。笔者通过从余杭法治指数的对外宣传资料考察,对其特点与发生逻辑可作如下三点总结:(1)余杭法治指数是量化评估的基础,其构建需要专家系统提供专业性的智识支撑,具体则体现在评估指标的构建、分值权重的设置、问卷设计、问卷调查与数据统计等一系列操作。(2)余杭法治指数从其本质来说是一种评估工具,即“兼具激励和约束功能的管理工具和技术手段”^[5],不过其功能的发挥

效果有待时间检验。(3) 余杭法治指数的实施是在政治精英的总体推动下,由学术团队精细加工、全民参与的一次地方性实验,即所谓的“余杭实验”是由政府牵头、专家系统主导、政府公众共同参与的评估实践,而且多元主体共同参与的原则在余杭法治评估体系的构建中便已考虑到。

不过,事实运作过程是否如其宣传所示还有待考察,即需从指标体系的构建、评估信息的获取、数据处理的方式三个方面进行评估内部操作层面的考察。

(一) 指标体系的构建

具体来说,余杭法治指数课题组在法治余杭评估体系框架下,从四个层面设计了一套详细的评估指标,即法治余杭评估总指标、区级机关指标、乡镇街道评估指标、村镇社区评估指标。其中,总指标依据法治余杭建设的九大面相细化为九大具体目标:“一、推进民主政治建设,提高党的执政能力;二、全面推进依法行政,努力建设法治政府;三、促进司法公正,维护司法权威;四、拓展法律服务,维护社会公平;五、深化全民法制教育,增强法治意识、提升法律素养;六、依法规范市场秩序,促进经济稳定的良性发展;七、依法加强社会建设,推进全面协调发展;八、深化平安余杭创建,维护社会和谐稳定;九、健全监督体制,提高监督效能”,从这9项一级指标又延伸出了26项二级指标,77项3级指标,且辅助了具体的考评标准(方便打分)。^{[1]98-112}区级机关指标则将余杭区的机构部分分为七大类,即党政机关部门、经济管理部门、司法部门、执法部门、社会服务部门以及公用事业部门,进而依各部门职责任务不同而分设评估指标,明确了各被考核部门及其相应的考核内容,具体的考核内容则操作化于“十一五”规划部署、“‘三五’依法治区、‘五五’普法教育”规划等法治余杭建设文件所规定的各项任务。乡镇街道层面则评估“组织制度建设、依法行政、依法管理、法制宣传”四个方面,共31项考核内容。村镇社区评估评估“组织制度建设、民主建设、法治建设”三个方面,其中农村社区各有25项存有差异的考核内容。

从指标结构来看,余杭法治评估指标体系实现了纵向指标与横向指标的有机结合。横向指标是主要针对区级机关重点部门的职能而分设评估指标,重在考察各部门的法治绩效。纵向指标包括总指标、区级各职能部门的评估指标、街镇的评估指标、村社区的评估指标,前三项指标主要反映政府职能履行的基本情况,而最后一项则反映基层的法治状态。在这四级纵向结构中,尤为值得注意的是最后一项及村社区的评估指标。由于村委和社区居委在我国政府体制改革下已经逐渐褪去政府职能派出机构的头衔,成为居民或村民自治的组织,因此对村社一级组织及基层状态的考察不是去考察政府职能命令的执行情况,而是看基层自治和基层民主的运转情况。由此看来,余杭法治评估体系在系统性与完整性方面是典型。

再从指标的生成过程来看,在《法治评估的实验——余杭案例》一书中,余杭法治评估的课题组对如何设计这一系列评估指标并没有详细说明,仅以“课题组为此进行了大量调查。经过多次讨论,最后形成了法治余杭评估体系。”然而,指标的生成过程决定了指标的质量,如何调查怎样讨论这都是指标真实有效的关键。尽管课题组强调了调查与讨论,但从指标的具体表述中我们可以发现,各操作性指标均是通过转化政府职能的表达而来的。当然这种表述方式也有其优势,确保了指标的针对性、适用性和可行性,为评估余杭法治状况提供了操作上的便捷性。作为一份评估法治政府工作绩效的指标,这是很完善的,但作为一份评估法治整体状况的指标,这是有欠缺的。因为这样一份指数框定了民众抑或是专家评定的范围,这个范围本身是政府有选择地呈现出来的,评估主体的参与方式尽管形式多样,但其思维方式是受政府制度所设定的。因此,尽管有第三方的独立评估,尽管有民众的大量参与,尽管有专家的专业评定,仍无法避免法治评估沦为政府进行政绩包装的工具。这就解释了为什么“余杭事件”呈现出来的余杭法治状况与法治指数报告所呈现出来的余杭法治状况存在差异。因为法治评估指标的构建是遵

循政府官方话语的逻辑,而非真实的法治运作逻辑。

再对具体评估指标进行考察发现,量化的指标更偏重政府工作步骤也就是程序层面,而对民主推动下的工作实效的测量所占比例较小。以法治余杭评估总指标为例,多以“参与比率”“达标率”“公布率”“办结率”“执结率”的数值考察为主,这些“率”的确提供了评判的客观性依据,也是量化法治评估生成法治指数的有效指标。这些指标反映了政府工作在量上面的成果,也说明政府工作是对程序的严格执行,但是缺少对政府效能的质的评估。以民主参与对党委决策影响为例,参与率是一个方面,参与中公众意见的采纳情况,对决策的影响效果则是另一个方面。仅以参与率来衡量,是难以揭示“走过场”的不良情形的。作为弥补,典型调查应成为一个可选方案,通过了解具体个案的发生、发展与结束,以弥补量化数据的信息缺陷。

(二) 评估信息的获取

余杭法治评估突破了封闭式内部评估单一信息来源的局限,搭建来一个由普通公众、政府系统的工作人员、企业家等市场主体代表、记者等媒体代表、律师等法律共同体代表,以及法学统计学等专家学者等多元主体共同参与的评估网络。无疑这是一个综合性的网络,对构建该综合网络的意义,课题组曾明确表示“能够较好地降低单一评价主体的主观性偏见带来的评价误差,最大程度地保证法治量化评估的客观准确性”^{[1]242}。可见,余杭法治评估的信息获取是内外结合的,是大众评判与专业评价相结合的。面对多元的评估主体,课题组借鉴香港法治指数评估的数据获取方法,将数据来源分为法律数据和调查数据^{[1]314-315}。

法律数据是官方统计的数据,具有两大作用,其一是直接整合进评估报告,其二是作为外部评估者进行评分的参考依据。政府无疑是国内最大的数据采集方和所有方,其所能提供的数据量也远高于其他民间统计机构,但是数据对公众的开放是有限的,所以才需要评估机构做专门的搜集供打分者参考。然而是否存在与纷繁复杂的评估指标一一对应的参考数据是值得怀疑的,这将致使打分者只能在有限的客观数据基础上进行相对主观印象的打分,而非完全客观中立的评分。

调查数据则主要来源于对民众法治满意度的主观性认知调查,为此课题组设计了九份不同主题的民意调查问卷,包括“人民群众对党风廉政建设的满意度”“群众对政府行政工作的认同度”“人民群众对司法工作的满意度”“人民群众对权利救济的满意度”“人民群众的社会法治意识程度”“群众对市场制度规范性的满意度”“人民群众对监督工作的满意度”“群众对民主政治参与的满意度调查问卷”“人民群众对安全感和治安满意度”^{[1]89-93}。具体的调查方式则有实地调查、电话访问和网上调查等,调查区域涉及党政军工农学商等各种单位、火车站、汽车站等^{[1]83}。主题丰富的调查问卷确实为信息的收集奠定了扎实的基础,多种调查方式也实践着实事求是和客观中立的原则。但是从问卷的结构中,我们看到了一个不可忽视的问题,那就是缺乏对调查对象基本资料的搜集。调查对象的基本信息是问卷结构中不可或缺的组成部分,通常作为社会调查中的重要自变量,缺少这一类信息,统计便没有针对性。至于课题组所声称的克服偏见的评估也缺少有力支持的证据。若从社会调查问卷的规范上来,正如风笑天教授所说,“如果一份资料缺少这些变量,实际上也就成了废卷。”^[6]

此外,据课题组介绍,“余杭地区的法治状况评估,正是通过大量定性和定量的研究才得以完成的。”^{[1]296}因此,除了法律数据、调查数据等量化数据之外,质性材料也是评估信息的重要类型。但是,从各项文本来看,我们都只能看到课题组对定量评估的绝对重视,而无定性研究理论储备的介绍。相较于详细的评估指标及问卷,定性研究的方法论、问题假设、研究提纲等均未在评估经验总结和介绍中有所突出。而即使是给人以精确感和复杂性的数据统计也存在不足,尤其是缺乏现象间关联的设计与揭示,哪怕是剔除特殊情形的普遍关联。

(三) 数据处理的方式

从信息处理方式来看,余杭法治评估既有单项的简单分析,也有项与项之间的对比分析,但总体来说描述多于原因解析。余杭法治指数报告中着重分析的数据有三类,一类是民意调查结果分析,一类是内外组评审结果分析,一类是专家组评审结果分析。(1)有关群众满意度问卷数据的分析,在报告中所呈现的是一个总体满意度得分,即通过固定的方程式计算出人民群众对“党风廉政建设、政府行政工作、司法工作、权利救济、社会法治意识程度、市场秩序规范性、监督工作、民主政治参与、社会治安”等九项的分值。一目了然的数值便于进行历年的对比,但具体问题的数据并未得到细致的分析,因此群众满意度问卷数据的分析显得单调苍白。且由于缺乏单个问题间的对比,所以难以发现历年差异何在。(2)有关内外组评审结果的分析相较于民众满意度问卷数据的分析而言稍复杂些,不仅有组内的纵向历史维度的对比分析,还有组间的横向不同评估主体意见的对比分析,且内外组间的差异点得到了明确的揭示。不过内外组之间的对比分析仅是简单的数值罗列,同质与差异所蕴含的重要含义、差异的原因同样尚未被深挖。(3)从专家组评审得分分析来看,2009年的法治指数报告才重视这一数据的分析。课题组称:“总之,专家组评审从专业化的角度给出法治指数,一定程度上平衡了与内外组评分产生的差异,也凭借专业化水准提示余杭法治建设应努力的方向。”可见,余杭法治评估课题组对专业评估的重视,对评估水平的提升起到了关键的促动作用,但要警惕把差异的平衡与余杭法治的未来完全寄希望于“专家”。尽管专家可以凭借理论储备对研究问题进行深度透视,但是专家在学术研究中积累的观点是否符合地区法治现实依旧是有待验证的。因为归根到底,问题的呈现还是需要依赖社会事实,也就是说专家的评判不能脱离实践进行理论空想,也应当以地区法治的特征与基本状况为基础。然而,从报告中我们发现前期收集的地区法治信息与专家的评审依据往往是两张皮。

此外,面对专家意见、民众意见、内外组成员意见的差异点,评估课题组没有充分的质性材料予以差异性根源分析,对评价主体所指出的法治问题也没有质性材料予以详细说明并解读根源所在。即是说,分析的局限产生于数据采集的方式过于定量化,缺乏定性的资料采集。进而形成的后续反应便是以数据分析为主的报告只能对政府工作提供就事论事的问题指导,评估建议难以具有长远的战略眼光。

三、从评估结论的对比分析看法治评估的结果效度

一直以来大家过多关注法治指数得分的变化,却鲜少关注历年评估报告的文本变化,本文以《中国法治指数报告(2007—2011年)——余杭的实验》为基础文本资料,进行评估报告的文本分析,尤其是问题和建议等评估结论部分。问题建议等结论部分是对前期所处理和分析数据的一种提炼,是评估的核心成果,体现了法治评估的民主性与专业性。从理论上讲,法治评估若能有效推进地区法治建设,将在评估结果的问题部分有所显示,比如问题减轻或消除。而问题的解决与评估建议的有效性又是息息相关的,建议部分的对比分析是判断建议是否专业,是否具有操作性,能否为政府所重视的可行方式。

(一) 问题的对比剖析

法治评估的出发点是为了“找问题补短板”,倒逼地区法治建设,因此问题剖析应当是法治指数报告中最为重要也最能体现评估成效的部分。简要来说,2007年的法治指数报告集中揭露了余杭在监督体系、司法公正、市场秩序、法治意识四个方面的问题,且在每个大问题下面又细分诸多小问题,力图将问题阐释清晰。2008年和2009年两份法治指数报告则从党委依法执政、政府依法行政、司法公平正义、权利依法保障、市场规范有序、监督体系健全、民主政治完善、全民素质提升、社会平安和谐九个方面全面分析和批判了余杭法治建设中存在的问题。2010年和2011年的法治指数报告所反映的问题则显得较为琐

碎,但焦点基本集中在监督、司法公正、市场秩序、法治意识等方面。可见,历年报告形式所有差异,但其大体框架还是较为一致的。下文集中选取监督体制、司法公正、市场秩序、法治意识四项在报告中得到一贯关注的主题,对具体问题进行分类整理。

1. 监督体制的问题。2007年的法治指数报告对这一主题的问题挖掘是最为全面的,包括“监督体系的形式庞杂、制度繁琐与实施效果的低下形成了鲜明反差”“惩治和预防贪污腐败的体系不成熟不完善”“监督体制没有从一开始就获得足够的权限,权力没有得到来自社会多元性的监督,不可能有很高的监督成效”“政府的责任追究和绩效评估制度不尽完善”^{[1]66-67}。2008年的法治指数报告在监督体系这一主题上所指出的主要问题有三项,即“行政监督主体地位不高,缺乏监督力度”“政府权力过大,权力运作不公开”“舆论监督没有发挥”^{[1]159-160}。2009年的法治指数报告指出监督体系的核心问题为“监督执行”^{[1]259}。2010年的法治指数报告指出的监督体制存在监督效能低下、监督意识薄弱、形式大于内容等^{[1]320}问题。2011年的法治指数报告指出的问题是“内部力量和民众力量均不足”^{[1]356}。

2. 司法公正问题。2007年的法治指数报告围绕司法权威指出了司法公正的三个问题,即外部环境不佳、法官检察官的业务素质和能力不过硬、人情官司后门官司等问题^{[1]68}。2008年的法治指数报告指出妨碍司法公正的主要问题是执行力度不够、行政干预以及司法工作者的素质等^{[1]153}。2009年的法治指数报告指出的司法工作方面的问题是执行难、党政干预司法^{[1]256}。2010年的法治指数报告指出司法权威受挫仍旧是当下整体所面临的难题^{[1]316}。2011年的报告则简单地用数据说明余杭司法公正权威指数的下滑,但并未真正说明原因。

3. 市场秩序问题。2007年的法治指数报告将影响市场秩序的问题直指薄弱的市场秩序诚信体系^{[1]69}。2008年的报告则指出了市场秩序治理中的三大问题,包括坑蒙拐骗等市场陋习、市场管理机关的运作问题,以及非诚信经营等^{[1]157-158}。2009年和2010年所指出的问题则与2008年的基本一致,2011年的市场秩序得分更是降到历年来新低。

4. 法治意识问题。2007年的法治指数报告总结“全区公众的法治意识还比较淡漠”^{[1]69}。2008年的报告指出全民法治素养的问题有三个,即法治意识淡薄、法制教育水平落后、法治教育方式需改进^{[1]163-164}。2009年的报告则指出的是法制教育辐射面有限的问题。2010年的报告对此未作评论,而2011年的报告则是简单给出了“处合理水平,但据优秀值还有距离”^{[1]357}的总体评论。

经由以上问题部分的对比分析可见,经评估挖掘的法治问题是集中且一以贯之的,但是问题在揭示之后却无显著改善。具体来说:(1)监督体制作为法治指数报告所要揭示的重点问题一直占据着报告的首位,尽管表达方式多样,但是都围绕着“监督执行难”来谈;(2)司法公信力下降的问题一直有呈现却从未有改善;(3)诚信等市场秩序问题一直未得到有效解决,甚至有恶化的趋势;(4)公众的法治意识作为地区法治的核心问题,亦未能在短期内得到显著提高。

(二) 建议的对比剖析

对应以上地区法治的重点问题,法治指数报告中也针对性地提出了以下建议:

1. 关于监督体制的建议。2007年的法治指数报告提出的建议是“采取多元化监督方式”“加强检察机关人员的职业道德修养”等^{[1]67}。2008年的法治指数报告的建议落实到具体的制度和机制上,包括区人大、区法院、区政府联合监督,完善政府信息公开制度,建立健全舆论监督机制^{[1]160-161}。2009年的建议进一步深化,不仅有强化监督网络、健全舆论监督机制,还有明确监督重点、责任监督工作体制等^{[1]259}。2010年和2011年的报告均尊重群众意见,建议加强体制内外的监督,提高监督效能。

2. 关于司法公正的建议。2007年的法治指数报告针对所发现的问题提出了营造良好的外部氛围、提升法官检察官素质等建议^{[1]69}。2008年提出的四点意见分别针对法院的审判和执行工作、法官素质的提

升、杜绝党政干预以及监督管理的规范化^{[1]154}。2009年的建议是严守司法程序确保判决的执行、坚持法律职业性^{[1]256-257}。2010年和2011年的报告仅总结群众所提的问题,并未给出具有建设性的实质意见。

3. 关于市场秩序的建议。2007年的法治指数报告指出为保障市场秩序,必须大力开展诚信体系建设^{[1]69}。2008年的报告所给出的建议较为系统,包括厘清政府管理部门的职责、加强对市场的引导和规范、借助媒体力量对公众进行诚信教育等^{[1]158}。2009年的三条建议中有两条与2008年的建议完全一致,增加的一条建议是要求完善各类法规^{[1]259}。2010年和2011年的报告均止于总结问题,而无针对性意见。

4. 关于法治意识的建议。2007年的法治指数报告建议首先要加强政府部门自身的法治意识,同时要通过普法教育提升群众的法治意识,并给出了较为具体的举措^{[1]71-72}。2008年的报告给出了三大建议,即着重提高领导干部的法治化管理水平、通过家庭学校单位社会多阵地联动来深化普法教育、调动律师媒体等外部力量创新普法方式^{[1]256-257}。2009年的建议则是扩大法制宣传辐射面,并将重点落在“权利义务对等观念”^{[1]261}。2010年和2011年的法治指数报告在法治意识部分的建议与其他部分情况一样较为弱化。

从以上四部分建议的梳理,不难看出这样两点:第一,2007年至2009年的法治指数报告均比较重视建议部分的内容,并较为细化,而2010年至2011年的报告则明显弱化了建议的地位。第二,即使在较为重视建议部分的2007年至2009年,无论是制度上的建议,还是操作层面的建议大多在重复提出。

(三) 问题建议对比分析的结论

结合以上问题部分和建议部分的对比分析发现,我们并不能从法治指数报告中找到法治评估倒逼地区法治的有力证据,相反是相同问题的一再揭示以及建议部分的老调重弹甚至日渐虚化这类评估事实。由此可推测:(1)无论是问题的揭示与分析,还是建议部分的专业指导都呈现出分量渐轻的情况,如图1、图2所示,说明在法治指数评估最初两三年间评估各方的热情最高、投入最多,然而到第四年便开始呈现出疲乏姿态,评估日渐形式化。(2)问题一直存在却未有实质性改善,说明与余杭政府在媒体宣传中对专业统计方法、科学方程式得出的法治指数数值的重视和强调形成鲜明对比的是,法治指数报告的具体内容却是被忽略的,问题未被重视、建议未被采纳,即评估结果缺乏有效接纳。(3)而前两点又共同说明法治评估的效度是很有限的,更多停留在形式意义上,而非实质促动,评估日渐呈现出“为评估而评估”的特性,评估从手段退化为了目的本身。

表1 历年法治指数报告问题数量

问题类型	2007 法治指数 评估报告 问题数量	2008 法治指数 评估报告 问题数量	2009 法治指数 评估报告 问题数量	2010 法治指数 评估报告 问题数量	2011 法治指数 评估报告 问题数量
监督体制	4	3	1	3	2
司法公正	3	3	2	1	1
市场秩序	1	3	3	3	3
法治意识	1	3	1	0	1

表2 历年法治指数报告建议数量

建议指向	2007 法治指数 评估报告 建议数量	2008 法治指数 评估报告 建议数量	2009 法治指数 评估报告 建议数量	2010 法治指数 评估报告 建议数量	2011 法治指数 评估报告 建议数量
监督体制	2	3	4	1	1
司法公正	2	4	2	0	0
市场秩序	1	3	3	0	0
法治意识	2	3	1	0	0

四、结论与反思

经由以上文本分析,笔者发现余杭法治评估既有成功之处,可资学习和借鉴的地方包括指标体系构

建的方法、信息采集的方法、凝聚地区法治共识等积极作用等,但亦有不足之处亟待反思和完善。现就文本分析总结如下:效度有限是法治评估发展初期的事实问题,而且成因复杂。具体说来,包括内外因共同制约下的评估深度有限和评估结果影响有限问题。

(一) 评估深度有限及原因

余杭法治评估尽管有着很强的专业外衣,但整个评估主要是简单描述式的,缺少因果关联式的分析,最终结果是仅能部分揭示面上的问题而未能解释问题的成因,评估深度之有限显而易见。追问原因,这既有评估方式上的问题,也有评估条件的问题。

第一,从评估方式角度来看,课题组试图通过量化的评估来给出科学的结论,也曾强调用定性评估予以补充,但是经过上文的梳理可以看出,有量化却没有相关性(乃至因果)的分析,有质性的信息采集技术却没有深入了解政府工作的具体运作逻辑。然而,真正有深度的评估需要理解政府的运作逻辑,需要理解政府治理模式与法治建设成效之间可能的因果关系。法治评估作为对法治进行全面检视的一项机制,应该承载其更多的职责,然而缺乏对与问题相关事件发生过程的深入探析,缺乏对政府运作逻辑的理解,是难以找出问题的症结所在的。

第二,评估所具备的现有条件也难以支撑评估的彻底性。这里既有第三方评估团队的自身条件,也有评估委托方所能提供的条件。由于法治已渗透政治经济社会生活的方方面面,对实践中的法治要能够形成全面的理解,必须借助法学、社会学、行政学、经济学、统计学等多门学科知识,因此为实现高质量的评估必须配套以知识结构比较完整的评估团队,然而余杭法治指数的课题组成员的学科背景相对来说还是比较单一的。即使在以上外部条件均具备的条件下,我们还需要政府部门予以充分配合,才能真正达致理想的评估深度。但是从现有的评估途径我们可以看到,评估方仅是对部分内部工作人员进行量表测定或访谈,而并没有能够真正深入政府的日常工作中,党政各部门并没有对第三方评估团队完全开放。缺乏这一开放条件,我们便无法论证话语所简单表述出来的内容是否与实际运作相一致,更无法了解并论证是否存在相悖之处。然而,期望完全开放也是不现实的,这具有难以一朝便能克服的阶段性,因此在有限的空间内,如何做到最理想的状态,对第三方评估机构来说的确是一大挑战。

(二) 评估结果影响有限及原因

从法治指数报告问题和建议部分的分析结果中我们可以获知,课题组历年所反映的问题相似,建议也相似,而且建议部分有越来越弱化的趋势。可见,相较于每年向外公布且较受追捧的指数值,具有战略意义的法治指数报告的实际影响力却并不理想。

首先要质问的是评估报告所提建议的含金量。法治指数报告并非出自实务部门,而是由具有专业权威的知识分子出具。知识分子往往以产出理念为己任,因此从前期的报告中我们也能看到一些具有创新性的制度和机制意见,这是值得肯定的。但部分建议往往给人空中楼阁之感,简言之“看去很美好,落地很困难”。因为在法治建设过程中会出现的种种问题和困扰并不能由简单的制度设计和策略创新就可得到解决,所出台的具体法治举措需要契合整个体制环境和社会环境,方能排除障碍有效推进法治建设。然而余杭的法治指数报告所提建议仅仅是依据表面问题进行的理论设想,而在问题本身尚未分析透的情况下,建议的治本性可想而知。2010年及2011年的法治指数报告更是将建议部分几笔带过,问题的叙述也仅是罗列被调查对象的话语,或作一两句简要总结,实在无法支撑评估报告发挥重大影响力。

其次要质问的是法治评估报告的接纳结果。年年反映的问题相似,年年提出的建议相似,这足以说明前期的问题及建议并未得到党政工作系统的有效接纳,可见评估与接纳始终是两张皮。尽管2008年和2009年的法治指数报告以评估结果对实际工作所产生积极的影响作为开篇,试图跳出就评估而评估的限制,说明并促进评估与接纳的有效对接,但在2010年和2011年的报告撰写中却并未得到延续。

这也从侧面补充说明了法治评估并未得到有效接纳。在实践中,法治评估结果已作为整体考核的一部分依据,但将一些问题与建议落到实处却要受制于当下的行政领导体制。作为其他各委办局的平行部门,司法局没有这样的权力与权威直接对其他部门的下达改进意见,下级部门直接服从上级部门,各部门的工作方向是受上级部门及本部门领导决定。实践中往往是领导或上级部门若重视,那么法治评估中所指出的问题和建议就能在本部门下一年度的工作中进行解决或落实。这样的话,法治评估对法治工作的推进便缺乏稳定性。因此,让法治指数报告得到有效接纳,对实际工作进行直接的推动还有很长一段路要走。需要实务界与学术界就构建出符合现有体制条件的法治评估接纳机制达成共识,并投入智识努力。

由此看来,法治评估初期发展阶段评估效能有限的问题并不能简单归因,而是多方面因素共同作用下的结果。加之没有明确的责任主体,更容易放任评估形式化的发展,因此,如何保障法治评估顺利完成制度化转型之路应当是下一步研究思考的重点。

参考文献:

- [1] 钱弘道. 法治评估的实验——余杭案例 [M]. 北京: 法律出版社, 2013.
- [2] 郭振纲. 法治指数的出台 政府花钱买数字? [N]. 东方早报, 2008-06-19 (03).
- [3] 志灵. “法治指数”无法衡量所有法治现状? [J]. 法制资讯, 2008 (4): 66.
- [4] 张建. 法治评估的地方实践: 动因、问题及反思 [J]. 云南师范大学学报 (哲学社会科学版), 2016 (1): 99—105.
- [4] 钱弘道. 余杭法治指数的实验 [J]. 中国司法, 2008 (9): 60—65.
- [5] 风笑天. 社会学研究方法 [M]. 北京: 中国人民大学出版社, 2009: 176.

The Validity of the Rule of Law Evaluation Practice

—Based on the Context of the Legal Index of Yuhang

Li Yuqing, Zhang Ling

Abstract: The validity of the legal practice evaluation becomes an important focus through the event of Yuhang Zhongtai waste incineration plant. The validity of legal practice evaluation includes the content validity and the results validity. The research finds the deviation between the validity and expression in the legal practice evaluation in Yuhang by analysing the content of index system, information collection, date process and the conclusion of the report. The lack of effective acceptance system and other reasons lead to above problem.

Key words: legal evaluation validity; legal index of Yu hang; Yu hang Event

(收稿日期: 2016-08-05; 责任编辑: 朱世龙)