

不完全考核： 模糊与精细混合的政府考核分析^{*}

练 宏 陈 纯

提要：本文通过田野观察一个税务部门连续三年的考核过程，揭示了考核的复杂运作图景，即受制于政府科层组织的独特制度环境，考核难以完全按照泰勒科学管理的思路来进行精细化考核，而是呈现出模糊与精细并存的混合考核实践。其中，精细化考核只是初步实现了考核目的，之后组织通过模糊运作获得权威控制，缓解了精细化考核的事本主义，具有“模糊的精确性”功能。本研究有助于重新解释考核中的偏差行为、审视单向度的科学管理和呼应国家当前对唯指标论和优绩主义的反思。

关键词：不完全考核 量化管理 混合考核 政府行为

一、研究问题

当代政府的考核愈发强调量化管理，由模糊性考核向精细化考核转变。模糊性考核 (ambiguous evaluation) 指具有较为主观、过程动态式、指标定性化、奖惩不严和结果部分公开等特点的考核。精细化考核 (accuracy evaluation) 则指具有相对标准化、固定化、指标清晰量化、奖惩严格和结果完全公开等特点的考核。精细化考核遵循泰勒的科学管理和福特主义思路，自现代以来被美国、德国、西班牙等多国采纳，成为一种广为接受的制度实践 (Guillen, 1994)。斯科特指出，现代国家发展的一个趋势是借助清晰化测量工具，实施一种简约和标准的度量政治学 (politics of measurement)，“如果没有标准化、固定性的度量单位，那么可能难以有效进行集中监督或者控制性比较” (Scott, 1998: 30)。与之类似，中国

* 感谢周雪光、刘世定、陈那波、黄晓春、刘玉照、陈家建、冯猛、艾云、狄金华、邓燕华、王冰杰等师友及两位匿审专家的宝贵意见。文责自负。

自改革开放以来也在推进精细化考核,由最初的年终考核演变为侧重日常的近距离考核;指标由定量和定性并重转向强调以“能量尽量”的量化考核为原则;由原来的趋同性考核转为差异化考核,强制拉开指标差距;考核结果走向公开化并做到奖优罚劣,推进干部能上能下,不能“干与不干一个样、干多与干少一个样、干好与干坏一个样”。^①

一个令人困惑的现象是,改革开放四十余年来,虽然政府不断推进精细化绩效管理,但在实践中模糊性考核现象仍普遍存在,甚至有些地方政府还主动或刻意保留模糊性。以本文关注的税务部门为例,税务在政府体系中属于较早开展绩效管理的单位。相比其他部门,它的考核体系相对完善,近几年还建立了数字人事信息系统,能够全面记录、量化计算和自动提取业绩数据。^②由此推测,税务应是最具备实施精细化管理条件的部门,但实际调查却发现存在诸多模糊与精细混合的现象:(1)定量指标比重趋于上升,但始终保留定性指标;(2)有些定性和定量指标可以灵活调整;(3)一些可调整指标在制定时被要求宜粗不宜细;(4)尽管定性指标总体减少,但领导评价定性指标依然保有重要位置;(5)组织内部默许轮流坐庄现象;^③(6)考核结果的部分公开而非完全公开,等等。图1展示了某税务部门考核指标分值变化,反映了税务考核既有精细化特征,也有模糊化图景。

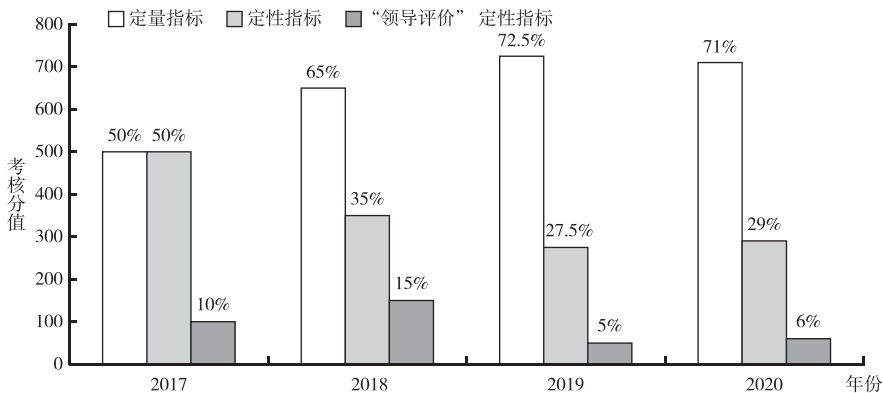


图1 某税务部门2017年至2020年的考核指标分值变化

① 中共中央组织部,2019,《公务员平时考核办法(试行)》,国家公务员局网站(http://www.secs.gov.cn/zcfg/202001/t20200108_16201.html)。

② 国家税务总局,2013,《国家税务总局关于实施绩效管理的意见》(税总发[2013]130号),国家税务总局网站(<https://fgk.chinatax.gov.cn/zcfgk/c102424/c469626/content.html>);国家税务总局人事司,2017,《数字人事管干部考核管理蹚新路》,国家税务总局网站(<https://www.chinatax.gov.cn/chinatax/n810219/n810724/c2834977/content.html>),9月26日。

③ 需要注意的是,“轮流坐庄”在本文只是一种中性意义上的描述,而非贬义用法。

考核的模糊化折射了制度设计与实际执行之间的张力。由此,本文提出如下研究问题:为何标准化和技术性较强的税务系统难以做到理想意义的精细化考核?模糊性生成的制度环境是什么?进一步而言,如果模糊性考核重要,为什么没有全部替代精细化考核?精细化考核的意义何在?

二、文献评述

为了解答上述问题,我们检索了相关文献,梳理现有研究关于模糊性考核现象的解释。经过分类,至少有如下两种类型的解释。

(一)技术角度的解释

这一角度的研究认为模糊性源于技术设计不足,具体分为三种情况:第一,技术信息不足。由于考核没有建立或利用各类客观数据库和信息库等,绩效评估缺乏科学依据,形成人为主观性的考核(尚虎平、赵盼盼,2012)。第二,技术统计不足。一些考核数据没有给统计部门汇总和审核把关,导致考核的主观性强,自话自说居多(崔述强等,2006)。第三,指标难以完全量化。蔡立辉和刘晓洋(2016)也认为实际考核过程很难把所有信息量化,使得考评者不得不凭主观感觉打分,损害考评的公平性。文宏和郝郁青(2017)从政府职能出发,认为政府部门具有职能多样性、工作复杂性等特征,使得很多工作无法被量化。

这一角度的研究以科学管理的精细化考核为参照,认为模糊性考核源于客观技术的不足,具有一定的启发性。但是,这一角度的有些观点值得商榷。第一,随着信息技术发展,当代政府的信息收集、信息统计及指标量化等能力均大幅提升,并不像历史上的政府那样不太擅长数目字管理(黄仁宇,2015),上述文献可能低估了现代政府驾驭统计信息的水平。第二,无法解释刻意模糊现象。这一角度的研究难以解释像上文税务部门所表现的刻意模糊行为,即税务部门在技术上明明可以把指标量化,却主动不量化。因此,技术角度的解释可能存在局限,正如诺斯等人所言,需要谨慎看待技术的作用,技术有时是结果而不是原因(North & Thomas, 1973)。

(二)权威角度的解释

与技术视角的解释不同,这一角度的研究认为模糊性并非技术不足,而是出于

组织的调控需要。倘若考核完全精细,将不利于组织对考核的灵活调整与内部控制,反过来会削弱组织权威。王汉生和王一鸽(2009)指出,不可量化的任务体现更为纯粹的政治权威关系;如果所有任务都被量化,则上下级政府的权威关系将被稀释甚至消解,而这不利于应对基层政权运作和地方社会治理的复杂性和多样性。这一观点从权威能动性角度解释了模糊性现象存在的原因。另外,还有一系列研究也注意到权威在政府行为中的作用。关于环保局考核的研究指出,上级政府通过内部权威关系尤其是目标设定权、检查验收权和激励分配权的不同组合变化来调节上下级关系,激励环保系统官员有效实现减排目标(周雪光、练宏,2012)。周黎安(2014)在行政发包制的研究中也指出,上级政府常常通过行政权控制、财政权分享与结果控制等方面的不同组合,实现对下级政府的有效管理。

上述权威视角的解释比技术视角更具启发性,它基于对政府内部运作的田野调查,揭示了模糊性的另一种动因,即组织通过模糊性获取权威,掌握考核主动性,进而形成组织抓手。当然,权威角度的解释也留有一些问题尚待研究。其一,既然模糊性考核有助于权威关系的建立,那么为什么没有完全替代精细化考核?精细化考核的作用何在?其二,考核中的权威有助于应对复杂多样的制度环境,那么复杂制度环境指的是什么?其三,模糊性运作既然可以增强权威控制与调节,那么背后的微观运作机制是什么?

本研究在权威视角基础上,试图做出进一步的推进,意在揭示模糊与精细并存背后的制度成因及其微观运作机制。文章接下来的内容安排如下:首先,介绍精细化考核的运作特征。其次,分析模糊性考核的修正调整。最后,分析考核面临的制度环境,以及在制度环境之下如何通过模糊考核获取权威控制。

三、奖优罚劣:精细化考核的运作分析

为便于理解,先介绍税务局考核的组织背景,然后展现精细化考核过程。

(一)研究背景

从2018年到2020年,我们的研究团队利用田野调查方法,在一个副省级城市的区税务局(简称“Y”局)进行了长达三年的蹲点观察。^① 研究人员以绩效办

^① 按照学术伦理规范,本文涉及的人名和地名等信息均进行了匿名化处理。

经办人员的身份,每年数月在 Y 局上班,参与了 Y 局绩效考核组织实施的各个环节,在日常工作环境中观察记录了上下级部门及内部同事的大量交往互动。Y 局具体的考核方案详见表 1。

表 1 Y 局的绩效考核方案(2018—2020 年)								
组织绩效	被考核主体	一级指标	二级指标	考核主体	2018 年	2019 年	2020 年	指标类型
	区机关科室 (1000 分)	基础指标	承接市局指标	市税务局	300 分	300 分	300 分	定量/定性
			共性指标	职能科室	350 分	425 分	410 分	定量指标
			基层评价	税务所	/	25 分	30 分	定性指标
			部门重点工作	全体局长	200 分	200 分	200 分	定性指标
			领导评价	全体局长	150 分	50 分	60 分	定性指标
		加减分	加减分项目	职能科室	表彰类等加分项目和 批评类等减分项目			定量/定性
个人绩效	平时考核 得分(70%)	人员类别	领导班子正职	领导班子副职		部门正职	部门副职	其他干部
		组织绩效 挂钩得分	单位组织 绩效(70%)	单位组织绩效(20%) 分管科室组织绩效(20%) 分管税务所组织绩效(10%)		部门组织 绩效 (50%)	部门组织 绩效 (30%)	部门组织 绩效 (10%)
		领导评鉴	10%	20%		20%	20%	50%
		现实表现	20%	30%		30%	40%	40%
	公认评价得分(30%)		所在单位办税服务厅的现场评价、纳税人满意度调查结果和行风评议结果等三个方面					

注:(1)基层评价指税务所对职能科室的工作进行满意度测评。(2)领导评价指局领导对各科室落实全面从严治党、重点任务落实、工作质效和工作创新等情况进行考评。(3)加分项目指通报表彰、批示表扬和其他加分;减分项目指通报批评、批示批评和其他减分。(4)领导评鉴指主管领导对干部个人进行综合评价;现实表现指对税务干部德、能、勤、绩、廉等五方面表现进行评价。

Y 局属于中央垂直管理体制,人事和财政均由市税务局负责。2018 年机构改革后,Y 局由原来的区国税局和区地税局合并而成,被定为正处级。在市税务局组织的历年考核中,Y 局总体成绩居全市前列。表 1 呈现了 Y 局的绩效考核方案,包括组织绩效和个人绩效两部分。考核流程如下:首先,国家税务总局在第一季度经过三次征求意见后,正式下发总局机关和省税务局的考核体系。其次,省税务局制定省局机关和市税务局的考核体系,市税务局制定市局机关和区税务局的考核体系。再次,区税务局考核科按照月考核、季度考核、半年考核和年度考核等时间节点,把考核成绩录入绩效系统。最后,区税务局按季度召开绩效讲评会议,年终考核结果由区税务局党委会议审议通过并公布。

(二)精细化考核的运作

精细化考核的初衷是借助分类和量化工具,形成相对公平和可比的考核体系,然后进行奖优罚劣,最终产生激励先进、鞭策后进和你追我赶的组织氛围。Y局的考核正是在这一思路下进行的。表2呈现了精细化考核的具体过程。

表 2 精细化考核的运作分析				
精细化维度	具体表现	2018 年	2019 年	2020 年
指标客观可测	定量指标比重	(650)65%	(725)72.5%	(710)71%
	过程性考核	月考核、季度考核、半年考核和年度考核,其中年度考核指标只占30%。		
	业绩数据客观化	以日常工作台账或系统截图作为佐证材料		
业绩分离比较	区分主办和协办	主办科室获得所有加分,协办科室获得加分分值的50%;减分同理。		
	区分组织和个人	分为组织绩效和个人绩效,个人绩效与组织绩效进行部分挂钩		
考核标准固定化	不可调整内容	承接市局指标+共性指标+基层评价		
	不可调整指标	(650)65%	(750)75%	(740)74%
	年内不再调整	年度中间原则上不可对绩效指标的完成时间和考评标准进行调整		
业绩结果差异化	强制拉差指标	(169)16.9%	(327)32.7%	(392)39.2%
	领导评价指标	(150)42.86%	(50)18.18%	(60)20.69%
	·优良合的间距	·100%—90%—80%	·100%—95%—90%	·100%—90%—80%
	·优秀比例	·优秀不超过50%	·优秀不超过10个	·优秀不超过50%
	部门重点指标	200分中的50分要求;	200分中的50分要求;	200分中的50分要求;
	·优良合的间距	·100%—90%—80%	·100%—90%—80%	·100%—90%—80%
	奖励考出差异者	对考评区局指标考出差异的市局处室加0.3分		
结果运用严格化	绩效讲评督促	每季度公布市税务局对区税务局的考核成绩,并由区税务局局长提出工作要求		
	奖优罚劣规则	在干部选拔、年度考核、评先评优和人才项目等方面进行奖优罚劣		
	结果完全公开	信息完全公开,让考核成员知晓全部排名情况,达到激励先进、鞭策后进的效果		

注:(1)括号内为指标分值,括号外数值为该指标所占比重。(2)领导评价指标占比是指领导评价指标占当年所有定性指标的比重,例如2018年领导评价指标是150分,定性指标是350分,则占比为42.86%,2019年和2020年的计算原则与此相同。

首先,形成一套可衡量、客观的竞赛指标。如果考核基于一些模糊和主观的标准,成员可能无所适从。为了防止指标模糊造成的激励效果弱化,税务系统采取了诸多措施:(1)提高定量指标的占比。如表2所示,2018—2020年的定量指标比重总体上升,由2018年占指标的65%上升稳定至2020年的71%。(2)增强过程性考核。考核节点分为月考核、季考核、半年考核和年度考核四种类型,其中年度考核的指标数量不超过指标总数的30%。(3)业绩数据的客观化。对那些有明确考核标准的绩效指标进行成绩录入时,考核科要求被考核主体提

供日常工作台账或系统截图作为佐证材料,作为业绩客观真实的保证。

其次,成员成绩相对可分离和可比较。有些政府工作需要通力合作,如果考核没有区分群体与个体的贡献,容易弱化激励。为此,税务系统做了两个方面的努力。(1)区分主办和协办。例如一个全局重点工作同时涉及主办和协办科室,若市局对区局这一工作加分,主办科室可获得所有加分,同时协办科室获得加分分值的 50%;如果加分项目涉及不止一个主办科室,则各主办科室需共享市局加分。若市局对区局扣分,则主办科室在承接失分的基础上,又会将失分按工作情况分配给协办科室,最终由主办和协办科室共同承担失分。(2)区分个人绩效与组织绩效。从表 1 可以发现,职位越高者,个人绩效与组织绩效挂钩比例越大,这可以让领导者站在组织整体考虑问题,为组织负责;职位越低者,个人绩效与组织绩效挂钩比例越小,但其三成的成绩由更了解情况的部门领导来评价。由此形成一种相互传导的效果,即领导为组织负责,个体为领导负责。

再次,成员努力与最终考核绩效关联。倘若成员无论如何努力,成绩与努力之间的关联很弱,激励效果则会降低。为了保证成员的努力与成绩存在较强关联,税务系统限定了两个内容。(1)限定上下级均不可调整的指标内容。表 2 显示不可调整内容包括承接市局指标、共性指标和基层评价,这些指标占比从 2018 年的 65% 上升至 2020 年的 74%,激励组织成员通过自身努力影响最终考核成绩。(2)除非出现不可控因素,否则不在年度内对考核标准进行调整。2019 年,市税务局下发的绩效考核规则显示,“年度中间原则上不可对绩效指标的完成时间和考评标准进行调整。因不可控原因确需调整的,由相关单位提出申请,绩效办审核并提出意见后,报分管绩效工作的局领导审核,再由绩效办提交绩效考评委员会进行审议”。^① 考核标准的固定化,限制了随意调整的可能性,有助于形成可置信承诺。

最后,拉开成员间的业绩差异。如果所有人的绩效都保持相同,那么大家都成了优胜者,可能诱导偷懒。为了防止这种情况出现,税务系统进行了精心设计:(1)设立强制拉开差距的指标(简称“强制拉差指标”)。从表 2 可以看到,强制拉差指标占比不断变大,从 2018 年的 16.9% 上升至 2020 年的 39.2%。(2)加大不同档次的得分差距并限定优秀比例。表 2 显示领导评价指标分为优秀、良好和合格,对应标准分值是 100%、90% 和 80%,并限定优秀比例或人数不得超过 50% 或 10 个。(3)对考出差异的部门加分。为了防止一些领导做

① 内容摘自《2019 年市税务局对区税务局组织绩效考评规则》(内部文件)。

“老好人”,不愿拉开下级考核结果差距,上级通过设置额外的加分项目,激励大家尽量考出差异。

(三)奖优罚劣:精细化考核的目标

以上考核运作旨在精细化管理基础上进行奖优罚劣,达到激励效果,具体包括如下措施。第一,开展绩效讲评督促。区局每季度召开绩效讲评会议,公布市税务局对区局的考核成绩,并由局长对下一阶段考核提出相应要求。第二,制定奖优罚劣规则并公布结果。区局根据考核绩效,在干部选拔、评先评优和人才项目等方面进行奖优罚劣,并公布所有信息。其中奖优指:(1)排名前两位,次年将被优先提拔,年度增加2个优秀名额;(2)排名连续两年上升且累计上升位次最大,次年优先提拔,年度增加1个优秀名额;(3)连续两年第1段,年度个人绩效第1段比例从40%提高到50%。罚劣是指:(1)排名末两位,次年减少1个科级晋升名额,年度减少2个优秀名额;(2)排名连续两年下降且累计下降位次最大,次年减少1个科级晋升名额,年度减少1个优秀名额;(3)连续两年均排名末位,部门主要负责人认定为不适宜担任现职的,予以调整。

从上面看到,税务部门连续三年的考核,试图解决指标量化程度不高、只考部门不考个人、随意调整标准、分数差距小等问题。这一精细化考核类似淘汰赛模型(Lazear & Rosen, 1981),其核心机制是相对绩效考核,即比较成员业绩的相对排名而非绝对排名,以剔除一些共同的、未被观察的噪音因素,增强考核精确度。在此激励模式之下,部门及成员为了赢得考核而竞相努力,以取得比其他考核者更好的排名或段位。

四、限高、稳中与托底:模糊性考核的修正调整

以上只是呈现了考核的部分故事,田野观察也揭示了另一部分图景,即Y局后续通过一系列模糊化运作,把原来精细化的奖优罚劣规则调整为限高、稳中与托底规则。接下来,先介绍模糊性考核的运作特点,然后展现限高、稳中与托底的微观案例。

(一)模糊性考核的运作

首先,定性指标的制定相对粗放。在上节展现的精细化考核中,定量指标占比虽有波动但趋于上升,定性指标仅剩三成,集中在部门重点工作、领导评价和

一部分加减分项目中。^① 田野调查发现,为了增强对这些指标的控制,上级要求标准制定宜粗不宜细。2018 年 1 月,Y 局召开领导小组会议审议绩效考核方案。党委书记表示今年因为是机构改革,所以是比较特殊的一年,无论是市局明确的政策倾斜要求也好,还是照顾原国地税干部的心理也好,都需要绩效管理作为调节工具。分管绩效考核的副局长问考核科科长今年有哪些指标可以用来调节,考核科科长表示按照往年情况,像领导评价指标和一些加减分项目是可以调整的,尤其领导评价指标对最终考核能够起到一锤定音的作用。随后党委书记要求考评标准不用写得太详细,不然到年底不好调整。最后,考核科把领导评价指标如全面从严治党、重点任务落实、工作质效和工作创新等只作大方向描述,没有明确规定细节,以增强后续调整空间(田野资料 20180125)。

其次,保证指标的灵活调整。在精细化考核中,承接市局指标、共性指标和基层评价属于市税务局有明确要求的指标,区局无法调整(如表 3 所示)。部门重点工作、领导评价和加减分等指标则像一种自选动作,市局没有明确要求,区局可以调整。另外,一般认为定量指标没有调整空间,但实际观察表明一些定量指标也可以调整。表 3 显示承接市局指标(定量与定性)、共性指标(定量)、基层评价(定性)均无调整空间,而领导评价(定性)、部门重点工作(定性)及加减分指标(定量与定性)均有调整空间。这反映了指标调整不完全由定量或定性类型决定,而是受到上级的规定动作限制。如果上级没有明确框定,则下级可以同时定量或定性指标进行调整。

表 3		不可调整和可调整的指标比较						
被考核主体	一级指标	二级指标	考核主体	2018 年	2019 年	2020 年	指标类型	调整空间
区机关科室 (1000 分)	基础指标	承接市局指标	市税务局	300 分	300 分	300 分	定量/定性	否
		共性指标	职能科室	350 分	425 分	410 分	定量指标	否
		基层评价	税务所	/	25 分	30 分	定性指标	否
		部门重点工作	全局局长	200 分	200 分	200 分	定性指标	是
		领导评价	全局局长	150 分	50 分	60 分	定性指标	是
	加减分	加减分项目	职能科室	表彰类加分和批评类减分等			定量/定性	是

再次,考核标准非固定,以应对政策环境变化。田野调查发现,考核标准很难像精细化考核所述的那样保持年内不调整,而是至少面临两轮调整:一是在全

① 加减分有两种,第一种是定量指标类,即考核标准清晰,加减分规则明确,要求被考核主体提供明确佐证材料。第二种是定性指标类,没有说明考核标准,只需通过绩效考核委员会议审议通过即可,操作空间较大。

国两会结束后的五六月份,根据两会精神,调整部分考核方案;二是在上半年工作已经结束的八九月份,根据经济发展等情况调整税收政策,修正考核方案。2020年,国家税务总局发现,即使每年有季度性的指标优化调整,也无法满足经常变化的实际工作需要,于是设立了动态调整机制。2020年4月,Y局召开绩效讲评会,考核科科长通报了上级动态调整机制的消息,分管考核的副局长表态道,“既然上级正式提出了动态调整,那我们就只能做好思想准备,积极应对每一次调整,各部门应该加强与市局对口处室沟通,第一时间获知指标的考评动向,变被动为主动”(田野资料20200401)。

最后,增强领导对业绩结果的可控性。如果完全按照精细化考核的走向开展考核,可能产生一些非预期后果,例如差距过大、公平性不足和淘汰实际努力者等。为了增强对业绩结果的可控性,Y局形成了一些隐蔽的微妙过程。(1)领导评价指标的权重调整。表4显示2018年主要局领导、分管局领导和其他局领导在领导评价指标上的评分权重各占1/3。2019年和2020年,三者权重调整为40%、30%和30%,局长的决策权重提高了。(2)维护主要领导的决策优势。如表4所示,在部门重点工作指标的考评上,主要局领导的决策权重仍是40%,同样强调主要领导的决策优势。(3)先由副局长评分再由局长评分。对部门重点工作和领导评价指标进行评分时,局长要求先由各副局长评分,这样局长不仅可以参考各副局长的打分情况,还可以对各部门的总成绩进行最终把控。表4呈现了模糊性考核的整体情况。

表 4 模糊性考核的运作分析				
模糊性维度	具体表现	2018 年	2019 年	2020 年
指标非量化	定性指标比重	350(35%)	250(25%)	290(29%)
	制定相对粗放	领导评价指标:大方向描述,没有明确规定细节,增强调整空间。 部分加减分项目;对未列明工作事项,由区税务局绩效考评委员会认定可给予加分。		
定量定性指标皆可调整	可调整指标	350(35%)	250(25%)	260(26%)
	可调整内容	部门重点工作+领导评价+加减分+领导评鉴		
考核标准非固定	年内多次调整	年度的第二季度和第三季度根据宏观经济环境调整税务政策和税务考核方案		
	动态调整机制	国家税务总局根据年内重要会议、重要文件和局领导批示指示,动态调整考核方案		
业绩结果可控性	领导评价权重比	33.33%—33.33%—33.33%	40%—30%—30%	40%—30%—30%
	部门重点工作	主要局领导、分管局领导和其他局领导决策权重比保持为40%—30%—30%		
	局长最后评分	各副局长评分后再由局长最终评分,由此局长可对各部门总成绩进行最终把控		

续表 4

模糊性维度	具体表现	2018 年	2019 年	2020 年
结果运用的 宽严并济	结果考核规则	优秀排位(限高)、中间排位(稳中)、落后排位(托底)		
	结果部分公开	只公布优秀者的排名,隐去末位者的排名,通过刻意模糊化,保护业绩落后者的尊严。		

(二)限高、稳中与托底:模糊性考核的目标

田野调查发现,以上一系列灵活运作旨在限高、稳中和托底。限高指奖励优秀者时,除非其实力不可撼动,否则让排位靠前的成员,轮流拿排名前两位,以提高前列优秀者的积极性。稳中指激励中间层,有时让排位处于中间的成员轮流获得第一段/档/等/梯队的业绩,提高中间层的积极性。托底指保护末位者,除非其触犯底线,通常适当照顾,即业绩差者轮流拿末位,防止其连续两年处于末位而被调岗。图 2a 和图 2b 的示意图对比刻画了这一过程。此外,考核也强调部分而非完全公开排名信息。例如 2020 年 9 月,Y 局召开上半年绩效讲评会议。在开会前,分管局长要求考核科在制作会议 PPT 时只公开成绩第一段的部门,其他部门自行到绩效管理系统查询成绩(田野资料 20200903)。这一选择性公开旨在刻意模糊化,有助于保护业绩落后者的“脸面”,使组织氛围相对融洽。

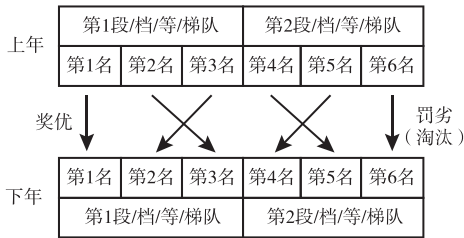


图 2a 理想意义的奖优罚劣

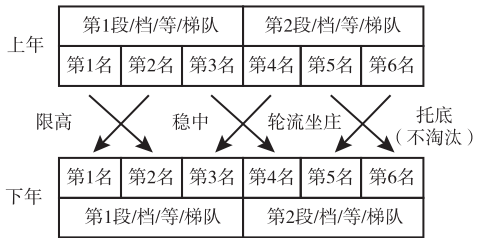


图 2b 实际运作的限高、稳中和托底

(三)案例呈现:限高、稳中和托底的微观运作

为便于进一步理解,接下来展现 2018 年至 2020 年我们在税务局的田野观察案例,以更好呈现考核的微观细节与过程。

1. 限高

案例 1:加减分方案的三易其稿。在 Y 局一次加分项目的分数分配中,A 科

获得加分 0.5 分,按照规则可以将 0.25 分分配给协办科室。^① 但是,在年终考核中,各科室之间存在直接竞争关系,因此,A 科在上报的第一稿加分方案中没有提报协办科室(如表 5 所示)。局长看到第一稿后电话联系 A 科室科长,询问加分事项的工作情况,得知除 A 科外,B、C、D 三个科室也参与了该项工作后,要求 A 科将这三个科室纳入协办,形成第二稿加分方案。2020 年 2 月,各部门总成绩出来后,局长又一次致电 A 科科长,“B、C、D 三个科室和你一样作为主办科室吧,这样协办科室可以列多几个部门,你把局里这些科室都加上,给大家都加分让大家开开心心的,这样明年做工作他们才有动力!放心吧,这个加分分配出去不会影响你的排名,就是因为不会影响排名我才让你改的”(田野资料 20200221)。在最后形成的第三稿加分方案中,A、B、C、D 四个科室共同作为主办部门,另有五个科室作为协办部门,共有九个部门得到加分。

表 5 三易其稿的限高方案		
方案	主办科室加分	协办科室加分
第一稿	A 科室 0.5 分	无
第二稿	A 科室 0.5 分	B、C 科室各 0.1 分,D 科室 0.05 分,共 0.25 分
第三稿	A 科室 0.2 分、B、C、D 科室各 0.1 分,共 0.5 分	E、F、G、H、I 科室各 0.05 分,共 0.25 分

案例 2:前面排位的轮流坐庄。业绩奖励通常有排名、划段与进步三个维度,其中划段比排名容纳的名额多,因此,当排名无法满足中上者的需求时,可以用段位间的轮流坐庄替代解决。某税务所所长透露,“第 1 段总共 5 个名额,常年处于优势的两个税务所占了 2 个,意味着剩下 7 个税务所要竞争 3 个第 1 段名额。领导考虑轮流坐庄,大家轮流拿第 1 段”(访谈资料 20200425)。其实这一限高逻辑不仅反映在区局,市税务局也有相似行为。2017 年,市税务局分上下半年对各区局的“信息公开”指标进行考核。由于 Y 局分管信息公开的副主任非常重视该项工作,上半年考评时 Y 局在全市 11 个区局中被评为第 1 段。下半年考评前,副主任按照考评标准计算出 Y 局积分依然遥遥领先,下半年也有望被评为第 1 段,但最终考评结果显示 Y 局下半年成绩仅被划为第 2 段。市税务局经办人员私人解释道,“该指标被领导拿来平衡了,你们 11 个区局轮流坐庄”(访谈资料 20180108)。

① 若市局对区局工作加分,主办科室可获得所有加分,协办科室获得加分分值的 50%。

2. 稳中

案例3:补偿机构改革的受损者。2018年机构改革后,国地税合并,原本两套领导班子合并为一套,领导编制数量减少。为照顾机构改革受损者的心理,市税务局要求各区税务局在绩效考核中对正职转副职(简称“正转副”)干部有所倾斜。2019年1月,考核科计算排名时发现,某“正转副”干部成绩排名靠后,划段位于第2段。考核科把这一难题汇报给局长,局长要求这位干部的分管局长调整领导评鉴分数,让其成绩进入第1段。考核科科长事后透露,“其实局长本意也不想让分管局领导调分,最好的办法是局长这边调分,但是这位干部的上半年成绩太低,只靠局长调整分数也救不了他”(访谈资料20190124)。另外,由于国税人员比地税人员多,为了照顾地税干部的心理,国税出身的局长在第1段的人数比例上也有所平衡。最终评分结果是原国税干部比原地税干部的第1段人数只多一个,这样不至于让两支队伍的差距太大,引起原地税干部的不满。

案例4:调整付出实际努力的部门排名。由于机构改革,从2018年第二季度开始,某后勤部门就着手筹备国地税合并事宜,工作业务涉及办公场地规划、工作设备采购和干部就餐协调等,加班时间远多于其他部门。这一后勤部门的工作属于辅助类业务,成绩排名长期处于第2段。由于工作量增加,局领导能够感受到该部门的实际付出,但是按照年初制定的考核方案,这些实际付出将不会得到体现。为此,局长决定进行调整,在进行领导评价指标评分时,特意给予该后勤部门优秀名额,使其三年来第一次登上第1段位置。

3. 托底

案例5:股员的轮流坐庄。2020年4月,考核科下发第一季度领导评鉴工作的通知,一个税务所所长打电话向考核科抱怨,“我觉得领导评鉴评分方式很不好啊!这个指标完全是把考核压力转移给主管领导嘛,直接由主管领导评鉴分数,来决定干部的排名和段次,这样会把矛盾全部集中到主管领导这里来的。天天都有人来我这里吵,我真的不知道要怎么办”(田野资料20200409)。实际上,许多税务所私下都采用轮流坐庄的办法,一位税务所所长在聊天中透露,“连续两年末位的又要换岗,那就只能我们主管领导去调整咯,去年你末位,今年就别让你末位”(访谈资料20200513)。^①

案例6:局长的苦恼。在2018年的年终考核中,某科长因不想承担责任,采取抽签方式决定部门末位人选。这一消息被其他部门知晓并扩散,在Y局引发

^① 《2019年市税务局对区税务局组织绩效考评规则》规定,如果各部门年度绩效考评连续两年均排名末位的,若主要负责人在该部门任职两年以上,经综合研判,可认定其为不适宜担任现职,予以调整。

了不小的轰动。这件事传到局长耳中,谁也想不到局长竟非常高兴!因为他一直苦于寻找科长末位人选,现在终于可以“枪打出头鸟”。随后局长在全局性会议上佯装不高兴,严肃批评这位科长,“这是原则性错误!”最终该科长当年的个人绩效被评为末位。2019年年终考核时,局长又陷入考虑末位人选的苦恼,在办公室来回踱步。临近考核截止日期,局长要求人事科把近两年已晋升的名单整理出来,从中挑选一位去年已晋升但今年工作完成一般的科长作为末位,并跟分管局长解释,“评为末位,不是代表他工作做得差,只是权衡后,选了一个影响最小的人作为末位,毕竟他去年刚晋升,近几年也不会再晋升了。如果今年晋升的是他,我肯定会考虑的,不会让他末位”(田野资料 20200106)。

以上一系列案例表明,组织的模糊运作为考核的修正调整提供了灵活操作空间。一方面,通过调整部门重点工作、领导评价和加减分等指标对一些特殊部门进行照顾,缩小部门间分差,减少部门间的考评矛盾。另一方面,通过调节领导评鉴指标对一些特殊个体进行补偿,体现其实际努力水平。由此可见,考核并不完全是精细化思路,它需要保留一定程度的模糊性,以对理想意义的奖优罚劣规则进行调整,形成更契合实际工作环境的限高、稳中和托底规则。

五、制度环境、刻意模糊与权威控制:一个解释框架

上面介绍了考核的灵活运作,但并未提供理论解释。一个需要回答的问题是考核为什么要留有口子、刻意保留部分模糊性?我们首先分析制度环境压力,进而揭示精细化考核的困境,最后考察模糊性考核的功能。

(一)制度环境压力

考核离不开制度环境的影响。政府在考核时至少面临三个方面的制度环境:相对绩效的弱比较、上下级人事关系的强连带和政策环境的多变性。这些特征使其区别于企业考核,反映了政府科层组织的特殊处境。

1. 弱比较的结构环境

在理想状态中,组织可通过制度设计保证各个部门进行相对比较,产生令人信服的考核成绩。但是田野调查表明,组织受到一些结构因素影响,难以做到精细的绩效比较。

例如在业务科室中,一些税务所因管辖企业较多,绩效考核具备先天优势。

此时,业绩与成员努力关联不大,而是跟所在组织结构中的地位有关。同理,在行政科室中,办公室、人事科等部门经常与领导接触,这一结构特性使得局领导更能够直观感受到他们的工作付出。与之相比,后勤、财务等部门每年工作大同小异,被局领导习惯性地视为辅助部门。另外,纪检部门一直处于结构劣势,在基层评价的考核中常年排名倒数。该部门科长表示,“没办法,我们部门的工作很容易得罪人,你看各种揭发检举,还有今年严查酒驾、醉驾,这些都是吃力不讨好的事情,税务所怎么会给我们高分呢?”(访谈资料 20191227)

上述内容表明,在相同努力的条件下,一些科室因常年占据结构的核心位置,比那些处于结构边缘位置的科室更易获得较好的绩效。这就意味着如果仅按照精细化考核思路,不采用其他手段调节,将导致成员的努力付出与回报不一致,引发不公平问题。

2. 强连带的人事环境

按照理想的组织形态,政府内部的上下级关系至少有两特点:一是人事流动是流畅的,可以按照奖优罚劣原则进行考核结果运用,推进干部能上能下;二是上下级是一种严格的科层或等级关系,下级对业务独立负责,如果表现不佳,上级不负有责任,全然由下级自负全责。

但是,在实际运作中这两个假定均不能成立。首先,与私人部门相比,政府等公共部门的人事流动性相对较低,呈现非违规不退出的“终身制”或“铁饭碗”。其次,上下级之间是一种相对扁平的利益和责任连带关系,而非纯粹的等级化关系。这一关系意味着上级否定下级可能会打击下级士气,削弱下级未来的政策执行配合或支持。当面临问责时,直接上下级是一个隐性的责任共同体,下级表现不佳,表明上级领导不力。在此约束下,上级会尽可能保护下级。

因此,在上述人事特征下,组织无法完全奖优罚劣,并严格做到人事能上能下,相反还要通过各种举措力求同事间的关系融洽。例如案例1中加减分方案三易其稿和案例2中的税务所轮流坐庄,均旨在削弱过度竞争,防止人事关系恶化。

3. 非稳定的政策环境

一般认为,相比企业面对不确定的市场环境而言,政府面对着更加稳定的政策环境。由此推之,在政府考核中,年初制定标准后,全年可以保持不变,后续按照既定考核政策稳定推进即可。

然而,这只是部分反映政府的环境特征,在税务部门实际考核中,年初制定的考核政策常常受到上级政府的政策调整影响。首先,政府每年会在不同季度进行政策调整。例如税务部门年初制定的考核方案在第二季度会根据全国两会

精神进行修正,在第三季度再根据上半年完成情况进行调整。另外,动态调整机制的出台意味着考核方案无论是时间节点还是调整内容都充满变化。除此之外,案例3和案例4表明机构改革也是一种不能忽略的政策环境,它通常在若干年之内会发生变化。一旦变化,除了当时的机构合并或增减,还牵涉到后续几年的政策磨合。这意味着考核制度无法保持静态、单一或稳定,而是需要不断修正调整以满足多方关切。

总之,上述内容反映政府内部的政策环境充满非稳定性,除了年内的季度政策变化、年内不定期的动态政策调整,还面临若干年一变的机构改革。这些政策环境的不稳定性意味着年初制定后不再调整的静态精细化考核在变化的环境面前将显得较为被动。

(二)精细化考核的困境

上述内容表明,在制度环境影响下,精细化考核面临一系列问题,包括相对绩效的弱比较引发不公平、过度竞争恶化同事关系、只关注优胜者导致整体激励不足、静态考核无法应对政策环境变化,等等。一般而言,当面临上述问题时,组织可以通过权威介入,以补偿处于结构劣势的部门、削弱过度竞争、形成整体激励和提高考核弹性等方式来解决。然而,在精细化考核背景下,一个管理困境是组织权威不仅不能增加,反而被不断削弱,这背后至少存在三个复杂而隐微的组织机制。

1. 规则的平等化机制

技术化的精细考核具有普遍主义原则和平等化倾向,不仅约束下级,也约束上级,这使得当负责考核的上级进行灵活调整时面临着规则约束。正如韦伯所指出的,“理性化的规则让正式化的非人格化因素占主导地位……不管对象是谁,每个个体都获得了正式过程的平等对待,科层体系的发展大大推动了身份地位的平等,这是理性化规则发展的一般趋势”(Weber, 1978: 225 - 226)。也有研究指出,“规则起到了使成员的身份地位极大地趋于平等的作用;那些运用规则并监督规则执行的人,同那些受规则支配的人一样会受到规则的制约”(孔飞力, 1999: 250 - 251)。法国社会学家克罗泽尔(Michel Crozier)强调,理性化的规则存在一种悖论,它在约束下级的同时,也在限制上级权力(Crozier, 1964)。

2. 由权威控制者变成数据统计员

一般认为,上级可以在精细化考核中掌握主动性。但是,在精细化考核背景下,决策已由数据主导,上级变成了一个普通的数据统计员,其权威遭到统计数

据的削弱。正如布劳对美国政府考核的研究所指出的,评估下属是上级的主要职责,如果评估完全依赖于统计指标,那么上级的职责就被简化成一个普通职员的工作,这不仅使上级的工作固化,还会破坏他在下属中的权威(Blau, 1955)。回到中国语境,陶然等人在批评周黎安的晋升锦标赛模型时强调,“中国这样一个重视集权的政治体系中,如果下级官员的政治升迁与客观的、可衡量的竞赛指标挂钩,那么上级领导将在很大程度上丧失其在官员任命上的最终控制权,而对下级官员任命的最终控制权,恰恰是当前政治体制的一个基本特征”(陶然等, 2010:17)。由此可见,精细化考核将使权威下移或平均化而非上移,上级的身份由监督控制者变成一个数目字管理员或文书操作者,权威由此旁落。

3. 下级适应性预期限制上级调整

为保证组织在常规环境中的稳定运作而设立的精细化考核在增强下级适应性预期的同时,也使得下级在环境变化、上级须调整考核方案时不愿意改变原来的游戏规则。换言之,在精细化考核条件下,成员熟悉了考核规则,知晓了自己的努力方向与范围,倘若上级突然修订或灵活调整,将破坏成员的适应性预期,由此触发下级抱怨甚至抵制。受制于科层压力,下级虽然不会公开抵制,但可以采取消极反抗或讨价还价作为应对措施。故此,精细化考核在增强下级稳定预期的同时,也限制了上级的灵活调整能力。正如孔飞力的研究所指出的那样,“规章条例在为下级的职责划定某种边界的同时,也为他们提供了一定的保护,使他们得以对抗来自上级的专断要求”(孔飞力, 1999:250)。

总之,不可否认,精细化考核对组织运作的贡献巨大,它是一种指挥棒和风向标,确保了成员的稳定预期和组织的常规运作。但是,精细化考核也会带来一系列问题。(1)对下级而言,相对绩效考核受到结构优劣影响,导致个人成绩与实际努力不完全匹配,引发不公平问题。(2)对上级而言,上级由权威控制者变成了数据统计员,权威旁落,无法根据环境对考核规则进行调整。(3)对组织而言,精细化考核偏重激励优胜者而无法实现整体激励,可能诱发过度竞争而破坏融洽的组织文化,更加麻烦的是组织的权威下移甚至平均化,使得组织缺乏抓手,降低了在复杂环境下的问题解决能力。

(三)引入模糊性考核:刻意模糊与权威控制

田野调查发现,当精细化考核面临问题时,组织主动或刻意嵌入模糊性,生成灵活性,用于权威调节,纠正精细化考核带来的诸种问题,提高应对环境的组织能力。下文将分析组织通过嵌入模糊性获取权威控制的微观机制。

1. 从信息对称到信息不对称

精细化考核意在保证上下级的信息对称,以生成下级的稳定预期和组织的常规运作。与之不同,模糊化考核则是有意让上级掌握一些不为下级知晓的信息,制造信息不对称。例如局领导要求定性指标制定宜粗不宜细、部分指标留有调整空间、考核标准非固定化等举措均反映了与传统的委托代理理论不同的地方(Holmstrom & Milgrom, 1991),后者意在消除上下级之间的信息不对称,而这里则是刻意制造上下级之间的信息不对称。

从信息对称到信息不对称,看起来是信息问题,实则是权威运作问题,背后反映了信息与权威的微妙关系。信息对称意在将纷乱复杂的考核过程标准化、固定化与清晰化,最终增强考核的可信度。但实际上还需要一个信息不对称的过程。通过信息不对称生成权威运作的空隙,便于上级对完全清晰标准化所带来的弊病进行纠偏,缓解环境变化带来的组织紧张。在这里,信息不对称生成制度缝隙主要借助于两个微妙的过程:一是信息类型由同质走向异质,造成多向度的信息过程,使上级掌握着不为下级知晓的关键异质信息,获得对整个考核的控制、调节与引导;二是信息加工由自动提取数据变成自动与人为双头并进,为上级的信息加工和信息选择提供机会。

2. 从事前约定到执行博弈

考核并非是事前约定后就不再更改,而是伴随着执行过程中的种种博弈。例如在案例中,为了让更多科室有干劲,局长在加分环节与那些想独享的科长博弈,希望他们让步;为了弥补机构改革的受损者,局长劝说下属调整分数,让受损者的排位上升;为了补偿在国地税合并中付出转多、贡献突出的后勤部门,局长需要调整领导评价指标。

以上表明,考核制度并非是静态的,而是一个动态运作过程。由于原初制度无法考虑和覆盖所有方面,只能让实施者进行博弈,从而赋予其足够的调适空间。否则,一切按单向度的制度规定,容易造成考核的固化和教条虚设,无法应对组织内外环境的变化与需要。因此,制度决策过程和执行过程是一个整体,事前约定的制度决策过程塑造了相对稳定的预期,而事中事后的制度执行过程通过政策的再界定与再调整,让制度具有充分的灵活性和弹性,保有组织在具体实践过程中的权宜性。正如马奇(March, 1988)所言,执行过程是组织决策过程的延续,一个没有考虑执行过程的决策是一个不完整的决策。

3. 从前台展演到后台运作

精细化考核类似于前台程序标准化运作的展演,而模糊性考核则是后台组

织调节的非制度化呈现。在上述案例中,考核结果的部分公开、增强领导者的决策权重以及副局长先评而局长后评等方式均体现了组织的后台运作逻辑。相比前台严格的形式化程序,后台博弈往往更加复杂,需要考虑平衡诸多因素,包括外部政策环境多变、业绩测量不准所引发的公平问题以及非流动不退出的政府特殊人事关系等。在此情况下,如何通过后台进行考核的局部调整、制度扩展和优化就显得非常重要。

这背后隐含的道理是,在复杂制度环境之下,组织运作需要同时拥有前后台设置(Goffman,1969)。前台的精细化考核将考核的激励方向以最为简化和明晰的方式呈现出来,最大程度减少复杂性、随意性和波动性。但是,只有前台是远远不够的,后台可以对信息进行重新辨识、筛选和组合,建立一套属于组织或上级有辨识度的考核系统。通过前后台相互补充的组织考核过程,最终造就一种多主体结构 and 多重性空间,增强考核制度运行中的变异性 and 权威控制。

4. 从事本主义到诠释解释

从案例中可以看到,在组织考核中,领导层之间和上下级之间会频繁沟通、互动和解释,例如领导小组的专门会议、末位人选挑选的商量和考核结果多次调整的沟通。这反映出组织考核过程并不仅仅是精细化考核的事本呈现(Weber,1978),还体现出组织自身的意义解释和诠释(Daft & Weick,1984)。

由此看到,不能把考核过程简单理解为一种量化文本的照念照搬,而要注意到在具体实践过程中嵌入了解释和诠释,它至少有三大功能。首先,领导层之间的沟通互动,有助于增进组织上层的共识,形成权威抓手,避免对一些重点问题做碎片化的分散决策。其次,上级对下级的日常沟通,通过运用关键事件和典型案例对一些矛盾、困境和难点进行重点讲解,有利于强化决策参与者和接收者的感受力和信任度。最后,上级对业绩不满者“做工作”,带有一种安全阀性质的制度化解功能,发送了上级对下级抱怨予以重视或关切的信号。通过一系列非事本化的组织解释过程,信息得以重新编码、筛选和组合,多级、多重和多向度的诠释得以展开,给组织留出了足够的空间。总之,这是一个文本再诠释或再演绎的过程,也可视为意义建构、意义衍生和权威控制过程,体现了实践的不确定性。

上文展现了一系列通过模糊运作获取权威控制的机制,它把可能陷入僵化的考核实践重新活跃起来,打破过度精细化考核的惯常边界,留出独特的韧性 or 弹性的空间,从而为整体激励赋予真正的气息和活力。由此也可以看出,考核并不是一个简单的考核问题,而是一种治理术(governmentality),它由信息重新编码、执行博弈、前后台交替和诠释解释等一套复杂操作组合而成,形成了信息场

域、博弈场域、空间场域和话语场域，为组织留出大量弹性空间和余地，这有助于提高地方政府在复杂人事环境下的问题解决能力。

六、总结：不完全考核

组织经济学的不完全契约理论指出，受到有限理性和机会主义等因素影响，组织订立的契约通常是不完全的（Hart & Moore, 2004, 2008）。从广义角度看，我们可以把真实世界的政府考核理解为一种契约，即上级基于精细化考核和下级签订了付出与回报相关联的激励契约。从以上分析看到，这一契约也是不完全的。受到制度环境（弱比较的结构环境、强连带的人事环境和非稳定的政策环境）影响，上级无法按照精细化考核达成契约，由此引入模糊性考核，借由信息对称到信息不对称、从事前约定到执行博弈、从前台展演到后台运作、从事本过程到诠释解释等一系列机制获取剩余控制权，解决由纯粹精细化考核所引发的一系列问题，如业绩不可比削弱公平、局部激励而非整体激励、过度竞争恶化同事关系、静态规则难以应对环境变化和上级权威削弱，等等。由此，我们可以把这一考核特征概括为“不完全考核”。图3是本文的解释框架。进一步看，这一研究可以带来三个方面的思考。

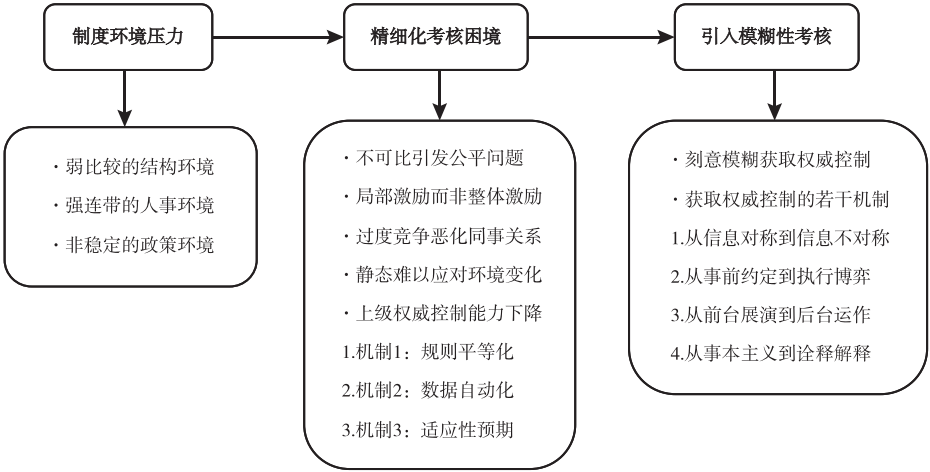


图3 制度环境、刻意模糊与权威控制：一个解释框架

第一,反思单向度的科学管理。在现代社会,理性的计算精神、道德中立的技术观念和工程化的社会管理模式已经渗透到各类组织之中,单向度的科学管理日益明显(Taylor,1911)。在这一技术理性不断扩张和蔓延时,现代组织越来越变成一个没有“人”的组织,大家只是遵照技术程序和规则行事,毫无独立思考和价值理念,呈现按部就班、谨慎行事和严守规章的非人格化特征,最终把手段当作目标,以工具价值代替最终价值。有学者在研究企业的绩效考核时指出,“没有一个测量系统能够完美捕捉绩效的所有重要方面和成员需要为组织付出并促进组织成功的所有行为。指标测量只是一个大纲,对行为起引导作用而已,它在实际执行中不可能有如此强大的作用,以至于可以替代组织所必需的判断与智慧”(Pfeffer & Sutton,2000:153)。除了企业领域,近年来国内外教育领域也出现了对于过度单一化指标评价体系的反思,例如,国家最近发文要扭转唯指标论的教育评价导向。^①另外,哈佛大学教授桑德尔(Michael Sandel)从道德哲学角度反思优绩主义给教育带来的非预料性后果,强调优绩主义的成功观虽然选拔了优秀人才,但也造成了精英的傲慢,使成功者深信其成功是自身努力的结果,忘记了背后帮助其成功的结构和运气,进而轻视比自己不幸和缺乏资格的人(Sandel,2020)。这一优绩主义将固化社会阶层、扩大贫富差距和助长阶层对立,因此应该反思教育的本质追求、工作的价值尊严和成功的意义所在。总之,本文所关注的精细化考核,虽然暂未导致上述组织后果出现,但至少也从另外一个侧面表明,需要反思当下过度技术化、统计化的现代管理趋势,应该把人的主体地位、创造性及价值理念重新找回来,发挥其主动性,以应对过度理性化所可能引发的系统性后果。

第二,重新看待考核的偏差行为。已有文献指出,政府考核过程常常存在许多偏差现象,例如轮流坐庄、模糊指标以及考核结果部分公开等,认为这些偏差行为降低了考核有效性(姜晓萍、马凯利,2005;刘昕,2007;唐健、方振邦,2017)。本文研究则表明,如果跳出规范视角进入地方政府的实际运作情景中,则会发现上述特征并非是简单的偏差行为,而是复杂制度环境的必然产物,甚至在某种意义上具有对技术主义进行纠偏的功能与作用。首先,轮流坐庄看似平均主义,实质是对前位、中位和末位进行适当调整,让不同层次的队伍都有盼头与期望,形成整体激励。其次,模糊指标表面看是一种考核的主观化,实质上是为了解决政策环境变化导致考核结果没有反映一些部门或个体实际努力的问题。组织通过

^① 中共中央、国务院,2020,《深化新时代教育评价改革总体方案》,中国政府网(https://www.gov.cn/gongbao/content/2020/content_5554488.htm)。

模糊指标提供了考核的再调整、再界定与再修复功能。最后,考核结果的部分公开并非仅仅为了“暗箱操作”,它在很大程度上是为了弱化比较、防止过度竞争和保护落后者,是一种组织关怀之举。因此,一些模糊过程与组织目标之间不一定是扞格不入的,反而会有助于达成组织的整体目标。

第三,揭示模糊性运用的边界条件。尽管模糊性在考核中不可或缺,但它也有适用条件。当考核指标与成员努力关联较强时,指标要清晰,自由裁量权要变得较小,这样才能形成强激励。周黎安(2007)的晋升锦标赛正是在这个意义上讲的考核激励。当考核指标与成员努力关联较弱时,指标要模糊,自由裁量权要变大些,这样才能灵活调整,根据实际情况弥补利益受损者。这类似于环境减排研究所揭示的弱排名激励(练宏,2016)。本文关注的税务部门考核贴近后者而非前者,强调的是在政府科层组织的复合环境影响下,精细化考核无法保证成员努力与回报的相对一致,为此嵌入模糊性,生成权威控制,根据实际情况进行平衡。因此,这一研究有助于明确模糊性运用的边界条件。^①

当然,限于篇幅,本文也还存在许多不足,例如没有进一步讨论精细化考核与模糊性考核的平衡问题。把握这两者的边界是一项高超的技艺,需要细致揣摩拿捏。如果产生过度的模糊性,或随意更改践踏,将使人们无所适从。进一步来看,本文沿着模糊性考核与精细化考核的路径推进,但可能过于侧重分析模糊性考核,对精细化考核的功能意义及其与模糊性考核之间的张力分析尚嫌不足,也没能细化分析“精细—模糊”和“强激励—弱激励”的各种组合关系。笔者期待未来能够进一步研究和拓展上述议题。

参考文献:

- 蔡立辉、刘晓洋,2016,《政府绩效评估中的博弈及其影响》,《社会科学战线》第3期。
- 崔述强、王红、崔萍、闫明、陈明,2006,《中国地方政府绩效评估指标体系探讨》,《统计研究》第3期。
- 黄仁宇,2015,《中国大历史》,北京:生活·读书·新知三联书店。
- 姜晓萍、马凯利,2005,《我国公务员绩效考核的困境及其对策分析》,《社会科学研究》第1期。
- 孔飞力,1999,《叫魂:1768年中国妖术大恐慌》,陈兼、刘昶译,上海:上海三联书店。
- 练宏,2016,《弱排名激励的社会学分析——以环保部门为例》,《中国社会科学》第1期。
- 刘昕,2007,《我国政府绩效管理中亟待梳理的几个关键问题》,《中国行政管理》第4期。
- 尚虎平、赵盼盼,2012,《绩效评估模式泛滥与绩效不彰困境——基于42个案例的分析》,《中国行政管理》第11期。
- 唐健、方振邦,2017,《公务员绩效考核:现实困境、国外经验与本土路径》,《行政科学论坛》第11期。

^① 感谢匿审专家提出的建议,启发本文进一步分析模糊性运用的边界或约束条件。

- 陶然、苏福兵、陆曦、朱昱铭,2010,《经济增长能够带来晋升吗?——对晋升锦标赛理论的逻辑挑战与省级实证重估》,《管理世界》第12期。
- 王汉生、王一鸽,2009,《目标管理责任制:农村基层政权的实践逻辑》,《社会学研究》第2期。
- 文宏、郝郁青,2017,《关键绩效指标考核下的地方政府行为选择分析——基于2011-2016年黑龙江煤监的数据资料》,《北京行政学院学报》第2期。
- 周黎安,2007,《中国地方官员的晋升锦标赛模式研究》,《经济研究》第7期。
- ,2014,《行政发包制》,《社会》第6期。
- 周雪光、练宏,2012,《中国政府的治理模式:一个“控制权”理论》,《社会学研究》第5期。
- Blau, Peter M. 1955, *The Dynamics of Bureaucracy*. Chicago: University of Chicago Press.
- Callahan, Raymond 1962, *Education and the Cult of Efficiency*. Chicago: University of Chicago Press.
- Crozier, Michel 1964, *Bureaucratic Phenomenon*. Chicago: University of Chicago Press.
- Daft, Richard L. & Karl E. Weick 1984, "Toward a Model of Organizations as Interpretation Systems." *Academy of Management Review* 9(2).
- Goffman, E. 1969, *The Presentation of Self in Everyday Life*. New York: Penguin Press.
- Guillen, Mauro 1994, *Models of Management: Work, Authority, and Organization in a Comparative Perspective*. Chicago: University of Chicago Press.
- Hart, Oliver & John Moore 2004, "Agreeing Now to Agree Later: Contracts that Rule Out But Do not Rule In." Working Papers of Harvard Institute of Economic Research.
- 2008, "Contracts as Reference Points." *The Quarterly Journal of Economics* 123(1).
- Holmstrom, Bengt & Paul Milgrom 1991, "Multitask Principal-agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics & Organization* 7(S1).
- Lazear, Edward P. & Sherwin Rosen 1981, "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89(5).
- March, James G. 1988, *Decisions and Organizations*. New York: Blackwell.
- North, Douglass & Robert Thomas 1973, *The Rise of the Western World*. Cambridge: Cambridge University Press.
- Pfeffer, Jeffrey & Robert I. Sutton 2000, *The Knowing-Doing Gap: How Smart Companies Turn Knowledge into Action*. Massachusetts: Harvard Business School Press.
- Sandel, Michael J. 2020, *The Tyranny of Merit: What's Become of the Common Good?* New York: Farrar, Straus and Giroux.
- Scott, James C. 1998, *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale University Press.
- Taylor, Frederick Winslow 1911, *The Principles of Scientific Management*. New York: Harper.
- Weber, Max 1978, *Economy and Society*. Berkeley: University of California Press.

作者单位:中山大学中国公共管理研究中心、
中山大学政治与公共事务管理学院
责任编辑:徐宗阳