

国家自然科学基金学科规划对学科发展的影响： 基于文献计量的分析与思考

陈思华^{1,2}, 邱焱¹, 霍红^{2*}

1. 江西财经大学信息管理学院, 南昌 330013;

2. 国家自然科学基金委员会管理科学部, 北京 100085

* 联系人, E-mail: huohong@nsfc.gov.cn

自20世纪50年代中国制定第一个科技发展规划, 即《1956~1967年全国科学技术发展远景规划》以来, 中国一直使用“规划模式”来推动科学技术的发展^[1]。

国家自然科学基金委员会(以下简称基金委)是我国科学研究的重点资助单位。在中国的科技资助体系中, 基金委有其独特的资助格局和规划路径, 重点支持基础研究, 同时又面向国家重大需求。为了使学科发展能更及时地适应时代的发展与变化, 并更精准地契合国家重大需求, 在国家制定5年发展规划时, 基金委各科学部及其学科处同步开展学科发展战略研究, 制定出未来5年的学科发展规划与若干个优先资助领域。然后, 围绕该规划陆续发布重点项目和重大项目指南, 以推动优先资助领域的发展。与此同时, 各科学部鼓励广大科研工作者结合当前国际科学难题与中国实际需求进行思考和自由探索, 并提交申请。这种自由选题类的研究主要以面上项目、青年科学基金项目(以下简称青年项目)、地区科学基金项目(以下简称地区项目)等形式予以资助。

科技发展规划在集中国内科研力量攻克国家重点科研难题、培养关键领域科研队伍与人才等方面发挥了重要作用^[2]。目前, 关于科技发展规划(包括基金委的学科规划)的研究大多集中在对规划制定的理论、方法和组织模式的探讨^[3-5], 鲜有对科技发展规划效果的定量评估, 而现有关于科技发展的定量评价研究又缺乏对科技发展规划的分析和思考^[6-8]。

对于基金委科学基金体系, 在每一轮5年学科规划执行结束后, 基金管理者应该思考的重要问题之一: 学科规划对学科各领域的发展产生了怎样的影响? 为了回答这一问题, 本文以基金委管理科学部下设置的4个学科之一——管理科学与工程学科(学科代码为G01)为例, 分析科学基金“十三五”学科规划对学科发展的影响。

管理科学与工程学科“十三五”规划于2016年开始实施, 2020年结束。期间资助了23项重大项目(含重大专项)、50项重点项目(含重点专项)、1101项面上项目、1106项青年项目和158项地区项目。如前所述, 获批项目的选题来源可大致分



霍红 国家自然科学基金委员会研究员, 环境科学与工程博士。曾在美国阿岗国家实验室、清华大学等单位从事能源与环境管理领域的科学研究工作, 现负责管理科学与工程学科的基金管理与学科战略规划工作。

为两类, 即学科规划和自由探索, 解析两者的差异有助于分析学科规划的影响。此外, 国内外学者在国际学术期刊上发表了大量的学术文献。可以这么认为, 基金项目是基金管理者与广大学者基于对领域前沿热点与中国实际需求的思考而形成的研究计划。学术文献作为科学研究产出最为重要的载体, 则真正体现了一个学科实际的发展趋势和方向。学科规划、学者个体思考、学术文献三者之间具有一定的关联并互相影响, 对其进行分析不但有助于从宏观视角把握学科研究热点的演化规律, 还能增强对科学基金资助特点的认识, 更能加深对学科规划意义的理解。这对基金管理者及本领域的学者均具有重要的启示意义。

综上所述, 本文首先作如下假设: 以重点和重大项目代表学科规划; 以面上、青年和地区(以下简称“面青地”)项目代表广大学者的个体思考; 以代表性国际学术文献代表国际学术前沿热点。然后, 通过定量分析重点和重大项目、面青地项目和学术文献三者之间的关系来探讨学科规划、学者个体思考、国际学术前沿热点三者的相关性, 研究基金项目与学术文献研究热点分布的契合度及其变化趋势, 探索学科规划对学科发展的影响路径与推动作用。

1 研究方法

图1为本文的研究路线: (1) 获取管理科学与工程学科代表性国际学术文献, 以及该学科基金项目的标题、关键词和摘要等信息; (2) 应用基于潜在狄利克雷分布(latent Dirichlet

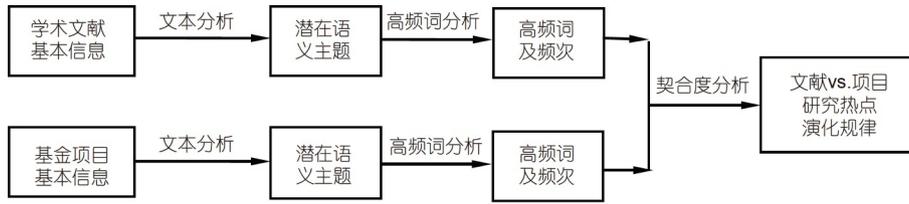


图1 研究路线

Figure 1 Research roadmap

allocation, LDA)模型的文本分析方法从上述信息中提取隐含在文本中的结构化主题,即潜在语义主题;(3)对潜在语义主题进行高频词分析,得到分组的高频词及其频次;(4)引入契合度概念,对学术文献与基金项目的研究热点进行契合度分析。

1.1 学术文献与基金项目的选取

本文采用的学术文献从管理科学与工程“十四五”学科发展战略项目课题组基于文献分析和专家调研等方式确定的46本管理科学与工程领域最具前沿性和代表性学术期刊中选取^[1](具体见表S1)。为了在更长的时间范围内探讨学术文献研究热点变化情况以充分挖掘基金项目与国际研究热点相关性的变化趋势,我们将学术文献分析的时间范围扩大至2011~2020年。由此,共选取了2011~2020年46本期刊中所有学术文献进行分析,共38332篇。

本文采用的基金项目为2016~2020年基金委管理科学与工程学科资助的面上、青年、地区、重点(含重点专项)和重大项目(含重大专项),共计2438项(具体见表S2)。由于本文选取的学术文献均为英文文献,为更准确地进行文本分析,对基金项目提取了项目负责人提供的英文标题、英文关键字与英文摘要,用于后续分析。

1.2 LDA文本分析方法

文本分析方法可对目标文档中的潜在语义主题进行自动识别、提取、总结和归纳。文本分析方法始于20世纪90年代,历经20年发展出潜在语义索引(latent semantic indexing, LSI)模型^[9]、概率潜在语义索引(probabilistic latent semantic indexing, PLSI)模型^[10]、LDA模型^[11]等多种模型。其中, LDA模型功能最为完善,不需要任何关于文本的背景知识,可以分析“一词多义”和“一义多词”的语言现象,能够进行超越词汇层次的逻辑分析,使得语义层次上的文本分析成为可能^[12]。LDA模型已成功应用于文本主题建模等多个领域^[13-15],也被广泛用于学术研究热点演化分析^[6,8]。本文选择LDA模型进行文本分析。

LDA模型应用贝叶斯方法计算目标文档的潜在语义主

题(以下简称主题)分布,以及每个主题中单词(以下简称主题词)的分布。LDA模型需预设目标文档的主题数量,为了获取其合理值,本文从总文档中随机抽取15%的文档进行测试,邀请管理科学与工程领域的教授、副教授以及博士研究生来验证抽取文档中最佳主题数量的范围区间。验证结果表明,当该值设置大于5时,会提取到一些与文档主旨不符的主题,小于5时,则会漏掉关键信息,因此本文将文档的主题数量设定为5。表S3为LDA模型生成的学术文献和基金项目的主题及主题词数量。

1.3 高频词分析

为了得到可表征研究热点的高频词,需对LDA模型生成的主题进行高频词分析,这一步骤包括高频词汇的判定与筛选,以及高频词汇的分组。

(i) 高频词汇的判定与筛选。首先,统计各学术文献与基金项目的主题中各单词和词组的总频次,再基于高频词汇的判定标准筛选出高频词汇。其中,在高频词汇的判定标准设定中,参考“秘书问题”的选取策略^[16,17],将高频词汇的阈值设定为 $1/e$ (约37%),即词频排在前 $1/e$ 的词汇即为高频词汇。由此得出,对于基金项目,高频词汇是词频大于10次的单词和大于5次的词组;对于学术文献,高频词汇是词频大于50次的单词和频次大于20次的词组。

(ii) 高频词汇的分组。由于学者写作习惯不同,高频词汇集合中含有许多具有相同意义、但文字形式不同的近义词和同义词,本文应用自然语言工具包(natural language toolkit, NLTK)的近义词词库工具对高频词汇集合进行无监督的分组^[18]。过程如下:首先建立一个由所有高频词汇组成的初始向量组,应用NLTK的近义词词库工具建立近义词关系矩阵,对高频词向量组进行循环的近义词判定与合并,直到得到稳定不变的高频词向量组。最终得到12个高频词汇频次较多的组和若干个高频词汇频次较少的组。对于12个高频词汇频次较多的组,根据其高频词汇体现出的研究领域特征,将其定义为GP01复杂系统组、GP02决策与博弈组、GP03预测与评价组、GP04金融工程组、GP05信息技术组、GP06管理统计组、GP07运筹与优化组、GP08心理与行为组、GP09风险管

1) 由于发表在国内知名期刊的研究大多数受国家自然科学基金资助,为避免陷入同义重复和自证循环,同时考虑到46本学术期刊足够代表国际学术前沿,本文未将国内期刊纳入学术文献分析中

理组、GP10供应链管理组、GP11系统工程组、GP12交通管理组。对于高频词汇频次较少的组,将其合并,定义为GP13其他组。GP13其他组的高频词汇频次数量占高频词汇总频次数量的4%~8%,由于其比例不高且领域繁多,后续分析将不对GP13其他组进行专门讨论。表S4统计了13个高频词汇分组的高频词汇数量和频次数量。

1.4 基金项目与学术文献研究热点的契合度分析

分组后的高频词汇可代表该组学术领域的研究热点,学术文献与基金项目中高频词汇的频次差异可体现两者研究热点之间的相似点与差别。为此,引入契合度(GSV)这一概念,用于表征与计算学术文献与基金项目研究热点的契合程度。GSV的算法基于学术文献与基金项目中高频词汇的频次差异(SV):

$$\begin{cases}
 SV_{y_1,y_2,y_3,y_4}(w_i) \\
 = \frac{[J_{y_1,y_2}(w_i) - P_{y_3,y_4}(w_i)]^2}{J_{y_1,y_2}(w_i) + P_{y_3,y_4}(w_i)}, \\
 GSV_{y_1,y_2,y_3,y_4}(GPx) \\
 = 1 - \sqrt{\frac{\sum_{i=1}^{C(GPx)} SV_{y_1,y_2,y_3,y_4}(w_i) [J_{y_1,y_2}(w_i) + P_{y_3,y_4}(w_i)]}{2}},
 \end{cases} \quad (1)$$

式中, y_1 、 y_2 分别为学术文献计算区间的起始年份和终止年份,当 $y_1=y_2$ 时,表示学术文献的选取年份区间为某一年; y_3 、 y_4 分别为基金项目计算区间的起始年份和终止年份,当 $y_3=y_4$ 时,表示基金项目的选取年份区间为某一年。GPx是编号为x的高频词分组, $x=01\sim 13$; w_i 为GPx分组中第i个高频词汇, w是GPx分组中学术文献和基金项目的高频词汇的并集。 $J_{y_1,y_2}(w_i)$ 表示 $y_1\sim y_2$ 年间学术文献中高频词汇 w_i 在其所在分组的频次比例; $P_{y_3,y_4}(w_i)$ 表示 $y_3\sim y_4$ 年间基金项目中高频词汇 w_i 在其所在分组的频次比例; $C(GPx)$ 为GPx分组的高频词数量。

$SV(w_i)_{y_1,y_2,y_3,y_4}$ 表示高频词汇 w_i 于 $y_1\sim y_2$ 年在学术文献与 $y_3\sim y_4$ 年在基金项目在所在分组比例的差异,该值越小,代表 w_i 在学术文献和基金项目中的词频比例越接近。

$GSV_{y_1,y_2,y_3,y_4}(GPx)$ 表示GPx分组的学术文献与基金项目的契合度。 $GSV_{y_1,y_2,y_3,y_4}(GPx)$ 体现了计算年份区间内GPx分组在学术文献和基金项目中的整体相似情况。由式(1)可以看出,对于某分组领域GPx,如果学术文献和基金项目的所有高频词汇比例都相同,则 $GSV_{y_1,y_2,y_3,y_4}(GPx) = 1$,为完全契合;如果学术文献和基金项目没有任何一个共同的高频词汇时,则 $GSV_{y_1,y_2,y_3,y_4}(GPx) = 0$,为完全不契合。

2 结果与分析

2.1 基金项目与学术文献的整体契合度

由于每年资助的基金项目数量不多,特别是重点重大项目数量更少,因此单年度资助项目对研究领域的代表性存在一定的局限性,为此将“十三五”期间的基金项目视为一个集合,考察5年资助项目与某一年份或某一时段学术文献之间的契合度,即在式(1)中, $y_3=2016$, $y_4=2020$ 。同时,这种较大的时间尺度(5年)可大幅度降低由于基金项目和学术文献之间可能存在的时间差而造成的分析误差。

表1为所有研究领域基金项目与每年学术文献的整体契合度。总体上,基金项目与每年学术文献的契合度在50%上下波动。两者契合度在2011年(47.5%)之后逐年上升,在2018年达到最高值51.6%,随后下降至2020年的47.6%。各类型基金项目与学术文献的契合度时间变化规律也大致如此。在契合度计算中,基金项目数据为2016~2020年,因此上述的契合度变化趋势显示,基金项目与学术文献在研究热点变化上保持着较为一致的步调,使得契合度峰值出现在2018年。

在4个项目类型中,重点重大项目与学术文献的契合度最高,比面上和青年项目高2~3个百分点,甚至在2017年提前

表1 “十三五”期间管理科学与工程(G01)学科资助项目与各年份学术文献的整体契合度^{a)}

Table 1 Overall fitting rates between the projects funded by the discipline of management science and engineering (G01) during the 13th Five-Year period and annual academic literature

基金项目类型	学术文献年份									
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
面上	46.4%	46.7%	47.8%	47.9%	48.5%	48.3%	49.1%	50.5%	49.0%	46.3%
青年	46.6%	46.8%	47.9%	48.0%	48.6%	48.5%	49.2%	50.6%	49.1%	46.4%
地区	35.3%	34.4%	35.6%	36.3%	36.6%	36.5%	36.4%	38.3%	37.0%	34.9%
重点重大	48.3%	49.8%	50.4%	50.3%	50.6%	50.6%	52.0%	52.3%	51.8%	49.1%
所有项目	47.5%	48.0%	49.0%	49.1%	49.6%	49.5%	50.4%	51.6%	50.3%	47.6%

a) 项目的资助年份为2016~2020年

达到了接近峰值的水平,体现了在立项阶段重点重大项目整体上不但比面上和青年项目更关注国际前沿问题,而且在科学问题凝练上更具有前瞻性。面上和青年项目与学术文献的契合度非常接近,说明这两类项目的申请人群体在设计研究内容时对国际前沿科学问题的思考方式和维度非常相似。地区项目与学术文献的契合度最低,在34.4%~38.3%之间,比面上和青年项目低11~13个百分点。

2.2 各高频词汇组基金项目与学术文献的契合度

图2为12个高频词汇组基金项目与学术文献的契合度。对于所有高频词汇组,两者契合度大致上遵循2011年开始增长、2018年左右达到峰值后开始下降的总体趋势,但是各组在契合度绝对值与波动幅度上呈现较大差异。这些差异可以

体现中国管理科学与工程各研究领域的发展特点以及在国际学术前沿的活跃程度,以下分类进行分析。

(1) 高契合度群。GP09风险管理、GP10供应链管理、GP12交通管理这3组的基金项目与学术文献的契合度最高,在60%~80%之间,表明这些领域基金项目关注的科学问题对国际学术研究热点有较高的覆盖率。特别是GP10供应链管理与GP12交通管理,基金项目与学术文献的契合度呈现明显的增长趋势,10年内上升了14个百分点,是12个领域中契合度上升幅度最大的两个领域。这表明,基金项目与这两个领域的学术研究成果具有很强的相互影响和相互带动作用,也预示着中国的供应链管理和交通管理研究将在国际学术舞台上居于越来越重要的引领地位。

(2) 中契合度群。GP02决策与博弈、GP04金融工程、

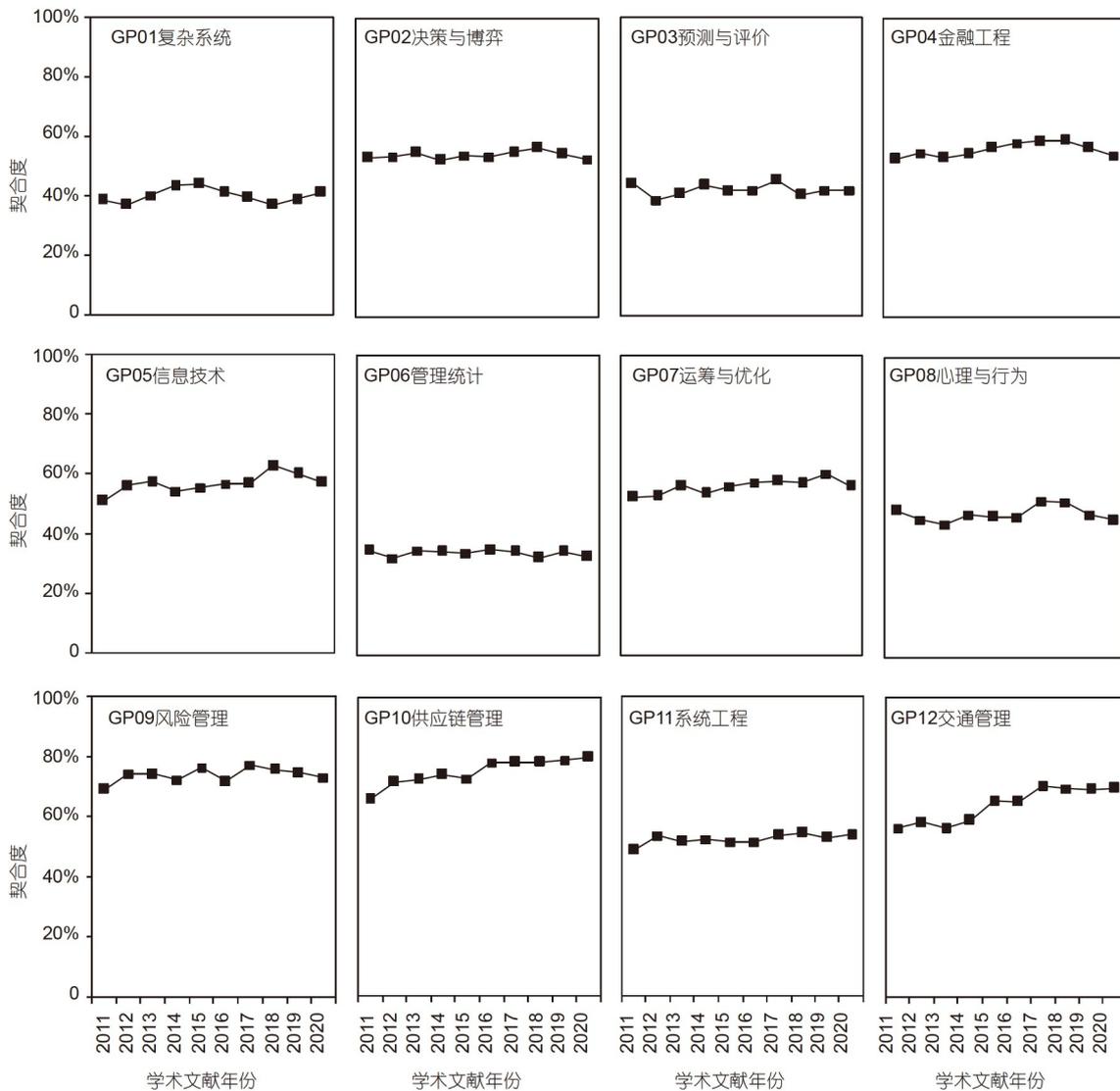


图2 “十三五”期间G01学科资助项目与每年学术文献的分组契合度

Figure 2 Fitting rates between 13th Five-Year G01 projects and annual academic literature by high-frequency word group

GP05信息技术、GP07运筹与优化、GP11系统工程这5组的基金项目与学术文献的契合度约为50%~60%。它们的契合度随时间有所上升,10年内上升幅度为4~12个百分点。

(3) 低契合度群。GP01复杂系统、GP03预测与评价、GP06管理统计、GP08心理与行为这4组的基金项目与学术文献的契合度在30%~50%之间。与高、中契合度群相比,低契合度群不但与学术文献的契合度值偏低,而且契合度的同期增长幅度也相对较小,说明这些领域的国内学者在其国际前沿的活跃程度弱于高、中契合度群。

综合上述分析,可以发现,偏重工程和应用的领域其基金项目与学术文献契合度较高,偏重理论和方法的领域其基金项目与学术文献契合度越低。这是中国管理科学与工程学科在“十三五”期间的发展特征之一,其形成原因可能有两个方面:(1) 国家重大需求牵引。“十三五”期间,中国在平台运营、信息通讯、移动支付、城际高速铁路等领域大跨步发展,为中国学者提供了领先世界的研究情景和研究机会,而挖掘并解决其发展过程中的管理科学问题更是国家重大需求。这些因素都牵引着学者在供应链管理、交通管理等领域开展深入研究。(2) 学科规划指引。“十三五”期间,管理科学与工程学科在战略布局上更加强化了应用与实践对于管理科学的意义^[19],围绕着高端装备制造、金融工程、城市群交通、平台供应链等研究方向资助了若干重大项目,围绕着互联网金融、运营管理、医疗服务、航空/港口运营等研究方向资助了一系列重点项目。这对中国学者在选择研究方向和研究问题上产生较大的指引作用。

然而应该重视的是,低契合度群中4个领域均为管理科学与工程学科的基础理论领域。它们较低的契合度说明,“十三五”期间基金项目中基础理论型研究在追踪或引领国际学术前沿方面比应用型研究相对滞后。我们在最近一次申请代

码调整工作中统计了2010~2019年中国学者(不含香港、澳门、台湾)在46本国际代表性期刊(表S1)的发文比例。结果表明,在基础理论领域,中国学者的贡献较低,例如,预测理论与方法为12%、博弈理论与方法为13%、管理心理与行为为13%、管理统计理论与方法为15%;但工程应用领域,中国学者的贡献较高,例如工业工程与管理为24%、交通运输管理为25%、工程管理为27%、系统可靠性与管理为52%^[20]。该结论与本文的契合度分析结果反映出中国学者在基础理论领域的投入和产出低于工程应用领域。

2.3 两大类基金项目与不同时期学术文献的契合度变化

如前所述,重点和重大项目选题大多来自学科规划,面青地项目则源于广大学者的自由思考。对比这两大类基金项目与学术文献的契合度变化可以发现学科规划和学者个体思考之间的共性与不同。将10年的学术文献分为两个时间段:“十二五”期间(2011~2015年)和“十三五”期间(2016~2020年),即式(1)中, $y_1=2011$, $y_2=2015$ 和 $y_1=2016$, $y_2=2020$ 。

图3展示了“十三五”期间面青地项目和重点重大项目与两个时间段学术文献的契合度变化。面青地项目与重点重大项目表现出不同的契合度变化特征。

(1) 对于大多数领域,基金项目与“十三五”期间学术文献的契合度要高于它与“十二五”期间学术文献的契合度。在GP12交通管理、GP05信息技术、GP07运筹与优化等领域,面青地项目的契合度增长幅度要大于重点重大项目。而对GP10供应链管理、GP02决策与博弈等领域,重点重大项目的契合度增长幅度更大。一方面,学科规划对热点领域研究起到助推作用;另一方面,广大学者基于自身对国际前沿问题与中国管理需求的观察和思考亦形成了强大的推动力。例

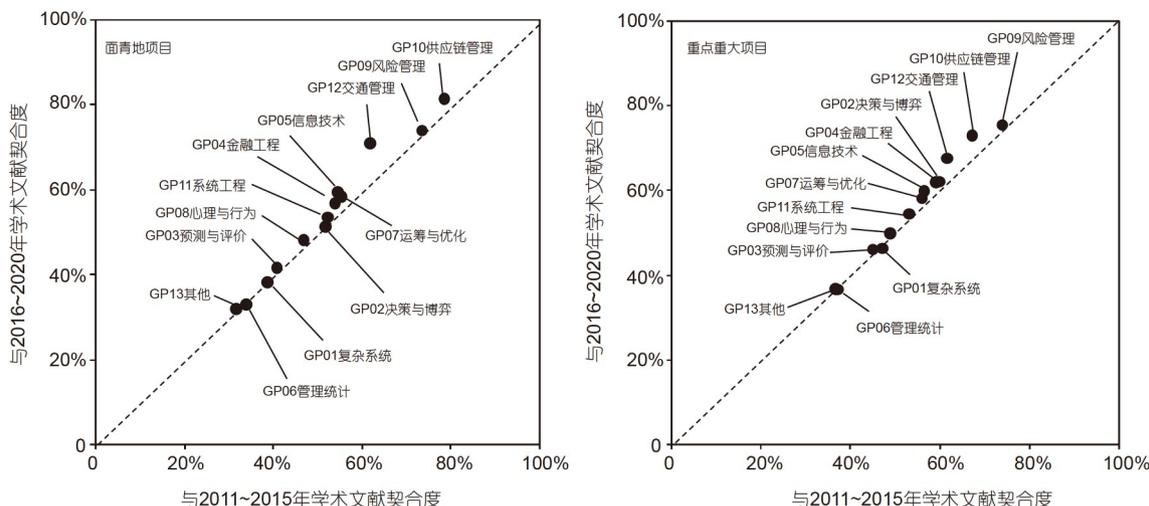


图3 “十三五”期间G01学科资助项目与“十二五”和“十三五”期间学术文献的契合度
Figure 3 Fitting rates between 13th Five-Year G01 projects and 12th Five-Year academic literature vs. 13th Five-Year academic literature

如,虽然管理科学与工程学科在交通管理领域资助了许多重点重大项目,但是该领域自由探索的面青地项目与国际学术热点契合度的增幅更为显著(面青地项目为9.5个百分点,重点重大项目为6.0个百分点)。

(2) 对于一些领域,两大类项目与两个期间学术文献的契合度无明显变化,例如GP06管理统计、GP08心理与行为、GP01复杂系统等这些基础理论领域,而且这些领域基金项目与学术文献的契合度也一直维持在较低水平。这些领域在学科规划中有所涉及,但强度不大,从图3可知,广大学者对它们亦未形成足够的研究热情。之前论及,工程和应用领域的基金项目与学术文献契合度较高,理论和方法的领域基金项目与学术文献契合度较低。从两大类项目契合度分析结果来看,这是学科规划与学者个体思考两者共同作用的结果。

2.4 各高频词汇组、各类型基金项目与学术文献的契合度对比

为了进一步聚焦各类基金项目与学术文献在不同研究领域的契合度差异,图4具体给出了各高频词汇组“十三五”期间重点重大、面上、青年和地区四类项目与“十三五”期间学术文献的契合度,即式(1)中, $y_1=2016$, $y_2=2020$ 。

总体上,12个领域的契合度结果显示两个共同的规律,即面上项目契合度等于青年项目契合度,以及地区项目在四类项目中契合度最低,与表1展示的全领域契合度特征相同。但是,重点重大项目与面青地项目的契合度对比结果存在明显的领域差异。基于这种差异将12个领域分成三类:(1)重点重大项目契合度>面上青年项目契合度,这类包括GP01复杂系统、GP02决策与博弈、GP03预测与评价、GP04金融工

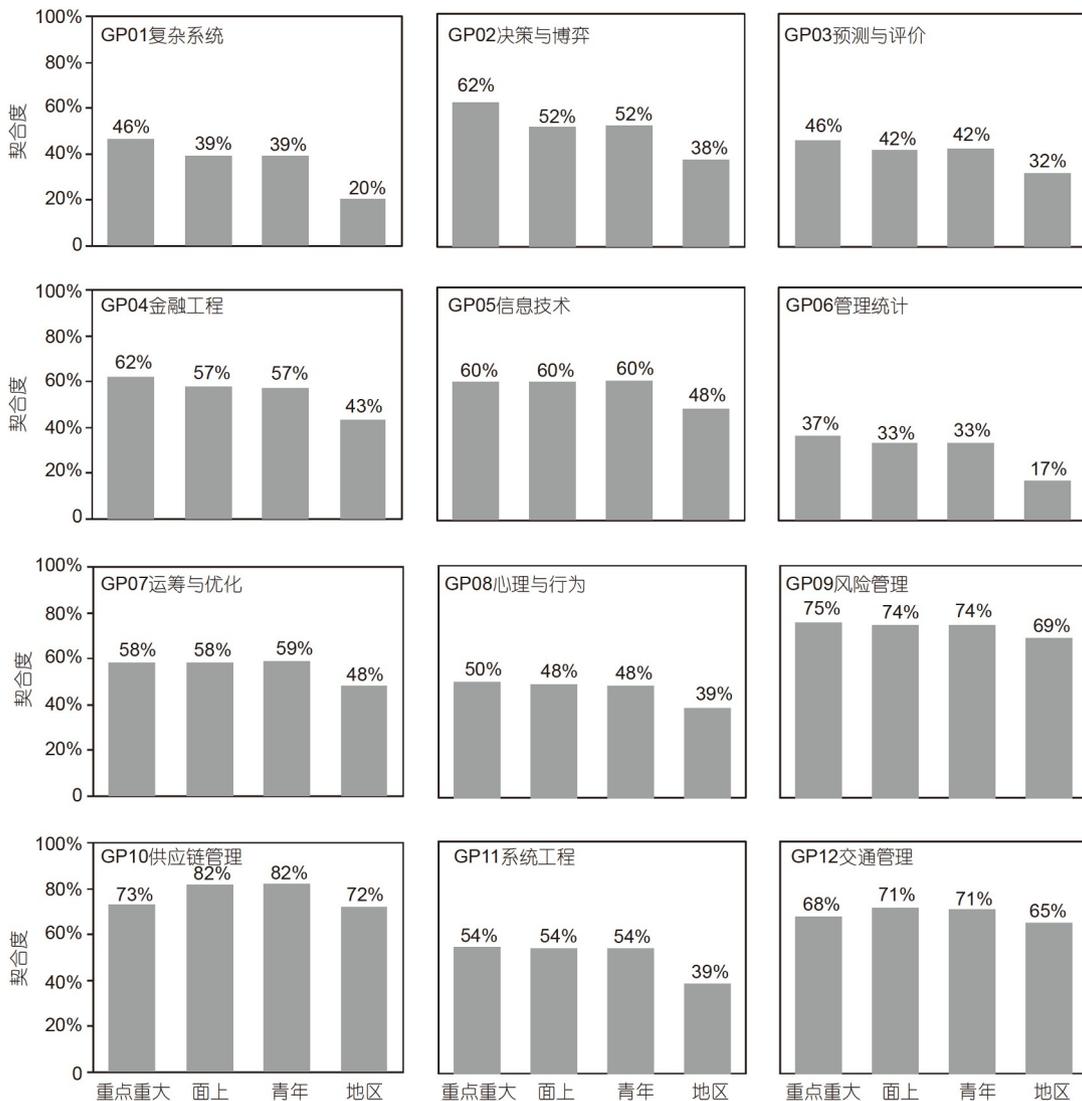


图4 “十三五”期间G01学科各类型资助项目与同时期学术文献的契合度

Figure 4 Fitting rates between 13th Five-Year G01 projects and 13th Five-Year academic literature by project type

程、GP06管理统计、GP08心理与行为; (2) 重点重大项目契合度 \approx 面上青年项目契合度(两者之差 $\leq 1\%$), 这类包括GP05信息技术、GP07运筹与优化、GP09风险管理、GP11系统工程; (3) 重点重大项目 $<$ 面上青年项目契合度, 这类包括GP10供应链管理、GP12交通管理。

这种差异反映了不同研究领域内学科规划与学者个体思考的相对力量, 对于后两类领域, 特别是第三类领域, 广大学者的自由探索显示出更为强势的科研力量, 实现了高于学科规划的国际学术影响力。面上和青年项目的负责人多数为青年学者(2016~2020年G01面上和青年项目负责人平均年龄为36.8岁, 40岁及以下负责人占到77%)。因此, 可以预见, 从这些领域申报国家杰出青年科学基金(以下简称“杰青”)和优秀青年科学基金(以下简称“优青”)的青年学者将会更具有竞争优势。

“十三五”期间, 管理科学与工程学科共资助了13个杰青项目和33个优青项目。将这46个项目的领域对应到12个高频词汇组里, 获取了它们的领域分布(图5)。46个杰青、优青项目中, GP12交通管理为12项, 占26%, GP10供应链管理为10项, 占22%, 两个领域的杰青、优青项目数量几乎占到学科总数的一半, 显示出强势的竞争力。这两个领域的基金项目不但与学术文献契合度最高, 更为重要的是, 其面上青年项目与学术文献的契合度高于重点重大项目与学术文献的契合度, 表现出强大的青年人才储备。

其他杰青、优青项目占比较高的领域还包括GP07运筹与优化(17%)、GP05信息技术(13%)等。这些领域基金项目与学术文献契合度的特点同样包括两个方面: 基金项目与学术文献契合度较高, 表明该领域处在国际较前沿的位置; 面上青年项目与学术文献契合度不低于重点重大项目与学术文献的契合度, 表明具备相当的青年人才储备。

总体上, 各领域资助的杰青、优青项目数量大致上和各领域基金项目与学术文献契合度高低分布一致。基本上, 契合度越高的领域获批的杰青、优青数量越多, 而对于面上青年项目与学术文献契合度高于重点重大项目与学术文献契合度的领域, 其青年人才的蓄势与增长潜力更加明显。由此可见, 契合度分析还可以预见青年人才快速成长的研究领域。

3 启示与思考

本文发现, 学科规划对学科发展产生了多方面的巨大影响。基金项目与学术文献契合度高的领域分布特征基本与“十三五”期间管理科学与工程学科的重点重大项目资助格局相吻合。不仅如此, 重点重大项目与学术文献契合度较高的领域, 其面上青年项目与学术文献的契合度也普遍较高, 说明学科规划对学者个体思考产生了很强的引领作用。这种作用是自上而下、由点及面的, 其影响深刻、广泛、而且长远, 它不但决定了学科当前的优势领域分布, 还影响着未来青年人才的领域布局。

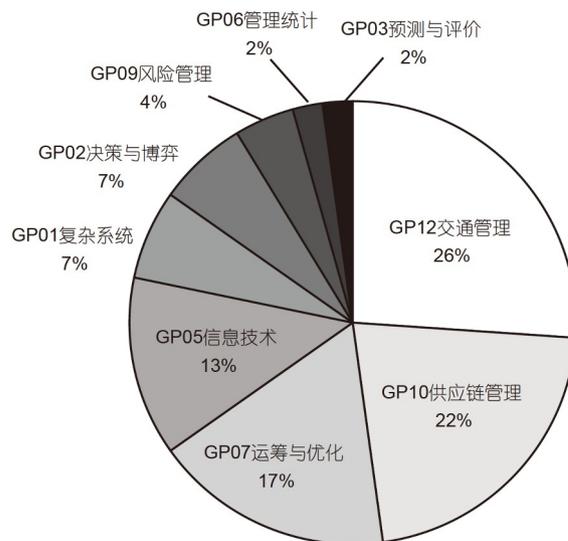


图5 “十三五”期间G01学科资助的46项杰青和优青项目领域分布
Figure 5 Academic field distribution of 46 projects of National Science Fund for Distinguished Young Scholars and Excellent Young Scientists funded by the G01 discipline during the 13th Five-Year period

学科规划中布局的一系列重点和重大项目对帮助青年学者在这些领域里扎根并明确自己的研究方向起到了重要的引领作用。重点和重大项目需要较大的研究队伍来支撑, 通常以一个或几个学术带头人带领若干青年学术骨干的团队模式来推进研究工作。有些青年学术骨干在团队协作的锻炼中逐渐形成了自己的学术思考, 并开始申请青年项目和面上项目, 挑战比重点和重大项目更为前沿的科学问题, 逐步成长为具有独立学术思想的优秀青年学者。这些青年学者在同龄人中广泛寻求合作, 一定程度上带动了领域内青年学者群体的学术热情, 提升了青年学者整体的学术创造力。从本文的契合度分析结果来看, 学科规划对人才成长的推动作用积极而明显的。

另一方面, 本文的结果显示, “十三五”期间管理科学与工程学科工程应用研究领域发展迅猛, 基础理论研究发展相对滞后。基础理论研究是一个学科立足的根本, 是体现一个学科原始创新能力与核心竞争力的重要学术阵地, 同时国家也对科技领域提出了加强基础科学研究、提高原始创新能力的要求。面对这一形势, 我们必须深入思考, 如何在“十四五”甚至更长远的时期利用学科规划继续扩大工程应用领域的优势, 同时激活基础理论领域。

由于学科发展规律具有相通性, 本文的研究方法和结论对基金委其他学科亦有借鉴意义。学科规划对学科发展的影响效果在不同学科将会表现出相似的规律性, 基金委一些学科对学科规划的效果给出了定性分析和结论, 例如, 数理科学部力学学科分析了“十二五”学科规划与学科发展的关系, 认为重点项目、重大项目和重大研究计划对推动关键领域发展具有重大战略意义^[21]; 化学科学部总结了“十三五”规划

的实施效果,充分肯定了学科规划对促进学科发展的重要引领作用^[22]。这些结论与本文的结论具有高度的一致性。而且,由于本文引入了文献计量方法,可用于定量评价学科规划中优先发展领域在国际前沿中的发展态势,将有助于更为细致、准确地把握学科规划对学科发展以及人才成长的影响

路径,并及时发现问题。

因此,我们必须充分认识并利用学科规划的这种巨大影响力,有计划地作出具有引导性的战略规划和领域布局,推动面向国家重大需求的研究领域快速发展,并瞄准关键基础领域提前布局、储备人才。

推荐阅读文献

- 1 Li A P. Milestone in China's science and technology development history: Twelve year long-term plan for science and technology development (in Chinese). *Sci News*, 1999, (28): 30 [李安平. 新中国科学技术发展史上的里程碑——十二年科学技术发展远景规划. *科学新闻周刊*, 1999, (28): 30]
- 2 Yang W L, Zhang M. Formulation, implementation and historical experience of the first science and technology development plan in China (in Chinese). *CPC Hist Stud*, 2007, (6): 42–49 [杨文利, 张蒙. 新中国第一个科技发展规划的制定、实施及历史经验. *中共党史研究*, 2007, (6): 42–49]
- 3 Zhang L H, Xu X X. Theory basis and methods of S&T plan (in Chinese). *Stud Dialect Nat*, 2005, 21: 68–73 [张利华, 徐晓新. 科技发展规划的理论与方法初探. *自然辩证法研究*, 2005, 21: 68–73]
- 4 Hu W J. The 12-year long-term science and technology development plan: Planning, effect and some enlightenments (in Chinese). *Policy Manage Res*, 2006, 21: 207–212 [胡维佳. “十二年科技规划”的制定、作用及其启示. *政策与管理研究*, 2006, 21: 207–212]
- 5 Cui Z F. Thoughts on major issues in compiling the 10th Five-Year Plan for science and technology development (in Chinese). *Sci Sci Manage S T*, 2000, 21: 29–30 [崔振凤. 关于编制“十五”科技发展规划若干重大问题的思考. *科学学与科学技术管理*, 2000, 21: 29–30]
- 6 Sun X L, Ding K. Research topics and evolution of management science: Based on the analysis of open repository of NSFC (in Chinese). *Sci Sci Manage S T*, 2017, 38: 3–11 [孙晓玲, 丁堃. 管理科学研究主题及其演化趋势——基于NSFC基础研究知识库的分析. *科学学与科学技术管理*, 2017, 38: 3–11]
- 7 Wang X F, Zhang S, Liu Y Q, et al. Forty years of research on science and technology evaluation in China: Historical and theme evolution (in Chinese). *Sci Sci Manage S T*, 2018, 39: 67–80 [汪雪锋, 张硕, 刘玉琴, 等. 中国科技评价研究40年: 历史演进及主题演化. *科学学与科学技术管理*, 2018, 39: 67–80]
- 8 He J X, Ge J Y, Zhang F. The world contribution of Chinese scholars to management research from the perspective of international collaborations, frontier hotspots and contribution approaches: A quantitative analysis based on a thousand of international English management journals (2013–2019) (in Chinese). *Manage World*, 2021, (9): 36–56 [何佳讯, 葛佳焯, 张凡. 中国学者管理学研究的世界贡献: 国际合作、前沿热点与贡献路径——基于世界千钟管理学期刊论文(2013~2019年)的定量分析. *管理世界*, 2021, (9): 36–56]
- 9 Deerwester S, Dumais S T, Furnas G W, et al. Indexing by latent semantic analysis. *J Am Soc Inf Sci*, 1990, 41: 391–407
- 10 Hofmann T. Probabilistic latent semantic indexing. In: *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Berkeley: Association for Computing Machinery, 1999. 50–57
- 11 Blei D M, Ng A Y, Jordan M I. Latent Dirichlet a location. *J Mach Learn Res*, 2003, 3: 993–1022
- 12 Zhang Y, Wei W. A jointly distributed semi-supervised topic model. *Neurocomputing*, 2014, 134: 38–45
- 13 Blei D M, Jordan M I. Variational inference for Dirichlet process mixtures. *J Bayesian Anal*, 2006, 1: 121–143
- 14 Vretos N, Nikolaidis N, Pitas I. Video fingerprinting using Latent Dirichlet Allocation and facial images. *Pattern Recognit*, 2012, 45: 2489–2498
- 15 Zhang H, Zhang X, Tian Z, et al. Incorporating temporal dynamics into LDA for one-class collaborative filtering. *Knowledge-Based Syst*, 2018, 150: 49–56
- 16 Tamaki M. The secretary problem with optimal assignment. *Oper Res*, 1984, 32: 847–858
- 17 Buchbinder N, Jain K, Singh M. Secretary problems via linear programming. *Math Oper Res*, 2014, 39: 190–206
- 18 Ramachandran D, Parvathi R. Analysis of twitter specific preprocessing technique for tweets. *Proc Comput Sci*, 2019, 165: 245–251
- 19 Huang H J, Liu Z Y, Yao Z, et al. Research Report for the 13th Five-Year Development Strategy and Priority Funding Fields of the Discipline of Management Science and Engineering (in Chinese). Beijing: Science Press, 2016. 31–61 [黄海军, 刘作仪, 姚忠, 等. 管理科学与工程学科“十三五”发展战略与优先资助领域研究报告. 北京: 科学出版社, 2016. 31–61]
- 20 Huo H, Yu Y G, Zheng S M, et al. Application codes and discipline layout of NSFC's management science and engineering: Towards fundamental theory and times transformation (in Chinese). *Manage World*, 2021, (12): 208–219 [霍红, 余玉刚, 郑圣明, 等. 国家自然科学基金管理科学与工程学科的申请代码与学科布局: 面向基础理论与时代变革. *管理世界*, 2021, (12): 208–219]
- 21 Zhan S G, Zhang P F. Review of NSFC projects on mechanics and the 13th Five-Year Development Strategy (in Chinese). *Chin J Theoret Appl Mech*, 2017, 49: 478–483 [詹世革, 张攀峰. 国家自然科学基金力学学科发展现状和“十三五”发展规划. *力学学报*, 2017, 49: 478–483]

- 22 Zhang G J, Xu X F, Zheng Q Y, et al. Chemical sciences transformation in China—Review of the 13th Five-Year Plan of Department of Chemical Sciences, NSFC (in Chinese). *Sci Sin-Chim*, 2020, 50: 681–686 [张国俊, 付雪峰, 郑企雨, 等. 转型中的中国化学——基金委化学部“十三五”规划实施纪行. *中国科学: 化学*, 2020, 50: 681–686]
-

补充材料

表S1 管理科学与工程领域的46本前沿性和代表性学术期刊

表S2 2016–2020年基金委管理科学与工程学科资助的各类项目数

表S3 LDA模型生成的主题及主题词数量统计

表S4 高频词汇分组的高频词汇数量和频次数量

本文以上补充材料见网络版 csb.scichina.com. 补充材料为作者提供的原始数据, 作者对其学术质量和内容负责.

Summary for “国家自然科学基金学科规划对学科发展的影响: 基于文献计量的分析与思考”

Impacts of the discipline plan of National Natural Science Foundation of China on discipline development: Analysis and implications based on the bibliometrics-method

Sihua Chen^{1,2}, Han Qiu¹ & Hong Huo^{2*}

¹ School of Information Management, Jiangxi University of Finance and Economics, Nanchang 330013, China;

² Department of Management Sciences, National Natural Science Foundation of China, Beijing 100085, China

* Corresponding author, E-mail: huohong@nsfc.gov.cn

The science departments of National Natural Science Foundation of China (NSFC) carry out strategic research on discipline development every five years, and formulate discipline plans and several priority funding areas for the next five years. Then they successively release guidelines for Key Program and Major Program (K&M) projects to promote the development of priority funding areas.

In order to understand the impacts of the discipline plan on discipline development, this study takes the discipline of management science and engineering during the 13th Five-Year Plan period as an example, and uses the approved K&M projects to represent the discipline plan, the General Programs, Young Scientists Fund, and Fund for Less Developed Regions (GY&L) projects to represent the individual thinking of general scholars, and the academic literature from representative international journals to represent the international frontiers and hot spots. This study uses the Latent Dirichlet Allocation (LDA) method, as well as high-frequency word analysis, and introduce a concept of fitting rate to analyze the relationship between NSFC projects during 2016 and 2020 and academic literatures during 2011 and 2020.

The results show that NSFC projects have higher fitting rates with academic literature in the academic areas with more K&M projects funded than in those with fewer K&M projects funded, and the academic areas with higher fitting rates between K&M projects and academic literature usually have higher fitting rates between GY&L projects and academic literature. Furthermore, academic areas with higher fitting rates between NSFC projects and academic literature tend to have more projects of National Science Fund for Distinguished Young Scholars and Excellent Young Scientists funded.

Therefore, NSFC's discipline plan has a profound, extensive and long-term impact on the discipline development. The impact goes from top to bottom, from points spreading to areas. It does not only determine the current distribution of priority areas, but also affects the potential distribution of young scientists in the future.

Because the development paths of sciences have similarities, the conclusions of this study could also be useful to other disciplines. We should fully understand and make use of these advantages of the discipline plan to promote the development of academic areas to meet national requirements, and aim at key fundamental fields to accumulate young talents in advance.

National Natural Science Foundation of China, discipline plan, text analysis, latent Dirichlet allocation, fitting rate, research hot spots

doi: [10.1360/TB-2021-0969](https://doi.org/10.1360/TB-2021-0969)