

我国大陆认知诊断研究的文献计量分析

戴海琦 谢美华 丁树良*

【摘要】 为了探索我国大陆认知诊断研究的现状及发展趋势,对收录在我国学术期刊网络出版总库(CAJD)中的94篇认知诊断研究文献(1980~2011)进行计量学分析,结果表明:(1)大陆认知诊断研究从2004年起呈现快速增长趋势但文献总量低于国外;(2)基础研究是主要的研究类型,诊断方法研究集中在近两年且主要采用模拟实验方法进行验证,应用研究较少报告诊断的信度和效度;(3)研究主要围绕Q矩阵理论、认知诊断模型改进及开发和结合认知诊断的计算机化自适应测验;(4)合作研究是认知诊断研究的主要形式,已形成核心作者群;(5)文献被引率为77.66%,10篇经典文献占被引总次数的58.89%;(6)文献主要分布于心理学、教育学和计算机科学;(7)大部分认知诊断研究受到各级项目资助;(8)心理科学、心理学报和中国考试是主要刊文期刊。

【关键词】 认知诊断研究;文献计量分析;中国大陆

一、引言

测验理论已经历了从标准测验理论阶段进入新一代测验理论阶段的发展。经典测验理论、概化理论和项目反应理论都属于标准测验理论。标准测验理论把所测心理品质当作一种纯统计结构,依据被试对题目的反应估计其能力大小。标准测验理论缺乏对能力所代表结构的详细描述,计量时只注重作答结果,忽视对被试作答过程的考察,具有较大的局限性。Mislevy(1993)将标准测验理论这种仅强调对被试宏观能力水平评估的范式称为能力水平研究范式。而新一代测验理论则强调心理测量学应与认知心理学相结合,强调测验应同时在能力水平和认知水平两个水平研究范式上进行。为此,能同时考察被试宏观能力和微观认知属性的认知诊断被视为新一代测验理论核心。认知诊断是认知心理学和现代测量学相结合的产物。广义的认知诊断指建立起观察分数和被试的内部认知特征之间的关系,其应用包括在心理学理论建构中的应用和在教育教学领域中的应用两个方面。狭义的认知诊断仅包括在教育教学领域中的应用,目的是按被试有没有掌握测验所测的技能或特质来

* 戴海琦,江西师范大学心理学院教授、博士生导师;谢美华,江西师范大学心理学院博士研究生;丁树良,江西师范大学计算机工程学院教授,博士生导师,330022。本研究得到国家自然科学基金(30860084,31160203,31100756)、江西省研究生创新基金(YC2011-B015)和江西省高校人文社会科学研究青年基金(XL1208)资助。

对被试加以分类。目前大陆学者更多关注的是狭义认知诊断,也即教育认知诊断。开展认知诊断评估需进行一系列的研究性工作,包括认知模型的界定、诊断测验的编制、诊断模型的参数估计、诊断结果的报告及有针对性的补救措施的制作,其中诊断测验的编制又包括确定诊断目标,根据诊断目标确定认知属性及属性间的阶层关系并加以验证,根据已确定的认知模型编制认知诊断测验并进行试测再进行正式测试。这一评估过程相对于标准测验理论阶段的测验而言是一项高难度、高复杂的专业性工作,主要体现在认知模型难以准确确定,认知诊断测验编制过程复杂,认知诊断模型参数估计较难实现,诊断评估过程需要多科专家参与等方面,因此,目前我国认知诊断更多地停留在研究阶段,而在实际教育考试中并未大规模开展起来。但认知诊断有助于人们更好地了解人类内部心理活动规律及其加工机制,探查个体的知识结构、认知策略及属性掌握状况,从而实现对个体认知强项和弱项的诊断,这将有助于学生本人、教师及家长对学生知识掌握情况的全面了解,有助于学生开展有效的补救学习和教师进行有针对性地补救教学,实现“促进学生发展”的真正目的,正因此,认知诊断成为当前国际心理测量学研究的一个重要热点。

国外认知诊断研究早于大陆^①,20世纪六七十年代,研究者已开始关注认知心理学与心理测量学的结合(Gentile, Kessler, & Gentil, 1969; Fischer, 1973; Macready, & Dayton, 1977),1980年代有了更进一步的基础研究并开始发表一些实际应用的研究(Tatsuoka, 1983; Ronning, Conoley, Glover, & Witt, 1987; Haertel, 1989),1990年代成果渐为丰富(Carpenter, Just, & Shell, 1990; Costa, & McCrae, 1992; Embretson, 1998; Buck, & Tatsuoka, 1998)且已出版认知诊断评估专集(Frederiksen, Glaser, Lesgold, & Shafto, 1990; Nichols, Chipman, & Brennan, 1995),21世纪以后至今进入了一个快速发展的

飞越时期,越来越多的研究者参与其中,有些研究者(Tatsuoka; de la Torre; Gierl; Henson; Leighton; Templin)发表了多篇论文,有大量学术论文发表,已开发出60多种认知诊断模型(Fu, 2005),出版多部专著(Leighton, & Gierl, 2007; Tatsuoka, 2009; Rupp, Templin, & Henson, 2010),曾召开过至少两次以认知诊断为主题的国际性大型会议。在教育实践中,美国、英国等发达国家非常重视认知诊断对教育改革做出的贡献,美国2001年正式通过法案《不让一个孩子掉队》,要求教育要有针对性地对每个孩子的学习过程作出评价。在我国台湾,有研究团队已开发出结合认知诊断的计算机化自适应测验并在教育实践中开展实验性应用,我国香港2003年已有博士开展结合认知诊断的计算机化自适应测验研究(文剑冰, 2003)。为了反映大陆认知诊断研究的进展状况,本文采取文献计量学方法对1980~2011年间大陆认知诊断研究成果进行分析和探讨,以期为进一步开展认知诊断研究提供参考。

二、研究方法

(一) 文献检索

本文以我国学术期刊网络出版总库(简称CAJD)为数据源进行分析。CAJD是世界上最大的连续动态更新的我国学术期刊全文数据库,收录自1915年至今出版的期刊,部分期刊回溯至创刊,核心期刊收录率达96%。本文文献获取采用专业检索方法,检索式为:SU = ‘认知诊断’ + ‘认知诊断评估’ + ‘规则空间模型’ + ‘DINA模型’ + ‘属性层级模型模型’,SU表示主题,“+”表示或者,检索年限为1980~2011年,采用模糊查询,检索结果为96篇,有效文献94篇,含1篇全英文文献(Yang, Ding, & Ding, 2010)。

^①本文仅分析CAJD收录的大陆认知诊断研究文献,不包含大陆作者发表在国外的文献,如辛涛(2005)、杨向东(2007)、杨淑群(2009)、陈平(2012)均在国外发表过认知诊断研究学术论文。

(二) 分析单元与方法

采用文献计量法,以频次、百分比对年载文量、研究内容及研究类型、关键词、被引频次、作者分布及合作度、研究学科、基金资助、发表源刊共 8 个分析单元挖掘我国大陆认知诊断研究的发展趋势、研究内容的规律、研究质量。

三、结果与分析

(一) 年载文献总量统计

1980~2011 年间,我国大陆、国外^①分别发表认知诊断研究文献 94 篇和 445 篇,大陆发文总量低于国外,发表年代分布见表 1。大陆的认知诊断文献始于 1984 年,起步晚于国外。从整体发展趋势看,文献量增长的快速发展期也晚于国外,国外在 2002 年以后发展较快,大陆在 2006 年以后发展迅速(见图 1),且国外文献量

自始至终比大陆多。1980—2001 年间国外文献总量 41 篇,年均 1.86 篇,我国同期文献总量为 4 篇,分别是 1984 年 1 篇(方富熹,1984)、1995 年 2 篇(余嘉元,1995;金庆跃,1995)、1997 年 1 篇(王萍,路志英,余家红,林孔元,1997)。表 1 表明大陆 2003 年前没有持续开展认知诊断研究。2002—2011 年间国外文献总量为 404 篇,年均 40.4 篇,大陆同期文献量为 90 篇,年均 9 篇,文献总量依然低于国外。大陆在 2004 年后随着项目反应理论及跨学科研究的发展,认知诊断研究逐步进入快速发展阶段,2004—2011 年间,大陆认知诊断文献持续增长率为 76.46%,呈现快速增长趋势。这里,持续增长率 r 的计算公式为: $r = \sqrt[n]{A/a} - 1$ (A 为文献累积量, a 为初始时刻的文献数量, n 为统计年数)(万晓霞,2009)。

表 1 大陆及国外认知诊断文献量统计

年份	2011	2010	2009	2008	2007	2006	2005	2004	2003	2002	1980—2001	总计
大陆	27	21	15	14	6	5	1	1	0	0	4	94
国外	88	90	77	59	32	19	12	17	8	2	41	445

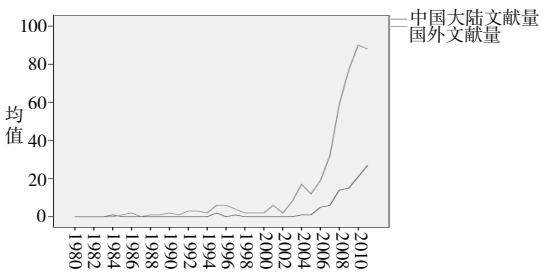


图 1 大陆、国外认知诊断研究文献量趋势

(二) 研究内容及研究类型

根据研究的目的不同,可把 94 篇认知诊断研究文献分为基础研究与应用研究两大类,其中基础研究 76 篇,占总数的 80.85%,应用研究 18 篇,占总数的 19.15%(表 2),可见我国大陆开展的认知诊断研究主要是基础研究。基础研

究包含理论介绍性文献和诊断方法研究。其中,介绍性文献有 36 篇,占总文献量的 38.30%,2006 年以前的文献主要是这一类,内容涵盖认知诊断理论的理论基础、认知诊断的基本流程、规则空间模型及属性层级模型述评、认知诊断模型的分类及模型进展以及教育领域对认知诊断的引介。诊断方法研究有 40 篇,占总文献数的 42.55%,这类研究集中于 2007 年以后,近两年来急剧增多,这类研究的内容涉及有:CD—CAT13 篇,其中选题策略 6 篇,在线属性标定 3 篇;模型提出、改进或比较 13 篇;诊断测验编制方法 6 篇;诊断分类方法 4 篇;等值 1 篇;其他 3 篇。可见,大陆认知诊断方法研究集中于 CD—CAT 及诊断模型研究。在 40 篇新方法或新模型的提出或改进研究中,采用实测数据进行验

^①国外文献来源于 Scirus 数据库,Scirus 是目前互联网上最全面、综合性最强的科技文献门户网站之一,文献检索方法及检索年限与大陆相同,共检索到相关文献 1180 篇,有效文献 445 篇。

证(钱锦昕,余嘉元.2010)或采用实测数据和模拟方法相结合的方法(陈德枝,戴海琦,赵顶位.2009;喻晓峰,丁树良等.2011)进行验证的研究较少,95%的研究采用蒙特卡罗模拟方法验证新方法的有效性。18 篇应用研究集中在 2008

年以后,诊断对象涉及有英语阅读问题解决、瑞文智力测验、小学儿童数学问题解决、汉语测试、图形或类比推理、初中化学、口腔内科、描述统计学习模式,但鲜有应用研究报告诊断的信度和效度。

表 2 大陆认知诊断研究类别分布

研究类别		1984	1995	1997	2004	2005	2006	2007	2008	2009	2010	2011	总	%
基础研究	理论介绍	1	1	0	0	1	5	3	7	7	5	6	36	38.30
	诊断方法	0		1	0	0	0	3	5	5	9	17	40	42.55
应用研究	方法应用		1		1				2	3	7	4	18	19.15
合计		1	2	1	1	1	5	6	14	15	21	27	94	100.0

(三) 关键词分析

关键词是为了文献标引工作从论文中提炼出来用以表示全文主题内容信息的单词或术语,对文献起着表达主题、反映本质和被检索的

作用。因此,对关键词进行统计分析,能够在一定程度上反映出我国认知诊断研究的方向和重点,见表 3。

表 3 关键词变化一览表

年份	关键词
1980—2005	规则空间模型(2)、项目反应理论、认知诊断、属性掌握模式
2006	认知诊断(3)、规则空间模型(2)、学业评价、心理测量理论、教育测量、知识结构、诊断性测验、新课程改革、属性掌握模式、信息技术、标准测验理论、新一代测验理论
2007	认知诊断(4)、CAT(3)、大规模考试、项目反应理论、项目特征曲线、等级反应模型、状态空间图、选题策略、认知诊断模型、计算机化自适应测验、IRT、状态转换图、概念格、多级评分
2008	认知诊断(7)、规则空间模型(4)、融合模型(2)、个人拟合、知识状态、认知诊断模型、分类法、简介、普适度、个人拟合指数、DINA 模型、认知属性、形式概念分析、测验实效、类比推理
2009	认知诊断(6)、规则空间模型(5)、认知诊断模型(2)、项目反应理论(2)、认知诊断理论(2)、属性层次模型(2)、认知心理学、规则空间方法、融合模型、属性掌握概率分类模型、认知诊断评价、DINA 模型、Q 矩阵、项目参数、选题策略、心理测量学
2010	认知诊断(12)、认知诊断模型(4)、规则空间模型(2)、知识空间理论(2)、等级反应模型、MC-MC 算法、神经网络、模式分类、可达阵、简化 Q 矩阵、认知属性、证据中心设计、项目反应理论、DINA 模型
2011	认知诊断(16)、选题策略(6)、计算机化自适应测验(5)、规则空间模型(4)、项目反应理论(3)、可达矩阵(3)、在线标定(2)、测量准确性(2)、新题(2)、Q 矩阵(2)、认知诊断计算机化自适应测验(2)、计算机化自适应诊断测验(2)、知识空间理论(2)、知识状态、题库使用均匀性、DINA 模型

注:关键词后括号中的数字是该关键词出现的次数,出现 1 次的关键词没有注明次数。

由表 3 进一步统计出高频关键词,结果为:认知诊断(48)、规则空间模型(19)、计算机化自适应测验或 CAT(13)、项目反应理论(8)、认知诊断模型(8)、选题策略(8)、DINA 模型(4)、知识结构(4)、Q 矩阵(4)、可达矩阵(4)、融合模型(3)、诊断(3)、认知属性(3)、知识状态(3)。由此可见,我国大陆认知诊断研究目前主要围绕认知诊断、认知诊断模型改进及开发、计算机化自适应认知诊断测验及在此基础上的选题策

略、认知模型构建的理论基础 Q 矩阵理论研究,这些研究是开展认知诊断的基础或关键技术,研究紧跟国际认知诊断理论研究的前沿,为中国引介了新一代测验理论并为进一步在教育测验领域开展应用研究奠定了坚实的基础。但从目前的研究成果来看,总体而言基础研究胜于实践研究;探讨过的认知诊断模型有限,主要有规则空间模型、属性层级模型和 DINA 模型;认知模型的构建方法尚没有新的突破;结合认知

诊断的计算化自适应测验尚没有服务于实践。不过,近年来认知诊断的高频出现及关键词种类日益增多均表明认知诊断研究日益受到重视且研究者们正在努力探索认知诊断的最新技术,表明认知诊断研究正在向深度和广度发展。

（四）论文被引分析

论文被引频次体现了研究成果的学术水平,也反映了作者的科研成果被同行关注的程度。因而论文被引次数越多说明其在该学科领域影响越大。认知诊断研究的 94 篇文献,共被引 647 次,每项平均被引次数 6.88,进一步统计发现,实际被引论文 73 篇,被引率为 77.66%,实际平均被引次数为 8.86。这 73 篇文献中 35.11% 的论文被引用 4 次以上(33 篇),分别有 9.57%、14.89% 和 18.09% 的论文被引用 3 次、2

次和 1 次,22.34% 的论文没有被引用过(表 5)。《科学评论索引》(ISR)认为,如果一篇论文的其他引次数在 50 以上,这篇论文就可以被认为是某一学科领域的经典之作(陈艳君,2012),如表 4 中前 3 篇,或一篇文献每年被引用 4 次或 4 次以上,则可列为“经典文献”(万晓霞,2009)。根据发表年限及被引次数,可以初步确定表中 10 篇文献为认知诊断研究的经典文献,这 10 篇文献共被引 381 次,即占总文献量的 10.64% 的 10 篇文献其被引次数占被引总次数的 58.89%,可见这 10 篇文献在大陆认知诊断研究中的影响力相当之大。这些经典文献集中于 2004 年之后,在较短的时间内取得较高的被引次数,说明 2004 年后是到目前为止是大陆认知诊断研究走向繁荣并取得阶段性成果的时期。

表 4 前 10 名高引论文

题名	引次	作者	机构	期刊	发表时间
规则空间模型在描述统计学习模式识别中的应用研究	64	戴海琦,张青华	江西师范大学	心理科学	2004
运用规则空间模型识别解题中的认知错误	61	余嘉元	南京师范大学	心理学报	1995
新一代测验理论—认知诊断理论的源起与特征	51	刘声涛,戴海琦,周骏	江西师范大学	心理学探新	2006
新课程背景下的学业评价:测量理论的价值	49	辛涛	北京师范大学	北京师范大学学报	2006
项目反应理论研究的新进展	37	辛涛	北京师范大学	中国考试	2005
具有认知诊断功能的计算机化自适应测验的研究与实现	34	林海菁,丁树良	江西师范大学	心理学报	2007
Tatsuoka Q 矩阵理论的修正	27	丁树良,祝玉芳,林海菁,蔡艳	江西师范大学	心理学报	2009
测量理论的新进展:规则空间模型	20	辛涛,焦丽亚	北京师范大学	华东师范大学学报	2006
可达矩阵在认知诊断测验编制中的重要作用	19	丁树良,杨淑群,汪文义	江西师范大学	江西师范大学学报	2010
求解简化 Q 矩阵的扩张算法	19	杨淑群,蔡声镇,丁树良,林海菁,丁秋林	福建师范大学	兰州大学学报	2008

表 5 表 4 以外的论文被引次数分布

被引次数	15	14	13	12	10	7	6	5	4	3	2	1	0	合计
篇数	1	2	2	2	2	4	6	1	3	9	14	17	21	84
%	1.06	2.13	2.13	2.13	2.13	4.26	6.38	1.06	3.19	9.57	14.89	18.09	22.34	89.36

（五）作者分析

1. 核心作者分析

核心作者群的形成对学科的发展起着导向

和推向作用,对学科的发展有着重要的影响。本文 94 篇文献共有 236 位作者(包括所有合著者),其年代分布见表 6,表 6 表明参与认知诊断

研究的作者人数近 5 年来急剧增加,从 2004 年的 2 人增长到 2011 年的 74 人,这 8 年的平均增长率^①(张敏强,2010)达 67.51%,说明认知诊断研究吸引着越来越多的研究者参与其中。进一步分析发现,这 236 位作者有 122 人次的重复计算,不计重复计算的作者人数是 114,即我国大陆 2011 年前共有 114 位研究者公开发表过认知诊断研究的学术论文,其中,有 71.05% 的人只发表过一篇文献,发表 2 篇和 3 篇论文的作者比例分别为 13.14% 和 7.02%,有 8.77% 的作者发文 3 篇以上(表 7),说明在 114 位作者中,只有极少数作者在持续研究认知诊断,绝大多数作者只是偶尔涉足认知诊断研究,还没有形成持续稳定的研究方向。

为了揭示作者人数及其发表文献数量间的规律,探明研究者的生产能力及对科技进步和社会发展所作的贡献,美国学者 A. J. Lotka 在 20 世纪 20 年代提出了描述科学生产率的经验规律,称为洛特卡定律,又称“倒数平方定律”,即 $f(x) = f(1)/x^2$, $f(x)$ 是发表 x 篇文献的作者量,如 $f(1)$ 是发表一篇文献的作者数量(叶鹰,潘有能,潘卫,2006),该定律指出写两篇论文的作者数量约为写一篇论文的作者数量的 1/4;写三篇论文的作者数量约为写一篇论文作者数量的

1/9;写 n 篇论文的作者数量约为写一篇论文作者数量的 $1/n^2 \cdots \cdots$ 。本文中, $x^2 = f(1)/f(2) = 81/15 = 5.40$, $x = 2.32$, $x^2 = f(1)/f(3) = 81/8 = 10.13$, $x = 3.18$,此两值分别与洛特卡定律的经验值 2、3(晏尔伽,朱庆华,2007)接近,说明洛特卡定律适应于认知诊断研究的作者分布规律。

为了确认真知诊断研究是否形成核心作者群,本文采用普赖斯定律(叶鹰,潘有能,潘卫,2006)进行进一步分析。普赖斯定律即核心论文量占有专业论文总和的一半即表示核心作者群的形成,核心作者最低发文数 m 值为: $m = 0.749 \sqrt{n_{\max}}$,其中, n_{\max} 指发文最多的作者发表的论文数。本文中 n_{\max} 为 27,故 m 取整数为 4。即发文 4 篇以上的作者为核心作者群成员。表 7 表明,发文 4 篇以上的作者有 10 人,这 10 人发表论文 61 篇^②,占总论文量的 64.89%,高于论文总数的 50%,说明认知诊断研究已形成稳定的核心研究群体。表 8 表明丁树良、辛涛、戴海琦、涂冬波、蔡艳、余嘉元、汪文义、杨淑群、陈平、丁秋林构成了我国大陆认知诊断研究的核心群体,他们是我国大陆认知诊断研究的引领者。

表 6 开展认知诊断研究的作者人数分布

年份	1984	1995	1997	2004	2005	2006	2007	2008	2009	2010	2011	合计
作者数	1	2	4	2	1	8	15	37	34	58	74	236

表 7 发表论文数作者分布

论文篇数	1	2	3	4	6	7	8	9	13	16	27	合计
作者数	81	15	8	2	1	2	1	1	1	1	1	114
%	71.05	13.15	7.02	1.75	0.88	1.75	0.88	0.88	0.88	0.88	0.88	100.00

表 8 核心作者及文献量

作者	丁树良	辛涛	戴海琦	涂冬波	蔡艳	余嘉元	汪文义	杨淑群	陈平	丁秋林	张敏强	赵顶位
机构	江师	北师	江师	江师	江师	南师	江师	福师	北师	南航	华师	江师
文献量	27	16	13	9	8	7	7	6	4	4	3	3

注:江师表示江西师范大学,北师表示北京师范大学,南师表示南京师范大学,福师表示福建师范大学,南航表示南京航空航天大学,华师表示华南师范大学。发文 3 篇的作者共有 8 位,按作者发文期刊的级别进行排序,表中所列两位为排名前两位。表中数字指作者参与研究的文献量,包含作者以第一作者、通讯作者或其他排位所发的文献。

①平均增长率: $\bar{x} = \sqrt[n-1]{x_n/x_1}$
②合作者不重复计算篇数,故此数字不是由表中数字直接相加所得。

2. 合作研究分析

合作研究是现代科学技术发展的必然趋势,也是开展认知诊断研究的内在要求。开展认知诊断研究须具备认知心理学、学科教育、统计与测量技术及计算机编程四方面知识,研究的合作度越大,合作率越高,学科的发展水平往往越高。本文 94 篇论文由 236 位作者(包括所有合著者)完成,平均合作度^①为 2.51,说明平均每篇论文有 2—3 名作者合作完成。在 94 篇论文中,独著 16 篇,占总数的 17.02%,合著 78 篇,合作率^②为 82.98%。其中,65.96% 的论文

有 2—3 人合作,13.83% 的论文有 4 人合作,3.19% 的论文由 5 人合作完成(表 9)。可见,合作研究是大陆认知诊断研究的基本形式。进一步结合作者的单位信息可知,在合作形式上,同一单位的内部合作有 54 篇,占合著论文数的 69.23%;不同机构间的外部合作的有 24 篇,占合著论文数的 30.77%,没有境外合作或国际合作^③。这表明认知诊断研究应当在更广阔的背景中寻求合作,不仅要加强内部合作,更要注重外部合作。

表 9 合作情况统计表

作者人数	1	2	3	4	5	合计
论文篇数	16	33	29	13	3	94
%	17.02	35.11	30.85	13.83	3.19	100.00

(六) 学科分析

我国大陆认知诊断研究分布于 5 大学科,即心理学、教育学、计算机科学、文学和数学,其中,心理学文献量为 46 篇,占 46.94%,居首位;教育学 24 篇,占 24.49%,处第二;最后是计算机科学、文学、数学,分别占总发文量的 22.45%、3.06% 和 3.06%。可见,大陆认知诊断研究文献主要分布在心理学、教育学和计算机科学领域,体现了认知诊断研究的跨学科性。学科教育中的研究相对较少,学科教育的认知诊断研究应该得到一定程度的加强。

数的 9.57%,可以看出,国家级资助力度最大,省部级资助力度也不小,可见,我国大陆对认知诊断研究非常重视,市、厅及学校应加强资助力度。

(八) 期刊分析

在我国大陆,载文量居前三位的期刊是《心理学报》、《心理科学》和《中国考试》,分别载文 11、10 和 9 篇,累计载文 30 篇,占发文总量的 31.91%,根据布拉德福定律(叶鹰,潘有能,潘卫,2006),刊载论文总数约占 33% 的期刊是该学科的核心期刊,所以可以初步确定这 3 种期刊是刊登认知诊断研究文献的核心期刊。其中《心理学报》和《心理科学》是我国心理学期刊的顶级期刊,共载文 21 篇,其中,2011 年 10 篇,2010 年 5 篇,2009 年 2 篇,2008、2007、2004、1995 年各 1 篇,即近两年《心理学报》和《心理科学》成为登载我国大陆认知诊断研究的核心期刊,一方面说明近两年我国大陆的认知诊断研究成果处于较高水平;另一方面也说明我国学术界逐渐重视认知诊断研究。

(七) 基金分析

论文被不同级别的基金资助说明该研究在国家的受重视程度,同时也说明该论文是否是同领域的新动向、新视角、新方法。对 94 篇正式发表的期刊文献的基金项目进行统计,受项目资助的有 63 篇,占总论文数的 67.02%,其中,受到国家级项目资助的有 40 篇,占总论文数的 42.55%,部级项目资助的有 14 篇,占总论文数的 14.89%,省级项目资助的有 9 篇,占总论文

①合作度 = 作者人数/论文篇数
②合作率 = 合作发表的论文数/论文总数
③指 2011 年前的大陆认知诊断研究文献中没有国际合作,但 2012 年心理学报已刊登有国际合作的认知诊断研究文献。

四、结论与展望

本文得出以下结论:(1)大陆认知诊断研究从2004年起呈现快速增长趋势但文献总量低于国外;(2)80.85%的研究为基础研究,95%的研究采用模拟实验方法进行验证,2006年以前主要是理论介绍性文献,诊断方法研究集中在近两年,应用研究的诊断对象较广,但应用研究较少报告诊断的信度和效度;(3)研究主要围绕Q矩阵理论、认知诊断模型改进及开发和结合认知诊断的计算机化自适应测验;(4)合作研究是认知诊断研究的主要形式,已形成核心作者群;(5)文献被引率为77.66%,10篇经典文献占被引总次数的58.89%;(6)文献主要分布于心理学、教育学和计算机科学;(7)大部分认知诊断研究受到各级项目资助,其中,42.55%的认知诊断研究受到国家级项目资助;(8)成果发表以心理学期刊为主,《心理科学》和《心理学报》和《中国考试》是主要载文期刊,《心理科学》和《心理学报》发文占22.11%。

尽管当前大陆认知诊断研究取得较为丰富的研究成果,但在以下方面仍值得进一步关注:(1)在研究内容上,首先应加强构建认知模型的构建方法研究,这是开展认知诊断研究的理论基础。Gorin(2007)认为目前还未开发适合于成就测验的认知模型。我国大陆目前主要采用专家认定法、口语报告法和文献法确定认知模型,却未见有采用眼动研究等实验法确定认知模型的研究。其次,大陆开展认知诊断研究采用的模型主要是规则空间模型、属性层级模型和DI-NA模型,这与国际上已开发出60多种认知诊断模型(Fu,2005)有一定差距,因此应加强对其他模型的研究并继续进行模型的改进和开发。第三,新方法或新模型的提出采用模拟方法进行验证是科学的也是合理的,但如果希望认知诊断研究能更好地服务于实践,今后应采集实测数据加以验证以增强研究的外在效度。第四,认知诊断是一门探测学生心理加工过程的

科学,其最终目的是在解释心理活动的基础上服务于教育实践,在研究的初期加强进行基础研究是必要的,但为了实现“促进学生发展”的真正目的,有必要加大力度开拓应用领域的研究。第五,认知诊断仅报告诊断结果难有说服力,应加强诊断信度和效度的研究。第六,已有研究对数学学习、推理及阅读过程进行过诊断,今后应进一步拓展认知诊断对其他学科学习的诊断研究。(2)尽管认知诊断研究近年来发展较快,但从文献的作者分布来看,核心作者过于集中,只有极少数人在持续研究认知诊断,绝大部分研究者只是偶尔涉足该领域,这与开展认知诊断研究需要具备多学科基础且难度较大有关,但开展认知诊断研究的力量亟待扩大范围。(3)认知诊断是集认知心理学、心理测量学、计算机科学、数学、各学科教学为一体的科学理论,认知诊断研究的开展需要各方专家的通力合作才能顺利完成,因此,这是一项特别强调合作的研究。尽管合作研究已是认知诊断研究的主要研究形式,但局限于内部合作甚至主要是校内合作,且缺少认知心理学专家和学科专家的参与,因此,今后应加强与外部机构的合作研究,甚至跨境或跨国合作研究,加强与各科专家的合作研究,以促进认知诊断的更快发展。(3)认知诊断是通过探测学生学习过程内部的微观的心理加工机制来判断学习内容的掌握状况,比传统的笼统的宏观的一个总分给教师和学生带来更丰富的有关知识掌握的信息,从而更好地开展下一步的学习或教学,实现促进学生发展的目的。所以,从这个意义上来说,认知诊断的生命力在于为教育实践服务,虽然现有文献中有24.49%分布在教育学科中,但这些文献局限于介绍认知诊断理论或仅仅是呼吁教育实践中需要这种测验理论却没有更深入地研究认知诊断,因此,教育领域的专家应更加重视认知诊断理论对促进学生学习 and 发展的巨大作用,主动参与研究认知诊断并进一步加大对认知诊断成果的推广力度,实现认知诊断在教育实践过程中的应用。(4)文献的增长速度、发文的核心期刊及项目的资助力度均表明,认知诊断研究

在中国大陆越来越受到前所未有的重视,但是中国大陆的文献总量却远远低于国外,持续开展认知诊断研究的作者十分有限,被引文献也过于集中,10篇经典文献占被引总次数的近60%,这说明认知诊断研究在中国大陆的影响

还没有达到如美国通过立法来推广这种评估方法,中国台湾实现了认知诊断计算化自适应测验的应用的理想程度。因此,认知诊断研究在中国大陆的发展依然任重而道远。

参考文献:

- 陈德枝、戴海琦、赵顶位,2009:《规则空间方法与属性层次方法的诊断准确性比较》,《心理科学》第2期。
- 陈艳君,2012:《近十年〈心理科学进展〉文献计量分析》,《首都师范大学学报》(自然科学版)第4期。
- 方富熹,1984:《关于儿童算术的认知诊断》,《心理学动态》第3期。
- Buck, G. & K. K. Tatsuoka,1998, “Applications of the rule-space procedure to language testing: Examining attributes of a free response listening test”, *Language Testing*,15(2), pp. 119 – 157.
- Carpenter, P. A. , M. A. Just & P. Shell,1990, “What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test”, *Psychological Review*,97, pp.404 – 431.
- Costa, P. T. & R. R. McCrae,1992, “Multiple uses for longitudinal personality data”, *European Journal of Personality*, 6, pp. 85 – 102.
- Embretson, S. E. ,1998, “A cognitive design system approach to generating valid tests: Application to abstract reasoning”, *Psychological Methods*,3, pp.380 – 396.
- Fischer, G. H. ,1973, “The linear logistic model as an instrument in educational research”, *Acta Psychologica*, 37, pp.359 – 374.
- Frederiksen, N. , R. L. Glaser, A. M. Lesgold & M. G. . Shafto (eds.),1990, *Diagnostic Monitoring of Skill and Knowledge Acquisition*, Hillsdale, NJ: Erlbaum.
- Fu, J. ,2005, *A Polytomous Extension of the Fusion Model and Its Bayesian Parameter Estimation*, Unpublished doctoral dissertation, University of Wisconsin, Madison.
- Gentile, J. , K. Kessler & K. Gentil,1969, “Process of solving analogies items”, *Journal of Educational Psychology*,60, pp.494 – 502.
- Gorin, J. ,2007, “Test construction and diagnostic testing”, in J. P. Leighton & M. Gierl (eds.), *Cognitive Diagnostic Assessment for Education: Theory and Applications*, Cambridge, UK: Cambridge University Press, pp. 173 – 201.
- Haertel, E. H. ,1989, “Using restricted latent class models to map the skill structure of achievement items”, *Journal of Educational Measurement*,26, pp.301 – 323.
- Macready, G. B. & C. M. Dayton,1977, “The use of probabilistic models in the assessment of mastery”, *Journal of Educational Statistics*,2, pp.99 – 120.
- Mislevy, R. J. ,1993, “Foundations of a new test theory”, in N. Frederiksen, R. F. Mislevy & I. Bejar (eds.) *Test Theory for a New Generation of Tests*, Hillsdale, NJ: LEA, pp. 19 – 39.
- Nichols, P. D. , S. F. Chipman & R. L. Brennan,1995, *Cognitively Diagnostic Assessment*, Mahwah, NJ: Erlbaum.
- Ronning, R. R. , J. C. Conoley, J. A. Glover & J. C. Witt,1987, *The Influence of Cognitive Psychology on Testing*, Hillsdale, NJ: Erlbaum.
- Rupp, A. J. Templin & A. Henson,2010, *Diagnostic Measurement: Theory, Methods and Application*, New York: The Guilford Press.
- Tatsuoka, K. K. ,1983, “Rule-space: An approach for dealing with misconceptions based on item response theo-

ry", *Journal of Educational Measurement*, 20, pp. 345 – 354.

Tatsuoka, K. K. ,2009, *Cognitive Assessment: An Introduction to the Rule Space Method*, Florence, KY: Routledge.

Yang, S. , S. Ding & Q. Ding, 2010, "Incremental augment algorithm based on reduced Q-matrix", *Transactions of Nanjing University of Aeronautics & Astronautics*, 27(2) , pp. 183 – 189.

(责任编辑:蒋永华)

A Bibliometric Analysis of Cognitive Diagnostic Research in Chinese Mainland

DAI Hai-qi, XIE Mei-hua, DING Shu-liang

Abstract: In order to explore the present situation and developmental tendency of cognitive diagnostic research in Chinese Mainland, the article conducts a bibliometric analysis of the literature amount, subject contents, subject distribution, authors, institutions, citation analysis, periodicals distribution and foundation aids on 94 papers about cognitive diagnostic from 1980 to 2011 included in China Academic Journal Network Publishing Database (CAJD). The results are as follows. (1) Cognitive diagnostic research showed a rapid growth from 2004 in Chinese Mainland but the literature amount was less than foreign countries; (2) The theoretical researches made up 80.85% of the total and the other were the application of the cognitive diagnostic method. Of the studies in which a new method or a reformed model was developed, 95% verified the new method through the simulation experiment. The researches before 2006 were mostly introductory, while the diagnostic techniques were developed in the past two years. The applied researches were carried out in many disciplines but most researches failed to report diagnostic reliability and validity. (3) The main content of the study were Q matrix theory, Cognitive Diagnostic Model and computerized adaptive testing with cognitive diagnosis. (4) Cooperative research has become the main pattern and the rate of coauthoring is 82.98%, the co-operation rate was 2.51. The key research group has been formed and Ding Shuliang, Xin Tao, Dai Haiqi, Tu Dongbo, Cai Yan, Yu Jiayuan, Wang Wenyi, Yang Shuqun, Chen Ping and Ding Qiulin are the leading authors. (5) The percentage of cited papers in all published articles was 77.66%, and the average rate of each original article cited by other researchers was 6.88, 30 articles (31.91% of the total papers) were cited 5 times or more and 58.89% of the total frequency of citation come from 10 classic theses. (6) The literature was mainly distributed in the basic subjects of psychology, education and computer science and the literature amount of three subjects are respectively 46, 24 and 22. (7) Most of the cognitive diagnostic researches were funded by various grants and 36.17% of all the papers were supported by National Natural Science Foundation of China. (8) Most articles were published in psychological journals, and 22.11% were published in *Psychological Science* and *Acta Psychologica Sinica*.

Key words: Cognitive Diagnostic research; bibliometric analysis; Chinese Mainland