

◎孙燕平 著

认知能力测验

在儿童青少年智力结构
模型检验中的应用

Renzhi Nengli Ceyan
zai Ertong Qingshaonian Zhili Jiegou
moxing Jianyanzhong de Yingyong



山东大学出版社

山东省高水平应用型立项建设专业群
——山东体育学院特殊教育专业群经费资助

认知能力测验在儿童青少年 智力结构模型检验中的应用

孙燕平 著

山东大学出版社

图书在版编目(CIP)数据

认知能力测验在儿童青少年智力结构模型检验中的应用/孙燕平著. —济南:山东大学出版社,2018.11

ISBN 978-7-5607-6236-4

I. ①认… II. ①孙… III. ①儿童—认知能力—能力倾向测验 IV. ①G449.5

中国版本图书馆 CIP 数据核字(2018)第 270340 号

责任编辑:唐 棣

封面设计:张 荔

出版发行:山东大学出版社

社 址 山东省济南市山大南路 20 号

邮 编 250100

电 话 市场部(0531)88364466

经 销:新华书店

印 刷:济南乾丰印刷有限公司

规 格:720 毫米×1000 毫米 1/16

19.25 印张 356 千字

版 次:2018 年 11 月第 1 版

印 次:2018 年 11 月第 1 次印刷

定 价:49.00 元

版权所有,盗印必究

凡购本书,如有缺页、倒页、脱页,由本社营销部负责调换

前 言

儿童青少年是国家人力资源储备的后备军,其综合素质的高低直接关系到国家未来的综合国力与国际竞争力。因此,保障与促进儿童青少年的健康、积极发展是实现国家持续繁荣发展的重要战略性基础工作。素质教育的根本目的也是为促进社会与儿童青少年更好地和谐发展,而素质教育不仅重视知识,更要重视智慧。智慧不完全依赖于知识多寡,而在于知识的应用。作为儿童青少年获取和应用知识的心理工具,认知能力在其学业成就中发挥着重要作用。一般认知能力指智力,具体认知能力涉及感知觉、记忆、言语、表象和思维等方面的能力。因此,认知能力的发展是儿童青少年心理发展中一个非常重要的内容,不仅显著影响其日常的学习与生活,而且对其各个方面的发展也具有重要意义。

要研究个体认知能力发展,前提是研究者需要可靠且有效的测量工具来测查认知能力,有了多个版本的认知能力测验,还需要从众多的测验中选择出参考或者利用价值较高的测验。本书选取了在国际知名大型调查中广泛使用的认知能力测验:英国能力量表系列、伍德科克-约翰逊(Woodcock-Johnson)认知能力测验系列、英国认知能力测验系列、韦氏儿童智力量表系列及认知评估系统系列,并且从理论基础、能力结构、心理测量学指标、测验修订完善以及测验结果报告等方面进行阐述,以便为国内认知能力测验的开发、修订和使用提供参考。

对于儿童和青少年来说,认知能力对其取得学业成功具有重要的影响和作用。本书探讨了对儿童和青少年学业成就有显著影响的具体认知能力,如知觉、记忆、思维、概念、言语和阅读、数量知识和智力等方面。有研究者可能认为概念、阅读理解和数量知识属于学业成就范畴而非认知



能力领域,但是纵览国际上的认知能力测验,多数认知能力测验都不同程度地测查了概念、词汇知识、阅读理解和数量知识的部分,而且这些任务的完成与认知能力密切相关,因此,本书将这些内容归入认知能力部分。另外,认知能力测验所测查的结果会影响对儿童青少年的能力评估,甚至会影响到其未来接受的教育内容和干预计划。研究者会利用认知能力测验结果对其认知能力发展异常进行诊断、分类、干预、治疗和预防,因此,认知能力测验的施测必须要科学客观,施测者应该遵循伦理、法律和职业规范,获得实施不同层次心理测验所要求的资质。

儿童青少年的认知发展与他们所处的文化息息相关。不同国家和民族的儿童不会发展出完全相同的心理内容,尽管他们会学习使用人类所特有的大脑和心理能力来解决问题,但他们与周围环境互动时行为模式与本土文化要求和价值观相一致。早在1978年,朱智贤先生就提出中国化的心理学研究。他说:“中国儿童和青少年,与外国的儿童与青少年的心理特点,既存在着普遍性,又具有其特殊性。只有研究出中国儿童与青少年心理发展的特点,才能在国际心理学界有发言权。”因此,有必要对我国儿童青少年的智力结构的较优模型进行验证,并与西方儿童和青少年的较优智力结构模型进行比较,探索出中国儿童青少年独特的智力结构模型。本书在了解不同文化背景下认知能力差异的基础上,采用验证性因素分析方法检验CHC理论的模型,不仅为CHC理论的丰富和补充提供实证方面的证据,也有助于理解中国特色文化环境下儿童青少年的智力结构。

最后,由于时间仓促以及作者水平所限,书中错误之处在所难免,敬请专家和读者不吝赐教。

孙燕平

2018年8月

目 录

第一章 认知能力测验	(1)
第一节 五种认知能力测验的理论基础及发展历程	(2)
第二节 五种认知能力测验的基本内容	(13)
第三节 五种认知能力测验测查的潜在能力	(24)
第四节 五种认知能力测验的结构框架及所测的具体能力	(35)
第五节 五种认知能力测验的心理测量学指标	(44)
第六节 五种认知能力测验内部及与其他测验的比较	(82)
第七节 认知能力测验的分数报告	(97)
第二章 认知能力及其测量	(165)
第一节 知 觉	(166)
第二节 短时记忆和工作记忆	(170)
第三节 思 维	(174)
第四节 概 念	(178)
第五节 言语与阅读	(182)
第六节 加工速度	(186)
第七节 数量知识	(188)
第八节 智 力	(194)
第九节 认知能力测量	(201)
第三章 跨文化背景下的认知能力比较	(216)
第一节 不同文化背景下认知能力的差异	(217)
第二节 智力的群体差异	(223)



第四章 智力结构的 CHC 模型的理论基础 (230)

 第一节 智力结构的 CHC 理论 (230)

 第二节 智力结构的 CHC 模型的构建 (237)

第五章 智力结构的 CHC 模型的检验 (249)

 第一节 中国儿童智力结构的较优模型的检验 (249)

 第二节 美国儿童智力结构的较优模型的检验 (263)

附 录 (279)

主要参考文献 (284)

第一章 认知能力测验

认知能力涉及感知觉、注意、记忆、想象、言语和思维等方面的能力。认知能力的发展是儿童青少年心理发展的重要方面,因此,认知能力的发展规律成为发展和教育心理学研究领域的重要课题。要探讨儿童青少年认知能力不同方面的水平高低,以及认知能力与其学业成就之间的关系,前提是研究者要具备能够测查具体认知能力的可靠且有效的工具,本章选取国际上知名且常用的五种认知能力测验系列作为测量工具,即英国能力量表系列(如 British Ability Scale-II, BAS-II & British Ability Scale-3, BAS-III)、伍德科克-约翰逊(Woodcock-Johnson)认知能力测验系列(如 WJ-III & WJ-IV COG)、英国认知能力测验系列(如 Cognitive Abilities Test: third Edition, CAT3&CAT4)、韦氏儿童智力量表系列(如 Wechsler intelligence scale for children-IV, WISC-IV)及认知评估系统测验系列(CAS)(Hill, Vivian, 2005; Taub, McGrew, 2004; Smith, Fernandes, Steve, 2004; Wechsler, Kaplan, Fein, 2004; Naglieri, Das, 1997)。尽管这五种测验都测查个体的认知能力,但是在理论基础、结构框架、适用对象和测查形式等多个方面存在差异。因此,本章主要从以下方面来描述五种认知能力测验的情况:第一节,五种认知能力测验的理论基础及发展历程;第二节,五种认知能力测验的基本内容;第三节,五种认知能力测验测查的潜在能力;第四节,五种认知能力测验的结构框架及所测的具体能力;第五节,五种认知能力测验的心理测量学指标;第六节,通过分析当前测验较以往的改进或优势之处,对五种认知能力测验内部及与其他测验进行比较;第七节,呈现认知能力测验的分数报告,目的是让校长、教师及家长能了解学生群体及个人的能力状况。



第一节 五种认知能力测验的理论基础及发展历程

一、五种认知能力测验的理论基础及发展历程

(一)英国能力量表系列的理论基础及发展历程

1. 英国能力量表系列的理论基础

埃利奥特(Elliott,1997)以霍恩-卡特尔(Horn-Cattell)理论为基础进行了BAS-II的编制工作,Elliott认为智力可以分为不同的部分,这些部分分别代表了意义不同的智力成分且相互之间存在着质的差异。Horn-Cattell理论模型源自Horn(1991)对于Cattell(1941)工作的拓展,在Cattell流体-晶体理论(Gf-Gc)理论上获得了共10个广泛意义上的能力:流体智力(Gf)、晶体智力(Gc)、短时获得和提取(Gsm)、视觉智力(Gv)、听觉智力(Ga)、长时储存和提取(Glr)、认知加工速度(Gs)、决策时间(Gt)、量化知识(Gq)和基于阅读和写作技能中理解和表达的因素(Grw)。其中模型中与三个信息加工系统(言语、视觉空间及非言语)相关最大的认知能力是:晶体智力(Gc)、视觉智力(Gv)和流体智力(Gf)。该理论示意图如图1-1所示。

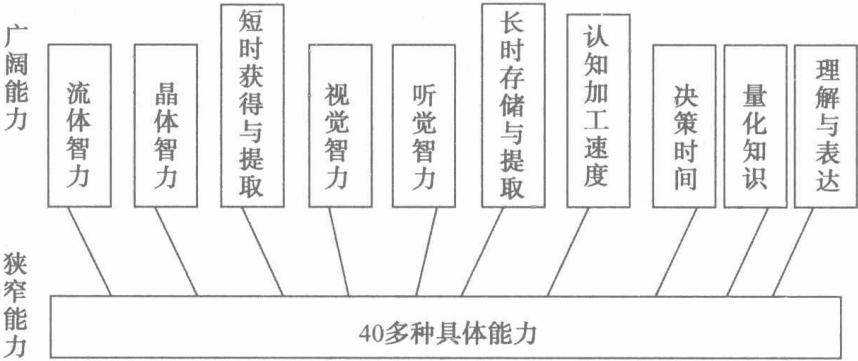


图 1-1 Horn-Cattell 能力结构理论示意图

除了以Horn-Cattell理论为基础,BAS-II的设计还基于其他关于人类能力的一系列心理模型、理论和范式,Elliott(1997)描述如下:

- (1)人类能力不是由单个认知因素(g)来解释的,或者甚至不是由两个或者三个更低级别的因素来解释。
- (2)人类能力由多种维度构成。儿童因在学习、获得成就和解决问题上存在思考方式的不同,使得可观察到的行为上存在差异,这些不同的行为构成了人类能力的不同维度。



(3)人类能力之间是相关的,但并非完全重叠的,因此,人类的能力之间存在着差异(Carroll,1993)。

(4)人类能力的各个部分表征了信息加工的多个相互连接的子系统。

(5)信息加工的子系统与中枢神经系统有相关,一些加工功能是定位的,而其他的是整合的。

2. 英国能力量表系列的发展历程

最初的BAS系列测验产生于1979年,心理测验专家编制该系列测验的原因在于:一方面,英国应用心理学界普遍不满韦克斯勒儿童智力量表(WISC)和斯坦福-比奈(Stanford-Binet)智力量表在英国的使用情况,他们不仅认为这些测验与英国本土的伦理、文化和社会存在巨大差异,而且对只是根据测验分数鉴别需要特殊教育的儿童和安置人员这种方法表示质疑。另一方面,这些测验花费时间较长,通过求和方式获得的智商(IQ)分数对于诊断也只能提供有限信息。因此,为更全面具体地了解英国儿童的学习风格和认知能力特点,进而依据儿童本身的个体特点对其进行干预,英国应用心理学家开始了适合本国不同种族、语言和文化背景儿童和青少年测验的编制工作,基于此目的,心理测验从最初的英国智力量表(British Intelligence Scale)改名为英国能力量表(British Ability Scale,BAS)。

最初的BAS系列测验主要测量对象是年龄范围在2岁6个月到17岁6个月的儿童青少年,涉及的主要信息加工领域为:速度、推理、空间图像、感知匹配、短时记忆和知识的提取与应用。比起美国的个别能力测验,如WISC和斯坦福-比奈(Stanford-Binet)智力量表,BAS测验包括的内容更广泛,其执行程序更严格。

基于BAS系列测验在英国的应用和发展,Elliott在美国心理公司的资助下着手差异能力量表(Differential Ability Scale,DAS)的编制工作。DAS是BAS在美国的发展,在咨询实践者关于测验的使用情况的基础上,Elliott编制DAS过程中对最初BAS的结构进行了改进,其中较为显著的差异是DAS剔除了原有测验中的6个量表,增加了4个新量表。这些测验的发展为BAS-II的发展奠定了基础。Elliott(1997)在已有研究的基础上指出Horn-Cattell模型是有助于构建测验的有效模型,它有相当大的因素分析效度(Carroll,1993;McGrew,1996)。因此,以Horn-Cattell理论模型为基础,Elliott(1997)构建了BAS-II。

BAS-II产生于1997年,经过了10多年的时间,BAS-III在2012年完成修订并出版。BAS-III距今不到10年时间,是目前英国最新且最成熟的认知能力测验。BAS经历了30多年的漫长历程,从最初的BAS测验,到BAS的修订,再到BAS在美国的发展与应用(DAS),BAS-III已经成为当前较为成熟且被广泛应用的测验。



BAS-III 是在 BAS-II 的基础上修订得到的,这一版本的认知能力测验便于使用,采用多重选择的项目形式并且修订了数字技能量表、拼写量表,增加了词汇阅读的复本测验。BAS-III 与 BAS-II 相比较,在图片命名、词汇命名和口语理解量表上改变较小。BAS-III 是建立在 BAS-II 的优势之上。尽管 BAS-III 对以往版本进行了修订和重新标准化以确保分数有效、可靠并且是跟得上时代发展的,这一版本的认知能力测验仍然拥有 BAS-II 的整体框架和量表特征,以便于熟悉 BAS-II 的测验人员能够高效使用它。

(二)Woodcock-Johnson 认知能力测验系列的理论基础及发展历程

1. Woodcock-Johnson 认知能力测验系列的理论基础

卡特尔-霍恩-卡罗尔(Cattell-Horn-Carroll,CHC)认知能力理论是 Woodcock-Johnson 认知能力测验第四版(WJ-IV)的理论基础。该理论由麦克格鲁(McGrew,1997)在 Cattell-Horn 理论和 Carroll 三层理论的基础上提出,他认为人类能力本质上是分层的:一般认知能力(g)处于模型的顶点;广阔能力因素包括流体智力、晶体智力、长时存储与提取、视觉智力等因素,处于模型的第二层;大量的具体能力处于模型的第一层,通常由具体测验测量获得。该理论能力结构如图 1-2 所示。

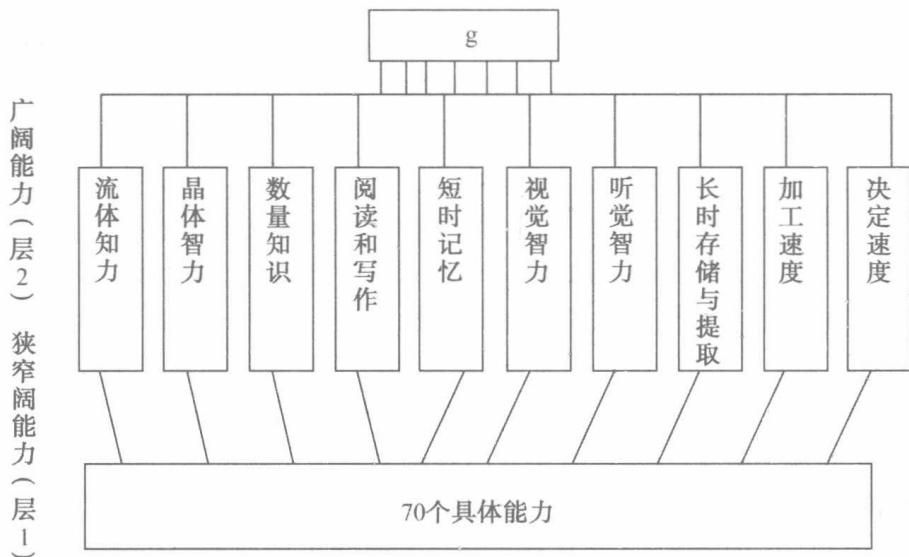


图 1-2 CHC 理论示意图

CHC 理论是基于 McGrew 对 Cattell-Horn 理论和 Carroll 三层理论的分析提出的,其中 Cattell-Horn 理论由雷蒙德·卡特尔(Raymond Cattell)和约翰·霍恩(John Horn)提出,该理论又称“流体智力-晶体智力理论”(Gf-Gc 理论)。流体智力和晶体智力最初由 Cattell(1941)进行区分。Horn(1965)以因素分析



研究结果为证据,鉴别出其他不同的广阔能力,包括短时记忆、长时提取、加工速度和视觉-空间思考能力。Horn 和斯坦科夫(Stankov,1982)识别并增加了听觉加工(Ga)。Horn(1988,1989)进行了进一步研究,在此研究的基础上又包括了不同于 Gf-Gc 模型中其他能力的因素——数量能力(Gq)。最后,Woodcock (1998)鉴别到了一个不同的阅读-写作能力(Grw)。

CHC 理论的另一个来源为 Carroll 三层理论,该理论在 John Carroll 进行现有的因素分析的次级分析基础上提出(Carroll,1993)。Carroll 提取了 461 个数据系列并对这些数据系列进行了重新分析。他使用探索性因素分析方式,进一步完善了“人类认知能力能以分层来定义”这一理论。Carroll 定义了 69 个具体或者狭窄能力,并将这些能力定义为层 I 能力。这些狭窄能力被分组成广阔的认知能力(层 II),如流体智力、晶体智力、一般记忆和学习、广阔的视觉感知、广阔的听觉感知、广阔的提取能力、广阔的认知速度和加工速度。在这个模型的顶端(层 III),Carroll 提出了一般智力(一般因素)或者 g。

CHC 理论通过精简完善已经超越了最初的版本(Schneider, McGrew, 2012)。此外,在整个 WJ-IV 的发展过程中,由 Cattell、Horn、Carroll、Woodcock 和同事提出的 CHC 理论架构已经被相关的研究进行了跨领域交叉验证和完善,CHC 理论结构更加清晰,而且反映了完善后最新的 CHC 理论。对于 WJ-R、WJ-III 和 WJ-IV 常模样本的分析提供了 3 个大型的多重能力数据系列,这些系列证实或者修正了最初的理论架构。WJ-IV 的模型在工作记忆、听觉加工、词汇通达速度和声音类型记忆的研究得到验证,WJ-IV 的模型也得到神经科学研究的支持(LaForte, McGrew, Schrank 2014)。

2. Woodcock-Johnson 认知能力测验系列的发展历程

施兰克(Schrank,2001)指出,从最初的 WJ 认知能力测验(Woodcock-Johnson Psycho-Educational Battery)(Woodcock,Johnson,1977),到 1989 年 Woodcock 对此版本认知能力测验的修订产生的 Woodcock-Johnson 认知能力修订版测验(WJ-R COG)(Woodcock,Johnson,1989),再到 WJ-III(Woodcock,2001),最后到 Woodcock-Johnson 测验第四版认知能力测验(WJ-IV COG,2012),WJ 认知能力测验的发展经历了四次较大的修订,以下分别来阐述 WJ 认知能力测验的四次修订。

(1)Woodcock-Johnson Psycho-Educational Battery(WJ)(Woodcock-Johnson,1977):WJ 测验系列包括三部分:认知能力测验、成就测验和兴趣水平测验。第一部分认知系列包括 12 个测验,测量言语和非言语功能。这个测验测查认知能力,从低水平加工如简单操作,到高水平加工如复杂的心理操作。Woodcock 首先发展的能力是视觉-听觉学习(Woodcock,1958),设计该能力预测个体阅读的情况。



接下来发展的能力是分析-综合,该能力被用来测查个体数学学习能力。Woodcock 构建了认知测验的其余部分,并测量了一系列广阔和复杂的认知能力。

Woodcock 在 1973 年开始发展 WJ。他采用科学-实验的方法发展测验,主要是使用一系列控制性实验来测量学习能力。WJ 系列在 1976~1977 年被标准化。标准化之后,Woodcock 分析了各因素和不同组,并定义了以下 4 个广阔的认知功能领域:知识-理解、推理-思考、记忆-学习和区分-感知。

20 世纪 70 年代,当人们对智商的有效性产生怀疑时,广阔认知能力被引入作为个体总体认知分数。该分数的获得主要是通过对 12 个认知测验进行不同加权得到,成为衡量个体认知能力的一种指标。

(2) Woodcock-Johnson 认知能力修订版测验(WJ-R COG)(Woodcock, Johnson, 1989): WJ-R COG 被分成两种不同系列:认知能力测验和成就测验。修订版的理论源自 Woodcock 对 Horn 关于 Gf-Gc 理论的操作性表征(Horn, 1991)的见解。Woodcock 计划并组织了这个测验使之更接近 Gf-Gc 理论(Horn, 1991)。为达到这个目标,他发展了 10 个新的认知测验即 WJ-R COG。这个新测验测量了以下 7 种广阔能力:长期提取(Glr)、短时记忆(Gsm)、加工速度(Gs)、听觉加工(Ga)、视觉加工(Gv)、理解-知识(Gc)、流体推理(Gf)。这些能力都包括在 Gf-Gc 理论中,在此意义上, WJ-R 认知能力测验是测量 Gf-Gc 理论中 7 种广阔能力的测验系列。

(3) WJ-III(Woodcock, 2001): 约翰·卡罗尔(John Carroll)于 1993 年提出三层理论,这个理论将人类认知能力分成三层层级。在第一个水平(层 I), Carroll 鉴别了 69 个具体或者狭窄能力,这些能力类似于由 Horn 识别的能力(Ekstrom, French, Harman, 1979)。Carroll 将这些狭义能力分组成为更广阔的认知能力分类,这些能力构成第二水平上的能力因素,第二水平(层 II)能力因素类似于 Horn 的 Gf-Gc 因素。第三水平(层 III)识别了一般智力能力(g)因素。

(4) WJ-IV COG(LaForte, McGrew, Schrank, 2012): Woodcock-Johnson 认知能力测验第四版(WJ-IV)是当代测量学生认知能力、口语和成就的测验中最综合性的测验系统。该测验是个别施测的心理教育测验。WJ-IV COG 是 WJ-IV 测验的子测验,除了该认知能力测验, WJ-IV 还包括口语测验(Woodcock-Johnson IV Tests of Oral Language, WJ-IV OL)和成就测验(Woodcock-Johnson IV Tests of Achievement WJ-IV ACH)(Schrank, Mather, McGrew, 2014)。整套 WJ-IV 测量了一般智力能力(g)、具体认知能力、口头语言能力和学业成就,能够评估个体在这三个方面的优劣势,有助于诊断学习问题并发展出针对个人的干预措施。常模建立在全国性样本基础上,共 7416 人,年龄范围在 2 岁到 80 多岁。



WJ-IV COG 测验基于人类认知能力的 CHC 理论,提供了最现代化的测量模型。与以往版本比较,最新版的 WJ 认知能力测验能够在现有的研究成果和专业实践需要上创造新的测验及其系列,测量最重要的认知能力;另外,该测验通过增加所选测验的认知复杂性,关注认知测验系列的生态学和预测效度;最后,该测验为了与认知加工测验、语言能力测验和学业成就测验进行对比来决定个体在不同领域的能力优势,增加了新的流体-晶体认知合成分数(Gf-Gc cognitive composite)。WJ-IV COG 提供了一个新的且具有全国代表性的大型样本,更新了项目,简化了测试的管理和解释程序,能够为年龄在 2 岁到 80 多岁的个体提供一般智力能力。该测验利用最先进的数据收集、测试开发和数据分析方法作为模型,促进应用测试领域的进步。

(三)英国认知能力测验系列的理论基础及发展历程

1. 英国认知能力测验系列的理论基础

CAT3 的理论框架是基于人类能力两个理论模型——弗农(Vernon,1961)分层理论模型(见图 1-3)和 Cattell(1987)流体-晶体智力理论模型(见图 1-4)。这两个模型都是因素分析后得到的分层模型。这两个模型都包括了从一般到具体几个分层的因素,其差异在于一般因素的数量不同。Vernon 提供了称为“g”的一般推理因素,这个因素支配着所有其他因素。Cattell 的模型提供了两个一般推理因素——流体-分析推理(通常由 Gf 表示)和晶体推理(通常由 Gc 表示),这两个因素支配着模型中的其他因素。这两个模型中,Cattell 的模型一般被认为是更具综合性且更完整的模型,尽管两个模型都是认知能力领域广为接受的模型,但哪个模型都没有被认定为人类能力模型。

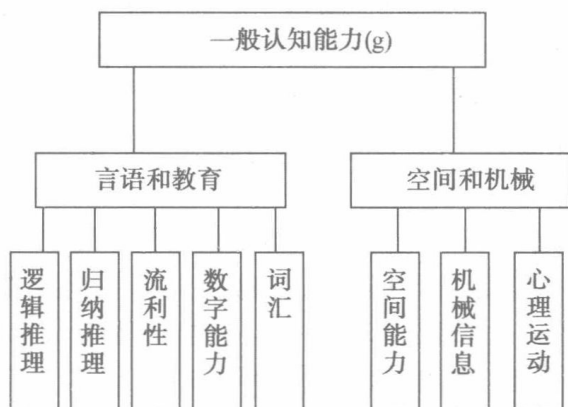


图 1-3 Vernon 分层理论模型

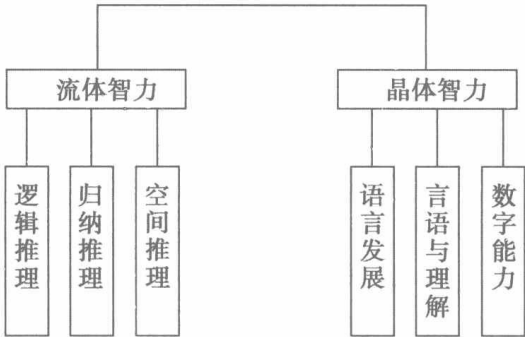


图 1-4 Cattell 流体-晶体理论模型

Vernon 的模型提供了两个主要的组因素,其中一个组因素是言语和教育。这个因素支配着更具体的因素,如逻辑推理、归纳推理、流利性、数字能力和词汇。另一个组因素为空间和机械因素。它支配着更多具体因素,如空间能力、机械信息和心理运动协调等。这两个主要组因素由一般推理因素 g 来支配。Vernon 指出,言语和教育与学科分数有更高相关,而空间和机械与技术科目分数有更高相关。他也指出,能力类型和结构随着个体接受的训练或者教育类型而改变。CAT3 的构建主要应用了 Vernon 模型中的一般推理因素 g ,这个 g 支配着所有组因素,这些组因素又支配着更具体的因素。

Cattell 将流体分析能力看成一般能力,他认为流体智力(Gf)表征个体获得知识的基本能力,要求归纳、推理或者空间推理能力的认知任务都是对 Gf 很好的测量;把晶体智力定义为个体通过文化传承(包括一般经验和正式学校教育)方式获得的能力(部分基于他或她的 Gf 水平),该智力包括如母语发展、词汇或者词汇知识、言语理解和数字能力等方面,相对于流体智力对晶体智力的影响,晶体智力对流体智力的影响更小。Cattell 和 Horn 指出使用非言语符号的测验任务是对 Gf 的很好测量,并且与言语和数量任务比较更少地受到文化传承的影响。

Cattell 的理论比起 Vernon 的理论更多地被用来构建 CAT3。基于 Cattell 和 Horn 的研究及流体-晶体智力的概念,测验编制者据此指导测验任务的选择。

2. 英国认知能力测验系列的发展历程

史密斯、费尔南德斯、斯特兰德(Smith, Fernandes, Strand, 2001)指出,从 20 世纪 70 年代早期认知能力测验(CAT)的最初形式在英国发展,到 20 世纪 80 年代认知能力测验的重新标准化,形成认知能力测验第二版(CAT2E),到 21 世纪初认知能力测验第三版本的完成(CAT3),最后,认知能力测验第四版出版



(CAT4), 认知能力测验经过了 40 多年的漫长历程。

(1) CAT 和 CAT2E: 认知能力测验第一版本的标准化发生在 1972 年, 标准化样本代表了全国范围内的学龄儿童, 这涉及水平 A 到 H (即 4 年级到 11 年级) 的儿童。12 年后, 研究者决定重新标准化这个测验, 并且考虑到使用第一版认知能力测验的学校的反馈, 测验的问题内容和顺序都没有变化, 但是水平 G 和 H 的测验部分被删除, 因为它们很少在第一版本中被使用。认知能力测验第二版本 (CAT2E) 在 1984 年秋季被标准化, 在给受测者执行主要的测验系列之前先进行练习测验。近 550 个学校和超过 11000 名学生参与了这次测验。

(2) CAT3 的发展: 为了使常模反映出 21 世纪早期英国儿童的能力, 1999 年研究者决定发展认知能力测验的新版本。这次的修订包括了 CogAT™ 在 CAT 和 CAT2E 基础上做出的一系列改进。CAT3 以在美国发展起来的 CogAT6™ 的测验形式作为基础, 该测验在测验内容上较 CAT2E 有较大的改变。

(3) CAT4 的产生: 英国的 GL 评估机构与学校合作 30 多年, 目的是帮助学生了解学生的发展能力、可能的学术潜力和学习偏好, 以促进学生取得更高的学业成就, 进而提高学生的学业成就。该机构修订完善了 CAT3 版本, 在 2011 年提出 CAT4。CAT4 是 GL 评估机构所开发最为完善的认知能力测试, 也是目前英国最广泛使用的推理测试能力。GL 评估机构认为, 对于学生的能力和成就评估, 除了老师的经验和判断力之外, 还应该利用测验为每个学习者提供全面客观的观点和认识。CAT4 能够提供三个领域的重要见解: ①能力方面: 这方面的分数表明学习者的潜在能力; ②成就方面: 该领域的分数显示学习者现在的位置以及同辈或者教师能做些什么来支持他们取得可能达到的最高成就; ③学习障碍方面: 该领域的分数帮助教师识别妨碍学习者的障碍所在, 并力图挖掘学生的最大潜力。

教师可以采用 CAT4 评估学生的推理能力, 不依赖于学生已有的学生成就、母语类型以及学校教育的课程材料, 提供了一种强有力的、标准化的认知推理能力测试方法。与全国平均水平相比, 通过 CAT4, 测试人员可以了解学生在语言能力、非语言能力、数学能力和空间能力的推理能力。教师可以根据测验结果调整教学方法, 以进行因材施教。CAT4 的改进在于能够通过电脑自动生成一系列的报告, 提供了更丰富和更全面的评估数据。使用者可以从 8 个报告中选择特定的报告。不同的报告根据使用者的身份不同, 呈现的内容也不同, 有的报告中叙述更详细, 使得报告更容易阅读和理解。报告中可以显示统计表、条形图和散点图等统计图表, 还包括对测试结果的清晰说明, 能解释学生的一般情况以及认知能力对教学和学习的影响。CAT4 另一个较大的变化是更加关注空间



能力的测查,增加了关于空间能力的新测验。为了使得 CAT4 更加规范,2011年秋季,对 25000 名中小學生进行了标准化测试。

(四)韦氏儿童智力量表系列的理论基础及发展历程

1. 韦氏儿童智力量表系列的理论基础

韦氏儿童智力量表第四版(WISC-IV)提供了对一般智力功能的测量(FSIQ)和 4 个指标分数。双重 IQ(言语 IQ 和操作 IQ)和在 WISC-III 中执行过的指标分数结构不再使用。新的框架是基于理论并得到临床研究和因素分析结果的支持。Wechsler(2004)指出,基于对能力理论的不斷完善和更新,WISC-IV 引入新的子测验测量流体推理、工作记忆和加工速度。有研究者指出,认知能力的理论强调了流体推理的重要性(Carroll,1997;Cattell,Horn,1978;Sternberg,1995)。现代研究表明,工作记忆是流体推理的重要部分,并且与成就和学习有关系(Fry,Hale,1996;Perlow,Moore,1997;Swanson,1996)。研究者指出,信息加工速度与阅读表现、阅读发展都有关系,信息加工被看成是认知能力因素分析研究中认知能力的重要方面(Carroll,1993,1997;Horn,Noll,1997)。WISC-IV 测验提供了大量的子测验测量个体拥有的重要认知功能(Wechsler,2004)。

2. 韦氏儿童智力量表系列的发展历程

施瓦兹(Schwartz,2004)表明韦氏测验从 1939 年韦克斯勒-贝尔维尤(Wechsler-Bellevue)测验的编制到 2003 年 WISC-IV 的形成,其发展历程较其他测验更漫长。韦氏测验发展历程的简表如表 1-1 所示。

表 1-1 韦氏测验的修订历程

韦氏成人智力量表	韦氏儿童智力量表的发展	韦氏学前儿童智力量表
W-B 1939	W-B II 1946	WPPSI 1967
WAIS 1955	WISC 1949	WPPSI-R 1989
WAIS-R 1981	WISC-R 1974	WPPSI-III 2002
WAIS-III 1977	WISC-III 1991	
	WISC-IV 2003	
	WISC-V 2014	

资料来源:[http://163.17.111.5/Manasystem/Question1/images/WISCIV%20powerpoint. pdf](http://163.17.111.5/Manasystem/Question1/images/WISCIV%20powerpoint.pdf)

Wechsler(2004)指出,最初的韦氏儿童智力量表(Wechsler,1949)改编自 Wechsler-Bellevue 儿童智力量表(Wechsler,1939)中的信息,如算术、类比、词汇、数字广度、理解、图画完成、图画安排、组块设计、物体装配和译码子测验。另外,迷宫子测验特别为韦氏儿童智力量表(WISC,1949)而发展,使得子测验的