

第五章 认知能力的评估

杨福义

博士 副教授

华东师范大学学前教育与特殊教育学院

E-mail: fy yang@spe.ecnu.edu.cn

Tel: 021-62232276

第一节 认知能力评估概述

一、认知能力概述

- 认知即认知过程的总称，包括感知觉，注意，记忆，想象，思维、问题解决、语言等一系列心理活动。
- 认知领域包含大量与智力相关的能力和技能，它们与注意力，辨别力、模仿、空间关系、时间关系，推理，分类，排序，序列和问题解决有关。

二、认知能力评估方法

- 标准化的发展测验
 - 常模参照测验
 - 标准参照测验
- 质性发展取向的评估方法
 - 实作评估
 - 动态评估
 - 课程性评估（如：卡罗莱纳特殊需要婴儿和学步儿课程（第二版））

- 顺序量表。即以项目间的层级关系为基础来选择认知项目，某一认知领域的项目是依据难度水平进行选择 and 排列的，是否进入到下一个项目或能力水平，则取决于儿童是否掌握了前一个项目。
- 观察访谈。如：儿童认知发展的主要特征是对物体的探索和操作，这可以通过观察儿童的吸允，扔东西和对物体的视觉追踪等行动而获得信息。
- 此外，游戏作为一种替代性方法也可以成为认知能力的评估方法

三、评估技术和工具的选择要点

- 认知评估中使用的技术因儿童的年龄和个人特点而异。具体要考虑以下因素：
 - 年龄差异
 - 儿童的特点
 - 年龄范围
 - 多领域或单领域

四、影响智力评估的因素

- 智力评估受到许多因素的影响，如果不对这些因素加以控制，可能会妨碍它在特殊教育中作用的发挥。具体而言，智力的评估受到以下因素的影响：

1、智力理论

- 有什么样的智力理论就会有怎样的智力评估。
- 如研究者支持智力是由多种因素组成的，他在编制智力评估测验的时候就会考虑如何测量这些因素。
- 愈来愈多的人接受动态结构理论，因此把评估的重点转向智力活动过程。

2、测验工具

- 一般来说，用诊断性测验较之于用筛查性测验获得的信息更多，结果也更可靠。
- 诊断性测验的信效度通常都会比筛查性测验高

3、评估人员专业素质

- 诊断性测验要求较高，如果评估人员不具备较扎实的心理学和测量学的知识，较熟练的测验操作技能，那么所得到的智力测验结果的科学性值得怀疑。

五、智力评估所引发的问题

- 虽然智力评估在特殊教育领域已获得了广泛的应用，但是由此也引发了一些问题，这些问题需要我们给予一定的关注和思考。

1、评估人员的资格

- 个别施测的评估测验对评估人员的素质要求较高
- 建立和完善评估人员培训制度和持证上岗制度势在必行

2、标签效应问题

- 当某个儿童经过智力评估被诊断为“智力障碍”时，这个名称或分类标志可能像标签一样一直跟随着他，随其身心发展产生极其不良的影响。
- 即使儿童被贴上“天才”的标签，这个标签所产生的作用也不一定是积极的。

3、测验滥用问题

- 测验的发行和使用上缺乏一套严格的管理制度，无论实施还是接受测验都是十分方便的。
- 频繁使用智力测验会产生练习效应，也会使测验者兴趣降低，最终都会导致结果不准确。

六、智力评估新进展

- 1、强调过程的评估（戴斯-纳格利尔里认知评估系统（**DN-CAS**））

2、强调动态的评估

- 一些心理学家如福尔斯坦，古斯克在维果茨基最近发展区理论的指导下制定了一系列动态评估系统。
- 动态评估的优点
 - 1、目的不在于分类，而在于发现学习潜能
 - 2、以成功为导向，在评估过程中给被评估者提供实质性的帮助。有助于帮助儿童的发展；
 - 3、给所有的被评估者都提供学习的机会，由进步的大小来判断智力水平的高低，

3、评估的范围在扩大

- 从20世纪80年代以来提出的智力理论可以看出，智力的外延有扩大的趋势。
- 例如在加德纳的多元智力理论中，音乐能力，身体运动能力等都被看成是智力。
 -

第二节 常用智力测验简介

- 智力和学习能力的评估主要借助心理测验来进行，这类测验又分为筛查性测验和诊断性测验两大类。
- 筛查性测验可以用快速测查方法对智力做出初步评定，其特点是测验内容少、时间短、工具简单，但内容上有些片面和过于简单，只能反应智力一个部分或一个侧面。

- 而诊断性测验的测验内容较为全面、广泛，测题数目较多，测验的编制过程严格遵循心理测量的技术要求，经过严格的标准化，结果趋向于定量化。诊断性测验具有较为准确、全面、信度和效度都比较高等特点。
- 经过筛查性测验选择出来的儿童需用标准化的诊断性量表进一步评估。

一、常见的筛查性智力测验

1. 分类测验 (Set Test)

- 适用于大年龄儿童。
- 要求儿童报出**10个城市、10种水果、10种颜色、10种动物**。
- 每报出一个计**1分**。小于**25分**为弱智可疑，小于**15分**则弱智可能性很大。

2. 绘人智能测验 (Draw A Man Test)

- 美国明尼苏达大学古德依纳夫1926年首次提出，后经哈里斯系统研究而成。有中国修订本。
- 该测验适用于4-12岁儿童，测试时要求被试画一个人体全身像，然后根据儿童所画像中头、眼、躯干、下肢、上肢、头发、鼻子、衣着、手、耳、口、足、脸、颈等17个方面进行评分，再将得分按量表规定换算成智商。

3. 皮博迪图片词汇测验 (PPVT-CR)

- 两位邓恩（Lloyd M. Dunn 和Leota M. Dunn）于1959年编制而成。
- 适用于2岁半-18岁儿童。
- 测试时，要求被试在听到主试读出的词汇后从一张卡片上的四幅图片中找出与该词对应的一幅图片来，答对一题计1分，再将总分换算成智商。
- PPVT的中国修订版共有175幅图片，按由易到难的顺序排列。适用于3岁半到9岁。

4. 丹佛发育筛选测验

(Denver Developmental Screening Test, DDST)

- 该测验的中国修订本 (DDST-R) 适用于0-6岁儿童的发育筛选，共104个检查项目
 - 应人能提示儿童对周围环境的应答能力和料理自己生活的能力
 - 应物能提示儿童看的能力和用手抓物与画图等精细动作的能力
 - 言语能提示儿童听声音、说话和理解语言的能力
 - 动作能提示儿童坐、走和跳等粗大动作的能力

5. 50项提问智能测验

- 原为美国儿科学院加州分院使用的“儿童入学准备测验”
- 评估内容涉及运动、常识、记忆、视觉、听觉、思维、语言等方面。
- 适用于4-7岁儿童，每位儿童测试时间仅为15-20分钟。

6. 瑞文测验 (CRT)

- 英国心理学家瑞文 (J. C. Raven) 在1938年编制的，它是一种非文字智力测验，主要测一个人的观察力及清晰思维的能力。
- 瑞文测验有4种版本：标准型、彩色型、高级型和联合型 (Combined Raven's Test, CRT)。

- 目前应用得较多的是联合型，已编制出三套中国常模：
 - 中国城市儿童常模（5-16岁）
 - 中国农村儿童常模（5-14岁）
 - 中国成人常模（17-75岁以上）

- 联合型瑞文测验一共有72题，全部由图形组成。
- 其测试结果是一个粗分智力等级，既可以团体施测，也可以个别施测。
- 它还也适用于沟通有困难的听觉障碍者和言语障碍者，也可以作为不同语种的跨文化研究的一个工具。

二、常用的诊断性智力测验

- 韦克斯勒儿童智力量表第四版（**WISC-4**）
- 斯坦福-比奈智力量表第五版（**SB-5**）
- 考夫曼儿童成套评估测验第二版（**K-ABC II**）
- 麦卡锡幼儿智能量表中国修订版（**MSCA-CR**）
- 希-内学习能力测验（**H-NTLA-CR**）
- 盲人学习能力倾向测验（**BLAT**）
- 团体儿童智力测验（**GITC**）

1、韦克斯勒儿童智力量表第四版 (WISC-4)

- 2003年，韦克斯勒儿童智力量表第四版出版
- 施测范围为6岁-16岁11个月的儿童，青少年。
- 将以往的言语和操作部分分成言语理解，知觉推理，工作记忆和加工速度四大领域。
- 该测验共包括15个分测验，其中10个为必做的分测验。其余5个为补充的分测验。

(1) 言语理解

- 该测验共分为5个类别，分别为：
 - 类同：概括出两个常见事物的共同点；
 - 词汇：受测者给每一张图片命名，并口头解释每一个单词的词义；
 - 理解：受测者回答一系列有关社会常识的问题以推测他的社会成熟度和道德判断水平；
 - 常识：这是一个补充测验，受测者回答一系列普通的常识问题；
 - 单词推理：这也是一个补充测验，主试提供一些线索，受测者猜测主试描述的是什么。

(2) 知觉推理

- 该部分共分为4个测验，其中必测测验3个，补充测验1个。
 - 矩阵推理：给受测者呈现一个矩阵图，图中有一小部分缺失，让其从五个选项中选出来填补这个缺失部分，使之成为一个完整的图形。
 - 积木：要求受测者用六面有全红，全白或者半红半白不同颜色的积木块，拼出和主试呈现的图像一样的图像来。

- 图形概念：给受测者呈现几组印有不同物品的图画，让他从每组图画中选出符合某个概念（如动物）的物品（如小鸟，牛等）。
- 填图：此为补充测验，给受测者呈现一系列有缺失部分的图片，要求他指出哪个部分缺失了。

(3) 工作记忆

- 该部分包含3个测试

- 数字广度：先让受测者将刚听过的一串数字按相同的顺序复述出来，再把刚听过的一串数字按相反的顺序说出来
- 字母-数字排列：给受测者念一串串由不同数字和字母混合编排成的序列，让其先把刚听过的数字依由小到大的顺序重说一遍，再把刚听过的字母依26个英文字母顺序重说一遍。
- 算术：给受测者念一道一道的算术题，让他在规定的时间内说出答案（不许用纸和笔来算）

(4) 加工速度

- 符号搜索：要求受测者用视觉扫描一组组符号，以确定是否出现了符号目标。
- 译码：要求受测者在规定的时间内把一系列与数字配对的几何符号都写出来。
- 删除图形：此为补充测验，先给受测者呈现若干目标图形，然后让他从一组组图形中找到这些图形并划掉。

- 该测验的原始分数要先转换成平均分为10，标准差为3的量表分数，然后再转换成4个因素指数和全量表IQ（平均分=100，标准差=15）。
- 此外，测验的原始分数还可以转换成百分等级。

- 韦克斯勒儿童智力量表第四版是韦氏儿童智力量表的最新版本，理论依据较新，编制技术成熟，信效度较高，因此受到了教育和心理学工作者的普遍欢迎。

2、斯坦福-比奈智力量表第五版 (SB-5)

- 自1905年比奈-西蒙智力量表发表以后，修订版本发表于1916年，成为斯坦福-比奈智力量表，该量表经过了5次修订，成为国际上使用最广泛的智力测验之一。
- 本课程主要介绍洛伊德主持修订的SB-5。

- **SB-5**的适用范围极广，2岁至成人人都可以使用该量表进行测试，完成整套测验时间大约为**45-75**分钟。
- 个别施测的标准化智力测验。
- 洛伊德将整个测验分为言语领域和非言语领域两部分
- 每个部分均包括流体推理、知识（晶体能力）、数量推理、视觉-空间信息加工和工作记忆**5**个因子。

言语领域

- 言语流体推理：让受测者描述图片中正在发生的事情的因果关系或人物之间的互动；做一些初级的推理；听主试说一段话，指出这段话中所包含的谬误；通过类比推理将一句不完整的话说完整。
- 言语知识：让受测者说出玩具的面部及身体特征，看图说出相应的词汇，解释单词的意义。

- 言语数量推理：让受测者数小玩具，小红点；说出数字的名称；做加减法运算和简单的文字题等。
- 言语视觉-空间信息加工：让受测者按照主试的指令放置物品，在听了若干有关方向转变的指令后，说出目前正确的朝向。
- 言语工作记忆：让受测者复述句子，回答主试在提问中说出的最后一个单词是什么。

非言语领域

- 非言语流体推理：给受测者呈现物品序列或矩阵，用手指一下缺少的部分和各个选项，然后指一下装有塑料片，积木和玩具的盒子，让受测者选择一个放在缺少东西的那个位置上。该测试可以确定受测者的能力水平和起测点。
- 非言语知识：用卡片或要求受测者做出某个姿势以表明他明白主试想让他做什么动作；让他仔细观察图片，指出图片中有哪些谬误。

- 非言语数量推理：让受测者选小圆点或积木块，用积木块，图片表示数字概念，序列，解决数学问题。
- 非言语视觉-空间信息加工：给受测者若干零部件，让其拼成完整的图形。
- 非言语工作记忆：把玩具放在某个塑料杯中，改变杯子的位置，让受测者辨认哪个杯子里有玩具。

- 该测验的大多数题目都是按照0或1计分的，但是有一部分难度较大的题目是按照0,1或2计分的。
- 将得到的原始分数按照平均分为10，标准差为3的量表分数进行转换，得到相应的量表分数。

- 根据卡特尔，霍恩和卡洛尔的智力结构理论编制而成的，反映了智力评估领域理论研究的新进展。
- 测验内容分为两个领域，使评估和诊断变得更为全面。
- 在施测程序上，它采用定位测验的方法
- 适用的年龄范围更广

3、考夫曼儿童成套评估测验第二版 (K-ABC II)

- 由考夫曼夫妇发表于1983年的儿童智力量表。
- 2004年，考夫曼夫妇发表了考夫曼儿童成套评估测验第二版
 - 鲁利亚神经心理学理论
 - 卡特尔，霍恩和卡洛尔的智力结构层次模型

- 该针对3岁-18岁11个月大的儿童进行个别施测的标准化智力测验。
- 共包括5个分测验：
 - 继时性加工量表包含了数字记忆，字词顺序，手部动作；
 - 同时性加工量表包含了数积木块，概念性思维，辨认面孔，漫游，三角形，完形闭合；
 - 学习测验包含了亚特兰蒂斯，亚特兰蒂斯回忆，画迷学习，画迷学习回忆；
 - 计划分测验包含了模式推理，完成故事；
 - 知识分测验包含了猜谜语，言语知识，词语表达等测验。

- 根据受测者的年龄，障碍类型以及解释测试结果时所依据的理论模式从上述18个分测验中选择一部分分测验进行测试
- 对学龄前儿童施测所用平均时间为45分钟，对学龄儿童施测所用平均时间为75分钟。

- 每个分测验的原始分数转换成平均数为10，标准差为3的量表分数。
- 将18个分测验的量表分数求和，转换成平均数为100，标准差为15的标准分数，即流体-晶体指数，以及相应的百分等级和年龄当量。
- 尽量减少对文字材料的使用，它非常适用于听觉障碍，言语障碍，情绪障碍，智力障碍，自闭症及学习障碍的儿童。

4、麦卡锡幼儿智能量表中国修订版 (MSCA-CR)

- 由美国儿童发展心理学家D.麦卡锡 (D. McMarthy) 于1972年编制的。
- 该量表主要是对幼儿和儿童的智能和心理发展做综合的测定和评价，其适用年龄为两岁半至8岁半。

- 该量表由18个测验组成，分属于言语（V）、知觉——操作(P)、数量(Q)、记忆(Mem)、运动(Mot)5个分量表。
 - 言语分量表包括图画记忆、词语知识、词语记忆、词语流畅、反义类推5个测验；
 - 知觉——操作分量表包括积木、拼图、连续敲击、左右方向、图形临摹、画人和概括归类7个测验；
 - 数量分量表则由数的问题、数字记忆和记数与数的区分几个分测验组成；

- 记忆分量表的测验分别来自V、P、Q三个分量表，包括图画记忆、连续敲击、词语记忆、数字记忆4个部分；
- 最后的运动分量表包括腿的动作、手臂动作、动作模仿、图形临摹和画人几个测验。
- 对于每个测题都有着它特定的记分方法，对所有的18个测验评出的分为粗分，经过加权以后按照各分量表所包含的测验各自合成分量表的总分，即“合成粗分”，将“合成粗分”经过常模分数转化以后得到各个分量表的“量表分数”和一般智能。

5、希-内学习能力测验 (H-NT-LA)

- 由美国内布拉斯州立大学的希斯基（**Hiskey**）编制的听力障碍儿童专用的智力测验，测出的智商称为学习能力商（**Learning Quotient**）。
- 1994年曲成毅、孙喜斌和张配瑛等人编制出《希-内学习能力测验中国聋人常模修订版》（**H-NTLA-RC**），适用于3-17岁儿童。

- 穿珠：测查儿童手眼协调及伴随记忆能力；
- 记颜色：测查儿童辨色能力及色彩记忆能力；
- 辨认图画：测查儿童知觉辨别与图形比较能力；
- 看图联想：测查儿童认识环境与思维联想能力；
- 折纸：测查儿童手眼协调与伴随记忆能力；
- 短期视觉记忆：测查儿童注意力及短时记忆能力；
- 摆方木：测查儿童空间定向和手眼协调能力；

- 完成图画：测查儿童分析综合和知觉想象能力；
- 记数字：测查数字记忆力；
- 迷方：测查儿童距离知觉与实物测量能力；
- 图画类推：测查儿童类同、比较和概念联想能力；
- 空间推理：测查儿童图形组合与抽象推理能力。

- **H-NTLA-CR**有两套指导语，一套是手势语的，用于聋童；一套是口语的
- 希-内学习能力倾向测验是国内第一套专门为聋童修订的智力测验。该测验已在全国范围内抽样，制定了标准化的聋童常模，并且从测量学性能来看已达到要求，因此适合在我们聋童中使用。

6、盲人学习能力倾向测验 (BLAT)

- 1969年，纽兰为评估盲人的学习能力，或智力而专门设计了一套量表，即盲人学习能力倾向测验。
- 该测验主要测量辨别、概括、序列、类比推理、完成图画，完成矩阵6种能力，测验结果用学习商数表示。

- 这套量表目前已在美国6-20岁的普通儿童中标准化，但它主要用于6-12岁的盲童。
- 该量表以浅浮雕的形式出现，测验中的点和线比盲文读物更容易辨别，除了指导语，所有的材料都是非言语的，也不要求受测者口头回答。

7、团体儿童智力测验 (GITC)

- 由华东师范大学金瑜教授编制的一套评估儿童智力的量表，发表于1996年。
- 该测验适用于9-18岁的中小學生。
- 由语言量表和非语言量表组成，共包括10个分测验
- 其中常识、类同、算术、理解、词汇5个分测验属于语言量表；辩异、排列、空间、译码、拼配5个分测验属于非言语量表。

- 该测验采用纸笔测验的方式，所有的题目均以多项选择题的方式呈现。
- 每个分测验施测的时间规定为6分钟，做完整个测验大约需要1小时20分钟。
- 测验完毕，根据受测者的原始分数和实足年龄就可以通过查常模表确定他的语言IQ，非语言IQ和全量表IQ。

- **GITC**已制定了上海市区常模和全国城市常模
- 研究数据表明，该测验具有较高的可靠性和有效性。