

论文·

认知发展水平诊断工具 IPDT 北京市常模的制订^{*}

方富熹 盖笑松 张丽锦 刘国雄 龚少英

【摘要】目的：制定儿童认知发展水平诊断工具（IPDT）的北京市常模，并对其心理测量学指标进行进一步的检验。方法：以北京市1600名7-14岁儿童为被试，进行了IPDT测验。结果：（1）IPDT测验总分随儿童年龄增长而提高，表现出良好的构念效度；（2）其中的18个子测验难度从易到难呈梯级分布，整个测验的平均难度系数为0.5，难度指标比较理想；（3）测验具有良好的区分度；（4）测验分数的负偏态分布特征表明该测验适合于诊断；（5）根据测试结果建立了北京市7-14岁儿童的百分等级常模以及诊断标准，并对测验的性质、用途和特点进行了讨论。结论：IPDT适合于7-14岁儿童的认知发展诊断。

【关键词】 儿童心理学 认知发展 心理测量学研究 IPDT 认知结构 皮亚杰 常模

Establishment of the Norm of Inventory of Piaget's Developmental Task in Beijing

FANG Fuxi¹, GAI Xiaosong^{1,2} ZHANG Lijin^{1,3} et al.

Key laboratory of Mental health, Psychology Institute in China Academy of Science, Beijing 100101

【Abstract】Objective: To revise the norm of Inventory of Piaget's Developmental Task (IPDT), and to have a further test on the fitness of IPDT. **Method:** 1600 children aged from 7 to 14 in Beijing participated the IPDT tests. **Results:** IPDT scale showed good construct validity, appropriate difficulty and discriminate ability. The norm of IPDT was compiled in the form of percentile rank, and the diagnostic indicators of children with cognitive delay were set up. The feature and function of IPDT were discussed, and suggestions for future research were provided. **Conclusion:** The results demonstrate that IPDT is suitable for demands of psychometrics and applicable in Chinese children.

【Key Words】 child psychology cognitive development psychometric studies IPDT Cognitive schema Piaget Norm

传统的儿童智力测验大多是基于成人智力结构理论的产物，我们在以往的应用中发现，虽然这类智力测验能够对儿童在智力活动中的表现做出有效的区分，但是其测量结果缺乏发展理论的解释，更注重筛选功能而较少诊断功能。

儿童认知发展水平诊断工具（Inventory of Piaget's Developmental Task, IPDT）是建立在皮亚杰认知发展理论基础上的诊断儿童认知结构发展程度的心理测验。皮亚杰认为，儿童的认知发展是通过同化和顺应不断建立和完善认知结构的过程，认知结构是指在活动中形成的一类彼此相似的认知经验构成的相应的认知组织^[1]，各种认知结构的形成和发展是儿童认知发展的一个重要方面，也是儿童成功完成相关认知活动的心理基础。

教育实践和儿童心理学研究都需要探查儿童认知结构发展水平的工具。教学应与儿童认知结构的发展程度相配合，才能取得良好的效果。认知发展落后儿童的矫正训练也需要以认知结构发展缺陷的诊断为

依据。皮亚杰曾经发展出许多精巧的实验任务用于考察儿童的认知结构，例如三山模型、守恒任务等。但是由于这些任务需要临床法操作并需要制作实物材料，给使用者带来了一定的困难。为了方便使用，Furth^[2]将一些经典皮亚杰任务发展为图形化的纸笔测验。Patterson等^[3]进一步将其标准化为儿童认知发展诊断工具（IPDT）。

我们过去对IPDT中文版信效度的考察结果^[4]表明，该测验具有比较理想的信度和效度。本研究的主要目标是建立IPDT测验的北京市常模，为实践运用提供参照标准，同时在大样本基础上对IPDT的特点、难度和区分度进行更详细的分析。

对象与方法

对象 被试的选取采取分层取样方式，从北京市4个区（宣武、朝阳、海淀、丰台）分别选取教育水平中等的中学和小学各2所，从中随机选取7-14岁儿童为被试。各年龄组取值跨度为1周岁，例如7岁

^{*} 本研究为国家科技部攀登计划项目“儿童脑高级功能开发与素质教育若干重要问题的研究”（95-专-09）的部分工作，也是国家自然科学基金会资助重点项目“儿童认知能力的发展与促进”（批准号：39730180）的部分工作。范珍桃和杨利利参加了本测验的修订工作。
作者单位：中国科学院心理研究所心理健康重点实验室（100101） 东北师范大学心理学系（130024） 宁夏大学教育科学学院（750002） 湖南师范大学心理学系（410000）

组取7岁零1天到7岁12个月30天的儿童。每个年龄组选取200名被试，平均分布在4个区，男女各

半。8个年龄组被试总数共1600人。被试的年龄情况见表1。

表1 IPDT北京市常模样本组的年龄分布

年龄组	7岁组	8岁组	9岁组	10岁组	11岁组	12岁组	13岁组	14岁组
平均年龄	7.5	8.5	9.5	10.4	11.5	12.4	13.5	14.4
(标准差)	(0.27)	(0.29)	(0.31)	(0.31)	(0.25)	(0.29)	(0.28)	(0.28)

测验材料 儿童认知发展水平诊断工具 IPDT: 包括5个问题领域，每个问题领域包含3个或4个子测验，共有18个子测验。每个子测验包含4个题目，整个测验一共有72个题目。每个题目有4个选择项，其中一个正确答案，选对得1分，选错得0分。测验总分为0到72分。IPDT的具体结构见表2。

表2 IPDT纸笔测验的结构

子测验编号	子测验名称	问题领域	诊断目标	子测验编号	子测验名称	问题领域	诊断目标
1	数量守恒	守恒	数量守恒	10	排列关系	关系	序列关系
2	水平面表征	表征	表征转换	11	旋转问题	规律	运动表征
3	顺序关系	关系	顺序关系	12	角度问题	规律	相互作用
4	重量守恒	守恒	重量守恒	13	投影表征	表征	观点采择
5	类比推理	分类	矩阵类比	14	类相交	分类	类交集
6	符号表征	表征	符号连接	15	长度守恒	守恒	长度守恒
7	观点表征	表征	观点采择	16	类包含	分类	类包含
8	运动表征	表征	运动表征	17	传递关系	关系	传递推理
9	容积守恒	守恒	容积守恒	18	概率问题	规律	概率判断

施测步骤: 对小学2年级以上被试采取全班团体施测，测验时间45分钟。对小学1年级被试进行小组团体施测（每组7-8人，以方便指导，由主试念指导语），小组团体施测分2次完成，每次30分钟。

结果

1. IPDT测验总分的分布特征分析

对测验总分分布特征的分析既是进一步数据分析的基础，也有利于对测验功能的定位。对IPDT总分进行正态分布检验，采用Lilliefors法对Kolmogorov-Smirnov统计量进行修正，得到统计值 $Z = 0.071$ ($df = 1600$, $P < .001$)，表明IPDT总分的分布与正态分

布之间存在显著差异。进一步的分析发现，IPDT总分分布的偏度系数为 -0.363 ，峰度系数为 -0.603 ，说明数据呈负偏态分布。图1是IPDT总分的数据分布直方图。

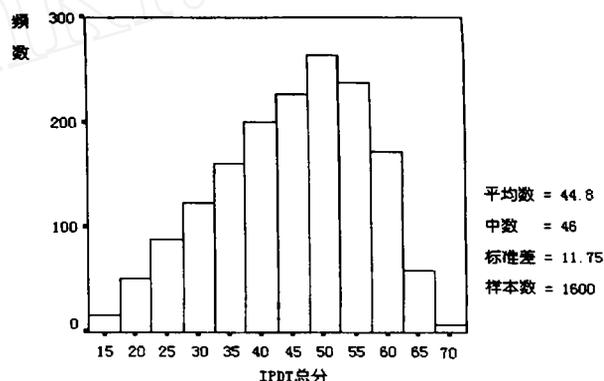


图1 IPDT总分的数据分布直方图

进一步对各年龄段被试的IPDT总分的正态分布检验结果分析发现，除7岁组外，其他年龄组的偏度系数都小于0并表现出与图1相似的负偏态分布特征。

分数呈负偏态分布的测验有利于将成绩较差的学生和中等程度的学生区别开，但不利于将中等程度的学生和成绩优秀的学生区别开。因此IPDT测验更适合于筛选认知发展水平较低儿童，符合诊断用途。

2. 儿童认知水平的发展趋势

由于IPDT总分呈负偏态分布，我们在描述各年龄组的集中趋势时，除了采用平均数以外，还提供了中位数和Huber M估计量。各年龄组的IPDT得分情况见表3。

表3 各年龄组在IPDT总分上的得分情况

年龄	被试数	最低分	最高分	平均数	标准误	标准差	中位数	M估计	偏度系数	峰度系数
7	200	14	55	30.84	0.57	8.02	31	30.67	0.24	-0.30
8	200	14	60	35.66	0.65	9.19	36	35.88	-0.04	-0.32
9	200	16	66	40.91	0.72	10.13	41	41.44	-0.27	-0.26
10	200	19	63	43.97	0.67	9.52	45	44.67	-0.36	-0.42
11	200	21	68	49.73	0.61	8.64	50	50.16	-0.48	0.32
12	200	23	68	50.44	0.59	8.32	51	51.02	-0.58	0.53
13	200	28	68	52.50	0.56	7.98	54	53.48	-0.84	0.65
14	200	30	70	54.25	0.54	7.66	55	54.94	-0.55	0.08

采用多个独立样本的非参数检验 (Kruskal Wallis Test) 对各年龄组总分的差异进行检验，结果发现

IPDT总分的年龄差异显著 ($\chi^2 = 721.49$, $df = 7$, $P < .001$)。进一步采用2个独立样本的非参数检验进行

的两两比较 (Mann - Whitney Test) 发现: 11 岁组和 12 岁组之间 IPDT 总分差异不显著 ($Z = -.836, P > .40$), 其他各组之间皆差异显著 ($P < .05$)。说明 IPDT 测验较好地体现了认知发展水平随年龄增长的趋势。

3. 难度和区分度的分析

表 4 IPDT 18 个子测验的难度系数

子测验名称	类相交	概率问题	容积守恒	旋转问题	距离守恒	角度问题	观点表征	类包含	投影表征	运动表征	传递关系	顺序关系	重量守恒	数量守恒	符号表征	类比推理	排列关系	水平面表征
难度系数	.13	.19	.22	.29	.33	.43	.46	.47	.49	.54	.56	.61	.64	.66	.70	.73	.78	.78

从表 4 看, 高难度的子测验 (难度系数低于 0.3) 有 4 个, 中等难度的子测验 (难度系数介于 0.3 - 0.7 之间) 有 10 个, 低难度的子测验 (难度系数高于 0.7) 有 4 个。全部子测验的平均难度是 0.5。根据心理测验理论, 当测验中的题目难度呈梯度分布且平均难度接近 0.5 时, 该测验潜在的区分效果最优。因此可以认为, IPDT 子测验的难度分布状况比较理想。

子测验的难度分析

因为测验题目是多项选择题形式, 所以采用吉尔福特矫正公式 $CP = (KP - 1) \div (K - 1)$ 计算每个题目的难度系数, 其中 P 代表该题目的正确率, K 代表选项数目 (4 个选项)。子测验的难度系数用子测验中 4 道题目的难度系数之平均数表示, 结果见表 4:

子测验的区分度分析

由于题目采用 0 和 1 计分, 所以通过计算题目与总分之间的点二列相关系数作为每个题目的区分度。然后把各个子测验中 4 个题目的区分度平均数作为该子测验的区分度。表 5 列出了 IPDT 中 18 个子测验的区分度。

表 5 IPDT 子测验的区分度

子测验名称	类相交	概率问题	数量守恒	距离守恒	投影表征	容积守恒	运动表征	旋转问题	水平面表征	类比推理	角度问题	符号表征	顺序关系	重量守恒	排列关系	类包含	传递关系	观点表征
区分度	0.18	0.25	0.27	0.27	0.28	0.31	0.34	0.34	0.38	0.40	0.40	0.43	0.44	0.45	0.45	0.49	0.50	0.51

18 个子测验的区分度介于 0.18 - 0.51 之间, 整个 IPDT 测验的平均区分度为 0.37, 根据 Ebel 提出的经验性评价标准^[5], 该测验整体区分度良好, 但若能对区分度小于 0.30 的子测验进行修改, 区分度将有进一步提高。子测验“类相交”由于难度较高而导致区分度较差。

4. 常模表的制作

百分等级常模的制作

为了使不同年龄儿童的 IPDT 得分具有可比较性, 我们制作了百分等级常模。由于原始分数呈负偏态分布, 我们首先采用非线性变换的方式对各年龄段的数据进行了正态化处理, 在此基础上制作各年龄段的百分等级表, 合并后的百分等级常模见表 6。

表 6 北京市 7 - 14 岁儿童 IPDT 常模表 (示例)

P	年龄组										Z
	7	8	9	10	11	12	13	14	14	14	
99	49	54	60	62	66	66	66	69	69	2.33	
98	48	53	59	61	65	65	65	68	68	2.05	
97	47	53	59	60	64	65	65	67	67	1.83	
...	
3	17	17	20	26	27	28	32	33	33	-1.88	
2	16	16	18	25	25	26	31	31	31	-2.05	
1	16	16	17	20	20	20	28	28	28	-2.33	

表中的数据为原始分, 各年龄组中每个原始分都

有与之对应的百分等级 (P) 和 Z 分数。查得的百分等级即是被试所得的标准分。百分等级的含义是在该年龄组中有百分之多少的儿童得分不高于该被试。

儿童认知发展水平诊断标准的建立

参照正态曲线的分布规律以及智力测验的一般做法, 根据前面的百分等级常模表把常模组被试的 IPDT 得分划分为 A、B、C、D、E 五个等级。其中 C 级包含中间 66% 的常模组被试, B 级和 D 级分别包括 C 级上下 12% 的常模组被试, A 级和 E 级分别包括两端 5% 的常模组被试。各级别的界限及含义见表 7:

表 7 认知发展水平 5 等级分类标准

认知水平发展程度	百分等级 (P) 范围	含义
A 级	P 95	超常
B 级	83 < P < 95	较好
C 级	17 P 83	中等
D 级	5 < P < 17	不足
E 级	P 5	落后

对于认知发展水平处于 D 级和 E 级的儿童, 可以进一步参考子测验常模以确定在哪些认知结构上落后于该年龄组平均水平。

讨论

研究结果表明, 常模组被试在 IPDT 测验上的得

分随年龄增长而显著升高,体现了儿童认知发展的上升趋势。IPDT得分在小学阶段年龄差异较大,中学阶段年龄差异相对较小,符合儿童认知发展先快后慢的规律。说明IPDT测验表现出了较好的构念效度。研究结果还发现,小学四年级左右测验分数标准差达到最大,可能意味着这一时期是儿童认知结构发展过程中分化较激烈的阶段。

18个子测验的难度从易到难呈梯级分布,整个子测验的平均难度为中等(难度系数0.5)。根据经典测量理论,具有这种难度分布的测验潜在的区分能力最高,因此测验的难度比较理想。对测验的区分度分析表明,测验整体区分度良好,其中“概率问题”、“数量守恒”、“距离守恒”和“投影表征”四个子测验的区分度稍低,若对其中的项目进一步调整可以获得更好的区分度指标。子测验“类相交”的区分度较低,这是由于该子测验具有较高的难度($CP=0.13$),在未来的研究中需要对其做进一步的修改。

IPDT的总分呈负偏态分布,这种分布特征是符合理论预期的。由皮亚杰发展出的此类认知发展任务是用于检验儿童是否建立了各种基本的认知结构,而不是考察儿童认知结构的复杂化程度。因此对于特定认知结构处于未成熟水平的被试,成绩表现出逐渐增高的特点。而对于该认知结构达到基本标准的被试,成绩会在高分处集中,不再在测验分数上表现出连续增长趋势。具有负偏态分布特征的测验对于中等及以下发展程度的儿童具有更大的区分能力,因此IPDT测验适合于认知发展不足儿童的诊断。

为了满足实践应用的需要,我们根据百分等级常模建立了认知结构发展程度诊断标准,把认知发展不足儿童和认知发展落后儿童的总数确定在低分端的17%。这样做的理由是:(1)这两部分被试的成绩大

约分别处于平均分的-1个标准差和-1.65个标准差之外,属于明显低分组。(2)我们以前的修订研究表明IPDT测验与智力测验有着很高的相关^[4],因此可以参照智力测验的划分方式。一般智力测验都把低端10%-25%左右的得分划为智力边缘或低下组^[5,6]。(3)参照国内关于学习困难儿童检出率的研究,在校儿童的学习困难检出率在17.4%左右^[7]。基于以上考虑,我们认为应把认知发展不足和认知发展落后儿童的人数总和确定在常模组的17%左右,使之不仅对认知结构发展状况的检测有价值,也对智力较低和学习困难学生的检测有参考作用。这种划分方法还有待于将来在弱智儿童和学习困难儿童等特殊被试样本中的研究来进一步检验。

参考文献

- 1 朱曼殊,许政援,刘范. 心理学论文选集. 北京: 中国经济出版社, 1998, 3: 6-7.
- 2 Furth, H. An Inventory of Piaget's Developmental Tasks. Washington, Dc: Catholic University, Department of Psychology, Center for Research in Thinking and Language. 1970.
- 3 Patterson, H. O., Milakofsky, L. A paper and pencil inventory for the assessment of Piaget's tasks, Applied Psychological Measurement, 1980, 4 (3): 341-353.
- 4 方富熹,盖笑松,龚少英,刘国雄. 对儿童认知发展水平诊断工具IPDT的信度效度检验. 心理学报, 2004, 36 (1): 96-102.
- 5 戴海崎,张峰,陈雪枫. 心理与教育测量. 广州: 暨南大学出版社, 2001. 3: 117-118, 282-282.
- 6 李丹,陈国鹏. 瑞文测验联合型(CRT)中国修订版手册. 华东师范大学心理学系, 1989. 6.
- 7 王耀华等, 儿童学习困难与精神卫生状况: 1787名在校小学生调查. 中国心理卫生杂志, 1989, 8 (4): 160-179.

责任编辑 胜利

(上接第277页)

试认为纸笔方式更具有人文性和被试的特殊身份——心理学系的学生有关。

结 论

1. 电脑测试和纸笔测试的结果一致, 电脑测试可以替代纸笔测试。2. 测试的先后顺序不影响测验结果。3. 被试更倾向于纸笔测试方式。

致谢: 感谢研究生阎春平和柴彩霞同学参与实验资料收集的部分工作。

参考文献

- 1 Masune Sukigara. Equivalence Between Computer and Booklet Administrations of the New Japanese Version of the MMPI. Education and Psy-

chological Measurement, 1996 (4): 570-584.

- 2 查尔斯·杰克逊 [英], 著. 姚萍, 译. 了解心理测验过程. 第1版. 北京: 北京大学出版社, 2000: 9-10.
- 3 Herman Aguinis, William H. Bommer, Charles A. Pierce. Improving the Estimation of Moderating Effects By Using Computer-Administered Questionnaires. Education and Psychological Measurement, 1996 (6): 1043-1047.
- 4 Walter P. Vispoel, Jaeyool Boo, and Timothy Bleiler. Computerized and Paper-and-Pencil Versions of The Roesnberg Self-Esteem Scale: A Comparison of Psychometric Features and Respondent Preferences. Educational and Psychological Measurement, 2001 (3): 461-474.
- 5 龚耀先, 主编. 艾森克个性问卷手册. 第1版. 长沙: 湖南地图出版社, 1992.

责任编辑 唐宏宇