

学习技能诊断测验的初步编制

程灶火¹, 陶金花², 刘新民², 袁国桢¹

(1. 无锡市精神卫生中心, 江苏 无锡 214151; 2. 皖南医学院心理系, 安徽 芜湖 241002)

【摘要】 目的: 编制一套学习技能诊断测验, 并检验其信效度。方法: 以理论法编制了一套用于小学生的学习技能诊断测验, 分低级(1-3 年级)和高级(4-6 年级)两个版本, 每个版本含语文分量表和数学分量表, 语文分量表由汉词拼音、汉词拼写和阅读理解三个分测验组成, 数学分量表由数量概念、四则运算和数学推理三个分测验组成。用新编测验对近 1000 名小学生进行测查, 重测样本 95 人, 效标效度样本 75 人。结果: 低级版的难度为 0.49-0.78, 高级版的难度为 0.52-0.86 之间; 低级版的区分度为 0.29-0.61, 高级版的区分度为 0.39-0.60。低级版重测信度为 0.73-0.90, 分半信度为 0.66-0.95, 系数为 0.70-0.97, 高级版重测信度在 0.73-0.89, 分半信度为 0.62-0.89, 系数为 0.79-0.95。分测验之间的相关系数为 0.36-0.97, 分测验与总分的相关系数在 0.71-0.95。分测验分数与华文认知能力量表智商的相关为 0.55-0.81, 与学校考试成绩的相关为 0.19-0.74。结论: 学习技能诊断测验的难度、区分度、信度和效度符合心理测量学的标准。

【关键词】 学习技能诊断测验; 小学生; 信度; 效度

中图分类号: R395.1

文献标识码: A

文章编号: 1005-3611(2007)05-0447-05

Preliminary Development of Academic Skill Diagnostic Test

CHENG Zao-huo, TAO Jin-hua, LIU Xin-min, YUAN Guo-zhen

Wuxi Mental Health Center, Wuxi 214151, China

【Abstract】 Objective: To develop a Academic Skill Diagnostic Test and examine it's reliability and validity. Methods: Academic Skill Diagnostic Test(ASDT) was developed by rational approach for primary school students, dividing into low-grade edition for 1-3 grade students and high-grade edition for 4-6 grade students. Every edition had two subscales: Chinese subscale consisted of phonetics, spelling, and reading comprehension, Mathematic subscale included quantitative conception, computation, and application. About 1000 students were administered the ASDT, among whom 95 students were re-tested after two months, and 75 students received Chinese Cognitive Ability Scale. Results: The difficulty and discrimination of the ASDT were 0.49-0.78 and 0.29-0.61 for low-grade edition, and 0.52-0.86 and 0.39-0.60 for high-grade edition. The stability coefficients, split-half reliability, coefficients were 0.73-0.90, 0.66-0.95, and 0.70-0.97 for low-grade edition, 0.73-0.89, 0.62-0.89, and 0.79-0.95 for high-grade edition. The correlations between subtests ranged from 0.39 to 0.97, and the correlations between subtests and total scores were 0.71-0.95. The correlations between subtest scores and IQs were 0.55-0.81, and the correlations between subtest scores and school examination scores were 0.19-0.74. Conclusion: The stability, internal consistency, and validity of the ASDT are good and meet with psychometric standard.

【Key words】 Academic Skill Diagnostic Test; Primary school student; Reliability; Validity

儿童学习障碍已成为家长、教师和社会关注的问题, 据估计有 10% 以上小学生存在不同类型和不同程度的学习障碍, 国内一些学者对学习障碍儿童的智力、记忆和神经心理特征等做了大量研究^[1-3], 但学习障碍的规范化诊断问题还没有解决。国外主要依据学业成就与智力潜能分离模式(即学业成就低于智力潜能 2 个标准差以上)或学业成就与年龄分离模式(即学业成就比同龄人低两个年级以上)。在成就-潜能分离诊断模式中, 标准化智力测验和标准化成就测验是必备的诊断工具, 多数国家都有政府或学术团体推荐的智力测验和成就测验。在美国一般用 Wechsler 儿童智力量表第三版(WISC-

)和 Wechsler 个别成就测验(WIAT), 因为它们有共同常模, 而且是最新的标准化测验^[4]。国内虽有标准化智力测验(如中国修订韦氏儿童智力量表), 但没有与智力测验配套的标准化成就测验或学习技能诊断测验, 所以在学习障碍的诊断和干预存在诸多困难, 阻碍了该领域的发展。要了解学习障碍是由哪种学习技能或认知过程缺陷引起的, 迫切需要用学习技能诊断测验, 国外有许多这样的测验, 但由于社会文化和教育背景的差异, 无法直接用于我国儿童学习障碍的研究和临床评估。国内近些年来也有学者编制了一些基本数学技能测验、阅读诊断测验和学业成就测验^[5-8], 但还没有成套学习技能诊断测验。基于这些原因, 我们于 2004 年开始着手编制《学

习技能诊断测验》(Academic Skill Diagnostic Test, 简称 ASDT), 试图测量小学生语文和数学方面的基本技能, 以量化分数评价基本技能的掌握程度, 通过错误类型分析了解学习技能缺陷的特征, 结合这两方面的指标对学习障碍做出定量和定性诊断。本文报告测验的编制情况和初步的信效度检验结果。

1 学习技能诊断测验的编制

1.1 测验编制的思考

学习技能诊断测验主要评估基本技能的掌握情况和存在哪些技能缺陷, 用于学习障碍的诊断、教育安置和干预指导。不同研究者编制的学习技能诊断测验所涵盖的内容有一些差异, 但所有成套测验都包含语文和数学相关的技能, 如拼音、拼写和阅读等语文技能, 数量关系、四则运算和数学推理等数学技能^[4]。因为这是小学生必须掌握的最基本技能, 不仅对语文和数学学习是必需的, 对其它学科的学习也是必要的, 所以我们把这些技能作为测验的基本内容, 分别构成语文分量表和数学分量表。

本测验试图用于小学儿童, 但从 1 年级到 6 年级涉及到知识和技能的跨度较大, 国外对不同的年级采用不同的版本, 我们拟编制两个版本: 低级版用于 1~3 年级学生, 高级版用于 4~6 年级学生。国外有些测验的不同版本间的分测验形式是不同的, 我们力求使两个版本的分测验保持一致, 只在条目难易程度做些改变, 而且高低两个版本的内容有部分重复, 使两个版本能更好地衔接。

测验条目力求反映全国统编教材要求掌握的基本技能, 同时也兼顾到地方教材的某些特殊要求。根据测验目标制定项目取样计划表, 确定不同技能的项目数, 每套测验项目量估计在 1 小时左右能完成。条目内容是教材为依据, 又不完全拘泥于教材, 但反映的技能和难度与教材相匹配。测验实施时既可采用个别进行, 也可用团体实施的纸笔形式。语文分量表包括拼音、拼写和阅读三个分测验, 数学分量表包括数量概念、运算和数量推理三个分测验。

1.2 测验内容、功能和记分

1.2.1 汉语拼音 36 个条目(给双字词注音)。我们在考察拼音时主要从以下几个方面进行测试: 反复考察常见声母和韵母的正确拼写, 同时尽可能涵盖一些生僻的声母和韵母。常用拼音的拼写规则。

翘舌音与平舌音的区分。特殊拼音的组合规律, 如 j、q、x 和 ü 相拼, ü 上两点省略。前后鼻音的区别。n、l 的区别。对相似的易混淆的声母的区别,

如 b 与 d, p 与 q。掌握韵母带调的方法, 准确的标注声调。记分方法: 一个词的拼音全对是 3 分, 全错是 0 分, 一个字声母对了得 0.5 分, 韵母对了得 0.5 分, 声调对了得 0.5 分, 最后累计整个词的拼音得分。

1.2.2 汉语拼写 36 个条目(根据拼音写出双字词)。在拼写测验中, 我们考察最多的是形声字, 因为形声字在汉字中约占七成。形声字总是分两部分, 一部分表意, 一部分表音。我们考察形声字尽量覆盖了以下类型: 左形右声, 右形左声, 上形下声, 下形上声, 内形外声、外形内声、形占一角、声占一角等, 还有一些特殊结构的字。通过测试我们可以了解学生对生字量的掌握情况, 对字体结构的正确书写能力, 区分形近字细微的差别能力, 对音同意不同的字的辨别能力等等。记分方法: 一个词拼写全部正确得 2 分, 一个字拼写正确得 1 分, 全错是 0 分。

1.2.3 阅读理解 六篇互相独立的短文组成, 每篇文章有五个条目, 共有 30 个条目。选择条目时主要从以下内容进行考察: 考察学生是否拥有丰富, 广泛的词汇量, 对音-形字的规范读法及书写。词语的理解及运用: 语言理解能力的考查, 要着重考查学生联系语言环境理解词句的能力, 废除孤立地解释词句意思。语言运用能力的考查, 要重点检查课文中出现的词语、句式能否灵活运用。句子的理解: 主要考察对短文的理解能力和概括能力, 对语言材料作出清晰而严密的逻辑推理能力。层次段落的划分及概括段意, 考察对短文结构的把握和概括能力。

归纳中心思想或拟小标题。考察对短文进行准确完整地概括能力。标点符号的一般和特殊用法, 含义。关联词的恰当运用。修辞知识的掌握。记分方法: 每个条目基本按 0、1、2 记分, 完全正确得 2 分, 全错是 0 分, 意思相近或者部分正确得 1 分。

1.2.4 数量概念 20 个条目。考察的内容主要有: 数的大小比较、认识钱币和时间、读数、数量关系的转换等。到了高年级, 随着数量知识的加深, 我们考察的内容更加具体化如: 分数与假分数、约数与倍数、质因素、小数位数数保留和百分数与小数的转化等。记分方法: 每个题目按 0、1、2 记分, 完全正确得 2 分, 全错是 0 分, 答案不完整或有小错误得 1 分。

1.2.5 四则运算 30 个条目。我们考察的基本运算能力有: 一位数和两位数的加减运算, 20 以内的加减运算(不进位、进位和不退位、退位), 20 以上 100 以内加减运算(不进位、不退位和进位、退位), 100 以内连加、连减、混合运算(不进位、不借位和进位、

借位), 2-9 以内简单乘法除法运算, 2-9 以内口诀三位连乘和连除运算, 2-9 以内口诀乘除混合运算, 100 以内除法有余数的运算, 100 以上万以内的加、减、混合运算(不进位、不借位和进位、借位), 一个因素是一位数的乘法运算, 除数是一位数的除法运算, 一个因素是两位数的乘法运算(不进位、进位), 除数是两位数的除法, 加、减、乘、除混合运算, 亿以内的加法、减法运算, 加、减或乘、除简易方程, 加、减或乘、除简便计算, 同分母的加、减运算, 小数的加、减、乘、除和混合运算, 分数和小数的加、减、混合运算, 分数的四则混合运算, 比例式的运算。记分方法: 主要按 0、1、2 记分, 完全正确得 2 分, 全错是 0 分, 运算法则正确得 1 分。

1.2.6 数量推理 12 个条目。我们选择了一些典型的问题进行考察: 行程应用题、工程应用题、相遇应用题、正、反比例应用题, 按比例分配应用题、以及平面图形和立体图形的求积应用题等。由于数量推理是两种能力的综合, 所以我们要分开记分, 每一个正确的关系式得 1 分, 相应的演算正确得 1 分, 错误记零分, 有多步计算的最后累计得分。

2 研究对象与方法

2.1 研究样本

2.1.1 基本样本 语文技能测验低级版样本 465 人(男 247 人, 占 53%, 女 218 人, 占 47%); 语文技能测验高级版样本 442 人(男 219 人, 占 49%, 女 223 人, 占 51%); 数学技能测验低级版样本 463 人(男 237 人, 占 51%, 女 226 人, 占 49%); 数学技能测验高级版样本 450 人(男 229 人, 占 51%, 女 221 人, 占 49%)。样本的年级分布、性别比例和平均年龄见表 1, 父母文化程度以初中、高中和中专为主体, 父母职业以劳动和商务居多, 管理、技术类和其它职业均涉及, 与我国的职业分布基本相符。

表 1 各年级的样本数、性别比例和平均年龄

年级	语文技能测验		数学技能测验	
	例数(男/女)	平均年龄	例数(男/女)	平均年龄
一年级	151(79/72)	7.2±0.44	158(80/78)	7.23±0.45
二年级	152(80/72)	8.19±0.33	148(74/74)	8.17±0.33
三年级	162(88/74)	9.25±0.45	157(83/74)	9.26±0.46
四年级	122(57/65)	10.21±0.42	122(59/63)	10.23±0.23
五年级	154(71/83)	11.22±0.48	174(89/85)	11.24±0.45
六年级	166(91/75)	12.24±0.49	154(81/73)	12.26±0.58

2.1.2 重测信度样本 语文技能测验低级版 47 人(男 24 人, 女 23 人), 平均年龄 8.19±0.18; 数学技能测验低级版 48 人(男 22 人, 女 26 人), 平均年龄 8.16±0.12; 语文技能测验高级版 54 人(23 人, 女 31

人), 平均年龄 11.09±0.94; 数学技能测验高级版 51 人(24 人, 女 27 人), 平均年龄 11.05±0.93。

2.1.3 效标样本 在基本样本中有部分学生同时接受了《华文认知能力量表》测试, 共计 75 人。一年级 15 人(男 7 人, 女 8 人), 平均年龄 6.53±0.74; 二年级 14 人(男 5 人, 女 9 人), 平均年龄 7.57±0.65; 三年级 14 人(男 8 人, 女 6 人), 平均年龄 8.51±0.78; 四年级 12 人(男 5 人, 女 7 人), 平均年龄 9.77±0.47; 五年级 10 人(男 5 人, 女 5 人), 平均年龄 10.87±0.70; 六年级 10 人(男 4 人, 女 6 人), 平均年龄 11.62±0.54。

2.2 研究方法

2.2.1 取样方法 采取分层按比例整群抽样原则在无锡市和芜湖市抽取了 2 所小学, 在每个学校的每个年级各选 2 个班, 共 24 个班, 排除患有严重躯体疾病、精神发育迟滞和有注意缺陷的儿童后, 获得有效问卷 1820 份。间隔二个月后在芜湖的王家巷小学的每个年级共抽取了 110 名学生作了重测。实测的时间为春季开始的第二学期, 2005 年 5 月~6 月。

2.2.2 学习技能测试 采用新编的《学习技能诊断测验》对入选对象做团体测试。测验时, 以班级为单位, 单双列交替分做语文和数学测验, 以避免受试者间的相互干扰。主试说明测验的目的和意义、读指导语和注意事项, 并解答被试提出的有关问题, 要求他们尽可能在 1 小时内完成, 年级老师维持课堂秩序。

2.2.3 效标资料的收集 采用华文认知能力量表^[9]对入选对象做个别测试, 计算被试的智力水平, 语文和数学成绩取第二学期期中考试成绩。

3 结 果

3.1 各年级测验得分比较

学习技能诊断测验低级版: 语文和数学分测验都是三年级得分较高, 一年级得分较低。学习技能诊断测验高级版: 语文和数学分测验都是六年级得分较高, 四年级得分较低。从结果可以看出, 各分测验得分随年级增长而升高的规律比较明显。见表 2。

3.2 测验的难度和区分度

表 3 中测验条目的难度为被试在每个条目上的平均得分率, 数值越大难度越低; 区分度为被试条目得分与分测验总分的相关。从表中可以看出, 无论是低级版还是高级版, 各分测验的难度随年级提升而降低, 低级版各分测验的平均难度在 0.49~0.78 之间, 高级版各分测验的平均难度在 0.52~0.86 之间; 各分测验在不同年级都有一定区分度, 低级版各分测验的平均区分度在 0.029~0.61 之间, 高级版各分

测验的平均区分度在 0.39~0.60 之间。

3.3 信度分析

本研究考查的信度指标包括重测信度、分半信度和 α 系数。无论是低级版还是高级版总分的重测信度都在 0.80 以上, 总分分半信度在 0.81 以上, 总分的 α 系数在 0.93 以上。详见表 4。

3.4 分测验间及与总分的相关

无论是语文测验还是数学测验, 其包含的分测验之间都有中等度以上的相关。分测验与总分呈高度相关。详见表 5。

表 2 各年级测验得分的比较

	低级版				高级版			
	一年级	二年级	三年级	F值	四年级	五年级	六年级	F值
汉语拼音	63.88	93.77	94.78	270.44	87.84	92.18	96.58	29.05
汉字拼写	24.30	44.98	57.69	585.04	47.07	57.71	62.39	74.97
阅读理解	19.01	29.35	39.04	229.08	30.20	35.13	39.66	42.12
数量概念	18.85	22.33	27.70	285.23	16.39	30.28	35.73	848.90
四则运算	21.38	33.22	45.08	514.92	30.96	40.55	51.25	279.78
数学推理	5.38	9.50	14.42	553.48	9.17	12.73	18.99	234.46
数学演算	5.06	8.84	13.48	509.14	7.30	10.93	15.88	192.47
语文总分	107.0	168.10	191.51	551.21	165.1	185.01	198.63	77.69
数学总分	50.62	73.89	100.68	819.32	63.83	94.43	121.8	597.33

注: 标准差省略, 所有 F 值的 P 值 < 0.01。

表 3 分测验条目的平均难度/区分度

	数量概念	四则运算	数学推理	数学演算	汉语拼音	汉字拼写	阅读理解
一年级	0.47/0.28	0.36/0.37	0.41/0.38	0.39/0.38	0.59/0.57	0.34/0.37	0.32/0.38
二年级	0.56/0.27	0.56/0.38	0.65/0.38	0.61/0.34	0.86/0.42	0.63/0.42	0.49/0.43
三年级	0.69/0.29	0.75/0.38	0.84/0.38	0.78/0.48	0.88/0.50	0.80/0.40	0.64/0.37
低级版	0.57/0.36	0.55/0.48	0.63/0.58	0.59/0.50	0.78/0.61	0.59/0.56	0.49/0.49
四年级	0.43/0.24	0.52/0.40	0.37/0.38	0.27/0.42	0.81/0.51	0.66/0.56	0.45/0.40
五年级	0.78/0.35	0.7/0.38	0.52/0.46	0.40/0.41	0.85/0.43	0.80/0.39	0.51/0.34
六年级	0.92/0.36	0.86/0.33	0.76/0.55	0.66/0.52	0.89/0.45	0.87/0.42	0.58/0.33
高级版	0.73/0.52	0.69/0.50	0.56/0.60	0.47/0.54	0.86/0.47	0.78/0.51	0.52/0.39

表 4 测验同质信度和重测信度

	α 系数		分半信度		重测信度	
	低级版	高级版	低级版	高级版	低级版	高级版
汉语拼音	0.96	0.89	0.95	0.89	0.76	0.76
汉字拼写	0.95	0.91	0.94	0.90	0.89	0.83
阅读理解	0.91	0.79	0.82	0.66	0.73	0.73
数量概念	0.70	0.88	0.67	0.87	0.74	0.89
四则运算	0.91	0.90	0.71	0.62	0.80	0.77
数学推理	0.81	0.82	0.66	0.67	0.90	0.75
数学演算	0.78	0.78	0.66	0.65	0.79	0.70
语文总分	0.97	0.93	0.86	0.81	0.83	0.80
数学总分	0.95	0.95	0.90	0.86	0.91	0.89

注: 所有相关系数的 P 值 < 0.001。

表 5 分测验之间的相关及分测验与总分的相关

	汉语拼音	汉字拼写	阅读理解	语文总分	数量概念	四则运算	数学推理	数学演算	数学总分
汉语拼音		0.67	0.36	0.83					
汉字拼写	0.73		0.45	0.89					
阅读理解	0.57	0.79		0.71					
语文总分	0.90	0.94	0.84						
数量概念						0.73	0.69	0.67	0.89
四则运算						0.72	0.68	0.67	0.92
数学推理						0.74	0.80	0.93	0.88
数学演算						0.74	0.81	0.97	0.87
数学总分						0.85	0.95	0.92	0.93

注: 下三角为低级版, 上三角为高级版

3.5 与华文认知能力量表的相关

华文认知能力量表与语文测验低级版的相关系数为 0.63(总分)、0.63(拼音)、0.60(拼写)和 0.55(阅读); 与语文测验高级版的相关系数为 0.71(总分)、0.81(拼音)、0.64(拼写)和 0.42(阅读)。华文认知能力量表与数学测验低级版的相关为 0.68(总分)、

0.62(概念)、0.66(运算)、0.62(推理)和 0.66(演算); 与数学测验低级版的相关为 0.64(总分)、0.56(概念)、0.64(运算)、0.67(推理)和 0.63(演算)。

3.6 与考试成绩的相关

语文分量表与各年级学生期中考试成绩的相关系数在 0.48~0.74 之间, 语文分量表各分测验与各年级学生期中考试成绩的相关系数在 0.37~0.72 之间, 多数相关值在 0.45 以上。数学分量表与不同年级学生期中考试成绩的相关系数在 0.30~0.57 之间, 数学分量表各分测验与不同年级学生期中考试成绩的相关系数在 0.19~0.58 之间, 多数相关值在 0.30 以上。

4 讨 论

4.1 学习技能诊断测验的结构

国外儿童学习障碍诊断主要依据智力潜能与学业成就的分离模式^[10], 即学业成就明显低于智力潜能, 但要分析学习障碍产生的原因、进行有效的干预, 就需要用学习技能诊断测验确定特殊的技能缺陷。针对这种特殊需要, 本研究编制了学习技能诊断测验, 结果分析采用两套评分系统: 定量分析和定性分析, 定量分析统计各分测验的正确得分(最后转化成标准分), 主要评估学生对基本学习技能的掌握程度; 定性分析统计各种错误类型的数量(最后确定各种错误类型的划界分), 主要评估学习技能缺陷的性质。

国外有许多成就测验^[1],如 Woodcock-Johnson 心理教育成套测验, Kaufmann 教育成套测验, Wechsler 个别成就测验, Metropolitan 成就测验;也有针对特殊技能的诊断测验,如 Stanford 诊断性数学测验, Gray 默读测验, Woodcock 阅读掌握测验-修订本。儿童学习涉及多方面的技能,多数测验只测量基本技能,如语音意识、单词拼写、组词造句、阅读理解、写作、数量知识、四则运算、数量推理等。本研究编制的学习技能诊断测验仅涵盖小学生最基本学习技能,即拼音、拼写、阅读理解、数量概念、四则运算、数量推理等技能,因为多数儿童学习障碍是由这些技能缺陷引起的。

4.2 测验项目的难度和区分度

统计结果显示:低级版各分测验的平均难度在 0.49-0.78 之间,高级版各分测验的平均难度在 0.47-0.86 之间,总体而言,多数分测验的平均难度都在 0.5 左右,尽管个别分测验偏易,但对诊断性测验来说还是可以接受的,其一,本测验是考查儿童的基本学习技能是否到达最低要求,而不是测量他的最大潜能,其二,本研究所取的样本绝大多数是学习正常的儿童,对学习障碍儿童来说还是有难度的。其三,无论是低级版还是高级版都包含 4 个年级以上的内容,年级跨度较大,对每个年级而言只有 1/4 的项目难度是适宜的,其它项目对低年级偏难、对高年级偏易。这些与我们的编制构想是符合的,与国内外同类测验的研究结果是一致的^[4,6,7]。

统计结果显示:低级版各分测验的平均区分度在 0.29-0.61 之间,高级版各分测验的平均难度在 0.39-0.60 之间,不同年级之间区分度的差别不大。Ebel 认为区分度在 0.20 以上的条目即达到可接受水平,0.30-0.39 为良好,0.40 以上为很好, Aiken 也认为区分度指数在 0.30 以上的条目较为理想,稍低于 0.30 也可接受^[12]。按这种标准来衡量,本研究编制的各分测验的区分能力是比较理想的。从条目的区分能力来看,不同条目对不同年级的区分能力是不同的,对低年级有鉴别力的条目对高年级不一定有鉴别力,但每个分测验在每个年级大约有 1/4 条目的鉴别能力是比较理想的。有些分测验在低年级的难度和区分度比较理想,如拼音,有些分测验在高年级的难度和区分度比较理想,如数量推理,阅读理解分测验在各年级的难度和区分度比较理想,这与不同技能获得的时期有关。

4.3 测验的信度和效度

统计结果显示:总分的重测信度都在 0.80 以

上,低级版分测验重测信度在 0.73-0.90 之间,高级版分测验的重测信度在 0.73-0.89 之间;总分的分半信度在 0.81 以上,分测验的分半信度在 0.62-0.95 之间;总分的系数在 0.93 以上,分测验的系数在 0.70-0.96 之间。按心理测量学标准,学习技能诊断测验的信度是可以接受的,与国内同类测验相比互有高低。本测验同质性信度和重测信度与国内外同类测验相当^[5-8,12],如 Stanford 成就测验的系数绝大部分在 0.80 以上,个别也有低于 0.80 的,广泛成就测验阅读和拼写分测验的分半信度在 0.92-0.98 之间,数学分测验在 0.85-0.92 之间。

本研究从三个方面论证了学习技能诊断测验的效度。作为学习技能诊断测验,它所测量的技能是明确的、也是比较容易定界的,每种技能都有对应教学内容,我们在编制时与教师做了深入的讨论,并以全国统编教材为依据,同时每个分测验同质信度也是比较高的,提示测验具有较好的内容效度。学习技能诊断测验考查的是各种学习技能,而不是考查某种特殊的心理结构,但作为正常学习儿童各种技能是协调发展,所以不同分测验之间会有一定的相关。本研究显示,语文各分测验之间以及与测验总分相关系数分别为 0.57-0.94 之间(低级版),0.36-0.89 之间(高级版);数学各分测验之间以及与测验总分相关系数分别为 0.72-0.97 之间(低级版),0.67-0.93 之间(高级版)。这些结果说明不同分测验测量的技能都有关,同时又有其各自的独特性。一个新编测验是否真正有效,仅凭内容效度和结构效度的证据是不够的,作为诊断测验更注重实证效度。学习障碍诊断的主要依据是学业成就与智力潜能分离,即正常儿童的学习技能水平与智力发展有较高的相关。本研究显示分测验与智力的相关在 0.42-0.81 之间,总分与智力的相关在 0.63-0.71 之间,这与韦氏智力测验与标准化成就测验之间的相关系数在 0.5-0.6 之间基本一致^[13],说明所编制的测验有一定的效度。另外我们考查了学习技能测验成绩与学校考试成绩的相关,结果显示语文测验与语文考试成绩的相关在 0.48-0.74 之间,数学测验与数学考试成绩的相关在 0.30-0.57 之间,国外标准化成就测验与学校成绩的相关一般也在 0.5-0.6 左右^[14],也说明本测验对学习成绩有一定的预测作用。

参 考 文 献

- 1 程灶火,解亚宁.学习困难儿童的神经心理学亚型分析.中国临床心理学杂志,1993,1(2):90-92

信念这一因子没有像预期那样聚合在一起,而是聚合成两个因子,根据这两个因子包含项目的涵义把这两个因子分别命名为“错误职业信念”和“错误决策信念”,这种结构的不一致可能说明职业决策困难问卷并没有全面反映大学生的职业决策困难^[1],但是具体原因还需要进一步研究。

3.2 信效度分析

本文考察了同质性信度,结果表明问卷有较好的信度,对于复测信度需要进一步验证。本文对问卷的结构效度、辨别效度和效标效度进行了考察:采用相关矩阵考察了结构效度,除了缺乏动机和错误职业信念两个因子之间相关不显著外,其余的相关系数均达到显著水平,这就说明因子间是一个有机的整体。各因子间相关系数在 0.004-0.62 之间,其中只有缺乏动机和错误职业信念、缺乏动机和犹豫不决两组因子间的相关系数低于 0.1。各分问卷与其组成因子之间的相关都比较高,而与其他分问卷的组成因子之间有着中等程度的相关,各因子与总问卷的相关在 0.42-0.79,各分问卷与总问卷之间的相关系数在 0.67 以上,这就表明各因子既有一定的独立性,又反映出相应的归属性,因此问卷具有较好的结构效度。与未接受过职业指导的大学生相比,接受过职业指导的大学生更少职业决策困难。本研究将大学生分为职业未定(undecided)和职业已定(decided)两类。结果表明两类大学生职业决策困难存在明显差异。表明大学生职业决策问卷具有较好的鉴别效度。

参 考 文 献

1 Campbell RE, Cellini JV. A diagnostic taxonomy of adult

(上接第 451 页)

- 2 CHENG Zao-huo, DAI Xiao-yang, LIU Shao-wen. The Patterns of Intellectual Functioning in Chinese Children with Learning Disability. Bulletin of the Hong Kong Psychological Society, 1994, 32/33:60-71
- 3 程灶火, 龚耀先. 学习障碍儿童记忆的比较研究: . 学习障碍儿童记忆功能的综合分析. 中国临床心理学杂志, 1999, 7(2): 89-93
- 4 程灶火. 学业成就测验. 见: 龚耀先. 心理评估. 北京: 高等教育出版社, 2003
- 5 杨志伟, 龚耀先. 汉语阅读技能诊断测验. 中国临床心理学杂志, 1997, 5(3):158-163
- 6 吴汉荣, 刘文龙, 武丽杰. 黑龙江省小学生数学能力测试量表及常模的制定. 中国校医, 2005, 19(1):1-3
- 7 范晓玲, 龚耀先. 4-6 年级多重成就测验的编制. 中国临床心理学杂志, 2005, 13(13):253-257
- 8 范晓玲, 龚耀先. 4-6 年级多重成就测验的编制: 信度考验. 中国临床心理学杂志, 2005, 13(13): 253-257

career problems. Journal of Vocational Behavior, 19: 175-190

- 2 Gati I, Saka N. High school student's career-related decision-making difficulties. Journal of Counseling and Development, 1996, 79: 331-340
- 3 Yahya H, Moshe T, Itamar G. Career Decision-Making Difficulties Among Israeli and Palestinian Arab High-School Seniors. Professional School Counseling, 2004, 8: 64-72
- 4 Mau WCJ. Cultural Dimensions of Career Decision-Making Difficulties. The Career Development Quarterly, 53: 67-77
- 5 Gati I, Krausz M, Osipow SH. A taxonomy of difficulties in career decision making. Journal of Counseling Psychology, 1996, 43(4): 510-526
- 6 Osipow SH, Gati I. Construct and concurrent validity of the career decision-making difficulties questionnaire. Journal of Career Assessment, 1998, 6(3): 347-364
- 7 Kleiman T, Gati I. Dysfunctional Thinking and Difficulties in Career Decision Making. Journal of Career Assessment, 2004, 12(3): 312-331
- 8 Albion MJ, Fogarty GJ. Factors influencing career decision making in adolescents and adults. Journal of Career Assessment, 2002, 10(1): 91-126
- 9 Gati I, Osipow SH, Krausz M, Saka N. Validity of the career decision-making difficulties questionnaire: Counselor versus career counselor perceptions. Journal of Vocational Behavior, 2000, 56(1): 99-113
- 10 Gati I, Saka N. High school student's career-related decision-making difficulties. Journal of Counseling and Development, 79: 331-340
- 11 杜睿, 龙立荣. 大学生职业决策困难问卷的初步研究. 中国临床心理学杂志, 2006, 14(3): 237-239

(收稿日期:2007-03-19)

- 9 程灶火, 孙金荣. 华文认知能力量表的理论构思. 中国临床心理学杂志, 2006, 14(4):340-342
- 10 American Psychiatric Association. Diagnostic and Statistical Manual of Mental disorders, the Fourth Edition(DSM-). Washington:APA, 1994
- 11 Aiken LR. Psychological Testing and Assessment. Allyn and Bacon, Inc, 1988
- 12 吴汉荣, 李丽. 小学生数学能力测试量表的编制和信效度检验. 中国公共卫生, 2005, 21(4):473-475
- 13 Hyeon JO, Joseph JG, Marley WW. Correct Interpretation of Latent Versus Observed Abilities: Implicating From Structural Equation Modeling Applied to the WISC- and WIAT Link Sampling, The Journal of Special Education, 2004, 38(3):159-175
- 14 William LD, Gerald H, Tim A. Coaching Effect on California Achievement Test Scores in Elementary Grade. Journal of Educational Research, 1987, 80(3):149-155

(收稿日期:2007-05-29)