

中学生学业能力倾向测验的效度研究

张月娟¹, 龚耀先²

(1. 河北师范大学教育科学学院, 河北 石家庄 050091; 2. 中南大学湘雅二医院, 湖南 长沙 410011)

【摘要】 目的: 对中学生学业能力倾向测验进行效度研究。方法: 确定效度检验指标, 采集数据, 使用 SPSS11.0 和 Amos4.0 软件对所得数据进行分析。结果: (1) 探索性因素分析表明各分测验对基本能力因子均有中度以上的负荷, 抽取的三个因子解释了总方差的 65.4% (A 式) 和 66.2% (B 式)。验证性因素分析的拟合指数 AGFI、CFI、NNFI 均在 0.96 以上, RMSEA 小于 0.05, χ^2/df 小于 3。(2) 测验总分与长-鞍团体智力测验总分的相关为 0.73 (A 式) 和 0.52 (B 式), 与教师评定的相关在 0.51~0.67。(3) 测验分数在不同年级之间、不同类学校之间、文理科班之间、男女生之间存在显著差异 ($P < 0.05$)。(4) 学习成绩的回归分析可见, 语文及数理分量表进入考试总成绩的回归方程, β 为 0.27、0.49 (A 式), 0.39、0.23 (B 式), 回归方程的决定系数为 0.46 (A 式) 和 0.29 (B 式)。结论: 中学生学业能力倾向测验具有理想的结构效度和较好的效标关联效度, 适用于中学生学业能力倾向的评量。

【关键词】 中学生学业能力倾向测验; 结构效度; 效标关联效度

中图分类号: G449.5

文献标识码: A

文章编号: 1005-3611(2004)03-0221-05

Validity of the Academic Aptitude Test for Middle School Student

ZHANG Yue-juan, GONG Yao-xian

Clinical Psychological Center, The Second Xiangya Hospital, Central South University, Changsha 410011, China

【Abstract】 Objective: To test the validity of a new instrument, the Academic Aptitude Test for Middle School Student (AATMS). **Methods:** Data from 1242 middle school students were analysed by SPSS11.0 and Amos4.0 program. **Results:** (1) Three factors were obtained in exploratory factor analyses, accounting for over 60% of the total variance. In confirmatory factor analyses, the fit indexes AGFI, CFI, NNFI were over 0.96, RMSEA was below 0.05, and χ^2/df was below 3. (2) Correlations between AATMS and CAITG (0.52~0.73), teacher's rating (0.51~0.67) were substantially positive. (3) There were significant differences among schools, grades, classes, and sex. (4) Liner stepwise regression analysis showed that language and math sub-scales were important predictors (β s = 0.27, 0.49 and 0.39, 0.23 for AATMS-A, B respectively) for the total final exam scores, R^2 = 0.46, 0.29 for AATMS-A, B respectively. **Conclusion:** The structure of AATME is reasonable, and criterion-related validity were supported in present study. AATMS can be used for middle school students.

【Key words】 AATMS; Construct-related validity; Criterion-related validity

中学生学业能力倾向测验(简称 AATMS)编制于 2000 年。该测验为纸笔团体测验, 包括 A、B 两个复本, 均由三个分量表(语文、数理、图形)含七个分测验(算术、填词、填数、填句、平面、阅读、立体)组成, 总计 128 个测验条目。已有文详细报告了测验的编制过程以及条目质量和信度考察结果^[1], 本文拟对该测验的各种效度研究做一详尽报告及分析。

1 方法

1.1 被试

共对石家庄市 6 所中学的初一到高三的 1424 名中学生施测 AATMS(A 式 713, B 式 711), 经 χ^2 检验, 各年级男女比例, A 式和 B 式的实测比例均无显著差异, 父母亲文化程度各年级无显著差异, 职业状况各年级之间差异显著。另外, 效标效度样本 363 例。

1.2 效度研究方法

采用 SPSS11.0 和 Amos4.0 统计软件分别进行数据的探索性因素分析和验证性因素分析来考察 AATMS 的结构效度。在本研究中, 将 A、B 式样本分成高中组和初中组两部分, 对高中组数据进行探索性因素分析, 对初中组数据进行验证性因素分析。

采用 t 检验、协方差分析、相关以及回归分析等方法, 分别以长鞍团体智力测验^[2]、教师评定、团体差异以及学习成绩为效标进行效标关联效度检验。

2 结果

2.1 结构效度研究结果

2.1.1 探索性因素分析 分别对 A 式和 B 式的高中组被试数据采用主成分分析与方差极大正交旋转法进行分析, 两组数据分析结果一致。各分测验均对一个基本能力因子中度以上的负荷, 三因子因素

分析发现,三因子模型恰好与 AATMS 结构吻合,填词、填句、阅读分测验均与语言能力有关,命名为语文因子,算术、填数分测验都与数的能力有关,命名

为数理因子,平面、立体分测验都与图形的知觉和想象有关,命名为图形因子。抽取的三个因子解释了总方差的 65.4% (A 式), 66.2% (B 式)。见表 1

表 1 AATMS 结构效度的探索性因素分析结果

	A 卷				B 卷			
	基本能力因子负荷	语文因子	数理因子	图形因子	基本能力因子负荷	语文因子	数理因子	图形因子
填词	0.60	0.66	0.25	0.33	0.54	0.58	0.29	0.33
填句	0.64	0.80	0.22	0.20	0.68	0.82	0.13	0.19
阅读	0.56	0.72	0.18	0.12	0.65	0.80	0.24	0.08
算术	0.74	0.30	0.83	0.30	0.60	0.38	0.62	0.21
填数	0.73	0.21	0.83	0.13	0.87	0.15	0.92	0.17
平面	0.59	0.27	0.15	0.70	0.62	0.20	0.37	0.76
立体	0.73	0.15	0.21	0.83	0.67	0.18	0.09	0.81

2.1.2 验证性因素分析 分别用 A、B 式的初中组被试数据对 AATMS 结构进行验证性因素分析,采用的拟合指标包括调整拟合优度指数 AGFI(adjusted goodness-of fit index)、相对拟合指数 CFI(comparative fit index)、非标准拟合指数 NNFI(nonnormal fit index)、近似误差均方根 RMSEA(root mean square error of approximation)、及卡方值与自由度的比值 χ^2/df 。结果见图 1 和表 2。图表中数据表明,因素载荷介于 0.51~0.78 之间,各拟合指数 AGFI、CFI、NNFI 均在 0.96 以上, RMSEA 小于 0.05, χ^2/df 小于 3,均符合心理测量学要求^[3]。

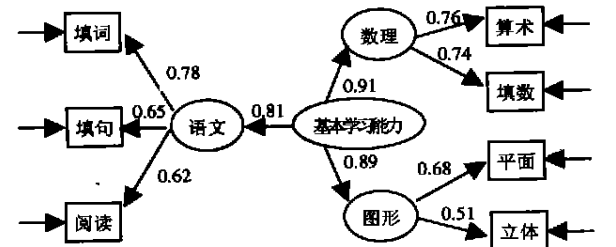


图 1 AATMS(A 式) 结构效度的验证性因素分析结果图示

表 2 验证性因素分析的拟合指数

	χ^2/df	AGFI	CFI	NNFI	RMSEA
A 式	1.336	0.971	0.982	0.991	0.030
B 式	1.243	0.971	0.992	0.996	0.027

2.2 效标关联效度研究结果

2.2.1 与长-鞍团体智力测验结果的相关分析 以高一和初二的 99 名学生(A 式 49, B 式 50) 为对象,计算两个测验分数间的相关。结果见表 3。

表 3 A 卷测验分数与长-鞍团体智力测验结果的相关(r)

	语文	数理	图形	总分
言语理解	0.56**	0.48**	0.33*	0.63**
知觉推理	0.44**	0.38**	0.22	0.49**
注意记忆	0.42**	0.28	0.35*	0.48**
总量表分	0.66**	0.53**	0.39**	0.73**

注: * $P < 0.05$, ** $P < 0.01$,下同。

2.2.2 与教师评定的相关分析 选取甲乙两所中学的两个初三班级(甲校 A 卷 30, B 卷 30; 乙校 A 卷 29, B 卷 28),请其班主任对学生的学习能力进行评定,计算其与测验分数的相关,结果列于表 4。

2.2.3 测验分数的年级差异比较 经方差分析,各分测验、分量表及测验总分在不同年级间存在显著差异,以父、母亲职业为协变量再做协方差分析,年级间差异依然显著。两两比较发现,许多相邻年级之间测验分数差异显著,若相隔两个年级,其测验分数几乎均有显著差异。见表 5。

2.2.4 不同学校及班级测验分数差异比较 我们比较了高中重点学校(男生 88 人,女生 89 人, A 式 90, B 式 87)、普通学校(男生 224 人,女生 192 人, A 式 210, B 式 206)和较差的学校(男生 57 人,女生 81 人, A 式 65, B 式 73)的测验分数。经 χ^2 检验,三类学校所做 A、B 式比例无显著差异,男、女生比例差异显著($P = 0.037$),故以性别为协变量对三类学校的测验分数进行协方差分析,结果三类学校的所有测验分数均差异显著。见表 6。另外,我们对一所学校同一年级一理科班(共 60 人,做 A 式、B 式各 30 份)与一文科班(共 42 人,做 A 式、B 式各 21 份)的测验成绩进行差异性检验,结果发现,理科班被试数理、图形分量表分以及测验总分显著高于文科班被

试,而文科班被试语文分量表分显著高于理科班被 试($P < 0.05$)。

表 4 测验分数与教师评价的相关(r)

教师评定		A 卷				B 卷			
		语文	数理	图形	总分	语文	数理	图形	总分
甲校	语文能力	0.62 ^{* *}	0.47 ^{* *}	0.27	0.63 ^{* *}	0.48 ^{* *}	0.16	0.06	0.20
	数学能力	0.56 ^{* *}	0.59 ^{* *}	0.30	0.66 ^{* *}	0.25	0.60 ^{* *}	0.44 [*]	0.63 ^{* *}
	总 能 力	0.61 ^{* *}	0.55 ^{* *}	0.30	0.67 ^{* *}	0.43 [*]	0.31 ^{* *}	0.33 [*]	0.53 ^{* *}
乙校	语文能力	0.50 ^{* *}	0.15	0.09	0.36 ^{* *}	0.55 ^{* *}	0.19	0.38 [*]	0.57 ^{* *}
	数学能力	0.35	0.48 ^{* *}	0.08	0.43 [*]	0.47	0.56 ^{* *}	0.48 [*]	0.52 ^{* *}
	总 能 力	0.55 ^{* *}	0.42 [*]	0.12	0.51 ^{* *}	0.63 ^{* *}	0.45 ^{* *}	0.52 ^{* *}	0.66 ^{* *}

表 5 A、B 式合并统计各年级测验分数差异比较

	初一 ($n = 248$)	初二 ($n = 227$)	初三 ($n = 218$)	高一 (285)	高二 ($n = 224$)	高三 ($n = 222$)	F	P
算 术	8.50	8.52	9.57 ^{★▲◆}	10.79 ^{★▲◆}	11.18 [★]	11.39 ^{★◆}	53.264	0.000
填 词	7.63	8.37 ^{★▲}	9.41 ^{★▲◆}	11.01 ^{★▲◆}	11.37 ^{★◆}	12.21 ^{★▲◆}	123.198	0.000
填 数	8.70	8.89	9.86 ^{★▲◆}	10.70 ^{★▲◆}	10.86 ^{★◆}	10.97 [★]	29.853	0.000
填 句	7.24	8.15 ^{★▲}	9.82 ^{★▲◆}	11.08 ^{★▲◆}	11.31 ^{★◆}	12.43 ^{★▲◆}	164.554	0.000
平 面	8.61	8.94	9.53 ^{★▲◆}	10.59 ^{★▲◆}	11.33 ^{★▲◆}	10.99 [★]	37.733	0.000
阅 读	7.74	8.59 ^{★▲}	9.39 ^{★▲◆}	10.94 ^{★▲◆}	11.30 ^{★◆}	12.04 ^{★▲◆}	99.470	0.000
立 体	8.66	8.76	9.48 ^{★▲◆}	10.57 ^{★▲◆}	11.22 ^{★▲◆}	11.32 ^{★◆}	42.379	0.000
语 文	22.62	25.12 ^{★▲}	28.63 ^{★▲◆}	33.04 ^{★▲◆}	33.98 ^{★◆}	36.67 ^{★▲◆}	218.208	0.000
数 理	17.20	17.40	19.43 ^{★▲◆}	21.49 ^{★▲◆}	22.05 ^{★◆}	22.36 ^{★◆}	54.537	0.000
图 形	17.26	17.70	19.02 ^{★▲◆}	21.16 ^{★▲◆}	22.54 ^{★▲◆}	22.31 ^{★◆}	60.648	0.000
总 分	57.08	60.22 ^{★▲}	67.08 ^{★▲◆}	75.69 ^{★▲◆}	78.68 ^{★▲◆}	81.34 ^{★▲◆}	172.396	0.000

注:★表示与初一年级比较 $P < 0.05$, ▲表示与前一年级比较 $P < 0.05$, ◆表示与前两个年级比较 $P < 0.05$ 。

表 6 三类学校测验分数的差异比较($\bar{x} \pm s$)

	较差	普通	重点	F	P
算术	9.27±2.28	10.80±2.35 ^{★★}	13.21±2.16 ^{★★}	84.463	0.000
填词	10.01±2.38	11.09±2.31 ^{★★}	13.56±2.42 ^{★★}	68.382	0.000
填数	9.66±2.46	10.61±2.54 ^{★★}	12.24±2.52 ^{★★}	38.214	0.000
填句	10.41±2.46	11.32±2.15 ^{★★}	13.02±1.78 ^{★★}	46.645	0.000
平面	10.43±2.73	10.67±2.57	11.98±2.49 ^{★★}	14.192	0.000
阅读	10.13±2.36	11.04±2.54 ^{★★}	13.16±2.36 ^{★★}	45.809	0.000
立体	10.13±2.39	10.77±2.77 ^{★★}	12.21±2.62 ^{★★}	24.673	0.000
语文	30.53±5.53	33.46±4.85 ^{★★}	39.73±4.74 ^{★★}	101.848	0.000
数理	18.93±3.84	21.42±3.98 ^{★★}	25.45±3.92 ^{★★}	85.388	0.000
图形	20.57±4.02	21.43±4.22 ^{★★}	24.19±4.09 ^{★★}	30.057	0.000
总分	70.06±8.38	76.31±9.32 ^{★★}	89.37±8.99 ^{★★}	136.547	0.000

注:★表示与差校比较 $P < 0.05$, ▲表示与前一组比较 $P < 0.05$

对于初中,我们比较了重点班(A 卷 97, B 卷 94, 男生 108 人,女生 83 人,共 191 人)与非重点班(A 卷 251, B 卷 251,男生 247 人,女生 255 人,共 502 人)的测验分数, χ^2 检验重点班和普通班所做 A、B 卷比例、男女比例均无显著差异。故把 A、B 式合并统计检验两种班的测验分数差异,重点班所有的测验分数均显著高于普通班。见表 7。

2.2.5 测验分数的性别差异比较 A 式和 B 式的性别差异比较结果基本一致,男生在数理和图形的

各分测验上的得分显著高于女生,在语文各分测验尤其是填句和阅读分测验上的得分女生高于男生,但差异不显著。见表 8(A 式表略)。

2.2.6 学习成绩的回归分析 以两所学校初二年级三个班的 147 名学生(A 式 66, B 式 81)施测 AATMS 约两个月后的期末统考成绩(语文、数学、英语、物理成绩及总成绩)为因变量,以 AATMS 各分测验或各分量表分数为预测变量,进行多元线性逐步回归分析,无论 A 式和 B 式,从分测验的回归看,填词分测验能有效预测语文和英语成绩(beta 为 0.36~ 0.65),算术分测验能有效预测数学成绩(beta 为 0.30),填词和算术分测验对统考总分也具有有效预测作用(beta 为 0.29~ 0.57)。从分量表的回归看,语文分量表能有效预测语文和英语成绩(beta 为 0.30~ 0.47),而数理分量表能有效预测数学成绩(beta 为 0.34~ 0.65)。语文和数理分量表能有效预测统考总成绩(beta 为 0.23~ 0.49),以统考总成绩为因变量,以语文和数理分量表为预测变量的回归方程的决定系数为 0.46(A 式)和 0.29(B 式)。表 9 详细列出了测验 A 式对各考试科目及考试总分起有效预测作用的分测验或分量表以及它们的各项回归指标(B 式表略)。

表 7 初中重点班与普通班测验分数的差异比较($\bar{x} \pm s$)

	重点班	普通班	t 值	P 值
算术	10.55 \pm 3.22	8.19 \pm 2.50	9.115	0.000
填词	9.42 \pm 2.48	8.05 \pm 2.40	6.633	0.000
填数	11.02 \pm 3.32	8.40 \pm 2.65	9.782	0.000
填句	9.09 \pm 2.70	8.07 \pm 2.73	4.393	0.000
平面	9.93 \pm 3.26	8.67 \pm 2.87	4.744	0.000
阅读	9.29 \pm 2.87	8.26 \pm 2.44	4.390	0.000
立体	9.89 \pm 2.87	8.59 \pm 2.81	5.405	0.000
语文	27.80 \pm 6.43	24.39 \pm 5.83	6.689	0.000
数理	21.57 \pm 5.84	16.60 \pm 4.26	10.737	0.000
图形	19.82 \pm 5.09	17.25 \pm 4.55	6.425	0.000
总分	69.20 \pm 13.92	58.23 \pm 11.49	9.695	0.000

表 8 B 式各分测验各分量表及总分的性别差异比较($\bar{x} \pm s$)

	男生($n=361$)	女生($n=350$)	t	P
算术	10.42 \pm 2.82	9.57 \pm 3.12	3.786	0.000
填词	10.17 \pm 3.03	9.82 \pm 2.97	1.558	0.120
填数	10.47 \pm 2.87	9.51 \pm 3.05	4.301	0.000
填句	9.85 \pm 2.98	10.16 \pm 3.02	-1.366	0.173
平面	10.28 \pm 2.93	9.71 \pm 3.05	2.548	0.011
阅读	9.90 \pm 2.99	10.10 \pm 3.01	-0.893	0.372
立体	10.30 \pm 2.99	9.70 \pm 2.98	2.674	0.008
语文	29.92 \pm 7.56	30.07 \pm 7.64	-0.256	0.798
数理	20.89 \pm 4.92	19.09 \pm 5.43	4.640	0.000
图形	20.57 \pm 4.92	19.41 \pm 5.03	3.108	0.002
总分	71.37 \pm 14.63	68.53 \pm 15.10	2.539	0.011

注: * $P < 0.05$, * * $P < 0.01$

表 9 考试成绩的逐步回归分析(A 式)

考试课程	分测验或分量表	B	beta	t	R^2
语文	填词	1.427	0.400	3.740* *	0.320
	立体	1.140	0.316	2.959* *	
数学	算术	2.195	0.303	2.377*	0.370
	填词	1.873	0.271	2.853* *	
	填数	1.832	0.270	2.469*	
英语	填词	2.438	0.335	2.996* *	0.391
	算术	2.088	0.274	2.628*	
	填数	1.765	0.248	2.304*	
物理	算术	2.638	0.398	3.813* *	0.327
	填数	2.211	0.356	3.411* *	
总分	算术	17.499	0.387	3.952* *	0.465
	填数	12.592	0.297	2.950* *	
	填词	11.350	0.263	2.503*	
语文	图形	0.907	0.392	3.590* *	0.357
	语文	0.532	0.325	2.974* *	
数学	数理	1.828	0.395	3.375* *	0.380
	语文	0.975	0.307	2.623*	
英语	数理	1.899	0.391	3.305* *	0.367
	语文	0.998	0.299	2.529*	
物理	数理	1.867	0.442	3.736* *	0.368
	语文	0.705	0.243	2.053*	
总分	数理	14.206	0.492	4.507* *	0.461
	语文	5.441	0.272	2.514*	

注: * $P < 0.05$, * * $P < 0.01$

3 讨 论

3.1 结构效度

本研究把被试分成初中组和高中组两部分, 分别对其进行探索性和验证性因素分析, 来检验 AATMS 的结构效度。从探索性因素分析结果来看, 各分测验均对一个基本能力因子中度以上的负荷, 抽取的语文、数理及图形三个因子共解释了总方差的 65.4%(A 式), 66.2%(B 式), 恰好解释了 AATMS 的结构。验证性因素分析通过数据与理论假设模型之间的吻合程度来表示一个测验结构效度的高低, 如果各项拟合指数都较好, 说明测验具有较好的结构效度。反之, 表明理论构想需进一步完善或测验有待修订。在对测验结构效度的考验中, 常用的拟合指数 AGFI、CFI 和 NNFI 的值在 0.95 以上表示模型拟合较好, RMSEA 的值小于 0.05 表示模型拟合得很好, χ^2/df 的值越小越好^[3-5]。本研究 A、B 式的各项拟合指数均符合上述要求, 表明本测验具有理想的结构效度。

3.2 效标关联效度

学业能力倾向虽与学业成就、智力不同, 但三者间密切相关, 故学习成绩和智力测验分数是考察学业能力倾向测验效度的常用效标^[6-7]。本研究以龚耀先等编制的长-鞍团体智力测验为效标的考察结果表明, AATMS 的语文和数理分量表与长-鞍团体智力测验的言语理解因素和知觉推理因素之间相关显著, AATMS 总分与长-鞍团体智力测验总分之间呈中度或以上的相关。正如有关专家所说, 学业能力倾向测验大多是测量了与学习密切相关的智力技能, 故学业能力倾向测验在内容和形式上与团体智力测验十分相似^[6]。因此, AATMS 与长-鞍团体智力测验的较高相关即在情理之中, 也说明 AATMS 测量了一般的能力倾向。

教师对学生学业能力的评定也是学能测验常用的效标之一。有学者指出, 智力测验分数与教师对学生的等级评定之间的相关系数只要在 0.35~0.50 之间就可以了, 因为教师的评价会受到与智力无关的其它因素的影响^[6]。林传鼎和张厚粲教授等改编修订的用于小学毕业生的学习能力测验, 其总分与教师对学生的学习能力评定之间的相关为 0.43 和 0.58^[8]。在本研究中, AATMS 测验总分与教师评定之间的相关在 0.51~0.67 之间, 表明两者具有相当程度的一致性, 这也成为 AATMS 的一个有力的效度支持。

个别智力测验关注测验分数在不同年龄间的差异, 常模多采用年龄常模, 而用于教育情景中的团体

测验多采用年级常模,更关注年级差异。本研究结果显示,AATMS 各分测验及各分量表及总分年级间差异显著,测验分数随年级增长呈稳定上升趋势。林崇德、莫雷等关于中、小学生学习能力随年级发展的系列研究表明^[9-10],运算能力随年级升高呈上升发展的趋势,初三年级的运算能力发展速度大大超过由初一到初二的速度,获得一个质的飞跃。数学逻辑推理能力到高中阶段趋向抽象和简化,在中学阶段数学逻辑思维能力发展还不成熟。空间想象能力经过初二一年的学习,到初三时迅速发展,到高二其发展进入成熟期。小学和中学生的语文能力的发展尽管也有阶段性,但其发展的连续性要比数学能力显得更明显。对文字的概括和表达水平及阅读能力从小学四年级直至高中在逐步发展。A、B 式的年级差异检验结果一致表明,数理和图形分量表的各分测验的分数,初二年级与初一年级差异不大显著,而初三年级显著高于初二年级,高一和高二年级显著高于初三年级;语文分量表的各分测验分数,几乎每个相邻年级均有较为显著的差异。这样的结果与上述对学习力发展的研究结果基本吻合。测验分数除了在年级间存在显著差异外,还在重点、一般和较差学校间以及文理科班之间存在显著差异,这也表明 AATMS 具有区分不同团体的效能。

智力的性别差异一直是许多研究所关注的问题,一般的研究结论认为:就整体以及大部分能力因素而言,性别间无差异。但在某些能力因素上存在性别差异,男生在立体空间关系能力和数量问题解决能力上优于女生,女生在某些言语技能尤其是言语流畅性上优于男生。所以,在强调言语和言语记忆的测验上女生优于男生,在有关空间和数学测验上男性优于女性^[6,11]。美国著名的学能测验 SAT 的效度研究表明,SAT 分数的性别差异主要表现在数学分测验(SAT-M)上,男生的分数明显高于女生^[12]。针对 SAT 分数上的性别差异引起了许多研究和争论,有些学者从测量学的角度解释,认为 SAT 本身可能有性别偏见,低估了女生的学习成绩^[13]。另外的学者则从一些认知和非认知因素的男女生差异来解释测验分数的差异,认为 SAT-M 的分数差异与诸如学习态度、自信心、课程选择和回避、父母和教师的期望,解决问题的策略,空间能力等等因素相关^[14-15]。本研究结果显示,男生在算术、填数以及立体分测验上的得分显著高于女生;在填句及阅读分测验上,女生高于男生,差异不显著。在 AATMS 分数上显示的性别差异与上述研究及同类测验基本

一致。但造成分数的性别差异的原因应是一个值得进一步研究的问题,以便对其做出科学慎重的解释。

诊断和预测是能力倾向测验的两种主要功能,而对学习成就的预测更是学业能力倾向测验的独特价值所在。复习文献可见,对学业能力倾向测验的效用的研究除同时效度外,更多的集中在对其预测效度的研究上^[16]。本研究收集了 AATMS 施测约两个月后的初二年级的期末统考成绩作为效标对 AATMS 的预测效度做了初步检验。从结果看,AATMS 总分对语文、数学、英语、物理及统考总分均具有显著的预测力,A 式和 B 式综合来看,对于分测验,填词分测验能有效预测语文和英语成绩,算术分测验能有效预测数学成绩,填词和算术分测验对统考总分也具有有效预测作用。对于分量表,语文分量表能有效预测语文和英语成绩,而数理分量表能有效预测数学成绩。这表明 AATMS 不同的分测验或分量表对不同的学科成绩的预测力是有差异的,这些预测上的差异说明利用 AATMS 不同的分测验或分量表来预测学生的不同学科的学习成绩是有效的。在此研究中,图形分量表未能进入统考总分的回归方程,考虑到此研究只是以初二学生为对象,而在初二的课程中涉及到图形尤其是立体图形的很少,所以不能就此否定此分量表的意义,其效能应进一步检验。另外,有些分测验对有些学科成绩的预测也并不很理想,这些均有待于对 AATMS 的预测效度进一步研究后对其进行改进。

综上,经因素分析检验,AATMS 结构效度良好,以各种不同效标的检验结果表明其具有较高的效标效度。但上述对 AATMS 的效度研究尤其是对其预测效度的研究还只是初步的,所以采用不同的方法,针对不同的团体,搜集提供 AATMS 的效度资料尤其是预测效度的资料将使我们今后工作努力的方向。

参 考 文 献

- 1 张月娟,龚耀先.中学生学业能力倾向测验的初步编制.中国临床心理学杂志,2002,12(1):1-5
- 2 龚耀先.长-鞍团体智力测验手册.湖南医科大学,1997
- 3 孟庆茂,侯杰泰.协方差结构模型与多层线性模型.北京师范大学心理计量与统计分析教研室,2001
- 4 Frank J. Floyd, Keith F. Widaman. Factor analysis in development and refinement of clinical assessment instruments. Psychological Assessment, 1995, 7(2): 286-299
- 5 Van JW, Prooijen, Van WA, Der Kloot. Confirmatory analysis of exploratory obtained factor structures. Educational and Psychological Measurement, 2001, 61(5): 777-792

的情绪耗竭感类似,并包含失效感的一些内容。因素二指向情绪的不安与不稳定性,命名为“焦虑”。MBI 中很少的测量情绪不稳定性的题目被划入情绪耗竭感。“焦虑”表现出本测验较为独特的方面,即个体在感觉到自身能量逐渐衰弱、消耗过度以及自我价值感降低的同时,还表现出对自我感觉到的这种状况的担心、焦虑与不安。这种现象在我国台湾学者进行的对发生职业过劳个体的访谈研究中也同样出现,他们称之为“在抵抗与耗损的两端中不断摆动与挣扎的现象”^[11]。精疲力竭、焦虑与 EPQ 分量表分数、内在与外在在工作满意度以及一般自我效能感之间的相关关系不仅表明职业过劳与人格特质、工作态度与自我评价存在密切关系,而且从内涵上将这两个因素之间的差别进一步区分开来。EPQ-N 分高的个体的一般特征是:焦虑、担忧,因此与本测验的焦虑因素相关很高。EPQ-P 高分则可能是孤独、缺乏情感和情感投入,偏外向者的职业过劳程度低,因此与精疲力竭关系密切,以上结果同国外相关研究的结果一致^[12]。精疲力竭、焦虑与实际销售业绩分组排名之间的关系模式,提示了我们此测验在应用方面具有一定价值。优等组,排名越前者越感到精疲力竭;中等、差等组,排名越后者越感到精疲力竭;焦虑也表现出类似的关系模式。看来最先者有领跑的压力,中后者担心淘汰,又不想放弃。与此相对,最轻松的是中间偏前的一部分人。这恰恰可能是国人中庸思想的一种体现。职业过劳与工作绩效之间的关系比较复杂,在这方面开展追踪式的研究很有意义。这样,不仅能够深化、细化职业过劳与其他变量之间的关系,而且可以为职业过劳的测量提供更加有利的效度证据。

参 考 文 献

- Maslach CM, Leiter MP. Burnout and quality in a sped-up world. *The Journal for Quality and Participation*, 2001, 24 (2): 48- 51
- 李 靖. 枯竭研究进展. *北京大学学报(自然科学版)*, 2003, 39(2): 286- 290
- Maslach CM, Schaufeli WB, Leiter MP. Job burnout. *Annual Review of Psychology*, 2001, 52: 397- 422
- Brotheridge CM, Lee RT. Emotional labor and burnout: Comparing two perspectives of "people work". *Journal of Vocational Behavior*, 2002, 60: 17- 39
- Enzmann D, Schaufeli WB, Janssen P, et al. Dimensionality and validity of the Burnout Measure. *Journal of Occupational and Organizational Psychology*, 1998, 71(4): 331- 351
- 钱铭怡, 武国城, 朱荣春, 等. 艾森克人格问卷简式量表中国版(EPQ-RSC)的修订. *心理学报*, 2000, 32: 317- 323
- Hirschfeld RC. Does revising the intrinsic and extrinsic subscales of the Minnesota Satisfaction Questionnaire short form make a difference? *Educational and Psychological Measurement*, 2000, 60(2): 255- 270
- Schwarzer R, Babler J, Kwiatek P, et al. The assessment of optimistic self-beliefs: Comparison of the German, Spanish, and Chinese version of the General Self-efficacy Scale. *Applied Psychology: An international review*, 1997, 46(1): 69- 88
- 黄光国, 杨宜蕙. 华人关系主义与情绪经验: 台湾大学生的案例. 台湾中央研究院民族学研究所行为研究组主办"述说、记忆与历史: 以'情绪文化'为核心的论述学术研讨会"论文. 1999, 台北
- Shirom A. Burnout in work organizations. In Schabracq MJ, Winnubst JAM, Cooper CL. *Handbook of Work and Health Psychology*. 1989, 51- 86.
- 李志鸿. 工作倦怠的理解与诠释: 从故事出发的初步探讨. 台湾辅仁大学应用心理学研究所, 硕士论文, 1994
- Mills LB, Huebner ES. A prospective study of personality characteristics, occupational stressors and burnout among school psychology practitioners. *Journal of School Psychology*, 1998, 36(1): 103- 120
- 郑日昌. 心理测量. 湖南教育出版社, 1987
- 洪 炜, 龚耀先. 一般行政能力倾向测验的建构及信度、效度研究. *中国临床心理学杂志*, 2000, 8(1): 1- 6
- 林传鼎, 张厚粲, 周 谦, 等. 少年儿童学习能力测验的编制. 北京师范大学教育研究所, 1979
- 林崇德. 学习与发展. 北京师范大学出版社, 1999
- 莫 雷. 中小學生语文阅读能力结构的发展特点. *心理学报*, 1992, 24(4): 346- 353
- Diane F. Halpern. Sex differences in intelligence, implication for education. *American Psychologist*, 1997, 52(10): 1091- 1102
- College Board. College-Bound Seniors: 1989 Profile of SAT and Achievement Test Takers. New York, 1989
- Leonard, David K, Jiang. Gender bias and the college predictions of the SATs: A cry of despair. *Research in Higher Education*, 1999, 40(4): 375- 407
- Beth CM. Mediators of gender differences in Mathematics College Entrance Test Scores: A comparison of spatial skills with internalized beliefs and anxieties. *Developmental Psychology*, 1997, 33(4): 669- 80
- Gallagher AM, De Lisi, Richard, Holst PC, et al. Gender differences in advanced mathematical problem solving. *Journal of Experimental Child Psychology*, 2000, 75(3): 165- 190
- James V, Mitchell JR. The Ninth Mental Measurements Yearbook. The Buros Institute of Mental Measurements of The University of Nebraska- Lincoln, 1985

(收稿日期: 2003- 12- 19)

(上接第 225 页)

- 6 郑日昌. 心理测量. 湖南教育出版社, 1987
- 7 洪 炜, 龚耀先. 一般行政能力倾向测验的建构及信度、效度研究. *中国临床心理学杂志*, 2000, 8(1): 1- 6
- 8 林传鼎, 张厚粲, 周 谦, 等. 少年儿童学习能力测验的编制. 北京师范大学教育研究所, 1979
- 9 林崇德. 学习与发展. 北京师范大学出版社, 1999
- 10 莫 雷. 中小學生语文阅读能力结构的发展特点. *心理学报*, 1992, 24(4): 346- 353
- 11 Diane F. Halpern. Sex differences in intelligence, implication for education. *American Psychologist*, 1997, 52(10): 1091- 1102
- 12 College Board. College-Bound Seniors: 1989 Profile of SAT and Achievement Test Takers. New York, 1989
- 13 Leonard, David K, Jiang. Gender bias and the college predic-

(收稿日期: 2004- 03- 20)