

# 成人词汇测验的初步编制

唐细容\*, 姚树桥\*\*

(中南大学湘雅二医院医学心理研究中心, 湖南 长沙, 410001)

【摘要】 目的: 编制成人词汇测验, 用于成人言语性学业成就和智能的简易评估。方法: 用编制的测验测试 450 名 16~65 周岁、按年龄和文化程度分层抽取的成人, 进行条目分析、信度和效度检验。结果: 大多数条目的难度分布于 0.25~0.75, 点二列相关和鉴别指数大于 0.30, 条目偏差绝对值在 0.10 之内, 信度指标多在 0.90 以上, 测验量表分与受教育年限相关 0.683, 与中国修订韦氏成人智力量表(WAIS-RC)言语智商、操作智商和全量表智商相关系数分别为 0.605、0.495、0.591。结论: 成人词汇测验条目基本符合要求, 分数可靠, 有一定的效度。

【关键词】 词汇测验; 成人; 智力; 心理测量学

中图分类号: G449.4

文献标识码: A

文章编号: 1005-3611(2003)01-0024-04

## Development of Vocabulary Test for Adults

TANG Xi-rong, YAO Shu-qiao

The Second Xiangya Clinical College, Central South University, Changsha, 410011, China

【Abstract】 **Objective:** To develop Vocabulary Test for Adult as brief instrument to assess verbal academic achievement and intelligence. **Methods:** Two sets of Chinese vocabulary were presented to 450 adults stratified by age and education, aged from 16 to 65, and oral interpretation of each vocabulary was required. Based on their responses, item analysis of the test was carried out. Reliability and validity of the test were also examined. **Results:** Most difficulty indexes were between 0.25 and 0.75; discrimination indexes were more than 0.30. There were no significant item biases. Reliability coefficients of the test were more than 0.90. Performance of the test showed significant correlation with years of education. Correlation coefficients between scale scores of this test and WAIS-RC verbal IQ, performance IQ, full scale IQ were 0.605, 0.495, 0.591 respectively. **Conclusion:** The Vocabulary Test for Adults has good psychometric properties.

【Key words】 Intelligence; Vocabulary test; Adult; Psychometrics

词是日常信息沟通的基础, 它的正确理解和运用是幼儿教育及学校基础教育的重要内容和基本要求, 作为概念的物质外壳, 其内在意义是客观事物和现象在人们意识中的概括的反映, 词的理解运用与相应概念的获得掌握及思维密不可分。因此, 个人词汇即个人在生活与学习中获得的词的数量和质量在一定程度上能反映他的教育情况和个人智能水平, 因而词汇测验倍受心理与教育测量专家重视, 智力测验和学业成就测验常有词汇测验的内容<sup>[1,2]</sup>, 还有一些专门的词汇测验<sup>[3]</sup>, 心理测验年鉴-9 收录的测验中有 6 项直接命名为词汇测验, 61 项测验提供词汇(vocabulary)分数, 此外还有测验包含物体命名等与词汇相关的内容<sup>[4]</sup>。

词汇测验能参与评估多种密切相关但不完全等同的心理功能, 在临床和科研中应用非常广泛。它能用于儿童心理发展水平的评估, 为发展心理学研究与精神发育迟滞的诊断等提供量化资料; 也用于评定语言课程的教学质量以帮助改进教与学<sup>[3]</sup>; 还

是学业成就、智力、学习能力倾向和社会交往相关职业能力倾向测验的重要内容和病前智力估计的常用方法<sup>[5]</sup>; 此外, 跨文化研究中也用到此类测验<sup>[6]</sup>。

自上个世纪八十年代初以来, 我国修订或编制了一些包含词汇测验的智力量表<sup>[7]</sup>以及专门的词汇测验<sup>[3]</sup>, 本研究试图编制有两个平行版本的成人词汇测验, 希望能为 16 岁及以上成人的言语性学业成就和智能评估提供简易工具。

## 1 材料与方法

### 1.1 词汇选择原则和方法

基本原则是测验对现代汉语常用词汇有代表性, 语义简约、确定以利于评分, 避免消极情绪和特殊职业可能带来测量误差。为此, 依据《现代汉语频率词典》, 采用等间距的方法从使用度不小于 10 的常用词、日常生活非常常用的口语词和使用度为 5~1 的不常用书面词中各抽取 260、20、20 个词。参照《现代汉语词典》选出只有一项语义或两项密切相关语义的词语, 去掉专业术语和消极情绪词语, 留下各有 65 个词性、语体色彩基本匹配的两组词语进行

\* 研究生, 现在南华大学附一医院工作。 \*\* 导师

预备测试。

1.2 测验取样本的确定

经过两次预试及条目筛选,初步建立各包含 40 个条目的测验取样版本。以常用词性计,甲式有名词 13 个、动词 17 个、形容词与副词 10 个,乙式依次为 15 个、16 个、8 个,连词 1 个。以使用度计,甲式高使用度(100 以上)者 5 个、中使用度(10 ~ 100)32 个、低使用度(10 以下)3 个,乙式依次为 6 个、31 个、3 个。

1.3 取样计划与实施

计划按年龄和文化程度两个变量分层抽取 16 周岁以上 65 周岁以下成人 400 名。年龄分为 5 个组:16 ~ 24 岁组(16 周岁以上且不满 25 周岁)、25 ~ 34 岁组、35 ~ 44 岁组、45 ~ 55 岁组、55 ~ 64 岁组,每组 80 人,甲乙版本各 200 人。文化程度的指标是受教育年限,分为 5 个层次:0 ~ 3 年(上学时间不足或仅有 3 年)、4 ~ 6 年(3 年以上但不足或仅有 6 年)、7 ~ 9 年、10 ~ 12 年、12 年以上。各年龄组的文化层次构成比按 1995 年全国 1% 人口抽样调查结果<sup>[8]</sup>粗略估计。抽样时适当考虑性别、居住地、职业等人口统计学变量,排除具有器质性脑疾病病史或精神发育迟滞史、现患有神经精神疾病、因各种原因正在服用精神药物及不愿意参加测试的人群。实际抽样于 2000 年 11 月至 2001 年 8 月在湖南省城乡三个地点进行,共抽取 450 人,各年龄组人数相等。所有测试均单独进行,主试是同一人,要求被试参照示范尽可能准确、完整、简要地口头解释所给词语,主试记录被试原话,评分另外择时进行。两个版本一般按抽样的先后顺序交叉进行。

1.4 评分原则与分数转换

各条目视被试的解释准确程度分别计 0 分、1 分、2 分,评分原则与 WAIS 词汇分测验类似,连续六

个 0 分则随后的条目均记为 0 分,各条目等量加权构成测验粗分。分年龄组进行分数正态化,评定粗分的等级时相同分数(ties)给予相同等级,均以相应的最高秩次,采用 Blom 公式估计累计概率,即:累计概率=(等级-3/8)/(最高等级+1/4),以此为中介,将粗分转换成平均数为 40,标准差为 10 的正态分布等值分数。为减少异常值的影响,以粗分为自变量,等值分数为因变量进行曲线拟合,对拟合函数值四舍五入取整数作为量表分。

1.5 测验的心理测量学检验

1.5.1 条目分析 难度采用得分率和难度指数估计,后者是前者的函数,是均数为 13、标准差为 4 的正态分布中累计概率等于条目得分率的标准分。区分度用点二列相关和鉴别指数估计,前者以 1 分定为通过与否的分界值,测验粗分为效标;后者是 65 名(约 29%)高粗分者与 65 名低粗分者之间条目得分率差。条目偏差采用标准化法计算<sup>[9]</sup>,按量表分将被试水平划分为 5 个水平:0 ~ 24 分、25 ~ 34 分、35 ~ 45 分、46 ~ 54 分、55 分及以上;计算条目性别偏差时,定女性为参考组,男性为目标组;计算条目城乡偏差定农村为参考组,城镇为目标组。

1.5.2 信度与效度 采用相关分析估计重测信度、评分者一致性、复本信度和同质信度,效度主要考验了效标效度,具体操作详见结果部分。

2 结 果

2.1 被试的一般情况

450 名被试中男 243 名、女 207 名,城镇 195 名、乡村 255 名,其中农民占 44%,平均受教育年限为 8.19 年,随年龄增长,年龄组的平均受教育年限逐渐下降,文盲逐步增多,甲乙版本的被试平均受教育年限和文化层次的构成基本一致(表 1)。

表 1 各年龄组被试的受教育年限

年龄组	平均受教育年限(年)			各教育层次的人数分布									
	甲	乙	合计	0 ~ 3 年		4 ~ 6 年		7 ~ 9 年		10 ~ 12 年		12 年以上	
				甲	乙	甲	乙	甲	乙	甲	乙	甲	乙
16 ~ 24	9.73±2.58	9.71±2.69	9.72±2.62	1	1	4	4	20	20	17	17	3	3
25 ~ 34	8.84±2.74	9.04±2.76	8.94±2.74	2	2	5	6	23	21	13	14	2	2
35 ~ 44	8.07±2.54	7.98±2.54	8.02±2.53	2	3	7	6	28	28	7	7	1	1
45 ~ 54	7.04±2.97	7.36±2.79	7.20±2.87	3	3	20	19	15	16	6	6	1	1
55 ~ 64	6.91±3.74	7.20±3.91	7.06±3.81	9	9	15	13	9	13	10	8	2	2
合计	8.12±3.11	8.26±3.11	8.19±3.11	17	18	51	48	95	98	53	52	9	9

2.2 不同年龄组的测验得分

年龄组的平均得分随年龄增长逐步下降,组内两个版本平均得分无显著差异(表2),方差分析发现年龄组主效应显著( $F=11.17, P=0.000$ ),而版本主效应不显著( $F=0.002, P=0.968$ ),也无显著交互作用( $F=0.290, P=0.884$ )。

2.3 条目分析

2.3.1 难度 甲乙版本测验难度平均值分别为0.497、0.496,它们大多分布于0.25~0.27,约85%的条目

难度指数位于10~16正负3/4个标准差内。

表2 各年龄组测验粗分的分布特征与版本间比较

年龄组	甲式		乙式		合计		版本比较 $P$ 值
	均数	标准差	均数	标准差	均数	标准差	
16~24	48.02	15.56	48.71	17.04	48.37	16.23	0.842
25~34	44.02	18.38	42.38	17.81	43.20	18.01	0.668
35~44	39.78	19.07	37.74	17.68	38.76	18.31	0.599
45~54	32.69	15.34	35.58	19.01	34.13	17.23	0.430
55~64	34.13	15.40	33.91	18.79	34.02	17.08	0.951

表3 条目分析

	难度指数					鉴别指数				性别偏差				城乡偏差			
	<10	<12	12~14	>14	>16	<0.2	≥0.2	≥0.3		<-0.1	-0.1~0.1	>0.1	>0.1	<-0.1	-0.1~0.1	>0.1	>0.1
甲式	3	9	18	7	3	2	2	4	32	2	37	1	2	35	3		
乙式	2	15	11	8	4	1	1	4	34	2	36	2	0	40	0		

2.3.2 区分度 除极难的少数条目,大多条目的鉴别指数在允许范围,两版本条目的点二列相关系数分布与鉴别指数类似。中数分别为0.565、0.574。

2.3.3 条目偏差 绝大多数条目偏差的绝对值小于或等于0.10(表2),甲式、乙式以及整个测验条目的性别偏差平均值分别为-0.001、0.012、0.005,城乡偏差平均为0.018、-0.003、0.010。

2.4 信度检验

2.4.1 重测信度 用甲乙版本间隔十天左右各重复测试15名被试,前后量表分分别相关0.95、0.93(已做校正,以下信度类同),后测分数平均高出前测分2.47、2.67。

2.4.2 评分者一致性信度 请一位临床测验经验丰富的专业人员参照评分原则对两版本各25份答卷进行评分,其评分与研究者的评分相关系数为0.979、0.976,分值无显著差异。

2.4.3 复本信度 采用三组被试来检验版本的平行性。第一组,100名35岁以下被试,其中高中文化60名,大学文化40名,分成文化程度基本对等的两组,分别测试甲式或乙式,结果两组被试的粗分均值是60.72、59.72、量表分49.10、47.64,版本间均无显著差异。第二组,30名被试同一次测试两个版本,版

本的测试顺序交叉平衡,结果两个版本分数相关系数0.90,均数无显著差异。第三组,30名被试相隔10天左右相继测试两个版本,分数相关0.92,两个版本分数均数几近相等。

2.4.4 同质信度 用整个标准样本的数据计算,甲乙版本的Cronbach  $\alpha$  系数分别为0.943、0.949。分半信度为0.927、0.901。概化理论 $p \times I$ 的研究显示,两个版本当前长度的 $G$ 系数分别为0.943、0.949,增加长度到60题, $G$ 系数分别为0.962、0.966,减少到20个条目, $G$ 系数降到0.887、0.897。

2.5 效度考验

2.5.1 内容效度  $\chi^2$  检验显示测验用词词性构成相同( $\chi^2=0.226, P=>0.80$ ),两个版本测验用词的使用度中数分别为23.5、25.5,秩和检验没有发现显著差异( $z=-0.318, P=0.751$ )。

2.5.2 测验成绩与教育限的关系 测验粗分、量表分均随受教育年限上升而上升,它们与受教育年限的相关系数分别为0.756、0.683( $P<0.001$ ),方差分析发现文化层次效应显著(表4),事后分析(post hoc test)结果显示任意两个不同文化层次均有显著差异( $P<0.01$ )。

表4 不同教育限被试的测验分数分布与比较

	0~3年	4~6年	7~9年	10~12年	12年以上	$F$ 值	$P$ 值
量表分	12.29±8.37	26.48±11.30	40.92±15.04	54.92±10.11	63.67±6.24	125.49	0.000
粗分	26.74±6.05	35.01±6.46	40.25±8.32	47.61±7.02	54.78±6.38	84.56	0.000

2.5.3 测验成绩与WAIC-RC的相关 对70名被试先后测试本测验(两版本人数各半)和WAIS-RC

全套测验或其词汇分测验,其中进行了WAIS-RC全套测验30人,主要为医科大学四年级学生。结果

本测验粗分与 WAIS-RC 词汇分测验粗分相关 0.956 ( $P<0.001$ ), 量表分与后者的年龄等值量表分相关 0.861 ( $n=70$ ,  $P<0.001$ ), 与 WAIS-RC 言语智商、操作智商和全量表智商相关系数依次为 0.605、0.495、0.591。

2.6 分数转换和常模形式

多次曲线拟合结果显示, 立方曲线的拟合度最好, 各年龄组拟合函数调整后的决定系数都在 0.99 以上(表 5)。四舍五入取整后所得的量表分与等值分对应关系良好( $r=0.997$ ,  $P=0.000$ )。最后本测验的 T 分形式常模依据各年龄组的拟合函数制订。

表 5 正态等值分的拟合函数

年龄组	B <sub>0</sub>	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	调整 R <sub>2</sub>
16~24	7.49	1.28036	-0.02394	0.000223	0.992
25~34	14.21	1.22376	-0.02765	0.000283	0.992
35~44	19.60	0.83497	-0.01506	0.000169	0.997
45~54	14.25	1.40132	-0.02641	0.000238	0.993
55~64	15.75	1.30742	-0.02578	0.000251	0.995

注: a. 拟合分 =  $b_0 + b_1 * r + b_2 * r^2 + b_3 * r^3$   $r$  = 测验粗分

3 讨 论

3.1 测验内容和条目分析

本研究旨在探索言语性学业成就和智能的简易评估方法, 希望测验成绩能同时反映学业成就与智能, 因此在安排测验内容时考虑了日常生活和正规基础教育的作用。现代汉语常用词汇与日常生活和基础教育都有联系, 对其不同掌握情况在一定程度上反映一个人通过生活与学习获得言语性知识的程度和能力, 因此测验用词对现代汉语常用词汇有代表性是测验编制的基本要求和出发点。参考《现代汉语频率词典》来编制条目, 采用等间隔的方法从其常用词表选取词语, 就是出于上述考虑, 而适当增加最常用的口头词语和生僻的低频书面词语则意在帮助鉴别极端文化程度的被试。通过筛选保留下来的条目, 有日常使用非常频繁的词语如土匪、猎人等, 也有依稀、蜿蜒等必须受一定教育才知道的词语。两个版本的测验用词词性构成大体相当, 主要是实词, 实词是日常生活用词的主干和基础语言教育的重点, 因此测验内容符合测验编制的预定要求。

本测验两个版本的平均难度都接近 0.50, 大多条目难度集中与 0.25~0.75(难度指数 10~16)之间, 这样能用较少的项目提供较大的分数变化, 是一些学者主张的做法[9, 10]。条目区分度令人非常满意, 除极端难度的几项外, 鉴别指数和点二列相关都

在 0.30 以上, 大多达到 R. L. Ebel 提出的优良标准<sup>[1, 9]</sup>, 而极端难度的区分度低是因为采用的区分度估计方法都不能排除极端难度的影响。条目偏差它是指能力相同或熟练程度相同的来自不同团体的个体正确回答给定条目的概率不同, 最初见于 1960 年代跨文化团体的偏差研究, 现有人称之为条目功能差异<sup>[9]</sup>, 进行条目偏差检测有助于发现和纠正测验的文化偏向<sup>[11]</sup>。曹亦薇等<sup>[12]</sup>对发现一些汉语词语具有条目偏差。本测验从条目水平看, 大多符合要求, 从测验的整体上看, 偏差非常小。

3.2 测验的信度

重复测验的前后分数相关显著, 大小与 WAIS-II 词汇分测验的重测相关接近<sup>[2]</sup>。可以认为测验成绩在时间维度上比较稳定。后测分数稍高于先测分数, 可能是因为重测间隔比较短, 存在练习效应。同样地, 结果显示不同评分者所给分数具有较好的一致性。甲乙版本平行性得到较好的验证: 第一、无论是整个标准样本还是各年龄组中, 两个版本的均数和标准差无显著差异。第二, 给文化程度对等的两组年轻的高学历者分别测试甲式和乙式, 结果测验粗分和量表分差异不显著。第三、无论是同一次测试还是相继测试, 被试在两个版本上的得分相关很高, 说明来自版本选择方面的误差很小。a 系数和分半信度高, 提示本测验内部同质性高, 内容取样相关的误差较小。G 系数显示当前测验长度测验分数在内容取样方面已有较好的稳定性, 增长测验长度测验性能会有所改善, 但收效不太大。

3.3 效度

效度被认为是测验编制应考虑的最重要、最基础的方面<sup>[2]</sup>, 恰当安排测验内容是保证测验效度的重要环节。从使用度和词性分布看, 本测验条目能代表现代汉语常用词汇, 吻合最初的设想本测验的。版本间不仅词语的使用度和词性基本对等, 测验分数经过严格检测基本等值平行, 按 Cronbach 的观点, 可以认为测验具有内容效度<sup>[1]</sup>。

词汇测验历来是 TOFEL 等语言水平考试、广泛成就测验和智力测验的常规内容, 研究发现词汇测验在一般智力因素和言语智力因素负荷量最高<sup>[2, 13]</sup>。因此, 通过词汇测验来评估言语性智能和学业成就的构想具有理论基础。本测验分数与受教育年限及全量表智商都显著相关, 进一步验证构想的合理性。测验粗分随年龄增长而下降, 可能与样本的文化程度岁年龄增长而递减有关。测验成绩与

领域的 $\alpha$ 系数在0.32~0.99之间,分半信度为0.92~0.99,表明该量表具有较高的内部一致性。各领域的重测信度为0.76~0.99,总分重测信度为0.98,表明该量表稳定性良好。各领域评定者一致性信度为0.92~0.99,总分评定者一致性信度为0.98。这些与国内外同类量表的信度水平接近<sup>[3,4]</sup>。

从效度检验结果来看,各分测验与总分的相关在0.88~0.95之间,4个适应技能领域之间的相关系数在0.69~0.88之间。另外在分量表水平上进行因素分析,只提取出1个因子,该因子能够解释总方差的91.5%。这些结果表明该量表具有较好的内部结构效度。

在效标效度方面,本研究考察了学龄期儿童适应技能评定量表与中国韦氏儿童智力量表、儿童适应行为量表的关系。大多数研究表明,智力与适应技能之间具有中等程度的相关。可以说,社会适应功能在一定程度上反映了智力功能,表现在MR人群中,一般社会商(SQ)常高于智商(IQ)。适应功能发展水平的上限和下限均受到IQ制约,IQ的最低限度为35左右时,适应功能才有可能发展。而且,适应行为越复杂,越依赖于智力。因而,人们一般倾向于认为二者是两个既有联系又相互区别的结构<sup>[5]</sup>。本研究对学龄期儿童适应技能评定量表与中国韦氏儿童智力量表的相关研究发现:适应技能总分与C-WISC的言语量表分、操作量表分及全量表分的相关系数分别为0.45,0.20,0.36,与C-WISC各分测验成绩的相关系数在0.01~0.54之间,呈低至中度

相关,与国内外有关研究结果一致<sup>[3]</sup>。本研究中智力测验的言语部分与社会化技能的相关较高,可能因为言语理解能力与获得性技能在日常生活环境中的应用关系密切<sup>[6]</sup>。与儿童适应行为评定量表的相关分析结果表明,二者总分的相关系数为0.67,本量表总分及各领域分与儿童适应行为量表8个分量表分的相关在0.51~0.89范围,与独立因子、认知因子、社会自制因子的相关在0.63~0.94,均呈中至高度相关,说明本量表与儿童适应行为评定量表具有高度的功能一致性。当然,由于评定内容和评定目的的不同二者不可能完全替代。

#### 参 考 文 献

- 1 姚树桥,周世杰,王丽,等. 幼儿适应技能评定量表的编制. 中国临床心理学杂志, 1999, 7(2): 81—85
- 2 AAMR. Mental Retardation Definition, Classification, and Systems of Supports. 9th ed. Washington: American Association of Mental Retardation, 1992
- 3 周世杰,姚树桥,王丽. 幼儿适应技能评定量表的信度效度研究. 中国临床心理学杂志, 1999, 7(4): 208—212
- 4 Spreat S. The AAMD Adaptive Behavior Scale: A psychometric review. Journal of School Psychology, 1982, 20(1): 45—56
- 5 Platt LO, Kamphaus KW, Cole RW, et al. Relationship between adaptive behavior and intelligence: Additional evidence. Psychological Reports, 1991, 68: 139—145
- 6 郭伯良,戴晓阳,姚树桥,等. 中国幼儿智力量表和幼儿适应技能评定量表的相关研究. 中国临床心理学杂志, 1999, 7(3): 143—146

(收稿日期: 2002—04—18)

(上接第27页)

WAIS-RC的言语智商相关强于操作智商的相关,显示本测验更倾向于测量到言语性智能。

本测验得分与WAIS-RC词汇分测验分数相关很高,在数值上接近WAIS-II和WAIS-R词汇分测验的相关(0.90,  $n=192$ )<sup>[3]</sup>,与WAIS-RC与三项智商相关也显著,其中与全量表智力、言语智商相关系数相对较高,说明本测验具有一定的效标效度。

#### 参 考 文 献

- 1 郑日昌. 心理测量. 湖南长沙: 湖南教育出版社, 1987. 30—425
- 2 Wechsler D. Wechsler Adult Intelligence Scale—Third Edition and Wechsler Memory Scale—Third Edition: Technical Manual. The Psychological Corporation, 1997
- 3 曹亦薇. 初中词汇理解能力量表的编制. 心理学报, 1999, 31(2): 215—221
- 4 Mitchell, J. V. The ninth mental measurement yearbook. The University of Nebraska—Lincoln, Lincoln, Nebraska: The Buros Institute of Mental Measurement, 1985

- 5 Burin DL, Jorge RE, Anzaga RA, et al. Estimation of premorbid intelligence: the word accentuation test—Buenos Aires version. J Clin Exp Neuropsychol 2000 Oct; 22(5): 677—85
- 6 申继亮,陈勃,王大华. 成人期智力的年龄特征: 中美比较研究. 心理科学, 2001, 24(3): 344—345
- 7 龚耀先. 中国修订韦氏成人智力量表手册. 长沙: 湖南医科大学, 1989
- 8 国家统计局人口资料和社会科技统计司. 中国人口统计年鉴 2000. 北京: 中国统计出版社, 2000 年
- 9 漆书青. 现代教育与心理测量学原理. 南昌: 江西教育出版社, 1998
- 10 Rust J, Golombok S. Modern Psychometrics: The Science of Psychological Assessment (2nd), Routledge, London, 1999
- 11 曾秀芹,孟庆茂. 项目功能差异及其检测方法. 心理学动态, 1999, 7(2): 41—47
- 12 曹亦薇,张厚粲. 汉语词汇测验中的项目功能差异初探. 心理学报, 1999, 31(4): 460—466
- 13 程灶火,唐秋萍. 韦氏智力量表因素构建的再分析. 中国临床心理学杂志, 1995, 3(3): 185—190

(收稿日期: 2002—06—10)