

MMPI 偏执分量表的计算机自适应测验可行性分析

任芬, 戴晓阳

(深圳大学师范学院心理学系, 广东 深圳 518060)

【摘要】 目的:对 MMPI 偏执分量表的 CAT 研究的可行性进行探索,并在普通人群和病人中检验 CAT 测验的结果。方法:首先采用项目反应理论对 Pa 分量表的单维性进行检验,其次依据项目反应理论挑选项目,形成用于 CAT 的删减版 Pa-D,最后在病人和普通人中验证 CAT 的可行性。结果:①Pa 和 Pa-D 都能满足单维性的假设,可以用 IRT 分析项目。②Pa-D 的项目参数在合理可接受的范围内。结论:Pa 和 Pa-D 均满足单维性,且删除题目未对测验的信息造成明显的影响。采用后验 CAT 模拟可以节省一半的题目。

【关键词】 偏执分量表; CAT; IRT

中图分类号: R395.1 文献标识码: A 文章编号: 1005-3611(2012)01-0015-03

Using the Computerized Adaptive Test to Analyze the Feasibility of MMPI Paranoid Subscale

REN fen, DAI Xiao-yang

Department of Psychology, Shenzhen University, Shenzhen 518060, China

【Abstract】 Objective: In this paper, the item response theory was used to analyze the Pa subscale of MMPI. And the item response theory was used to explore the feasibility of the post-hoc CAT of this subscale. In order to test the effect of the post-hoc CAT, the normal people and the patients were the sample used in this study. **Methods:** There were three steps in total. First, the unidimensionality of the scale was assessed and the item parameters were estimated. Second, according to the parameters, good items were selected to form the Pa-D scale. Third, the effect of the post-hoc CAT was tested in patients and normal people. **Results:** Both Pa and Pa-D could meet the unidimensional assumption, and could be analyzed under the IRT framework. Pa-D item parameters were in the acceptable range. **Conclusion:** Pa and Pa-D are unidimensional, and the test information is not obviously influenced by those deleted items. The post-hoc CAT simulation can save half items.

【Key words】 Paranoid subscale; CAT; IRT

1939年 Hathaway 和 McKinley 以经验法编制的明尼苏达多相人格问卷 MMPI 得到广泛应用,但在应用过程中人们发现 MMPI 一定的问题。首先是施测时间的问题。MMPI 最初的题目在 550 个左右,施测时间 40 分钟到 1 小时。2003 年^[1]国内肖利军等编制了中国简式,将条目从 399 减至 307,然而测试时间并没有显著地减少。Dahlstrom 等^[2]确立了 180 个题目的简版 MMPI-2,结果显示简式的 MMPI-2 只适用于特定的情况。Carlton 等^[3]在神经心理学的背景下检验了 MMPI-2 简式的心理测量学特性,结果显示 MMPI-2-180 可以提供的信息是有限的,在预测临床编码类型上面存有不稳定的偏差。这些简式或多或少会对全量表的信息有所损失。如果能在不损失信息的情况下,对 MMPI 进行高效率的施测是非常重要的。随着计算机技术的发展,计算机自适应测验的出现就满足了上述的要求。

计算机自适应测验(Computerize Adaptive Test, CAT)在测验领域中得到了非常多的应用,最开始是

在能力测验中,后来逐渐用到人格测验中,但是鉴于项目反应理论的前提假设,不论是能力测验还是在人格测验中都是应用到单维的测验中^[4-6]。本研究拟运用项目反应理论(Item Response Theory, IRT)和技术对 MMPI 的偏执(Pa)分量表的项目进行分析,探讨在不违背原测量目标的情况下其项目是否符合 IRT 理论分析的条件,如果符合则运用 IRT 技术筛选和分析项目的各种参数,并在此基础上将这个分量表建成 CAT 测验,并在正常和异常人群中检验新形成的 CAT 测验的效果。

1 对象与方法

1.1 对象

研究对象共 1674 人,其中普通人群 1269 人,精神分裂症病人 405 人。普通人群来源于 2001~2004 年企业求职人员。病人组为精神病院确诊的精神分裂症患者,均符合《中国精神障碍分类与诊断标准第三版》(CCMD-3)有关精神分裂症的诊断标准。

1.2 工具

宋维真修订的明尼苏达多相人格问卷中文版 R

式,包含4个效度量表和10个临床量表,固有题目399题,采用“是”、“否”方式回答,评分方式为0、1计分。本研究只选了Pa分量表40个题目。

1.3 统计方法

单维性检验的方法是全息项目因素分析(full-information factor analysis, FIF),使用的软件为TESTFACT4.0。Pa分量表的分析所用软件为BILOG-MG,最后用POSTSIM2.0进行后验CAT模拟。

1.4 研究过程

第一步是考察Pa分量表是否符合IRT要求的单维性假设,Pa为0、1计分模式,研究将采用FIFA对量表进行因素分析,抽取因子。第二步是挑选性能良好的题目,形成Pa分量表的删减版Pa-D。第三步是用CAT模拟的方法来检验Pa-D的效果。

2 结果

2.1 单维性检验

多数的研究者检验量表的单维性采用探索性因素分析或验证性因素分析,但是MMPI属于依据临床经验法编制的问卷,在使用传统的方法进行因素分析时,得不到清晰的量表结构。在本研究中,用探索性因素分析法分析Pa分量表时,碎石图显示因子分散,主因子不明显。有学者指出,线性因素分析对二元计分的量表进行探索因子时容易产生虚假因子,因为线性因素分析假设项目-特质间是线性相关的,然而对于大多数人格和态度量表而言,这一假设并不能得到满足。FIFA(全息项目因素分析)是一种直接以项目反应理论为基础的因素分析方法,利用项目参数以及被试对项目的反应矩阵,最大限度地利用了反应数据。

对偏执分量表的40个题目进行全息项目因素分析,限定抽取一个因子。结果显示大于1的特征根有2个,第一特征根为11.767,第二特征根为2.272,二者之比为5.178;第一因子解释的方差为27.741%;第一特征根与条目数的比值(第一特征根/条目数)为0.29(大于0.2)。该结果符合大拇指规则要求的3个条件,可以说Pa分量表满足必要的单维性这个要求。各条目的点二列相关在-0.330到0.763之间,因素负荷在-0.336到0.877之间。结果显示40个条目的Pa量表符合单维性检验的结果,可以用项目反应理论进行分析。但是有些条目的点二列相关和因素负荷较低,相应的区分度参数也较低,也就是说这些条目贡献的项目信息量也低。

2.1.1 Pa分量表的原始数据和项目参数的描述性统计 由表2结果可见,Pa分量表的难度参数分布

范围比较广,其中有极端值存在,区分度参数相对比较集中,但是其中有负值存在。在IRT中,通常认为项目的区分度的取值范围为0.3到4.0之间。Pa的40个条目的区分度参数中有4个项目的区分度为负值,有1个项目的区分度参数小于0.3,其余35个项目的区分度参数都落在0.3到4之间。对于测验的难度参数,参数的分布范围越宽越好,难度值分布越宽说明测验的项目可以适用于不同特质水平的被试,测验的适用性越广。在Pa的40个条目中,难度参数小于0的项目有4个,其余难度参数值较大,说明接受项目的被试,特质水平较高。

表1 Pa原始分的描述性统计

分组	人数	原始分 M ± SD	原始分 最小值	原始分 最大值
普通人	1269	17.16 ± 5.34	3	25
病人	405	11.58 ± 2.91	4	40

表2 Pa各个参数的描述性统计

参数	M ± SD	最小值	最大值
难度	0.678 ± 4.454	-24.890	6.634
区分度	0.413 ± 0.566	-0.532	1.298

2.1.2 Pa分量表删减版Pa-D条目的筛选及其单维性检验 由于Pa分量表中有些条目的区分度和难度参数与实际情况不匹配,一般来说项目的难度在-3到3之间,区分度参数在1左右为佳,共删除条目4个,剩余36个,形成Pa分量表的删减版(Pa-D)。对删减版再次进行单维性检验,结果见表3。由表3的结果可见,Pa-D比Pa更能满足单维性的假设,可以使用IRT进行项目分析。此外,Pa-D的36个条目的点二列相关在0.240至0.945之间,因素负荷在0.257至0.803之间,整体上优于Pa分量表。对比Pa的参数情况,Pa-D的难度参数的平均值有所上升,标准差有所减少,原来难度小于0的题目全部被删除,在Pa-D中的所有题目的难度参数都大于0,说明Pa与Pa-D一样难度偏高。Pa-D的区分度参数有所上升,标准差有所减少,说明留下的都是区分度较好的题目。

表3 Pa与Pa-D的单维性检验结果

量表	条目数	第一特征根	第二特征根	二者之比	解释方差
Pa	40	11.767	2.272	5.178	27.741%
Pa-D	36	10.428	1.016	10.259	31.129%

表4 Pa-D的参数估计

参数	M ± SD	最小值	最大值
难度	1.763 ± 1.047	0.296	5.380
区分度	0.681 ± 0.318	0.113	1.320

2.2 Pa与Pa-D的比较

表5列出了原偏执分量表与其删减版经IRT分析后的各种测量信息,以及两个分量表原始分的相关。

从Pa与Pa-D的测验信息量可知,2个版本的量表的信息量均在25以上,说明2个量表的测验性能良好。减少了4个题目,信息量减少了8.471,误差均方根较小。删除的4个题目本身的项目信息较大,对整个测验信息量的贡献较大,删除的题目是区分度为负的题目,但是其绝对值较大,而信息量与区分度参数的平方是成正比的。题目删除后,测验的信息量有所损失,但仍高于测验信息量的标准。因此,题目的删除基本未影响量表对偏执这一特质的测量。

Pa与Pa-D的原始分的相关为0.947,在0.01水平上达到显著性。提示删除版得分与原版分量表得分的一致性非常高。

表5 Pa与Pa-D各种信息量的比较

量表	测验信息	信度	信息量最大时的?	RMS	原始分r值
Pa	75.153	0.924	0.95	0.137	0.947**
Pa-D	66.682	0.781	1.23	0.336	

注:** $P < 0.01$

表6 不同结束规则下的题目使用情况

结束规则	使用题量最小值	使用题量最大值	与全量表相关
18题	18	18	0.926
SE=0.3	30	40	1.00
18题或SE=0.3	18	18	0.926
信息量小于0.1	4	26	0.801

2.3 CAT模拟的结果

为了比较不同结束规则下的题目使用情况,分别选取了4种结束规则,4种结束规则下的题目使用量及与全量表的相关见表6。

由表6可见选取的结束规则为固定18题时是比较理想的,这种结束规则下可以节省一半的题目,且CAT的结果与全量表的得分之间的相关达到了0.926。CAT的结果是有效的,可以在保证测量精度

的条件下,节省较多的题目。

3 讨论

在项目反应理论的框架下分析MMPI的Pa分量表发现其满足拇指原则的3个标准是符合单维性假设的,可以适用IRT来进行分析。

剔除溢出边界的项目参数也就是删除异常的区分度参数和难度参数之后形成之后的Pa的删减版Pa-D,通过与未删减的量表进行比较发现虽然条目减少了,但是并没有对整个测验的信息量造成太大影响。在比较了Pa与Pa-D之后发现二者是测量的同一个特质。

对挑选出的测量性能较优的题目进行后续的CAT模拟研究,选取4种不同的结束规则,发现Pa分量表可以进行CAT测验,可以节省一半的题目。在CAT模拟测验中,CAT得到的特质值与全量表得到的特质值之间的相关在0.01水平上达到显著性,说明CAT测验可以在有效减少施测条目的情况下,依然可以保持测验的性能良好。

参考文献

- 肖利军,苗丹民,王京生,等.中国简式明尼苏达多相人格问卷与新兵社会适应性评价.第四军医大学学报,2003,24(21):1996-1999
- Dahlstrom WG, Robert PA. A Shortened Version of the MMPI-2 Assessment, 2000, 7: 131
- Carlton SG, Cherry AL. MMPI-2 Short Form: Psychometric Characteristics in a Neuropsychological Setting Assessment, 2001, 8: 213
- Forbey JD. Validation of the MMPI-2 Computerized Adaptive Version (MMPI-2-CA) in a Correctional Intake Facility Psychological Services, 2009, 6(4): 279-292
- Biskin BH, Kolotkin RL. Effect of computerized administration on scores for the MMPI. Applied Psychological Measurement, 1997, 1: 543-549
- Roper B, Ben-Porath Y, Butcher J. Comparability of computerized adaptive and conventional testing with the MMPI-2. Journal of Personality Assessment, 1991, 57: 278-290

(收稿日期:2011-08-04)

(上接第14页)

参考文献

- Skinner BF. Superstition in the pigeon. Journal of Experimental Psychology, 1948, 38: 168-172
- 叶浩生. 心理学理论精粹. 福建教育出版社, 2000. 212
- 陈永艳,张进辅,李建. 迷信心理研究评述. 心理科学进展, 2009, 17(1): 218-226
- Tobacyk J. A revised paranormal belief scale, International Journal of Transpersonal Studies, 2004, 23: 94-98
- Zimbardo PG, Boyd JN. Putting time in perspective: A valid, reliable individual-differences metric. Journal of Personality

and Social Psychology, 1999, 77(6): 1271-1288

- 杨凤岗. Quantifying religions in China, seventh annual conference for the social scientific study of religion in China. Beijing, 2010, 7: 26-27
- 孙建国. 当前我国学生中的鬼神迷信思潮. 科学与无神论, 2003, 6: 38-41
- 乐国安. 封建迷信与社会稳定. 赣南师范学院学报, 1998, 1: 73-79
- 韦嘉. 中学生数字偏好现状初步调查与分析. 四川省干部函授学院学报, 2011, 11(1): 3-116

(收稿日期:2011-06-29)