

探索性因素分析——最近 10 年的评述*

范津砚**

(Ohio State University, USA)

叶 斌

(华东师范大学心理咨询中心, 上海, 200062)

章震宇

(上海中医药大学基础医学院心理学教研室, 上海 200032)

刘宝霞

(University of California at Berkeley, USA)

摘 要 目的:(1)介绍国外心理统计学界对探索性因素分析中几个重大问题的基本观点;(2)系统地评述过去 10 年里(1991~2000 年)我国心理学研究者对这一技术的使用情况;(3)强调运用这一技术时值得注意的一些事项,以期这一技术对我国心理学研究发挥更大的作用。方法:作者检索了《心理学报》和《心理科学》在 1991~2000 年间与探索性因素分析有关的文章,对其中以探索性因素分析为主要研究方法的重点文章(feature article)进行了编码,统计了频数及百分比分布。结果:我国心理学研究者在积极使用探索性因素分析这一先进统计技术时还存在一些问题,主要表现在:(1)在确定因素个数时,倾向于机械地依靠某个单一方法来作决定,(2)大量使用正交旋转,(3)过于依赖 SPSS,(4)对因素分析过程中的重要信息/结果报告不够。结论:探索性因素分析在过去十年得到了广泛的应用,如能吸收国外同行的一些观点,探索性因素分析这一技术在我国心理学研究中必将获得更广泛有效的应用。

关键词 探索性因素分析, 评述, 因素个数, 因素旋转。

分类号 B841.2

1 引言

过去十几年里,探索性因素分析在国内心理学研究中开始得到广泛的应用,其在心理学文献中出现的频率也有急速增长之势。这对国内心理学研究起到了一定的推动作用。然而本文作者在与国外的心理统计专家的交流的过程中注意到,在对这一统计技术的理解和具体运用上,他们的观点与我们颇有不同之处。这引起了作者对这一领域的研究兴趣并最终有了本文。本文的目的在于:(1)介绍国外心理统计学界对探索性因素分析中几个重大问题的基本观点;(2)系统地评述过去 10 年里(1991~2000 年)我国心理学研究者对这一技术的使用情况;(3)强调运用这一技术时值得注意的一些事项,以期这一技术对我国心理学研究发挥更大的作用。

2 探索性因素分析中几个重要问题的澄清

2.1 关于因素个数的抽取

因素个数的确定是探索性因素分析中的一个非常重要的决定,因为一旦因素个数确定,模型的拟合度

收稿日期:2002-09-02

* 鉴于验证性因素分析在 1991~2000 年间出现于国内心理学文献中的频率过低(《心理学报》5 次,《心理科学》7 次),无法进行全面评述,故本文只考虑探索性因素分析。文中若不加说明,“因素分析”和“探索性因素分析”在文字上是混用的。

** 第一作者的通讯地址为: Jinyan Fan, I/O Psychology Program, Department of Psychology, Ohio State University, Columbus, OH 43210, USA。电子邮件地址为: fan.61@osu.edu 或 fanjinyan@yahoo.com

也就确定了,不随以后因素的旋转而改变。因素分析的目的在于解释观察变量之间的相关关系,而相关关系是由共同因素造成的,因此抽取的因素越多,对相关的解释就越充分。在极端情况下,对于 $P \times P$ 的相关矩阵,若抽取 $P-1$ 个因素,我们可以“完美”地解释观察变量之间的相关关系。然而这种做法无异于用一个复杂的结构去解释另一个同样复杂的结构,无助于加深我们对原来复杂结构的理解。另一种极端情况是只抽取一个因素。这样虽然最大可能地简化了相关关系,但并没有充分解释大部分的相关关系^[1]。事实上,我们“永远不可能”找到或确定最正确无误的因素个数,我们所能期望的是找到比较适宜的因素个数 m ,以满足以下几个条件:(1)若抽取 $m-1$ 个因素,模型对相关结构的解释度显著下降;(2)若抽取 $m+1$ 个因素,模型对相关的解释度无明显改善,而且(3)经过旋转后,这 m 个因素意义清晰,易于解释^[2]。显然, m 的确定有很强的主观性。为确定因素个数,心理统计学家提出了多种方法,这些方法与模型参数估计(model fitting)方法密切相关。若模型参数估计采用主因素法,则有三种主要确定因素个数的方法:(1)Kaiser法(“特征根 >1 ”法),(2)Scree Test(碎石检验),和(3)Parallel analysis“平行分析法”^[1]。若模型参数估计采用最大似然法,则不但照样可以用上述三种方法,还可以借助许多其它信息来帮助我们做决定,这是因为最大似然法可以提供更多的模型拟合度指标。不过,该方法有个前提假设,即所有观察变量必须呈正态分布。在最大似然法的框架下现在比较流行的有3种确定因素个数的方法:(1)perfect fit test(完全拟合检验)^[1];(2)Tucker—Lewis法^[3]和(3)RMSEA法^[4]。面对诸多确定因素个数的具体方法,我们应当如何取舍,从而最终确定因素个数呢?简单的回答是根据多种信息,综合决定。Browne提出了以下的程序^[1]:(1)考虑研究者在理论中是否事先假设了因素个数;(2)考虑一些简单方法,如Kaiser法、Scree Test所提供的信息;(3)考虑由最大似然法所产生的模型拟合度的信息;(4)根据以上三方面的信息,将可能的因素个数压缩到一个比较小的范围内;(5)根据(4)分别抽取不同个数的因素,比较旋转后因素负荷的可解释性,以作出最终决定。这是一个相对比较全面的程序,研究者可以批判性地采用。总之,因素个数的确定并不存在着“唯一”、“正确”、“客观”的答案,我们需要利用各种信息进行主观判断。最坏的做法是根据某一个“客观的”、“无须思考的”单一方法来确定因素个数(如Kaiser法或统计软件上预设的某个“统计分析键”),因为这时研究者将至为宝贵的“控制权”、“思想权”让给了“死”的规则或电脑软件。

2.2 关于因素的旋转

因素个数确定之后,就要进行因素旋转。文献中出现了很多种因素旋转的方法,这些方法总体上可以分为两大类,即正交旋转和斜交旋转。正交旋转假设各个因素之间是相互独立的,没有相关,而斜交旋转则没有这一假设。很多研究者采用正交旋转,因为他们认为正交旋转比较简单,概念上也比较清晰。有些研究者甚至认为因素分析的结果就是抽取 m 个相互独立的因素。这些看法是片面的^[1]。首先因素分析模型并没有规定因素之间必须独立。其次,心理学研究中很多概念或范畴(construct)之间是相关的,硬性将它们限制成相互独立的因素不见得符合事实。再次,正交旋转人为地设置了多余的限制,导致了旋转后的因素负荷矩阵的“简单性”、“清晰性”、“易解释性”,以及整个模型的拟合度都要比相同条件下斜交旋转的结果要差。Fabrigar及其同事^[3]在最近的一篇文章中用具体的数据清楚地显示了这一点。最后斜交旋转比正交旋转能提供更多的信息,即因素之间的相关矩阵。这一信息很可能与研究者的理论是密切相关的。另外如果因素之间相关较高的话,还意味着可能存在着“高阶”因素,因此还可以进行更高阶的因素分析。这一点正交旋转是办不到的,因为它已事先限制了因素之间是互相独立的。因此,在绝大多数情况下,因素旋转时采用斜交旋转是更好的选择。

2.3 关于简单结构

因素旋转的目的是将初始因素负荷矩阵旋转至“简单结构”以便于解释。现在占统治地位的“简单结

构”是由 Thurstone 于 1947 年提出的^[5]。Thurstone 的这一定义常常被研究者误解，其中最流行的误解是，不少研究者认为每一观察变量只能有一个大的因素负荷，同一行的其它负荷必须为 0*。这种理解的错误在于（1）因素分析模型完全允许一个观察变量受多个因素影响；（2）Thurstone 的定义只不过指出，对于旋转后的因素负荷矩阵中，一个观察变量受多个因素影响的情况不能多次出现。与这一误解密切相关，不少研究者在报告旋转后的因素负荷矩阵时，往往只报告绝对值大的负荷，如大于 0.40 的负荷；同时若同一行出现两个负荷较大的负荷，往往只报告绝对值较大的那个负荷。这种做法形成了因素负荷矩阵“开天窗”（未报告全部数据）的现象。这种看上去似乎清晰、简单的“天窗”结构，有可能对此报告的其他研究者产生误导，应当加以避免。最好的做法是将因素负荷矩阵的全部数据加以报告。

3 对过去 10 年（1991~2000 年）探索性因素分析使用情况的评述

本文作者从过去 10 年在《心理学报》和《心理科学》上发表的文章中，检索出以探索性因素分析为主要方法的文章，并对这些文章进行了系统分析。

3.1 方法

3.1.1 样本的确定

样本全部来自发表在《心理学报》和《心理科学》上的文章，因为这两本杂志是国内心理学界的权威期刊，反映了国内心理学发展的水准。本文作者检索了这两本杂志 1991~2000 年的所有文章，凡是被认为运用了探索性因素分析的文章均被列入初始样本。鉴于不少文章出现了多个因素分析，本文作者决定评估建立在因素分析的水平上，而非文章的水平上。从表 1 中可以清楚地看到，在过去 10 年里探索性因素分析的应用有快速增长的趋势。为了全面、客观、公正地考察我国心理学研究者使用这一统计技术的状况，在筛选最终样本时，特制定了以下的入选标准：（1）只评估重点文章（feature article），因为短篇报告受篇幅所限，不可能详细描述因素分析的过程；（2）同理，只评估将探索性因素分析作为主要研究方法的文章；（3）关于因素分析的方法论文章不作评估。若在筛选样本时作者意见不一则通过讨论达成一致。经过第二、第三作者的独立筛选和共同讨论，形成了包含 95 个因素分析的最终样本。其中《心理学报》有 36 个，《心理科学》有 59 个。

表 1 探索性因素分析的频率分布

刊名	1991~1995	1996~2000	增长率 (%)
心理学报	17	27	58.8
心理科学	36	54	50.0

3.1.2 样本的编码

本文第一、第四作者分别对最终样本中的每一个因素分析进行独立编码。编码的内容如下：问卷的信度、项目数与因素数之比、样本大小与项目数之比、样本大小、因素个数的确定、因素的旋转、软件的选择和研究结果的报告。编码的一致率达 91.7%，当出现不一致时，则根据上述原则经讨论形成一致意见。

3.1.3 统计方法

用简单的频数分布和百分比分布来揭示有关问题。

* 0 指的是那些数值小，接近于 0 的负荷。

3.2 结果与评估 (见表 2)

表 2 结果汇总

	心理科学		心理学报		总和	
	N	%	N	%	N	%
问卷信度						
<0.6	1	1.5	0	0	1	1.1
0.6~0.7	1	1.5	0	0	1	1.1
0.7~0.8	8	12.3	5	13.9	13	13.7
0.8~0.9	8	12.3	3	8.3	11	11.6
>0.9	1	1.5	2	5.6	3	3.2
未报告	40	61.5	26	72.2	66	69.5
项目数：因子数						
<3	15	23.1	3	8.3	18	18.9
3~4	12	18.5	4	11.1	16	16.8
4~5	13	20	2	5.6	15	15.8
5~6	1	1.5	4	11.1	5	5.3
>6	18	30	22	60	40	42.1
未报告	0	0	1	2.8	1	1.1
样本数：项目数						
<5	12	18.5	8	22.2	20	21.1
5~10	17	26.2	8	22.2	25	26.3
>10	27	41.5	19	52.8	46	48.4
未报告	3	4.6	1	2.8	4	4.2
样本规模						
<100	8	12.3	1	2.8	9	9.5
101~200	14	21.5	9	25	23	24.2
201~300	13	20	4	11.1	17	17.9
301~400	8	12.3	2	5.6	10	10.5
>400	13	20	20	55.6	33	34.7
未报告	20	46.2	15	41.7	45	47.4
因子数目的确定						
特征根>1	25	38.5	11	30.6	36	37.9
碎石检验	6	9.2	21	58.3	27	28.4
模型拟合度	1	1.5	1	2.8	2	2.1
理论假设	4	6.2	1	2.8	2	2.1
因子可解释性	2	3.1	11	30.1	13	13.7
方差解释率	18	27.7	4	11.1	22	23.2
多种方法	8	12.3	19	52.8	27	28.4
其它方法	1	2	0	0	1	1
未报告	19	29.2	9	25	28	29.4
因子的旋转						
正交旋转	32	49.2	26	72.2	58	61.1
斜交旋转	4	6.2	2	5.6	6	6.3
未旋转	8	12.3	1	2.8	9	9.5
未报告	16	24.6	7	19.4	23	24.2
软件的选择						
SPSS	24	36.9	8	22.2	32	33.7
SAS	0	0	1	3	1	1
其它	3	4.6	2	5.6	5	5.3
未报告	32	49.2	25	69.4	57	60
结果的报告						
相关矩阵	1	1.5	0	0	1	1.1
特征根	28	43.1	12	33.3	40	42.1
共同度	13	20	4	11.1	17	17.9
旋转后的因子负荷矩阵	42	64.6	30	83.3	72	75.8
负荷矩阵“开天窗”*	22	52.4	25	83.3	47	65.3

注：*该行百分比的计算是以报告了旋转后的因子负荷矩阵的因素分析的数目为基数的。

3.2.1 问卷信度

这是因素分析中的一个重要指标,问卷信度是共同度的上限^{*}。若信度不高,共同度就不会高,其后果是参数(因素负荷)的标准误较大,即稳定性较差。换言之,对于同一人群中不同样本进行因素分析得到的因素负荷相差较大。表 2 显示在报告了问卷信度的因素分析中,其信度是比较好的,一般都大于 0.7。但是近 70%的因素分析没有报告信度,对这部分研究我们无从了解其因素负荷的稳定性。

3.2.2 样本

样本规模直接影响因素分析参数(因素负荷)的稳定性。虽然原则上样本越大因素负荷越稳定,但是样本越大,所消耗的资源也就越多。研究者通常希望既能获得足够的样本以满足因素分析的要求,又能节约有限的资源。目前比较流行的评估样本的指标有 3 个:(1)绝对样本规模,200 为最低要求;(2)样本数与项目数之比,一般要求大于 5;(3)项目数与因子数之比,要求大于 4^[6]。从表 2 中可以看到,62%的因素分析,其绝对样本规模超过了 200,75%的因素分析的样本数与项目数之比大于 5,63%的因素分析的项目数与因子数之比大于 4。总体来说,大部分因素分析的样本规模是适当的。但是也应当看到,还有近 40%的因素分析样本过小,对于这部分因素分析,其结果的稳定性应受到质疑。

3.2.3 因数个数的确定

总体上看最流行的确定因素个数的方法是 Kaiser 法、碎石检验和方差解释率。有意思的是具体的确定因素个数的方法在两本杂志中的分布是不一样的。发表于《心理科学》上的因素分析用 Kaiser 法的占 38.5%,用方差解释率的占 27.7%,碎石检验占 9.2%;而发表于《心理学报》上的因素分析用碎石检验的占 58.3%,Kaiser 法占 30.6%,因子可解释性占 30.1%和方差解释率占 11.1%。前文提过,每种方法各有优劣,研究者应利用多渠道信息综合决定因素个数。根据这一标准,《心理科学》和《心理学报》中分别仅有 12.3%和 52.8%的文章在使用因素分析方法时考虑了多重信息来决定因素的个数。另外,近 30%的因素分析没有报告如何确定因素个数这一重要的问题。

3.2.4 因素的旋转

总体上仅 6.3%的因素分析使用斜交旋转,高达 61.1%的因素分析使用的是正交旋转,近四分之一的因素分析没有报告因素旋转的方法。

3.2.5 软件的选择

现在一般大型软件包如 SPSS、SAS、SYSTAT 等都可以进行因素分析,而表 2 显示,SPSS 显然占了统治地位,1/3 的因素分析使用了该软件。如果不考虑未报告软件选择的因素分析(60%),SPSS 所占比例是相当高的。事实上,SPSS 是存在着一些缺点的,比如全部选择“缺省值”的话,使用者实际上将进行的是主成分分析,而非因素分析,所以研究者应避免因为对统计知识理解不深入而对 SPSS 产生误用。本文作者在此特别推荐一种新近开发的 CEFA 的软件^{**}。该软件的优点之一是不设“缺省值”,强迫使用者就几大关键问题进行思考,作出抉择,否则无法运行。这将促使研究者深入思考,而非盲目地点击几个键,便以为大功告成。

3.2.6 研究结果的报告

科学研究结果的报告应当全面客观,应当有利于知识的积累以及其它研究者对自己的研究进行检验和

* 信度=(共同方差+特定方差)/总方差,而共同度=共同方差/总方差,所以信度=共同度。

** CEFA 是 Comprehensive Exploratory Factor Analysis 的简称。它是由美国俄亥俄州立大学心理系的 Browne 教授开发的,可以从以下网址免费下载: <http://quantm2.psy.ohio-state.edu/browne/>

批判。为达到上述目的,研究者在报告因素分析的结果时,应当全面报告相关矩阵、特征根、共同度以及完整的旋转后的因素负荷矩阵^[1]等数据。从表 2 中,我们可以看到,仅 1.1%的因素分析报告了相关矩阵,17.9%的因素分析报告了共同度,42.1%的因素分析报告了特征根,75.8%的因素分析报告了旋转后的因素负荷矩阵,其中仅 26.3%的因素分析在报告了旋转后的因素负荷矩阵时没有数据缺损。虽然《心理学报》中的因素分析报告旋转后的因素负荷矩阵的比例(83.3%)比《心理科学》(64.6%)高,但其中数据缺损的比例(83.3%)也比后者(52.4%)高。为了检验因素分析的数据的可靠性,我们需要知道观察变量相关矩阵。在全部论文的 95 个因素分析中,按照这一标准,仅有一个因素分析允许其它研究者对其过程进行检验。这种状况显然是需要加以改善的。

4 正确运用探索性因素分析需要注意的事项

针对上述种种问题,本文作者认为研究者在使用探索性因素分析的过程中应注意如下事项:(1)在问卷设计阶段,对于理论中假设的每一个因素都应确保有足够的项目来充分描述它;(2)在问卷测试阶段,应对项目慎加取舍,保证最终的问卷有较高的信度;(3)在确定抽取因素个数时,应根据多渠道信息作合理的决定,而不是依赖某一机械的、貌似客观的方法;(4)在因素旋转时,无特殊理由应采取斜交旋转;(5)在报告结果时,必须报告样本大小和相关矩阵,并同时报告特征根、共同度,以及完整的旋转后的因素负荷矩阵,以方便其它研究者评判;(6)在软件的选择上也可考虑使用其它软件,而不过分依赖 SPSS,尤其要防止对统计软件的误用。

5 结论

本文介绍了国外心理学界对探索性因素分析中的几个重要问题的基本观点,系统地评述了在过去十年(1991~2000年)中我国心理学研究者应用这一统计技术的情况。总的结论是,在过去十年里虽然探索性因素分析得到了越来越广泛的应用,但在应用质量上还是存在一些问题的,其主要表现有:(1)在确定因素个数时,往往只依靠某一个方法机械地作出决定,没有综合考虑多渠道的信息;(2)在进行因素旋转时,大量使用正交旋转,而忽视使用斜交旋转;(3)在使用软件时,过多地依赖 SPSS,而且往往全部选择“缺省值”进行运算,有可能导致统计方法的误用;(4)在报告研究结果时往往不能对全部的信息和结果(如问卷信度、模型的选择、因素个数的确定、因素的旋转方式、软件的选择、数据的相关矩阵、共同度、旋转后的因素负荷矩阵等)加以报告,而这些信息对于研究而言其实是相当重要的。针对这些问题,本文强调了正确使用这一技术应当注意的一些事项,同时建议(1)加强心理统计领域的国际学术交流,及时将国际上最新最先进的统计技术介绍到国内并应用到心理学研究中去;(2)国内高校心理学系和心理学研究机构要努力提高统计学课程的质量,以提高学生应用复杂的、先进的统计技术的能力;(3)有关机构应考虑明确规定研究报告在提供数据资料时的规范,比如必须提供全面充分的数据资料,以便同行复核。美国心理学会(APA)对此是有明确规定的,在这一点上我们应考虑学习他们的做法,与国际接轨,以建立有效的同行监督机制,进一步提高学术研究的质量。

参考文献

- [1] Fabrigar L R, Wegener D T, MacCallum R C, Strahan E J. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 1999, 4: 272~299
- [2] Velicer W F, Jackson D N. Component analysis versus common factor analysis: some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, 1990, 25: 1~28

- [3] Tucker L R, Lewis C A. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 1973, 38: 1~10
- [4] Browne M W, Cudeck R. Alternative ways of assessing model fit. *Sociological Methods and Research*, 1992, 21: 230~258
- [5] Thurstone L L. *Multiple factor analysis*. Chicago: University of Chicago Press, 1947
- [6] MacCallum R C, Widaman K F, Zhang S, Hong S. Sample size in factor analysis. *Psychological Methods*, 1999, 4: 84~89

Exploratory Factor Analysis: A Literature Review Between 1991 and 2000

Fan Jinyan

(Department of Psychology, Ohio State University, Columbus, OH 43210, USA)

Ye Bin

(Psychological Counseling Center, East China Normal University, Shanghai, 200062)

Zhang Zhenyu

(Division of Psychology, College of Basic Medicine, Shanghai Chinese Medicine & Pharmacy University, Shanghai, 200032)

Liu Baoxia

(School of Optometry, University of California at Berkeley, Berkeley, CA 94720)

Abstract: The purposes of this article was to: 1) clarify some confusing issues in exploratory factor analysis (EFA) among Chinese psychologists, and 2) critically review how Chinese psychologists have been applying EFA in their research during the last decade (1991~2000). In the first part of this article, the following issues were discussed: (a) estimating the number of common factors, (b) the rotation problem, and (c) the myth of simple structure. The authors searched *Psychological Science* and *Acta Psychologica Sinica* between 1991~2000 and identified 95 EFAs from feature articles that used EFA as their major quantitative method. All the identified EFAs are coded on the following characteristics: (a) scale reliability, (b) ratio of item number to factor number, (c) ratio of sample to item number, (d) sample size, (e) methods to determine number of factors, (f) rotation method, (g) software, and (h) report practice. Simple statistics such as, frequency and percentage, revealed the major problems many Chinese psychologists have in employing EFA include: (a) exclusive reliance on some mechanical methods to determine factor numbers, (b) using orthogonal rotation without justification, (c) overuse of SPSS, and (d) failure to report major information and results so that peer-evaluation is essentially impossible. The possible reasons for the current situation are discussed.

Key words: exploratory factor analysis, review, factor number, and factor rotation.