

诊断认知策略的几何类比推理测验题的特征及其编制研究^{*}

刘声涛^{1 2} 戴海崎¹

(1. 江西师范大学 教育学院, 南昌 330027; 2. 福建师范大学 教育科学学院, 福州 350007)

摘 要 :该文在新一代测验理论的视角下,以几何类比推理测验为研究对象,以认知策略的诊断为目的,研究更能引发被试的规则构建策略或选项剔除策略的项目的特征。研究结果表明,项目中元素数量是影响被试使用规则构建策略或选项剔除策略的关键因素,元素数量越多时越倾向于用规则构建策略,而元素数量越少时倾向于用选项剔除策略。研究结果可直接用于测验设计,使得测验能更多在某一策略的框假下分析描述被试特征。

关键词 :新一代测验理论 认知诊断 几何类比推理 策略

中图分类号 :B841.2

文献标识码 :A

文章编号 :1003-5184(2007)02-0069-04

1 引言

新一代测验理论^[1]的核心是将测验的能力水平的研究推向测验的认知水平的研究,以实现测验的认知诊断功能,提高测验效度。

在测验的能力水平的研究中,困扰心理测量学研究者的问题是测验测量的究竟是什么。因素分析方法分析出的特质还仅是一个“统计结构”,并不能给出满意的答案。对于智力,就有研究者不无讥讽地指出“智力就是智力测验测得的那些内容”。人们并不知道智力是什么,也不知道智力测验测的是什么,只能用一个概念去解释另外一个概念。

认知心理学的发展让心理测量学研究者看到了摆脱困扰的希望所在,心理测量学研究者可以学习认知心理学的研究方法,以用于测验内容的分析。也可以直接吸收认知心理学的研究成果。“认知和测量相结合”^[2]将测验置于认知水平的研究中,能弥补心理测量能力水平研究的局限,使得研究者能够明确测验对信息加工、策略、成分、知识等认知变量实施了怎样的测量,以及怎样的测验能更好地测量各认知变量。这样的测验效度更高,能实现对被试的认知诊断,即描述被试的内部心理特征^[3]。

被试在问题解决中使用不同的认知策略,这对于心理测量来讲可能是受欢迎的,也可能是不受欢迎的。因为不同的策略背后有不同的心理机制。假如在某测验项目中,被试在解决问题时有两种策略,一是空间策略,一是言语策略。那么,测验对用空间策略的被试测到的是空间能力,而对用言语策略的被试测到的是言语能力。这种情况说明,确定被试

在问题解决策略中的差异,才能准确定义被试能够做什么。反过来,从测验目的来说,如果测验的目的是被试解决该类问题的总的能力,那么就不必区分被试到底用的是什么策略;如果测验的测量目的是测量到被试的空间能力而不是测量言语能力,在这种情况下,因为不能控制被试的策略选择影响因素,所以只能从测验设计出发,去选择那些更能引发被试空间策略的项目。不管是哪种情况,在认知诊断框架中,测验编制者都应该说清楚测验中的策略选用情况,怎样的项目可以引发怎样的策略。这一点对于解释测验和解释被试非常重要,这也是测验高效度的标准之一,也是实现测验在认知策略方面的认知诊断,即对被试认知策略使用的特征的深入描述的前提。

几何类比推理测验是心理测验中常用的一种测验形式,这种类型的测验所测的能力处于一般智力的核心^[4]。目前在几何类比推理的测量实践中,仍旧是给被试一个几何类比推理上的总的能力值,并没有更多、更深入的质的分析。实质性心理学对几何类比推理的认知策略有较深入的认识。Bethillo-Fox, 等^[5]的研究表明,在几何类比推理测验中有两种问题解决策略:规则构建策略与选项剔除策略。规则构建策略指被试根据题干主动构建项目的理想答案,然后把理想答案和各个选择项相比较,选出和理想答案最符合的选项。选项剔除策略是被试在作出选择前,反复地看题干和选项,按一定的规则剔除不合适的选项。

Vodegel-Matzer^[6]设计了项目分屏测试的方法

^{*} 基金项目:高等学校博士学科点专项科研基金项目(20050414001)。

来区别被试使用了规则构建策略还是选项剔除策略。分屏测试的方式是:分别在前后两屏中呈现题干和选项,被试可以在两屏之间来回点击。对于规则构建策略,Vodgel - Matzen 预测:被试在题干屏上会用更长的反应时,在选项屏上会用相对少的反应时,一旦进入选项屏,就很少或没有回看题干屏。对于选项剔除策略,Vodgel - Matzen 预测:被试在题干屏上会用相对少的反应时,在选项屏上会用相对长的反应时,而且会频繁地回看题干屏。Vodgel - Matzen 的研究证明了这一点。

在上述研究的基础上,可能对被试在几何类比推理测验中的作答表现获得比一个总的能力值更多的测量信息。从项目来讲,如果能区分更能引发被试的规则构建策略项目和更能引发被试的选项剔除策略的项目,那么就能在测验中按测验目的提供项目,以便在某一策略框架内更好地分析、描述、比较被试。

项目刺激特征(item stimulus feature)即项目构成中,对被试解决问题的认知加工有影响的那些特征。Mulhullan, T. M.^[7]将几何类比推理项目的项目刺激特征区分为 A 元素的数量及 A 到 B 变换的数量,认为这两方面的特征影响项目难度。所谓元素即指构成 A 项的基本组成部分,如点、线、圆等,这些基本组成部分可以独立地变换。在 Mulhullan, T. M. 等的研究的基础上,Whitely, S. E.^[8]认为可以进一步将 A 到 B 的变换分为两类,一类是移置(displacement transformations),一类是变形(distortion transformations)。所谓移置是指一个或多个元素的方向发生变化,具体的包括位置置换、旋转、翻转,这类的操作被认为有视觉的或心理的想象。所谓变形是指元素要经过某种变化才成为另外的状态,具体包括数量、大小、颜色、形状的变化,这类的操作不需要心理想象。这样共有七种变化,再加上 A 元素的数量,项目刺激特征共分为八个方面。

研究在前人研究的基础上系统研究几何类比推理测验项目策略选择和项目刺激特征之间的关系。目的是提供更易引发某一策略的项目,以便在某一策略框架内探讨被试的心理特征。

2 方法

2.1 被试

随机抽取 6、7、8、9 年级被试共 79 名,男生 38 名,女生 41 名。这些被试都上过计算机课,会电脑的基本操作。被试在学校的计算机机房中接受测试。

2.2 实验工具及程序

研究所用的几何类比推理测验项目是 A :B :: C (五个选项)的格式,自编了 40 道几何类比推理测验项目及几何类比推理分屏测试计算机系统。该系统由指导语、模拟测试、正式测试三部分组成。模拟测试模块让被试熟悉操作过程,以免被试因操作失误而影响测试成绩。40 个项目出现在正式测试部分。在题干屏中,仅出现项目的题干,题干下有三个按钮,分别为“题干”、“选项”、“下一题”。在题干屏中,仅有选项按钮可激活。点击“选项”按钮即可进入该题的选项屏,选项屏给出了项目的五个选项。选项下也有“题干”、“选项”、“下一题”三个按钮,此时“题干”、“下一题”两个按钮可激活。被试可以点击“题干”按钮再次进入该题的题干屏,也可以选择“下一题”按钮再次进入下一题。

几何类比推理分屏测试系统记录被试在每个项目的题干、选项上每一次注视的时间,以及被试的选择项。如,计算机可记录下被试在某项目的题干屏注视了 5 秒,接下来进入选项屏注视了 6 秒,再回到题干屏注视了 7 秒,又进入选项屏注视了 8 秒,最后选择了 A。

在被试进入计算机测试程序之前,首先向被试介绍几何类比推理项目及项目的作答方法,并说明在电脑上题干和选项是分前后两屏呈现。待学生都听明白后给每个被试一个事先准备好的编号,登陆进入几何类比分屏测试系统。被试阅读指导语后,再进行模拟测试,最后进入正式测试。

2.3 实验数据的处理与分析

在计算机测试系统记录下的每个被试在题干屏和选项屏的每一次注视时间的基础上,分别统计每个被试在题干屏和选项屏的注视时间的总和以及被试进入选项屏后回看题干屏的次数,即回看次数。按照 Vodgel - Matzen^[6]的研究,依据下列标准判断每个被试在每个项目上所用的策略:1)如果总的反应时小于或等于 4 秒,则认为被试在猜测;2)如果回看次数小于或等于 2 次,则认为是规则策略;3)如果在题干上的反应时短于选项上的反应时,则认为是选项剔除策略;4)如果在题干上的反应时长于选项上的反应时,则认为是规则策略。

判断每个项目上每个被试使用的策略,以更高频率被使用的策略做为该项目所能引发的策略,这样就可将项目按所能引发某种策略分为规则构建策略项目组 and 选项剔除策略项目组。基于被试在分屏

测试中的作答模式用逻辑斯谛克模型估计各策略组项目的难度,估计软件是江西师范大学统计测量开发室自主研发的 ANOTE 软件。

按 Whitely S.E.^[8]对几何类比推理的项目刺激特征的分类对研究所用到的 40 个项目的项目刺激特征做详细的记录,以项目在每个方面特征的取值

来记录每个项目的项目刺激特征。以项目 1 为例,它的 A 项所含元素的数量是 1,从 A 到 B 项发生了一个变换,即翻转。则元素数量和翻转取值为 1,其它特征取值为 0。项目 1 的项目刺激特征记录如下所示。

表 1 项目 1 的项目刺激特征记录

序号	元素数量	变 换						
		移 置			变 形			
		置换	翻转	旋转	数量	大小	颜色	形态
		1	2	3	4	5	6	7
1	1	0	1	0	0	0	0	0

其它的项目也以这种方式记录项目刺激特征。通过建立各策略项目组项目难度对各刺激特征的回归方程来分析项目刺激特征对策略选择的影响。

3 结果与分析

建立规则构建策略组项目难度对项目刺激特征的回归方程。使用的回归方程建立方法为向后回归,得到的回归方程的常数及影响因素的有关信息如下表所示。

表 2 规则构建策略组项目难度对项目刺激特征的回归分析

常数及影响因素	系数值	p
常数	- 1.174	0.01
元素数量	0.196	0.09

回归方程表明,元素数量对规则构建策略组的项目难度在 0.09 的显著性水平上有影响。

建立选项剔除策略组项目难度对项目刺激特征的回归方程。使用的回归方程建立方法为向后回归,得到的回归方程的常数及影响因素的有关信息如下表所示。

表 3 选项剔除策略组项目难度对项目刺激特征的回归分析

常数及影响因素	系数值	p
常数	- 1.168	0.00
元素数量	0.222	0.00
数量	- 0.181	0.08
大小	0.269	0.00

回归方程表明,元素数量、大小变化对规则构建策略组的项目难度在 0.00 的显著性水平上有影响,数量变化在 0.08 的显著水平上有影响。

综合表 2 和表 3 发现,元素数量是规则构建策略和选项剔除策略共同的影响因素。进一步统计各元素数量值的项目使用各策略的概率,结果如下表所示。

表 4 不同元素数量项目使用各策略百分比

项目中元素数量	1	2	3	4
使用规则构建策略的百分比	0	0	28	86
使用选项剔除策略的百分比	100	100	72	14

从表 4 可看出,当项目中元素数量等于或少于 3 时,被试倾向于用选项剔除策略;当项目中元素数量等于 4 时,被试倾向于选用规则构建策略。

4 讨论

4.1 与已有研究比较

研究发现,元素数量是导致被试选用不同策略的主要因素,当元素数量越多时,被试更倾向于使用规则构建策略。

实验研究结果和 Diehl,K.D.^[9]分析被试在瑞文标准推理测验中的表现时的发现是一致的。他发现,当项目越复杂时,回看的次数越少。回看次数少正是体现了规则构建策略的使用。研究认为,当元素少时,错误的答案很容易剔除,因为不符合的地方很容易看出,因此被试用选项剔除策略能力节省心理资源。当元素数量多时,不容易看出不符合的地方,选项剔除实际上是要消耗更多的心理资源。因此,被试在面临更多元素的项目时,使用规则构建策略,更快捷。

实验研究在刺激特征的层面发现了项目策略选用的主要影响因素:项目中元素的数量,这一结论能直接用于测验设计。

4.2 对测验设计的启发

当项目中元素数量少时,被试倾向于用选项剔除策略,当项目中元素数量多时,被试倾向于选用规则构建策略。另外,进一步数据分析还表明当项目中元素数量等于或少于 3 时,被试倾向于用选项剔除策略;当项目中元素数量等于 4 时,被试倾向于选

用规则构建策略。这两个数值可以做为编写不同策略组项目的参考值。

4.3 在被试特征分析中的应用

尽管项目可以分为不同的策略组,但这仅是从项目的角度,从统计学意义上讲的。当不同的被试面临同一批项目时,仍然会有策略选用的差异,这时是加入了被试的因素。

编写不同策略组项目的含义不是编写每个被试都用同一策略的项目,这实际上是不可能完成的任务。实验研究一开始做的数据分析已表明,在项目上的策略选用不仅受到项目的影响还受到被试的影响。编写不同策略组项目的含义是编写出的项目更可能被被试用某种策略完成。前面的数据统计分析也能说明,确实不同的项目更能引发的策略不一样。因此编写出更可能被被试用某种策略完成的项目是可以完成的任务。

这样做的意义在于,能够设计出更可能引发某种策略的项目,在这样一种条件下去深入描述被试的特征。实际测试环境下,几何类比推理可能更多的是要测被试主动构建规则的能力。因此,更多地是希望在被试规则构建的策略下评价与分析被试,假如为这个目的施测,而编写的项目都是更能引发选项剔除策略的项目,那么就可能造成诊断评价的失真,造成了工作上的难度。

参考文献

1 Mislevy R J. Foundations of a new test theory. In :N. Frederiksen et al. ,Eds. Test theory for a new generation of tests. Hillsdale, NJ :LEA ,1993. 19 – 39.

2 漆书青,戴海崎,丁树良. 现代教育与心理测量学原理. 南昌:江西教育出版社,1998. 307.

3 刘声涛,戴海崎,周骏. 新一代测验理论—认知诊断理论的源起与特征. 心理学探新,2006,4: 73 – 77.

4 Snow R E. Aptitude processes. In :R. E. Snow,Pat – anthony Federico, Montague W E, Eds. Aptitude, learning and instruction – Volume 1 :Cognitive process analyses of aptitude. Lawrence Erlbaum associates, 1980. 27 – 63.

5 Bethell – Fox C E, Lohman D F, Snow R E. Adaptive reasoning : componential and eye movement analysis of geometric analogy performance. Intelligence, 1984, 8: 205 – 238.

6 Vodegel Matzen L. Performance on Raven 's progressive matrices : what makes a difference ? Unpublished dissertation, universiteit van Amsterdam. Amsterdam, Netherlands, 1994.

7 Mulholland T M, Pellegrino J W, Glaser R. Components of geometric analogy solution. Cognitive psychology, 1980, 12: 252 – 284.

8 Whitely S E, Schneder L M. Information structure on geometric analogies :A test theory approach. Applied Psychological Measurement, 1981, 5: 383 – 397.

9 Diehl K A. Algorithmic item generation and problem solving strategies in matrix completion problems. Unpublished Dissertation, university of Kansas, 2002.

Research on the Cognitive Strategy Diagnosis
Function Development of Geometric Analogy Test

Liu Shengtao^{1, 2}, Dai Haiqi¹

(1. Education College, Jiangxi Normal University, Nanchang 330027; 2. Education College, Fujian Normal University, Fuzhou 350007)

Abstract New generation of test theory relies on the psychological research, it aims at cognitive diagnosis. Geometric analogy tests the ability of the core of general intelligence, where substantive psychology has done in – depth research already. This research of this thesis focuses on the cognitive strategy of geometric analogy. It answers a detailed question : 1. what kind of item can arose the rule – constructing strategy of the subject, or rule – deleting strategy? To the question, it 's found that the number of element of the items is the key factor which influences the subject 's rule – constructing strategy or rule – deleting strategy. The number of element tends to rule construction strategy when it gets more and more, to the item deleting strategy when it gets less and less. The results of the research can be used directly in the test design and make the test better construe thorough description of the subject 's properties under the framework of some strategy.

Key words test theory for a new generation of tests ;cognitive diagnosis ;geometric analogy test ;cognitive strategy