

# 平衡秤任务复杂性的事前与事后分析

张 丽 辛自强

(北京师范大学发展心理研究所,北京 100875)

**摘 要:**任务复杂性的分析和评估是心理测量学和认知心理学都非常关注的重要主题。以 264 名小学四、五、六年级儿童为被试,平衡秤任务为研究材料,考察任务在未旋转时第一个因素上的载荷(事后分析)能否作为评价任务复杂性的一个指标,以及关系一表征复杂性模型对平衡秤任务复杂性分析(事前分析)的有效性两个问题。结果表明:平衡秤任务施测后所得所有项目的因素载荷与其难度之间没有显著正相关,即因素载荷的高低没有反映平衡秤任务复杂性的多少;而基于关系一表征复杂性模型对任务的事前分析所确定的任务等级复杂性和知识经验对任务难度的解释率为 95.0%,可见,关系一表征复杂性模型提供的分析任务复杂性的思路和方法是较为合理的。

**关键词:**关系一表征复杂性模型;任务复杂性;任务难度;因素分析

## 1 问题提出

任务复杂性是任务本身具有的特点,反映了问题解决必需涉及的元素及其关系的复杂程度,它是影响个体认知能力测量和评估的重要因素,因此任务复杂性的分析和评估是心理测量学和认知心理学都非常关注的重要主题。关于任务复杂性的分析,测量学家们通常采取事后分析的思路,如使用通过率和反应时作为任务复杂性的指标。一般来讲,个体在任务上的通过率越低,任务难度则越大,任务复杂性则越高;个体在任务上的反应时越长,任务复杂性也越高。

近年来,有些研究者提出依据因素载荷的大小来评估任务的复杂性。根据 Crawford 的看法,斯皮尔曼提出一般智力是共同因素 g 的观点,直接来自他的一个发现,即不同测验在一般智力(即 g 因素)上的载荷不同,而任务复杂性概念正是用来解释此现象的<sup>[1]</sup>。具体讲,不同任务在一般智力因素上的载荷不同是因为任务的复杂性不同。这启发人们思考的一个问题是,任务的复杂性程度是否可根据其在一般智力因素上的载荷来判断呢? Stankov 及其同事对有关任务复杂性和智力关系的问题进行了很多探讨<sup>[2~5]</sup>。其研究表明,复杂性水平不断增加的任务在流体智力因素上的载荷也不断增加,他还提

出可根据因素载荷的变化来评估研究者是否采取了复杂性不同的任务<sup>[6]</sup>。与此同时,根据 Arend 等人的观点,Jansen 曾提出可根据因素分析未旋转时提取出来的第一个因素的载荷来判断任务的复杂性,复杂任务比简单任务在这个因素上的载荷要高<sup>[7]</sup>,即任务复杂性和因素载荷是正相关。Arend, Colom 和 Botella 等人使用传递性推理任务进行的研究表明<sup>[7]</sup>,任务难度(指不能正确解决问题的人数百分比)和任务在未旋转时第一个因素上的载荷有显著相关,因为通常任务复杂性越高其难度也越大,因此结果间接验证 Jansen 的观点。然而,正如他们自己所言,此观点仅在传递性推理任务上得以验证尚具有一定的限制性,还有待在更多任务领域进行广泛证实。这构成了本研究的第一个研究目的,即使用平衡秤任务进一步检验 Jansen 等人的观点。

根据通过率、反应时或者因素载荷大小判断任务的复杂性,都要在任务测验完成后依据所获得的数据进行,可以视为对任务复杂性的事后分析;除此之外,还可以进行事前分析,即在任务施测之前就根据理论思路和逻辑分析等大致确定任务复杂性。因为事前分析在提高任务设计的合理性,减少“事后诸葛亮”式的麻烦等方面都有其重要意义,因而成为研究者非常关注的问题,如辛自强的“关系一表征复杂性模型”就提供了事前分析任务复杂性的思路和方

\* 基金项目:国家自然科学基金资助项目(30500162)。

通讯作者:辛自强,北京师范大学发展心理所副教授,博士·E-mail: xinzhiqiang@sohu.com

法<sup>[8,9]</sup>。具体来讲,该模型认为事物的本质在于关系,且应从等级和水平两方面界定任务的关系复杂性。研究者可以对设计好的或选用的实验任务进行这种复杂性的分析,这种分析不需要依赖于被试的实际表现,而完全是研究者的一种逻辑的分析,旨在确定被试完成任务通常必须进行的最低层次的理解或者确定被试必须理解的任务中所包含的关系。首先,所谓关系的“等级”是强调理解低一级关系的输出是理解高一级关系的输入,高一级关系的理解必须以低一级关系的理解为前提,二者有迭代和嵌套关系。而等级复杂性反映的就是这种关系的迭代和嵌套程度。最简单的任务只有一个要素,这时无所谓关系,称为零级关系。如果有两个或更多要素同时有关联,就构成一级关系;然后一级关系又与同样级别的关系或要素发生关系,就形成了二级关系,以此类推可以有三级、四级关系,等等。在水平方向上,每级关系可能要同时理解的其所包含的下级(即低于该级的)关系(或要素)的数量就是该级关系的水平复杂性。问题的解决可能涉及多个等级关系,整个任务的水平复杂性由水平复杂性最大的那个等级的水平复杂性来表示。除了任务的等级和水平复杂性会影响任务复杂性以外,任务的知识背景和经验要求也对任务复杂性有很大影响,甚至是决定性的影响。以上便是关系—复杂性模型的基本内涵。目前,该理论模型已得到初步验证,如辛自强以小学高年级学生为被试,以长方形面积问题方面的三种测验收集的数据对该模型的内部和外部效度进行了检验,结果表明它能够有效解释问题难度,并区分优、中、差三类学生所能达到的表征复杂性<sup>[8]</sup>。此外,该模型也能较好地解释问题解决中图式与策略的关系<sup>[10]</sup>。然而,该理论仍需要更多的实证研究拓展并检验模型的合理性与有效性。本研究的第二个研究目的即是考察该理论在预测不同平衡秤任务复杂性时的有效性。

平衡秤任务即在平衡秤两侧将不同重量的物体放在距支点不同距离的柱子上,要求儿童判断平衡秤的哪一端会下沉。作为经典的皮亚杰任务,研究者基本是围绕诸如平衡秤任务上儿童的心理发展机制、发展年龄、发展的个体差异等发展问题开展的研究。如,Siegler 提出了平衡秤任务的决策树模型来刻画不同年龄儿童使用的规则复杂性特点<sup>[11]</sup>,Richardson, Baughman 等人提出了平衡秤任务的计算模型来解释个体间和个体内的发展多样性问题<sup>[12]</sup>,Shultz 和 Schmidt 提出层叠—关联模型模拟平衡秤任

务上人类的认知发展过程和现象<sup>[13]</sup>,Van der Maas 和 Jensen 则使用高级统计技术探讨了儿童在平衡秤任务上的规则使用发展特点及规则使用的认知过程<sup>[14,15]</sup>,其他研究者如 Dawson 和 Zimmerman 提出了联结主义模型<sup>[16]</sup>,van Rijn, van Someren 和 van der Maas 模拟了平衡秤任务上不同阶段间的发展过渡机制<sup>[17]</sup>。然而,发展研究中,研究任务的选择是非常重要的环节,皮亚杰正是选择了合适的任务证实了自己的理论假设,但也正是这些任务使他的理论备受批评。因此,有必要对平衡秤任务本身进行细致的分析,这将有助于回答有关平衡秤任务的个体心理发展问题。

综上,本研究总的研究目的有二:(1)考察任务在未旋转时第一个因子上的载荷能否作为评价平衡秤任务复杂性的一个指标;(2)考察关系—表征复杂性模型对平衡秤任务复杂性分析的有效性。这两个问题的探讨是以任务难度和反应时作为任务复杂性的效标。通过对这两个问题的探讨,可从事前和事后分析两个角度更全面的了解任务复杂性的分析方法,为以后任务的选择和心理能力的评估奠定基础,具有重要的方法学意义。

## 2 研究方法

### 2.1 被试

从北京某小学四至六年级选取 264 名儿童作为被试。其中四年级三个班 118 人,男生 66 名,女生 52 名,平均年龄 10.4 岁。五年级三个班 90 人,男生 47 名,女生 43 名,平均年龄 11.5 岁。六年级两个班 56 名,男生 38 名,女生 18 名,平均年龄为 12.3 岁。

### 2.2 研究材料

实验采取自编的计算机程序呈现平衡秤任务,改编自 Siegler<sup>[11]</sup>和 Jansen 和 van der Maas<sup>[14]</sup>。本研究在 Siegler 所设计的六类问题基础上,去掉了平衡秤两侧重量和距离均相等的平衡问题,增加了重量距离问题。每类问题的题目特点和实例见表 1。

平衡秤任务采取图片呈现方式。如图 1 所示,平衡板上面有 8 个小柱子,每个小柱子的重量及相邻小柱子的距离都相等。每根小柱子上都可以套重量相同的小铁圈,但最多只能套 6 个。例如图 1 中的平衡秤,左侧的第二根柱子上套有两个小铁圈,右侧第二根柱子上套有一个小铁圈,这就构成了一个重量问题。平衡秤任务共包括 51 个项目,其中前三类问题均有 6 个项目,冲突重量问题有 9 个,冲突距离和冲突平衡问题均有 12 个。所有题目随机呈现。

项目呈现的同时,给儿童提供 3 个选项,即平衡秤右侧下沉、左侧下沉和处于平衡状态 3 种答案,儿童在认为正确的答案下方用鼠标点击即可。

表 1 六类问题的题目特点和实例

六类问题	题目特点	实例			
		左边(重量 距离)		右边(重量 距离)	
重量问题	天平两侧的重物到支点的距离相等,但两侧的重物重量不等	2	3	5	3
距离问题	天平两侧的重物重量相等,但重物到支点的距离不等	3	2	3	1
重量距离问题	天平一侧的重量较大,同时离支点的距离较远	3	1	4	2
冲突重量问题	天平一侧的重物比较重,而另一侧的重物距支点的距离较远,最终物体较重的一侧向下倾斜	4	2	1	3
冲突距离问题	天平一侧的重物比较重,而另一侧的重物距支点的距离较远,最终距支点较远的一侧向下倾斜	3	2	2	4
冲突平衡问题	天平一侧的重物比较重,而另一侧的重物距支点的距离较远,但两侧的力矩相等,平衡秤保持平衡	2	2	1	4

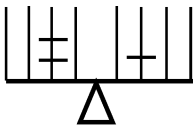


图 1 平衡秤

2.3 实验任务复杂性的事先分析

关系一表征复杂性模型是从等级复杂性、水平复杂性和知识经验三方面来分析平衡秤复杂性的。举例来讲,表 1 所示的冲突距离问题,此类任务的解决首先需要理解重量和平衡秤的状况存在一定关系,即重量大小比较中重量大的一端下沉,同时距离和平衡秤的状况也存在一定关系,即距离大小比较中距离远的一端下沉,这是两个关系,但这两个关系在逻辑上无先后之分,因此并列形成了一级关系;在此基础上,还需要理解两个维度是如何协调共同影响平衡秤的情况,这形成了二级关系。具体讲,左侧的重量大但右侧的距离远,这时可比较两侧重量差(左侧重量-右侧重量)和两侧距离差(左侧距离-右侧距离)的绝对值,因为距离差的绝对值大,因此根据距离判断平衡秤的情况。或者也可比较两侧的力矩差,即比较两侧重量和距离乘积的大小,除此之外还可以采取加法原则,即比较两侧的距离和重量之和的大小。不管采取何种方法,都需要儿童对重量和距离两个维度进行权衡,这是任务对儿童问题解决的基本要求,因此这类任务的最高等级为二级关系。关于水平复杂性,一级关系中需要分别比较两侧的重量大小和两侧的距离大小,其中不管是重量大小的比较还是距离大小的比较,需要同时理解加工的都为两个元素,即左右侧的重量或距离,因此该等级的水平复杂性为 2。对于二级关系,以使用第一种方法为例,儿童需进行两个操作步骤,首先计算重量差和距离差,然后比较差的大小,其中第一步

的重量差或距离差的计算需同时理解加工的元素均为三个,即左侧重量(距离)、右侧重量(距离)和重量差(差距离),因而水平复杂性为 3,而第二步的大小比较需同时理解加工的元素为两个,即距离差和重量差,水平复杂性为 2。整个任务的水平复杂性以最大的水平复杂性来表示,因此该任务的水平复杂性为 3。

对于重量问题,儿童只需理解如前所述的一级关系,即重量和距离都影响平衡秤的平衡状况,不用对重量和距离两个维度进行权衡。因为距离相等,儿童只需根据重量大小就可判断出平衡秤向哪侧下沉。因此这类任务的等级复杂性为一级关系,水平复杂性为 2。

对于距离问题,它与重量问题类似只需理解一级关系,不用对重量和距离两个维度进行权衡。因为重量相等,儿童只需根据距离大小就可判断出平衡秤向哪侧下沉。因此其等级复杂性亦为一级关系,水平复杂性亦为 2。然而,这两类任务的复杂性是存在很大差异的,这就涉及知识经验对任务复杂性的影响。很多研究表明儿童首先能够解决只需根据重量判断平衡秤状况的任务,然后才是那些只需根据距离判断平衡秤状况的任务<sup>[11, 14, 18]</sup>,之所以有这样的结果,正如 Siegler 将重量称为优势维度(dominant dimension)<sup>[19]</sup>,是因为直观的知识经验中重量似乎与平衡秤的平衡状况有更大的关系。考虑到这一点,本研究也将知识经验主要是重量优势效应作为评价平衡秤任务复杂性的一个重要指标。具体讲,如果平衡秤任务只需考虑重量维度,即比较两侧重量大小就能作出正确回答,那么任务要求的知识经验记为 0;相反,如果只考虑重量维度不能作出正确回答则记为 1。举例来讲,前述提到的冲突距

离问题,如果只根据重量维度判断平衡秤的平衡状况,答案为左侧下沉,但正确答案应为右侧下沉,即只根据重量维度判断会得到错误答案,因此在任务要求的知识经验方面记为 1。关系一表征复杂性模型对六类问题复杂性水平的预测详见表 2。

表 2 关系一表征复杂性模型对每类问题复杂性水平的预测

问题类型	基于关系一表征复杂性模型的任务分析		
	等级复杂性	水平复杂性	知识经验
重量问题	1	2	0
距离问题	1	2	1
重量距离问题	1	2	0
冲突重量问题	2	3	0
冲突距离问题	2	3	1
冲突平衡问题	2	3	1

2.4 研究程序

使用学生的计算机课程上机时间在计算机网络教室进行团体施测。由心理学专业的研究生担任主试,实验前首先对主试进行培训,以统一指导语和一些意外情况的处理。具体施测时,主试首先向儿童介绍平衡秤的构成情况,以便儿童熟悉任务情境。之后儿童进行三个项目的练习,以便熟悉任务的作答过程。最后,儿童进入正式的测验阶段,计算机会自动记录呈现项目的编号、项目呈现至儿童作出选择的时间、儿童所选择的答案以及答案的正确与否。正确记为 1 分,错误记为 0 分。

3 结果

3.1 六类问题的难度

以问题类型为自变量,任务难度为因变量进行方差分析。其中,任务难度值为被试在该类任务上(每类任务均有至少 6 个项目)没有通过的项目数占总项目数的百分比,图 2 中每类任务的难度为所有被试平均后的任务难度。结果表明六类问题的难度存在显著差异,  $F_{(5, 45)} = 383.13, p < 0.001$ , 如图 2 所示。进一步对六类问题的难度差异进行事后检验,结果表明:除了重量问题和距离重量问题的难度不存在显著差异外,其余问题之间差异均非常显著,  $p < 0.01$ 。具体来讲,六类问题的难度序列从低到高依次是:重量问题和重量距离问题、冲突重量问题、距离问题、冲突距离问题、冲突平衡问题。任何问题的难度都显著高于序列中其前面的问题,同时显著低于序列中其后面的问题。任务难度作为评价任务复杂性的重要指标,以上结果说明重量问题和重量距离问题、冲突重量问题、距离问题、冲突距离问题、冲突平衡问题的复杂性水平不断增加。

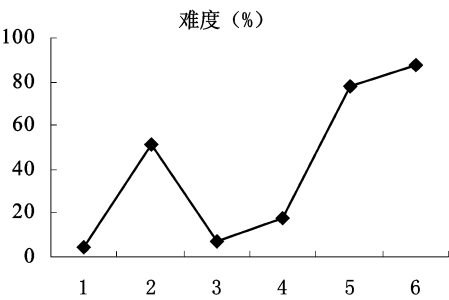


图 2 六类问题的难度

1=重量问题 2=距离问题 3=重量距离问题  
4=冲突重量问题 5=冲突距离问题 6=冲突平衡问题

关于儿童对每类问题的反应时,我们分析了得分较高的被试和得分较低的被试其反应时是否存在差异。以儿童是否能够正确解决每类问题中 50% 以上的项目为标准,将其分成高分组和低分组。因冲突重量问题共有 9 个项目,以儿童是否能够正确回答 5 个以上的项目为标准进行了划分。结果表明:(1)重量问题上,高分组和低分组的反应时没有显著差异;(2)重量距离问题和冲突重量问题上,高分组的反应时显著少于低分组,  $F$  值分别为 4.36、53.53,  $p < 0.05$ ; (3)距离问题、冲突距离和冲突平衡问题上,高分组的反应时均显著多于低分组,  $F$  值分别为 26.54、7.19、12.1,  $p < 0.01$ 。可以看出,所有任务上反应时与任务难度的关系并不是很一致。因本研究中六类任务的正确率有很大差异,最低才 10% 左右,这时反应时未必是正确回答时的反应时,而完全做出正确回答的被试又很少,这时选取这些被试分析其反应时意义也不大,因此以下分析主要以任务难度作为任务复杂性的效标。

3.2 因素载荷与任务复杂性

基于上述的分析以及任务难度是人们普遍接受的用来评价任务复杂性的指标,对任务在未旋转时第一个因素上的载荷和任务复杂性关系的探讨,主要是通过考察因素载荷和任务难度的关系而实现的。对儿童在平衡秤任务 51 个项目上的得分进行主成分因素分析,并考察每个项目在因素分析未旋转时抽取出来的第一个因子上的载荷与项目难度(没有正确解决该项目的人数百分比)之间的关系。结果表明,两者之间不存在线性关系,但是可分成两种情况来讨论。

根据因素分析未旋转时抽取出来的第一个因子上的载荷,可以将所有项目分成两类:首先,重量问题、重量距离问题以及冲突重量问题的载荷均为负值,在 -0.23 至 -0.72 之间,共 21 个项目,可归为

一类;其余三类问题(距离问题、冲突距离问题和冲突平衡问题)的载荷均为正值,在 0.31 至 0.75 之间,共 30 个项目,可归为一类。依据以上的分类,分别考察每类问题中项目的载荷与其难度之间的关系。结果表明,第一类问题的 21 个项目在未旋转时抽取出来的第一个因子上的载荷与其难度之间无显著相关,  $r = -0.35, p > 0.05$ ,这说明第一类问题中的项目的因素载荷与其难度之间不存在必然的关系;而第二类问题的 30 个项目在因素分析未旋转时抽取出来的第一个因子上的载荷与其难度之间的相关极其显著,  $r = -0.81, p < 0.001$ ,进一步回归分析表明,项目的因素载荷对其难度变异的解释率为 65.1%,这说明第二类问题中的项目的因素载荷越高,难度则越低。

3.3 关系一表征复杂性模型对平衡秤任务复杂性的预测

因本研究平衡秤任务的特殊性,这些任务的等级复杂性和水平复杂性恰好完全相关,加上等级复杂性是任务的关系复杂性分性的基础,因此进行回归分析时,仅选择了等级复杂性和知识经验作为自变量,同时选取任务难度为因变量,采用分层逐步回归分析考察了基于关系一表征复杂性模型对平衡秤任务复杂性的分析是否能够较好地预测任务的难度。因为该理论强调的核心是等级复杂性与水平复杂性,因此进行分层回归分析时,等级复杂性作为第一层自变量,而知识经验则作为第二层自变量。分析结果表明,关系一表征复杂性模型对任务难度的解释率高至 95.0%,其中等级复杂性的解释率为 35.8%,标准化回归系数为 0.29,知识经验的解释率为 59.2%,标准化回归系数为 0.83。此结果说明关系一表征复杂性模型对平衡秤任务复杂性的分析能够较好地预测平衡秤任务的难度,换句话说,能较好地预测任务的复杂性。

表 3 基于关系一表征模型的任务分析  
对任务难度的回归分析

变量	B	$\beta$	$t$	$R^2$
等级复杂性	21.17	0.29	8.62***	0.36
知识经验	57.73	0.83	24.20***	0.59

注:\*\*\*  $p < 0.001$ .

4 讨论

4.1 因素载荷与任务复杂性

本研究对任务在未旋转时第一个因素上的载荷和任务复杂性关系的探讨,主要是通过考察因素载

荷和任务难度的关系而实现的。之所以选择任务难度而不是反应时作为任务复杂性的效标,主要原因在于当被试作出的是错误回答时,这时的反应时长短并不能反映任务客观的复杂程度,这种情况尤其发生在难度较大的任务上。具体到本研究,我们发现难度较大的距离问题、冲突距离和冲突平衡问题上高分组的反应时均显著多于低分组。之所以产生这样的结果,是因为这三类题目的正确解决都需要儿童能够同时考虑重量和距离维度,但很多被试采取简单的策略来完成,如只根据重量维度迅速完成题目或者因为任务太难完全靠猜测完成,这就导致被试的反应时很短但正确率却很低。除非任务难度较低,正如本研究中的重量问题,儿童在这类任务上的正确率非常高,这时的反应时就能够反映任务客观的复杂性。也就是说,反应时是否能够反映任务客观的复杂程度与任务难度有很大关系,因此我们选择了任务难度作为任务复杂性的效标。

以任务难度为效标,对任务在未旋转时第一个因素上的载荷和任务复杂性的关系进行探讨。结果表明,重量问题、重量距离问题以及冲突重量问题中项目的因素载荷与其难度之间没有显著相关,而距离问题、冲突距离问题和冲突平衡问题中项目的因素载荷与其难度之间存在显著的负相关,这些结果与 Jansen, Arend 等人认为的任务复杂性和第一个因素上的载荷存在正相关的观点是不一致的。不过, Stankov 的一项研究结果也曾显示因子载荷和任务难度未必是正相关<sup>[6]</sup>。他以大学生为研究对象,考察了不同 Swaps 任务和不同 Triplets 任务的复杂性,并采用结构方程模型探讨了这两类任务和瑞文推理等其他智力测验之间的关系。其中,复杂性依次增加的四个 Swaps 任务在其公共因子上的载荷分别为 0.46、0.49、0.68、0.62,而复杂性依次增加的四个 Triplets 任务在其公共因子上的载荷分别为 0.48、0.43、0.56、0.37。可以看出,这两种任务中复杂性较大的任务未必因子载荷就高。此现象该如何解释呢?

仔细分析可以发现,原因可能在于本研究所选取的任务的难度与以往研究有很大不同。比如, Arend, Colom 和 Botella 是以参加专业培训课程的成人作为研究对象对传递性推理任务的复杂性进行了考察,这些任务的难度在 0.06~0.43 之间,平均难度为 0.24,而本研究探讨的是 10~12 岁儿童在平衡任务上的表现,其中第一类问题的任务难度在 0.04~0.27 之间,平均难度为 0.11,第二类问题的难度均

在 0.46~0.92 之间,平均难度为 0.76。可以看出,与 Arend 等人的研究相比,本研究中的第一类问题对于儿童来讲有些过于简单,第二类问题则有些过于困难。而儿童在那些相似的均过于简单或者均过于困难的任务上的认知过程复杂性可能并不是通常情况下那样,完成较复杂任务总比完成较简单任务的认知过程要复杂。

具体来讲,本研究第一类问题中的重量问题、重量距离问题以及冲突重量问题,都可以靠直观的生活经验,即依据重量决定平衡秤的平衡状况这一生活经验迅速作出正确回答,因此儿童在这三类任务上的认知过程复杂性基本相同,这从儿童在三类问题上的正确率均在 80% 以上以及儿童对这三类问题的反应时不存在显著差异也可推断出。而第二类问题中的距离问题、冲突距离问题和冲突平衡问题尤其是后两类冲突问题,如果依靠直观的知识经验即重量决定平衡秤的平衡状况将作出错误回答,加上 10—12 岁的儿童都没有学过力矩知识、任务采用电脑呈现方式<sup>[20]</sup>以及要求儿童尽快作出回答,这些都导致任务的难度增加,使得儿童很可能使用一些不正确的简单方法来解决这些问题尤其是难度较大的冲突平衡问题,比如仅根据重量判断平衡秤的状况。本研究对儿童的回答模式进行具体分析就发现,有 49% 的儿童在解决所有问题时仅根据重量判断平衡秤的状况,表现在重量问题、重量距离问题以及冲突重量问题上正确率非常高,但距离问题、冲突距离问题和冲突平衡上回答正确率非常低。这意味着,第二类问题中较难的冲突问题可能和较简单的距离问题的认知过程复杂性基本相同,甚至冲突问题的认知过程更为简单,比如完全靠猜测。因此第二类问题中那些复杂的冲突问题反而可能比简单的距离问题的认知过程复杂性要低。

而因素载荷的高低正是任务需要的认知过程的复杂性或能力的高低的反映<sup>[6, 7, 21]</sup>,即因素载荷高的任务比载荷低的任务的认知过程复杂性要高。比如,Stankov 认为使得因素载荷增加的实验任务之所以更为复杂,是因为它们考察了广泛的基本认知过程<sup>[6]</sup>。Snow, Kyllonen 和 Marshalek 认为复杂性意味着特定任务需要的认知成分的数量在增加<sup>[21]</sup>。Arend, Colom 和 Botella 则提出未旋转时抽取出来的第一个因子上的载荷代表任务需要的资源数量(过程和/或者能力)<sup>[7]</sup>。结合上述对本研究中第一类问题的认知过程复杂性和第二类问题的认知过程复杂性的分析,第一类问题中项目的因素载荷与其难度

之间没有显著相关以及第二类问题中项目的因素载荷与其难度之间存在显著的负相关就不难理解。

此外,任务过于简单或者过于困难,从统计学角度来讲,儿童在任务上的成绩变异或标准差会较小,也会造成因子载荷比较低。以本研究中第二类问题为例,距离问题、冲突距离问题和冲突平衡问题的任务难度标准差分别为 0.43、0.30、0.19,这与三类问题的任务难度恰好是负相关。关于这一点,Raykov 曾进行过专门探讨<sup>[3]</sup>,他从统计学的角度分析了任务复杂性和相关系数、因子载荷的关系,指出如果题目过于简单或者过于困难从而出现天花板效应和地板效应,这时就不会有高的因子载荷,而且过于复杂、过于简单和中等难度的题目测量的不是潜在的一种特质。这提示我们,Jansen 等人的观点可能适用于中等难度的任务,即因素载荷和任务复杂性的正相关关系可能是有限的。

#### 4.2 基于关系一表征复杂性模型的任务分析

研究结果表明,基于关系一表征复杂性模型的任务分析能够较好地预测平衡秤任务的难度,其解释率达到 95.0%。可见,就平衡秤任务而言,关系一表征复杂性模型提供的分析任务复杂性的思路和方法是比较合理的。

然而,基于关系一表征复杂性模型对任务复杂性的分析,有两点需要强调:首先,本研究中的平衡秤任务其等级复杂性和水平复杂性呈现完全相关,是任务特殊性使然,并不是普遍现象。该模型认为,任务复杂性的分析首先是分析等级复杂性,在此基础上再分析水平复杂性。其中等级复杂性对任务难度起决定作用,任务需要理解的关系等级可以是一级关系、二级关系、三级关系、四级关系甚至更高级关系,等级越高其难度也越大。基于此,本研究在进行回归分析时将等级复杂性纳入了回归方程。结果也如所料,等级复杂性对平衡秤任务难度的解释率达到 35.8%,能够显著地预测平衡秤任务的复杂性。而水平复杂性对任务难度起限制性作用,因为个体能够同加加工处理的元素是有限的,如 Pascual-Leone<sup>[22]</sup>和 Halford<sup>[23]</sup>都认为,我们同时只能加工四元关系,即相互关联的四个变量或元素,这是人类认知加工能力的限制。这意味着当等级复杂性不断增加时,水平复杂性可能一直为两个元素或三个元素,因此等级复杂性和水平复杂性未必就是线性关系。

其次,知识经验对任务难度有很大影响。本研究中直观的知识经验即重量优势效应对平衡秤任务的难度解释率接近 60%,可以说解释率是比较高

的。不同领域的任务,知识背景和知识要求都是不同的,即便在同一领域每个任务要求的知识点也可能不同。这种知识要求的不同,必然会影响问题的难度。例如,辛自强曾使用过的长方形面积任务,都涉及到了面积公式,但是其中部分问题还要求使用周长公式,从而增加了任务难度<sup>[8]</sup>。一般来讲,在任务要求知识背景较多的领域,知识经验对任务难度的影响相对较大,如学科问题中。而在任务要求知识背景较少的领域,等级复杂性和水平复杂性对任务难度的影响则相对较大,如逻辑推理问题。本研究中的平衡秤任务为物理问题,因此知识经验对其难度有较大影响。进一步的研究需要将关系一表征复杂性模型应用在更多的任务领域尤其是逻辑推理问题上,考察基于该模型的任务分析主要是等级复杂性和水平复杂性对任务难度的影响。

综上,本研究通过对因素载荷提供的事后分析思路和基于关系一表征复杂性模型进行的事前任务分析思路进行初步探讨,获得了一些很有价值的结果,比如任务在未旋转时第一个因素上的载荷不能反映平衡秤任务的复杂性,基于关系一表征复杂性模型的任务分析具有一定的有效性,这些都有助于今后对任务的选择和分析,有助于更好地对个体心理能力进行评估和测量。然而,进一步的研究需要扩展到更多的任务领域拓展和验证上述的结果,比如选取难度适中的任务考察任务在未旋转时第一个因素上的载荷是否和任务难度有正相关,还可以对 Stankov 的观点进行考察和验证即任务在流体智力上的因子载荷是否能够反映任务的复杂性,而关系一表征复杂性模型也有待考察其中等级复杂性、水平复杂性和知识经验对任务难度的相对影响。

## 5 结论

关于平衡秤任务复杂性的事前和事后分析,我们可以得到以下结论:

(1)事后确定的任务在未旋转时第一个因素上的载荷未能反映平衡秤任务的复杂性;

(2)关系一表征复杂性模型提供的事前分析平衡秤任务复杂性的思路和方法是较为有效的,任务等级复杂性和知识经验两者对任务难度的解释率为 95.0%。

## 参考文献:

[1] Crawford J D. Intelligence, task complexity, and the distinction between automatic and effortful mental processing. In H A Rowe (Ed.),

Intelligence reconception and measure. Hillsdale, New Jersey: Erlbaum, 1991, 120—144.

[2] Stankov L, Crawford J D. Ingredients of complexity in fluid intelligence. Learning and Individual Differences, 1993, 5(2): 73—111.

[3] Raykov T, Stankov L. On task complexity and “simplex” correlation matrices. Australian Journal of Psychology, 1993, 45: 81—87.

[4] Stankov L. The complexity effect phenomenon is an epiphenomenon of age-related fluid intelligence decline. Personality and Individual Differences, 1994, 16(2): 265—288.

[5] Stankov L, Raykov T. Modeling complexity and difficulty in measures of fluid intelligence. Structural Equation Modeling, 1995, 2(4): 335—366.

[6] Stankov L. Complexity, metacognition, and fluid intelligence. Intelligence, 1999, 28(2): 121—143.

[7] Arend I, Colom R, Botella J, et al. Quantifying cognitive complexity: evidence from a reasoning task. Personality and Individual Differences, 2003, 35: 659—669.

[8] 辛自强. 关系一表征复杂性模型的检验. 心理学报, 2003, 4: 504—513.

[9] 辛自强. 问题解决和知识建构. 北京: 教育科学出版社, 2002.

[10] 辛自强. 问题解决中图式与策略的关系: 来自表征复杂性模型的说明. 心理科学, 2004, 27(6): 1344—1348.

[11] Siegler R S. Three aspects of cognitive development. Cognitive Psychology, 1976, 8: 481—520.

[12] Richardson F M, Baughman F D, Thomas M S C. Computational modeling of variability in the balance scale task. 7th. International Conference on Cognitive Modeling. Trieste, 2006, April.

[13] Shultz T R, Schmidt W C. A cascade-correlation model of balance scale phenomena. Paper presented at the Thirteenth Annual Conference of the Cognitive Science Society. 1991, 635—640.

[14] Jansen B R J, van der Maas H L J. The development of children's rule use on the balance scale task. Journal of Experimental Child Psychology, 2002, 81: 383—416.

[15] van der Maas H L J, Jansen B R J. What response times tell of children's behavior on the balance scale task. Journal of Experimental Child Psychology, 2003, 85: 141—177.

[16] Dawson M R W, Zimmernan C. Interpreting the internal structure of a connectionist model of the balance scale task. Brain and Mind, 2003, 4: 129—149.

[17] van Rijn H, van Someren M, van der Maas H. Modeling developmental transitions on the balance scale task. Cognitive Science, 2003, 27: 227—257.

[18] Halford G S, Wilson W H, Phillips S. Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. Behavioral and Brain Sciences, 1998, 21: 803—831.

[19] Siegler R S. Developmental sequences within and between concepts. Monographs of the Society for Research in Child Development, 1981, 46: 1—84.

[20] Birney D P, Halford G S. Cognitive complexity of suppositional reasoning: An application of the relational complexity metric to the knight —

knave task. *Thinking and Reasoning*, 2002, 8(2): 109—134.

[21] Snow R E, Kyllonen P C, Marshalek B. The topography of ability and learning correlation. In R J Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol 2). Hillsdale, NJ: Erlbaum, 1984.

[22] Pascual-Leone J A. A mathematical model for the transition rule in

Piaget's developmental stages. *Acta Psychologica*, 1970, 32: 301—345.

[23] Halford G S, Wilson W H, Phillips S. Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, 1998, 21: 803—831.

## Prior Analysis and Post Hoc Analysis on the Complexity of the Balance Scale Task

ZHANG Li XIN Zi-qiang

(*Institute of Developmental Psychology, Beijing Normal University 100875*)

**Abstract:** The aims of this study were to explore if the loadings on the first un-rotated factor can be taken as a way to evaluate the task complexity of the balance scale task and to explore the validity of the prior analysis on the balance scale task based on the Relational-Representational Complexity Model. 264 children from grade 4, 5 and 6 were presented with a computerized balance scale task. There were 51 items and six item types, including weight problems, distance problems, weight distance problems, conflict weight problems, conflict distance problems, and conflict balance problems. Results showed that: (1) the loadings of all balance scale items on the first un-rotated factor and the difficulty of those items had no significant positive relation. It might be that task complexity could be quantified by the loadings on the first un-rotated factor only when the difficulty of task was moderate. (2) The analysis of task complexity based on the Relational-Representational Complexity Model could predict task difficulty ( $R^2=0.95$ ), which indicated that the model was valid in analyzing the complexity of the balance scale task.

**Key words:** Relational-Representational Complexity Model; task complexity; task difficulty; factor analysis