

基于潜在类别分析的6~15岁儿童 平衡秤任务认知规则探索*

赵玉¹ 戴海琦^{**2} 刘铁川²

(¹赣南医学院心理学系,赣州,341000)(²江西师范大学心理学院,南昌,330022)

摘要 本研究使用自编平衡秤测验测试468名6~15岁儿童,通过潜在类别分析对认知规则进行分类。结果发现,被试使用了重量规则、重量距离规则、补偿规则、力矩规则、距离优势规则等六种规则;6~9岁儿童主要使用重量规则;10~13岁儿童主要使用补偿规则;14岁以上儿童主要使用力矩规则;13岁到14岁之间是掌握力矩规则的关键时期。与规则评估技术相比,潜在类别分析用于认知规则研究具有明显优势,最后对运用此方法的前提假设与局限进行了讨论。

关键词 平衡秤任务 潜在类别分析 认知规则 解题策略

1 引言

不同被试可能采用不同的策略解决面临的问题,完成认知任务的灵活性是智慧行为的重要表现。研究表明,儿童在代数运算、因果推理、决策、阅读理解、序列回忆和空间推理等多个领域都存在使用多策略进行问题解决的现象(Siegler, 1991)。在发展心理学研究中,平衡秤任务(图1)被广泛用于研究儿童比例推理能力发展的阶段性特征与差异(Jansen & van der Maas, 1997, 2002)。Siegler(1981)认为不同发展阶段儿童问题解决过程的根本差异在于所使用规则(Rule)的差异。规则为被试解决问题时采用的一系列心理操作,即策略,文中二者同义。

Siegler(1981)认为儿童解决平衡秤任务存在四种基本规则。为读者理解和行文方便,重新命名如下。a. 使用规则I的儿童只考虑重物,认为平衡秤重量更大的一端会下降,若重量相等则保持平衡,此规则可称为重量规则。b. 使用规则II的儿童在两端重物重量不等时只根据重量判断,若重量相等,则倾向于根据距离判断,可见儿童根据任务特征不同倾向于根据不同维度进行判断,此规则可称为重量距

离规则。c. 使用规则III的儿童同时考虑重量和距离,当重量或距离至少有一个维度相等时,他们能正确判断;但是当重量与距离维度的信息存在冲突时,如某一端重物的重量更大,而另一端重物与支点距离更大时,则凭猜测解题,此规则可称为无冲突规则。d. 使用规则IV的儿童掌握了重量与距离相乘的力矩运算法则,通过力矩大小来判断平衡秤的下降方向,此规则可称为力矩规则。研究显示,儿童也可能使用质性比例规则(Normandeau, Larivée, Roulin, & Longeot, 1989)、补偿规则(Boom, Hoijsink, & Kunnen, 2001)等其它规则完成平衡秤任务。

1.1 Siegler的规则评估技术

为确定被试解决平衡秤任务时具体采用了哪种规则,Siegler(1981)提出了规则评估技术(Rule Assessment Methodology, RAM)。使用RAM的第一步是对四种基本规则解决各类平衡秤任务的正确概率进行理论分析。平衡秤任务一共有平衡(Balance)、距离(Distance)、重量(Weight)、冲突重量(Conflict - Weight)、冲突距离(Conflict - Distance)、冲突平衡(Conflict - Balance)6种类型。表1呈现了使用四种基本规则解决6类平衡秤任务的预期正确率。

表1 四种基本规则解决各类平衡秤任务的预期正确率

使用规则	题目类型					
	平衡	重量	距离	冲突平衡	冲突重量	冲突距离
重量规则	1	1	0 ^a	0 ^b	1	0 ^b
重量距离规则	1	1	1	0 ^b	1	0 ^b
无冲突规则	1	1	1	0.33 ^c	0.33 ^c	0.33 ^c
力矩规则	1	1	1	1	1	1

注:^a 错误地预测平衡秤保持平衡。^b 错误地预测平衡秤向更重的一端下降。^c 通过猜测作答。

* 本研究得到高等学校博士学科点专项科研基金联合资助课题(20103604110002)、国家自然科学基金项目(31100756)和江西省高校人文社会科学青年基金项目(XL1003)资助。

** 通讯作者:戴海琦。E-mail: daihaiqi@eyou.com

将被试的作答正确率与表1比较便可判定被试所使用的规则。由于作答过程中的猜测和失误,观察到的作答概率不会与表1的数据完全匹配,Siegler(1981)认为答对80%或85%左右的题目便符合100%的理论概率。但此标准为主观制定,所以RAM缺乏观察模式与理想模式匹配程度的统计检验信息(Boom & ter Laak, 2007)。另外, RAM需要事前界定全部规则可能出现的反应模式,所以很难应用RAM来发现新规则,从而可能导致对被试规则使用的误判。

1.2 潜在类别分析

近年来一些研究者开始使用潜在类别分析方法(Latent Class Analysis, LCA)研究平衡秤任务中的使用规则(Boom et al., 2001)。对于探索性LCA,通常依据Pearson卡方、似然比卡方、AIC和BIC等统计指标确定潜在类别数量。若增加潜在类别数量,而AIC与BIC减小不明显,则认为已有潜在类别足以解释数据。在可能的作答模式非常多、被试数量相对较少的情况下,卡方统计量与对数似然比皆不服从卡方分布,使用自助法似然比检验(Bootstrapped Likelihood Ratio Test, BLRT)比较k个类别模型与k-1个类别模型的拟合程度差异,可以较准确地确定潜在类别数量(Nylund, Asparouhov & Muthén, 2007)。

LCA具有明确的统计基础,克服了RAM的不足,受到了研究者的广泛欢迎(Boom & ter Laak, 2007)。也有研究者认为LCA需要大量被试,且容易产生被试数量少且难以解释的潜类,很难找出真实潜在类别的数量,所以无法取代RAM(Shultz & Takane, 2007)。van der Maas, Quinlan和Jansen(2007)认为上述缺点正是LCA的优势,正因如此,LCA才可探查理论分析所忽略的解题策略。

辛自强,张丽和林崇德(2008)应用RAM技术对10~12岁儿童的研究发现有51.9%的儿童使用了重量规则,他们认为重量规则仍是10~12岁儿童使用的主要规则。而国外研究中,重量规则主要是4、5岁儿童使用此规则(Siegler, 1981)。国内关于儿童在平衡秤任务上使用规则的研究较少,对上述国内外研究结果的巨大差异尚缺乏相关研究来验证。而且,当前儿童规则使用研究中,现代统计方法已逐渐取代了传统的RAM技术。因此,本研究将通过目前广泛应用的LCA方法对中国儿童在平衡秤任务上的策略使用情况进行分析,以确认上述差异的可能性;平衡秤任务被认为是研究儿童比例能力

的最佳任务(Jansen & van der Maas, 2002),在平衡秤任务上的使用不同的规则体现了儿童认知发展水平上的差异,而以往研究中被试年龄往往相对集中,本研究选取了年龄跨度较大的6~15岁儿童完成平衡秤任务,以较全面地探索儿童的认知发展特征。

2 研究方法

2.1 被试

由赣州大公路小学一至五年级、以及赣四中初一至初三年级,在各年级随机抽取一个班的学生,共468人。本研究中作答完整的有457人,其中男性253人,女性204人,被试年龄范围为6岁到15岁,各年龄分别有37(.57)、60(.52)、47(.60)、58(.53)、65(.70)、41(.44)、46(.50)、44(.61)、52(.46)、7(.71)人,括号内为男性所占比例。

2.2 测量工具

自制实物木制平衡秤用于演示,以有助于儿童理解测验中题目的含义。另自编23题平衡秤任务测验,包括3个练习题目和20个正式题目。正式题目包括重量、距离、冲突平衡、冲突重量等四种类型,每类5题,在测验中出现次序随机,均以图形方式呈现,每题有三个选项代表平衡秤的平衡、左端下降、右端下降三种状态。图1为第1个重量题目。

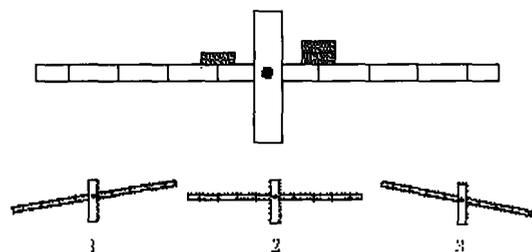


图1 平衡秤任务测验示例

2.3 施测过程

在班级内进行团体测试,每个儿童发放一份自编测验,同时向全体儿童展示实物平衡秤。通过演示、解释和练习,大部分儿童都能够理解测验试题的含义,之后开始正式的图形平衡秤任务测验,要求儿童尽量不要遗漏。整个测试耗时15分钟左右,回收作答完整、有效的测验457份。评判正误后,Cronbach系数为.819,显示自编测验内部一致性较好。

2.4 统计分析

研究使用Mplus6.0通过LCA方法分析各类题目,得到儿童的使用规则。

3 结果

儿童使用不同的规则,会体现在各类题目的作

答之上。分别处理各类题目无法准确区分各种规则,研究将同时分析各类题目,以有效探查儿童使用何种规则完成平衡秤任务。但未掌握重量规则的儿童无法顺利完成本测验,会造成分类混乱,研究通过对重量题目的分析剔除此类被试作答。

表3 重量题目的潜在类别模型

模型	df	Pearson χ^2	似然比 χ^2	AIC	BIC	BLRT 检验 p 值
W3	210	129.056(1.000)	87.051(1.000)	1140.325	1170.757	.000
W4	199	82.322(1.000)	62.983(1.000)	1138.257	1179.150	.000
W5	188	75.472(1.000)	53.397(1.000)	1150.671	1202.025	.600

用 L、B、R 分别表示平衡秤左端下降、保持平衡及右端下降三种状态,则测验中 5 个重量题目的正确反应模式为 [RRLRR]。通过分析 LCA 得出的期望反应模式与正确模式的差异,可以确定各类儿童是否掌握了重量规则。第 1 类儿童的期望反应模式为 [BRLRR], 作答概率不高,只有第 3 题为 .81,其它题目概率在 .33 ~ .66, 此类儿童不理解题意或未掌握重量规则,所占比例为 4.38%。第 2 类儿童的期望反应模式为 [RRLRR], 作答概率为 .97 ~ .99, 已掌握重量规则,所占比例为 86.43%。第 3 类儿童的期望反应模式 [LLRLL], 其作答概率在 .89 ~ 1.00, 与第 2 类相反,选择的是重物更轻的那一端,可以称之为 SWD 规则 (Smallest Weight Down Rule), 此类儿童未掌握重量规则,所占比例为 1.97%。第 4 类儿童的期望反应模式为 [RRLRR], 与正确模式一致,但作答概率在 .58 ~ .80, 并不高,所占比例为

3.1 重量题目的潜在类别分析

表 3 显示,4 类别模型 (W4) 的 AIC 值最小,其 BIC 比 3 类别模型 (W3) 略大,但 BLRT 检验显著,显示增加一个类别可以显著改善模型拟合程度。因此,最终采用模型 W4。

7.22%。

检查第 4 类儿童在距离题目上的作答,发现此类儿童倾向于选择“平衡”或根据距离维度作答,说明此类儿童可能是开始做题时未掌握做题的技巧,而他们在距离题目上又表现出特定的作答模式,所以保留第 4 类儿童的作答。将第 1 类与第 3 类的被试作答剔除,共删除了 29 人,后续统计分析将使用 428 人的数据。

3.2 距离、冲突平衡、冲突重量题目的潜在类别分析

剔除未掌握重量规则儿童的作答后,再考虑其他三类题目。为同时分析各类题目和避免作答模式稀疏导致的参数估计问题,从各类题目中选择第 4 题和第 5 题代表同一类型的项目。表 4 列出了此 6 题的基本性质及一些其他规则解决这些题目的正确概率。

表4 距离、冲突平衡、冲突重量的第 4、5 题的基本特征

项目	左端		右端		答案	补偿规则	最近距离规则	距离优势规则	质性比例规则
	重量	距离	重量	距离					
D4	4	2	4	1	L	1	0 ^R	1	1
D5	4	3	4	4	R	1	0L	1	1
CB4	2	2	4	1	B	0 ^R	0 ^R	0 ^L	1 ^B
CB5	2	3	3	2	B	1	0 ^L	0 ^L	1 ^B
CW4	1	4	3	2	R	0 ^B	1	0 ^L	0 ^B
CW5	2	2	1	3	L	0 ^B	1	0 ^R	0 ^B

注: ^R 预测平衡秤右端下降, ^L 预测平衡秤左端下降, ^B 预测平衡秤保持平衡。

比较结合不同类型题目的潜在类别模型 (表 5), 其中 6 个类别的模型 C6 的 BIC 最小, 其 AIC 比

模型 C7 略大, 根据模型简化原则与 BLRT 检验结果, 最终选择 6 类别模型。

表5 三类题目的潜在类别模型

模型	df	Pearson χ^2	似然比 χ^2	AIC	BIC	BLRT 检验 p 值
C5	664	567.911(.997)	278.077(1.000)	3336.692	3393.378	.000
C6	651	496.098(1.000)	234.038(1.000)	3318.652	3386.854	.000
C7	638	326.335(1.000)	204.766(1.000)	3315.381	3395.097	.120

第 1 类儿童的期望反应模式为 [LRRBBB], 其作答概率为 .62 ~ .94, 与补偿规则的作答模式一致, 所占比例为 23.36%。第 2 类儿童的期望反应模式为 [LRRRRL], 其作答概率为 .71 ~ .90, 与重量距离

规则的作答模式一致, 所占比例为 12.15%。第 3 类儿童的期望反应模式为 [BBRRRL], 其作答概率为 .92 ~ 1.00, 与重量规则的作答模式一致, 所占比例为 29.91%。第 4 类儿童的期望反应模式为

[BBLLLR],其作答概率为.67~.89,他们对距离题目和冲突重量题目会通过重量进行判断,而对于冲突平衡题目倾向于通过距离进行判断,说明此类儿童未完全意识到距离维度的应用情境,无法解决距离题目,此类儿童处于重量规则和重量距离规则之间阶段,称为重量规则,所占比例为4.44%。第5类儿童的期望反应模式为[LRBBRL],其作答概率为.91~1.00,与力矩规则的作答模式一致,所占比例为17.29%。第6类儿童的期望反应模式为[LR-LLLR],其作答概率为.71~.94,与距离优势规则的作答模式一致,所占比例为12.85%。

3.3 规则使用的年龄特征

现在考察分组后儿童的年龄特征。图2中横坐标表示各年龄阶段,纵坐标表示各规则类别中各年龄阶段儿童的人数频次。

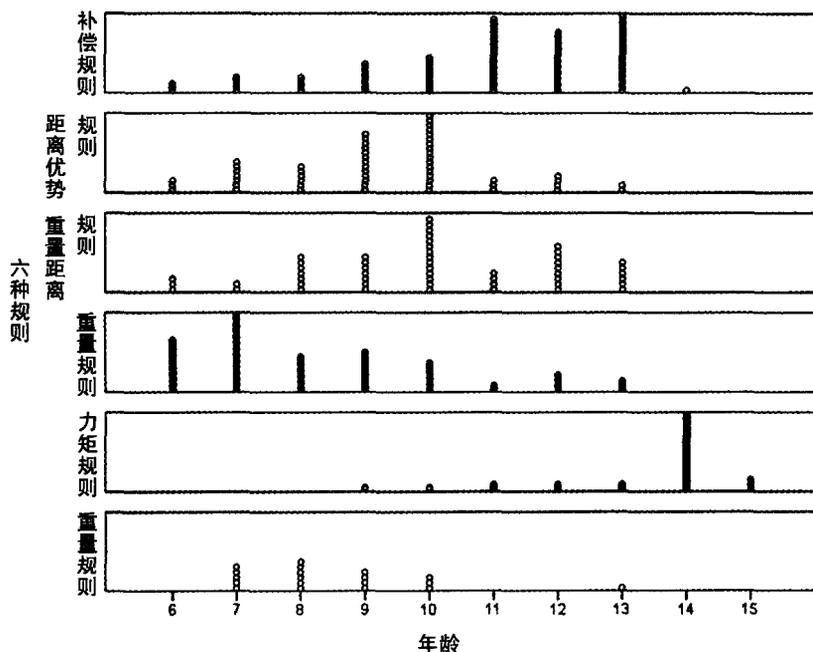


图2 六种规则的年龄分布图

4 讨论

本研究通过潜在类别分析发现被试主要使用六种规则,其中三种不在 Siegler (1981) 提出的四种基本规则之列。

平衡秤任务反映儿童的比例推理能力,使用不同规则的被试比例应具有随年龄变化的趋势。发展性研究中阶梯模型 (Staircase Model) 假定解题规则间的绝对年龄顺序,即某年龄儿童只可能使用一种解题规则。而潜在类别分析方法的理论与叠波模型 (Overlapping Waves Model) 更接近 (Boom & ter Laak, 2007),被试可能使用多种策略,具体使用何

可以很直观的看到,使用重量规则主要集中在6~9岁的儿童,6岁儿童中有73.53%使用此规则,随着年龄增加使用人数逐渐减少,10~13岁的儿童中只有20%左右使用此规则。使用重量规则的主要是7~10岁儿童,但所占比例不高,在10%左右,11岁之后几乎没有儿童再使用此规则。使用重量距离规则的主要是8~13岁的儿童,随着年龄增加,使用人数有减少的趋势。使用距离优势规则的儿童较多的集中在7岁到10岁之间,在11岁之后只有不到10%的儿童使用此规则。使用补偿规则的主要是9岁到13岁的儿童,随年龄增加,使用者也增加,11~13岁的儿童中有半数均使用此规则,但在14岁之后,几乎全部被试(98.08%)都使用了力矩规则。

种策略解题由后验概率确定。本研究虽然无法探索个体的策略发展轨迹,但基于全部样本也可以说明这一点(图2),某年龄儿童可能使用多种策略,但低年龄儿童使用简单规则者更多,而高年龄儿童使用复杂规则者更多,从理论上来看更符合叠波模型的假设。

本研究发现14岁前后掌握力矩规则的儿童数量相差非常大,这与他们测验分数差异相一致,说明力矩规则存在跳跃式的发展。结合以往研究结果 (Boom & ter Laak, 2007; Jansen & van der Maas, 2002),应可判断14岁为掌握力矩规则的关键时间。此外,辛自强等人(2008)研究发现10~12岁儿童主

要使用重量规则完成平衡秤任务(51.9%),而 Jansen 和 van der Maas(2002)发现 10~12 岁儿童中最多 30% 使用重量规则,他们发现这个年龄段中使用加法规则的儿童随年龄增加越来越多,直至 15 岁; Boom 等人(2001)也发现 11 岁之后几乎没有儿童再使用重量规则。本研究结果显示 10~13 岁儿童主要使用补偿规则,而重量规则是 6~9 岁儿童使用的主要解题策略,与国外研究结果比较一致。本研究结果与辛自强等人(2008)存在差异的原因可能来自多个方面,样本中被试年龄范围、规则分类方法以及所使用的平衡秤任务类型都可能影响分类结果。由图 2 可以看出本研究选取年龄跨度较大的被试样本是合理的,如此才能观察到规则使用的完整的年龄变化趋势。

本研究未发现儿童使用最近距离规则和质性比例规则。Jansen 和 van der Maas(2002)观察到了最近距离规则,但并未发现其随年龄的变化趋势,所以这可能并不是儿童的一种稳定的解题策略。无冲突规则是 Siegler(1981)提出的四种基本规则之一,但本研究与其它一些使用 LCA 的研究并未得到明确的可解释为无冲突规则的潜在类别(Boom et al., 2001; Boom & ter Laak, 2007)。事实上,无冲突规则代表了一类处于重量距离规则和力矩规则之间阶段的一类解题策略,有研究者认为这个概念过于宽泛,以至于包罗万象(encompassing)(Boom & ter Laak, 2007)。补偿规则、质性比例规则、最近距离规则等都处于力矩规则前的过渡阶段,所以此阶段的儿童并非单凭猜测解决平衡秤任务,本研究结果说明了儿童很少用猜测,而更多使用补偿规则。所以能否用单一的无冲突规则来指代此过渡阶段存在多种解题策略的儿童,应值得商榷。

使用 LCA 研究平衡秤任务的一步重要工作是对得到潜在类别进行理论上的解释,虽然本研究发现的潜在类别都可以根据其条件概率得到明确的理论解释,但潜在类别分析的结果并非总是如此,统计发现的潜在类别与理论上被试可能使用的规则很可能并不完全对应(Shultz & Takane, 2007)。Boom 和 ter Laak(2007)试图通过将工作记忆与年龄作为协变量纳入对 LCA 来避免这个问题,但只改善了模型拟合程度,而对被试规则分类影响不大。所以,LCA 成功应用关键可能还是在于被试样本的代表性,以及题目区分不同规则的有效性。

5 总结

根据研究结果,得到如下结论:(1)应用 LCA 方法于平衡秤任务认知规则研究,发现 6~15 岁儿童中使用了重量规则、重量规则、重量距离规则、补偿规则、力矩规则、距离优势规则。(2)6~9 岁儿童主要使用重量规则,随年龄增加,使用重量规则、重量规则、重量距离规则的儿童逐渐减少;10~13 岁儿童主要使用补偿规则;14 岁以上儿童主要使用力矩规则。(3)由 13 岁到 14 岁之间,使用力矩规则的儿童数量呈跳跃式增加。

利用 LCA 方法的优势,本研究对中国儿童在平衡秤任务上的规则使用进行了研究,但国内平衡秤任务相关实证研究不多,需要更多研究验证上述结论。另外,使用 LCA 方法分析认知规则有一假设,被试始终使用同一认知规则解决不同类型的平衡秤任务。有研究者试图将作答数据转换为正误后,通过限制 LCA 的正确解答条件概率可以在部分程度上检验这一假设(Jansen & van der Maas, 1997),但数据转换会损失大量策略使用信息,不利于规则分类。对于问题解决中的认知规则转换问题,可能需要使用更复杂的计量模型(Rijkes & Kelderman, 2007)加以研究。

参考文献

- 辛自强,张丽,林崇德.(2008). 10~12 岁儿童在平衡秤任务上的规则使用及其复杂性. *心理科学*, 32(1), 62-69.
- Boom, J., Hoijtink, H., & Kunnen, S. (2001). Rules in the balance: Classes, strategies, or rules for the balance scale task. *Cognitive Development*, 16(2), 717-735.
- Boom, J., & ter Laak, J. (2007). Classes in the balance: Latent class analysis and the balance scale task. *Developmental Review*, 27(1), 127-149.
- Jansen, B. R. J., & van der Maas, H. L. J. (1997). A statistical test of the rule assessment methodology by latent class analysis. *Developmental Review*, 17(3), 321-357.
- Jansen, B. R. J., & van der Maas, H. L. J. (2002). The development of children's rule use on the balance scale task. *Journal of Experimental Child Psychology*, 81(4), 383-416.
- Normandeau, S., Larivee, S., Roulin, J. L., & Longeot, F. (1989). The balance-scale dilemma: Either the subject or the experimenter muddles through. *Journal of Genetic Psychology*, 150(3), 237-250.
- Nylund, K. L., Asparouhov, T., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14(4), 535-569.
- Rijkes, C. P. M., & Kelderman, H. (2007). Latent-response Rasch model for strategy shifts in problem-solving processes. In M. von

- Davies & C. H. Carstensen (Eds.), *Multivariate and mixture distribution Rasch models* (pp. 311–328). New York: Springer.
- Shultz, TR, & Takane, Y. (2007). Rule following and rule use in the balance – scale task. *Cognition*, 103(3), 460–472.
- Siegler, R. S. (1981). Developmental sequences between and within concepts. *Monographs of the Society for Research in Child Development*, 46(2), 1–84.
- Siegler, R. S. (1991). Strategy choice and strategy discovery. *Learning and Instruction*, 1(1), 89–102.
- van der Maas, H. L. J., Quinlan, P. T., & Jansen, B. R. J. (2007). Towards better computational models of the balance scale task: a reply to Shultz and Takane. *Cognition*, 103(3), 473–479.

Latent Class Analysis of 6 to 15-Year-Old Chinese Children's Rule Use in Balance Scale Task

Zhao Yu¹, Dai Haiqi², Liu Tiechuan²

(¹ Department of Psychology, Gannan Medical College, Ganzhou, 341000)

(² School of Psychology, Jiangxi Normal University, Nanchang, 330022)

Abstract The balance scale task has been frequently studied in developmental psychology. In those studies, Siegler's rule assessment methodology (RAM) was often utilized to identify which rule children used to solve each type of balance scale task. However, RAM was frequently criticized for lacking statistical background. It can be solved by latent class analysis (LCA), by which we can statistically test how many rules are needed to fit the data and what these rules are.

In this paper, a test made of 20 items, which belong to four different types of balance scale task, was developed, and the latent structure of Chinese children's performance on the test was analyzed. The 457 Chinese children aged 6 years to 15 years finished the entire test and the data was analyzed through the latent class approach with Mplus 6.0. Before the analysis of the combinations of items, it is essential to eliminate the ones who could not use the weight rule (RULE I). Those children were found by latent class analysis of five weight items. After that, there were 428 people in our dataset for further analysis of the combined items.

The results showed that (1) the LCA had several advantages in classification of the cognitive rules through statistical analysis of combinations of distance, conflict-balance and conflict weight items, two of each type. (2) Model of six classes well fitted the data, in which classes corresponded to weight rule, weight and distance rule, compensation rule, torque rule, distance dominant rule. (3) Chinese children aged from 6 to 9 mainly used weight rule, and with the increase of age, fewer children chose weight rule, weight and distance rule, to solve the test items. Children aged from 10 to 13 mainly used compensation rule. Almost all children older than 14 used torque rule during testing. (4) The increase of the proportion who used torque rule was between age 13 and age 14. The results of the present paper can be well explained by the Overlapping Waves Model.

The rules utilized by Chinese children are discussed and compared to other researcher's findings. One assumption of using LCA to investigate the cognitive rules is that subjects should utilize the same strategy during testing. However, the assumption is difficult to test with LCA and more complex strategy shifting model should be considered.

Key words balance scale task, latent class analysis, cognitive rule, solution strategy