DOI:10. 16382/j. cnki. 1000-5560. 2018. 01. 017

非连续性与异质性*

——多阶段混合增长模型在语言发展研究中的应用

刘　源[1]　刘红云[2]

(1. 西南大学心理学部暨认知与人格教育部重点实验室,重庆 400715;
2. 北京师范大学心理学部,北京 100875)

摘　要:多阶段混合增长模型(Piecewise Growth Mixture Modeling,PGMM)是近几年新兴的同时关注群体的发展阶段非连续性和潜在异质性的统计模型。它将多阶段增长模型和潜类别增长模型进行整合,可以描述同时存在发展转折点和不同发展类别的描述群体增长趋势的数据。文章以早期儿童的追踪研究(幼儿园版)为例,运用 PGMM 模型探索其增长趋势,得出:(1)两阶段混合增长模型能最有效地描述学生阅读能力的发展,转折点在一年级,随着年龄的增加,发展速度变慢;(2)发展趋势分为三类,大部分个体起点低、发展快,小部分个体起点高、发展慢,到三年级以后两个类别差距越来越小,另一部分整体发展都比较缓慢;(3)教师对学生行为的评价比父母的评价更能有效预测学生阅读成绩的类别和趋势。

关键词:多阶段混合增长模型(PGMM); 非连续性; 潜在异质性; 模型拟合

一、引言

追踪研究在教育学研究中应用广泛,其最大的优点是可以合理地推论变量之间的因果关系,为揭示事物的发生、发展和变化提供有效的途径。与此同时,随着结构方程模型(Structural Equation Modeling,SEM,参见 Kaplan, 2000)和多层线性模型(Hierarchical Linear Modeling,HLM,参见 Raudenbush & Bryk, 2002)等统计技术的发展,追踪研究的分析方法得到了极大的推广,而且研究设计上的限定条件越来越弱,解决的问题不断深入,越来越符合实际的情景。

潜变量增长模型(Latent Growth Modeling,LGM)就是在 SEM 的框架下,帮助研究者同时探讨整体的发展趋势和趋势之间的个体差异及其影响因素的方法(Collins & Sayer, 2001; Singer & Willett, 2003; 刘源, 骆方, 刘红云, 2014)。然而,LGM 关于"发展趋势同质(即所有群体的发展趋势相同)"的假设条件有时并不能描述特质的实际发展情况,这一假设的不合理性进而导致了结果解释上的困难。后来,研究者提出了潜类别混合增长模型(Growth Mixture Modeling,GMM,B. Muthén, 2004)来描述发展趋势存在不同质的潜类别群体,并逐渐在教育心理学(Boscardin, Muthén, Francis, & Baker, 2008; Liu, Liu, & Hau, 2016; Schaeffer, Petras, Ialongo, Poduska, & Kellam, 2003; 张洁婷, 焦璨, 张敏强, 2010)、认知神经科学(Hunter, Muthén, Cook, & Leuchter, 2010)、医学(B. Muthén & Brown, 2009)等领域得到应用。例如,Lei 等人(2011)采用混合模型的方法探讨了汉语语言认知任务上儿童不同的发

* 基金项目:中央高校基本科研业务费专项资金资助(SWU1709379)。

展类别;Ferrer 等人(2015)的研究采用混合模型,分析了儿童在语言发展上的群组差异。研究指出,语言发展的群组差异可能比想象中出现更早,在小学之前,发展迟滞的个体就呈现出明显的能力落后的现象。而且发展迟滞的群体赶上正常个体的几率也很小(Ferrer et al., 2015)。国内有学者使用混合增长模型描述学生同伴关系发展趋势的异质性,并指出大部分学生的同伴关系在三年级到六年级期间相对稳定,但是约有 8% 的学生在同伴当中的相对位置会处于不稳定状态(刘红云, 2007)。

另一方面,LGM 往往假设个体特质在某一时间段内的发展趋势连续。随着研究的不断深入,研究者注意到实际中某些特质在不同阶段的发展并不相同——即发展趋势非连续。如 Chou 等人(2004)和 Li 等人(2001)都发现,吸烟、吸毒等行为存在不同的发展阶段,并探讨了干预前后发展轨迹的变化;McCoach 等人(2006)发现学生阅读能力的发展速度在入学前后不同,增长速度受社会经济地位、性别等因素的影响。这类研究有助于更细致合理地分析个体发展的差异及原因,对特殊发展阶段的甄别和深入分析有非常重要的意义(Berninger, Abbott, Nagy, & Carlisle, 2010; Blozis, Feldman, & Conger, 2007; Choudhury, Kasman, & Plowman, 2010)。多阶段增长模型(Piecewise Growth Modeling, PGM, Raudenbush & Bryk, 2002)的提出为解决这一类问题提供了切实有效的工具,也使得追踪研究所能探讨的问题更深入和细化。

个体语言能力的发展需要经历一些关键期,如幼儿期(学龄前儿童)词汇、句法等都处于巅峰状态,而上学以后,经过学校的教育,小学期的语言朝复杂的思维能力、理解能力方向发展,词句的发展进而变缓。一项对美国儿童长期的追踪研究(ECLS - K)证实了儿童阅读的发展存在以小学入学时期为转折的发展趋势,入学前的儿童的阅读能力提高速度非常快,甚至是入学后阅读能力发展的三倍;而在上学之后,阅读能力的发展变缓,特别是在小学三年级以后,阅读能力的发展尤为缓慢(Kieffer, 2011; Liu et al., 2016; Votruba - Drzal, Li - Grining, & Maldonado - Carreño, 2008)。这些关键转折点的存在是教育者根据个体的不同情况选择恰当的教学内容和教学方式的理论依据。另一些关于阅读能力发展的研究则侧重于语言发展的群体(Grimm, Ram, & Hamagami, 2011; Palardy, 2010):有一小部分个体的发展明显迟滞于大部分个体。这些差异可能是由个体以外的(如家庭的社会经济地位、是否离异以及种族和语言等)因素造成(Dawson & Williams, 2008; Liu & Hou, 2017; Kieffer, 2011; Palardy, 2010; Sun & Li, 2011),也可能由群体平均的社会经济地位、教师教学行为等群体因素(Palardy, 2010)造成。由此可知,语言能力在所有时间段的发展趋势不一定都相同,可能存在一定的发展阶段和转折点;同时,并非所有个体都遵循同样的发展模式,而是存在不同的潜类别。所以,如果要同时考察语言能力发展的阶段性和类别特征,就需要运用新的统计手段。

综上所述,GMM 和 PGM 方法的应用使得追踪研究探讨的实际问题更加深入更加切合实际。然而,在具体的研究中,如果研究群体发展阶段不连续,研究群体又存在潜类别,且不同的潜在类有不同的转折点,这时单独使用上述某一个模型对数据的估计是不合适的。这时需要将上述两个模型结合,形成多阶段混合增长模型(Piecewise Growth Mixture Modeling, PGMM)。Li 等人(2001)的研究首次使用 PGMM 的思路去描述青少年酒精成瘾的状况,结果发现用两个类别的两阶段线性模型的结果比用线性模型、二次增长模型和不定义曲线增长类型的模型都要好;但是在其研究中没能就不同发展类别的影响因素进行更深入的分析。McAuley 等人(2011)对老年人锻炼自我效能的发展阶段和发展类别进行了研究,其研究中首先采用发展阶段模型探讨了发展的阶段,然后再采用潜在类别模型探讨潜在类别。这是一种两阶段的分析方法,如果先分析潜在类别再分析发展阶段有可能会得到不同的结果。Uher 等人(2010)采用 PGMM 模型对抑郁病人在治疗过程中抑郁症状的发展变化轨迹进行了研究;然而在其研究中,治疗方案的不同就可以对病人的发展类别做出较好的分类和解释,而一般心理特质的变化和发展原因可能更为复杂。总体来看,对 PGMM 模型的探讨目前尚处于起步阶段,如何采用 PGMM 模型确定发展阶段和发展类别需进一步探讨。本文结合儿童阅读能力发展的追踪研究数据,采用 PGMM 模型对阅读能力发展同时具有阶段类和连续类的情况进行分析,并对可能造成这些差异的影响因素进行

探讨。

二、方法

(一) 数据

采用美国国家教育统计中心(National Center for Education Statistics,NCES)提供的早期儿童的追踪研究—幼儿园版(Early Childhood Longitudinal Study – Kindergarten cohort,ECLS – K,参见 Tourangeau, Nord, Lê, Sorongon, & Najarian, 2009)。该追踪研究调查了 1998 – 1999 年入学的 21409 名学生,用长达 9 年的时间追踪调查了他们从幼儿园到八年级的智力发展情况和来自不同层面的影响因素。数据测量了 7 个时间点上学生的成绩:幼儿园秋季学期,幼儿园春季学期,一年级秋季学期(30% 的样本),一年级春季学期,三年级春季学期,五年级春季学期和八年级春季学期(Y_1 – Y_7)。测量的指标有直接认知能力(包括阅读成绩、数学成绩和科学成绩)、间接认知能力(包括教师对学生认知能力的整体评价,采用行为量表计分)、学生的自我描述以及学生的自我概念和控制点。影响的因素包括学生个体层面、家长层面、教师层面以及学校或社区层面的变量。

由于第三次测量(一年级下期,Y_3)的时间只对其中 30% 的被试进行了调查,本研究中没有采用这一次的调查数据,只使用了其余六次测量的数据。ECLS – K 调查采用了复杂的抽样,对于用于不同目的分析的数据库提供了不同的权重,可以通过合理加权分析总体的特征。对于这六次测量,采用数据库中提供的这六次追踪的权重(C1_7FC0)对数据加权。本研究考察的目标变量上均不含缺失值的有效样本 7803 名,其中男生占 51.9%,女生占 48.1%。第一次测试时孩子的平均年龄 68.47 个月,标准差 4.21 个月,最大值 79 个月,最小值 54 个月。

(二)测量变量

1. 结果变量:阅读能力

阅读能力主要测量了学生阅读理解方面的技能,从幼儿园到八年级,主要包括:(1)字母认读,(2)发音,(3)词汇,(4)词义理解与解释,(5)文章及文学评价,(6)复杂句法和高级词汇的理解与评价。随着学生阅读能力的发展,不同年级的学习内容不断加深,某些基本的能力如字母认读在高年级不再考核,而是逐渐增加阅读材料和阅读任务的难度。

为了考察阅读能力随着年龄增长所发生的变化,需要将不同年龄学生的阅读能力放在同一的量尺上,因此需要建立可以跨年龄比较的发展量尺。在测验设计上,相邻年龄的测试题目有一定的重叠,以保证不同次的测试结果可以进行链接(对应数据库中的变量 C1R4RTHT_R – C2R4RTHT_R、C4R4RTHT_R – C7R4RTHT_R)。关于阅读测验设计进一步的信息可以参考相应的技术报告(Tourangeau et al., 2009)。

2. 影响阅读能力的协变量

(1)学习方法的评定

采用社会评定量表(Social Rating Scale,SRS)由教师和家长分别对孩子与学习有关的行为和做事方法进行了评定。评定采用"1(从不)"到"4(经常)"四点李克特式量表。

家长评定量表中的题目主要有:(1)直到完成任务之前都一直学习;(2)对很多事情都感兴趣;(3)专心于任务之上,不受到其他干扰;(4)帮助做日常杂务;(5)对于学习新东西很热情;(6)在学习和玩耍中都很具创造性。第一次评定中这几个题目得分的平均值作为家长对孩子学习方法的初始评价(数据库中第一次测量的 P1LEARN 变量,转折点测量的 P4LEARN)。

教师评定量表中的题目主要有:(1)把东西收拾整齐;(2)对学习新事物具有热情;(3) 独立完成作业;(4)能轻松适应日常生活的变化;(5)对完成任务坚持不懈;(6)注意力集中。第一次评定中这几个题目得分的平均值作为教师对孩子学习方法的初始评价(数据库中的 T1LEARN 变量,转折点测量的 T4LEARN)。

(2)背景变量

所考察的协变量除了家长和教师对孩子学习方法的初始评定外,还包括孩子的背景特征,主要有性别(GENDER)、第一次测试时的年龄(R1_KAGE)、在家是否说英语(WKLANGST)以及社会经济地位(WKSESL)。其中社会经济地位是通过对父母访谈得到的综合反应家庭社会和经济水平的指标。主要测查了父亲教育水平、母亲教育水平、父亲职业、母亲职业和家庭收入五个方面的内容。

(三)研究假设

假设 1:对于阅读能力的发展趋势,假设其分为两阶段:第一阶段(幼儿园到小学一年级)阅读能力呈线性增长趋势,一年级为阅读能力发展的转折期;第二阶段(一年级到八年级)阅读能力呈非线性增长趋势,增长速度随着年龄的增长有减慢的趋势。

假设 2:阅读能力的发展存在不同的潜在类别,不同类别的发展水平和增长趋势存在差异。

假设 3:在控制性别、年龄、家庭交流使用语言和社会经济地位后,幼儿时期教师对孩子学习方式的评价对阅读能力的发展类别有显著的预测作用,而家长对孩子幼儿期的评价对阅读能力的发展类别和趋势预测作用不显著。

假设 4:在控制性别、年龄、家庭交流使用语言和社会经济地位后,第一阶段教师对孩子学习方式的初始评价和家长对孩子学习方法的初始评价对阅读能力的发展水平和发展速度有显著的预测作用,第二阶段教师对孩子学习方式的初始评价和家长对孩子学习方法的初始评价对阅读能力的发展速度有显著的预测作用。

(四)数据分析方法

首先采用 LGM 探讨阅读能力随时间发展的趋势。本研究定义两种模型:二次性发展趋势模型,第一阶段为线性发展、第二阶段为二次性发展的阶段发展模型(转折点在小学一年级)。然后根据所选择的增长趋势模型,采用 GMM 探讨发展存在的不同潜在类别。最后采用多分类变量 Logistic 回归(Multinomial Logistic Regression)分析协变量对不同类别的预测作用;采用含有预测变量的潜变量混合增长模型探讨预测变量对发展趋势的影响。

以上分析均采用 Mplus7.0 软件(L. Muthén & Muthén, 2012),应用 MLR 估计方法分析数据。

(五)分析模型

对于阅读能力的发展,两阶段增长模型的一般定义如图 1 所示。

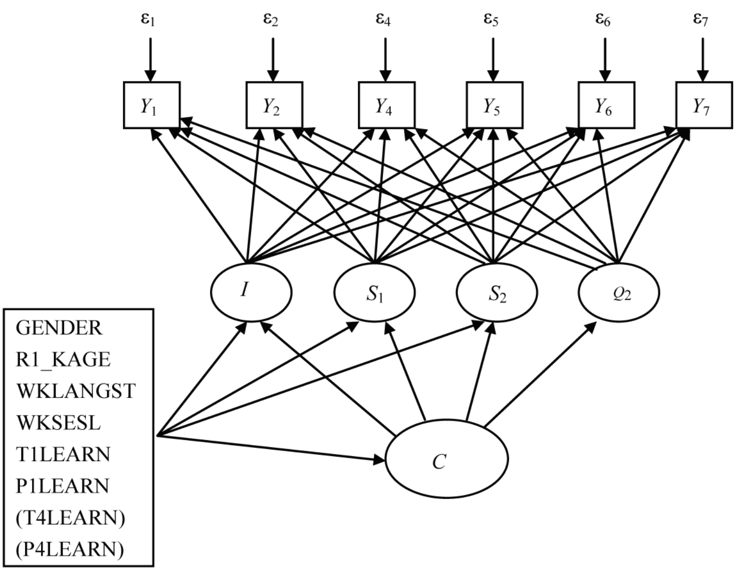

图 1 两阶段混合增长模型

图中 $Y_i(i=1,2,4,5,6,7)$ 表示第 i 次测试得到的阅读能力，截距 I 表示阅读能力的初始水平，S_1 表示第一阶段增长直线的斜率，S_2 表示第二阶段二次增长曲线的线性部分，Q_2 表示第二阶段二次增长曲线的二次型部分。结合本研究的实际数据，截距 I 对应于六次测量的路径系数固定为(1, 1, 1, 1, 1, 1)，S_1 对应于六次测量的路径系数固定为(0, 0.5, 1.5, 1.5, 1.5, 1.5)，S_2 对应于六次测量的路径系数固定为(1.5, 1.5, 1.5, 3.5, 5.5, 8.5)，Q_2 对应于六次测量的路径系数固定为(2.25, 2.25, 2.25, 12.25, 30.25, 72.25)。即：

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1.5 & 2.25 \\ 1 & 0.5 & 1.5 & 2.25 \\ 1 & 1.5 & 1.5 & 2.25 \\ 1 & 1.5 & 3.5 & 12.25 \\ 1 & 1.5 & 5.5 & 30.25 \\ 1 & 1.5 & 8.5 & 72.25 \end{pmatrix} \begin{pmatrix} I \\ S_1 \\ S_2 \\ Q_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{pmatrix}$$

其中 I，S_1，S_2 和 Q_2 是连续的潜变量，用来描述增长趋势，这部分的模型相当于一个潜变量增长曲线模型。不同的增长趋势的模型可以通过路径系数的定义得到。如线性模型可以通过将截距对应的路径系数固定为(1, 1, 1, 1, 1, 1)，斜率的路径系数固定为(0, 0.5, 1.5, 3.5, 5.5, 8.5)得到；二次增长曲线模型可以通过在线性模型的基础上增加一个路径系数固定为(0, 0.25, 2.25, 12.25, 30.25, 72.25)的二次型潜变量得到。

C 表示一个潜在分类变量，即描述变化类别的分类变量(Trajectory class variable)，用来描述变化趋势可能存在的类别。本研究中进一步考察协变量对潜在分类 $C(k=1,2,\cdots,K)$ 的影响，由于因变量为分类变量，可以采用 Logistic 回归对协变量的影响进行分析：

$$P(C_i = k \mid x_i) = \frac{e^{\beta_{0k}+\beta_{1k}x_i}}{\sum_{s=1}^{K} e^{\beta_{0k}+\beta_{1k}x_i}}$$

以最后一类(第 K 类)为参照类，第 $k(k=1,2,\cdots K-1)$ 与第 K 类发生概率比值的对数为：

$$\log\left(\frac{P(C_i = k \mid x_i)}{P(C_i = K \mid x_i)}\right) = P(C_i = k \mid x_i) - P(C_i = K \mid x_i) = \beta_{0k} + \beta_{1k}x_i$$

因此 β_{1k} 表示协变量每增加一个单位，第 k 类相对于第 K 类对数发生比的增加。

由于因变量为连续变量，协变量对增长速度的影响采用回归模型进行分析：

$$\begin{pmatrix} I \\ S_1 \\ S_2 \\ Q_2 \end{pmatrix} = \begin{pmatrix} \beta_{0I} \\ \beta_{0SI} \\ \beta_{0S2} \\ \beta_{0Q2} \end{pmatrix} + \begin{pmatrix} \beta_{1I} \\ \beta_{1SI} \\ \beta_{1S2} \\ \beta_{1Q2} \end{pmatrix} X_1 + \cdots + \begin{pmatrix} \beta_{mI} \\ \beta_{mSI} \\ \beta_{mS2} \\ \beta_{mQ2} \end{pmatrix} X_m + \begin{pmatrix} \zeta_I \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{pmatrix}$$

三、结果

(一)阅读成绩及协变量的描述统计结果

表 1 给出了六次测量阅读测验成绩、家庭社会经济地位、第一次测试中父母对孩子学习方式的评定以及教师对孩子学习方式的评定之间的相关矩阵，以及各变量的均值和标准差。

(二)阅读测验发展趋势及潜类别分析结果

1. 增长趋势分析

表 1 六次阅读成绩以及协变量的相关矩阵与均值、标准差

	阅读 1	阅读 2	阅读 4	阅读 5	阅读 6	阅读 7	SES	父母评价	教师评价
阅读 2	.793**								
阅读 4	.669**	.777**							
阅读 5	.603**	.657**	.762**						
阅读 6	.584**	.639**	.721**	.852**					
阅读 7	.542**	.559**	.604**	.746**	.795**				
SES	.147**	.132**	.129**	.157**	.162**	.178**			
父母评价	-.019**	-.014**	-.028**	-.012**	-.027**	-.024**	.540**		
教师评价	.269**	.264**	.244**	.240**	.233**	.204**	.092**	.050**	
M	-1.277	-0.715	0.128	0.798	1.046	1.293	0.014	3.492	3.086
SD	0.506	0.491	0.453	0.315	0.299	0.383	0.789	1.505	0.960

** $p < .01$

为了考察阅读能力的发展趋势,定义如下发展趋势:

模型 1:连续线性发展趋势模型,即假设阅读能力的发展不存在转折点,从幼儿园到八年级呈线性增长趋势;

模型 2:连续二次型发展趋势模型,即假设阅读能力的发展不存在转折点,从幼儿园到八年级呈二次型增长趋势;

模型 3:两阶段线性发展趋势模型,即假设阅读能力的发展存在转折点,第一阶段(从幼儿园到一年级)和第二阶段(从一年级到八年级)的发展都呈线性增长趋势,但是两个阶段的发展速度存在差异;

模型 4:两阶段二次型发展趋势模型,即假设阅读能力的发展存在转折点,第一阶段(从幼儿园到一年级)呈线性发展趋势,第二阶段(从一年级到八年级)的发展呈二次型增长趋势。

表 2 给出了四个备选模型的整体拟合结果。

表 2 四个备选模型与数据的整体拟合

模型	χ^2	df	CFI	RMSEA	AIC	BIC	ABIC
模型 1	52797.438	16	0.000	0.612	56593.443	56670.016	56635.060
模型 2	18906.476	15	0.433	0.449	26187.379	26291.796	26244.129
模型 3	11728.063	12	0.648	0.354	17145.410	17249.827	17202.160
模型 4	4415.809	7	0.868	0.284	8426.164	8565.386	8501.830

从表 2 的结果可以看出,在所比较的模型中,模型 4 的整体拟合指标 CFI 最高,而 AIC、BIC 的值也最小,对数据的拟合最好。

表 3 零模型(模型 4)参数估计结果

	固定部分估计值	*SE*	*t*	随机部分方差	*SE*	χ^2
截距	-1.715	0.008	-226.817***	0.337	0.008	43.815***
线性斜率 1	0.960	0.004	257.225***	0.017	0.002	10.318***
线性斜率 2	0.368	0.003	143.902***	0.004	0.001	4.646***
曲线斜率 2	-0.022	<0.001	-94.611***	<0.001	<0.001	<0.001

*** $p < .001$

整体来看,初始水平越高,第一阶段的增长速度越慢(截距与线性斜率 1 的协方差为 -0.051,$p < .001$),第二阶段的线性增长速度也越慢(截距与线性斜率 2 的协方差为 -0.050,$p < .001$),而第二阶段二次型变化越快(截距与曲线斜率 2 的协方差为 0.004,$p < .001$)。

初始状态阅读能力的估计值为 -1.715,第一阶段儿童的阅读能力存在显著的线性增长,斜率为 0.960;第二阶段存在显著的二次型变化趋势,说明第二阶段的增长速度明显低于第一阶段,且增长速

度随着年龄增长有减缓的趋势。截距和第一阶段与第二阶段增长的斜率均存在显著的个体间差异。

2. 阅读能力发展类别分析

对于混合增长模型,常常采用信息指标 AIC、BIC 和 ABIC 来确定模型中包含的潜在类别,BIC 的值越小说明模型与数据拟合越好。但是对于样本量较大的情况,采用 BIC 指标往往倾向于选择类别较多的模型,Logan 和 Petscher(2010)建议采用碎石图的方法来确定潜在类别的个数。图 2 给出了潜变量的类分别为 1 – 5 时,BIC 的变化趋势。从图中可以看出,增加到三个类型以后,BIC 的变化趋势很小,因此对于阅读能力的发展趋势可选择三个类别模型。

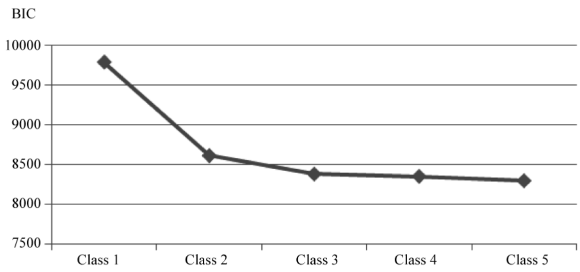

图 2　不同类别混合增长模型 BIC 的变化

对于三类别的模型,77.0%(6009 名)的学生的阅读能力的发展属于正常组(中等组),11.5%(894 名)的学生的阅读能力的发展水平较低(低分组),另有 11.5%(895 名)的学生的阅读能力的发展水平相对较高(高分组)。表 4 给出了不同发展组增长参数的估计结果。图 3 给出了三类不同发展组六次阅读能力发展的模型估计的结果和实际观测结果。

从表 4 和图 3 的结果可以看出,对于三个不同发展组,第一阶段发展速度都较快,但是相对来讲低水平能力组的增长速度稍慢;而到了第二阶段,相对于正常发展组,能力水平较低的组发展速度略微高于高水平发展组,而高水平发展组,到了第二阶段,发展速度反而减慢。从图 3 可以看出,到了七年级以后,三个组之间的水平差异相对初始状态减小。对于随机部分,所有的类别中,截距、斜率的方差均显著,说明在各类的内部,个体之间的发展是存在差异的。

表 4　不同发展组的增长曲线参数估计结果

	缓慢发展组			中等发展组			快速发展组		
	估计值	*SE*	*t*	估计值	*SE*	*t*	估计值	*SE*	*t*
截距	–1.768	0.048	–36.914***	–2.416	0.054	–44.821***	–1.236	0.105	–11.783***
线性斜率 1	0.977	0.013	76.282***	0.873	0.05	17.327***	0.983	0.028	35.220***
线性斜率 2	0.378	0.014	26.955***	0.530	0.02	25.883***	0.256	0.024	10.575***
曲线斜率 2	–0.023	0.001	–20.330***	–0.034	0.002	–20.741***	–0.013	0.002	–5.540***

*** $p < .001$

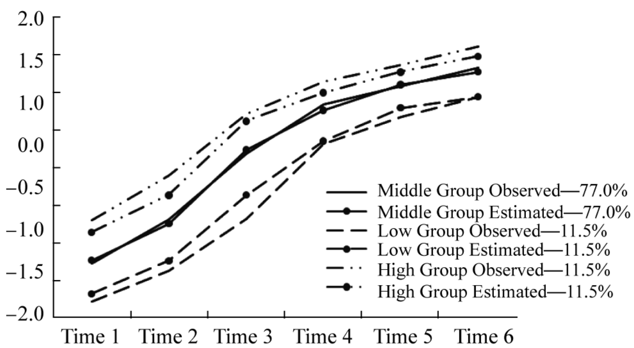

图 3　阅读能力发展类别及趋势

(三)阅读发展潜类别的预测

下面进一步考察学生背景变量性别、年龄、在家中交流所使用语言、家庭社会经济地位、父母对学生幼儿时期学习方法的评定和教师对学生幼儿期学习方法的评价对发展类别的预测(Liu & Hou, 2017)。对于分类变量,给出每个类别上预测变量不同取值人数所占的比例,对于连续变量给出不同类别上预测变量的均值和标准差(表 5)。

表 5 三个不同类别上预测变量的描述统计结果

	中等发展组		缓慢发展组		快速发展组	
	频次	百分比	频次	百分比	频次	百分比
性别(男)	2970	73.4%	613	15.2%	463	11.4%
性别(女)	3040	81.0%	282	7.5%	432	11.5%
家庭语言非英语	663	74.9%	167	18.9%	55	6.2%
家庭语言英语	5173	77.4%	686	10.3%	825	12.3%
	M	*SD*	*M*	*SD*	*M*	*SD*
年龄(月)	68.446	4.136	67.700	4.649	69.378	4.076
社会经济地位	0.270	1.661	-0.042	2.063	0.593	1.357
父母评价	3.488	1.475	3.626	2.067	3.383	0.911
教师评价	3.034	0.639	2.489	0.627	3.299	0.580

分类 Logistic 回归分析结果表明(参数估计结果见表 6),总体来看,预测变量对阅读能力的发展类别有显著的预测作用($\chi^2(12)=1201, p<.001$),Pseudo $R^2=0.198$。预测变量与阅读能力发展潜在类别之间关系的似然比检验结果表明:性别对阅读能力的发展类别有显著的预测作用($\chi^2(2)=43.260$, $p<.001$),男生阅读能力发展为缓慢组的发生比是中间组的 1.703 倍,说明男生阅读能力发展缓慢的人数比例相对女生较多;阅读能力发展为较快组的发生比是中间组的 1.192 倍,与 1 很接近,说明男生阅读能力发展较快的人数比例相对女生差异不大;除此之外,初始测试时的年龄($\chi^2(2)=39.686$, $p<.001$)、家庭交流使用语言($\chi^2(2)=11.125, p=.004$)、学生社会经济地位($\chi^2(2)=360.216$, $p<.001$)对阅读能力的发展类别有显著的影响,且后两者的影响更显著。在控制了性别、年龄、家庭交流使用语言以及社会经济地位后,家长对孩子初始学习方式的评价对阅读能力发展类别的预测作用不显著($\chi^2(2)=2.708, p=.258$);而教师对孩子初始学习方式的评价对阅读能力发展类别的预测作用显著($\chi^2(2)=418.959, p<.001$),教师对学生学习方法的评定每增加一个单位,学生阅读能力发展为缓慢组的概率是正常组概率的 0.340 倍,为高能力组的概率是正常组的 1.852 倍。

表 6 阅读能力发展类别预测的参数估计结果

		标准化系数	*SE*	Wald 检验值	*p*	发生比
缓慢发展组						
	截距	2.942	0.665	19.591	<.001	
	社会经济地位	-0.704	0.057	150.117	<.001	0.495
	年龄	-0.037	0.01	14.024	<.001	0.964
	父母评价	0.020	0.028	0.535	.464	1.020
	教师评价	-1.080	0.066	269.708	<.001	0.340
	性别	0.532	0.086	38.741	<.001	1.703
	家庭语言	0.305	0.109	7.805	.005	1.357
快速发展组						
	截距	-6.83	0.647	111.434	<.001	
	社会经济地位	0.633	0.049	167.602	<.001	1.884
	年龄	0.043	0.009	22.539	<.001	1.044
	父母评价	-0.059	0.043	1.89	.169	0.942
	教师评价	0.616	0.066	86.118	<.001	1.852
	性别	0.175	0.077	5.195	.023	1.192
	家庭语言	-0.251	0.154	2.666	.102	0.778

注:对于阅读能力发展,参考组为中间组;对于分类预测变量,性别的参照组为女生,家庭交流所使用语言的参照组为使用英语组。

(四)发展趋势的预测

假设阅读能力的发展水平和速度受学生背景特征和学习方法的影响,分析结果见表7。

表7　阅读能力发展水平和速度影响因素分析结果

	缓慢发展组		中等发展组		快速发展组	
	标准化系数	t	标准化系数	t	标准化系数	t
初始水平(截距)						
性别	-.138	-4.184	-.014	-1.138	-.017	-.467
家庭语言	.052	1.599	.133	10.382	.012	.337
初始年龄	.243	7.322	.142	11.239	.139	3.847
社会经济地位	.340	10.424	.203	15.583	.192	5.156
父母评价	.010	.302	.006	.513	.177	4.944
教师评价	.133	4.025	.296	22.786	.146	4.096
第一阶段斜率						
性别	.140	4.163	.044	3.170	-.080	-2.280
家庭语言	-.017	-.520	-.039	-2.767	-.060	-1.690
初始年龄	-.202	-5.980	-.139	-10.151	-.208	-5.973
社会经济地位	-.341	-10.273	-.108	-7.610	.303	8.495
父母评价	.019	.579	-.001	-.101	.068	1.977
教师评价	-.139	-4.149	-.100	-7.064	.119	3.462
第二阶段斜率						
性别	.068	1.839	-.013	-.925	-.045	-1.189
家庭语言	-.104	-2.856	-.110	-7.793	-.037	-.987
初始年龄	-.193	-5.171	-.116	-8.370	-.121	-3.247
社会经济地位	-.016	-.447	.077	5.424	.003	.067
父母阶段评价	.011	.311	.005	.333	-.170	-4.623
教师阶段评价	-.008	-.219	-.102	-7.167	.009	.245

对于初始水平,缓慢发展组男生成绩显著高于女生($t=-4.184$, $p<.01$)。而家庭语言只对中等发展组有影响($t=10.382$, $p<.001$),英语为母语的家庭阅读成绩高。对于所有的组,初始年龄越高、社会经济地位越高,学生的成绩越高。在控制了背景变量的条件下,教师对行为的评价越高,学生的阅读能力越高,而父母评价只对快速发展组起到预测作用($t=4.944$, $p<.01$)。对于第一阶段的斜率,控制了背景变量的情况下,教师评价越高的个体,缓慢发展组和中等发展组的发展速度越慢($t=-4.149$, $p<.01$; $t=-7.064$, $p<.01$),快速发展组与前两组的趋势相反,即教师评价越高的个体,发展速度也越快($t=3.462$, $p<.05$);然而父母的评价不能预测第一阶段的发展速度。对于第二阶段的发展速度,对中等发展组的预测有效,教师的评价越高,中等组的发展速度越慢($t=-7.167$, $p<.01$);父母评价只对快速发展组有预测作用,评价越高,发展速度也越慢($t=-4.623$, $p<.01$)。

四、讨论

PGMM可以同时考虑追踪研究中存在的不同质的发展类别和非连续的发展趋势的问题,因此从方法上解决了传统两阶段分析方法的弊端。这一方法的应用可以为心理学中关于特质的发展趋势和类别的研究提供更加合理的分析,尤其是对发展研究中关于个体差异、不同类别及其原因的分析,可以帮助研究者更加深入细致地探讨心理特质发展的规律和特征,从而更有针对性地制定干预策略和方案。由于PGMM将发展趋势和发展类别统一起来考虑,虽然在本文的数据中不同潜在类别的发展趋势都是二次型的阶段发展模型,转折点也相同,但是方法本身并没有这一限制。也就是说,研究者可以在不同类别中定义不同的发展趋势(如一类别是线性发展趋势,而另一类别是非线性发展趋势),也可以允许不同的类别有不同的转折点(如一类别的转折点在第三次测量,而另一类的转折点在第四次测量)。

对于PGMM而言,一个非常重要的问题是模型的选择。分类数目越多,所能描述的子群体越多,一

致性越高。但是有研究指出,潜在类的类别数量会影响估计的准确性和类的解释的问题,所以潜在类的数量并非越多越好(Nylund, Asparouhov, & Muthén, 2007)。对于潜在类别分析模型、混合因素分析和潜变量增长模型这类"潜类别"模型而言,有众多决定分类个数的信息指标(Information Criterion, IC),Mplus 软件里常列出的有 AIC、BIC 和 ABIC。Nylund 等学者(2007)建议使用 BIC,因为它在不同的潜类别分析当中具有一定的稳健性,优于其他指标,而且相对于 AIC,BIC 是经过样本矫正的。Masyn 等学者(2010)运用"尺类谱"(Dimensional - Categorical Spectrum, DCS)进行潜在结构探索时考虑了众多模型判断的标准,认为 BIC 并不是选择模型的唯一标准,需要根据实际数据的情况,如相关大小、潜在类别的大小等因素综合考虑。然而对于实际的解释而言,分类过多,会对各类别的解释造成困难,甚至出现交叉类别的情况。尽管 BIC 拟合指标能随着类别数目的增加而减少,但是其减少程度会随之减小,这意味着如果需要达到更小的 BIC,需要分出更多的类别,需要估计的参数也越多,模型越复杂。所以 Logan 和 Petscher(2010)建议采用 BIC 碎石图考察 BIC 减小的程度是否随着类别数量减小。与此同时需要考虑其他的分类指标,比如分类判别准确率、不同类别所占比例等。所以,分类是一个需要付出代价的过程,它并不是完全靠数据驱动的,BIC 只是一个最常用的参考指标。在统计上需要遵循从简原则,即用简单的模型解释复杂的现象,要在符合数据的基础上选择最简单最易于解释的模型。应该特别注意的是,分类的结果应该具有实际的意义,应该从理论上是可以解释的。最近有研究指出,可以先采用 BIC 作为首选分类标准选出"真实的"模型,进而根据熵值(entropy)和返还指数(recover index)来确认"更好的"最终模型(刘源等, 2014)。

关于发展趋势的问题,以往研究提及的阶段发展都比较符合阅读能力发展的实际情况,但是大多只考虑阶段线性增长的模型。从本文所采用的模型估计结果可以看到,阅读能力的发展第二阶段二次型发展的趋势能比线性发展更好地拟合数据,这就更进一步解释了不同阶段发展趋势的差异。Grimm 等人(2010)利用相同的数据提出了学生阅读成绩发展的混合 Gompertz 模型,该模型和 Logistic 模型都常用在种族生长的生物学领域,来描述个体的发展受时间、固有生长率和一定区域内总体数量影响的情况。Gompertz 曲线所描述的主要发展特征是初期增长缓慢,以后逐渐加快,当达到一定程度后,增长率又逐渐下降。这与本文探讨的阅读能力发展趋势有一定的共同之处,不过初期的缓慢发展阶段应该发生在幼儿园以前。本研究所探讨的这一发展时期(幼儿园到八年级),主要发展特征是先快速发展,然后增长速度逐渐下降。从模型拟合和可解释的角度来看,一些非线性模型(如 Gompertz 曲线,指数增长曲线或对数增长曲线)虽然也可能得到好的结果,但是这一类模型往往过于复杂,要求有较多的测量时间点,在教育学和心理学的研究中并不实用,同时也无法直观解释阅读能力的发展特征。且该类曲线模型仍然是连续发展模型,不能分析出发展曲线的转折点。温忠麟等学者提出增长模型中的交互效应分析框架,来针对较为复杂的协变量模式对增长曲线的影响(温忠麟, 刘红云, 侯杰泰, 2012; 温忠麟, 侯杰泰, 张雷, 2006)。这些模型进一步将教育学和心理学的研究问题细化,通过理论建构和实际数据,来探究个体如何变化,以及受到哪些外在因素的影响。一个实际的发展问题,其发展趋势往往非常复杂,需要灵活考虑阶段发展、非线性发展以及不同类别发展的情况,因此要使用"合适"的模型,并且要易于解释。基于不同发展类别,可以进一步考虑每个类别各自的影响因素,从而更精确地解决学生在发展当中的问题,针对个别的发展实施个别的教育,这才是教师和研究者应该关心的问题。

PGMM 在参数估计上随着模型复杂程度的提高(如潜在类的增多),可能会存在模型收敛、估计方差协方差矩阵非正定、似然函数局部最大化等问题(Hipp & Bauer, 2006)。在实际应用中,这些问题可以通过在模型中增加一些限定条件解决,如限定不同类增长趋势的方差协方差矩阵相同。Hipp 和 Bauer(2006)在其研究中指出,对于一组初始值,只运行一次可能会出现偏差的结果,为了避免局部解的问题,往往需要采用不同组的初始值,多次运行结果。实际应用者可以通过在估计方法中改变初始值的组数(如在 Mplus 中采用 starts,并增大 starts 的数量)来解决局部解的问题。但是如何更加有效简便地使用这一解决策略有待进一步在方法上进行研究。

五、总结与展望

本研究用 PGMM 对发展群体进行了分析，基于 ECLS - K 的样本，得出：(1)两阶段混合增长模型能最有效地概括学生阅读能力发展的趋势，转折点在一年级，第一阶段为线型趋势，第二阶段为二次性趋势，且随着年龄的增加，发展速度变慢；(2)发展趋势分为三类，大部分个体属于起点低、发展快的类别，小部分个体起点高、发展慢，到三年级以后两个类别的差距越来越小，第三类整体发展都比较缓慢；(3)教师的评价比家长的评价更能有效预测学生的阅读成绩的类别和趋势。

本研究为同类研究提供了一种新的统计方法和新的研究思路。采用多阶段混合模型不仅可以针对发展模型不连续、发展群体不同质的问题进行解答，而且能将许多实际问题考虑在内，为相关研究提供了一种简单且易操作的模型。同时，本研究强调了分别对待小部分群体的重要性，为教育者提供了理论上的数据支持，为今后教育朝更小、更精、更人性的方向发展提供了依据。但是，本研究仅仅只考虑了个体层面的变量的影响，并没有涉及群体层面变量的影响，比如学校和教师的氛围、社区的经济状况等。本研究也没有考虑随时间变化的变量对发展趋势的影响。基于此，多水平的结构方程模型的应用可以为进一步探讨这类复杂问题提供方法。

PGMM 将为教育学和心理学领域中关于发展趋势和发展类别的研究提供更加合理的分析方法和工具，有助于从实证研究的角度进一步深化心理学的研究。另外，含有预测变量的 PGMM 的应用将从更加合理和综合的角度对影响个体发展趋势和导致个体差异的因素进行分析和探讨，从而为教育和实验干预提供一些有价值的信息。

参考文献

刘红云．(2007)．如何描述发展趋势的差异：潜变量混合增长模型．*心理科学进展*，15(3)，539 - 544．

刘源，骆方，刘红云．(2014)．多阶段混合增长模型的影响因素：距离与形态．*心理学报*，46(9)，1400 - 1412．

温忠麟，刘红云，侯杰泰．(2012)．*调节效应和中介效应分析*．北京：教育科学出版社．

温忠麟，侯杰泰，张雷．(2006)．有中介的调节和有调节的中介变量．*心理学报*，38(3)，448 - 452．

张洁婷，焦璨，张敏强．(2010)．潜在类别分析技术在心理学研究中的应用．*心理科学进展*，18(12)，1991 - 1998．

Berninger, V., Abbott, R., Nagy, W., & Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in Grades 1 to 6. *Journal of Psycholinguistic Research*, 39(2), 141 - 163.

Blozis, S. A., Feldman, B., & Conger, R. D. (2007). Adolescent alcohol use and adult alcohol disorders: A two - part random - effects model with diagnostic outcomes. *Drug and Alcohol Dependence*, 88, *Supplement* 1(0), S85 - S96.

Boscardin, C. K., Muthén, B., Francis, D. J., & Baker, E. L. (2008). Early identification of reading difficulties using heterogeneous developmental trajectories. *Journal of Educational Psychology*, 100(1), 192 - 208.

Chou, C. - P., Yang, D., Pentz, M. A., & Hser, Y. - I. (2004). Piecewise growth curve modeling approach for longitudinal prevention study. *Computational Statistics & Data Analysis*, 46(2), 213 - 225.

Choudhury, K. R., Kasman, I., & Plowman, G. D. (2010). Analysis of multi - arm tumor growth trials in xenograft animals using phase change adaptive piecewise quadratic models. *Statistics in Medicine*, 29(23), 2399 - 2409.

Collins, L. M., & Sayer, A. G. (2001). *New methods for the analysis of change*. Washington, DC: American Psychological Association.

Dawson, B. A., & Williams, S. (2008). The impact of language status as an acculturative stressor on internalizing and externalizing behaviors among Latino/a Children: A longitudinal analysis from school entry through third grade. *Journal of Youth and Adolescence*, 37(4), 399 - 411.

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. E., Michaels, R., & Shaywitz, S. E. (2015). Achievement gap in reading is present as early as first grade and persists through adolescence. *The Journal of Pediatrics*, *Aug* 2015.

Grimm, K. J., Ram, N., & Estabrook, R. (2010). Nonlinear structured growth mixture models in Mplus and OpenMx. *Multivariate Behavioral Research*, 45(6), 887 - 909.

Grimm, K. J., Ram, N., & Hamagami, F. (2011). Nonlinear growth curves in developmental research. *Child Development*, 82(5), 1357 - 1371.

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods*, 11(1), 36 - 53.

Hunter, A. M., Muthén, B. O., Cook, I. A., & Leuchter, A. F. (2010). Antidepressant response trajectories and quantitative electroencephalography (QEEG) biomarkers in major depressive disorder. *Journal of Psychiatric Research*, 44(2), 90 – 98.

Kaplan, D. (2000). *Structural equation modeling: Foundations and Extensions*. Thousand Oaks, CA: Sage.

Kieffer, M. (2011). Before and after third grade: Longitudinal evidence for the shifting role of socioeconomic status in reading growth. *Reading and Writing*, 1 – 22.

Lei, L., Pan, J., Liu, H., McBride – Chang, C., Li, H., Zhang, Y., et al. (2011). Developmental trajectories of reading development and impairment from ages 3 to 8 years in Chinese children. *Journal of Child Psychology and Psychiatry*, 52(2), 212 – 220.

Liu, Y., & Hou, S. (2017). Potential reciprocal relationship between motivation and achievement: A longitudinal study. *School Psychology International*, 014303431771057. doi: 10.1177/0143034317710574.

Liu, Y., Liu, H., & Hau, K. T. (2016). Reading ability development from kindergarten to junior secondary: Latent transition analyses with growth mixture modeling. *Frontiers in Psychology*, 7:1659.

Logan, J. A. R., & Petscher, Y. (2010). School profiles of at – risk student concentration: Differential growth in oral reading fluency. *Journal of School Psychology*, 48(2), 163 – 186.

Malone, P. S., Van Eck, K., Flory, K., & Lamis, D. A. (2010). A mixture – model approach to linking ADHD to adolescent onset of illicit drug use. *Developmental Psychology*, 46(6), 1543 – 1555.

Masyn, K. E., Henderson, C. E., & Greenbaum, P. E. (2010). Exploring the latent structures of psychological constructs in social development using the dimensional – Categorical spectrum. *Social Development*, 19(3), 470 – 493.

McAuley, E., Mailey, E. L., Mullen, S. P., Szabo, A. N., Wojcicki, T. R., White, S. M., et al. (2011). Growth trajectories of exercise self – efficacy in older adults: Influence of measures and initial status. *Health Psychology*, 30(1), 75 – 83.

McCoach, D. B., O'Connell, A. A., Reis, S. M., & Levitt, H. A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *Journal of Educational Psychology*, 98(1), 14 – 28.

Muthén, B. (2004). *Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data*. Thousand Oaks, CA: Sage.

Muthén, B., & Brown, H. C. (2009). Estimating drug effects in the presence of placebo response: Causal inference using growth mixture modeling. *Statistics in Medicine*, 28(27), 3363 – 3385.

Muthén, L., & Muthén, B. (2012). Mplus (Version 7.0). Los Angeles, CA: Muthén & Muthén.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(4), 535 – 569.

Palardy, G. J. (2010). The multilevel crossed random effects growth model for estimating teacher and school effects: Issues and extensions. *Educational and Psychological Measurement*, 70(3), 401 – 419.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*(*2nd ed*). Thousand Oaks, California: Sage Publications.

Schaeffer, C. M., Petras, H., Ialongo, N., Poduska, J., & Kellam, S. (2003). Modeling growth in boys' aggressive behavior across elementary school: Links to later criminal involvement, conduct disorder, and antisocial personality disorder. *Developmental Psychology*, 39(6), 1020 – 1035.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Sun, Y., & Li, Y. (2011). Effects of family structure type and stability on children's academic performance trajectories. *Journal of Marriage and Family*, 73(3), 541 – 556.

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., and Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of* 1998 – 99 (*ECLS – K*), *Combined User's Manual for the ECLS – K Eighth – Grade and K* – 8 *Full Sample Data Files and Electronic Codebooks* (*NCES* 2009 – 004). National Center for Education Statistics, Institute of Education Sciences, U. S. Department of Education. Washington, DC.

Uher, R., Muthén, B., Souery, D., Mors, O., Jaracz, J., Placentino, A., et al. (2010). Trajectories of change in depression severity during treatment with antidepressants. *Psychological Medicine*, (40), 1267 – 1277.

Votruba – Drzal, E., Li – Grining, C. P., & Maldonado – Carreño, C. (2008). A developmental perspective on full – versus part – day kindergarten and children's academic trajectories through fifth grade. *Child Development*, 79(4), 957 – 978.

(责任编辑　胡　岩)

Evolution of Adult Education in Britain in the first Half 20th Century: From Working Class Education to Mass Adult Education

LI Li
(School of Education, Zhejiang University, Hangzhou 310028, China)

Abstract: After the Industry Revolution, emphasis was placed on educating the working class in British adult education, which met their needs for education to a certain extent. From the beginning of 20^{th} century to 1956, British adult education was shaped by the competition among three main actors: Labour Colleges, the Workers' Education Association, University Adult Education. The independent working class education offered by Labour Colleges turned into the training means of Trade Unions' executives while the liberal adult education led by the Workers' Education Association and challenged by university adult education, graduated into mass adult education from working class education. Meanwhile, general higher adult education dominated by the universities, achieved its own institutionalization by competing against the Workers' Education Association and ended up with thriving. The changes in the three forces were closely related to the changes in the British labor movement at the same time. The change of the natureof labor movement directly affected the development of adult education in the first half of the 20th century and it contributed to the evolution of adult education from working class education to mass adult education.

Keywords: working class education; mass adult education; Labour Movements; Labour Colleges; the Workers' Education Association; university adult education

Non-Continuity and Heterogeneity: Application of Piecewise Growth Mixture Model in Language Development Study

LIU Yuan[1] LIU Hongyun[2]
(1. Faculty of Psychology & Key Laboratory of Cognition and Personality, Southwest University, Chongqing 400715, China; 2. Faculty of Psychology, Beijing Normal University, Beijing 100875, China)

Abstract: In recent researches, the piecewise growth mixture model (PGMM) has been used in longitudinal studies to detect the non-continued growing trend and heterogeneous population simultaneously. The present study used the data from Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K) as an example to illustrate the use of PGMM. An ideal model of PGMN is a two-piece growing model, with the turning point at Grade One, linear trajectory in the first period and quadratic trajectory in the second. The result showed that there should be a crucial turning point in the development of reading ability, with a rapid growing rate from kindergarten to Grade One and then a sharp-decline rate after entering the primary school. Furthermore, a three-class model was selected where the heterogeneous sample-based population was essential in describing the growing pattern. Finally, the result indicated the teachers' assessment of children's behavior was more likely to predict the latent class than that of the parents' with the control of the background effects.

Keywords: Piecewise Growth Mixture Modeling (PGMM); non-continuity; latent heterogeneity; model fit