课程研究前沿　总主编　崔允漷

Evidence-based Research on Curriculum and Instruction

柯政　周文叶　编

基于证据的课程与教学研究

华东师范大学出版社

课程研究前沿

总主编　崔允漷

基于证据的课程与教学研究

柯　政　周文叶　编

华东师范大学出版社

图书在版编目(CIP)数据

基于证据的课程与教学研究/柯政,周文叶编. —上海:
华东师范大学出版社,2017
ISBN 978-7-5675-6992-8

Ⅰ.①基… Ⅱ.①柯…②周… Ⅲ.①课堂教学—教
学研究—文集 Ⅳ.①G424.21-53

中国版本图书馆 CIP 数据核字(2017)第 249309 号

基于证据的课程与教学研究

编　　者　柯　政　周文叶
策划编辑　王冰如
审读编辑　翁晓玲
责任校对　王冰如
装帧设计　刘怡霖

出版发行　华东师范大学出版社
社　　址　上海市中山北路 3663 号　邮编 200062
网　　址　www.ecnupress.com.cn
电　　话　021-60821666　行政传真 021-62572105
客服电话　021-62865537　门市(邮购)电话 021-62869887
地　　址　上海市中山北路 3663 号华东师范大学校内先锋路口
网　　店　http://hdsdcbs.tmall.com

印 刷 者　南通印刷总厂有限公司
开　　本　787×1092　16 开
印　　张　19
字　　数　381 千字
版　　次　2017 年 10 月第 1 版
印　　次　2017 年 10 月第 1 次
书　　号　ISBN 978-7-5675-6992-8/G·10666
定　　价　48.00 元

出 版 人　王　焰

前言

根据马克思恩格斯的历史唯物主义，评价一个时代的生产力水平，最好的指标是生产工具。生产工具从某种意义上说，就是这个社会创造财富的方式方法。似乎是基于同样的逻辑，皮亚杰、科尔伯格等人也认为，衡量一个个体的（道德）认知水平的高低，最好的指标是他思考问题的方式，而不是其结论对错。比如在著名的两难问题——海因兹是否应该去偷药中，从儿童回答是否应该这个结论中，我们不能看出他的道德认知发展水平，但从他得出这个结论的思维方式中，我们就可以看出他的道德认知发展水平。方式方法在很多时候更能说明问题。

那一个时代或国家的课程与教学研究水平怎么来体现呢？当我们在说现在我国课程研究水平比改革开放之前高（或者相反），或者我国的课程研究水平比某个国家的低（或者相反），背后事实上有一个评判中国课程研究水平高低或者发展阶段的标准。很多人可能感觉，这个标准好像模模糊糊地存在着，但要具体说起来，颇为困难。在我们看来，与前面所说的用生产工具来评判时代的发展水平，用思维方式来评判学生的认知水平同一个逻辑，衡量一个时代或国家的课程与教学研究水平，主要还是要看它主导的研究方式方法是什么。研究方法比研究观点更能表达一个时代或者一个国家的教育或课程研究水平。

作为一个课程研究者，我们的学术使命当然是希望能为人类提供更多更加可靠的课程知识。但从一个群体的层面来看，若没有方法的更新升级，知识是很难更新升级的。一般来说，人类获得知识的方法大致可能有这么几种。最多的可能就是来自权威，即有人告诉你这个知识，然后你获得了这个知识。一种是来源于生活经验，我的生活经验比你丰富，见到过更多的事物，我就比你有更大的机会获得更好的知识。一种是依靠直觉、灵感，这很难捕捉，但很多人经常依赖这种方式得出结论，在很多时候，这些知识也都是有用的。一种是推演，即从一个已知或者公认的知识或理论中推演出来，比如我知道杜威的理论，那么我就有更多的关于课程应该怎么开发的知识，因为很多东西我可以从杜威的理论中推演出来。也有可能是依据一些现代社会很难认同的非理性方法，比如通过占卜、梦示。还有一种就是大家所说的科学的方法，即主要通过实验、观察、调查等实证的方法。当然，以上说的这些方法大多也是概念上的区分，在实践中经常是交叉在一起的。比如，可能某个学者就是因为某种灵感和直觉获得了一个假设，然后通过实证的方法确认了这个知识。

可能每种方法都有其价值和适用性。但人类经过几百年的持续比较，科学已经充分地证明了这点，那就是虽然不同的方法都有诸多类似或者关联的地方，但不同的方法的确存在质

的区别。虽然科学的方法也经常犯错，但科学方法具有更强的纠错能力；虽然通过其他方法获得的知识也经常被证明是对的，但通过科学方法获得的知识犯错误的概率更低。若从一时一人的角度来看，这些区别可能都不是很明显，但从长时段大范围的角度来看，科学研究的这些优势足以制造出巨大的差异。

而对科学研究方法来说，它的最基本原则或者唯一信奉的原则就是证据原则（principle of evidence），即评判一个观点和结论对不对，唯一的标准就是它是否获得了足够有力的证据支持。哪怕一个观点最终证明是错的，只要它当时比其他竞争性观点有更好更多的证据支持，我们就会相信它。而且也只有当竞争性观点获得更多更有力的证据支持时，我们才会知道它是错的。这就是所有科学研究的逻辑。这个逻辑在物理学、生物学、经济学、教育学等所有科学学科中都是一样的。

受多种因素影响，我国课程与教学研究中基于证据的研究还是偏少，甚至可以说是严重不足。这已经明显地制约了我们的课程与教学研究水平的进步。课程与教学问题是一项极为复杂的问题，不同的人基于不同的立场得出不同的结论，是极为普遍的。如果大家不是基于证据来交流，就必然会出现，谁的权力大谁就对的情况。在基于证据的科学原则没有确立之前，几乎所有学科都是这样的。

基于此考虑，华东师范大学课程与教学研究所在几年前就决定把第十四届上海国际课程论坛的主题定位"基于证据的课程与教学研究"，希望借此办法和平台，推动国内课程与教学研究领域逐渐形成基于证据的研究文化，激励更多的学者从事基于证据的研究。我们深信，这是中国课程与教学研究的方向。

这次会议得到了国内外同行的积极响应，诸多作者都向大会提交了会议论文。我们挑选了其中部分论文，编辑出版了本书。本书是根据几大研究主题来进行组编的，其中有些论文，可能也不算是好的实证研究，甚至不算是基于证据的研究，但这可能就是当前阶段我国课程研究的真实样态，我们讨论后还是把它收录其中。这是我们的"第一张小板凳"，丑点没关系，关键是要有进步。

本书能顺利出版，毫无疑问，首先归功于各篇的作者。另外，我们作为主编虽然付出了一些劳动，但更多更麻烦的支持性工作是由崔允漷所长和付黎黎主任完成的。华东师范大学出版社一如既往的专业支持，更是不可或缺。赵玉翠同学做了很多具体的行政支持工作。在此一并致谢。

也让我们共同期待，越来越多的基于证据的课程与教学研究。

柯政　周文叶

目录

基于大数据的课程与教学研究

Multilevel Modeling of Educational Data
…… Todd Milford 003
国际大型教育研究的过去、现在与未来：台湾经验
…… 陈柏熹 011
什么样的教学才能形成良好的课堂学习品质
…… 崔允漷 董泽华 雷 浩 031
新课改理念下的学生学习结果类型及其生存环境
…… 陈依婷 杨向东 042
课堂评价对作业负担的影响有多深：基于 16 141 份数据
…… 郑东辉 叶盛楠 057

课程变革与学校课程领导力

Revolution en Cours du Paradigme Curriculare Français
…… Roger-François Gauthier 071
面向社会与性别民主化的法国课程改革与冲突
…… [法]尼古拉·莫斯蔻妮 张 丹 076
论学生的课程权力
…… 夏永庚 085
“新劳动教育”：为孩子的幸福人生奠基——“新劳动教育”课程探索
…… 章振乐 095
我国教科书研究的新世纪图景——基于 CiteSpace 知识图谱的分析
…… 张铭凯 靳玉乐 109
我国课程领导研究热点和演进——基于 2002 年以来 195 篇 CSSCI 期刊论文知识图谱分析
…… 袁 娇 孙芙蓉 王庆超 123

三世说与课程变革
…………………………………………………………………… 刘良华 131
指向核心素养的课程改革前瞻与思考——香港高中通识教育科的启示
…………………………………………………………………… 肖 驰 139

促进核心素养发展的学与教

How Best to Know What Our Students Know
…………………………………………………………… Alison L. Bailey 149
基于问题解决的数学合作学习教学模式建构
…………………………………………………………………… 裴昌根 171
构建“学习共同体” 促成“境脉学习”——苏州市平江中学校“课程六对话”经验
……………………………………………………………… 徐燕萍 邓大一 181
布卢姆认知目标分类学(修订版)的教学应用:方式与问题
…………………………………………………………………… 王小明 215
促进学生核心素养发展的课堂教学形态——源自国际文凭课程的经验与启示
…………………………………………………………………… 赵士果 224

课改背景下教师的课程理解与专业发展

多元文化背景下的教师文化身份认同——基于民族地区“外来”教师的案例考察
…………………………………………………………………… 王艳玲 239
社会性别理论视角下女性教师专业发展的内涵、困境及其解决路径
…………………………………………………………………… 王素月 261
教育硕士专业学位培养质量评估:毕业生跟踪调查的视角
…………………………………………………………………… 汪贤泽 267
“辩证意识”与现实的“课程理解”
……………………………………………………………… 余闻婧 吴刚平 276
集团教研:一种新的教研型态的实践探索
……………………………………………………………… 范建成 王英豪 288

基于大数据的课程与教学研究

Multilevel Modeling of Educational Data

Todd MILFORD

University of Victoria, Canada

Abstract: The role and impact on education of International Large Scale Assessments (ILSAs) such as the Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS) has grown considerably in the past few decades. ILSAs of student achievement offer a glimpse into broadly defined concepts of literacy and provide information on student and school level factors that have been shown to be associated with learning. This paper provides an overview of one such ISLA (i. e. , PISA) as well as some examples of the research that can be accomplished with these data. Through the use of Hierarchical Linear Models (HLM), the results from PISA will be used to model relational patterns of student, home and school characteristics for Canadian students across these broadly defined concept of science literacy. It is anticipated that other researchers will find much from these models is transferable to China.

The most recent iteration of the Program for International Student Assessment (PISA) was completed in 2015 and represented the second full cycle (i. e. , 6th iteration) of the assessment. PISA is a tri-annual International Large Scale Assessment (ILSA) that was first conducted in 2000 and is developed by the Organization of Economic Cooperation and Development (OECD) to provide participating countries with internationally comparative mean literacy scores in reading, mathematics and science for students nearing the end of compulsory schooling. It is anticipated that results from the most recent iteration of PISA will be released on the 6th of December, 2016, at 11:00 Central European Time. This will be followed by a series of media reports associated with different nations reactions to the school and country rankings — the so called league tables. These rankings are typically reported in terms of mean performance and often are associated with media sensation and political statements. Seldom are they ever used to inform public policy or instructional decisions (Shelley, 2009). What makes PISA more useful in providing information that might be used to inform public policy or to guide instructional decisions is the additional information on student, home and school characteristics collected and made available by the OECD. This information allows for the modeling of school effects as they point to the relationship between school characteristics and schooling outcomes (O'Connell & McCoach, 2008). In this way, PISA can be used to go beyond the simple but appealing ranking of nations (i. e., the league tables) to viewing results as outcomes of schooling that can be modeled from school characteristics.

The history of school effects research can be traced back to the initial studies by Coleman et al. (1966) and Jencks et al. (1972) who argued that once student background characteristics (e.g., SES and prior achievement) were accounted for, little variance in student achievement remained (Creemers, 2006). These authors were suggesting that manipulated variables at the school level such as per-student funding, teacher ratios and/or the nature of the curriculum had little impact on student outcomes when compared to the more fixed and less pliable variables such as student background. The conclusion taken from this research was that schools did not make a difference in student outcomes. As a direct response to these findings, school effectiveness research (SER) emerged over the subsequent decades. Advances in both the statistical modeling and analysis of school effects combined with the additional variables related to the student and school have led to the general recognition that schools do have some additional impact on educational outcomes.

This paper provides an overview of how PISA might be used to explore school effects research using the most fundamental multilevel procedures. In general, school-effects research explores the differences in student educational outcomes among students within schools and across schools by indicating the relationships between these characteristics. O'Connell and McCoach (2008) classify these characteristics into context and climate variables. Context variables are those associated with the physical structure and 'hardware' of the school; physical aspects (school location, resources), student body (socio-economic status or proportion of immigrant students) and teachers (experience, level of qualifications). Climate variables are those associated more with the 'software' of the school; learning environment (organizational structure, attitudes values and expectations of students, teachers and parents). The later of these variable groups (i. e., school climate) is often the focus of school-effects research as it is under the control of students, teachers and parents. For example, student self-efficacy beliefs have been associated with positive academic outcomes and school wide efforts can be made to target and encourage these beliefs. On the other hand, context variables are typically not under the direct control of these groups but can be used as a control for adjusting climate variables.

With this logic in mind, school-effects using multilevel modeling can be demonstrated here with selected results for Canada from PISA 2006 (the last year science was the major domain) using Hierarchical Linear Modeling (HLM) software (Raudenbush & Bryk, 2002). PISA operates on a 3-year cycle and all three of the achievement domains are assessed with each cycle; however, there is one domain of focus each iteration. For example in 2000, the first year of PISA, reading was the domain of focus with mathematics and science being minor. In 2003, the domain focus was mathematics and in 2006 it was science. Since the most recent iteration in 2015, PISA has complete this cycle of domain focus twice. HLM software is an extension of the standard statistical analysis of regression based analysis that explicitly incorporates into the analysis the hierarchical structure that is common to educational data sets such as PISA. In PISA, students are grouped or nested within schools within countries. The data required for these analysis consists of both outcome (achievement or performance) and predictors from the students (level 1) and the school (level 2). With these grouping effects, students are no longer independent but their responses are correlated; resulting in a loss of independence among observations — a violation of a key assumptions for the analysis (Snijders & Bosker, 1999).

At level-1, HLM allows us to describe the linear relationship of literacy achievement (any of the three domains in PISA) and student level characteristics such as gender, socioeconomic status (SES), immigrant status, attitudes towards school, motivations *et cetera*. This level-1 model can be represented as a familiar linear regression model, in this case, modeling science achievement (Science for student i in school j) with student gender:

$$\text{Level 1: } Science_{ij} = \beta_{0j} + \beta_{1j}(Gender)_{ij} + r_{ij}$$

Here, each student's science score is modeled on the intercept (β_{0j} — roughly the mean score for science for each of the j schools) plus the weight (β_{ij}) associated with gender and individual error. Unlike multiple regression, the weights are subscripted by j, signifying that a weight (β_{ij} associated with gender) is calculated for each of the j-schools in the data set. For example, if the weight for gender is 1.5 for a particular school and males and females are coded as 0 and 1 respectively, then on average females score 1.5 points higher than males in that school.

HLM explicitly models variance in the level-1 predictor (in this case gender) across schools and evaluates if it is different from 0. This can be done for every coefficient at level-1 and is modeled in a second set of regression equations which are termed the level-2 models. For example, looking at the intercept (β_{0j} — the conditioned mean score for science) the school variance on the intercept is modeled (error_{0j}) as well as school level traits such as school size can be included in the equation:

$$\text{Level 2}: \beta_{0j} = \gamma_{00} + \gamma_{01}(\mathit{SchoolSize}) + u_{0j}$$

Here the intercept is modeled with a level-2 intercept (γ_{00} — constant for all schools in the data set) plus a weighted (γ_{00}) measure of school size and a school-level error term. This models the average school science score as a function of the overall school science score and school size. HLM then tests the significance of the residual error to evaluate variation in school mean science scores (the intercepts- $\beta_{0j}s$) one science achievement has been conditioned on school size. If the error variance is significant, it can be interpreted to mean that there is still significant variation in average school scores after conditioning on school size. A non-significant error variance term suggests that once school size is accounted for, school means scores do not significantly vary from one school to the next.

Similarly, the gender slope in the level-1 model can be modeled with level-2 (school level) variables. This is a unique feature to multilevel models. For example, student level gender may be significantly and positively related to science achievement; however, there may be significant difference between schools for this relationship (the β_{1j}). HLM will explicitly estimate and evaluate these relationships for each school and allow researchers to model school slope variation with school traits. For example, if the gender slope (β_{1j}) varies significantly across schools, they can be modeled with school traits such as those done with the intercept:

$$\text{Level 2}: \beta_{1j} = \gamma_{10} + \gamma_{11}(\mathit{TeacherMorale}) + u_{1j}$$

If gender is related to science achievement, and teacher morale has a negative relationship (γ_{11}) to this slope then it suggests that schools with higher levels of teacher moral would dampen the relationship in terms of gender. This finding would suggest that teacher morale moderates the relationship of gender to science achievement; schools with higher teacher moral do not see as big a difference in science achievement grades between genders. A policy implication could be that if teacher moral is improved, student gender is not as strongly related to achievement in science as in schools with lower teacher moral. Additionally, as school-level variance is explicitly modeled we can determine is any significant variation remains to be potentially modeled by additional school level variables.

A final important outcome of HLM analysis is the Intraclass Correlation Coefficient (ICC) generated by running the null or unconditional model. This statistic is an index of the proportion of variance in the outcome measure that can be accounted for by level-2 units.

Some Results from PISA 2006 and 2009

The objective of research on school effects is to uncover how the context and climate variables influences schooling over and above the influence of students personal and background variables. In this section, science achievement will be modelled as an outcome measure and the effects of school climate and context through a variety of models.

Null model. The null model contains only an outcome variable and no predicator variables other than the intercept. The null model can be represented as

$$Y_{ij} = \beta_{0j} + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{0j}$$

The null model estimates the grand mean of science achievement with adjustment for clustering in school sand different sample sizes across schools and estimates the variance component at the student and school level. The proportion of variance in the outcomes that is between schools is the intracalss correlation and can be calculated as the between school variance (τ_{00}) divided by the total variance ($\sigma^2 + \tau_{00}$). This proportion ranges from 0 to 1 with higher values indicating greater between school variance.

Table 1 ***Statistical Results of the Null Model of School Effects on Science Achievement***

	Fixed effects			
	Coefficient	SE	t-ratio	p
Intercept (science achievement) γ_{00}	524.08	1.90	275.59	.001
	Random effects			
	Variance	df	Chi-square	p
Between school variability (intercept)	2 202.69	811	5 346.24	.001
Within school variability	7 340.90			
Reliability (intercept)	0.85			
Intraclass Correlation	0.23			

Student model. Adding student level variables in models is to control for characteristics of students attending schools. The student model contains no school level variables and can be expressed as

$$Science_{ij} = \beta_{0j} + \beta_{1j}(SES)_{ij} + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Table 2 ***Statistical Results of the Student Model of School Effects on Science Achievement***

	Fixed effects			
	Coefficient	SE	t-ratio	p
Intercept (science achievement) γ_{00}	517.33	1.69	305.27	.001
Student level variables				
ESCS γ_{10}	26.58	1.19	22.35	.001
	Random effects			
	Variance	df	Chi-square	p
Between school variability (intercept)	1 497.88	805	3 306.45	.001
Within school variability	7 050.03			
	Proportion of Variance Explained			
At the school level (between schools)	0.32			
At the student level (within schools)	0.04			

Contextual model. The contextual model is defined as a regression model with individual level variables and aggregated contextual variables. The contextual effects of socioeconomic status are often of interest to researchers and students nested in schools are effected by both the individual level SES as well as the mean SES of the school they are attending.

$$Science_{ij} = \beta_{0j} + \beta_{1j}(SES)_{ij} + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(SchoolMeanSES) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Table 3 ***Statistical Results of the Contextual Model of School Effects on Science Achievement***

	Fixed effects			
	Coefficient	SE	t-ratio	p
Intercept (science achievement) γ_{00}	504.81	3.57	141.19	.001
Student level variables				
ESCS γ_{10}	22.71	1.85	2.25	.001
School level variables on intercept				

续 表

School mean ESCS γ_{01}	45.12	9.04	4.98	.001
	Random effects			
	Variance	df	Chi-square	p
Between school variability (intercept)	1 101.56	804	2 914.58	.001
Within school variability	7 048.11			
	Proportion of Variance Explained			
At the school level (between schools)	.50			
At the student level (within schools)	0.04			

Evaluative model. The evaluative model builds from the contextual model by adding variables of school climate to explain the variance in school mean science achievement after controlling for school contextual effects. The purpose of this model is to evaluate the impacts of school climatic variables on schooling outcomes.

$$Science_{ij} = \beta_{0j} + \beta_{1j}(SES)_{ij} + \beta_{2j}(SCIEFF) + \beta_{3j}(SCSCIE) + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(SchoolMeanSES) + \gamma_{02}(PCGIRLS) + \gamma_{03}(SCHSIZE) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Table 4 ***Statistical Results of the Evaluative Model of School Effects on Science Achievement***

	Fixed effects			
	Coefficient	SE	t-ratio	p
Intercept (science achievement) γ_{00}	509.05	9.91	51.36	.001
Student level variables				
ESCS γ_{10}	10.94	1.06	10.23	.001
SCIEEFF γ_{20}	26.54	1.24	21.27	.001
SCICIE γ_{30}	20.29	1.04	19.45	.001
School level variables on intercept				
School mean ESCS γ_{01}	38.70	3.60	10.75	.001
Proportion GIRLS γ_{02}	−30.82	19.03	−1.32	.109
School SIZE γ_{03}	0.010	0.003	3.19	.002
Between school variability (intercept)	1 024.89	780	2 782.99	.001
Within school variability	5 291.99			

		Proportion of Variance Explained
At the school level (between schools)	.53	
At the student level (within schools)	.29	

From the brief model building exercise detailed above, the potential for some useful and meaningful policy relevant research should become clear. For example, SES has been shown to be consistently positive and significant; however, models such as these can uncover the variance in the relationship between one country and another. A further exploration of those schools with a more equitable (i. e., flattest SES gradient) to see if there are key school traits that might be associated is worthy of exploration. That data may well lie within data sets such as the one detailed here.

References

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Winfield, F., & York, R. Equality of educational opportunity [R]. Washington: US Government Printing Office, 1966.

Creemers, B. M. P. The importance and perspectives of international studies in educational effectiveness [J]. Educational Research and Evaluation, 2006,12(6): 41 - 59.

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. Inequality: A reassessment of the effects of family and schooling in America [M]. New York, N. Y.: Basic Books, 1972.

O'Connell, A. A. & McCoach, D. B. Multilevel modeling of educational data [M]. Charlotte, N. C.: Information Age, 2008.

Shelley, M. C., II. Speaking truth to power with powerful results: Impacting publicawareness and public policy [C]//M. C. Shelley II, L. D. Yore, & B. Hand (Eds.). Quality research in literacy and science education: International perspectives and gold standards. Dordrecht, the Netherlands: Springer, 2009: 443 - 446.

Snijders, T. & Bosker, R. Multilevel analysis: An introduction to basic and advanced multilevel modeling [M]. London, U. K.: Sage, 1999.