《編者的話》
使大型教育資料庫發揮最大效益

緣起

　　多年來，大型教育資料庫逐漸受到教育研究人員與決策者的重視，主要是因為大型資料庫能提供充沛的資料，讓其他研究人員從資料中設想議題進行次級分析（secondary analysis）。這些分析所帶來的成果往往出乎意料的豐富，大大降低研究成本。以美國教育部全國教育統計中心（National Center for Education Statistics (NCES), U.S. Department of Education)在 1980 年代所執行的**高中及高中後追蹤調查方案**（High School and Beyond）為例，其建置的資料庫廣為大眾使用，產出的研究成果非常豐碩，包括專題研究報告、政策分析、期刊論文、博士論文、碩士論文、研究技術報告等（詳情請上 http://www.nces.edu.gov 查閱）。另外，此方案為追蹤調查，資料有前後時序之分，先頭的背景與其他資料可設為因，後頭的資料可設為果，讓研究者模擬在自然情境中的教育實驗，其結果非常珍貴。

　　雖然大型資料庫不全然都是為了追蹤研究而建置的，很多是針對某一群對象做一次性資料收集（cross-sectional studies），但是這些資料對教育評鑑與研究的價值也是眾所周知。比如 IEA 贊助的**國際數理趨勢評鑑**（**Trends in International Mathematics and Science Study** — TIMSS）和**國際閱讀能力評鑑 Progress in International Reading Literacy Study**（PIRLS）以及美國 NCES 推動的**全國性教育進展評量 National Assessment of Educational Progress**（NAEP）和**全國性教師與校長調查 National Teacher and Principal Survey**（NTPS）都是很好的例子。這些研究方案每隔幾年會重覆做一次，因此可與先前的資料做比較，進一步檢視其變化，提供相關教育人員及決策者寶貴的資訊。

　　不過，大型資料庫的建置不是一件易事，需要花費很大的工夫，而且所費不貲。因此，任何大型資料的收集必須要有周詳的規劃、設計與執行，以確保達成其最高價值與品質。為達成此目的，得借鑒前人的經驗與歷程。

本特刊之目的

因此，本特刊的目的在於彙集前人在大型教育資料庫上的設計、執行、與應用的集體智慧。更具體來說，我們想要瞭解樣本及資料收集工具的設計過程與需要考慮的面向與因素，也想要瞭解分析這類資料需要的正確分析方法、新的分析軟體與相關資料運用經驗，以確保分析結果的品質。最後，也希望能更進一步瞭解資料庫該如何建置與傳播，以增進資料的運用。希望此特刊彙集的資訊能讓許多研究人員在運用既有資料和開創新的大型教育資料庫方面能增進興趣與信心。

特刊文章簡介

非常感謝投稿本特刊的作者，雖然有些稿件因為主題或性質不適，未能採用，但有這麼多人投稿此一相對冷門的主題，證明德不孤，必有鄰。相信未來將會有越來越多教育研究人員與決策者支持並促進教育大型資料庫的建置與應用。

本特刊一共刊登五篇文章，第一篇文章標題為 **Establishing a Comprehensive Large-Scale Data Infrastructure for Educational Research: The Example of the German National Educational Panel Study（NEPA)**。由德國 Leibniz Institute for Educational Trajectories（LIfBi）的 Jutta von Maurice、Daniel Fuß 與 Hans-Günther Roßbach 共同撰寫。主要在說明建置與傳播 NEPA（德國全國性教育長期追蹤調查）資料的歷程。NEPA 是德國政府近年全力支持建立的大型教育長期追蹤資料庫。首先介紹 NEPA 的學生抽樣設計以及針對學生能力的發展與教育過程的議題選擇的考量，兩者都考慮到相關學習環境與社會平等、正義因素。接下來說明如何建置資料檔，包括資料錯誤檢驗與更正、資料編碼（coding）與變項建構以及資料檔案說明書撰寫與資料庫管控等步驟。同時也介紹資料安全保密與傳播措施以及運用方式和未來發展願景。此文對長期追蹤研究方案資料的規劃、設計以及執行程序與策略提供良好的指引與規範。

第二篇文章標題為 **Data Analyses With IEA's TIMSS and PIRLS International Databases**。作者為 TIMSS and PIRLS International Study

Center 的 Pierre Foy 與 Liqun Yin。一般大規模教育評鑑通常都需要周詳與精密的評量工具、複雜的抽樣設計以及先進的測驗理論（item response theory）的應用，以達成評鑑分析目標。這些設計上的要求增加了分析上的困難，需要使用適當的方法與程序，才能獲得正確的結果。本文以 TIMSS 和 PIRLS 為例，闡釋三項必須使用的分析方法：（1）做統計分析時必須使用樣本加權指數（sampling weights）讓樣本回歸原母體群的分佈，以獲得正確可靠的結果。（2）使用特殊方法（the Jackknife Repeated Replication technique）去推算抽樣誤差值（sampling variance）（詳情請見內文）。（3）正確運用個人 5 項成績推估值（plausible values）來推算學生表現。內文有範例做詳細的示範與說明。

第三篇文章標題為 **U.S. National Teacher and Principal Survey（NTPS）as a Valued Resource for Teacher and Principal Studies**。由美國內布拉斯加大學林肯分校的 Jiangang Xia、西密西根大學的 Xingyuan Gao 和 Jianping Shen 共同撰寫。作者先提示說明 NTPS 將取代原先的 Schools and Staffing Survey（SASS），兩者都是由美國教育部全國教育統計中心執行的方案。因為資料性質大致上相同，所以作者以 SASS 的文獻來分析 NTPS 資料的分析潛力。結論是 NTPS 的資料是非常獨特、豐富的教育研究資源，可用來探討許多教師與校長的議題。目前在許多國家，這些議題常被忽略。因此，作者建議應建置類似 NTPS 的資料庫，並鼓勵多做有關學校教師與校長的研究。

四篇文章標題為**Challenges and Opportunities for Estimating Effects With Large-Scale Education Data Sets**。作者為美國密西根州立大學的 Guan Kung Saw 和 Barbara Schneider。由於大型資料庫有可能用來模擬與實驗相近的情況，而不需運用傳統隨機分派參與實驗者到實驗組與控制組的方法。這種能量使大型資料庫的分析能用來推估教育方案與措施的效果。其結論對於教育決策人員如何做出正確的抉擇有很大的幫助，非常有價值。不過，也有一些挑戰需要克服。本文特別檢視這類效益分析的條件與可能碰到的統計問題。相信本文所呈現的資訊對於欲嘗試運用大型資料進行因果關係分析的研究員應該有啟示與參照之用。

第五篇文章題目是 **Using R to Analyze International Large-Scale Educational Assessment Data**。作者為成功大學教育研究所的紀馥安及許

清芳。目前處理大型教育資料庫常用的工具大多為商用的統計軟體。其資料管理與分析的流程不易與其他研究者分享，結果也較難被覆驗，以便累積研究成效的信度來幫助指引教育政策的改進方向。本文提供解決此問題的方法：以開放軟體 R 來管理與分析大型教育資料庫的檔案。本文以包括台灣在內一共十五個國家中十五歲學生的數學成就表現和多項影響該表現的相關因素為案例，示範如何運用 R 軟體去管理與分析資料。除了使用 R 的基本功能之外，還特別嵌入針對處理大型教育資料庫發展的 intsvy 套件並結合 maptools 等相關套件，將分析結果呈現在地圖上以利比較。另外，文中提及的檔案管理及資料分析過程，也可使用附錄的 R 程式碼來複製。

結語

　　總而言之，上述論文皆直接或間接地說明大型資料庫是珍貴的研究資源，能提供決策者更加正確、可靠的資訊作為決策的依據。不過，大型資料庫需採用複雜的抽樣設計與資料收集方法，以確保有足夠的樣本數供分析、檢測所有重要的研究對象，也有充分的資料來探討重要議題。另外，這些結構複雜的資料也需要運用一般不常見的統計方法來處理與分析。這些條件與要求往往帶給資料建置與使用者挑戰。所幸這些年來已有許多相當成功的例子。我們可以從這些案例中學習知識與經驗。加上一些分析與測驗方法的研發，應可加速精進大型資料庫的建置與運用。因此，期待在不久的將來，大型教育資料庫的建置與運用將成為一般慣例，大大地提升教育評鑑與研究的品質與成果。

<div style="text-align: right">

彭森明
特約主編

邱皓政、溫福星
執行編輯

2015 年 10 月

</div>

《Editors＇ Notes》
MAXIMIZING THE VALUE OF LARGE-SCALE DATABASES IN EDUCATION

Background Statement

Large-scale databases in education have been increasingly drawing attentions from researchers and policymakers, primarily because of their analytic power to produce in-depth studies and cost-saving benefits in the long run. To be more specific, a well-designed large-scale database will not only be able to effectively address the originally intended study issues but also could become an important resource for numerous secondary analyses because of its elaborate sample design and comprehensive data contents. Such capability and productivity consequently will lower the unit cost for each study. Take **High School and Beyond**, a longitudinal database of the 1980 high school sophomore and senior students sponsored by the National Center for Education Statistics (NCES), U.S. Department of Education, for example. It has been widely used by researchers and policymakers, resulting in a large number of publications, including journal articles, research reports, policy analyses, dissertation/theses of graduate students, and technical reports (http://www.nces.edu.gov). Furthermore, because of its longitudinal nature, an individual's prior background, attributes and performance can be used as controlled variables in estimating the impact of educational programs or practices. Such analyses are the so-called controlled studies in a natural setting. They are greatly valued in educational research.

Although not all large-scale databases are longitudinal in nature, their values for comprehensive and efficient evaluation and research studies, nevertheless, are widely recognized. Several well-known studies such as IEA sponsored **Trends in International Mathematics and Science Study** (TIMSS) and **Progress in International Reading Literacy Study** (PIRLS) as well as NCES administered **National Assessment of Educational Progress** (NAEP) and **National Teacher and Principal Survey** (NTPS) are good examples of this nature. Those studies, when repeated every few years, also enable trend analyses to further examine and monitor any changes over time, providing valuable information for program or policy decision-making.

However, the development of a large-scale database is not a simple task. It takes a lot of efforts and it is not cheap. Thus, any large-scale data collection must be carefully planned, designed and implemented to assure that its values can be maximized. To this end, experience and lessons from prior studies would be extremely helpful for all future studies.

Objectives of this Special Issue

This special issue of Contemporary Educational Research Quarterly (CERQ) is, therefore, intended to compile collective wisdom about the design, implementation and applications of large-scale databases in education. Specifically, we would like to know what steps and factors should be considered in designing a study sample and data collection instruments. We would also like to know what statistical analysis techniques would be required in mining the database, and how the database should be constructed and disseminated to facilitate its use. We hope that the information assembled in this special issue would further inspire and encourage the development and use of large-scale databases in education.

Overview of the Articles

We have accepted five articles for this special issue.

The first article, **Establishing a Comprehensive Large-Scale Data Infrastructure for Educational Research: The Example of the German National Educational Panel Study (NEPA)**, written by Jutta von Maurice, Daniel Fuß and Hans-Günther Roßbach from Leibniz Institute for Educational Trajectories (LIfBi), presents some key processes in preparing and disseminating a rich empirical database to researchers from different disciplines. It starts with an overview of the NEPA's design of sample and selection of research topics on competence development and educational processes, both taking into account the relevant learning environments as well as issues of social inequality, the special situation of migrants and the various returns to education. The article then describes the processes of creating the database, including data cleaning and editing, coding and variable generation, documentation and metadata management, as well as data enrichment. The data protection and dissemination strategies are also explained. Furthermore, the article also gives some basic information about data usage as well as an outlook on future developments within the NEPS. This article clearly illustrates the steps and considerations required for planning, designing, implementing the development of a resourceful longitudinal database and its final applications for research and evaluation in education.

The second article, **Data Analyses With IEA's TIMSS and PIRLS International Databases**, by Pierre Foy and Liqun Yin from TIMSS and PIRLS International Study Center, states that large-scale assessments in education generally rely on sophisticated assessment instruments, elaborate sample designs, and leading-edge item response theory to meet their analytical objectives. To analyze such databases, proper analytic techniques and processes would be required. The authors used TIMSS and PIRLS data to illustrate the required analysis procedures, including the use of sampling weights to produce accurate and reliable results, the application of the Jackknife Repeated Replication technique to derive proper estimates of sampling variance, and the correct handling of student achievement reported as sets of five plausible values to estimate the students' performance. Following those procedures, researchers and users of the TIMSS and PIRLS databases should feel confident in the results of their analyses.

In the third article, **U.S. National Teacher and Principal Survey (NTPS) as a Valued Resource for Teacher and Principal Studies**, by Jiangang Xia from University of Nebraska-Lincoln and Xingyuan Gao and Jianping Shen from Western Michigan University, these authors first give a brief review of NTPS which will replace the traditional Schools and Staffing Survey (SASS), both administered by the National Center for Education Statistics, U.S. Department of Education. Since the data collected by NTPS will be similar to those collected by SASS, the authors outline the potential studies using this database by reviewing and discussing how the SASS data have been utilized for educational research and policy analyses over the years. It is concluded that the NTPS database will be very unique and resourceful for studies on school teachers and principals, an area that has often been neglected in many countries. The authors strongly recommend that such a database be developed and more studies about teachers and principals be conducted.

The fourth article, **Challenges and Opportunities for Estimating Effects With Large-Scale Education Data Sets**, by Guan Kung Saw and Barbara Schneider from Michigan State University, describes the capability of large-scale databases to simulate near-experimental conditions without employing traditional methods that require randomization of units (e.g., students, schools, districts) to treatment and control situations. Such a capability for making robust inferences regarding the effect of educational programs and/or practices is very important and helpful for decision-making in education. This article examines the opportunities and potential statistical problems with estimating effects with large-scale databases. The information presented in this article should be valuable and helpful to researchers in using large-scale database analyses to make inferences of the effect of educational programs or practices.

Finally, the fifth article, **Using R to Analyze International Large-Scale Educational Assessment Data**, by Fu-An Chi and Ching-Fan Sheu from National Cheng Kung University, offers a solution to researchers who do not have access to the procedures of analysis by commercial software to replicate results others have reported in publications. The authors demonstrate how to use the free, open-source R computing environment to manage and analyze international large-scale educational assessment data. The example is comprised of mathematics achievement and covariates for 15-year-old students from 15 countries, including Taiwan. Also included in the demonstration is the use of the "intsvy" package developed by Caro and Biecek to manage the data and the maptools package to link numerical summaries to geographical boundaries of countries examined in the illustration. R codes to perform analysis of data from PISA detailed in the paper is provided to facilitate reproducibility.

Concluding Remarks

In summary, as directly or indirectly shown in those articles, large-scale databases offer great opportunities to researchers and policy-makers. However, most of such databases require an elaborate sample design and comprehensive data collection methods in order to ensure all important subgroups are covered and all relevant issues can be sufficiently addressed. Also required is special statistical techniques to properly analyze the data to obtain unbiased results. Such requirements often pose challenges to data developers and users. Fortunately, over the years many large-scale databases have been developed. Lessons and experiences gained from those studies and new technological skills and tools developed in recent years have helped the advancement of large-scale data for educational evaluation and research studies. As a well-designed database can facilitate many high-quality research and policy studies, it is hoped that such large comprehensive databases, instead of scattered small and localized databases, will become a common practice in the near future to advance educational evaluation and research studies.

<div align="right">

Samuel S. Peng
Guest Editor

Hawjeng Chiou & Fur-Hsing Wen
Managing Editor

October, 2015

</div>