

大數據時代的基本學力檢測未來趨勢 —以臺北市國小基本學力檢測為例

陳清義

臺北市市政顧問、臺北市國小校長協會理事長、臺北市福星國民小學校長、
臺北市立大學教育行政與評鑑研究所博士

王淑珍

臺北市國小基本學力檢測專案教師、臺北市立大學教育行政與評鑑研究所博士

壹、前言

2012年《紐約時報》的專欄文章「The Age of Big Data」宣告了「大數據時代」的來臨，在科技進步與社會快速變遷的時代下，資料的積累量與速度已遠遠超越過去，呈現大數據（big data）的樣態。大數據，也有譯稱「巨量資料」，它是高容量（volume）、高速度（velocity）、多樣性（variety）的訊息資產，此一定義源自 Gartner 公司的首席資料分析師 Laney 在 2001 年為了協助客戶了解如何識別，及因應大數據的挑戰，所提出的「3-D 數據管理：控制數據量，速度和多樣性」，而後也為大數據提供了一個比較完整的 3V 架構定義（Laney, 2012）。雖然後續也有一些數據科學家、研究人員添加更多的 Vs，如 Patgiri 和 Ahmed（2016）的相關研究指出大數據的定義除了包括 3V 架構之外，亦應包含顧客量（vendee）以及啟用大數據的基礎需求（vase）；IBM（無日期）也提出大數據的 5V 特點，比 Laney（2012）提出的 3V 架構多了真實性（veracity）及價值（value）兩項……等，但大多還是以 Laney 所提出的 3V 架構為主。

大數據打破了以往數據資料的規模限制，為經濟社會各領域帶來全新的價值認知，在教育上也掀起了一股熱潮，如何運用大數據來重塑教育，提升教師的教學品質、提高學生的學習效果，更是關心教育的人士共同的期盼。為了掌握教育品質與教學成效，世界各先進國家與國際組織莫不重視學生的基本學力，因而陸續開發多項評量計畫來測量與探究學生的學習成效。例如：經濟合作暨發展組織（the Organization for Economic Cooperation and Development，簡稱 OECD）自 2000 年開始推動，每三年進行一次的國際學生能力評量計畫

(Programme for International Student Assessment, 簡稱 PISA), 以及國際教育成就評鑑協會 (International Association for Evaluation of Educational Achievement, 簡稱 IEA) 每四年辦理一次的國際數學與科學教育成就趨勢調查 (the Trends in International Mathematics and Science Study, 簡稱 TIMSS) 和每五年辦理一次的促進國際閱讀素養研究 (Progress in Reading Literacy Study, 簡稱 PIRLS) 皆是為瞭解學生的學習狀況與發展趨勢而設, 顯見基本學力檢測的運用在世界教育改革的趨勢中已然佔有一席之地。

以各項國際評量計畫所取得的檢測數據資料而言, 皆可從事大數據的分析與應用。然而, 不論是 PISA、TIMSS 或是 PIRLS 每三到五年舉辦一次的評量結果分析, 目前僅止於整體現象的調查, 並無法直接使每個學生、每個家長從中受惠。有鑑於此, 臺北市為了解國民小學學生的學習趨勢, 臺北市政府教育局遂於 2005 年成立學習評量工作小組, 著手規劃臺北市國民小學學生基本學力檢測工作 (以下簡稱臺北市國小基本學力檢測), 並開始實施國語文基本學力檢測, 2006 年加入數學領域, 2009 年再增加英語學科。而拜科技之賜, 基本學力檢測的結果可以大量、快速且結合問卷從不同面向分析, 同樣是進行測驗數據的分析, 但此等大規模在以往的小學測驗中則是未曾有過。

臺北市國小基本學力檢測訂於每年 10 月施測, 各科檢測結果則是分析每個學生、每個班級、每個學校, 而且每年皆進行測驗, 實施迄今已邁入第 16 年, 檢測成果除了掌握學生的學習成就趨勢之外, 更提供教師教學回饋, 同時也作為教育局制定教育政策的參考。然而, 在大數據引領的時代下, 教育決策機關或學校如何運用大數據來精進教學或制訂政策, 讓學生可以從測驗中獲得實質的回饋與學習成長, 又能如何透過測驗協助教師調整、精進教學方式, 來提升學生學習的成效, 都是本文所欲了解與探討的內容, 期能為未來的教育測驗發展更添新氣象。

貳、大數據時代的激盪與驅動

具有「大量紀錄、快速取得、來源多樣」等特質的大數據正在進入教育體系的所有層面, 顛覆傳統教學模式, 為教育帶來更多的可能性, 美國牛津大學教授 Mayer-Schönberge 和《經濟學人》雜誌編輯 Cukier 在《大數據：教育篇》一書中提出, 大數據所改變的不只是學習方式, 更包括學習內容, 而大數據的

運用之所以能讓學習活動徹底改頭換面，主要因為具備以下提升學習效果的三大特點（林俊宏譯，2014）：

一、即時與雙向受惠的回饋方式

傳統教學所提供的回饋是學習分數，常被用來判斷一個學生在學校的學習表現，且往往是單一方向，從學校、教師提供給家長和學生。而大數據正在改變這樣的回饋模式，透過資訊科技系統的結合，讓學習過程也能資料化，然後導回教學來提升學生的理解程度和表現，雙向受惠。以電子書平臺為例，當學生在閱讀電子書時，系統可以同步紀錄閱讀時間、不同閱讀內容的停留時長、線上測驗的成果、反覆瀏覽與重複測驗的紀錄……等資訊，線上平臺的測驗結果能提供學生立即回饋、反覆學習的機會，也能分析學習的歷程各種紀錄，讓教師了解學生的學習樣態，從中反思如何精進教學來提升學生的學習成效。

二、落實因材施教的個人化模式

傳統教育在同樣的時間、空間（教室）中呈現完全相同的內容，大多是為了常態（鐘型）曲線「中等程度」的學生所設計，而在曲線兩端的學生，反應快的在課堂上可能被迫無聊度日，反應慢的則可能是在後面苦苦追趕不及，如此「一種規格」想要全體適用的模式顯然不符學生的學習需求，如今，在大數據的演算之下，適應性學習軟體已然崛起，這類系統能夠依據學生的答題狀況，決定後續要問的問題，讓學生覺得針對學習困難處加以練習，更加落實因材施教的需求，而結合網路平臺，打破時間、空間限制，讓學習不再只是侷限於教室裡的課堂時間，個人化模式為學生帶來了更多學習的機會與成長的契機。

三、接近真實世界的可能性預測

大數據可以讓我們看見一群人是如何學習，也能看到每一個人各自如何獲取知識，運用在教育上，可以看見諸多「可能性」的預測，例如：某個單元或某個題目類型對於 80% 的學生來說可能是偏難的、某樣教材有 95% 的機率可能提升某些學生的測驗成績……等，這些機率的預測雖然不是確定的結果，但在面對不確定的情況下，也足以讓我們做為參考的依據。而隨著資料蒐集越多，透過大數據的相關性分析找出「正是如此」的實際狀況，讓「可能性預測」越來越準確，藉此改善教學活動。

大數據為學習帶來了三大改變，不僅能結合各種平臺大量蒐集過去無法蒐集或是高成本的資料，還能將學習個人化、針對個別學生的需求來調整，更能運用可能性預測調整出最好的學習內容、學習時機和學習方式（林俊宏譯，2014）。數位時代來臨後，過去無法處理或是所費不貲的大量學習資料都能被記錄與運用，而透過大數據的運算分析，更能回應我們所欲了解的教育相關問題。

以臺北市當前的國小基本學力檢測為例，為了分析全市國小五年級近 2 萬名學生的檢測資料，提供學習回饋，工作小組採用試題反應理論（Item Response Theory，簡稱 IRT），來分析學生考試成績的模型，並就試題進行單題分析、各向度或全題本分析，就學生進行個人分析及團體分析（如表 1）：

表 1 基本學力檢測的分析與解讀面向

學生 \ 試題	單題分析	向度 / 題本分析
個人	學生在單一試題的反應情形。	學生在特定概念的學習反應。
團體 (班級、學校、行政區、全市)	反應系統性的問題。 例如：教師可能忽略某一類試題的教學，或是學生在某一類試題未能有效學習。	反應教學問題、教材問題。 例如：教師的教學無法讓學生學會特定概念，或是教材的選擇或設計與課綱指標的概念未盡相符。

資料來源：作者自行整理。

透過學力檢測的數據分析，我們可以了解學生個人在單一試題的反應情形，知道單一試題對學生而言是困難或簡單，其在特定向度的概念是否學習精熟，而從學生團體（各班、各校、各行政區、全市）的角度來看，可以呈現學生普遍學習的概況，也能從中找出學生的學習瓶頸或教師教學的潛藏問題，以提供教師、學校或教育主管機關作為精進教學或教育政策制定的參考。為了讓讀者對臺北市國小基本學力檢測有一概括性的認識，下段茲就檢測計畫緣起、政策運用、配套措施、資料運用、釋出管理等層面簡要介紹。

參、臺北市國小基本學力檢測的實施

一、檢測緣起

臺北市政府教育局於 2005 年成立「提升學生國語文能力工作組」及「九年一貫課程推動工作組」，下設「國民小學學習評量工作小組」辦理臺北市國民小學基本學力檢測，並逐年增加檢測學科，以了解本市學生學習發展趨勢（陳清義主編，2006）。而根據 2005 年至 2012 年度六年級學生檢測結果，本市國小學

生學習趨勢穩定，為能進一步將檢測結果作為教師進行補救教學之參考，自 2013 年度起考量學生學習內容、學習壓力及補救教學之需求，調整檢測對象為五年級學生，俾使學生於進入國中階段前，由教師衡酌學生需求進行補救教學（陳清義主編，2019）。

二、政策運用

臺北市國民小學基本學力檢測於 2005 年度首次辦理後，逐年賡續辦理並建置學生學習資料庫，追蹤學生學習趨勢，作為政策及教學之參考。歷年來依據檢測結果與問卷調查分析後，所推動教學政策如表 2。

表 2 臺北市參考檢測分析結果所推動之教學政策

年度	推動政策	說明
2006	調整課程規劃	各校實施每週增加 1 節國語課。
2008	教師命題研習	辦理全市教師命題研習，提升教師命題知能。
2009	精進開發計畫	針對弱勢學生比例較高的學校，推動精進開發計畫。
	補救教學計畫	推動補救教學計畫，成立攜手激勵學習潛能工作組。以資源整合的方式，建立更完善的低成就學生課後補救教學系統，讓班級學習低成就學生，依實際需要接受課後補救教學之協助。
	數學診斷測驗	開發數學適性化診斷測驗工具，協助教師瞭解學生平時在各能力指標的學習差異，及時調整教學重點。
2010	教師教學增能	結合臺北市立大學辦理教師教學增能學分班，針對學習瓶頸的指標進行研討。
2011	提倡閱讀教育	以「閱讀精進年：愛在悅讀·精進 100」為題，與遠見雜誌合作舉辦閱讀學習論壇，提倡閱讀教育。透過「小小說書人」、「晨讀 15 分鐘」等主題閱讀活動，推廣多元閱讀。
	加強英語能力	國小中、高年級每週增加 1 節英語課。
2012	各科命題研習	辦理全市國語文、數學、自然、社會、英語五科教師命題研習。
2013	國語診斷測驗	編製國語診斷測驗，作為補救教學前檢測工具。
	注音符號先修	開設小一新生注音符號先修班。
	檢測結果說明	將全市分為 4 區，針對檢測結果運用，以校長、教務主任、學科領域召集人為研習對象，辦理檢測結果分析說明會。
	研究應用分析	學力檢測工作組增設「研究應用小組」，建立學生學習資料庫及問卷之應用分析。
2015	建置檢測網站	建置基本學力檢測網站，落實資訊數位化，定期公告檢測相關訊息，提供各校自行下載成績報表。

年度	推動政策	說明
2016	診斷測驗數位化	數學、國語診斷測驗數位化，並將診斷測驗與本市酷課雲系統結合。
	兩班三組計畫	推動英語兩班三組計畫，重視學生個別差異，讓學習較慢學生獲得充分關注並培養英語基本能力，達到適性與補救教學雙贏的目標。
	輔導團到校協作	推行輔導團到校協作方案，檢視學校運作、教師教學與學生學習情形，並彙整報告與建議。
2017	英語診斷測驗	英語數位診斷測驗開發。
	數學能力重建	學生數學能力重建工作坊。

資料來源：作者自行整理。

三、資料釋出

每年於學力檢測完成後，針對學生的檢測成績表現情形，由本市檢測小組進行統計分析，製作學校版及班級版成績報表，內容包含各向度及逐題答對率等資訊，各校可透過完整的成績報告書瞭解學生學習狀況，作為教師補救教學之用。資料庫由本市基本學力檢測小組負責管理，並作為政策制訂的參考依據。2011年起訂定「臺北市國民小學基本學力檢測資料使用辦法」，提供相關研究單位、大學教授或研究生申請資料，作為教育研究之用。

四、追蹤管理

本市對五年級學生實施之基本學力檢測，係藉以診斷學生學習成效及瓶頸，俾協助教師精進教學之參考；同時作為瞭解本市國小五年級學生整體學習趨勢及擬定教育政策之參考，惟尚無提供個別學生學習歷程表現之功能。

臺北市國小基本學力檢測訂於每年 10 月施測，實施迄今已邁入第 16 年，在過去 15 年的檢測成果報告以及分析結果，提供了學校及教師作為學生補救教學的參考，也對教育政策的訂定有實質的貢獻。然而，對照大數據時代「大量、快速、多樣性」的需求，關於學生個人即時回饋、各項學習歷程的記錄與追蹤目前還未能加以連結，更難以據此進行未來可能性的預測，這些未及的目標尚待未來教育與各界繼續共同努力。

肆、基本學力檢測的未來趨勢

測驗只能反映學生在某個特定時刻的表現，並無法真正反映學生的學習成效，除非我們能增加測驗的頻率，否則無法大幅提升測驗所能帶來的價值（林俊宏譯，2014）。基本學力檢測是一種完成性的測驗，所能反映的即是學生在測驗當下的學習狀態，而在數位引領的時代下，因應大數據時代的即時回饋、個人化模式和可能性預測三大特點，臺北市國小基本學力檢測的未來發展或許可從以下幾個方向出發，運用大數據的概念，結合其他系統平臺及資源，藉以整合並深化檢測數據的意義與價值：

一、建置線上系統，放大測驗平臺

基本學力檢測是單一總結性的測驗，倘能建置線上測驗系統，或是配合既有的學習平臺，如：臺北市酷課雲、PaGamO、臺北市數位診斷測驗系統……等，即可跳脫時間與空間的限制，放大測驗的平臺，增加診斷測驗的次數，藉由各種測驗平臺蒐集大量測驗的資訊，亦即運用大數據的高容量與多樣性的特質，從原本單一測驗所獲得的單一點，逐漸串連成線，最後交織成面，呈現出學生較完整的學習面貌。

二、搭配行動載具，記錄測驗資料

在講求速度及高容量的大數據時代中，為了快速且大量記錄學生的測驗資料，行動載具將是不可或缺的工具。學生若能透過手機、平板、電腦……等行動載具的搭配，即可從事各種形成性測驗，並且完整記錄學習歷程。

三、即時記錄回饋，發展適性教學

基本學力檢測的結果應可採取彌封方式或是更快的速度推播至家長的智慧型行動載具中，提供家長及學生個人化的回饋，此外，臺北市正在發展線上測驗系統，學生所進行的測驗除了可以如實的被系統記錄下來，同時也能即時從系統中獲得測驗結果的回饋，若能再配合適應性學習軟體的應用，測驗系統將可變得更加適性化，因應學生的學習狀態個人化提供不同的試題，教師也能從系統紀錄資料中探查學生的學習需求，進行差異化的教學。

四、長期縱貫研究，看見成長軌跡

符應大數據高容量與價值性的特質，透過數位工具的使用，學生的學習將可進行長期的記錄，而從這些學習記錄資料的積累中，則可看見學生學習成長

的軌跡。倘若基本學力檢測能針對每一位學生做長期縱貫的追蹤，將可瞭解個別學生及整個團體成長的趨勢，並從事相關研究的探討，找到學校及教師可以改變學生的因素，介入去改善將可促進學生長期的成長，這樣的辦學對學生來說就更有其價值意義。

五、結合校務系統，整合學習紀錄

大數據的運用主要是管理問題，而非技術問題，當記錄資料各執一處時，大數據的分析將會窒礙難行。以測驗記錄來說，不論是國小基本學力檢測、國中會考、甚至是大學入學考試的成績資料，在謹守個資保護的原則下，倘能與教育單位的校務系統加以結合，將個人的學習記錄加以整合，當運用高容量的大數據加以分析時，才能愈發精確的提供更多的可能性預測，從大數據的抽絲剝繭中體現真實性。

六、結合問卷分析，擴大影響層面

數字是冰冷的，要加上人的溫度才能發揮它的效益。當我們獲得數據分析的結果時，如何讓數據深化是很重要的。以基本學力檢測為例，透過問卷項目與大量的學生檢測成績分析，除了呈現學生學習的樣態，也能讓教師從中反思教學是否有效，而學校取得研究報告時，也可據以檢視學校效能，提出改進策略，更進一步提供政策制定的參考，顯見測驗數據的回饋不單單只是反映學生的學習成果，其影響力已然擴及教師、學校、甚至是教育主管機關等層面。

七、善用巨量資料，共創未來價值

當數據被分享、當數據被深化，這樣的數據才具有價值性，觀諸國內測驗資料庫的建置，其實已擁有許多大數據資料，但釋出量尚不多，著實可惜。反觀基本學力檢測的資料，若能在不違背個資處理的原則下，將資料釋出、提供研究，才是基本學力檢測影響力的擴大。數據會說話，唯有數據釋出才能完成更多的研究提供政策參考，未來可以獎勵學術研究的方式，鼓勵大學或研究單位善加運用這些大數據進行研究，共同參與地方教育建設的腳步，貢獻 know how，落實社會責任，共創未來價值。

伍、結語

「數據戰爭早已揭開序幕，當一切人類與非人類活動都可以被量化，能掌握並善用大量數據的人等於掌握優勢。」（臺灣資料科學協會，2016），大數據

可以讓我們擺脫過去的試誤學習方式，透過大量、快速取得、多樣性的資料分析，找出各種可能的答案在真實情境中的價值，然後再從中找出最佳解答。

測驗結果的分析即是一種大數據的應用，它是反映學生學習成效的重要指標，也是提供親師生了解學生學習狀態的回饋方式，而在大數據時代的驅動之下，教師所扮演的不再是我講你聽的角色，是轉變為幫助孩子思考的引導者，當數位工具及線上學習或測驗系統可以普遍運用後，教師的工作非但不是被取代，反而是擁有更多的時間去顧及每個學生不同的需求。由於大數據的資料蒐集與分析，可以提供教師更多關於學生學習樣態的回饋，就能協助教師調整出更適合個人化的教學方法。

學生從測驗中獲得即時反饋，教師、學校、教育主管機關從資料的多樣來源及長期記錄積累中看見學生學習成長軌跡與發展的可能性預測，這些都是大數據為教育所帶來新契機，但在大數據帶來的利多背後，也隱藏了倫理性的「標籤化」、「以過去決定未來」以及保密性的資安風險。然而，造成此些隱憂的癥結點並非資料本身，而是資料的使用方式。因此，只要在大數據的運用之下，設有嚴謹的保護措施，為資料的保存建立適當的監督保管機制，在資料的使用上符合倫理限制與規範，將可避免大數據的不當應用，並從大數據的分析中持續的學習與改進。

參考文獻

- 林俊宏（譯）（2014）。V. Mayer-Schönberger & K. Cukier 著。大數據：教育篇——教學與學習的未來趨勢（Learning with big data: The future of education）。臺北市：遠見天下。
- 陳清義（主編）（2006）。臺北市 94 年度國民小學國語文基本學力檢測計畫成果報告書。臺北市：臺北市政府教育局。
- 陳清義（主編）（2019）。臺北市國民小學 107 年度基本學力檢測計畫成果報告。臺北市：臺北市政府教育局。
- 臺灣資料科學協會（2016）。大數據（1）——什麼是大數據。取自 <http://foundation.datasci.tw/what-big-data/>
- IBM (n.d.). *Extracting business value from the 4 V's of big data*. Retrieved from <https://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data>
- Laney, D. (2012). *Deja VVVu: Others claiming gartner's construct for big data*.

Retrieved from <https://blogs.gartner.com/doug-laney/deja-vvvue-others-claiming-gartners-volume-velocity-variety-construct-for-big-data/>

Patgiri, R., & Ahmed, A. (2016). *Big data: The V's of the game changer paradigm*.

Retrieved from https://www.researchgate.net/publication/311642627_Big_Data_The_V's_of_the_Game_Changer_Paradigm