

天下文化

大數據

教育篇

教學與學習的未來趨勢

Learning with Big Data

The Future of Education

by Viktor Mayer-Schönberger and Kenneth Cukier

麥爾荀伯格、庫基耶 著
林俊宏 譯

| 第2章 |

CHANGE

改變

大數據重塑教育的最大優勢

就在於它能落實提升學習效果的三大核心概念：

回饋（即時讓學生知道學習成效）

個人化（真正做到因材施教）

可能性預測（讓學生面對真實的世界）

雙贏的巧妙設計

不論外表或行為，馮安（Luis von Ahn）就像是典型的美國大學生。他愛打電動、愛開著一臺藍色跑車四處奔馳，也像是《湯姆歷險記》書裡那個湯姆的現代版，老愛叫別人幫忙做事。但是，可別讓他的外表給騙了。馮安其實是全球頂尖的資訊工程教授，而且可說有十億人都曾經幫他做過事。

十年前，他還只是個二十二歲的研究生，就參與創造了一套圖片驗證機制「CAPTCHA」（英文全名是 Completely Automated Public Turing Test to Tell Computers and Humans Apart，意指「能夠分辨電腦和人的完全自動化公共杜林測試」）：網路使用者如果想要註冊免費電子郵件之類，就得先輸入圖片上幾個歪歪曲曲的英文字母或阿拉伯數字，證明自己是人類、而不是垃圾訊息機器人（spambot）。

後來，馮安讓這個機制升級成「reCAPTCHA」，並且賣給了谷歌（Google）；從此，網路使用者要辨別、輸入的字，不再是純粹為了驗證刻意設計的扭曲英文字母或阿拉伯數字，而是「谷歌圖書掃描計畫」裡機器無法辨識的英文字。這種使用資料的方法太帥了，完全是一石二鳥：既能完成線上註冊程序，又能協助谷歌辨識文字。

在這之後，馮安這位卡內基美隆大學的教授，還是繼續尋找這種一石二鳥的做法，由使用者來提供可以擔綱兩種用途的資料。馮安的想法在2012年搖身一變，成為新創公司 Duolingo，推出網站和智慧型手機的App，能夠協助使用者學習外語（這點馮安深有體會，畢竟他小時候住在瓜地馬拉，英語也不是他的母語），但不同之處在於，這套教學系統背後有個絕妙的設計。

這家公司將文章切成一個一個短句，再請使用者

翻譯、或是評估和修正別人的翻譯。一般翻譯練習，用的都是人工設計的句子，但是Duolingo的句子來源都是實際需要翻譯的文件，而且他們也會從中收取翻譯費。等到有足夠多的使用者都將某個短句翻譯了、並且評定某個翻譯版本為最佳譯法之後，系統就會接受這句譯文；最後再將所有句子的最佳譯文，集合成完整的文件。

現在，有些媒體公司，譬如CNN和BuzzFeed，都已經成為Duolingo的客戶，用以翻譯海外新聞的內容。就像reCAPTCHA圖片驗證機制，Duolingo也是一種愉快的「雙贏」設計：學生能夠免費學習外語，而且同時產生經濟價值。

好處還不只是這兩項。所有使用者和網站之間的互動，都會產生所謂「資料廢氣」(data exhaust) 這種副產品——在網路上，這個術語講的就是用戶的線上動作，例如：游標滑過哪裡、使用者點擊哪裡、在同

一頁面停留多久、打了什麼字等等。而Duolingo將這些資料廢氣全部蒐集起來，就會知道像是：使用者要花多久時間，才能夠在某種語言的某方面達到精熟？多少的練習量最理想？如果漏了幾天課會如何？

馮安瞭解到，只要把這些資料進一步處理，就能讓他看到什麼是最好的學習方式。這在非數位環境下可不是一件容易做到的事。然而，光是在2013年，Duolingo每天就有大約一百萬名使用者上線使用、每次停留超過三十分鐘，所以馮安已經擁有了相當巨大的研究母體。

關鍵是：誰在學習？

目前，馮安最重要的一項發現就是，以下這個問題根本就問錯了：「學生該怎麼學習，才是最好的學習方式？」

真正的關鍵並不在於「學習方式」，而是「學習者是什麼樣的人」！在馮安看來，目前對於外語的理想教學方式，很少是基於實證研究。雖然有很多理論，像是有人就主張應該先教形容詞、再教副詞，但背後並沒有具體的資料做為佐證。馮安也提到，就算真的有資料背書，常常也只是幾百名學生的小規模研究，要拿來推論整體的情況，實在有點危險。如果能運用數千萬名學生多年學習的資料來得到結論，不是更好嗎？有了Duolingo，這個夢想正在成真。

馮安分析Duolingo的巨量資料，有了一項重大發現：理想的語言學習方式，會因為學生的母語和外語的組合而有所不同。舉例來說，如果母語是西班牙語的人要學習英語，一般會先教代名詞，像是 he（他）、she（她）、it（它、牠），但馮安發現，由於西班牙語無法直接簡單表達 it 的概念，這個字對學生通常會造成困擾和焦慮。於是，馮安跑了幾個測試。

他發現，如果先教 he 和 she，但是讓 it 晚個幾星期再學，就能大幅提高撐下去繼續上課的人數。

馮安有些結論和我們傳統的認知正好相反。舉例來說，女性學運動詞彙的成績其實比男性更高，而男性學烹飪和食物相關詞彙的成績又優於女性。又例如在義大利，女性學英語的成績也打敗了男性。許多這樣的新發現正不斷浮現。

提升學習效果的三大核心概念

從Duolingo的故事，我們可以看到，巨量資料重塑教育的最大優勢所在。巨量資料告訴我們，提升學習效果的三大核心概念，分別是：回饋、個人化、可能性預測。

先談「回饋」。從幼兒園到大學，我們的正規教育都會提供各種回饋：不論是作業、課堂參與、報告