

教師對於「數值添加」評鑑模式尚有疑慮

駐舊金山辦事處教育組

依據統計學的觀點來說，現今的公立教育有點像 90 年代的棒球，對球員天分評鑑的機制並不準確。90 年代時，棒球隊伍們紛紛使用全壘打數與打點來歸類球員，這使得他們常將平均型球員歸類為明星球員。今天的教師評鑑系統認為所有人都是符合資格的，卻無法鑑別出真正的 MVP（最有價值的球員）。

20 年前開始，棒球界轉而採取不同的球員評鑑機制，將球員的各項表現放入評分並分析。這個機制在球界運行良好，教育研究員也以此重新省思教師評鑑的統計方式。研究者能夠依據統計分析，預測出學生一年後的學習成長結果。而一年後學生的真正學習表現與預測的差異，就是一個教師一年中使其增加或減少的變動「數值」，這個系統被稱做「數值添加模型」（Value-Added Models or VAM）。舉例來說，當一個學生在九年級的數學成績在學年的前百分之七十，統計將預測她在十年級的數學成績也將在前百分之七十。如果在州立測驗後，她的成績提升為前百分之五十，則理論上她的數學老師將被評鑑為增加二十個百分點的「數值」。

「數值添加模型」是個簡單分辨教師教學成效的評量方法，並且不會假設所有教師都在平均值上，讓學校上級能評鑑他們的教師於精確的百分比排名。這個評鑑同時表現一個教師統合性的一年性成績，而不像傳統的教師評鑑系統，校長只能依靠少數的觀摩活動評量教師表現（例如作者威廉·愛格的學校以往只使用兩小時的教學觀摩來評鑑教師表現）。更精確的「數值添加模型」數據，則需要伴隨更多的測驗。

儘管教師評鑑的方式已有改進，美國教育部長愛尼·當肯（Arne Duncan）卻宣布各州可自行決定是否延遲在教師評鑑中使用這種評量依據。這可能是教育部門擔心「數值添加模型」在實際執行上的精確度，尤其許多學校正在統一教綱的轉換過渡期。除此之外，還有其他理由讓教師們對此評量方式存疑。這個評量設定所有教師擁有一樣的教具質量與設備質量，但即使在同一個學校中，每個教師的教室設備與教具都不同，甚至每一班學生的情況都不同。例如作者威廉·愛格的同年級教學，同樣的教師、一樣的教材教室，一班學生比統計預期值進步更多，但另一班學生卻總是沒達到統計預期值。

教學是教師與學生互動的結果，而這種互動則和教師與班級組成有關。舉例來說，一班有熱心幫忙的學生會比另一班沒有的學生進步多。當然教師能創造樂於助人的學生，或激勵不合群的學生，但這種教師在「數

值添加模型」中，將和幸運教到許多用功學生班級的教師被放到同一個等級。美國統計協會回顧學術文章，發現教師品質在「數值添加模型」中只有 1 到 15 百分比的差異。在實質評量上，這個評量方式還有缺陷。即使在數據上，這個系統能顯示教師的差異性，但還是可能誤判。好的評量系統應該能更精確評定每個教師的表現，以及幫助他們成長。

附註：Model 在教育領域的學術翻譯為「模型」，但不曉得對於大眾是否該使用其他翻譯。（例如：模組。但模組在學術領域有較嚴格的定義，無法隨意使用。）

資料來源：2014 年 8 月 27 日 教育週報 William Eger 報導

連結網址：

http://www.edweek.org/tm/articles/2014/08/27/fp_eger_valueadded.html

原文標題：Education Is Not 'Money ball': Why Teachers Can't Trust Value-Added Evaluations Yet

