美國共同核心評量測驗的省思

曾建銘

國家教育研究院測驗及評量研究中心，測驗及評量組

壹、議題重要性

2008年，歐巴馬在競選總統期間，曾抨擊美國各州課程標準不一，是導致美國基礎教育在世界排名中持續滑落的主要原因。因此，他當選後，於2010年頒布首部《各州共同核心課程標準》，標誌著各州採用並實施全國統一課程標準的開始。目前，現已有45個州採用此標準，希望提高美國基礎教育品質、提升美國學生的國際競爭力。

目前美國採納《各州共同核心課程標準》（CCSS）的各州都正在為實施《各州共同核心課程標準》及「共同評量」（Common Assessment）作準備，多數採納共同標準的各州分別加入二個不同的聯盟規劃「共同評量」，這二個聯盟為「智慧平衡評量聯盟」（Smarter Balanced Assessment Consortium, SBAC）及「大學與職涯準備度評量夥伴聯盟」（Partnership for Assessment of Readiness for College and Careers, PARCC），這二個聯盟已經在2014年春天針對各會員州進行實地的試測，正式測驗的實施將從新的2014－15學年度開始。

這兩套共同評量法是根基於改革之後的共同核心標準，希望能夠更正確地呈現學生們的學識能力，取代聯邦之前飽受批評的「有教無類法案」（No Child Left Behind Act）所使用的評量方式。採行相同的評量法是為了能比較各州學生成績表現，做為教育改革成效的參考。

雖然幾年前各州在聯邦鼓勵下，有45個州及華盛頓特區雄心勃勃地加入積極落實共享的評量（shared assessments）這樣的理念及行動，但是，這些採用共同評量的各州逐漸減少，目前只剩下約26個州將在2014－15年採用共同評量，其他的則另採行其他測驗。會做此決定之考量原因包括：測驗的時間長度、測驗花費、聯邦與地方的政治問題，截至2014年5月中，只有17個州還在「智慧平衡評量聯盟」（SBAC），10個州（特區）在「大學與職涯準備度評量夥伴聯盟」。

這樣一個新的評量政策，基於提高美國基礎教育品質、提升美國學生的國際競爭力。從當初贊成的大多數而開始規劃、預試，到明年即將實施的日期漸近，卻發

生各州、教育學者或教師有不同意見，導致有些州擬延遲實施，甚或退出計畫的情況發生，這其中或許有些過程或細節值得我們省思。

貳、美國共同核心課程與評量計畫現況與挑戰

一、美國共同核心課程

共同課程標準是一套以K-12的英語與數學教學標準，各州都可以參考使用。標準訂立時的想法是基於各州標準不一，落差過大，不如統一制定各級學生應有的學習目標。全美目前簽署採用此項課程標準者，包括華盛頓特區和其他45州，這也是1992年起全面檢討K-12教育體系計畫的重要里程碑。

新的標準強調批判性思考以及理解力。學生在課堂或在標準化測試中，將同時以口語和寫作應答。「各州共同核心課程標準」減少了教條式的標準，並讓教師能夠了解班級中進度落後學生的需求，但也同時能夠照顧學習優良學生的需求。

自2010年起，沙加緬度學區助理教育局長泰勒（Iris Taylor）在她的學區內致力推動新的教學標準。她說「各州共同核心課程標準」鼓勵教育者「深入探究幾項教學標準，而非試圖推動多項標準。」新的標準著重於「培養所有學生進入大學和職場的應變能力。」

以數學學科而言，「各州共同核心課程標準」將減少數學教學的內容和份量，這樣有助學生深入了解某些數學課題，並利用數學的方式以口語表達他們的想法。

加州大學聖克魯茲分校（UC Santa Cruz）數學教育教授尤莫思塔（Judith Moschkovitch）認同經由「各州共同核心課程標準」將平衡學生的理解力和計算能力。

莫思塔說學生將在課堂中和老師及同學討論數學，經由這樣的方式將有效幫助學生理解和記憶數學概念。

在語言和文學方面，「各州共同核心課程標準」將幫助學生學習非科學類的學科，這些學科將提升學生的寫作能力，這亦是學生在未來職場上必須具備的能力。學生將閱讀難度較高的教材，並以口語和寫作的方式對這些題材進行批判性的思考。

某些批評者擔心教師不熟悉新的教學標準，這將擴大各州已經存在的成績差距問題。然而新標準推行成功與否，將取決於學校和學生是否有更多機會取得「高品

質的教材」和「優秀的教師」。

教育政策中心執行長、同時也擔任此研究主持人的Diane Stark Rentner表示：
「全面實施共同核心課程標準是一項耗時耗力的大工程，牽涉到教育體系中許多不同的層面。目前最重要的挑戰是充足的資源。政策制定者必須了解，資金短缺可能會導致各州刪減或暫緩這些計畫。」

二、評量實施計畫

以下說明兩大評量大致實施計畫（Doorey, 2014；PARCC, 2014；SBAC, 2014）：

1、實施年級與時間：根據「大學與職涯準備度評量夥伴聯盟」（PARCC）的規劃，2015年春天將針對會員州的學校3年級到11年級正式實施這套新的評量。而「智慧平衡評量聯盟」（SBAC）的各州則是將針對3—8年級及11年級，各州也可以另行要求9或10年級接受測驗。

2、共同評量測驗形式：這二個聯盟主要都是採總結式的測驗，也都是規劃在學年結束前3個月左右施測，其長度大約有7—10小時的時間，也都有不同的選擇題題型，雖然二個聯盟的測驗設計有所不同，但原則上都會包含表現本位（performance-based）的評量及年終總結評量的內涵。

3、實施測驗之階段：「大學與職涯準備度評量夥伴聯盟」（PARCC）的時程規劃，該聯盟下的學校將有二個階段的測驗窗，最多各為20天，第一個測驗窗將在學期進行超過75%後打開，作為學校實施表現本位（performance-based）評量的指引，第二個測驗窗將在學期進行超過90%後打開，來實施總結性評量。各州將與學區與學校來決定最合適的20天實施期限，並加強學校人員施測的訓練。「智慧平衡評量聯盟」（SBAC）的規劃為，將把表現本位任務測驗及電腦總結式測驗都放在同一個測驗窗，該測驗窗將開放12週的時間，在最後33%教學日之間開放，11年級則是在最後20%教學日期間開放。

4、長度與配時：「大學與職涯準備度評量夥伴聯盟」（PARCC）的測驗長度規劃，大約是3年級8小時、4—5年級9小時又20分鐘、6—8年級9小時又30分鐘、9—10年級9小時又45分鐘、11年級9小時又55分鐘。而「智慧平衡評量聯盟」（SBAC）則是3—5年級約7小時（包含1小時教室活動）、6—8年級7小時又30分鐘（包含1小時教室活動）、11年級8小時又30分鐘（包含1小時教室活動），該教室活動將由教

師引導進行表現本位之任務評量活動。

三、評量與績效責任

新的「測驗本位績效責任制度」(test-based accountability)，是指將即將實施的、對應於《各州共同核心課程標準》的「共同評量」，包括「智慧平衡評量聯盟」(Smarter Balanced Assessment Consortium, SBAC)及「大學與職涯準備度評量夥伴聯盟」(Partnership for Assessment of Readiness for College and Careers, PARCC)考試結果作為學校、校長及教師評鑑依據之一的績效責任系統。這些各州實施新的測驗本位績效責任制，原因起源於聯邦政策因素以及政治考量因素，其一方面是聯邦推動「沒有孩子落後彈性豁免計畫」(No Child Left Behind Flexibility)審核各州的教育改革計畫後賦予各州以「大學與職涯準備度」及新的績效責任制度來換取鬆綁豁免實施「沒有孩子落後」(No Child Left Behind)相關規定，另一方面則是美國近來積極推動教師評鑑改革，加諸教師提升學生成就及學業水準之責任，各州乃因而加速採納新的指標或是大修對於學校及校長及教師評比的方式與制度；此外，又因逢《各州共同核心課程標準》及「共同評量」的公布與上路，各州爰群起以採納該新課程及測驗作為宣示改革之決心，並且規劃以新測驗結果作為績效責任指標。

雖然《各州共同核心課程標準》(Common Core State Standards, CCSS)及其對應的「共同評量」(Common Assessments, CA)將在新的學期實施，不過，2014年以來，因為基於相關因素影響，包括時程急促、學校及教師反彈、配套不足、議會質疑等，各州面臨著相當複雜的政策決定：何時實施新績效責任制度？何時將測驗結果反映在人事決定？何時可以讓校長及教師喘息？這些複雜的問題讓許多州又紛紛立法或政策決定延遲或調整原本即將上路的測驗本位績效責任制度(test-based accountability)，延後將新的「共同評量」測驗結果作為學校評比及教師評鑑之依據。

四、美國教改共同核心課程的標準測試現況 (Smarter Balanced Assessment 官方網站, 2014)

經過四年的推動，共同核心課程的標準測試，也稱為「智慧平衡評量」(Smarter Balanced Assessment Consortium)，這個學年度將在部分地區(州)正式施行。去年許多州都已經試行智慧平衡評量，測驗成績也將會作為今年修正的重要憑據。

今年共同核心課程仍然有許多方向待努力。但與此同時有些州開始退出這個體

系，如愛荷華州7月底宣布，即日起退出智慧平衡評量聯盟。未來該州將採行Iowa Core 愛荷華核心課程，為該州教育方向直接負責，顯示該州對於共同核心課程的導向缺乏信心，即使會喪失聯邦補助經費也在所不惜。8月份路易斯安那州長金道爾（Bobby Jindal）也控訴美國教育部違憲。他指出教育部及鄧肯部長侵犯各州州權，擅用經費來誘使各州接受共同核心教材。

雖然如此，共同核心教材的標準測試，甫於上週增加十個盟友，包括夏威夷，愛達荷，密蘇里州，蒙大拿州，俄勒岡州，南達科他州，華盛頓和美國維爾京群島，簽署了備忘錄同意加入智慧平衡評量。該聯盟也宣布，預計近期將有另外9個地區計畫加入聯盟，也就是會有更多州更多地區開始這套標準測試。

五、共同核心測驗的迷思

聯邦教育部長鄧肯（Arne Duncan）多次公開支持這項2014-15學年度上路的共同核心測驗，推崇這項共同評量是公立學校革命性的創舉（absolute game-changer），將徹底改革十多年來推行的制式選擇題標準測驗。但是美國公平及公開測驗中心（The National Center for Fair & Open Testing, Fair Test）則不這麼認為，其提出的共同核心測驗迷思包括如下（Fair Test, 2013; Strauss, 2013）：

迷思一、共同核心測驗將比現在的考試更好，有更多考題可以測出高層次能力（Higher Order Skills）

事實：新的測驗仍然以傳統多項選擇題（Multiple Choice Questions, MCQ）為多數，雖然支持者指出這項新的測驗可以協助教師強化批判思考，但實際上該項新的測驗除了增加一小部分的表現任務（short performance tasks）或短篇文章寫作，其他仍與目前測驗形式相當。

迷思二、採納實施共同核心測驗將可終止《沒有孩子落後》考試的過度殺傷威力（overkill）

事實：在《各州共同核心課程標準》底下，仍將會有很多考試及誤用情形發生。

《沒有孩子落後》法案造成過去十年來全美如海嘯般的測驗狂潮，同樣地，《各州共同核心課程標準》也將如帶領各式測驗如浪潮般淹沒教室。目前這二個測驗聯盟都規定中小學部分的3-8年級實施英語文及數學二個科目測驗，這個部分與《沒有孩子落後》一樣。但其測驗將比現行各州測驗來得長，因為在高中部分，「大學與職涯準備度評量夥伴聯盟」（Partnership for Assessment of Readiness for College and

Careers, PARCC) 規劃高中三個年級都將實施閱讀及數學二個科目測驗, 「智慧平衡評量聯盟」(Smarter Balanced Assessment Consortium, SBAC) 則是在10-11年級實施。「大學與職涯準備度評量夥伴聯盟」(PARCC) 的各州高中生也將接受說與聽的考試, 同時, 從幼稚園到2年級也都提供有形成性測驗讓學校使用。這二個聯盟也都鼓勵每年多實施額外的2-3次考試。與《沒有孩子落後》同樣, 這二個聯盟的測驗也都將被用在高風險(high-stakes)的相關決定上, 包括教師評鑑與學校績效責任等, 同樣有可能產生過度使用或誤用。

迷思三、新的多州合作之評量將可以節省納稅人的錢

事實: 多數州的測驗費用將不減反增, 學校也將花更多錢在電腦設備升級上。由於經費考量, 有五個州已經在本年8月決定退出評量聯盟, 「大學與職涯準備度評量夥伴聯盟」(PARCC) 坦承該聯盟中半數以上的州將花費比過去更多經費在新的測驗上

迷思四、新的評量聯盟將能真正設計出比其他測驗公司更好的評量。

事實: 評量聯盟仍是由同樣營利的測驗公司, 包括Pearson, Educational Testing Service and CTB/McGraw-Hill等設計試題, 這些公司長久以來仍多所誤失。

迷思五、各州一定要實施共同核心評量

事實: 不, 他們不一定要。高品質的評量能改進教與學並提供豐富的資訊給學校運用, 有用的評量應該是課程的一部分, 包括形成性評量、表現評量、檔案評量或學生作業任務等, 學校也可以運用不同的證據來源來評估其辦學優劣, 而非只是單一測驗。

六、挑戰與困境

(一) 共同核心課程實施的正反面意見

保守派認為, 「歐巴馬核心」代表聯邦政府打算接管公共教育事務, 左派則擔心那會引來更多標準化測驗, 教師也得接受更仔細地審視。共同核心(Common Core) 教學標準充滿政治疑慮, 很容易讓人忘記, 為何一開始會有那麼多州採行。

6年前, 各州採行不同標準, 高中文憑亦失去意義; 多達4成的大學新鮮人需要補習課程。共同核心建立齊一標準以確保學生的能力; 學生每年都得學會特定技能, 但學校可以自行設計課程計畫。2010年, 45個州和華盛頓特區決定採行共同核

心；2014—15學年開始，新測驗將測試學校是否有達成這些標準。

多數教師支持此標準，但考試日期將近，反彈聲浪也越來越大；印第安那率先退出共同核心，幾個共和黨州亦跟進。阿拉巴馬、亞歷桑那、佛羅里達、喬治亞、堪薩斯、密西西比、田納西和威斯康辛，也都因為共同核心出現了激烈的爭執。不過，這些州最後並沒有退出。部分原因在於，主流保守派傾向於支持這套新標準；企業大多支持，因為它們希望能有教育程度更好的勞動力。此外，退出的州也遇上不少問題，例如，印第安那最終就採行了與共同核心十分相似的標準。

大部分的州保留了共同核心，接下來的問題則是，它們會不會全心推動。許多州表示，其課程與共同核心相符，但拒絕實施新測驗。問題在於，沒有合適測驗，學校就沒有改變課程所需的動機和資訊。

（二）費用問題（高等教育記事報，2012）

許多州也面臨了商業上的問題，像是採買課綱的材料、專業培養、評量工具、與科技。這些都是幫助校方符合新的學術指南與新測驗標準的需求。雖然與這些廠商高達數百萬的合約金都是有補助的，但這些標準與評量一直影響著州政府與學區如何使用他們的經費。然而，在資源更新與替換資源方面，花費該如何運作則尚未有明朗的分析。

共同課綱、測驗與學區預算之間的複雜關係也在許多州中可見。像是俄亥俄州在過去幾年已經把經費幾乎都花在購買共同課綱相關的產品，從評量測驗系統到無線科技都有。舉例來說，俄亥俄州與其他州就預計在2015年春季在線上測驗化花費數千萬美元。

目前這兩套評量法遭受批評的主要原因，在於高額的成本及設備需求。PARCC今年七月公布，每一個學生預計要花29.5美元來參加數學及語言評量考試。對於使用PARCC評量法的各州來說，這個金額要比他們之前的聯邦考試費用高出快一半。

不過目前最艱難的挑戰，其實是各州經費不足的問題。共有21州認為2011-12學年度實施共同核心課程標準時的重大困難之一，就是資源有限。除此之外，各州大多認為現有教學科技設備不足，恐怕無法依照共同核心課程標準施行相應的線上評量：20個州表示必須增加學校電腦數量，15個州現有無線網路設置及頻寬流量均不敷使用，14個州擔心缺乏專業知識處理線上測驗的相關技術問題。

參、我國現況概述

目前我國十二年一貫課程的規劃正進行中，是否訂立類似美國的課程核心標準，以及對核心課程標準施予電腦化評量，來檢核學生於十二年基本教育過程中與畢業後所學得的知識技能，似乎還在規劃中。

而攸關升學考試分發的大學學科能力測驗、大學指定考試則已進行多年研究、規劃，並以開過多次研討會與公聽會，美國共同核心評量到目前的準備過程與引發的爭論可作為借鏡與省思。

肆、對我國的啟示或建議

一、是否與績效責任連結？

若以測驗的目的，將測驗結果單純運用在學生學習成效、教師教學改進與學生補救，而不與學校辦學績效、教師教學績效相連結，所引起的反對聲浪會較小。

二、是否與入學考試連結？

而根據美國教育週報（Education Week, 2012）報導，大學理事會（College Board）於2012年5月16日宣布David Coleman為下一任大學理事會主席。分析指出，這項新的人事任命意味著未來「學術能力測驗」（Scholastic Aptitude Test, SAT）將可能加強與「各州共同核心標準（Common Core State Standards）」的連結（alignment）。

目前大學招生及入學研究方案已將未來十二年國教課程的規劃納入參考，未來若我國有共同核心標準測驗，是否納入參考值的審思。

三、該用什麼指標去衡量共同核心計畫的成效呢？

共同核心計畫的成果取決於如何反思編制教學課程及其教授方式。如果依然用舊有的課程和教學來教新標準，如在教師專業培訓工作坊只告訴教師：「看，這個就是核心，跟著教吧！」，或如果我們只管用平板電腦（iPad）來美化試卷，那麼這計畫也是徒然的。

但如果教師能善用標準去將教學變得多樣化以配合不同學生的需要、能利用科技協助教學、教師間能增加合作、解決問題，而看到學生更投入學習、看到他們英語及數學水平的提升，這就是我們的量度指標。由於經舊制度的洗禮，轉制對高中

生未必會有顯著的影響。但共同核心對年幼學生的影響是非常值得觀察。

四、哪些共同核心值得家長注意？

實施後如果成功，學生的成績會提昇，在升大學和就業方面有更好的預備；其次設立新的考試標準會花上數年時間，屆時會有很大改變，我們要為此預備。家長要明白當有新制實施時，學校表現和孩子的成績都會改變。最後就是要向家長說明如何預備新課程、設立新試制度，還有如何理解新成績。

五、教師的培訓

新課程標準的師資培訓工作也將是新標準成功與否的一大考驗。然而根據美國教育周刊（2014）研究中心調查顯示：教師接受越來越多關於共同核心標準的訓練，但對於教授共同核心標準課程感到準備不足。極少數教師對於由「為大學及職業做準備的評鑑夥伴（PACC）」及「智慧平衡評量聯盟（SBAC）」合作團隊共同研發的共同核心標準評鑑感到熟悉，僅一半的教師表示熟悉團隊的數學項目測驗，而2/3教師表示熟悉英語/語言項目測驗。因此，我們也要引以為戒，實施前，就做好教師培訓工作，並能對學生、家長與外界說明。

參考文獻

高等教育記事報（The Chronicle of Higher Education）(January 25, 2012)。

<http://chronicle.com/article/Despite-Budget-Problems/130442/>

教育周刊（2014）。教師尚未準備好教授共同核心標準課程。駐波士頓辦事處教育組，黃瑋婷翻譯。

Doorey, N. (2014). The Common Core Assessments: What You Need to Know. *The Educational Leadership*, 71(6).

Using Assessments Thoughtfully, Pages 57-60.

Retrieved from

<http://www.ascd.org/publications/educational-leadership/mar14/vol71/num06/The-Common-Core-Assessments@-What-You-Need-to-Know.aspx>

Education Week (2012). 大學理事會新任主席希望強化SAT與共同核心課程的連結。駐美國代表處

文化組，

張佳琳翻譯。報導連結網址：

[http://www.edweek.org/ew/articles/2012/05/16/32collegeboard.h31.html?
tkn=NMSFjWEBlii7bmNF56NZ2ErEZl%2FssZCW1Av4&cmp=clp-edweek](http://www.edweek.org/ew/articles/2012/05/16/32collegeboard.h31.html?tkn=NMSFjWEBlii7bmNF56NZ2ErEZl%2FssZCW1Av4&cmp=clp-edweek)

Fair Test. (2013, September 3). *Common Core Assessment Myths and Realities: Moratorium Needed From More Tests,*

Costs,Stress.

Retrieved from <http://fairtest.org/common-core-assessments-factsheet>

Partnership for Assessment of Readiness for College and Careers (PARCC). (2014).

Official Website. <http://www.parcconline.org/>

Smarter Balanced Assessment Consortium (SBAC). (2014). Official Website.

<http://www.smarterbalanced.org/>

Strauss, V. (2013, September 4). *Seven facts you should know about new Common Core tests.* The Washington Post,

The Answer Sheet. Retrieved from

<http://www.washingtonpost.com/blogs/answer-sheet/wp/2013/09/04/seven-facts-you-should-know-about-new-common-core-tests/>