

學生藝術評量及藝術課程評鑑中教師關係重大

查理士·董
佛羅里達州立大學

目前美國至少有卅八州都在大力推行學齡前至十二年級 (K-12) 的教學改革，2002年加州、德拉瓦、愛達荷、肯塔基、密西根、奧克拉荷馬、奧立岡、田納西、德州、猶他、維吉尼亞、懷俄明、佛羅里達等州，相繼宣佈畢業條件的新規定、新的教學成效模式、新的檢定與畢業要求、所有學科的全州測驗、成就成效模式與補救教學等 (NAEA 1999)。美國三所大學所推行的「藝術成就評量模式」(MAAP) 研究專案，正是針對上述的教改措施，從2000年到2001年進行一年的研究，對象包括佛羅里達、伊利諾、印地安那三州的70位藝術教師、一千多位K-12學生，研究資料則包括8000件學習檔案。本文為這項研究案的研究報告。

全美各地的藝術教師，目前都面臨藝術評量的難題，再加上缺乏標準化的藝術測驗，也沒有全學區的評量計畫，無法經濟、準確的評量各校的藝術課程。例如佛州教育廳目前就沒有任何藝術評量計畫。一般認為藝術教育界對於該教什麼、如何評量，意見仍然太過分歧。如果沒有適當的測驗及務實的學區評量計畫，美國多數中小學的藝術課程，很可能永遠沒有評量，而目前的大環境，普遍認為無法評量的東西也就不可能教，因此，未來藝術甚至可能會被排除在課程之外。

改革運動

Cusic (1994) 指出，改革派一般認為中小學出現了權力真空，必須以授權及規範來控制教師，還將各種教育問題歸咎於「教師中心」的班級結構，認為教師必須接受改革或管制。改革派主張州政府當局的權力應該提高，減少班級教師的權力，這種主張一部分的原因就在於改革派乃來自校園之外，多半沒有什麼教學經驗。

另一方面，學校教師通常抗拒改革，但教師其實是任何改革運動真正的決定因素，Cusic認為教師應該有權自行選擇是否參與改革，他更指出教師之所以選擇教學，正是因為個人的詮釋與選擇在教學專業上是非常重要的，尤其是教師是許多的個人，不是一個集體。他認為教師需要的不是加強管理控制，而是有機會加入改革行動，不受州政府的強制要求。這一點正符合Cusic所謂教師是半自治的個人，能夠獨立自律、自我評量，訂定自己的標準。然而，美國各中小學的藝術教師若要做到自律、訂定自己的標準，還有許多困難必須克服，其中有些困難正是源於「目標2000」教改方案，包括「全國教育目標」、「全國及各州教材標準」，以及「全國教育進步評量」(NAEP) 目前所推行的全國評量辦法。

全國教育目標

首先，藝術教師必須瞭解，「全國教育目標」是「目標2000」的基礎，其目的是做為公共政策，大部分與課程的美學或藝術目標無關。這些做為公共政策的全國目標，主要涉及地方、各州及聯邦政府的社會、政治、經濟目標，其關切重點是法治、就業、商業等議題。藝術雖然有助於實現一些社經或政治目標，但這些目標卻非藝術的必要目的或充分目的，也不是藝術最能發揮成效之處。

全國視覺藝術標準

參與《目標2000教育美國法案》制訂工作的藝術專業人士，在1990年代初訂定了六項全國視覺藝術標準，隨後各州所訂的標準中，或多或少都加入了這六項全國標準，例如佛州納入其中五項，佛蒙特州納入三項。

這些標準明確規定藝術科所應評量的基本成就，確實有其用處；但我們必須注意，這些標準在課程與評量方面，也反映了學科取向的立場，而學科取向教學，到現在已經受到許多批評，有人認為是失敗的課程，而且最初其實是針對通識教育所設計的在職進修計畫 (GCEA 1994)，本身並不是要做為課程的 (Kaagan 1990)。教師也應該注意，全國教材與成就標準，並沒有告訴我們該教哪些藝術技巧，該強調什麼標準，或是各項標準的重要性是否有輕重之別。

NAEP藝術測驗

「全國教育進步評量」(NAEP) 藝術測驗，是由Getty贊助，1995年與1997年實施了「全國藝術能力」試測。這項評量的開發成本大約1000萬美元，最初預定要針對約4500名四年級、八年級、十二年級學生實施。然而，1997年針對2999名八年級學生所做的試測，結果似乎未能提供我們所要的答案，反而凸顯出更多問題 (Persky, Sandene & Askew 1998)。

試測結果公佈之後，各界對NAEP藝術測驗提出了許多質疑，其中之一是全美符合測驗資格的學生可能高達6000萬人，但只有5000人左右接受測驗，要等到2007年現行試題淘汰後，才會對所有學生全面施測。有些專業人士則指出，有關美學、藝術批評、藝術知識的選擇題和問答題，考試時間為60分鐘，但成就表現的考試時間只有30-40分鐘，並不符合實際情況，因為事實上學校的藝術課程，可能多達90%都是創作活動。

有些人擔心這項測驗的設計是標準化測驗，每個學校的每一位學生，都必須回答相同的答案，因此也就假定美國所有的藝術教師，都應該教授同樣的藝術活動、採用同樣的教學方式，也假定所有學校的所有學生，都能接受同樣份量、同樣品質的藝術教學。例如有些測驗結果發現51%的受試學生，未能說明Romare Bearden的一幅構圖，58%拼貼技巧不及格，95%不懂得如何使用塑膠黏土，97%無法使用鐵線和塑膠黏土創作雕塑作品，45%無法舉出當代西方藝術的例子，75%不認識立體派初期的作品，而且與其他研究的結果相反，藝術測驗中發現女生表現優於男生，白人與亞裔優於黑人與西語裔學生，最後，學生測驗成績與父母的教育程度有直接的關聯 (NAEA 1999)。

一項全國測驗，如果半數學生或半數以上都答不出來，顯然試題並沒有反映出全國藝術教師對學生所預期的知識或技巧。NAEP測驗是根據「全國目標」，而「全國標準」似乎並沒有測出學生在多數藝術課程中學到的東西。也可能NAEP測驗與全國標準，反映的是改革者認為所有藝術教師「應該」教的內容，但卻忽略了在實際情況中，學生的美學能力有高有低，各校的藝術教學品質也不相同。

藝術教師教些什麼

要達成全國目標，我們首先必須檢視藝術教師實際上教些什麼。如果我們根據藝術教師實際上的教學內容，來訂定教學結果，而不是以行政命令做由上而下的改革，會有以下優點。第一，教師能夠自己掌握學習與評量過程；其次，如此教師就必須負責將自訂的目標清楚傳達給學生和學校；第三，這樣的評量過程，對學生和家長都很公平。

此外，如果將評量直接與教學內容相關聯，教師會覺得不需要再根據別人的規則，而是依自覺的努力，讓學生知道應該要做到什麼。這樣一來，教師就要負起完全的責任，讓自己的教學與目標相配合，教學成果也要明確反映預訂的目標。藝術教學若要做到這一點，那麼學生所學的藝術創作、藝術史、藝術批評，就必須清楚反映在書面、口頭與視覺的教學產物中，包括表達與認知兩方面都是如此。此外，我們如果開始考量教師實際上教些什麼，學生又學到什麼，教師就會明白，有效的教學取決於本身要知道自己哪些地方教得好。

評量應符合真實的情況

真實的評量是要讓學生從事有用、重要、有意義的活動，這種評量在形式上屬於學習活動，涉及概念和更高層級的思考技巧，也需要連結數種不同形式的知識。這樣的評量也會具體說明對學生作品的評斷依據，本質上其實是在建立標準，而不是在測試標準。因此，真實的評量，最終目標在於促使學生能夠透過教學活動，學會教學內容且達到成就標準 (Dorn 1999)。

藝術上真正的學習，也代表能夠有目的、有意義的應用相關訊息，而非只是學會事實知識而已。這種學習也能帶動課程實務上的改變、評量過程的改變，以及藝術的改變，教學結果不再要求所有學生都必須以同樣的方式學

會同樣的東西。

考學生學過的東西

如果教師的教學內容，主要都在於學生的藝術表現，那麼測驗自然也應該著重在學生的藝術。藝術表現的評量，指的是我們在學校藝術課程中一般常見的活動，也就是創作，同時在過程中評斷作品。雖然表現評量並非全新的東西，但評分程序著重於明確界定工作、每項工作有一定範圍的分數，卻是新的做法。

真實評量以及下文會討論的MAAP專案，所用的評分標準稱之為量表(rubric)。量表能夠做為評分的工具，利用序數量尺或Likert類型的量尺，將表現排列出高下順序。所用的量尺通常是「標準參照」，具體說明某一年級的學生一般應該能夠達到的表現等級，而非根據一些籠統或絕對的藝術水準。因此，從量表所得的分數，比較可能顯示學生的成就是否符合某一年齡、某一年級所應達到的標準，或是高於標準或低於標準。因此，藝術表現評量應注意以下四點：

- 應基於藝術課程中學生所學過的內容，或是基於教師希望學生知道或能夠做到的東西。
- 評量藝術表現的整體品質相對於同班或同校其他學生，位於什麼等級。
- 應解讀為反映特定學生在特定班級中的學習成果，而非用於學生之間、教師之間或學區之間的比較。
- 根據評量結果以決定學生應該學習的內容、學習的成效如何、教室環境需要做何調整，才能夠促使學生更上層樓。

應該評量什麼

在藝術教學中，有三大重點應該評量，也就是(1)表達、(2)知識與技巧、(3)概念之形成。首先，我們需要評量「表達」，才能知道課程及學生的學習，在哲學上是否符合藝術的手段與目的，是否具有足夠的學科效度，以呈現藝術探究的產物和探究的手段方法。其次，需要評量知識與技巧，才能推知課程的哲學效度，也就是有關人類成長與發展、學習、個別差異等學理。這一點需要利用分析評量，回答「應該教什麼」、「什麼時候教」、「應該教

誰」、學生的的想法、兒童的思維推理如何成長改變等等問題。

「評量藝術表現之模式」(MAAP) 專案

2000-2001年的MAAP專案，對象包括三個州70位K-12的藝術教師，以及1000名學生，共同參與一項藝術教學真實評量研究，藉此呼籲學校行政人員及議員重新考慮全國測驗政策，不應採用一套預定的教育標準和評量。MAAP由NEA贊助，由三所大學藝術教育教師及美國十一個學區共同成立「全國教育基金會」，以發展數項K-12藝術評量模式，希望日後能夠推行至全美各學校。基金會有三大活動獲致相當的進展：(1)教師訓練與評量發展研究機構；(2)將研究推行至學校藝術課程中；(3)將研究成果推廣至藝術教學專業領域。負責訓練及督導研究過程的三所大學（佛羅里達州立大學、普度大學、北伊利諾大學），以及佛羅里達、印地安那、伊利諾三州參與專案的十一個學區，藝術教學成績都深受全國肯定，對整個藝術教學專業貢獻良多。

各校的研究發展，著重於以下的重要需求：(1)協助教師瞭解如何進行真實的評量模式，藉此評量班上學生的作品；(2)協助教師根據自己的班級或學校，發展適用的評量計畫；(3)設計一個資料蒐集系統，能夠配合藝術學生、學校、各州及全國藝術評量標準的需求。

專案活動包括：(1)訓練教師使用藝術量表，以評量K-12學生的藝術表現；(2)由教師同仁體驗「盲目評分法」，以瞭解教師對學生作品之評分的效度；(3)訓練教師使用真實評分的學生作品，做為課程工具，以改進藝術教學；(4)針對特殊需求，發展評量檔案及分析量表；(5)發展評量工具的方法，以及記分的方法，以配合學生需求及「目標2000」中各州及學區的標準。各校講師包括藝術家、課程與評量專家、藝術師資教師等。藝術家提供美學及技巧方面的知識，協助教師增進本身的表達能力；參與的學區包括佛州的Pinellas和Dade郡學區、印地安那州的Washington和Wayne兩鎮，以及伊利諾州的十一個學區。

研究專案的詳細結果，在2002年邁阿密的NAEA會議中發表，也會收錄在〈評量表達行為〉一文(Dorn, Madeja, and Sabol, 2002)。關於學生學習檔案的評量過程，從將近2000個學習檔案、16000件學生藝術作品所得到的資料中，分析結果證實以下數點：

- (1) 教師受過訓練後，能夠對學生的創作表現進行量化、可靠的評量。

(2) 藝術教師受過適當訓練後，能夠自我規範，建立本身的標準，對學生的表現進行有效且可靠的評量。

(3) 在這些真實評量環境中，採用專案量表，且教師都熟悉藝術創作形成的本質，如此所進行的評量，能夠有效測量出學生的表達成果。專案提供的量表，是根據兒童的發展所編訂，同時也考量教師本身對藝術思維與藝術創作的直覺知識。

(4) 學生的藝術表現與進步情形，各班級、各年級、各學區都不相同，也就顯示學生與教師的能力，以及學校環境，並非一致均等的，因此教師之間、學生之間、學校之間、學區之間做比較，既沒有實用，也無助於提升教學。

(5) 學生表現有所進步，與教師工作坊、學生的年級、學生的表達能力，都可能正相關。整體而言，在各年級之間、接受相同或不同工作坊訓練的教師之間、表達能力不同的學生之間，學生表現的進步情形並不均等。

無數學區與各州教育廳，都堅持藝術教師必須評量學生的表現，然而短期內各州或各學區並沒有藝術測驗或學校藝術評量的計畫，因此需要類似MAAP專案所顯示的新方法。我們應該組織全美的藝術教師，共同決定學校本位評量該走的新方向。因此我們建議，各州藝術教師應該開始：(1)發展自己的真實評量工具；(2)透過同仁審查的過程，發展學校與學區評量計畫；(3)設計適當可行的方式，以記錄學生的進步情形，同時能夠符合學區或各州的預定目標。

總結而言，採用真實的方法，以評量K-12的藝術教育，同時讓藝術教師負責實際的評量工作，應該是可以推廣於實務的。要做到這一點，藝術教育專業人士不能只是批評各州統一的紙筆測驗，或只是抱怨各州的強制規定與績效要求，導致學生、教師、學校、學區之間都成了競爭對手。

各州教育當局如果認為課程發展就是規範教師，而不是協助教師自我規範、建立自己的標準，當然就不太可能贊同這樣的改變。官僚往往認為他們的首要職責，就是執行各項規定，要求各校進行議員所決定的改革；政治人物關切的是選民的要求，要將毒品趕出校園，要提高畢業標準。若是信任藝術教師能夠做到這項工作，就等於剝奪了官員的規範管理功能，他們也就沒有存在的理由了。所幸官僚都很清楚，教師既不團結，也不信任測驗，所以他們可以繼續要求毫無意義的是非題／選擇題測驗，看看哪些學生、教師或

學校考得最差，而不是努力讓學校成為更理想的教學環境。

要克服測驗專業人士的偏見，也同樣困難。測驗如果考慮到課程目標不同、學習環境不均等，對於統計人員來說，就難以決定誰通過了、誰不及格；在他們的觀念裡，測驗並不是用來找出學校、學生、教師能夠如何提升教學，也不認為不同學科需要不同的評量策略，以評量不同形式的知識。每人的興趣、求知慾、學習方式各不相同，但這些個別差異，卻是測驗發展人員最討厭的，他們總希望統計變數能夠明確、便於操控。

我們最需要的是具體的證據，證明由教師所設計、執行的測驗，不但跟常模參照測驗一樣有效、可靠，而且更可能鼓勵學校不斷改進，不會因為不同而有不利的影響。最後我們必須瞭解，美國的中小學生美學能力並不相同，各校所提供的美學機會也不平等；但我們至少可以假定，這兩個情況都不是完全不可改變的，那麼最理想的發展應該是決定孩子需要學些什麼、需要學會哪些能力，以此做為學校與教育改革的首要重點。

參考文獻

- Cusic, P.A. (1994). Teachers regulating themselves and owning their own standards in *The Future of Education Perspectives on National Standards in America*, 205-257. Cobb, Nina (ed.). New York: College Entrance Examination Board
- Dorn, C. (1999). *Mind in Art: Cognitive Foundations in Art Education*. Mahway, NJ: Lawrence Erlbaum Associates Publishers.
- Dorn, C., Madeja, S., and Sabol, R. (2002). *Assessing Expressive Learning*, Mahway, NJ: Lawrence Erlbaum Associates Publishers (in press)
- The Getty Center for Education in the Arts (GCEA) (1994-95). *Summer Blues*, J. Paul Getty Trust, Los Angeles.
- Persky, H. R., Sandene, B.A., and Askin, J.M. (1999). *The NAEP arts Report Card: from the National Assessment of Educational Progress*, Washington DC: U.S. Department of Education Office of Educational Research and Improvement
- Kaagan, S.C. (1990). *Aesthetic Persuasion: Pressing the cause of arts in American Schools*, Getty Center for Education in the Arts, Los Angeles.
- NAEA (1998). NAEP and the visual arts, framework, field test and assessment. *Art Education*.
- NAEA (1999). Student achievement in the arts fall short, *NAEA News*, 41 Feb 1, 1999, 1-3.