



多元化的教學評量

圖 張清濱

壹、引言

教學評量乃是教學過程中重要的一環。學生有必要瞭解他（她）們學得如何，教師也有必要知道他（她）們教得如何。教學評量就是一種手段，它是用來判斷教學目標是否達成並且據以改進教學。Michaelis（1980）即指出：評量的過程包括三個階段：擬訂目標，蒐集、組織並解析達成目標的證據，利用解析的結果據以改進教學並報告學生進步的實況。然而，目前學校教師過份熱衷於考試，學生、家長却過份計較於分數，殊少用之於改進教學。顯然，教學評量已經產生了偏差。本文擬就教學評量的基本原則，探討教學評量的方法與技術，以供教師們改進教學的參考。

貳、教學評量的基本原則

要使教學評量發揮效果，教師至少必須遵循下列十個原則，今略述如次：

一、評量應根據教學目標

評量的主要功能在於判斷教學目標是否達

成。在任何評量的過程中，評量的目標均必須具體、明確。教學目標通常都以行為目標的方式敘寫；評量的內容自應涵蓋教學的目標。

二、評量應兼顧認知領域、技能領域及情意領域

教師的觀點往往左右評量的重點。如果教師注重學習單元中知識的獲得，則評量的內容偏重於認知領域。如果教師注重態度與價值觀念的培養，則評量將着重於學生在這些態度與價值的程度。但是，教學是一種完整的活動。知識、思考過程、技巧、態度、價值及行為的型態等一樣重要，不分軒輊。雖然某些教學目標有輕重緩急之分，教學評量的內容却不宜有所偏廢。有些教師偏重於知識的灌輸，忽略技能及情意的培養，因而只造就一些五育不全的學生。即使認知領域，部份教師僅評量學生的記憶、理解、應用而已，殊少評量到分析、綜合、評鑑等高層次的認知領域。難怪許多學生不善於批判思考。

三、評量應適應學生的個別差異

部分學校，教學評量純粹為應付升學考試，全校同年級月考試題相同，期考試題也相同。考試的結果，當然會有相當高比率的學生不及格。更嚴重的是有些學校使用坊間測驗卷，超出課程範圍。學生拼命用功，即使挑燈夜戰，也甚難及格。如果屢次「滿江紅」，就會喪失學習的信心。學生一旦放棄課業，就會衍生另外的問題行為。因此，教學評量應由任課教師根據班上學生的程度，命擬合適的試題或採用合適的方式，評量學生的學習結果。基於此認識，教學評量宜採標準參照評量（*criterion-referenced evaluation*）而非採用常模參照評量（*norm-referenced evaluation*）。評量應以激發學生的學習動機與熱忱為考量。

四、評量是繼續不斷的過程

教師在教學前、教學中及教學後都可評量學生。教學前，教師應施以診斷性評量（*diagnostic evaluation*）以辨別學生個人及團體的需要。教學中，教師應施以形成性評量（*formative evaluation*）以看出學生每天進步的實況並驗證教學歷程有無缺失。教學後，教師應施以總結性評量（*summative evaluation*）以瞭解單元目標是否達成。這三種評量的方式，教師在教學的過程中，應相機採用。

五、評量應兼顧過程與結果

這是說評量應重視學生的學習過程，也要重視學習的結果。有些教師評閱試卷或測量時，只看結果，不看過程，無法正確判斷學生的思考過程，亦無法衡量解題的品質。譬如兩位學生的答案均正確，但解題過程却有程度上之不同，給分自應有所不同，始能正確測出學生

的學習成就。教師也才能知道學生的優缺點在哪些地方。

六、評量應在各種不同的情境中實施

學生態度、興趣、觀念的改變及知能的增進在團體設計、討論、報告中可予以評估。學生在戲劇、韻律及角色扮演所表現的行為也可顯示其學習增長的情形。學生在課堂內、課堂外、校內、校外都會表現出一些行為。教師應在各種不同的場合，評量學生的表現，始能判斷其學習是否產生良好的改變。

七、評量應運用各種不同的方法

教師可利用許多不同的工具及評量的技術，蒐集有關教學結果的資料。常用的評量方式與技術有：觀察、討論、辯論、面談、個案會商、個案研究寫作、研究報告、口頭報告、實驗、自傳、教師自編測驗、師生合編測驗、標準化測驗、問卷、社交距離測量、語意區分測驗、圖表、繪畫、查核法、等級量表、活動紀錄、工作樣品展示、學習日誌、日記、軼事紀錄、行為紀錄及其他紀錄、表現等。教師使用何種方式評量，要看評量的目標如何以爲定。教師不宜僅使用一種方式去評量某一目標。即使教師使用查核法、等級評量法或測驗，他也可以同時採用觀察法。混合使用各種方法比單獨採用一種方法要好得多。譬如測量游泳，如果以紙筆測驗，至多僅能測出知識而已，無法測出其技能，尚應在游泳池中，測出其速度及技巧。

八、評量是師生共同合作的過程

學校行政人員、教師、學生，甚至家長都應參與評量計畫的工作，因為他（她）們對於教學改進均有密切的關係。教師與學生共同評量有助於目標的澄清。團體評量與自我評量有賴於教師的指導及學生的合作。家長也應多參與討論其子女的學習狀況。學校行政人員及視導人員應多提供協助，並與教師合作，俾設計並採用有效的評量計畫，協助教師改進教學。

九、教師應提供學生自我評量及同儕評量的機會

透過自我評量，學生可分析自己的知識、技能、態度、行為的優、缺點與需要。當他（她）們評估個人與團體努力的結果時，他（她）們也就培養個人的責任觀念。自我評量促進自我學習。同儕評量則可提供見賢思齊的機會。

十、評量應與教學密切結合

評量應該是教學的一環。二者有如天衣無縫，密不可分。成功的教師會觀察並記載學生學習進展的情形。他（她）們會將評量的結果告知學生及家長，提供立即回饋。他（她）們也會建立教學評量的監控系統，根據評量的結果，檢討教學的利弊得失，改進教學。

綜上所述，教學評量已趨於多元化，包括評量的目標、內容、方式、情境及標準。評量的方法與技術也更趨於多元。教師究應採取何種有效的方法與技術去評量學生的學習結果，必須深思熟慮，妥善運用。

參、多元化教學評量的理論基礎

依據Gardner（1983）的研究，人類智慧

的結構是多元的。他提出多元智慧的理論（Theory of Multiple Intelligence），包括音樂的、語文的、數理邏輯的、空間的、肢體運動的、知己的及知人的智慧等七類。他的理論應用到學校，兒童可以顯現各種不同的智慧，不一定與傳統學校的課程學科相結合。傳統的評量方式常無法評量到這些能力。

Gardner的理論提供了多元化教學評量一個很好的思考支柱。從多元智慧的理論來看，兒童的學習也是多向度的。為了確切、真實評估兒童的學習成長，美國Illinois州Crow Island School即設計一套學習經驗表，內容包括：語文、科學、數學、社會學科、藝術、音樂、體育、閱讀、人際關係及人格成長等十個向度，充分顯示兒童的學習活動是多元化的，教學評量自應多元化（Herbert, 1992）。

當代心理學家Perkins（1986）對於智力的研究，亦持類似的觀點。他認為智力能否藉學習予以增進，可分為三種不同的看法：能力論（power）、心術論（tactics）及知識內容論（content）。今分述如後：

- (一)能力論者主張智力係依賴大腦的神經效能作為處理資訊的機關。Jensen（1984）認為智力間接測量了大腦的基本效能。大體上，個人的能力取決於遺傳，也就是自己原有的才能（original equipment）。
- (二)心術論者認為智力就是一種策略（Baron, 1978）。一些智能不足者或學習遲緩者表現低劣，而且沒有心機，他（她）們沒有記憶及解決問題的策略。如果教師們能教他（她）們一些策略、方法，則其表現將大為改觀。這種結果反駁了智力是由有機體所決定的

觀念，而支持智力是可以學習的說法。

(三)知識內容論者主張原則上智力係以豐富的智識為基礎。此類人士認為精通某一特定領域如數學、物理、社交技巧……等知識可作為此一領域有效思考的基礎。問題的解決有賴於豐富的知識 (Glaser, 1984)。知識豐富的人較能觸類旁通，靈機應變。

這三派的支持者都有其立論的根據。但是智力畢竟不是單純的事，而是許多影響力的組合。職是之故，Perkins (1986) 認為智力就是能力、心術及知識內容的總合，他提出了一套智力的等式：

智力 = 能力 + 心術 + 知識內容

智力的結構既是多元化，教學評量的內容、方法及場合也應多元化，不宜局限於某一狹小的領域或單一的方法與情境中，始能測出其能力，進而發展其潛能。

另一方面，教學評量也深受學習及認知心理學的影響。良好的教學評量大都建立在當前的學習及認知理論及將來學生需具備何種知能的論點上。

依據認知學派的研究，有意義的學習是反省性、建構性及自我調適性 (Herman, 1992)。學生知道某事，不僅要接受訊息，也要能解讀訊息，更要把它與其他已獲得的知識發生聯結。晚近的學習與動機統整的研究亦強調情意與後設認知技巧的重要性。這說明學習涉及認知、技能及情意領域，也包含德、智、體、群、美等五育的發展。從教育的理念言之，教師要因材施教，教學評量就應多元化，尤其在方法及標準方面，始能長善救其失，達到人盡其才的理想。

肆、多元化教學評量的新趨勢

教師教導學生，不僅要讓學生知道些什麼，這是知 (knowing) 的層次，也要讓學生會做些什麼，這是行 (doing) 的層次，更要讓學生能想些什麼，這是思 (thinking) 的層次——涉及感覺 (feeling) 及反省 (reflection) 等層面。晚近的教學評量方式更趨於多元化、人性化及彈性化。世界教育先進國家普遍採用實作評量 (performance assessment) 及個人檔案紀錄評量 (portfolios assessment) 即為明證，茲分別述之如下：

一、實作評量的背景及涵義

美國行之多年的選擇式標準化測驗，把知識技能細分成更小的部分，鮮能測出學生的思考技巧，或綜合內容及解決問題的能力。學校教師感受到一股「為測驗而教學」的壓力。學校常依測驗的結果，分成等級，未能通過測驗的學生無法取得畢業文憑。

近年來，美國教育界普遍存在一種看法，認為學生畢業應具備一些基本知能，如解決問題的能力、團體合作的精神及綜合各學科知識的能力等 (O'Neil, 1992)。二十一世紀的國民不是只會在試卷作答的人，而是要有「分析、預測及適應能力的人」。簡而言之，就是能為生活而思考的人 (Hiebert & Calfee, 1989)。實作評量乃應運而生。

實作評量旨在運用各種方式，評量各種能力及技巧，要求學生展示知識的應用，而非僅展示知識的本身 (Long & Stansbury, 1994)。教師可要求學生撰寫一篇論文、團體做科學實驗、以寫作方式提出申辯如何解答數學問題

、或保存最好作品的個人檔案紀錄。相形之下，標準化的紙筆測驗，通例只要求學生個別作答，從選擇題中選出答案，似乎不適合這些需求。

實作評量不是教學評量的一種新策略。以往善於教學的教師即經常採用觀察、實驗、寫作及實際操作等方式判斷學生進步的情形。目前許多學校採取有系統的轉變，擺脫選擇式測驗，改用實作評量的方式，作為測量教學及績效的工具。Khatti, et al. (1995) 指出：實作評量可以提供教育的範型，協助教師發展有效的教學技術，也可以提供有關學生進步的綜合訊息，包括學生的優缺點，更可引起學生的學習動機，促進學生主動學習。

實作評量與真實評量 (authentic assessment) 常交互使用，惟二者有別。依 Meyer (1992) 的研究，前者着重在學生接受測驗時的反應種類；後者則強調學生接受測驗時的反應情境，亦即在現實的生活情境中產生。

在實作評量中，學生必須完成或展示評量者所要測量的行為。通常，都有起碼的論斷程度。譬如，測量的行為是寫作，那麼學生就考寫作，不必完成有關句子及段落的選擇題，而要完成高度的推論學生寫作能力的題目，是否能從別人的文章中，找出錯誤之處 (Meyer, 1992)。

在真實評量中，學生不僅要完成或展示所要測量的行為，而且要在現實的生活情境中實施。現實生活的情境可能以學生的期望（如在教室）或成人的期望安排之。真實評量的寫作標準可能取決於學生的控制場所 (locus of

control)，亦即由學生自行決定主題、時間、步調及條件等 (Meyer, 1992)。

二、實作評量的特點

實作評量包含一系列的歷程。這些歷程具有下列四個部分 (Airasian, 1994)：

- (一)學生必須展示所教的歷程。
- (二)展示的歷程細分為較小的步驟。
- (三)展示的歷程可直接予以觀察。
- (四)按照小步驟的表現，判斷其成績。

基於上述的認知，實作評量必須符合下列四個特點 (Airasian, 1994)：

- (一)應具有明確的目的

實作評量首應確定評量的目標是什麼，通常以行為目標的方式敘寫，並且要能含蓋主要的教學目標。

- (二)辨認可觀察的實作行為

目標確定後，教師應考慮以何種行為最能展示學習的歷程及結果。這些行為必須客觀的、可觀察的及可測量的。

- (三)能提供合適的場地

可觀察的行為必須在合適的場地進行，也許在教室、禮堂、實驗室或工廠等。教師應設計合適的場地，評量學生展示的行為。

- (四)備有預擬的評分或計分標準

譬如演說的實作評量，可把演說的行為細分為五個部分：1. 眉目傳情 (making eye contact)；2. 口齒清晰宏亮 (speaking clearly and loudly)；3. 抑揚頓挫 (changing voice tone to emphasize points)；4. 呈現論點 (presenting arguments) 及 5. 總結論點 (summarizing main points)。這五種行為就成

為評量演說的標準，裁判可用來觀察並判斷演說的表現。

大體上，實作評量具有下列各項功能（Airasian, 1994），茲列表如下：

表一 實作評量與紙筆測驗的比較

實作評量	紙筆測驗
· 涉及學生把知識轉化成可觀察的表現行為或成品的能力。	· 主要涉及學生的知識及資訊的獲得。
· 設計及施測費時，但評量表可針對同一或新的學生重複施測。	· 設計費時，但可同時施測許多學生，同組學生僅能使用一次。
· 學生表現不佳，可予診斷及補救，可監控學生進步實況。	· 除論文式及開放式數學題之外，甚少提供方向，指示如何改進表現。
· 教學首重表現及過程。	· 教學着重內容知識。

資料來源：Airasian (1994), P.236.

表二 各種評量類型的比較

	客觀式測驗	論文式測驗	口頭發問	實作評量
目的	以最大的效率及信度，測驗代表性的知識。	評估思考的技巧及或知識結構的瞭解程度。	教學時，評估知識。	評估知識及瞭解化為行動的能力。
學生的反應	閱讀、評量、選擇。	組織、寫作。	口頭回答。	計劃、建構、及表達原始的反應。
主要優點	效率—在測驗時間內可測驗到許多項目。	可測量複雜的認知結果。	使評估與教學結合。	提供充分的表現技巧。
對學習的影響	過度強調回憶，鼓勵背誦記憶，如能適當出題可促進思考技巧。	鼓勵思考及寫作技巧的發展。	刺激學生參與學習，提供教師立即回饋了解教學是否有效。	強調運用知識、技巧於實際的問題情境。

資料來源：Airasian (1994), P.229.

三、實作評量的要領

教師設計實作評量的試題時，可參考Airasian（1994）的研究，茲略述如下：

(一)確認所要評量的表現或工作，親自操作一番，或自己想像一下如何操作。

教師不妨自行設想：為了完成這件工作，我必須做哪些事？我必須遵循哪些步驟？哪些

是表現或成品最重要的部分？教師也可以一邊觀察學生操作，一邊確認實作評量的重要成分。

(二)舉出表現或作品的重要部分

完成某一表現或作品的最重要屬性或具體的行為是什麼？把它們列舉出來。在教學上強調哪些行為？這些具體的行為或屬性就成為實作評量的指標。

(三)設法限制實作評量的準則（criteria），俾學生操作時，予以觀察。

評量少數的準則要比評量很多的準則容易得多。每次只評量一個準則。準則以不超過10～15個為宜。

(四)可能的話，組成教師小組，集思廣益，找出實作評量的重要行為。

譬如國小低年級語文科，各班教師都實施朗讀，那麼同年級的教師共同擬訂朗讀的準則，則可節省時間，並能建立更完整的準則。

(五)以可觀察的學生行為或作品的特性，表達符合準則的程度。

實作評量的準則應針對學生所做的或作品的特性而設計。評量的準則要具體明確。譬如：不可寫成「學生跳起來」，應寫成「學生從起跳點，向前跳了至少二碼。」要注意：實作評量是以可觀察的行為準則來衡量的。

(六)避免使用模稜兩可的字眼，以致模糊了評量準則的意義

評量準則切忌使用形容詞轉變而成的副詞如正確地、合適地、優雅地等字，因為觀察者的解讀可能因時、因人而異。譬如「正確地說出」，可能解釋為「唸得很正確」，或「室內都可聽得到」，或「內容很清楚」。

(t)依序列出準則，俾便觀察

此一方式，教師可節省時間並可集中注意於實作上。

(v)在建構自己的準則之前，先檢討現行使用的實作評量準則。

現行的準則既經多人深思熟慮研究出來，必有其道理。有些準則可以繼續採用，毋須大費周章，重新設計。

四、實作評量的例證——以數學為例

美國數學教師委員會（the National Council of Teachers of Mathematics, NCTM）於1980年確定數學教育的重點在於解題，認為數學是一種解題（Mathematics as Problem Solving），數學是一種推理（Mathematics as Reasoning）及數學是一種溝通（Mathematics as Communication）。這顯示美國的數學教育已由規則性、被動性的解題轉趨於主動性、積極性的學習，強調推理與溝通的解題過程（Szetela & Nicol, 1992）。

問題解決是一種過程。要成功地解決問題有賴於後設認知的過程（metacognitive processes）。典型的解題過程如下（Szetela & Nicol, 1992）：

- (一)掌握問題的情境。
- (二)考慮合適的策略。
- (三)選擇並採用最有可能的解題策略。
- (四)監控並執行與問題有關的條件及目的。
- (五)獲得並溝通所要達成的目的。
- (六)評估答案的允當性及合理性。
- (七)如果答案錯誤或不適當，重新澄清問題之所

在，改用新方法或策略，找出程序上或觀念上錯誤之所在。

這些後設認知的過程殊難評估，但評估可藉創造問題的情境，促進學生思考與溝通。學生解題時不加說明就寫出答案，或只寫答案沒有演算過程，均無法顯示解題者的思考本質。因此，評量解題，只看答案是否正確，不看解題的過程是不恰當的。如果教師能設計一套方法，引導學生的解題過程，就可更有效評量學生的能力。這種評量可測出學生思考的品質。

Szetela & Nicol（1992）指出：教師可提供一些情境，讓學生思考如何解題。其方式如下：

- (一)呈現一則包含各種事實及條件的問題，要學生寫出合適的問題，解答完整的問題，並寫出其對答案允當性的感受。
- (二)呈現一則只提供部分答案的問題，要學生完成作答。
- (三)呈現一則包含不相干事實的問題，要學生評論該問題的品質或修改問題，除去不一致性。
- (四)呈現一則沒有數字的問題，要學生提供合適的數字，估量答案並解決問題。

教師可使用下列例子，評量學生思考、解題反應的品質：

籃子內有10粒柳橙及橘子。柳橙每粒5元，橘子每粒10元，總共70元。請問籃子裏面，有幾粒柳橙？

假設小華是這樣解題的：

$10 \times 5 = 50$	$8 \times 5 = 40$
$2 \times 10 = 20$	$3 \times 10 = 30$
_____	$4 \times 10 = 40$

70 $6 \times 5 = 30$

答：籃子裏共有30粒柳橙。

請看一看小華的解題方式，再回答下列問題：

1. 小華的解題方式好嗎？

答：好。

請說明解題方式好或壞的原因。

答：因為它告訴你可能的答案，但她沒仔細看清題目。

2. 小華得到正確的答案嗎？

答：沒有。

請解釋她有或沒有的原因。

答：因為籃子裏只有10粒水果。

3. 正確的答案是多少？

答：籃子裏面有6粒柳橙。

五、個人檔案紀錄評量的產生及涵義

美國Rockefeller的研究計畫之一推進計畫（PROPEL）結合了Pittsburgh公立學校教育測驗服務中心及Harvard教育研究所的零計畫（Project Zero），共同進行研究，結果顯示：以嚴謹而不曲解的方式，評估藝術及人文思考歷程的特徵是可能的。他們認為教師可設計學生學習評量的方式並發現學習成長的方式。學生可成為充分了解並考慮周詳的評估者（Wolf, 1989）。

為實現這些目標，PROPEL的研究群邀請藝術家、音樂家及作家，說明他（她）們如何取樣作品並評判其一生的作品。研究結果發現：思想家及發明家大都保存其一生縱式系列的理念、手稿及問題等。他（她）們運用這些檔

案，作為日後創作的可能。研究人員終於發展成視覺藝術、音樂及寫作的檔案紀錄評估系統（System of portfolio assessment）。

個人檔案紀錄評量係指多角度、多向度的評量歷程，包括一切有關的人員、事項及資訊的蒐集等。它具有下列各項功能：

（一）激勵學習的興趣

Knight是美國一所中間學校的數學教師。她要求班上的學生寫出個人檔案紀錄應包括哪些東西？怎樣才能顯示學生們的努力及數學成績？哪些活動最有意義？結果發現學生的個人檔案資料包羅萬象，包括：每天的日記、個人的財務計畫、彩券計畫、按比例繪圖、最好的測驗、最壞的測驗、本週的難題、課堂上的筆記及家庭作業等。然後她要求學生從個人檔案紀錄中，找出五項最能代表數學知能及努力成果的東西。隨即全班討論良好檔案紀錄的格式及版式。最後，大家決定：版式要工整，打字，要有封面及目錄。此外，每一部分的開頭，都要寫出個人的看法，表示其重要性。待全班同學的個人檔案紀錄蒐集齊全之後，教師把它們發給不同的同學互相觀摩、互相學習、互相評分。教師發現：班上同學都興高采烈，學習情緒高昂。個人檔案紀錄顯然可以激勵學習的動機（Knight, 1992）。

（二）建立學生的自信心

Frazier是美國Oregon州一所小學四年級的語文教師。她採用個人檔案紀錄，在學年開始時，即告訴學生們：她將記錄學生在個人檔案上所寫的一切資料。學年結束，教師要求學生寫一封信給別班或校外的人，說明其個人檔案裏學習到哪些東西。學年開始不久，有一位學生這樣寫着：「我想，我的故事是一個好的

70 $6 \times 5 = 30$

答：籃子裏共有30粒柳橙。

請看一看小華的解題方式，再回答下列問題：

1. 小華的解題方式好嗎？

答：好。

請說明解題方式好或壞的原因。

答：因為它告訴你可能的答案，但她沒仔細看清題目。

2. 小華得到正確的答案嗎？

答：沒有。

請解釋她有或沒有的原因。

答：因為籃子裏只有10粒水果。

3. 正確的答案是多少？

答：籃子裏面有6粒柳橙。

五、個人檔案紀錄評量的產生及涵義

美國Rockefeller的研究計畫之一推進計畫（PROPEL）結合了Pittsburgh公立學校教育測驗服務中心及Harvard教育研究所的零計畫（Project Zero），共同進行研究，結果顯示：以嚴謹而不曲解的方式，評估藝術及人文思考歷程的特徵是可能的。他們認為教師可設計學生學習評量的方式並發現學習成長的方式。學生可成為充分了解並考慮周詳的評估者（Wolf, 1989）。

為實現這些目標，PROPEL的研究群邀請藝術家、音樂家及作家，說明他（她）們如何取樣作品並評判其一生的作品。研究結果發現：思想家及發明家大都保存其一生縱式系列的理念、手稿及問題等。他（她）們運用這些檔

案，作為日後創作的可能。研究人員終於發展成視覺藝術、音樂及寫作的檔案紀錄評估系統（System of portfolio assessment）。

個人檔案紀錄評量係指多角度、多向度的評量歷程，包括一切有關的人員、事項及資訊的蒐集等。它具有下列各項功能：

（一）激勵學習的興趣

Knight是美國一所中間學校的數學教師。她要求班上的學生寫出個人檔案紀錄應包括哪些東西？怎樣才能顯示學生們的努力及數學成績？哪些活動最有意義？結果發現學生的個人檔案資料包羅萬象，包括：每天的日記、個人的財務計畫、彩券計畫、按比例繪圖、最好的測驗、最壞的測驗、本週的難題、課堂上的筆記及家庭作業等。然後她要求學生從個人檔案紀錄中，找出五項最能代表數學知能及努力成果的東西。隨即全班討論良好檔案紀錄的格式及版式。最後，大家決定：版式要工整，打字，要有封面及目錄。此外，每一部分的開頭，都要寫出個人的看法，表示其重要性。待全班同學的個人檔案紀錄蒐集齊全之後，教師把它們發給不同的同學互相觀摩、互相學習、互相評分。教師發現：班上同學都興高采烈，學習情緒高昂。個人檔案紀錄顯然可以激勵學習的動機（Knight, 1992）。

（二）建立學生的自信心

Frazier是美國Oregon州一所小學四年級的語文教師。她採用個人檔案紀錄，在學年開始時，即告訴學生們：她將記錄學生在個人檔案上所寫的一切資料。學年結束，教師要求學生寫一封信給別班或校外的人，說明其個人檔案裏學習到哪些東西。學年開始不久，有一位學生這樣寫着：「我想，我的故事是一個好的

參考文獻

- Airasian, P. W. (1994) . Classroom assessment, 2nd Edition, New York: McGraw-Hill, Inc.
- Baron, J. (1978) . Intelligence and general strategies. In Strategies in information processing, edited by G. Underwood, London: Academic Press.
- deBono, E. (1985) . The CoRT thinking program. In J. W. Segal, S. F. Chipman & R. Glaser (Eds.) , Thinking and learning skills, Vol. 1, Relating instruction to research. Hillsdale: Erlbaum.
- Frazier, D. M. & Paulson, F. L. (1992) . How portfolios motivate reluctant writers. Educational Leadership, 49(8), 61-65.
- Gardner, H. (1983) . Frames of minds : The theory of multiple intelligences. New York: Basic Books.
- Glaser, R. (1984) . Education and thinking : The role of knowledge. American Psychologist, 39, 93-104.
- Hansen, J. (1992) . Literacy portfolios: Helping students know themselves. Educational Leadership, 49(8), 66-68.
- Herman, J. L. (1992) . What research tells us about good assessment. Educational Leadership, 44(8), 74-78.
- Hiebert, E. H. & Calfee, R. C. (1989) . Advancing academic literacy through teachers' assessments, Educational Leadership, 46(7), 50-54.
- Jensen, A. R. (1984) . Test validity: Generality versus the specificity doctrine. Journal of Social and Biological Structures, 7, 93-118.
- Khatti, N.; Kane, M. B.; & Reeve, A. L. (1995) . How performance assessments affect teaching and learning. Educational Leadership, 53(3), 80-83.
- Knight, P. (1992) . How I use portfolios in mathematics. Educational Leadership, 49(8), 71-72.
- Long, C. & Stansbury, K. (1994) . Performance assessments for beginning teachers, Phi Delta Kappan, 76, 318-322.
- Meyer, C. A. (1992) . What's the difference between authentic and performance assessment? Educational Leadership, 49(8), 39-41.
- Michaelis, J. U. (1980) . Social studies for children: A guide to basic instruction. 7th ed. Englewood Cliffs, New Jersey: Prentice-Hall, Inc. 448
- Perkins, D. N. (1986) . Thinking frames. Educational Leadership, 43, 4-10.
- O'Neil, J. (1992) . Putting performance assessment to the test. Educational Leadership, 49(8), 14-19.
- Szetela, W. & Nicol, C. (1992) . Evaluating problem solving in mathematics. Educational Leadership, 49(8), 42-45.
- Wolf, D. P. (1989) . Portfolio assessment: Sampling student work. Educational Leadership, 46(7), 35-39.