

我國工業教育教學評量之檢討

鍾瑞國

壹、前言

課堂教學的主要目的在幫助學生達到預期的教學目標（蕭錫錡，民七十八）。根據葛雷斯（R. Glaser）之基本教學模式，教學評量之功能包括：一、了解學生的起點行為；二、建立確實可行的教學目標；三、確定教學目標達到的程度；以及四、改進教材教法、增進教學效果。職是之故，如何透過教學評量以了解教學之成效，實有其積極意義。

工業教育為我國整體教育的一環，旨在培養健全之工業基層技術人才，強調專業技能之學習，務求學習後具有一技之長，奠定學生創造、適應變遷及自我發展之能力。惟長久以來許多教育專家都批評工科之教學評量偏重認知領域（郭生玉，民七十六；何英奇，民七十八；陳明照，民七十二），或對技能評量不甚熟悉，或忽略情意之探討，終無法

有效了解實際之教學成效。因此，本文擬從教學評量的意義、功能、種類、方法以及檢討當前工業教育教學評量之得失，提供教育工作者之參考。

貳、教學評量的意義與功能

評量與評鑑一詞常被混淆，實有加以探討之必要（何英奇，民七十八；蕭錫錡，民八十；李隆盛，民八十二；Wolansky, 1985; Finch & Crunkilton, 1989）。所謂評量（evaluation）乃是經由測量（如測驗、評定量表、檢核表、問卷晤談、觀察等）而得量化的（quantitative）資料，再根據這些可靠的數據資料，參照合理的衡量標準，加以比較分析和綜合研判的系列過程。至於評鑑（其英文與評量同，只是範圍更大）係指為特定人士有系統地蒐集某些人事物之活動、特性和結果方面的資訊，以供此等人士能對這些人事物增進瞭解、改善效能和作成決定的程序。

因此，教學評量係指教師透過測量的方法，將教學過程（含教學前、中、後）所收集到的資料加以綜合統整，進行價值判斷，並據以作成決定的過程。因此，教學評量在教學前、教學中及教學後三階段中各有其不同之實施內容要項。教學前評量注重目標的擬定、學生特性的掌握及前測的實施；教學中評量注重教師自編測驗以診斷分析學生學習的強弱處、心理錯誤轉機所在，以及學習過程中能力的變化；教學後評量則注重總結性的效標參照評量，並作為教學計畫的回饋資料。

教學評量的功能很多，歸納起來，可臚列如下（黃國彥，民七十一；張瑞村，民七十五；溫世頌，民七十五；李大偉，民七十五；Wolansky, 1985; Madus, Scriuen & Stufflebean, 1989）：

- 一、對教師而言，可幫助教師明瞭學生的學習起點行為；瞭解教學的效率；診斷學生的學習，實施增廣或補救教學，以及作為教師改進教學內容與教學方法的依據。
- 二、對學生而言，可幫助其瞭解學習的目標，激發學習動機，增進自我的了解；改變學習方法與態度；使學生瞭解自己學習的進程與技能的精熟度，以及培養學生良好的工作習慣與態度。
- 三、對學校行政管理而言，可做為評量教師教學成果的依據；改進課程結構與教材內容的參考，改進技術教學的

參考，以及為學校提供一種教育品質管制的辦法。

四、對家長而言，使家長明瞭子女學習情形，以及做為升學或就業之參考。

叁、教學評量的基本原則

工業教育教學評量之有效發揮上述功能，必須於實施時，把握下述幾項基本原則（簡茂發，民七十六；張瑞村，民七十五）：

- 1 符合性原則（Congruent）：教學評量最重要的是依據教學目標，擬定具體的標準以評量教學成效是否符合既定之目標。
- 2 廣泛性原則（Comprehensive）：教學評量應以學生身心成熟度為依據，且評量的範圍與項目宜力求廣泛，並應配合運用多種評量方式。
- 3 繼續性原則（Continuous）：教學評量為一動態的歷程，而非靜態的實體，更不是教學歷程的結束，而是持續互動，回饋修正之歷程。
- 4 合作性原則（Cooperative）：教學評量工作宜由教師、學校當局、學生家長合作協力完成。
- 5 描述性原則（Descriptive）：評量結果宜妥善運用，並採敘述性方式，將學生學習歷程加以分析和診斷。
- 6 統整性原則（Integrative）：以總結性評量評定學生的學習成果，以整體瞭解學生是否達成了認知、技能、情意等目標。
- 7 個別性原則（Individual difference）：評量須符合學生的個別差異。
- 8 一致性原則（Consistency）：針對技能操作表演測

驗方式進行評量，測驗情境與使用工具、儀器、機器設備的性能必須要求一致，以保證測驗之正確性。9 安全性原則 (Safety)：針對實際操作技能評量時，必須考驗受試者在評量過程中之安全問題，以防意外傷害。

肆、教學評量的類型

工業教育之教學評量方式依用途或功能與參考點之不同，而有如下之分類（李咏吟，民七十；何英奇，民七八；簡茂發，民七十六；廖鳳池等，民八十）：

一、依用途或功能而分

(一)安置性評量 (Replacement Evaluation)

安置性評量應用於教學前，其關心者為學生之學習起始行為，亦即學生在教學前所需具備之知能為何？學生精熟預定教學內容的程度如何？學生的興趣、工作習慣和人格特質適合某種教學模式的程度如何？因此，安置性評量的目的旨在決定最有利於學生之教學順序與教學模式。

(二)形成性評量 (Formative Evaluation)

形成性評量應用於教學過程中，其所關心者為學生的那些學習已令人滿意？那些尚需要幫助？因此，形成性評量的目的旨在不斷提供教師和學生回饋，使他們知道學習的成功與失敗，有助於學生學習錯誤之修正，以及教師調整教學和

實施團體與個別的補救教學措施。

(三)診斷性評量 (Diagnostic Evaluation)

診斷性評量應用於教學進行中，或在教學結束時，其所關心者為學生持續學習困難之原因與解決之道為何？因此，這種評量的主要目的在於確定學生的學習困難和原因後，提出學生補救教學的措施。由文獻歸納知，常見低學業成就原因之診斷約有後述數種：學習活動失敗；基礎學力不足；缺乏正確的學習方法、態度與習慣；性格上的問題；智能缺陷；學習興趣與欲望的喪失；對學校、班級的不適應；對教師持消極態度；親子關係不良；以及同儕關係不良等。總之，形成性評量在提供一般性學習問題的處理，而診斷性評量則針對嚴重學習問題。

(四)總結性評量 (Summative Evaluation)

總結性評量應用於教學的課程或單元結束時，其所關心者為教學目標達成的程度？學生精熟預期學習結果的程度？因此，這種評量的目的較偏重評定學生的成績等級，與認定學生精熟的程度，而不在發現學習困難和改進教學。通常學校定期舉行的週考、月考和期考，均可視為此種評量（郭生玉，民七十四）。

二、依參考點而分

評量因參考點 (reference points) 的不同，通常可分

為效標參照評量 (Criterion-referenced evaluation) 和常模參照評量 (Norm-referenced evaluation) 兩種 (蕭錫錡, 民八十)。常模參照評量係以學生的學習表現與某一特定的參考團體相比。亦即, 此種評量係把測量當作比較學生各種特質的工具, 它並不在乎學生實際上到底懂多少或能做多好 (李大偉, 民七十五)。效標參照評量則是以教學前決定的絕對性標準為評量學生學習行為的依據。換言之, 教師要求學生在學習後達到所訂的標準, 否則給予學生更多的時間學習或補救教學, 因此, 此種評量係把測量用來瞭解學生的成就是否達到某些既定的標準, 而不是在比較某一學生與其他學生學習成就上之異同。效標參照評量與前述常模參照評量之特性比較可如表一所示 (蕭錫錡, 民八十)。

由上述之分析與比較, 吾人可得知工業教育之教學評量宜以效標參照為參考點。事實上, 從單位行業時代之教學策略到能力本位教育, 以迄群集教育, 技能及相關專業科目之教學, 實為重點所在, 整個教學過程無不依著既定之教學目標 (特定標準) 而不斷測量與回饋, 而學生畢業時的技術能力要求 (丙級技能檢定或相當丙級技術士之技能測量通過) 更是公認的教學品管目標之一。職是之故, 工業教育之教學評量宜以效標參照為主, 在教學前、中、後適時發揮評量的功能, 不僅能提升整體教學效果, 且能促進師生之互動、回饋, 達成既定教學目標的要求。

表一 常模參照與效標參照評量一般特性比較

常 模 參 照 評 量	效 標 參 照 評 量
1. 常模參照評量的主要功能在確定學生在一個常模團體中的相對地位。	1. 效標參照評量的主要功能在評估學生是否業已熟練特定的效標或行為表現的標準。
2. 在編製常模參照的評量時, 可詳述一般概念的結果或精確的目標。	2. 在編製效標參照的評量當中, 詳述完整的教學目標。
3. 在使用常模參照評量時, 熟練的效標通常不加詳述。	3. 為了用於效標參照的評量, 熟練的效標須加詳述 (例如, 教學目標)。
4. 常模參照評量的測驗題目, 其編製的目的在鑑別學生。	4. 效標參照評量的測驗題目, 其編製的目的在測量事先確定的熟練程度。
5. 分數的變異性質是合意的, 有助於富有意義的解釋。	5. 變異性質是無關的; 這不是滿意的效標參照評量的必要條件。
6. 常模參照評量測驗的結果, 可轉換為傳統的評分制度 (甲、乙、丙、丁、戊等)。	6. 效標參照評量測驗的結果, 表示一種二元的制度 (即滿意—不滿意, 及格—失敗)。然而, 效標參照評量的測驗結果 (依照一套特別編製的規則), 亦可改換為傳統的評分制度。

伍、教學評量之方法

工業教育之教學評量爲一目標導向之評量，唯有明確的教學目標，教學活動才有清楚的指引，教學評量也才有依據。因此，評量方法也要以教學目標爲依歸，一般而言，教學目標可分認知、情意與技能三方面，不同的領域之評量方法所適用的工具不一，本文擬先分別探討各領域之評量方法，再做綜合分析。

一、認知評量的方法

教師常用的認知成就評量的命題方法通常有二種，即主觀式命題與客觀式命題，前者包括口試與問答題，後者包括簡答題、填充題、是非題、選擇題，以及配合題。上述各種方法都有其優點與限制，因此，教師選用評量方法時，應根據評量設計與評量情境，經過專業判斷加以選擇，期能正確而有效地評量學生的成就。下述爲主觀式命題與客觀式命題的適用範圍以及編製時之注意事項（康自立，民七十八；李大偉，民七十五；簡茂發，民七十六；Wolansky, 1985; Crocker & Algina, 1986; Sax, 1980; Gronlund & Linn, 1990）：

（一）問答題的適用範圍包括：1 比較二者或二者以上的事物；2 贊成或反對某些問題；3 對現象成因與效果之探

討；4 將知識或原理應用於新情境；5 在特定領域中資訊的結論；6 列舉說明規則、原理、概念、情境等；7 對先前學習經驗的推論；8 判斷與批判某些事物之正確性或妥切性，以及9 鑑別事物價值、事物手續與理想。而問答題編製時之注意事項如下述：1 除了客觀式命題無法評量時才使用問答題；2 題意要明確，敘述文句要清晰明白；3 儘量給學生有充分的作答時間；4 避免使用選題作答方式。

（二）口試與問答題一樣用以評量學生較高層次、較複雜之學習結果，對回憶、組織、綜合及評鑑之能力的評量有較佳之效用。此外，它比問答題更適用於1 可彌補文筆不順之困擾；2 爲手足殘障或弱視目盲者最佳之評量方式。至於口試之最大限制在於它的不客觀，常因主試者與受試者互相認識，而影響評量的公正性，故設計命題時應注意下述事項：1 口試問題應限於一個主要的思想或主題；2 問題必須主題明確、容易溝通；3 必須有一定範圍，避免過於空泛；4 應有計畫進行。

（三）填充題與簡答題常用以評量學生對事實、原理、及操作步驟等知識之了解，編製時應注意下述事項：1 應使答案簡短、明確，而且只有一個絕對的答案；2 避免從書本或其他文獻斷章取義或從中截取片段支語爲試題；3 避免題中有多個空格，以至題意不清；4 避免暗示作答；5 儘可能採用「完全敘述句」之形式敘寫。

(四)是非題常用以要求學生判斷定義、定理、名詞或事實現象的真偽，比起其他評量方法，命題較容易，省時省事，編製命題應注意：1 應使答案對或錯的題目各佔一半；2 避免直接摘錄課文；3 避免使用否定敘述；4 避免學生猜答之做法：①答案不要有規則性；②使用「答錯倒扣」及「將敘述為錯的題目加以修改為正確時才得分」的計分方法；③使用「題組」型態，增加學生的整合能力。

(五)選擇題常用以評量學生之理解、辨別、應用或解決問題的能力，比是非題具較高的鑑別力；編製時應注意下述事項：1 命題要符合實際之生活情境，並使經驗與學習互相驗證；2 儘量使用圖解命題；3 每題之選項數目宜一致，至少要有四項為佳；4 如選項為一串數字，應依照次序排列，且這些數字要互斥，不可重複；5 避免答案具有規律性；6 儘量少用否定敘述。

(六)配合題用以評量學生對事物或事件的相關、因果的瞭解，編製時分兩部分，一為命題，另一為配合項或反應項，注意事項如下：1 應使配合項目數多於命題項目數；2 適當的命題與配合題至少應有五項以上，但不超過十二項；3 命題與配合題應安排在同一頁上；4 作答說明應明確，例如要說明清楚配合項目可選用一次或重複選用。

二、情意評量的方法

工業教育之情意教學目標在於注重人格修養、文化陶冶、培養學生敬業、負責、勤奮、合作等職業道德。具體而言，情意教學旨在提升學生對於技職工作的興趣，使其學習上能正面肯定自我並樂於協助群體，且能高興的與他人合作，進而培養學生對工業產品的鑑賞力（林建仲、方榮爵，民七十八）。情意教學之層次，依據布魯姆（Bloom）的看法，可依序分為：接受、反應、價值評定、價值概念之組織與建立、以及品格形成。這些層次的發展與培育必須透過長期的陶冶，教之以道德規範、操作規範、行為規範，使學生在施展其行業技巧時能注意到或意識到應該遵法的行職業規範。因此，在情意領域中，學習內容的界定、教學方法的選用、以及測量學習成就，均較認知領域來得複雜。

職是之故，林建仲與方榮爵（民七十八）建議參考之情意教學目標如下：1 遵守各種安全規則，工場規則；2 主動保持工作環境之清潔，並隨時進行機器設備之維護；3 節省材料、愛惜公物；4 尊重別人的權益；5 養成有組織、有計畫的工作習慣；6 能主動進行工作；7 保持自信、自尊、和成就感；8 隨時保持儀容整潔；9 謹守工作時間上的一切要求；10 重視技術品質、不偷工減料。而常用以評量之方法則如下所述（李大偉，民七十五）：

(一)直接觀察法（Direct Observation）：本法必須在實際的環境中實施，通常著重於學生對某事或某物之反應情

形。實施時，應注意觀察對象不能太少，常提供必要的刺激，以及避免「霍桑效應」。因此，為求觀察結果之可靠性，吾人常藉觀察檢核表（observer checklist）、數字評量表（numerical scales），以及圖表評定量表（graphic-rating scales）三種測量工具來實施，其主要特性如下：

1 觀察檢核表：首先在表格上詳列某些行為或活動，而後由主試者（可能是老師、同學、或自己）根據觀察某個人在某段時間內的表現，在表上劃記。實務例子如：實習操作習慣、安全習慣等。

2 數字評定量表：由評審人員就學生的行為表現，給予一些特定的數字（或分數），以代表其好壞、高低、滿意不滿意、或同意不同意的等級。其實例有：學習動機量表、各種態度量表等。

3 圖表評定量表：此量表除了如上表使用數字外，同時也使用一些敘述，來描述不同數字所代表的意義。常用者如同儕相處情形量表、學生工作特質量表等。

(二)面談或面試法（Affective Interviews）：常用以測量被詢問者對某些事物的情意反應，有時也可深入探討行為發生的原因。一般實施面談或面試時，可分結構式面談法（structured interviews）與非結構式面談法（unstructured interviews）兩種。

1 結構式面談法：先將欲面談之內容，製作成各項可選

擇性問題，再於面談中勾選適當選項。

2 非結構式面談法：不必事先擬好表格，主要採一種近似於漫談之方式，由訪談者針對事先擬好之幾個主題來引起話題、引發談話。實施應注意儘量使面談氣氛友善、輕鬆、愉快，才能暢所欲言；其次，面談者要扮演良好的傾聽者角色，並儘可能態度保持中立；另外，面談者須細心地注意受訪者之態度、手勢、語調等細部反應。

(三)問卷或調查表法（Questionnaires）：為最常用之評量方法。主要優點是省時、省力、易實施、不具名、客觀性高。常用之工具有自我檢核表、選擇題、填充題、評定題（如來克特量表，Likert Scales；桑斯通量表，Thunstone Items；差異性語意量表，Semantic-differential Items）。

(四)投射法（Projective Techniques）：係依據羅夏克墨漬測驗（Rorschach Ink-Blot Test）的原理（亦即受測者在測驗情境中對兩邊對稱但形象模糊的圖片的刺激所引發的反應，會投射出其人格組織的全貌）而編製的測驗。應用此種投射式人格測量之技術可概分為1聯想技術（association techniques）：要求受試者以因測驗刺激所引發的第一個字詞或意象反應，墨漬測驗和字詞聯想測驗即屬此類；2構造技術（construction techniques）：受試者須針對某種情況或圖片建構出一個故事，主題統覺測驗為其例

子；3 完成技術（completion techniques）：受測者須將測驗刺激之未完成語句或其他作業加以完成，常見之語句完成測驗屬此類；4 表達技術（expressive techniques）：包括繪畫測驗和遊戲技術等，由受測者實地去進行一項表達性活動（葛樹人，民八十）。

（五）暗地測量法（Unobtrusive Measures）：係在受測者不知道被觀察之情況下進行，或利用間接資料，對受測者的情意行為加以評定。應用實例如教師可從工場意外事件發生的頻率看出學生對工場安全的關心程度；或教師可從學生實習時在工作檯上的刻字、在牆上寫字、工具的損壞、失竊、打架等瞭解學生的品行（李大偉，民七十五）。

三、技能評量的方法

工業教育之技能領域教學目標在於培養健全的工業基層技術人力，傳授各類科之實用技能（包括擔任機械之製圖、製造、操作、管制、檢驗、維護等工作之技能）（教育部，民七十五）。因此，技能教學之內容就包括了相關知識與技能操作；所以，評量內容自然也要學科與術科並重。有關相關知識之評量方法與前述認知領域的評量方法一致，惟測驗內容應著重與技能相關性高的知識，例如機械製圖科目中的繪正投影視圖能力，其相關知識應包括正投影原理、視圖排列的方法、線法、字法、尺度標註等知識。這種相關知識可

利用一般認知領域之評量方法，此處不再贅述。而術科之技能部分的評量方法，常用者有作品評量、實物測驗、以及操作表演測驗（康自立，民七十八；張瑞村，民七十五；李大偉，民七十五）：

（一）作品評量：係要求學生運用知識、經驗，設計並製造一件完成的作品；教師依成品的功能及精確度等特性予以評定成績。

（二）實物測驗：係讓學生識別實際工具、材料、或樣品，寫出或說出其名稱，回答關於實物之有關問題。

（三）操作表演測驗：係讓學生在實際或控制的情境下，完成一項或多項的技能操作，此為最適當、最理想之測驗方式。

此三種評量方法之適用性、優點與限制、以及編製時之注意事項，如表二所示（張瑞村，民七十五）：

此外，當教師在進行技能測量時，無論使用何種方法，吾人皆可利用有關之評量表，以協助將學生之技能成就予以量化（李大偉，民七十五），這些評量表包括：1 評等表（Ranking）：為目前使用較多之量表，係將學生作品依其好壞排列成等級，因此，這是常模參照之評量。2 成品對照表（Product Scales）：係將學生作品拿來與預先建立之成品對照表比對一番，再看此作品與對照表上那一個等級的作品相近似，再據以評定等級或分數。此法同時具有常模參

表二 常用技能評量方式之比較

評量方式	用 處 與 優 點	缺 點	編 製 時 應 注 意 事 項
作品評量	1.可測定學生計劃工作之能力。 2.可測定統合應用知識經驗之能力。 3.可引出學生之創造潛能。 4.可間接測出學生之藝術涵養。 5.學生可以分工合作，協力完成作品，培養合作精神。 6.可培養學生實驗研究的精神。	1.如不觀察學生完成作品的過程，則無法確定學生之真正能力。 2.管理不易，無法顧及所有學生製作成品的過程。 3.二人以上共同完成一件成品，難以評定個人成績。	1.成品製作宜適當的予以分組。 2.給予成品設計的架構，由學生提出作品藍圖。 3.教師與學生共同研究作品藍圖。 4.指導學生進行的程序與方法。 5.觀察學生設計與製作成品的過程，並予記錄。 6.編寫評分要項，依此評分。 7.要學生說明並操作其完成之作品。
實物測驗	1.可測定識別工具、材料、零件之種類、大小和等級，以及選擇的能力。 2.可測定識別在工場所使用之各種接合、形狀、割切、及指出其名稱之能力。 3.可測定對工具之調整與機械裝配之瞭解和知識。 4.測定執行操作中特定步驟之能力。 5.判別學生優劣之能力高。	1.若測驗時間太短，可能造成受試者的情緒緊張。 2.測驗情境不一致。 3.需要相當長之時間。 4.管理不易。	1.各測驗問題要短，敘述簡潔明確。 2.答案紙應附註作答方法。 3.各試題回答時間要大致相等。 4.廣泛利用機械、設備、工具、材料等樣品作為題材。 5.準備評分用標準答案。
操作表演測驗	1.用表演測驗測定學生執行操作之能力，比其他方法客觀、可靠而有效。 2.它供給分析和測驗學生精通技術各要素之程度。 3.表演測驗可測出教學直接應用之結果。	1.需要相當長之時間。 2.編製不易。 3.管理困難。 4.常造成受試者的緊張情緒。 5.無法讓全部學生在同一條件下(如時間、環境)工作。	1.選擇重要操作編入表演測驗。 2.詳加分析各操作。 3.選擇或設計適當工作。 4.準備工作之工作圖和明細表。 5.列舉所有測驗目的，編製檢查表。 6.準備學生應遵守之方法和指示。 7.準備管理測驗之方法。 8.選擇或作檢查學生完成工作之裝置。

資料來源：張瑞村(民75)工業教育技術教學評量方式之比較研究。教育學院學報，11期，頁333。

照與效標參照評量之特性。3 檢核表、數字評定量表、圖表評定量表之使用特性已見前述情意評量文中，此處不再贅述。

綜合認知、情意、技能三個領域之評量方法、性質、以及適用性，可得表三之比較。

陸、當前教學評量問題之檢討與改進

經由前述工業教育之教學評量文獻與理論的探討，包括評量的源起、意義、功能、分類、方法，倘能再針對實際教學評量之缺失，加以改進，必能提升教學的成效，達至既定之教學目標。

一、教學評量存在之問題

國內學者（郭生玉，民七十六；陳明照，民七十二；李克明，民七十二；吳天元，民七十七；陳海雄，民七十三；邱錦昌，民七十；康自立，民七十八；李大偉，民七十五）認為當前存在之教學評量問題有：評量目標、評量方式、評量時機、評量結果的解釋、以及命題觀念等問題。

（一）評量目標的問題

工業教育之教學評量為一目標導向之評量，所依據者應為教育部頒訂之教學目標。可是目前的教學已漸傾向升學導向，於是考試領導教學，學校所實施的教學評量，幾乎是

表三 各種評量方法之適用性與特質比較

項 目 評 量 方 法	評 量 用 途 或 功 能				特 質	
	安置 評量	形成性 評量	診斷 評量	總結性 評量	實際情境 之掌握 是否深入	客 觀 性
標準化測驗	○			○	否	良
教師自編測驗	○	○	○	○	否	良
問卷法	○	△	○	○		良
口試	○	○	○	○		
觀察評定法		○	○	○	是	差
作品評定法				○	是	差
實作表演法	△	○	○	○	是	差

註：1. “○”表適用；“△”表尚適用。

2. 資料來源：修改自何英奇（民78）評量的基本概念，載於康自立等，專科學校教師教學評量手冊，頁18。

爲升學而評量，至少顧及學科的教學目標或單元教的教學目標，至於技能教學也被忽略，遑論情意評量了。其次爲求應付升學，學校的評量往往偏向艱深，使學生經驗太多挫折感，導致厭惡學習或逃避學習；而評量偏重機械式的記憶，也扼殺了學生創造思考的能力，使學生成爲「背多分」的工具。

(二) 評量方式的問題

沒有一種評量方法能夠涵蓋學生所有的重要學習結果，如前所述，各種評量方法有其適用範圍，因此，教學評量之方式須視教學目標而選用最適當者行之。可惜的是，目前學校的教學評量在認知領域中，幾乎傾向採用紙筆測驗一種，而技能領域中，也多數採用成品評分法。前者導致學生只知死啃書，應付考試，而不能利用各種學習方法，主動求知或探索解決問題的方法；後者由於不重視過程評量，學生只要繳交作品即可，易產生他人代勞或不安全、不正確之操作，實質上，學生之學習多未能達到教學目標之要求。至於情意領域之評量，幾乎很少有教師積極從事，如何能使學生具有行職業正確之工作規範與職業道德呢？

此外，有關評量的信度與效度的要求，也付之闕如。評量最要緊的是講求信度與效度，前者係指測驗分數的一致性（consistency）或穩定性（stability），亦即相同的個人在不同時間，以相同的測驗測量，或以複本測驗測量，或在

不同情境下測量，所得結果的一致性；後者乃是一個測驗最重要的特徵，比信度更受重視，亦即高效度代表高信度，但反之不然，因此，效度是指一個測驗在使用目的上的有效性，亦即測驗之效度愈高，表示它愈能測量所欲測量之特質（郭生玉，民七十四）。不幸的是，目前學校教師不論採用那種評量方法，都很少注意信度與效度的問題，特別是教師自編測驗多取材市場參考書籍，也很少對測驗是否能測量出學生的水準加以探討。另外，教師對命題本身之難易度及鑑別度，同樣也甚少關懷，一方面缺乏評量知識，另一方面缺乏動機，結果也就無法達成評量的目標了。

(三) 評量時機的問題

工業教育之教學評量多數以教學後實施總結性爲主，雖然於能力本位教育時代極力推行教學前的診斷評量（用以決定學生在學前之能力以幫助學生選擇他的學習活動）、以及教學中的過程評量（用以評量每個單元是否達成預定的能力水準，用以診斷學習困難，以及對學習能力強的學生給予學習動機的增強），可惜的是，潮流一過，目前學校中已較少教師真正的執行能力本位教學了。因此，大體而言，工業教育之教學評量多是用以決定學生成績的高下，談不上診斷學生學習的困難。

事實上，今日職校多忽視教學前之評量，因而教學往往缺乏學生起點行爲的依據（郭生玉，民七十六）。由於工業

教育之科目，特別是技能實習科目，其知識技能係成學習階梯結構，亦即前後有關連，前面學不會，後面就不會，因此，在教學前了解學生的學習起點行為實有其意義。至於前後有密切關連的教材或課程，如機械製圖的正投影原理及正投影視圖的畫法之學習，未達相當水準時，必會影響後續的工作圖進度，職是之故，形成性評量在工業教育中確實佔重要地位，今後職校之教學評量宜針對上述缺失力求改進。

(四) 評量結果的解釋問題

如前述，教學評量依參考點亦即評量結果的解釋，分為常模參照評量與效標參照評量，前者指評量結果的解釋，係依據個人分數在團體中所佔的相對位置而定，亦即，個人分數的意義是和別人比較而得的；後者則指評量結果的解釋，係依據教學前教師所設定的標準，凡是達到標準者，表示學習「精熟」，而未達標準者，表示學習「未精熟」。今日職校教學評量普遍存在的問題，在於偏向將評量結果只重在和別人比較，過分強調個別間的競爭，因此，評量結果只重名次排列，特別晚近升學主義風氣已吹亂工業教育的陣營，都會區學校的教學評量已甚少再以效標參照評量為主了；於是乎，教師既難於了解學生學習的困難所在，也不知道學生到底學會那些知能，同時，對自己的教學效率，也無從回饋。

(五) 命題觀念的問題

「考試領導教學」已是不爭的事實，教學評量的命題方

式，影響學生學習的方式甚鉅。長久以來，工職學校評量命題最引人批評者如下所述：

1 試題偏重記憶測量，強調「背多分」，忽略綜合能力與思考能力之測量。

2 試題斷章取義，直抄課本或市售教科書。

3 試題型式與預定評量結果不相配合。評量時，應考慮到那一類型的學習結果應該用那一類型的題目，自不可用固定單一的型式去評量各類型的學習結果。

4 技能評量試題與學生的能力水準不符，難易度及鑑別力教師多無法探知，遑論試題分析了。

二、教學評量問題之改進

針對上述教學評量的問題，提出可行之修正看法，臚列於下：

(一) 評量應以教學目標為導向

教學評量應以工業教育既定之教學目標為導向，明確指出認知、技能、情意教學目標，一切教學設計、活動均以此目標為依歸，時時評量檢核是否偏離目標，以及提供增廣或補救教學，最後，總結性評量更應以教學目標為準，適時回饋整個教學系統，如此，方能提升整體教學成效。

(二) 評量方式多元化

學生個別差異大，為求教學成效直接回應學生需求並培

養學生各方面能力之發展，教學評量應採多種方式進行，不應侷限在筆試一種，宜配合教學目標，利用各種適當評量方式進行評量。

(三) 加強能力本位之評量概念

能力本位的概念非常適合技能實習科目之教學評量，亦即，重視教學前、中、後之各階段評量，以及學生學習前啓始行為之界定，增廣或補救教學之實施以及學生知能精通程度之判定，以確定教學目標是否達成。

(四) 多以效標參照評量解釋評量結果

依據教學目標，慎選評量參考點，多以效標參照評量來解釋評量結果。以清楚得知學生誰學會那些，誰的困難有那些？

(五) 增進教師命題之技術

1 於師資培育機構增設教學評量課程，強化教師此項教學能力。

2 辦理在職教師教學評量研討會或研習班，灌輸正確命題技術、試題分析、評量結果解釋等能力。

3 學校教師自編測驗時，建議依循下述步驟可得較佳之測驗卷：(1)教材分析；(2)分析施教情境；(3)書字測驗雙向細目表 (test map 或 test specification)；(4)書寫評量之行爲目標；(5)評量設計；(6)命題；(7)施測；(8)試題分析；(9)回饋修正等（康自立，民七十八）。

4 建立題庫，此題庫應具備相當之信度與效度，而試題本身應有相當之鑑別力。

5 熟悉各種試題之優點與限制，以及編製時之注意事項。

柒、結語

工業教育為整體教育的一環，其教育目標旨在培養健全之工業基層技術人才，在認知領域方面為傳授各類科之基本知識，在技能領域為傳授各類科之實用技能，而情意領域則注重人格修養、文化陶冶、以及培養敬業、負責、勤奮、合作等職業道德，終能奠定學生創造、適應變遷及自我發展之能力。因此，工業教育之教學設計與活動不僅受外在科技發展、環境變遷、社會需求、學生個人需要的影響，更要以教學目標為導向，並透過教學評量隨時的、動態的回饋，修正整體教學系統，以提升教學成效，獲至既定之教學目標。

教學評量宜依據教學目標，秉持教學評量之九大原則，在教學前、中、後適時實施安置性評量、診斷性評量、形成性評量、以及總結性評量，如此才能有效得知學生學習前啓始行為，那些學生須要接受增廣或補救教學，教師本身之教學效率，學生學會那些知能，精通學習之程度如何，以及教學目標達成水準。此外，教學評量結果之解釋，宜以效標參照評量為參考點，才能實際回應教學目標，符合工業教育之

本質與特性。

教學評量的方法宜依據教學目標之領域不同，測量目的、教學情境，乃至於學生特質而選擇適當方式進行。而教師自編測驗時倘能依據系統化過程，追求試題信度與效度的提高，並確定命題本身之難易度與提升鑑別力，則必能充分發揮評量之功能。

總之，工業教育之教學評量旨在透過有系統地運用科學與技術的方法，搜集學生的學習行為與成就之證據，以評定教學系統之輸出是否達成預期之教學目標，並提供學生學習診斷與改進教學系統之主要依據。本文探討教學評量的源起、意義、原則、功能、種類，並檢討評量有關問題，以及改進建議，期能對從事工業教育工作者有所助益，進一步改善教學評量的方法，以提升教學成效，達成整體教育目標。

參考書目

一、中文部分

- 何英奇（民七十八）。評量的基本概念。載於教育部編，專科學校教師教學評量手冊，頁一—頁二八。
- 李大偉（民七十五）。技職教育測量與評鑑。臺北：二民書局。

李克明（民七十二）。情意領域的評量。國教輔導，第二四卷第四期，頁五一頁七。

李咏吟（民七十五）。教學原理。臺北：遠流出版社。

李隆盛（民八十二）。國中工藝／科技教育評鑑模式之研究。臺北：中華民國工藝教育學會。

李聰明（民五十四）。教育評價的理論與方法。臺北：幼獅書局。

吳天元（民七十七）。工科學生實習成效的評量模式。教育學院學報，第十三期，頁一四一—頁一六三。

林建仲、方榮爵（民七十八）。情意評量的方法。載於教育部編，專科學校教師教學評量手冊，頁九一—頁一〇八。

邱錦昌（民七十）。如何使教學評量的結果能真正代表學生實際能力。師友，第十一卷一七三期，頁一六—頁二二。

教育部（民七十五）。高級工業職業學校課程標準大綱暨設備標準。臺北：正中書局。

陳海雄（民七十三）。藝能學科教學評量之研究。臺灣教育，四〇三期，頁一四—頁二〇。

陳明照（民七十二）。試評目前班級教師教學評量的缺失及其補救之辦法。臺灣教育，三九五期，頁二四—頁二七。

郭生玉（民七十四）。心理與教育研究法。臺北：大世紀出版事業公司。

郭生玉（民七十六）。當前學校教學評量工作的檢討與改

進。現代教育，第一〇卷八期，頁七五—頁八一。

黃國彥（民七十一）。測量與研究。臺北：正昇教育科學社。

康自立（民七十三）。工業教育能力本位課程發展之理論與實際。臺北：全華圖書公司。

康自立（民七十八）。認知評量的方法。載於教育部編，專科學校教師教學評量手冊，頁五一—頁九〇。

溫世頌（民七十五）。教育心理學。臺北：三民書局。

張瑞村（民七十五）。工業教育技術教學評量方式之比較研究。教育學院學報，十一期，頁三二三—頁三四三。

葛樹人（民八十）。心理測驗學。臺北：桂冠圖書公司。

廖鳳池等（民八十）。教育心理學。臺北：心理出版社。

簡茂發（民七十六）。心理測驗與統計方法。臺北：心理出版社。

蕭錫錡（民七十八）。教學目標與評量。載於教育部編，專科學校教師教學評量手冊，頁二九—頁五〇。

蕭錫錡（民八十）。機工類技能檢定評量模式之研究。高雄：復文出版社。

二、英文部分

Crocker, I. & Algina, J. (1986). Introduction to classical & modern test theory. N.Y.: Holt, Rinehart

and Winston, Inc.

Finch, C.R. & Crunkilton, J.R. (1989). Curriculum development in vocational and technical education.

Massachusetts: Allyn and Bacon, Inc.

Gronlund, N.E. & Linn, R.L. (1990). Measurement and evaluation in teaching. N.Y.: Macmillan Publishing Company.

Madaus, G.F., Scriven, M. & Stufflebean, D.L. (1989). Evaluation Models. M.A.: Kluwer-Nijhoff Publishing.

Sax, G. (1980). Principles of educational and psychological measurement evaluation. San Francisco: Wadsworth, Inc.

Wolansky, W.D. (1986). Evaluating student performance in vocational education. IA: The Iowa State University Press.

【作者簡介】鍾瑞國先生，臺灣省南投縣人，現任國立彰化師範大學工教系副教授兼職教學院實習工廠主任。

