

## 第二章 文獻探討

此章呈現有關教學評量的相關概念與研究，第一節呈現教學評量的概念與教學評量的方法，第二節呈現有關幼兒園教師實施教學評量情況之相關文獻。

### 第一節 教學評量

#### 壹、評量的相關名詞與發展趨勢

教學與評量具有密切關係，在教學的過程中透過評量的回饋作用，可以促進教學成效，因此評量一般被視為是達成教學目標的重要手段，是整個教學過程中重要的一環（何英奇，1992）。然在談及教學評量之前，有幾個名詞須先釐清，即：測驗（test）、測量（measurement）、評量（assessment）、評鑑（evaluation）這四個名詞，有些文章將之視為同意詞而混著使用，也有不少文章加以區別、界定，但因各家的翻譯不一、說法不同，容易讓人混淆。台南師院測驗發展中心（1998）就曾指出，特別是在課堂教學上，國內常將「評量」視為「測驗」和「測量」的同義詞，互相使用；而且不管是中文的「評量」與「評鑑」，或是英文的「assessment」與「evaluation」都常互用，概念也相近。茲先對此四個名詞加以指述。

測驗：狹義的觀點認為測驗就是用以測量幼兒行為樣本的工具；廣義的觀點則認為測驗是行為樣本的系統化測量程序，是對行為樣本所做的一種客觀的標準化的測量（楊銀興，1999）。

測量：是指「採用一套標準的刺激、對個人的特質作客觀、測量的有系統程序」（林順發，1995，p23）。以數量來描述個人具有擁有的屬性、特徵、行為表現，指出其在所屬團體中的相對位置（江啟昱，1994；McMillan, 1997）。

評量：是採科學的方法與途徑，多方面蒐集適切的事實性資料，再參照合理的評量標準，加以分析並做綜合研判。所以，評量比測量多了價值判斷的工作，會根據一定的標準，對所蒐集的事件、物體、或個體做價值判斷。根據 Anderson（2003），評量是收集資料以作有效決定的過程，在評量前，必須先知道為何要評量（目的）、需要那些資料去作決定（基礎）、何時需要資料（時間）、如何選擇最佳資料（方法），然後對所蒐集的資料進行價值判斷。

評鑑：評鑑的定義比較廣泛，包含獲得幼兒學習訊息的各種方法，例如傳統的紙筆測驗、擴展式反應的題目（如申論題）以及真實情境的表現（如操作實驗），然後根據所獲得的訊息，對幼兒學習進步的情形做價值判斷的工作。

從上面四個名詞的定義來看，可見「評鑑」和「評量」的涵義相近，比「測量」廣且深入，而測驗只是測量的一類。吳毓瑩（1993）認為這四個概念從單純到複雜的順序為：測驗、測量、評量、評鑑。但有人主張評量和評鑑的概念比較接近，如劉安彥（2003）提出評量同評鑑一樣，都深具主觀性，都是以測量結果為根本依據，在一既定的某些條件和準則進行評量的活動，因此，評量與評鑑的概念是可以互用的。他並進一步指出「教學評量」（educational assessment）的「評量」乃是傳統測量的延伸和多元化：延伸

指的是除了客觀的測量外，再加入評鑑的功夫，多元化則是運用各種不同的方法與方式來進行評量。綜合上述眾多評量相關名詞的釋義，本研究採用大多數學者的意見，認為評量與評鑑在概念上相近，因而統以「評量」稱之。從檢視有關評量的論述中，可以發現評量的定義日趨廣泛與深入，以更多元的角度進行判斷與解釋。

由於晚近標準化測驗受到越來越多的質疑，加上新的教學評量典範（質的評量）一一出現，學者們開始強調評量時應考量各種相關的整體情境，從各種可行的途徑，蒐集全面性、多元化的資料，再從各個角度和不同觀點加以比較分析與綜合研判，進行整合性的詮釋，獲致充分的了解（王文中等，1999；李坤崇，1999；吳裕益，2000；周珮儀，1996；楊銀興，2000；簡茂發，2002；Guddemi, 2003；Ward & Ward, 1999；Worthen, Sanders, & Fitzpatrick, 1997）。

從上述學者的論述中，茲歸納出近年來評量的趨勢：

- （一）靜態評量變為動態評量：傳統上，教師常以考試成績來評量幼兒的學習成果，較不關心幼兒改變的情況，而關心某段時間之後，幼兒的整體表現到底如何。相反的，在動態評量裡，主張學習是動態的，會因時間不同，而產生質和量的變化，因此評量的重點擺在隨時間而產生的改變，教師關心的是幼兒學習的軌跡和改變的情況。近年來盛行的卷宗評量（portfolio assessments，也有譯為檔案評量）就是基於動態評量的理念。另外，幼兒主要是透過遊戲學習，在遊戲中統整學習，在遊戲中展現學習，評量因此須在遊戲中進行，以合乎動態評量的精神。

(二) 機構化評量變為個人化評量：就評量的目的而言，傳統的評量通常是為了支援機構，如學校透過評量以了解教育目標達成的程度、比較各校間的差距、評估課程設計的優缺點、教師的教學效能等，但是無形中取代了非正式觀察的診斷，也往往忽略了幼兒的個別差異。因此，近年來轉向個人化評量理念，視幼兒個人的成長為評量的主要目的，希望透過評量幼兒的學習狀況，改進教學的方式，設計適合幼兒的課程規劃和學習策略，提供幼兒輔導與諮商，以解決其學習障礙等。為了要達到個人化評量的目的，教師要採持續性動態評量，隨時掌控幼兒的學習狀態，設計適當的課程和補救措施；而且因為每位幼兒有他自己的發展速率、家庭、文化和經驗背景，所以一種評量標準要套用在所有的幼兒身上是無法滿足大部分幼兒的需求的。

(三) 單一評量變為多元評量：傳統上，評量只強調認知，忽略情意和技能，而認知也常只強調零碎的記憶、片段的知識而已；而且傳統的評量常指用紙筆測驗，甚至只用紙筆測驗中的是非題和選擇題，因此評量的內容與方式是單一的。相反的，在多元評量裡，體認到教學目標和學習成果的多樣性，因此除了認知的評量外，也要兼顧推理、問題解決、創造的能力，因此評量的內容是多元的。評量方式除了紙筆測驗外，可使用實作評量、口試、師生間日常溝通與觀察，蒐集幼兒的學習狀況。彈性的使用多種評量方式才能全面的理解幼兒，因此評量方式是多元化的。

(四) 虛假評量變為真實評量：傳統上，評量的題材嘗試利用一些虛擬

的資料和事件，由於題材的虛假和幼兒的真實世界差距太大，以致往往困惑幼兒，降低學習興趣，也難以判斷幼兒是否可以有效地應用知識。然而，學習除了必須和個人的經驗結合外，現代的教育心理學也指出學習不能獨立於文化情境體系之外，因此測驗與評量當然也要加以配合，在測量的材料方面，要與幼兒的認知和經驗有關。真實評量，也就是所謂真實情境的教學評量，即是配合高層次思考能力的目標，以課程與教學為基礎，並且以日常生活情境的應用為著眼點所採用的評量方式。

由上述教學評量的發展趨勢評量可知，教學評量的方式趨向「動態化」；視幼兒「個人化」的發展與成長需求為評量的主要目的，教師要彈性的使用「多元的」評量方式以瞭解幼兒各方面「真實的」的能力與學習狀況。

## 貳、教學評量的功能與目的

一般而言，教學評量是運用科學方法和技術蒐集有關幼兒學習行為及其成就的正確可靠的資料，予以分析、研究和評斷的一系列工作（林順發，1995）。為進一步了解評量與教學的關係，圖 2-1 呈現 Glaser（1962）所提出的基本教學歷程，其中清楚指出教學評量在整個教學活動中的關鍵性地位。

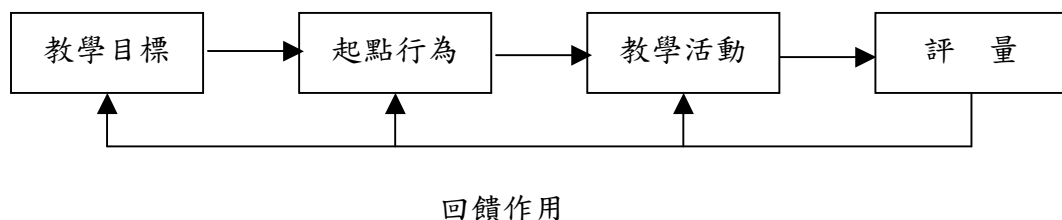


圖 2-1 基本教學歷程圖

資料來源：出自何英奇（1992，p.5）。

由圖 2-1 可見，教學評量的作用在於評估幼兒的起點行為以及教師教學成效與幼兒學習成果，以了解幼兒學習的進程與困難，提供幼兒進一步學習的訊息，並從而作為教師改進教學、修正課程設計的參考（林進材，1993；林順發，1995；簡茂發、單文經，1995）。評量對教師、幼兒具有回饋與決策的功能，如下所述（何英奇，1992）：

#### 一、教學決定的功能：（對教師而言）

- （一） 了解幼兒的起點行為：透過預備性評量，教師可以了解幼兒的起點行為，以作為訂定目標的參考。
- （二） 建立確實可行的教學目標：教師了解幼兒的起點行為，對照原先設定的教學目標，據以調整教學目標。
- （三） 確定教學目標達到的程度：透過形成性或總結性評量以了解目標達到的程度。
- （四） 改進教材教法，增進教學效果：評量的結果可回饋給教師，做為改進教材教法之參考。

#### 二、輔導與諮商決定的功能（對幼兒而言）

- （一） 增進幼兒瞭解教學目標。
- （二） 增發幼兒學習動機。
- （三） 協助幼兒了解自己的能力、性向、潛能與人格特質等。
- （四） 診斷幼兒的學習與適應困難。
- （五） 輔導幼兒做最佳學習方式的選擇。

#### 三、行政決定的功能

- （一） 選擇決定：如入學考試的選擇，在幼兒教育中此功能較少為用。

(二) 安置功能：把學生安置到適當的班級。

(三) 分類決定：指將幼兒分配到不同類別的學習組別。

(四) 課程與教育計畫決定：對於新課程、教法、教育革新計畫與研究等，需經實驗、評量後才能決定是否可行。

郭碧喙（1991）也提出，對幼兒進行評量工作，有下列幾個目的：(1) 找出有關學習或行為問題的本質；(2) 找出幼兒在學習能力方面的長處與短處；(3) 評量行為問題；(4) 考慮幼兒的能力、人格、教師及家庭情形，進而幼兒發展一個合適的教育計畫；(5) 評量幼兒對輔導處理方式的反應及(6) 在課程及安置方面做建議及修正。

基於此，我們更可以理解凡是與幼兒教育有關的人都需要了解幼兒的發展，才能來了解自己花在幼兒身上的努力是否對其未來成就有所幫助，才能衡量幼兒發展的程度以及幼兒對教學的反應如何，以督導教育人員促進幼兒發展。

晚近，學者逐漸重視到評量對家長的功能。例如：葉殷（1995）提出評量幼兒的三大目的，有：(1) 提供教師教學效果，計畫適合幼兒學習的課程；(2) 紀錄幼兒成長與學習階段；(3) 協助家長瞭解幼兒學習進度，求取家長合作，增加幼兒學習成果，就多了家長的層面。同樣地，黃淑苓（1995）在指述教學評量目的時，也將「幫助教師與家長溝通幼兒的學習表現與學習成就」列為目的之一，也增加了評量對家長的功能。

因此教學評量的目的需考慮對於教師、幼兒、園所（行政人員）與家長的回饋，不斷的檢視與調整適合幼兒學習的教學活動，以增進幼兒的學習與發展。

### 參、評量的基本原則與特性

教學評量要如何執行呢？張清濱（1996）提出要使教學評量發揮效果，教師至少必須遵循下列十個原則，略述如下：

- （一） 教學應根據教學目標：評量的主要功能在於判斷教學評量是否達成，評量內容自應涵蓋教學目標。
- （二） 評量應兼顧認知、技能、情意領域：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都一樣重要，教學評量的內容卻不宜有所偏廢。
- （三） 評量應適應學生的個別差異：教學評量應由擔任教學的教師根據班上學生的程度，選擇合適的方式，評量學生的學習結果。基於此，教學評量宜採標準參照評量，而非採用常模參照評量，且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、教學後都可評量學生，且教師在教學過程中應視情況採用。
- （五） 評量應兼顧過程與結果：評量應重視學生學習過程，也要重視學習的結果。
- （六） 評量應在各種不同的情境中實施：學生在團體活動、小組活動、個人操作時間，或是戶外活動時都會表現一些行為，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：教師可利用各種不同的工具及評量



的技術，蒐集有關教師教學與學生學習結果的資料。教師使用何種方式評量，要看評量的目標如何而定，但教師不宜僅使用一種方式去評量某一目標，混合使用各種方法比單一採用一種方法更能看見學生各方面的行為與表現。

- (八) 評量是師生共同合作的過程：學校行政人員、教師、學生、甚至是家長都應參與評量計畫的工作，因為他們對於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，家長也應多參與討論其子女的學習狀況。
- (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。當他們評量個人或團體努力的結果時，他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：評量應是教學的一環，二者有密不可分的關係，教師會觀察並紀錄學生的學習進展情形，會將評量的結果告知學生與家長，提供立即回饋；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。

林進材（1993）歸納出教學評量的幾點特性：

- (一) 持續性：整個教學歷程中階段的不同須配合各種不同的評量，以發揮不同的作用。例如在教學前所實施的預備性評量旨在了解學生的起點行為和舊經驗以及本身所具備的能力，在教學進行階段中實施的形成性評量旨在了解教學效果及效率問題，並依據評量結果診斷出問題的癥結所在，擬定補救方案；最後在教學活動後

進行總結性評量，以確定教學效果。

- (二) 有效性：教學評量對於改進教學活動應該是積極的，而不是消極的，因此，它要能適切地評定教學進行的情形，掌握整個教學的脈絡，了解教學活動與教學目標是否相互一致、教學評量與教學目標是否相互配合。
- (三) 多樣性：教學評量不應僅侷限於某一教學角落裡，而應該掙脫傳統的桎梏，採用多樣的評量型式，教學評量才不易僵化而流於單一形式。
- (四) 全面性：教學評量不但要縱的前後連繫，也要橫的上下連續；除了兼顧認知、技能和情意三領域的學習，在每個層面和領域均應有其交織點，如此教學評量才能因全面性需要評量實際學習活動和教學過程。
- (五) 發展性：教學評量應著重整個評量的發展性，評量結果應以過去的、上次的評量為基準，與現在的、此次的相互比較才有意義。

在幼兒教學評量方面，我國於民國 76 年頒布的「幼稚園課程標準」，以「參、實施通則的第四點教學評量」對於教學評量也有所規定，羅列於下：

- (一) 教師應根據教學目標進行教學評量，以作為改進教學的依據。
- (二) 教學評量應包括前評量、教學活動中的評量、後評量和追蹤評量。
- (三) 教學評量的方法有觀察、紀錄、口述、表演、操作、作品等，教師可視教學評量內容，相機配合運用。
- (四) 教學評量的工作可從教師方面與幼兒方面分別進行；其評量內

容宜依課程性質加以訂定。

(五) 教學評量的結果，需與家庭教育配合。

林淑玲(1997)也提出教學評量的一些重要關注點，並據以評論國內的幼兒評量方式，她指出：

(一) 評量是「決定」的歷程，而非「價值判斷的歷程」：評量所要做的是決定行為樣本達成目標的程度，並沒有要對行為樣本進行好壞、優劣的價值判斷。但現在幼稚園的教學評量中，常見到的是教師以與好壞、優劣有關價值評斷方式來呈現幼兒的學習表現與發展情況，同時有些教師為了避免家長看到幼兒學習的評量結果不理想，都評「優」、「良」，而不敢真實呈現幼兒的學習狀況，若真如此，也就失去了評量的意義和功能了。

(二) 評量的對象是「行為樣本」，而非「幼兒本人」：也就是對事不對人，教學評量要評的是幼兒在教學歷程中的行為，而不是幼兒這個人的籠統概念或他給別人的印象。而教師教學的過程中，如果不是當場記錄或引用視聽媒材的輔助，大多數教師事後對幼兒的紀錄都添加了他自己對幼兒的某種印象，不一定能呈現幼兒真實的行為樣本。不過當場紀錄與引用視聽媒材對幼兒園教師來說都是需要再花時間與心力的，因此在教學現場的可行性如何，或是否有其他可替代方式是教師也該一併考量的。

(三) 教學評量要用各種工具及方法，而非單一的方式：由於幼兒發展的特性、語言表達能力、適應能力及認知理解能力的限制，教師只用口語問答、事後回憶或紙筆評量等方式，往往無法獲知幼

兒學習結果，因此幼兒階段的教學評量應採多元化的方式進行。

(四) 幼兒的教學評量應是系統化的過程，而非臨時起意的作為：所謂系統化的歷程是教學評量應事先做好計劃，對於所有應用表格或方式，都應是在教學前教師經過充分討論後執行。

(五) 幼兒教學評量應屬於課程本位評量：教師教什麼，提供什麼經驗給幼兒，就由此進行評量。每一所幼稚園提供給幼兒的經驗理論上都是不一樣的，因此幼兒教學評量所用的方式與表格應該都是由教師自行研擬，不應沿用坊間的評量本或他園所發展應用的評量工具。

綜合言之，幼兒園教師在實施教學評量時，為了使教學評量發揮效果，宜把握以下原則與特性：(一) 評量應有計畫、有系統地與教學密切結合、根據教學目標進行符合課程本位的評量；(二) 評量內容顧及全面性，不僅涵蓋認知、情意、技能，也需兼顧教學前、中、後的學習過程與結果；(三) 評量應考慮幼兒的個別差異；(四) 評量應運用多元的方式、持續的、有效的瞭解幼兒在各種情境中、各方面的發展與表現。(五) 評量可以是師生與其他有關人員共同參與的過程；(六) 評量結果的呈現是真實、具有教育意義，且配合家庭教育的功能。而家長參與在過去的評量是常被忽略的，因此盧美貴(1999)也提出可以運用真實性的評量開始師生過去的評量及家長共同參與教學評量的觀點，評量的呈現不再只是給家長的層面，更希望家長參與可以使教育的資源更為豐富，再加上幼兒的表達能力可能無法詳實呈現與表達在學校的學習情況，所以家長若與教材成為評量的夥伴，對於幼兒的學

習將帶來更積極正向的教育效果。此外，教師使用坊間教材的適切性如何，依課程相關性來看，使用坊間教材提供的評量一定不行嗎？值得再進一步深思與探討。

## 第二節 幼兒教育的教學評量內容與方式

Guddemi (2003) 提出，近幾十年來，幼兒教育團體不斷提出標準化評量不適用於幼兒教育階段，認為學前幼兒和大一點的兒童或大人的學習內容和方式大不相同：幼兒學習和建構知識的方法是經驗性的、有互動的、具體的和實際操作的，透過各種不同的媒介、用不同的方法在不同的環境中接觸和操作實體來學習。因此，傳統紙筆測驗的單一評量內容及單一評量方法並不適用於幼兒。其次，每位幼兒的學習與發展速率也不一，需要父母或教師一對一的對幼兒進行評量，才有可能獲致有意義的評量結果，再加上幼兒的注意力時間短，只有在能反映質性教學結構的高品質評量（quality assessment）出現時，評量和教學間才會變得幾乎天衣無縫、完整地互相地提供資訊。因此，美國幼兒教育協會（National Association for The Education of Young Children[NAEYC]）在 1987 時代表一萬名幼教專業人員的心聲提出一項有關幼兒評量的聲明，聲明中指出高品質評量不是以紙筆測驗的方式進行，而是以適合幼兒的年齡和發展階段且重視幼兒技能和知識展現的方式進行，強調評量者扮演的角色不再是以單一測驗的分數來解釋評量，而重在收集資料、並能敏感地從資料的收集中說明和覺知幼兒個別不一致的發展速率。NAEYC 並在 1995 年訂出「3-8 歲幼兒合宜課程內容和評量指引」

(Guidelines for Appropriate Curriculum Content and Assessment in Programs Serving Children Ages 3 Through 8)，提供具體的指引，建議適合幼兒的課程的目標、範圍、標準及各種有系統性的、持續性的評量工具。

接著，在晚近動態性、多元性及真實性評量的趨勢下，幼兒的教學評量也逐漸走向多元性與真實性的真實性評量 (authentic assessment)，幼兒的教學評量要考量三個層面：(一) 高表現：指的是幼兒主動建構答案的程度高，幼兒在評量中能呈現出主動的思考歷程及創意；(二) 高真實性：指的是評量項目或評量情境愈與真實生活相近的程度，評量項目愈接近真實生活，愈能引起幼兒的興趣，對幼兒的意義也愈大，幼兒主動想參與或自己進行評量的動機也會因此提高；(三) 與課程內容的相關性高：指的是評量內容與教學內容的一致程度，課程相關性高的評量，重視評量項目是重要而有意義的，且反映教師課程中所教的事實與概念(廖鳳瑞、陳姿蘭編譯，2002)。換句話說，就是評量項目能逼近真實生活的活動或工作，讓幼兒運用真實生活中會用的知能來建構答案，而評量的內容又能與教學課程有相關聯。

由於幼兒不會寫字或閱讀，幼兒的發展也挑戰著一般評量的策略，評量幼兒的策略必須配合幼兒每個生理、心理和社會發展階段，同時依據不同的評量目的，使用不同的評量方法與策略 (Wortham, 2001)。所謂工欲善其事必先利其器。以下從幾個不同的層面來詳細說明評量幼兒的方法：

## 壹、評量的時機和功能

評量可依教學時程的前、中、後三階段來區分，分為「預備性評量」(preparative evaluation)、「形成性評量」(formative evaluation) 和「總結

性評量」(summative evaluation)，而在不同階段，教師評量幼兒的責任和內涵也隨之有所變化(張景媛，1992；劉安彥，2003)：

- (一) 教學前的評量(或稱預備性評量)：指教師在進行某一單元或主題之前，先給幼兒實施的教學評量，主要目的是要了解幼兒的基本能力與起點行為，以作為設計教材內容及決定教學方法的參考；其次，則是了解幼兒的個別差異，以決定如何因材施教。通常教師運用觀察、晤談和報告等方法來收集相關資料，並對此等資料進行解析，以期了解幼兒的文化背景、興趣、技能以及有關學習其他相關活動的能力，同時，哪些幼兒在基本能力上有所欠缺，然後針對這些缺失進行補救教學。
- (二) 教學中的評量(或稱形成性評量)：指教師在教學過程中所實施的教學評量，一方面是讓教師了解其教學成效，另一方面則是考核幼兒學習的優缺點。主要目的乃是監控幼兒邁進學習目的的程度，測定幼兒的學習成就以及學習困難，幫助教師調整教學活動，以提供適合的回饋給幼兒。
- (三) 教學後的評量(或稱總結性評量)：指教師在單元教學結束後所進行的教學評量，描述個別幼兒達到短期或長期學習目標的程度，讓家長、教師等人了解幼兒的優點及其需要改變、改善之處。總結性評量的範圍較大，在認知方面涵蓋的題目應包括知識、理解、應用、分析、綜合及評鑑等六方面，且應顧及技能及情感兩方面。

若要教師的教學與幼兒的學習能有效，此三種評量皆不可忽視的。

## 貳、評量工具與方法

大部分幼教機構和幼兒園教師同意幼兒評量應包含許多質性的評量工具，當一起使用時，這些工具可以創造出一個評量系統，這一些評量的工具包括：觀察和檢核表、軼事紀錄、持續性觀察、文件檔案、家庭訪問、發展檢核表、診斷性評量和標準化評量等（Guddemi, 2003）。王佩玲（1995）提出幼兒評量的方式有觀察法、訪問法、測驗法與個案史的方式；一些學者指出國內幼兒園現場的教師自陳在教室內試行的評量方式有：幼兒行為觀察紀錄、幼兒發展檢核評量表、教師自編學習單、教學活動評量表、幼兒作品集（檔案）、活動照片、錄音帶與錄影帶等（何佳盈，2003；呂翠夏，2000；蔡其綦，1997）。分述幼兒園教師較常使用的幾種評量方法，包括：觀察法、檢核表、檔案評量、學習單、評量系統等，還有在國內幼兒園教師較少使用的紙筆測驗：

### 一、觀察法

觀察法是幼兒研究發展中最早使用的方法，是讓教師「知道、看見」了解幼兒各方面的發展的方法，是評量幼兒發展與成就的基礎，也是教師設計教學計劃的依據，是符應幼兒的教育需求最基本的評量工具（廖鳳瑞、陳姿蘭編譯，2002）。教師與所教導的幼兒一天要相處好幾小時，他們經常觀察幼兒，而透過有系統的觀察，教師蒐集所需的資訊去評估幼兒發展的過程，以發現兒童的學習情況，及與其他幼兒、教師、大人間的相處情形，這樣持續的、以資料為基礎的決定對幼兒教育計畫攸關重要（馬燕譯，2002）。所



以觀察提供做決定時所需要的重要資訊，所做的觀察對於幼兒、幼兒園教師、課程、教材和設備及計畫和政策的細節都很有幫助。通常觀察是個不斷進行的過程（賴慧玲譯，2000）。

而在進入觀察情境之前，準備各種有組織的步驟是相當重要的（林青青譯，1995）：

1. 確定觀察的目標：有各種理由值得去觀察幼兒，而當要進行觀察前，心理總要有個明確的架構存在，明確地訂定觀察目標是非常重要的，因為不可能在同一時間裡觀察所有我們認為重要的行為。所以在觀察前要先能夠回答幾個問題：觀察的重點是什麼（社會能力、動作技能或語文能力）？要花多少時間在這些觀察行為上（一小時、數分或只要一發生這個行為便進行觀察）？要觀察的對象（一個團體、某位特殊幼兒、一些特殊行為的幼兒）？要如何記錄所觀察到的行為（檢核表、持續性的記錄、採樣記錄或時間取樣）？將採取什麼方式來說明所看到的行為（根據何種理論或事實等）？
2. 確定觀察的地點：觀察目標必須與環境所能提供的資訊相配合，例如在公共遊樂場有較多年齡層幼兒遊玩，於是觀察上便有較多的彈性。
3. 確定觀察情境中可以做的事：知道在某種觀察情境中哪些是你可以做、哪些不能做是相當重要的。有些教師希望離開教室後再記錄，在這種情形下你只能仰賴記憶直到離開教室後，趕緊拿出筆記本記下得到的資料，在這種情形下慎選記錄工具便非常重要。

觀察記錄有以下幾種主要的方式：

1. 採樣紀錄法（specimen records）：使一種正式的觀察法，又稱「描述

性的敘述法」，是指觀察者要持續不斷，且盡可能詳細的紀錄下被觀察者本身或被觀察者與他們接觸者所做的每一件事和所說的每一句話。主要目標在於取得沒有經過推論解釋或評量的，詳細而客觀的行為紀錄，其重要性在於不僅要對於行為的紀錄很詳細，且對於行為發生時的背景情境和場所以及行為發生的先後順序也許也都有很詳細的紀錄；其優點在於提供了觀察當時所發生的全部實況，是在完全自然、毫無人為影響或實驗控制的情況下，捕捉了幼兒的種種行為、背景和風格。此法在很多情況都可以使用，不需要是先預備觀察表或符號紀錄表，也不需亦任何特殊的語言或專門的術語，但缺點在於會花掉觀察者很多的精力和時間，如果想要立即獲得代表性行為，這種觀察法是很沒有效率的（劉慈惠譯，1995）。

2. 時間取樣法 (time sampling)：是指觀察者僅在事前設定的時間內，以規律性間隔或隨機性間隔觀察和紀錄所選擇的行為，在這些時段以外所發生的行為和事件則不加以紀錄。觀察和紀錄的長短、以及間隔的分配，則因觀察者的目的、在觀察地點時間的長短、以及某特定行為發生的頻率等因素的不同而有很大的變化。所以確定要觀察的行為具有代表性是非常重要的。此法的優點在於對觀察的行為沒有限制，可以從幼兒的想像行為到幼兒之間的友誼和爭吵、生活習慣等都包括；所需要的時間和精力也是符合經濟效益的，因為所觀察的內容和所需的時間都是事先設定好的，加上使用事先設計好的符號紀錄表，減少了觀察者在判斷和推論上的變異性，使得不同觀察者之間具有一致性，觀察者之間的內在信度也提高。可以同時使用符號紀錄表和描述

性的方式也此法之優點，不過此法缺乏連續性、背景的完整性和取樣事件的自然性，而某特定觀察行為發生的頻率也是此法容易受到限制的因素（劉慈惠譯，1995）。

3.事件取樣法（event sampling）：「事件」是事件取樣法的重點，和時間取樣法不同的在於其只選擇某一特定的事件作為紀錄的對象，不限特定的時間。而「事件」是指可以歸類於某特殊範圍的行為，記錄的技巧就在於立即記下某些事件發生的情形，要詳細地描述行為的細節，可以選擇符號記錄表或敘述性描述或兩者並用。此法之優點在於很實用，尤其適用於記錄那些並不經常發生的行為，且是由行為和情境的自然單位所構成的，可以讓觀察者研究行為和環境的關係。而且缺點在於中斷了行為連續性，因此無法保持行為的完整性（王莉玲譯，1995）。

4.軼事紀錄法（anecdotal records）：是教師常用的非正式觀察法，軼事紀錄並非限於記錄幼兒的新行為，而是記錄下所有觀察者覺得值得記錄的行為。教師可由紀錄中做推論以增進對幼兒人格特質及行為的了解。此法有五個主要的特性：第一、直接觀察幼兒，而不以道聽塗說為記錄的根據。第二、針對某特殊事件做迅速、正確和詳細描述的記錄。第三、提供了幼兒行為發生的前後關係，說明了行為的背景及情境，這種前後關係包括了幼兒或其他參與者所說的話。第四、如果觀察者要做任何推論或解釋，則無需再做客觀的敘述，因記錄的本身非常的清晰。第五，所觀察的行為可以是一般的或特殊的行為，如果是不常見的行為則需做說明。此法之優點在於能提供教師瞭解幼兒行為

的詳細資料，也可以不斷地用來比較幼兒的行為，以了解幼兒的成長變化；此法是最容易的記錄法，因為不需安排特別的情境、範圍或時間；不過其缺點在於不易撰寫。也被批評容易，受偏見影響而選擇所記錄的行為，還有不恰當的文字記錄會導致閱讀報告者對幼兒產生錯誤的解釋或價值判斷等（王莉玲譯，1995）。

5. 日記式記錄法（diary records）：是一種非正式記錄的觀察法，是研究幼兒發展所最早採用的方法。它是對於幼兒做一段長時間的研究，以日記的方式記載幼兒成長與發展中某特定範圍的情形；目的是有次序地記錄個體行為發展中所產生的新行為，尤其是當個體正值嬰兒或幼兒時期，所強調的是記錄幼兒所表現的新行為。使用時需要觀察者與幼兒間保持親密的關係，以做長時間的接觸，因而似乎也只有父母或監護人能做這點。此法之優點在於具有詳實性、廣度性和能與其他的觀察資料做比較的永久性；缺點在於限制了觀察者的人選，即使是教師也無法和幼兒維繫這種關係（王莉玲譯，1995）。

6. 次數統計法（frequency counts）：「次數」是此法的主要特性，每當特定行為出現時，觀察者只要在觀察記錄表上做記號即可，持續性記錄法（running records）是次數統計法的另一種變化。此法需事先對所欲觀察記錄的行為予以界定。優點在於其簡易性，記錄表本身容易操作，能立即提供量化的資料，可以用於發現某行為在時間內的改變；缺點在於它們的閉鎖性，無法得知行為的細節及其前後關係，只可用來記錄那些不需要很多描述性的片斷行為（王莉玲譯，1995）。

綜上可見，每一種觀察紀錄法都有其不同的特性、不同的使用方法、不

同的優缺點，也都有其最適合運用的特定情況和目的。

外國學者曾為文論述在觀察過程中的三個重要因素，能幫助觀察者更正確且可信度更高（林青青譯，1995）：（1）客觀的描述；只儘可能將所觀察到的事物確實而完整的記錄下來，不加入私人成見、解釋、價值及印象，無論任何人來做觀察都可以得到相同的資料。（2）說明與解釋：除了客觀的描述外，試著給予所觀察到的事物解釋或賦予某些意義。（3）評估：在觀察過程中，評估所要冒的風險可能是最大的，因為我們會不由自主的將自己的評價或態度加諸在幼兒的行為、特質或人格上，甚而形成一種刻板印象。不過一旦有了刻板印象就很難有客觀的評量，任何評價無論是正面或負面，都必須植基於經常性、具有代表性且客觀性的行為描述中。因此使用各種觀察法的同時，可以利用照片、錄影帶、錄音帶來豐富資訊；而要有系統的觀察，教師要先列出問題，決定想要看什麼、哪一些幼兒想多瞭解一些、觀察的重點是什麼，然後再選擇合適的紀錄方式。

## 二、檢核表法

檢核表法（checklist）非常容易使用，是能顯示一些現象是否出現的一種記錄法。檢核表是在特定情境中，紀錄幼兒有無出現特定行為的方法，再根據表上所列一系列明確具體的行為，評量者根據觀察，勾選行為的有或無。檢核表之表格編制容易，填寫者也容易勾選。檢核表有兩種：第一類是為靜態性描述性項目，是針對研究對象或情境中具有穩定特性的項目（如年齡、性別、物理環境特性和時間等），用有系統的符號來檢核者。第二類是活動式檢核表，以行為，為觀察的重點。典型的活動式評量是在觀察時段內

記錄特定行為的出現，只要幼兒在觀察時段內出現這些行為項目，觀察者即在表上做記號。檢核表也可以用於記錄被評量者是否能表現一些被要求的行為，此時檢核表就成為一種測驗（王莉玲譯，1995）。

檢核表法通常具有的評分表的功能，評斷幼兒學習和發展的價值時，需要評分表的準備和使用的目的。如果設計的得當，通常有三種功能：(1) 引導對行為特定向度的觀察；(2) 提供在同一個特質上，對所有幼兒表現評分的架構；(3) 提供記錄觀察者判斷的便利方法。而檢核表的評分標準應該在設計評量之初就加以定義，至少也要暫時地定義。一個健全的評分標準，包括下列六個特徵：重要性、逼真、概化、發展的適切性、易接近性與實用性（呂金燮，1999）。至於要怎麼樣去評學生的反應，一般來說，有的人採用三階層（3-level）的分類來評量幼兒反應層次的好壞，將評分層次分為高、中、低；有的則採用四階層（4-level）的方式分為不可接受、需要改進、好、非常好的評分層次，至於要如何認定所表現的要達到什麼程度才可稱為「好」或「非常好」呢，則須再以文字敘述，或舉出具體例子作為評分的參考。

因此，檢核表的實施需要周詳的計畫，所要檢核的行為項目必須事前清楚地界定。但檢核表之優點在於教師及觀察者可不受時間和情境的限制，隨時記錄，容易使用且易達到效果，因其能以簡單的符號提供可計量的資料，可減少複雜的描述性資料。但是此法之缺點在於缺少原始資料，缺乏觀察行為及情境的細節，所觀察的記錄主要是幼兒一些片斷的動作或對幼兒零碎的印象，無法獲知事情的前因後果，容易陷於主觀與籠統，評定結果容易失去準確度（王珮玲，1995）。

而另外一種評定量法（rating）也是將所要測量的特質編製成題目，由

觀察者根據觀察的結果，在題項中勾選，除方便使用外，也可以量化，且在評量之後，可加以相互比較，但其限制在於觀察者之主觀因素，會導致對評量項目有不同的定義，就可能造成誤差，例如一般人在選擇答案時，會有選擇中間答案的趨勢。因此，使用評定量法時，須對量表有一致的共識，以提高客觀性（王珮玲，1995）。所以評定量表在使用之前需要對行為有先行的判斷，應尋求嚴謹的範疇和可觀察的行為（馬燕譯，2002）。

因「檢核表法」與「評定量法」在使用上均是針對評量項目進行勾選檢核的方式進行評量，差異僅在於勾選的選擇數的不同，因此本研究中將此兩者歸為同一類型的評量方式，以「檢核表法」統稱之。

學前的幼兒園教師可使用檢核表來評量和紀錄幼兒各方面發展的成長與進步，包括身體動作、認知、語言、社會等發展方面，且他們通常利用檢核表去篩選和決定幼兒適合何種課程或活動；檢核表也可幫助教師去設計符合幼兒經驗和能力的課程，和紀錄與評量幼兒的行為表現與學習進展（Wortham, 2001）。

### 三、檔案評量

檔案評量（portfolios）是一種以表現為依據的評量類型，以有目的的蒐集幼兒的作品，展現出幼兒在一個或數個領域內的努力、進步、與成就（吳毓瑩，1999；廖鳳瑞、陳姿蘭編譯，2002）；此方法以觀察紀錄構成文件檔案的基礎，而要豐富觀察紀錄，就需要增加蒐集每個幼兒的工作，包括幼兒的素描、故事、讀過的書單及其成果。一段時間下來，文件檔案可以提供一個有關幼兒的進步情況報告。比起正式篩選測驗和其他測驗，文件資料裡蒐

集的兒童實踐活動更具動態，檔案評量一個很好的用途是提供家長資料，教師和家長一起閱覽所蒐集的資訊，除了可以看到孩子的進步，家長也可以具體的瞭解教師的教學（馬燕譯，2002）。根據何嘉盈（2003 在實際運用檔案評量後，表示幼兒園教師運用檔案評量的三個重要理由為：針對幼兒個別需求提出指導、教師經此擁有專業成長、豐富家庭參與學校教育的形式。

檔案評量的主要功用是紀錄幼兒所學的歷程與結果，所以課程的內容決定了幼兒檔案中應選擇什麼樣的作品（鄭英耀、蔡佩玲譯，2000）。Meisels 及其同事針對學前幼兒的檔案評量，提出詳細的實施的步驟（廖鳳瑞、陳姿蘭編譯，2002），描述於下：

### 1.開學前

- (1) 為每位幼兒準備一個作品檔案，並決定如何保存作品集檔案及檔案存放的地點。
- (2) 設計蒐集幼兒作品的程序。
- (3) 思考呈現幼兒學習的方法。
- (4) 計畫核心項目的蒐集。

### 2.蒐集期間

- (1) 規劃課程時將作品的蒐集考慮在內。
- (2) 教導幼兒有關作品集的事情。
- (3) 與幼兒討論他們的作品。
- (4) 規律地蒐集作品。
- (5) 定期選擇作品。
- (6) 為作品作註解以增加它的意義。



(7) 追蹤整理所選擇的作品。

### 3. 蒐集期結束

(1) 幼兒回顧並評鑑作品集。

(2) 教師回顧並評鑑作品集以撰寫綜合報告。

(3) 與幼兒家長分享作品集。

盧美貴(1999)指出檔案評量是有計畫的蒐集幼兒作品，並將檔案的組合方式，整體為四類：(1) 陳列類 (showcase portfolio)：可由教師或幼兒挑選優良的作品，加以陳列，這是最開放的方式，幼兒可以自由選擇其所喜愛的作品，並依照自己的想法歸類。(2) 文件類 (documentation portfolio)：由教師或幼兒依參與的時間，蒐集其學習進程中的各項成果或作品。(3) 過程類 (process portfolio)：顯示幼兒在一項大型的工作過程中的學習表現。(4) 評鑑類 (evaluation portfolio)：可由多人組成評量人員，針對既定的目標評量學習成果。

檔案評量的特色有：(1) 多元化的評量方式；(2) 增加互動及參與的機會；(3) 瞭解幼兒的能力；(4) 實現開放的精神：行動和思想自由、參與均等、資源和訊息流通、個別差異得到尊重、重歷程也重結果；(5) 注重自我比較；(6) 強調幼兒學習的整體表現；(7) 評量者特性等，尤其強調家長在評量中的角色，由過去評量成績或資料的被動接受者，轉變為資料及回響的提供者，也可以在評量過程中便成為活躍的參與者(盧美貴，1999)。但檔案評量也有其問題陳金祝(2001)對於檔案評量的省思有二：(1) 認為教學評量的內容，大多以教師的意見為主，使學生參與學習的機會減少、教師應考量學生個別差異的需求，斟酌加入學生的意見，以培養學生自我負責的正

確態度。(2) 顯而易見的缺失是將學生所有書面資料，未經過濾全數納入檔案中，教師應和學生共同研討分析，找出具有意義與保存價值的資料，才能使學習檔案真正成為記錄學習的寫真集。

#### 四、學習單

「學習單」(worksheet) 最早是從 worksheet 這個字發展出來的，已經採行很久；意涵為：教學者根據教學計畫中所列學習目標，設計學習者可以完成該學習目標之指導內容。學習單的設計強調與教學活動的配合，盡量作統整概念性的呈現，提供學生反省與創造的空間，因此，學習單可以是學習過程的記錄，也可以是教學後的課後作業或心得分享等，主要在補充現有教學之不足，記載幼兒學習的歷程和結果，讓幼兒掌握學習的脈絡與腳步，完成完整的學習（陳金祝，2001）。其應具備的功能有：1.引導學生學習；2.評量的功能；3.發展性；4.活動性，不只是停留在寫字的功能上（林秋榮，2002）。

而學習單的運用情況，盧美貴（1996）在台北市幼稚園與小學銜接的實驗與推廣中提出，為加深加廣或精熟學習內容，評量方式採「學習單」形成性評量，包括認知、情意、技能的多元化內容，隨時與家長保持密切聯繫，以協助幼兒充實、快樂而無障礙的學習。同樣的在作品檔案系統中也運用學期單作為呈現幼兒學習的方法（廖鳳瑞、陳姿蘭編譯，2002）。而陳金祝（2001）依據在國小的實施經驗表示，課程後省思的學習單的立意雖佳，但過多的學習單反而成為學生和家長的重大負擔，「談學習單則色變」的情形層出不窮。因此，對於學習單的使用，陳素櫻（1997）提出可行的建議：（1）設計學習

單的最高原則為協助幼兒順利進行學習而製作。(2) 事先檢視教學內容，避免設計重複內容的學習單。(3) 有些學習單的規劃，是使學生達到精熟學習的目的，可考慮改為「作業單」或「練習單」等名稱，以免過度氾濫。

## 五、紙筆測驗

以時代脈絡的觀點來看，傳統的紙筆測驗，基本上是出現在結構思想和操作主義盛行的時代，在此思潮中，知識被認為是依經驗建構可獨立的經驗的某種理論模式，因此利用結構好的學習評量，就可以充分檢視受測者的知識、能力。這種方式就檢驗教學成果的目標而言，尤其是內在效度方面，具有一定的功能（陳素櫻，1997）。

然基於下列因素，測驗在幼兒教育中成為非常缺乏效率的教學評量工具：(1) 幼兒閱讀寫字能力有限；(2) 幼兒難以適應「標準答案」的反應模式；(3) 幼兒難以語文反應抽象觀念；(4) 幼兒的專注力不易持久；(5) 幼兒容易疲累；(6) 幼兒不易瞭解考試的影響；(7) 幼兒反應容易受情境、情緒的影響（黃淑苓，1995）。

因此測驗在幼兒教育領域不是最佳的教學評量方式。如果幼兒園教師要在教學評量使用測驗，必須自身具有測驗使用、解釋能力，必須能編制或選用品質良好的測驗，必須安排適當情境，並且在解釋、應用測驗結果時要參考平日有關幼兒的其他評量資料（廖鳳瑞等譯，2005）。

## 六、評量系統

有些學者結合幾種不同評量的方法組成評量系統，提供有系統的評量幼

兒的步驟、方法、表格等，例如作品取樣系統（廖鳳瑞譯，1995；陳姿蘭，2001；廖鳳瑞、陳姿蘭編譯，2002）、光譜計畫（葉嘉青譯，2002）。以下針對此兩系統做簡單介紹：

#### （一）作品取樣系統（Work Sampling System）

因感於傳統式成就測驗所提供的資訊往往誤導、脫離生活情境，密西根大學教育學院的 Dr. Meisels 與其同事共同研發而提出作品取樣系統，它是應用表現評量的一種評量工具，適用於學前幼兒至國小五年級的學童，在美國已有相當多學區的學校與幼教機構使用這套評量系統。而從其名稱可以看出其特質：

- （1）作品（Work）：指對幼兒的作品或表現有興趣，概括幼兒在學校所作的。
- （2）取樣（Sampling）：只從幼兒所有的表現中選取代表性的樣本。
- （3）系統（System）：指這套評量方法是一套系統，包含發展指引與檢核表、卷宗、綜合報告三個部分。

作品取樣系統是以教師日常教學情境內觀察幼兒積極從事與創造作品的經驗為基礎，提供一套方法來檢視及監督幼兒在社會、情緒、體能及認知方面的進步與成長；所含涵蓋的學習領域有：個人與社會發展、語言與文學、數學思考、科學思考、社會文化、藝術、體能發展等七個領域，每一個領域又各涵蓋數個功能分項及表現指標，而每個領域都會在觀察、資料蒐集、綜合評論的運作機制下表現出來。其運作方式分為：

- （1）幼兒的觀察：教師運用「發展指引與檢核表」觀察幼兒的行為與表現。「發展指引」是一套依據美國課程標準與兒童發展知識訂出來

的觀察的指標，呈現對每個年齡層幼兒發展上合宜的期望，提供教師觀察的架構，讓不同地方的教師用同一種標準來執行專業的判斷，評鑑幼兒的行為、知識、成就。教師對幼兒的觀察就紀錄在「發展檢核表」內。

- (2) 資料的蒐集：教師利用「作品集」有目的地蒐集可以展現幼兒努力、進步與成就的作品，並以結構性的方法來蒐集兩種作品：核心項目與個人項目。「核心項目」用來顯現幼兒跨時間的成長以及在不同課程領域作品的品質；「個人項目」則用來展現幼兒的獨特特質，以及用來反映幼兒統整課程領域的作品。幼兒與教師共同參與作品集的設計、選擇及評量。
- (3) 表現的總整理：將前述兩項資料綜合於「綜合報告」，一年填寫三次，每次教師要結合「發展檢核表」及「作品集」的資訊，再加上兒童發展的知識，最後作出有關幼兒表現與進步情況的結論。然後教師在「綜合報告」勾選適當的表現等級，簡短描述幼兒的優點，並指出還需改進的地方。

作品取樣系統客觀地紀錄幼兒的學習與教師的教學，以提供有意義的評鑑與真正的績效考核，其優點有以下幾點：

- (1) 增強幼兒的學習動機：強調幼兒所能作的，而不是幼兒所不能作的；它讓幼兒參與評量的過程。
- (2) 增進教師對幼兒如何學的了解：要求教師要紀錄與評鑑幼兒在多個領域上的成長與發展，以及幼兒多樣化的學習風格。
- (3) 提供向家長報告幼兒進步情況的有效方法：包含資訊是個別化的，

是專門針對每一個幼兒的。對家長而言，這種內容是他們所熟悉的，而且對他們也是有意義的。

(4) 適用各種幼兒：適用的範圍包括一般幼兒、特殊幼兒或不同文化、經濟背景的幼兒。

(5) 提供溝通與合作的機會：以美國三歲到五年級的課程標準為基準，以三種互有重疊的紀錄方式（發展指引與檢核表、蒐集幼兒作品的作品集、綜合報告），提供不同階段或前後年級的教師一個共通的評量語言及標準，也提供教師、行政主管或家長一個合作的機會。

## （二）光譜計畫

光譜計畫是一個長達十年的研究計畫，由提出多元智慧理論的哈佛大學 Dr. Gardner 和泛文化理論的塔夫特大學 Dr. Feldman 共同主持，主要在發展一種不同的課程與評量的方法，希望教師、父母和幼兒自己，能夠清楚了解到可以用許多種不同的有效方式來表現幼兒多元的潛能，讓每位幼兒所表現出來的智慧、風格和特質如光譜般廣闊，而關鍵就在如何評量出幼兒身上的多元性。

為了系統化地引導幼兒進入多元智慧的領域，光譜採用了學習中心的方式，涵蓋的領域有語言、數學、自然科學、機構與建構、藝術、社會理解、音樂、和肢體動作等八種，每一種學習中心都準備各種吸引幼兒的教材教具，和許多幼教課程所用的學習中心一樣具備的特質有：動手操作的教材教具、小組活動、以及幼兒自由選擇想參與的活動。光譜學習中心獨特的地方，就是基於多元智慧理論而設計出系統化的方式以鑑定幼兒，並且支持幼兒自

不同領域的專長和興趣的活動；同時 Gardner 和 Feldman 發展出八組關鍵能力 (key ability)，是指幼兒在每一種知識領域中，需要成功表現出來的工作能力或認知技能，例如：在音樂方面，關鍵能力包括音樂知覺能力、唱歌和作曲。關鍵能力在學習中心的學習與評量方面扮演了重要角色。教師可以將關鍵能力當作觀察的原則，注意幼兒在不同領域中的專長和弱點。

王為國 (2001) 曾為文論述光譜計畫對於目前課程與評量現況的啟示，其中關於評量方面的有：

- (1) 教學與評量是一體兩面密不可分：目前的教學與評量往往是分割的，教師通常將教學和評量視為是兩回事，然而光譜取向認為教學與評量之界線是模糊的，評量並非單獨存在，光譜計畫的教學活動就是評量活動，教師在進行教學活動之同時，也在針對幼兒之活動表現以及展現之優勢、弱勢領域加以觀察並予以紀錄。
- (2) 評量是在真實情境脈絡之下進行：目前幼兒評量結果的好壞和實際情境中解決問題的能力較無關，因此，光譜計畫強調在真實情境下進行評量的做法，值得參考。
- (3) 發掘幼兒優勢智能以啟導弱勢智能：從光譜計畫實施經驗來看，發掘幼兒獨特的優勢智能，不僅可以讓幼兒從中得到自信，也可以藉以優勢智能為媒介，啟發其弱勢智能。若教師能找出最適合幼兒個人的教學方式，將有助於幼兒的學習與發展。
- (4) 提供完整詳實的評量報告：光譜計畫於每學年末均會提供幼兒多元智能發展的報告，內容詳實且具參考性，國內幼兒園教師若能針對個別幼兒紀錄下詳實而完整的評量報告，協助家長瞭解幼兒

的狀況，也容易獲得家長的信任。

(5) 尊重幼兒的個別差異：光譜計畫的精神顯示了美國的文化傳統，一是，這套教學計畫的基礎理念重視教材與科技，幼兒應該有很多機會可以接觸許多的材料；二是，強調幼兒個別差異的重要性，每位幼兒應該儘可能地發展其獨特的長處；三是，雖然光譜計畫鼓勵團體工作，但基本上每位幼兒都是個別的學習者。若要光譜計畫的理念成功地在台灣這種一向強調全體關係的文化中推廣，教師必須學習去尊重幼兒個別的目標與需求才行。

(6) 父母的參與：光譜計畫重視家長參與幼兒的活動與評量，而設計許多課後活動，要求家長與幼兒一起完成，藉以幫助家長瞭解自己的幼兒，因此國內的教育機構應加強親職教育，引導家長建立正確的教育理念，以及安排適當的在家活動與評量，以協助促進親子的瞭解。

## 七、小結

由上可見，本節呈現出教學評量對於幼兒園教師而言，不應只是促進其教學成效，幫助其了解幼兒的學習與成長情況的重要工具，更是在整個教學歷程不可缺少的環節，透過系統性、持續性、多元化的評量方式，收集幼兒身心各方面發展與學習狀況的資料，統整、組織、分析後，提供教師、行政主管或家長了解幼兒優缺點、學習表現與發展情況。不同的方法會提供不同的訊息，有些特質無法用量化的工具收集或表示，不只要收集有關兒童不足之處的資料，也要收集有關長處能力的資料；在資料的來源上，最好能採自



不同的來源，除了兒童本身之外，與兒童生活息息相關的成人（例如父母、教師等），都是兒童資料的最好來源。大部分的學者認為應採用多元的評量方式來進行教學評量（王珮玲，1995；郭碧喆，1991），評量兒童的方法各有優缺點，可互相彌補彼此的不足，也較能提供客觀且足夠的資料。因此，教師必須要將不同方法及不同來源獲得的資料統整組織在一起，然後據此作合適的建議。幼兒園教師若能充分地理解教學評量與有效地運用，將使幼兒的學習與表現獲致最大的效果與發揮，自然也達到教學評量的最終目的。

### 參、評量資料的解釋

郭碧喆(1991)表示評量結果的書寫就像評量過程一樣，要仔細且謹慎；它不只報導測驗分數，還要將行為觀察及其他評量方法得來的發現和有意義的資料一起加以考量，而書寫評量報告應該表現出對幼兒個別性的尊重。

從評量資料的解釋而言，教學評量可分為「常模參照評量」（norm-referenced evaluation）和「標準參照評量」（criterion-referenced evaluation）（何英奇，1992；林惠雅，1990）：

- （一）常模參照評量（相對比較）：是以同年齡或其他條件相若的一群幼兒在某項成績上的分配情形，取其平均數或中位數為參照點，分析幼兒之間相對的優劣。許多標準化測驗屬於這種方式。此種評量的優點在於提供個人在團體的相對地位，及其發展程度為何，這對篩選發展遲緩的幼兒有其重要的貢獻。但是幼兒的教學評量，最好能和他的教學目標、課程內容、教學方法等聯結，從中了解兒童的發展情形，但是常模參照評量卻比較不能達到這樣的功能。

(二) 標準參照評量 (絕對比較)：以事前決定的標準為衡斷的依據，考驗個別幼兒的知能是否達到要求的程度，用來檢定幼兒學會哪些技能或熟練的程度；以幼兒發展評量為例，這種方式的評量是把要評量的發展技巧依其難易度加以排列，而後設計相關活動，從中評量幼兒在哪些發展技巧中已會做什麼，還不會做什麼。此種評量的優點在於能確認幼兒具體的行為或能力、可利用教學活動進行評量、能配合課程與教學、能確認幼兒的進步與困難、能確認幼兒所需之輔導。總而言之，標準參照評量與教學可互相配合，並達到兒童發展能力與回饋教學活動的設計與進行，比較適合用於幼兒的教學評量。

王珮玲 (1995) 指出在於解釋評量結果該注意下列事項：(1) 以各種評量所得結果解釋幼兒行為：如果只根據單一種評量方法所得的結果來評斷幼兒的發展，可能會有偏差存在。(2) 任何一次評量只是幼兒在某個時間及情境所表現的行為：任何一次的評量只是行為的取樣，不能代表幼兒全部的行為，因為幼兒會在不同的時間、情境表現出不同的行為。(3) 避免將評量結果當作一種標記；評量過程中有一些影響評量情境的因素，很難加以控制，所以不能根據評量所得的結果來評斷幼兒的一切發展，為其貼上標籤。(4) 資料必須保密。

林淑玲 (1997) 以針對其臨床教學經驗，提出其處理教學評量結果的可行做法有三：(1) 將評量結果轉換為書面資料以進行親師溝通；(2) 提供家長參考的幼兒評量結果是以每位幼兒自己的前後表現做比較，而不是與他們比較或全班比較的；(3) 運用評量結果隨時作為進行親師座談的參考。

由上可知，評量的基本目標是要正確的獲取和應用評量結果的資料，並且解釋各種評量資料；教學評量重視評估、敘述、預測及診斷等功能，解釋評量結果會因功能及測驗目的不同而做不同分析。要能使評量結果達到預期目標的功能，評量結果的解釋須考慮下列原則：(1) 解釋評量結果必須了解評量的目的及內容；(2) 要了解評量內容所轉換的指數及常模團體的分佈狀況；(3) 要解釋受試者的教育發展水準；(4) 要解釋受試者生長及進步情形；(5) 評量結果的解釋應注意受試者在各方面的能力組型；(6) 評量結果的解釋應小心謹慎（陳李綢，1992）。

### **第三節 幼兒評量相關研究**

本節分三部分呈現國內有關幼兒評量的研究：第一部分呈現由於幼教領域對於幼兒園教師教學評量現況的研究付諸闕如，因此本節先呈現有關中小學教師教學評量研究，作為本研究之參考與學習；第二部分呈現國內幼稚園試驗新評量方式的實作研究；第三部分呈現國內政府所辦理的幼教機構評鑑中，幼兒園教師的評量實施狀況。

#### **壹、中小學教師對於評量的理解與實施現況**

以往國內探討教師實施教學評量現況的研究，主要對象以國中小學教師為大宗，大多以量化問卷調查的方式對於中小學教學評量的現況作整體性的調查（王明源，1998；陳秀美，2003；張達田，1994；楊銀興，2000）。

張達田（1994）和王明源（1998）探討國中、國小教師對於教學評量

的實施狀況，發現國內中小學教師都肯定教學評量是教學歷程中的一部份，認為教學評量最主要的目的在「評定師生的教學活動是否達到教學目標」，教學評量工作應由任課教師和學生共同參與，且大多教師的做法與理論上所述的方法相合；但國中、國小教師很少在教學前對每一單元教學編擬教學活動設計，且雖然知悉學校禁止使用坊間出版的測驗卷，但國中平日還是常使用它來評量學生學習成就，而國小教師雖然贊成與反對使用坊間測驗卷者各半，但是還是有三分之一的教師，表示學校大多數人仍在使用。不及二成的國中小教師在實施教學評量後會實施診斷性測驗，以診斷學生的學習困難所在。

楊銀興（2000）以自編之「國民小學教師實施教學評量方式」問卷，探討教師實施傳統評量（如有標準答案的各式紙筆測驗、是非題、選擇題等）與新式評量（如實作評量、真實性評量、變通評量、檔案評量、軼事紀錄等）的差異，並進一步瞭解國小教師實施新式評量對教師教學、學生學習及家長觀念影響的看法與實施所遭遇的困難，結果發現在大學課程中曾學過「教學評量」課程者及主任也比級任或科任教師好；城市地區的教師及有參加過相關評量的教師對新式評量的認知情形比鄉鎮地區及沒參加過的教師清楚；相對的，在傳統評量的逐題分析上，年紀大、年資深的教師及主任瞭解情形比年紀輕、年資淺的教師及科任或級任教師好。整體而言，平均有三分之一到二分之一國小教師樣本不了解新式評量中的檔案評量、工作分析、行為觀察、軼事記錄、檢核表應注意的原理原則之瞭解，顯示國小教師對於新式評量的認知仍有待加強。不過，此研究並未探究使用「傳統評量」之教師到底使用了哪一些評量方式，是觀察、紙筆測驗還是檢核表；也只探討了不同

背景的教師對於新式評量清楚程度的了解，但是對於教師對於新式評量哪裡不清楚、為何不清楚，並無法得知。在研究的最後楊銀興以開放式問卷的方式呈現教師樣本對傳統與新式評量的看法，表示新式評量在實施上可能遭遇的困難依序有：學生人數太多，不可能一一有實作機會；施測費時；教材份量太多沒有時間讓學生實作；評分主觀；對新式評量各項實施技術、過程、評分方式的不熟悉以及教師不願改變傳統評量的方式。此外，家長注重學習成績的觀念、質疑新式評量的評分的公正性；學生課堂秩序難以掌握；場地設施不足；教師心態無法調適、教師工作份量過重、雜務太多等都是實施新式評量阻礙的來源。

陳秀美（2003）以問卷調查探討 138 位國小英語教師的多元評量信念及實施情形，問卷內容包含教師對多元評量的認識、信念及實施情況三方面。她的研究發現國小英語教師在多元評量概念及實施要領的認知和信念均良好；實施英語科多元評量時，教師最傾向採用測驗式評量及同儕評量模式；而他們所遭遇的困難為多元評量施測費時、教師任教班級數過多且班級人數過多使實施困難、增加教師負擔等。

少部分研究以個案研究的方式深入探討教師實施教學評量情況（吳毓瑩，2003；莊明貞與丘愛鈴，2003）。吳毓瑩（2003）發現三位個案教師對於教學評量各有不同的著重點，而教師評量信念的核心會反映出評量策略的選擇：就評量方式而言，有的教師喜歡用寫的，而採用詮釋與歷程檔案的評量方式；有的喜歡以加深加廣素材刺激學生學生的思考能力，便採用實作評量；有的則強調評量結果的公正性與行為的觀察。吳毓瑩進一步發現影響教師評量信念的原因牽涉師資養成之訓練、個人能力之強弱、人格特質的傾

向、教師專業的認定、人生的經驗、對學生的期待等。莊明貞與丘愛鈴(2003)以一位國小教師為對象，了解一位個案教師身處在學校評量改革的氛圍下，運用課程統整與檔案評量、實作評量等課程與評量革新的情形，不僅改變傳統課程的組織，使學生習得知識更為活用，並藉學生檔案的結果作為更有效達到教學目標的預測。但是他們也發現評量改革仍面臨一些困難，如：所費時間與花費較大、課程能力指標的解讀與轉化困難、真實評量的規準難訂。儘管如此，莊與丘認為不可否認的是，教師實踐多元智慧學習評量有助於展現學生學習的完整認知圖像，而分科課程結構對學科課程統整與檔案評量之實施有所侷限。

在國外的研究中，McNair, Bhargava, Adams, Edgerton 和 Kypros (2003)以訪談法對於 PreK-2 和 3-4 年級兩組的教師進行實施評量的追蹤研究，目的是希望了解 PreK-2 和 3-4 年級兩組的教師如何決定所進行的教學評量形式、使用頻率、和評量工具，特別是關心教師對於評量工具形式的使用，包括多久使用一次紙筆測驗、多久寫一次觀察紀錄和如何處理所觀察的紀錄，是否使用檔案來評量，還有其教學是否受到州政府測驗的影響。研究結果顯示兩組教師對於評量工具的使用明顯地不同，如在紙筆測驗方面，高達 92% 的 3-4 年級教師使用（其中有 20% 使用自編測驗，13% 使用坊間測驗，67% 兩者都使用），而只有低於 36% 的 PreK-2 年級教師使用（其中 28% 的 PreK-2 年級會自編測驗，32% 使用坊間的測驗卷，40% 是兩者都使用）；同時 91% 的 PreK-2 年級和 98% 的 3-4 年級會以紙筆測驗做為總結性評量；在觀察法的使用方面，兩組的教師都使用觀察去觀察所有的幼兒，主要也是做為總結性評量。其中有 47% 的 PreK-2 年級與 52% 的 3-4 年級教師使用自編

的檢核表用於總結性評量；在檔案評量方面 95% 的 PreK-2 年級和 88% 的 3-4 年級的教師宣稱他們使用檔案評量，主要是為了總結性評量的使用。

綜合上述，研究者認為幾位對於國內中小學教師教學評量所實施的問卷調查之問項過於簡單、籠統，並無法深入了解教師使用教學評量的看法與態度，若是以訪談的方式，使受訪者在毫不受限情況下，充分與詳細的說明使用情況或表達意見，可彌補其回答受限的部分。但是，綜前述之國中小教學評量可見，教師或行政人員的學歷（教師類型）、目前擔任職務、任教地區、年資、大學是否修過相關課程、參加過相關研習情況等對評量工作的認知和實際做法的表現有顯著相關；吳毓瑩（2003）、莊明貞與丘愛鈴（2003）的研究也發現個人專業背景、學校脈絡、職場專業發展經驗、個人特質及家長的支持等會直接或間接地影響教師評量的實踐，值得在本研究中做進一步的探討。相較之下，國外的訪談研究，能簡單清楚地呈現出不同年級的教師對於教學評量的使用情況。因此，本研究擬採用訪談的方式瞭解不同背景（教師類型、目前擔任職務、任教地區、年資、大學是否修過相關課程、參加過相關研習情況等）的幼兒園教師其使用教學評量的情形如何？是否會有顯著的差異？困難何在？

## 貳、幼兒園試驗新評量方式的實作研究

近年來國內逐漸出現多元化的評量方式之實作研究，主要試驗的對象以工作取樣系統和多元智能之評量（光譜計畫）為主。

### 一、工作取樣系統之實作研究

呂翠夏（2000）在一所公立幼稚園針對四歲幼兒的班級試驗「工作取

樣系統」的教學評量，採用下述方式蒐集幼兒學習與發展的證據：

1. 發展檢核表與摘要報告：檢核表包括六大領域與表現指標，教師定期將軼事紀錄整理到檢核表上，學期末依據幾次的檢核作最後的勾選，依三個層次（尚未發展、發展中、熟練）來檢核，並敘寫綜合的報告，以一段文字將各領域的觀察結果摘要出來，並加上課程上如何幫助幼兒的計畫。
2. 軼事紀錄：每天教師都隨身攜帶筆紙，在幼兒之間穿梭觀察並作紀錄。軼事紀錄能幫助教師具體看見幼兒的思維與行動。
3. 教學活動評量表：教師事先設計教學活動時，同時計畫各活動蘊涵了哪些學習與發展的機會，將有關的領域與表現指標填寫於「教學活動檢核表」，此表能讓教師在主持活動時或幼兒練習時抓住短暫時間做評量紀錄，此表也提供全班幼兒能力與技巧的團體摘要，供教師了解整體的學習狀況，並據以設計適當的教學活動。
4. 作品集：包括圖畫、文字、照片以及摺紙等作品，平面的可黏貼於同樣大小的紙上，便於收藏，立體的建構如積木、黏土、廢紙工等則以照片保存。每一件作品都盡量為當時的情境留下註解，包括幼兒的想法、命名、夥伴、自發性等。每一個月底教師利用自由遊戲的時間，一次邀請一位幼兒整理他工作櫃的作品，和幼兒一起欣賞和回顧，相似的作品則與幼兒討論，選擇最特別與最喜歡的，隨著日期累積作品集顯現幼兒的發展情形。同時以開學前與期中各一次的家長座談會、日常溝通與定期刊物及期末家長個別會談的方式，讓家長參與和了解。



經過一年的嘗試，呂翠夏發現參與評量的三位教師，在經驗、教育背景和人格特質都不同，對於同一群幼兒的期望和標準歧異，需要透過經常分享軼事紀錄與案例反覆討論與印證，逐漸拉近標準的距離與取得共識。她建議從選擇與設計一個評量系統起，教師即時時刻刻應用自己的專業判斷，且結合理論、研究與專家的看法，不能仰賴個人的意見；教師更須努力尋求訓練與資源，充實自己，提升專業素養。鑑於該研究中的教師雖實作工作取樣系統，但對於評量標準卻有此見，因此本研究也將針對教師對於所使用的評量的標準一事的情況進行了解。

陳姿蘭（2001）以參與觀察的方式，探討一位國小附幼教師實施作品取樣系統的卷宗評量的歷程。由於該國小附幼原來的評量方式有些缺點（如期末要另闢時間進行評量、評量內容與課程內容不符、評量項目不夠明確、客觀性受質疑、未能確實評量出幼兒的能力、教學與評量難以同時兼顧、評量無法協助教師改進教學、無法比較幼兒前後的進步、作品的收集沒有事先的規劃與事後的分析，無法發揮真正的作用等），使得該研究的個案教師亟欲修改其教學評量方式。在嘗試幾種不同的評量方法之後，個案教師決定採用作品取樣系統法的「作品集」（檔案）作為教學評量的方法。該教師實施作品取樣系統的作品收集和分析的程序後，覺得解決了舊有評量分析時「專業能力不足」與「主觀問題」的困難，且可以同時兼顧讓教學與評量，不用另闢評量時段；在分析時，有較清楚的分析方向，較能比較幼兒的進步，不會因作品過多而眼花潦亂。同時，該研究也發現即使是幼稚園的幼兒，也具有自我反省的能力，幼兒自評的能力會隨著時間與練習機會的增加而增進。換句話說，卷宗評量發揮了「展現幼兒進步與成長的情形」、「讓幼兒參與評

量」並「與自己比較」、「展現幼兒的最佳能力與獨特特質」、「提供質性資料」以及「反映教師的教學與教室活動」的功能。

## 二、多元智能評量（光譜計畫）之實作研究

王正珠（2001）以一所國小附幼之兩位教師為對象，探究其多元智能理論於運用實際教學之歷程。個案教師運用多元智慧理論於課程教學與評量，發現在所進行的六個主題中，語文智能領域的課程活動出現最多，而音樂智能領域和內省智能領域的出現最少。至於評量方面，他們透過教學日誌、檢核表、幼兒觀察記錄評量表、闖關活動及在教學檢討會議等不同方式進行評量，讓原本教室內較片面且不聚焦的評量發展為全面而明確、且趨近真實評量之方式，教師並得由評量更全面而深入的瞭解班上每位幼兒。同時，王正珠提出影響教師運用多元智慧理論的最有利因素在於「教師的內在動機」，此外，教師對理論的認知、自我能力的限制、搭檔的共識、備課時間、研習、物質與環境的資源、人力資源、行政主管單位的支持、家長的配合等，也都影響著教師運用多元智慧理論的深度，尤其是「搭檔的共識」和「備課時間」更是幼兒園教師特有的困擾。

莊雯心（2002）是以自己所任教的台北縣國小附幼班級進行「光譜計畫」的領域活動與評量之行動研究。她在期初、期末和家長共同利用「多元智能觀察評量表」為幼兒進行評量，再據此量表所得分數，繪製出幼兒各項智能曲線圖，以了解在幼兒的成長在哪裡，學習困難點在哪裡，並透過師生共同討論決定檔案收集的內容，包括：生活剪輯、觀察紀錄、評量紀錄、學習區活動、大單元活動、親子習作等幾個部分，以了解幼兒的成長。結果她發現

以光譜計畫設計的活動與評量，確實能找出每個幼兒在智慧上的優勢，藉由收集一學期的活動單，看出幼兒成長的變化。

陳依湘（2003）探究花蓮一班公幼的兩位教師及實習教師應用光譜計畫社會領域的活動與評量，並利用光譜計畫活動中的「教室模型觀察活動」（教師透過「教室模型觀察表」，觀察幼兒操作配有小家具以及貼著同學、教師照片的木製人偶的教室模型的情況，並搭配教師填寫光譜所發展的「社會互動圖」，得出「教室模型摘要表」的分數），及「同儕互動活動」（透過同儕互動活動遊戲中對幼兒所做的「同儕互動檢核表」之觀察，可依幼兒的行為表現分為五種社會角色：領導者、促進者、獨行俠和團隊成員，及另一個角色—「轉變中的幼兒」；然後將三位教師所填的同儕互動檢核表做成統計，填入每個幼兒的「同儕互動摘要表」以檢視哪些行為符合哪些社會角色）兩種評量工具來檢視幼兒了解自我、了解他人及了解社會角色的社會智能。研究發現「教室模型觀察活動」與「同儕互動活動」兩種評量工具皆可讓該班教師了解幼兒社會智能發展的情況，由此研究可知光譜計畫所發展的評量工具和標準主要是運用過觀察與檢核表這兩種評量方式適切地搭配與運用，能測得幼兒的社會智能。

同樣是光譜計畫的試作，黃美津（2004）以高雄市某國小附幼的 18 名幼兒為研究對象，以光譜計畫語言領域為評量系統，評量內容共分為故事板活動的評量、影片報導活動的評量、週末新聞活動的評量三種，結果發現三種方法均可幫助教師了解幼兒語言能力的發展情況。

### 三、其他系統

許琇粧（2004）以嘉義市某公幼大班 30 位幼兒為行動研究對象，為改善原有幼兒學習評量方式而改採用軼事觀察記錄方法對幼兒進行觀察，並根據 High/Scope 幼兒學習評量的六大重要經驗（key experience）做分類，以作為瞭解幼兒學習狀況及教學活動的參考依據。結果發現教師經由對幼兒仔細觀察與記錄，可以聽到、看到及瞭解到幼兒的興趣所在與需求，也可透過與其他教師共同討論幼兒觀察記錄資料，檢視當日教學，分析幼兒的重要經驗，再依據對幼兒學習能力與興趣的瞭解來安排教學活動，可以提昇教師對整體教學脈絡的掌握。在研究中，她依據自己行動研究後的經驗提出對 High/Scope 幼兒學習評量的二項建議：包括在觀察紀錄表上，以代碼符號有系統界定觀察記錄內容；及可輔以錄音、照相等影音設備，解決來不及記錄的困擾，減輕觀察記錄的負擔。

蕭玉佳（2002）也是以行動研究的方式探討自己身為一位幼教教師如何從教學實務上體認「學習歷程檔案」的意義，在實際應用檔案於教學評量的歷程中，她發現因為她讓幼兒自行決定要放入檔案的作品，且幼兒學習成果的收集方式過於片段零碎，以致於學習歷程檔案無法有效作為改進課程與教學的依據，也無法窺見完整清晰的幼兒成長軌跡。但她仍認同檔案能真切反映出教師的教學風格與人格特質。因此，檔案對她而言，「紀念價值遠勝於作為一種評量方式的價值」。

上述有關評量系統的實作研究，顯示評量是有系統、有規劃、持續地進行，且使用的評量的方式不只一種多元時，不僅可看見幼兒的學習歷程與成果，以及其進步，也可改進教師的課程與教學。

### 參、國內政府所辦理的幼教機構評鑑中幼兒園教師的評量實施狀況

教育部及各縣市地方政府每年均舉辦幼稚園評鑑或托兒所評鑑，以確保幼教機構的品質，在國內各縣市幼稚園評鑑或托兒所評鑑手冊中，「評量」是教保活動的評鑑項目之一。例如「九十三年台北市托育機構評鑑自評表」中，將托兒所評鑑分為「行政組織與經營管理」、「教保活動」、「衛生保健與環境安全」三方面來進行，而「幼兒個別紀錄/評量」為教保活動的評鑑重點之一，表 2-1 呈現九十三年台北市托兒機構評鑑中「幼兒個別紀錄/評量」之項目、內容、說明與訪評的方式為例。

整理近幾年所出版的台北市公私立幼稚園教育評鑑報告中有關評鑑人員對該年幼稚園評鑑結果的評論(台北市政府教育局，1997、1998、2000、2002、2003) (詳細評鑑報告之整理見附錄一)，均發現「評量」一直被列為教學中「最弱」的一環，大多幼稚園教師在評量方面的表現差。從評鑑報告中可歸納出這幾年評鑑結果對幼兒園教師在評量方法的使用現況與問題所在的意見如下：

1. 教師近幾年來用來評量幼兒學習的方法，多以單元綜合學習評量表、幼兒學習檔案(作品集收集)、發展評量表(檢核表)、角落或學習區觀察評量等；有不少的幼稚園購買坊間的各種評量表，疊床架屋的讓教師紀錄不少重複性的內容。
1. 幼兒作品的收集沒有事先規劃也沒有系統地整理：部分教師會展示幼兒活動作品，有的期末會整體讓幼兒攜帶回家與父母分享，但較少進行作品評量。
2. 採用坊間出版的評量本，無法配合實際進行的課程活動內容來評量。
3. 評量內容與項目有些不適宜。

表 2-1

九十三年台北市托育機構評鑑自評表「幼兒個別紀錄/評量」之評量項目表

評鑑項目	評鑑內容	說明及訪評方式	評分	備註
57. 幼兒的學習評量	<input type="checkbox"/> 針對教保活動的計劃目標，定期實施幼兒的學習評量 <input type="checkbox"/> 以多元化的方式，評量、記錄、比較幼兒不同階段的學習情形	1. 查閱幼兒教學評量資料 2. 多元化的方式，指利用幼兒作品、說話紀錄、照片、錄音...等方式		
58. 幼兒的發展評量	<input type="checkbox"/> 每年對全所幼兒實施發展篩檢 <input type="checkbox"/> 定期針對全所幼兒的發展變化，進行整理討論	發展篩檢指對幼兒發展里程檢核之「台北市學前兒童發展檢核表」		
59. 幼兒的輔導與追蹤	<input type="checkbox"/> 能夠依據學習與發展的評量結果，發現需要輔導的幼兒個案 <input type="checkbox"/> 針對輔導個案，能夠實施具體的輔導策略與定期追蹤	1. 查閱幼兒個案輔導與追蹤的相關資料 2. 包括早療個案或是行為需要個別輔導的幼兒		

資料來源：出自台北市政府社會局（2004）。

而盧美貴（1999）也指出在該年度的教保評鑑中，「評量」項目應被列入「最低景氣指數」，儘管有些幼稚園提供了一大堆資訊，不是反覆性過多，就是不夠深度與省思性，其問題如下：

1. 幼稚園大致都有幼兒行為問題、情緒問題簡略輔導的個案記錄，但往往礙於專業的不足，表現在紀錄上的較為表面且粗陋的書面資料；分析行為表現背後的原因及解決問題的方法往往不夠深層。
2. 教師會做幼兒記錄，但是否依據教學目標期望或活動主題則有待專業的琢磨；其次，平常的記錄流於一般性的「流水賬」，亦即幼兒在學

習區、分組做的活動敘述，而非幼兒活動的選擇、進行過程或解決問題的分析。

3.大部分教師會展示幼兒活動作品，有的期末會整理讓幼兒攜帶回家與父母分享，但對作品評量似乎付諸闕如，顯然幼兒園教師在此方面的專業仍須加強與努力。

少數教師會深度或與幼兒討論系列作品的表現與進步情形，但大部分在幼兒活動完畢後僅做展示與簡單的分享。

由上可窺見，近幾年來幼教機構教學評量的不足，也呈現出幼兒園教師在實施教學評量的概況與問題，顯示幼兒園教師對於評量概念似乎有待釐清，評量方法有待改進。

綜合以上，可見探討幼兒教育實施評量狀況的相關研究相當缺乏，而從幼托機構的評鑑報告中，也發現教學評量始終是教保評鑑結果最需加強的一環，不過近年來有不少新式評量的實驗性研究，都顯示評量如以有規劃的系統性、持續性評量，最能發揮其功效；同時在國中、小學教師的相關調查研究中發現影響教師實施教學評量的有顯著差異情形的背景變項，可作為本研究探究幼兒園教師實施教學評量情況和體認之參考與比較。

