



從新式評量論合宜的幼兒道德發展評量

游振鵬／耕莘健康管理專科學校幼兒保育科專任助理教授

蕭立成／國立台灣師範大學化學系碩士生

一、前言

教育自Aristotle提倡智、德、體三分說，並主張以德育培養意志，發展德性便成為教育的主要目的之一（伍振鷺等，1998：285）。既然道德及道德教育是教育的重點，則道德教學與評量即為緊接而來的重要課題，因為道德教學可養成學生的品德，而評量旨在瞭解學生的道德發展。然而，德可教嗎？早在古希臘時期的Socrates就問過這樣的問題，本文亦試著問：德可評嗎？若依伍振鷺等（1998：297-298）的分法，道德教學可分成道德認知、道德行為、道德情操等層面，則幼兒道德發展評量即可評量上述層面。

至於幼兒是否可以被評量？這是一個見仁見智的議題。有學者認為幼兒不應接受評量，因為幼兒年紀小，語言表達能力有限，集中注意力的時間短暫，容易產生疲勞；且施測時常在教室跑來跑去，並問施測員一些無關緊要的問題，加上幼兒情緒不穩定，不適合被評量。此外，評量所採用的紙筆測驗，未必能評出孩童真正的能力，因而評量的結果未必能解釋孩童的發展。然而，亦有論者認為人類發展有一定模式可循，且幼兒早期發展是後期發展的基礎，「預防重於治療」，若能及早發現幼兒的問題，予以輔導治療，可避免日後產生困擾（王珮玲，2006：16）。

儘管學者對幼兒是否適合接受評量的論見不一，然並不影響幼兒道德發展評量的本質，亦即，或許傳統較偏向紙筆測驗的評量未必能解釋幼兒的道德發展；惟當代新式評

量的內涵、理論基礎均有別於傳統的量化評量，故較能瞭解幼兒道德發展的道德認知、道德行為、道德情操等面向。以下先闡述評量的理論基礎及角色定位，再探究當代新式評量之內涵，進而從新式評量討論合宜的幼兒道德發展評量。

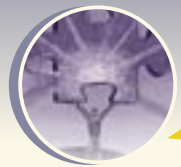
二、評量的理論基礎與角色定位

近來對於評量的定位、內涵、功能之研究日多，因而出現多元化、多樣態的評量類型。儘管評量的種類繁多，大抵可歸為傳統評量與新式／變通性評量二大類。依王逸慧（2001：74-76）的研究，傳統評量多指在學校所實施的紙筆式評量中具有標準答案之各類測驗方式，該測驗之試題類型包含選擇題、是非題、配合題、填充題、簡答題等，其施測情境是在同一時間對班級所有學生進行評量。惟傳統評量有其限制，有學者認為傳統評量常集中於容易測得的部分，較少評量學生高層次的認知能力，而偏重記憶與背誦等較低層次，造成教學僵化、學習窄化且未能真正反映學生的能力。新式評量即為試圖解決傳統評量的缺失而出現，主張以另類方式替代標準化紙筆測驗，相信這些評量方式更能真實反映學生的學習成果，並激發其高階思考能力。

綜上所述，評量可分成傳統評量及新式評量二大類，在探討新式評量與幼兒道德發展評量前，宜先敘明評量的理論基礎及角色定位如下。

（一）評量的理論基礎

針對評量方法的論戰，Hickey與Zuiker



(2002)認為是植基於過去對認識與學習的三種觀點：

1.實證主義觀點：以Skinner行為主義為理論基礎

知曉某事物 (know something) 意指與外在客觀事物的聯結 (association)，知識被假定為以階層安排，實證主義承繼化約主義 (reductionist) (以較小的要素構成複雜行為或概念) 及加成性 (additivity) (較小要素能組合到複雜實體的精確表徵中)。

其基本假定是以常規評量取向所採用的效度為主，只要假定低層次聯結能建立到高層次知識，則在標準測驗表現良好的學生，對測驗所涵蓋的領域是有見識的 (knowledgeable)。透過競爭之後，比較學生能認可的聯結數目與性質來評價教學成效，其所使用的項目回應理論 (IRT) (考生在某一測驗試題上的表現情形，可由一組因素來加以預測或解釋，這組因素叫作潛在特質或能力，且考生的表現情形與這組潛在特質間的關係，可透過一條連續性遞增的函數來加以詮釋)，具有高信度、試題互換度、量表標準化。

2.理性主義觀點：以Piaget認知與建構主義模型為理論基礎

知識是人類對環境賦予意義 (理性化) 而生，依賴人類處理與轉換資訊的能力。其基本假定是有見識的人需獲得概括的推理基模群、領域特定概念以及特定過程，因為知識係個體建構的，個體的領域知識是獨特且特異的 (idiosyncratic)。學習是透過內在賦予意義的過程來獲得高層次認知結構，所以評量係確認學習環境所建構的基模群能在不同脈絡下解決新的疑難問題。開放式的構念一回應的項目能評估較高層知識，且可用高信度來評分與管理，包含一些探究活動的表現評量也是其方式。

3.情境觀點：以Vygotsky社會文化模型為理論基礎

Vygotsky運用Hegel的假設 (勞動工具定義人類)，論證文化發明的工具將定義人類心靈，儘管他聚焦於語言，實已擴大認識與學習的範疇，包含獨特知識領域的事實、概念、數值、語言學與方法論工具，此可統稱為知識實踐 (knowledge practice)。其基本假定是有見識的個體能成功參與知識實踐，限制度 (constraint) 與賦給性 (affordance) 涉及參加活動的自然與社會面向。

社會文化觀點對學習的看法不同於協同教學，對於施教的原理則依循其觀點賦予領域知識若干特色。此觀點主張科學知識係處在科學社群的實踐中 (而非認識者或環境中)，進而導出見習制 (apprenticeship) 或社群取向 (community-oriented)，這些取向假定學校特定的領域知識居於工具的使用 (語言、測量單位和實驗室工具) 以及日常實務 (帶領及比較探查和實驗) 中。科學進展也發生於探查者在協同者間首度協商其結論，而後在學術會議及公開場合討論。所以設計問題解決活動較能實現真正科學探究的社會面向。

針對評量時情境的影響，Hickey與Zuiker (2002) 認為，實證主義的分析係假定一種機械式的世界觀，當摒除環境的關聯度時，能對學習環境有高度的控制。理性主義透過對人類處理資訊的構念而將社會群體活動加以概念化，例如Bandura (2000) 過去二十年研究自我效能，現進展到以集體效能的意念作為社會群體活動的特徵，這取向論證沒有任何實體能獨立於個體的信念與行動之外而運作，對集體活動的分析是來自個體表現評量的累積分數。Hickey與Zuiker (2002) 繼而表示，目前結合理性主義與社會文化看法的動態評量，主要聚焦於給予進入近側發展區所需的輔助。

(二) 評量的角色定位

R. Thomas (2003) 聲稱後現代學習不



僅關心教師與學生的關係，也包含行動與省察式判斷的相互遊戲（interplay）。所謂的省察式判斷有別於決斷式判斷（determining judgment），它是想像的、流動的、啟蒙的（enlightened），常在欠缺規則中運作。後現代評量因其產生知識的多元性以及欲評量的對象（object）是新的事物，而不建立在決斷式判斷（多數實作與科學典範的基礎）的觀念中。不同範圍的兩個評量概念有時可進行劃分，一種注重教學與學習的成就／成果，此類評量係屬追溯式取向（retrospective approach），另一種則強調學習與教學的過程，此類評量係屬判斷而非評價。

以往藉著觀看學生與教師投入的程度、收集資訊的方法、評量的目的來區分評量架構，其中以檢視師生在評量過程所扮演的角色尤為重要。一般而言，角色所扮演的功能至少需符合下列四個定義之一（林建福，2001）：

1. 廣義類別概念，概稱一群具有某些共同屬性的個人。
2. 扮演的身份，自己自視為該身份，也期望他人據此待他。
3. 角色占有者具有履行的功能與所做的貢獻。這訴諸於生物學類推，為了維持社會系統在永續狀態下，從社會功能的角度來使用角色。
4. 作為法律制度的一群權利與義務。

Serafini（2001）參考Short與Burke在1994年所描述的課程的三個典範：課程為事實、課程為活動、課程為探究，進而提出評量的三個典範：

1. 評量為測量

此典範緊密連結實證與現代主義對實體與知識的看法，採大規模、常模參照的測驗，課堂教師主要關心是否這些評量提供每日教學與課程決策，所要求的必備資訊。相信知識是分離學習者而存在的，學生的工作

是習得而非建構，視學習為從教師到學生的知識傳遞，相信意義處在文本（text）中，在標準測驗中只接受一種詮釋或判斷。在此典範下客觀性、標準化、與信度優先於教師與學生的投入，學生在測驗方策（program）中主要角色是接受者，少有機會進行自我評鑑與省察。設計標準測驗為比較大規模教育方策與提供公眾利益關係人（stakeholders）績效責任，而非支持教室教學，所收集的資訊以及評鑑資訊的手段通常是由教室外的權威所決斷，甚少使用檔案夾評量。

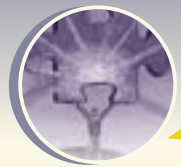
2. 評量為程序

此典範注重評量程序，認識論上，此典範和評量為測量典範有緊密關連，仍相信知識是可客觀測量的。教師與學生仍未直接包含在決策評量程序與目的之間，縱使藉著新方法收集資訊，但主要是關心資訊的報告。通常是由學區管理者初始設計許多檔案夾方案，教師僅僅收集學生作品、評定分數，提交給上級部門。檔案夾變成教室活動，而非促進師生省察及引導課堂決定的媒介。評量本身以程序為終結，教師未包含在新方法的设计中，未能提供孩童學習新的洞見。

3. 評量為探究

此典範植基於建構主義理論、學生中心學習、以及探究過程，教師為探究學習者及其學習過程而使用質性與量化評量技巧，此乃探討與詮釋的過程，用於促進學生理解、態度與素養能力。上述評量為程序的典範關心的是收集資訊的程序，僅改變教學層次；此典範旨在深化理解特定脈絡下個別的學習者，從外部權威到教師、家長、學生。

評量視為社會脈絡下的詮釋活動，相信知識是藉由個體建構，而非透過傳遞或教學技巧而獨自習得。鼓勵多重詮釋，每個學習者與不同脈絡和創造意義的世界交會，教師與學生被視為知識的積極創造者而非被動接收者。教師使用教室本位評量程序來助長學



習、引導課程決定、與家長和學生進行更有效的溝通。其中，檔案夾可作為促進學生與教師省察、自我評價的媒介，旨在理解學生興趣、能力、需要與價值。評量被視為學習過程的一部份。檔案夾試圖聚焦於學生強項（strength）而非不足，旨在提供學生概念架構與觀看世界的方式。

三、新式評量之內涵與詮釋

經由前文有關評量的理論基礎與角色定位之剖析後，可知欲完成有意義的學習不僅需形成學習意向，亦需透過評量顯現學習情形、結果和成就，傳統的評量常是由教師測量學生的學習結果，導致與所期待的教學和學習取向相抵觸，這是鼓勵教師為評量準備，而非讓評量傳達教學成效（Avalos, 2001）。

傳統的量化評量，僅能概述幼兒的學習情形，不易衡量幼兒解決疑難問題、推論及建構知識的能力，因此，發展新的評量形式與規準（criterion）是當前幼兒教育的重大課題。目前引人注目的是另類評量，以質性的形式呈現，翔實記載與幼兒學習有關的重要資訊，其所構成的文化包含四個特點（Wolf et al, 1991；Birenbaum, 1996；引自 Brian, 2001）：

- （一）教學與評量實務應是整合的。
- （二）在發展評量程序的過程中，學生是積極參與者，包含判斷其表現的規準。
- （三）評量事務的成品及過程應該被評估。
- （四）評量結果的報告應是質性輪廓的形式，而非單一分數或數量化。

Hargreaves與Earl（2002）研究29位加拿大安大略省經校長認定的教師，這些教師曾實施新課程政策、定義學習成果、另類評量報告，而提出技術、文化、政治、後現代四種觀點來審視新式評量：

（一）技術觀點

聚焦於組織、結構、策略以及相關技

能的議題，強調量度技術、發展教師評量素養與管理組織才能。技術觀點使我們注意到擬定與精鍊量度的有效形式之困難，所面臨的困難包含各種評量技能與策略的習得，此將挑戰教師的以往面相；需要使學校與家長對評量的期望能和諧一致；時間與資源的議題能輔助（或阻止）新評量實務引入校園例行公事中。惟這些實施所引發的技術議題，往往是不當使用、政治與科層干預（bureaucratic interference）或制度等因素所造成。

（二）文化觀點

文化是群體的規範、價值、信念、期望與相沿成習的活動，文化觀點檢視在學校的社會文化脈絡下，詮釋與整合另類評量的方式，將評量視為學習的一部分，關注在學習階段中，讓師生藉由評量過程的對話，發展共同的瞭解與評量規準，而能注重學生重要觀念和解決問題能力的心智表徵。真實性評量（authentic assessment）是多方向、直接、深入且依賴教師的判斷。教師控制學習綱領並正向使用回饋，鼓勵教師兼重評量與教學。Hargreaves與Earl（2002）追溯1995年安大略省的教育與訓練（Ontario Ministry of Education and Training）報告中之建議，認為此與真實性評量的某些原理若合符節：

1. 評量必須包括各種方法的使用，使學生成就的評價盡可能精確。
2. 評量、評鑑及報告是教師的責任，其須考慮個別學生的需要並與他們及家人密切配合。教師擬定有關學生學習方策與進步的決策時，應包括學生與家長等層面。
3. 評量、評鑑與報告是課程及教室實務連續且不可或缺的一環。綜言之，在真實性評量的文化觀點中，教育者不在迎合有關評量的大眾成見或假定，而是加深每個人對學習與評量議題的理解（understanding）。

（三）政治觀點



教育革新的政治觀點包括群體間的權力、權威與競爭利益之交相互作用和協商，認為所有的評量包含權力的運作，實施新式評量除了技術一致性與人際溝通外，亦包含學校與社會的意識型態及利益團體的權力爭鬥。評量有兩個取向：聚斂性評量（*convergent assessment*），明瞭幼兒是否能認識、理解或做預定的事；而擴散性評量則重視學習者的理解而非評量者的認知架構，學生需接受某些學習的責任，教師則承擔創造產生此過程的條件。教室評量另類取向的哲學為評量是學習過程的一部份，評量的判斷是協商的運作，其過程是多方向（學生—學生、學生—教師、家長與教師）。此外，政治觀點不僅呈顯傳統與另類評量的差異，也凸顯另類評量的政治風險與超量（*excessess*），尤其評估情意似乎等同於無時無刻監測學生行為，許多教師評估學生情意領域都使用考核表（*checklists*）評估學生肢體語言、作品的總量、學生對伙伴的評述、教室的專注程度、對學科的正向態度等。

（四）後現代觀點

後現代情境重塑公共教育及教育變革，而涉及下列三項評量變革的議題：

1. 複雜性（電子化衍生知識與資訊過多和混亂，挑戰教學精粹內容的假定）、分歧性（學生的文化差異增大，挑戰既有課程中的西方知識與信念）、不確定性（學校教育是多元智能的提升者與傳統內容標準的防衛者，而受到多元文化者、殖民主義等之間不確定的抨擊，在一些地方政府有時打轉於不確定與複雜性中，重視多元智能、分歧學習風格、以過程為主的課程，然在不確定性的浪潮下，有些政府則有力地肯定，迎合家長的學校圖像，躲到有效性的程序假象中回到標準考試）。
2. 電子模擬與擬像改變學生應該學的技能、重新組織教與學。市場—選擇原理侵入公

共教育，家長變成個人消費者，不再重視讓所有社會孩童共享公益。

3. 家庭結構更複雜，孩童的真正家庭與雙親常常不是那麼清楚，家庭的結構不再是單一型態，智能不再是單一、固定與可預測的，在文化分歧的教學場域中，教師難以掌握或預知學生行為。

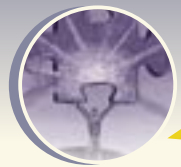
這樣的後現代觀點建基於今日複雜且不確定的世界觀，認為後現代典範下的新式評量有多重來源，是立體派畫作，以多角度重新產生實體。有時以檔案夾反應學生成就，或進行雙親與學生的三方式訪談，在初始點上未必會比傳統評量的回應更真誠。其中檔案評量可能是模擬（*simulate*）而非引發（*stimulate*）學習成就，誤使教師與學生重視形式而非實質，重視光鮮封面、優雅字形、多彩圖形的影像。當檔案夾定義學生的成就時，學生可能會為了良好的檔案夾，而非為道德價值而執行社區服務與課外活動，此將影響學習的本質。

由上可知，當評量變得複雜且較接近真實世界時，公正性的問題便日益重要，是故教師對新式評量的信心與看法是評量變革能否成功的關鍵因素。教師對評量的熟悉度也會影響教學模式的使用，Veenman等學者（2002）指出，許多有經驗的教師並不樂意採用合作學習，其原因在於：1. 擔心教室失控；2. 對部分教師缺乏信心；3. 害怕教學內容不完整；4. 害怕學生勞逸不均；5. 不熟悉新式評量的技巧。

四、從新式評量析論合宜的幼兒道德發展評量

經由以上對評量的理論基礎及角色定位、新式評量之內涵的闡釋後，合宜的幼兒道德發展評量可從下列層面來探討：

- （一）評量係發現幼兒道德發展序階的程序
傳統的評量傾向鼓勵教師為評量作準



備，而非讓評量顯現教學成效；反之，新式評量是以各種不同的方法蒐集幼兒道德發展的資料。就Kohlberg的道德認知發展理論而言，幼兒身處於不同的道德判斷發展階段，反映其理解社會事實（social facts）與選擇道德價值的社會觀點（social perspective），Kohlberg以對社會習俗尊重、瞭解的程度，或個人與社會成規發生關聯時可能採取的三個觀點，將個體的以檔案評量道德判斷發展歷程分為前習俗道德期、習俗道德期、後習俗道德期等三個發展層次及六個發展階段（引自郭靜晃，2005：542-545）。以幼兒檔案評量及道德問題評量而言，教師宜透過若干道德兩難問題引發幼兒思考、發問、討論，從而瞭解其道德發展序階，藉此蒐集幼兒個別的道德發展檔案。申言之，教師可採團體評量的方式，讓幼兒分組進行簡單的討論，各自說明自身的看法，教師從中進行個別評量。

（二）評量宜深度理解幼兒建構道德知識的歷程

幼兒道德發展評量的重點之一在探究與詮釋幼兒的道德判斷及道德認知情形，幼兒的道德知識係在社會文化脈絡下，與他人互動而逐步建構。以社會建構論而言，道德知識係幼兒在成人的引導下主動建構而成。因此，評量可視為對幼兒道德發展的詮釋活動，而質性取向的另類評量可深度理解幼兒推論及建構道德知識的能力，進而形塑幼兒的觀念架構與視野。

至於教師應如何引導、協助幼兒建構道德知識呢？依林建福（2006：216-217）之見，在進行道德教育時，教師應確定學習者是否已具備正確的道德相關概念，且須兼重道德概念的學習與道德經驗的體悟。因此，教師可以配合學習者的理解能力教導道德概念，並透過角色扮演、故事、文學、藝術等方式，讓學習者體悟道德經驗，從而豐富其

道德知識。

（三）評量規準是藉由師生對話而共同建立

道德是社會的產物，社會的狀況不同，道德強調的重點就有別（林逢祺，2004：103）。基此，道德與文化、規範、價值、信念、期望、習俗等有關，幼兒園亦可視為社會的縮影，而幼兒道德發展更是其人生全程動態發展的連續性歷程。因此，評量是幼兒道德發展的一部份，教師和幼兒藉由評量過程的討論，開展共同的評量規準，例如運用真實性評量，深化教師及幼兒對道德議題的理解。

（四）以多維面向構築幼兒道德發展評量

在目前的後現代社會中，幼兒道德發展評量可能會受到多元文化、多元智能、多樣態的學習風格、資訊媒體的發達等因素，導致評量的內涵較不明確。是以，教師宜體認到新式評量有多重類型，具變通、另類、可供選擇的意涵，教師需以多維角度再建構幼兒道德發展的實際情形，例如使用各種自我評量讓幼兒自我探索、自我省察。此外，適時運用幽默策略也是評量的相關配套措施，Mallard（引自王逸慧、游振鵬，2006：72-73）發現幽默是降低緊張情緒的有利策略，幽默的範例、測驗內容、指導語可減輕智性作業所導致的焦慮。同時，依McMorris、Boothroyd、Pietrangelo的建議，評量時可使用一些幽默技巧，然需注意這樣的考試不要有時間限制，或是有時間限制但對學生而言有充裕的填答時間時，幽默才有正向功能與助益。

五、結語

Dewey認為，兒童有高度的模仿力且容易受到暗示，惟一切活動不宜由模仿而產生，活動須由兒童本身而起，然後再提供模仿的對象以幫助兒童產生較明確的想像，使其知道何者是真正需要的。因此，模範（model）的價值在使概念更清晰而充分，不



在於使行動有所遵循（林寶山、康春枝譯，1990：82）。準此，評量並非教師對幼兒的獨斷式評價，而是二者在相互遊戲中，透過新式評量等多元方式、多重角度，瞭解幼兒的道德發展序階，並深度理解幼兒推論及建構道德知識的能力，進而深化教師及幼兒對道德議題的理解。

綜上，本文經由敘明評量的理論基礎和角色定位、剖析新式評量之內涵，從而歸結合宜的幼兒道德發展評量有四項要點：1.評量係發現幼兒道德發展序階的程序；2.評量宜深度理解幼兒建構道德知識的歷程；3.評量規準是藉由師生對話而共同建立；4.以多維面向構築幼兒道德發展評量。

參考文獻

- 王逸慧（2001）。九年一貫課程基本能力評量政策之研究。國立政治大學教育研究所碩士論文，未出版，台北。
- 王逸慧、游振鵬（2006）。「幽默」概念及其在課程教學實務應用之探究。當代教育研究，14卷2期，頁53-80。
- 王珮玲（2006）。幼兒發展評量與輔導。台北：心理。
- 伍振鶯、林逢祺、黃坤錦、蘇永明（1998）。教育哲學。台北：五南。
- 林寶山、康春枝譯，Dewey原著（1990）。學校與社會、兒童與課程。台北：五南。
- 林建福（2001）。教育哲學：情緒層面的特殊觀照。台北：五南。
- 林建福（2006）。德行、情緒、與道德教育。台北：學富。
- 林逢祺（2004）。教育規準論。台北：五南。
- 張銀富（2001）。學校行政：理論與應用。台北：五南。
- 郭靜晃（2005）。兒童發展與保育。台北：威仕曼。
- Avalos B. (2001). Global demands for quality : realities and myths in the improvement of initial teacher preparation. *Education and Society*, 19, 33-53.
- Brian, K. L. (2001). Rethinking assessment from a critical perspective. *Language Testing*, 18(4), 351-372.
- Gallagher, C. J. (2003). Reconciling a tradition of testing with a new learning paradigm. *Educational Psychology Review*, 15, 83-99.
- Hargreaves, A. & Earl, L. (2002). Perspectives on alternative assessment reform. *American Educational Research Journal*, 39, 69-95.
- Hickey, D.T. & Zuiker, S.J. (2002). A new perspective for evaluating innovative science programs. *Science Education*, 87, 539-563.
- Moni, K.B, Christina E., van Kraayenoord, C.E. & Baker, C.D. (2002). Students' perceptions of literacy assessment. *Assessment in Education*, 9(3), 319-342.
- Serafini, F. (2001). Three paradigms of assessment ; measurement, procedure, and inquiry. *The Reading Teacher*, 54(4), 384-393.
- Thomas, R. A. (2003). Learning and assessment in postmodern education. *Educational Theory*, 53, 313-327.
- Veenman, S., van Benthum, N., Bootsma, D. ,van Dieren, J.&van der Kemp, N.(2002). Cooperative learning and teacher education. *Teaching and Teacher Education* ,18, 87-103.

