

學科知識結構之評量研究—— 以「教育測驗與評量」學科知識為例*

余民寧**

國立政治大學教育學系

摘要

本研究的目的，旨在探討「路徑蒐尋網路分析」方法學所建構出來的知識結構評量模式的可行性。73 名政大教育系大三學生接受「教育測驗與評量」師資培育課程的教學後，接受以知識結構評量模式作為評量該學科知識結構的工具。研究結果顯示，該模式不僅可以一種客觀的相似性指數（即 PFC 指數），來彰顯學生的知識結構與教師的知識結構間的一致性程度，更可以區別出不同學習成就學生間的學習差異，以及解釋為什麼會有學習成就上的差異。此種評量方式，不僅讓教學評量活動變得較有意義；同時，可以圖形表徵方式來彰顯學生看不見的知識結構，並且增進評量結果的說服力；更可以直接指明知識結構的缺失及有待改進之處，以便提供後續有效的補救教學措施。最後，本研究亦針對未來教學評量與研究涵義提出討論與建議。

關鍵詞：知識結構、路徑蒐尋網路分析、學科知識

* 本論文為國科會補助專題研究案之衍生報告，作者要感謝國科會支助部分經費（專題研究案編號：NSC-87-2511-S-004-001）。

** 作者的通訊處：「台北市木柵郵政 1-433 號信箱」或「台北市文山區指南路二段 64 號 國立政治大學教育學系」。電話：(02)2938-7199（研究室）。E-mail：mnyu@nccu.edu.tw。

壹、緒論

教學與學習(instruction and learning)原是一體兩面；無論是教師的教學或是學生的學習，都需要師生雙方共同參與的一種活動歷程。教師在預定的教學目標指引下，運用各種方法，循序漸進，以期望學生的學習行為能夠隨著教學的進展而有所改變，並且是朝著既定的教學目標邁進。這種把整個教學歷程（包括所有會影響教學成果的因素在內）做有系統的處理，便是所謂的「教學模式」(teaching model)（張春興、林青山，民

73）。至今，教學模式已發展出多種範本，而且，熟悉其中幾種教學模式也已列為當今師資培育的必修課程之一。

Joyce & Weil (1996)曾將 22 種教學模式歸納成四大類：(1)訊息處理模式(information-processing model)；(2)個別化模式(personal model)；(3)社會互動模式(social-interaction model)；和(4)行為模式(behavioral model)。雖然，教學模式的種類有許多種，但卻都共同具備下列一個最簡單、最基本的教學模式雛型(Glaser, 1962)：

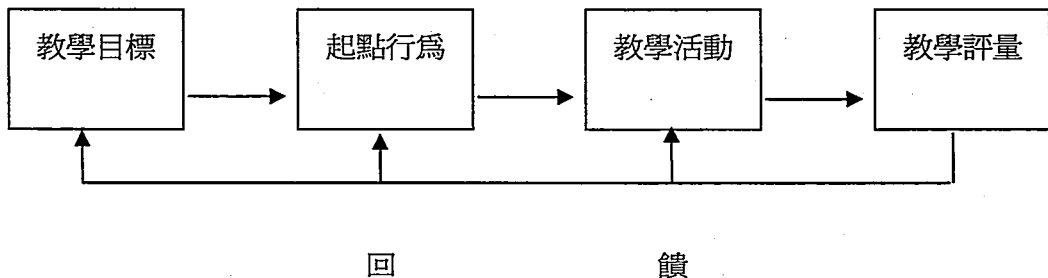


圖 1 基本的教學模式

由上述的基本教學模式可知，整個教學歷程包括四大部分：教學目標、起點行為、教學活動、和教學評量。其中，又以教學評量一環最為重要，因為它不僅可以提供回饋(feedback)訊息給前三者，更可以將整個教學歷程統整在一起，促進最大的教學與學習的效果（余民寧，民 86a；邱淵等，民 78）。

然而，無論使用什麼樣的教學或學習方法及策略，當今國內各級學校和各種重大考試（如：升高中的基本學力測驗、大學推薦甄試的學科基本能力測驗、公務人員高普考試、及各種資格檢定考試等），仍然慣用傳統式的紙筆測驗(paper-and-pencil test)（尤其是客觀測驗的選擇題）作為主要的評量方法，聽來

實在有點令人沮喪。這種紙筆測驗，雖然有其方便性，但缺點是不僅不容易測量出學生較高層次的認知能力（如：綜合與評鑑能力），更不容易測量出學生的語意思考與組織表達的能力，甚至無法表徵所習得的知識結構(knowledge structure)或認知結構(cognitive structure)為何；此外，若測驗編製不當，則更容易只測量到最低層次的認知能力（如：記憶），招致「考試領導教學」之譏，更遑論說要測量出學生的認知結構及其改變。因此，在傳統的考試領導教學情境下，呆板的考試方式便會導致無趣的教學與學習方式：教師的教學方法一成不變，既不生動又不活潑；學生只為考試而學，學習似乎變成只是一種機械式的記憶(rota memory)活動，既無聊又沒意義。這種導致學生學習動機低落、學習成果不如預期理想、教師教學方法欠缺活潑變化等現象，頗值得我們深思反省；其中，吾人相信傳統的教學評量方法應負一部份責任。

近年來，研究者試圖從教學評量觀點出發，致力於改善教學評量成效的研究，試圖將學習活動從機械化的學習轉變成為有意義的學習，將評量方法從計算總分式的傳統紙筆測驗轉換成以評量知識或概念結構為主的圖形評量方法；其中一項作法，即是把新概念或新知識的學習建立在舊概念或舊知識的既有基礎

上，將其聯結成一種有意義的網路脈絡形式，以作為新學習的開始，並以 Novak & Gowin (1984) 的概念構圖(concept mapping)作為教學評量的架構和工具，於課堂結束後進行該學科的學科知識結構評量，其成效業已獲得局部的肯定和證實：學生在學習前後的概念結構(concept structure)確實產生明顯的變化，並且，概念構圖確實也是一種有效表徵學生概念結構變化的評量工具（余民寧，民 86b；余民寧、陳嘉成、潘雅芳，民 85；余民寧、陳嘉成，民 85；余民寧、潘雅芳、林偉文，民 85；陳嘉成，民 87；陳嘉成、余民寧，民 87）。

然而，以概念構圖作為一種教學評量的工具或方法，並不表示該方法就是完美無缺、毫無限制。在評量學科知識結構的應用上，概念構圖可以算是一種相當實用與簡便的工具，然而，其最大的限制便是：評分的程序並未標準化；換句話說，拿同樣一張概念圖給不同的人來評分，可能會獲致不同的評分結果。這是由於評分時，評分者需要針對每張概念圖的組成要素（即：關係、階層、交叉聯結、舉例等項目），進行主觀的判斷，判斷該要素的存在是否是有意義、正確的聯結及正確的使用聯結語、階層分類是否正確或適當、屬於一般的聯結或交叉聯結、舉例是否恰當等問題，再根據 Novak & Gowin(1984)所發展

的計分原則進行評分；這些繁複的程序，都需仰賴評分者自由心證式的主觀判斷，想當然爾，同一張概念圖拿給不同的評分者評分，自然會出現不同的評分結果。所以，嚴格說來，概念構圖雖然可以改進教學評量的成效，作為傳統紙筆測驗外的另一種評量替代策略，但就成爲一種客觀、公正、與科學化的評量工具而言，它仍然稱不上一種嚴謹的工具。因此，我們有必要另尋新方法，或者是改進舊有的評量方法和策略。本研究即是先從改進舊有的評量方法和策略方面著手，至於發現另類新式的評量方法則留待後續研究。

由於受到「建構知識論」(constructivist epistemology)思潮的影響，近代教學評量的發展趨勢，亦從紙筆式的單科總結性評量發展到注重知識建構過程的學科(content)或主題(subject)知識的評量(Airasian, 1994; Guba & Lincoln, 1989; Hart, 1994; Stiggins, 1994; Stiggins & Conklin, 1992)。在此發展潮流下，如何兼顧足以表徵學生知識結構及其變化的概念構圖優點(Novak, 1990a, 1990b, 1991, 1993; Novak & Musonda, 1991; Wallace & Mintzes, 1990; West & Pine, 1985)，並改進其評分的主觀性缺點，但又不放棄當今學校已經爛熟使用、編製簡便、評分客觀的主流紙筆測驗方式的優點，已成爲當前「認知診斷

評量」(cognitively diagnostic assessment)學術領域所要努力解決的重點問題(余民寧, 民 84; Nichols, 1994; Nichols, Chipman, & Brennan, 1995)。本研究即爲研究者針對這個思考問題及其應用課題，蒐尋並評閱相關研究文獻後，所找到足以解決並應用認知診斷評量策略於此一問題的方法學，盼能達成拋磚引玉，吸引學界共襄盛舉的目標。

要妥善解答此一問題，首先，我們必須先了解學生是如何學習和被評量的？

持認知學習論的學者認爲，學習者的「先備知識」(prior knowledge)是決定能否進行有意義學習的主要關鍵，更是決定學習成果能否長期保留的重要因素之一(Ausubel, 1968, Ausubel, Novak, & Hanesian, 1978; Novak, 1990a, 1990b; Novak & Gowin, 1984)。而學習者的先備知識，卻從個人誕生開始，隨著年齡的增長，即已逐漸建立中。在這過程中，有的人建立起錯誤的概念或知識(即 misconception)，有的人則欠缺一個明確、穩固、與結構良好的基礎知識架構，而有的人則是建立起接近專家程度的知識結構，這些現象決定每個人學習成就的高低，以及是否有能力再接受不同深度的後續學習活動。但不管怎麼建立，這些先備知識通常都被學者專家們稱爲認知結構(cognitive structure)；在此，所

謂的認知結構，即是指個人在知覺及理解客觀事實的基礎上，在腦中所形成的一種心理結構，它係由個人過去的認知經驗所組成的（朱智賢，1989）。所以，一般學習者的學習歷程，即是不斷將新獲取的訊息（如：知識或概念）同化或調整到個人既存的認知結構裡，以改變舊的並產生新的認知結構來適應環境的歷程；亦即是認知心理學家所認為的人類的學習歷程(Mayer, 1981, 1987)。

然而，上述這種學習歷程，是一種內隱的歷程，我們不太容易具體地看見人類是否真的改變其認知結構，只能憑人類所表現出來的外顯行為間接地推測；而最常用的推測方法，即是根據「知識表徵」(knowledge representation)來推論得知其知識結構。在此，所謂的表徵，即是指將外在現實世界轉換為內在的心理事件，而以概念代替實物的歷程（張春興，民 78）。而所謂的「知識表徵」，即是指外在知識結構存在於人類心理的內在表徵，亦即是人類對知識編碼(encoding)、儲存(storage)、及提取(retrieval)的形式；換句話說，個體經由認知活動而於內在心理所產生的各種心象即是（郭姿媚，民 76；Solso, 1998）。

由此可知，人類是透過內在表徵的形式來處理訊息，如果在教學與評量時，能夠透過學生在學科知識上的表徵形式來著手研究，將能有助於我們（尤

其是教師和研究者）了解：學生是如何組織知識的？以及在學生的認知結構中，知識（或概念）與知識（或概念）之間是如何架構、關聯、和組成的？如果能夠明瞭這一點，研究者相信對學生的真正學習困難所在提供一種適當、適時的補救教學或學習輔導策略而言，將變得十分重要且有意義；這對改進教學與評量成效來說，亦具有重大的教育涵義，並且能夠發揮十分具體的貢獻。

而在分析知識結構時，我們必須先明瞭知識的組成單位是什麼？基本上而言，知識是由概念與概念間的關係所組成的有意義單位（稱作「命題」(proposition)）所構成，而不是由單獨存在的概念集合體所組成（Shavelson, 1974）。因此，認知心理學家在探索研究概念的形成、語文理解、思考運作、與記憶等主題時，也僅能由「知識結構的內在表徵」問題來探究得知。例如，Yoshino(1991)即認為知識表徵至少可以分成下列兩種方式：

(1) 命題式的表徵 (propositional representation): 即能夠使用文字敘述概念及概念間關係的表徵方式。

(2) 圖畫式的表徵 (pictorial representation): 即指圖畫式的表徵方式而言，就如認知圖一般。

但 Rumelhart & Norman(1985)則將知識表徵系統簡單的分成三類：

(1)命題式表徵系統(proposition-based representation system):即認為知識結構係由許多具有意義的最小單位——命題,所組成的一種語意網路。

(2)類比式表徵系統(analogical representation system):即認為人類是使用影像的方式來直接表徵知識結構。

(3)程序式表徵系統(procedural representation system):即認為人類是以實際作業或動作程序來表徵含有過程性的知識結構。

而 Paivio(1986)也認為大部分的影像或程序性知識,都可以使用「命題」的方式來表徵。因此,命題式表徵系統所能表徵的知識涵蓋面較廣,並且,由於命題式表徵系統發展較早,也較為完整,與中小學的學科領域知識也比較接近,所以比較受到重視。研究者亦同意 Paivio 的看法,認為應將研究重點集中在命題式表徵系統的相關議題上。

至於對知識結構和表徵之間的偵測、評量、和聯結,首先面臨的就是該如何進行的問題,不同領域的學者可能運用不同的評量方法。但是,不約而同地,學者們的焦點都聚集在如何發展一套能夠真正代表及評量學生內在認知結構的方法或工具上。因此,睽諸近二十年來的研究文獻,有關這方面問題所使用的研究方法,大致上可以分成下列幾種(邱上真,民 78; Jonassen, Beissner, &

Yacci, 1993):

(1)字的聯想(word association):在作法上,先給受試者一組概念,再要求受試者在時間限制內作自由聯想,實驗者便從這自由聯想的組合來推論受試者的知識結構。

(2)自由回憶(free recall):這個方法即是運用受試者在回憶概念組合時,其彼此間的時間差距,進行分析及推論受試者潛在的知識結構。

(3)概念卡分類(card sorting):這個方法即是要求受試者將一套概念卡片,按概念與概念間的關係疏密程度來加以分類或分組,進而分析瞭解受試者的知識結構。

(4)概念配對比較(pairwise comparison):在作法上,先給受試者一對對的概念,並要求受試者就每一對概念給予 1 到 9 的分數,用以評估受試者對每一對概念彼此間相關程度的認知,從而作為分析受試者知識結構的依據。

(5)概念構圖(concept mapping):這個方法即是在教學前後各給予受試者一組概念,然後要求受試者將這些概念用適當的聯結語把這些概念聯結起來,以成為一個概念圖,並從中偵測學習者的知識結構及錯誤概念。

郭姿媚(民 76)在分析 1970 年代以後有關知識結構與知識表徵研究的文獻指出,這方面所使用的研究方法至少有

兩大典範：第一種典範是利用「反應時間」(response time)來推論知識結構；其實，這種研究方法隱含著一個假設，即屬於同一「類別歸屬」的核心動作或核心基模，在反應時間上會比非核心動作來得快(胡志偉，民 78；Galambos, 1983；Galambos & Rips, 1982)。另外一種典範，則是利用「相似性判斷或接近性判斷」(similarity or proximity scaling)來推論知識結構；在方法上，它先要求受試者針對兩兩物件或概念間比較其相似性或接近性，然後得到一個有關這些物件或概念的相似性資料矩陣，再透過「多元量尺法」(multidimensional scaling，簡稱 MDS)與「群集分析法」(cluster analysis)來分析這個資料矩陣，從而獲得相關知識的結構表徵。

因此，拾起行為學派所揚棄的「黑盒子」，我們想要瞭解人類如何進行訊息處理的奧祕，就不得不從人們知識結構中訊息的貯存方式(即知識表徵)來著手。就如前述，「命題」乃是知識結構中的最小單位，而許多相關的命題匯集而成，即可組合成一個所謂的「語意網路」(semantic network)。當這些語意網路具有某一特定主題時，則可稱之為「基模」(schema)，而許多相關的基模聚集在一起，即可組合成所謂的「基模群」(schemata)。除了基模和基模群之外，還有其他相似的概念，如：「基架」(frames)、

「腳本」(script)、「活動基模」(胡志偉，民 78)、和「心智模式」(mental models)等，這些概念也都可以用來表徵較高層次及較複雜的知識結構(邱上真，民 78)。

近來，Solso(1998)歸納文獻分析所得，將知識表徵結構大致區分成五種不同的模式：

(1) 集群模式(cluster model)：即認為學生是將概念組合成「集群」，再以集群的方式表徵出其知識結構。常用的評量方式，是由一些不相關的字來做自由回憶，以發現這些被歸類在一起的概念集群。

(2) 集合理論模式(set-theoretical model)：即認為學生是以「集合」的方式來組織並表徵概念結構，該集合可以包括一個類別中的許多範例、屬性、和特徵。

(3) 語意特徵比較模式(semantic feature-comparison model)：即認為學生是以多向度的空間來組織知識，並以「語意特徵」來表徵該知識結構。任意兩個特徵之間係以命題的意義來進行聯結，而特徵又分成定義性特徵(defining feature)和特質性特徵(characteristic feature)兩種，可以分別聯結不同的概念而形成有意義的命題。

(4) 網路模式(network model)：即認為學生是以各個獨立的概念節點聯結在

一起，而形成一個複雜的網路關係的方式來儲存及表徵知識結構。

(5) 神經認知模式 (neurocognitive model)：即認為學生的知識是以神經網路 (neuronetwork) 的組織方式來表徵的，知識即存在於成千上萬的神經節點間的聯結。

上述這些分類，研究者認為網路模式是一個研究趨勢，而且是未來的發展方向。從早期 Collins & Quillian (1969) 針對人類的語文理解與語言行為的電腦模擬研究開始（該方法稱做「可教的語文理解者」(teachable language comprehender)，簡稱為 TLC 模式)，到 Collins & Loftus (1975) 提出一種改良的「蔓延激發模式」(spreading activation model，簡稱為 SAM 模式)，再到 Anderson (1983) 提出的「人類聯想記憶」(human associative memory，簡稱為 HAM 模式) 和「思考的調控模式」(adaptive control of thought，簡稱為 ACT* 模式) 等，都是延續語意網路的知識表徵方式而來（鄭昭明，民 83）。因此，研究者覺得有必要建立起這兩者間的橋樑，讓實驗室中的研究成果，也能夠順利地應用到一般的教室情境中。

近年來，由於人工智慧 (artificial intelligence) 及神經網路科學 (neuronetwork science) 的興起，並提出「平行分配處理」(parallel distributed

processing，簡稱為 PDP) (Rumelhart, McClelland, & et al., 1986) 的概念，認為人類處理訊息的方式是利用複雜的神經網路上成千上萬的節點，同時以一種平行分配的方式進行處理。並且，有學者結合圖形理論 (graphic theory) 和語意網路計算科學的方式，發展出用來建構和分析知識結構的「路徑蒐尋量尺化算則」(pathfinder scaling algorithm)，來測量和計算有關學生知識結構的參數 (Schvaneveldt, 1990, 1994)。

Schvaneveldt (1990, 1994) 及其同僚們 (Cooke, 1992; Cooke, Durso, & Schvaneveldt, 1986, 1994; Cooke, Neville, & Rowe, 1996; Cook, & McDonald, 1987; Cooke, & Schvaneveldt, 1988; Goldsmith, Johnson, & Acton, 1991; Gomez, Hadfield, & Housner, 1996; Gonzalvo, Canas, & Bajo, 1994; Housner, Gomez, & Griffey, 1993; Johnson, Goldsmith, & Teague, 1994; Jonassen, Beissner, & Yacci, 1993) 在美國新墨西哥州立大學 (New Mexico State University at Las Cruces) 所發展出來的「路徑蒐尋量尺化算則」，主要是用來建構和分析知識結構的一種圖形表徵工具和方法，目前並已開發出一套專門用來測量知識結構的電腦軟體程式，簡稱作 KNOT (Knowledge Network Organizing Tool)。這套軟體程式，目前多運用在測量知識結構方面的研究，以探討知識結

構與學業表現間的關係、比較專家與生手間的知識結構差異、及分析學習對知識結構的影響。另外，「路徑蒐尋量尺化算則」亦可運用在呈現及改變知識結構方面的研究，Schvaneveldt 曾將這方面的應用成果，彙編成爲一本專書，取名爲：「Pathfinder associative networks: Studies in knowledge organization」(1990)，以有系統地介紹這種方法學的理論及其實證研究結果。

根據上述發展趨勢和研究取向之文獻評閱及研究者近期針對概念構圖的研究心得，研究者認爲結合概念構圖與「路徑量尺化算則」所發展出來的方法學及電腦軟體（如：Schvaneveldt, 1994），頗適合作爲表徵與分析知識結構的評量方法或策略。「路徑蒐尋量尺化算則」是一種十分類似概念構圖的知識結構分析與表徵的網路模式，它不但可以量化表徵學生的知識結構，更可以概念圖形式描繪學生腦海中隱藏的知識（或認知）網路結構，在幫助一般教師明瞭學生於學習終了時的知識結構是什麼、學習過程中的知識結構是如何變化、以及發現學習缺陷並據以擬定適當的補救教學措施上，深具研究發展潛力和應用價值。因此，本研究的目的，即擬嘗試應用此一方法學於研究者任教的「教育測驗與評量」課程的教學評量上，企圖結合傳統的紙筆測驗方式和「路徑蒐尋量尺化算

則」表徵知識結構的方式，以改進概念構圖評量方法的計分不夠客觀及傳統紙筆測驗僅計算總分等缺點，並達成能用圖形表徵學生於課程學習後的知識結構、補充說明紙筆測驗所欠缺提供的資訊、及找出學生的可能學習缺陷所在等目標，作爲擬定補救教學之參考及評量學生學習成就的另類方法。

貳、研究方法

一、樣本

根據上述研究目的，本研究擬以 88 年度國立政治大學教育系三年級 73 名學生爲樣本，他們均是修習研究者開設「教育測驗與評量」必修課程的學生。由於這門課程是政大教育系三年級的必修科目，修課學生在一、二年級時都已修過諸如「普通心理學」及「教育統計學」等先備知識的課程，但尚未修過類似「測驗」與「評量」等方面的相關課程。因此，這些樣本在學習「教育測驗與評量」課程前的先備知識可以視爲雷同，彼此間的起點行爲亦可視爲接近或類似。

二、研究工具

(一)知識網路組織程式(KNOT)

在資料分析方面，本研究使用根據「路徑蒐尋量尺化算則」所研發出來的

結構化評量方法(Schvaneveldt, 1990)及電腦軟體程式「知識網路組織程式」(Knowledge Network Organizing Tool, 簡稱 KNOT) (Schvaneveldt, 1994), 作為主要的評量方法和資料分析工具。

(二)期末成就測驗

為了提供受試者在正常化班級教學後的成果評量, 並驗證 KNOT 程式的可行性, 研究者復根據課本教材及測驗編製原理(余民寧, 民 86a), 編擬一份含有 50 題單選的選擇題期末成就測驗。受試者在該測驗上的得分, 即代表他們在本課程的學習成果之一, 分數愈高, 即代表學習成果愈好; 該項得分並作為受試者的效標變項資料, 供作後續的統計分析之用。

(三)概念相似性判斷測驗

本研究復根據教材內容, 即以整本「教育測驗與評量」(余民寧, 民 86a) 教科書作為取材範圍, 從中挑選 20 個基本概念, 編擬成一份概念相似性判斷測驗(參見附錄)。然後, 要求受試者針對兩兩配對的概念組合進行五點評定量表式(rating scale)的判斷; 受試者若判斷認為每一配對概念間最相似的話, 則給予 5 分, 若判斷的相似程度依序遞減的話, 則分別給予 4、3、2 分, 若判斷為最不相似的話, 則給予 1 分。由於本項工作

需要進行 $n(n-1)/2$ (即 $20(20-1)/2=190$) 次的判斷, n 為概念數, 整個施測作業需要花點時間才能完成。

三、研究程序

由於研究者只任教該班「教育測驗與評量」一科課程, 沒有再任教其他班級相同的課程, 因此無法取得足供對照用的控制組班級資料, 該班僅能當成實驗班級, 所以只能根據實驗班級在期末成就測驗的考試得分作為分析用的效標變項資料。

由於這些樣本的起點行為大致一樣, 若使用前測資料當作控制變項的話, 比較沒有參考價值, 因此, 本研究擬採單一樣本後測的實驗設計方式來進行。

首先, 受試者接受為期 18 週(每週三小時)的正常化演講式教學(lecture instruction), 參考教材則為研究者出版的「教育測驗與評量」(余民寧, 民 86a) 整本教科書; 並在課程結束後, 隨即進行期末考試。該期末考試包括兩部分: 一為 50 題單選的選擇題成就測驗, 另一為針對本課程所擬 20 個概念的概念相似性判斷測驗(參見附錄)。受試者在該 50 題選擇題成就測驗上的作答得分, 即當成本研究的效標變項資料; 而他們在 20 個概念上兩兩配對間的相似性判斷資料, 即作為本研究的預測變項資料。

同時，研究者本身亦填寫一份概念相似性判斷測驗，以作為評量受試者在該判斷測驗上作答結果的參照標準。接著，在受試者作答資料收齊後，隨即進行「路徑蒐尋量尺化算則」的分析，以 KNOT 程式分別計算出他們與研究者在該 20 個概念上判斷資料間的三種相似性指數：PFC、PRX、GTD 等，以表徵他們與研究者的知識結構間的相似程度。該三種指數的用途均類似，只是在計算指數的公式不同而已，它們的數值愈高，即代表受試者與研究者間的知識結構愈相似。

最後，以該班受試者的三種相似性指數作為預測變項資料，並針對他們的期末成就測驗成績（當作效標變項）進行回歸分析，以期找出最佳的預測模式和解釋關係。同時，根據 s-p 表分析的方法學（余民寧，民 86a），找出全班中居於兩種極端的學習類型學生，分別是學習穩定型和學習異常型兩類，並比較該兩類學生在此三種相似性指數間的差異，並且，亦呈現不同成就程度學生的知識結構圖，以供作輔助解釋研究結果

的參考。

本研究所使用的方法，是在現有教務行政的規範及不更動原有班級的常規下進行，雖然在設計上有其限制所在，然而卻能兼顧到生態效度 (ecological validity) (Cronbach, 1990)，因此，研究成果比較能夠推論到一般正常教學下的班級裡。在全部受試者正式施測完畢後，即以其後測成就測驗成績和概念判斷測驗成績作為分析資料，分別檢定概念判斷的相似性指數是否具有預測學業成就的效果，或是否能產生顯著相關，並深入去比較不同學業成就的受試者們的三種相似性指數是否有所差異，以作為結論與提供建議的參考。

參、結果與討論

在受試者期末考試完畢後，旋即進行資料的登錄、轉檔、和統計分析，並進行 Tester 程式的 S-P 表分析（余民寧，民 86a），茲將 73 名受試者的三種相似性指數、期末成就測驗成績、及其學習類型等資料，陳列於表 1。

表 1 73 名受試者的三種相似性指數和期末成就測驗成績

座號	PFC	PRX	GTD	期末成績 (類型)	座號	PFC	PRX	GTD	期末成績 (類型)
St001	.300	.389	.234	70(B')	St038	.415	.463	.222	82(A)
St002	.270	.294	.260	72(A')	St039	.213	.381	.570	70(B)
St003	.239	.336	-.129	66(C)	St040	.300	.317	.170	72(B)
St004	.266	.225	.143	74(B')	St041	.275	.420	.461	70(B)
St005	.122	.154	.015	70(B')	St042	.226	.233	.050	66(C)
St006	.242	.290	.169	72(A')	St043	.208	.240	.081	74(B)
St007	.329	.451	.270	76(A')	St044	.259	.280	.084	70(B')
St008	.188	.191	.016	74(B)	St045	.328	.230	.219	74(B)
St009	.271	.315	.082	72(B')	St046	.510	.200	.064	82(A)
St010	.213	.422	.003	68(B')	St047	.323	.349	.298	72(B)
St011	.343	.314	.334	82(A)	St048	.366	.351	.206	80(A)
St012	.167	.067	.135	70(B')	St049	.100	-.081	-.079	66(C')
St013	.278	.272	.341	74(A')	St050	.229	.260	.130	70(B')
St014	.155	.154	.169	70(B')	St051	.232	.264	.367	66(C)
St015	.274	.232	.172	68(B')	St052	.231	.299	.189	64(C')
St016	.200	.178	-.007	68(B')	St053	.143	.112	-.168	66(C')
St017	.262	.366	.056	70(B')	St054	.217	.190	.022	74(B)
St018	.261	.342	.152	72(B')	St055	.286	.286	.315	72(B)
St019	.267	.186	.247	66(B')	St056	.322	.290	.278	72(B)
St020	.241	.311	.292	76(A')	St057	.307	.285	.085	68(B')
St021	.337	.286	.273	72(A')	St058	.257	.280	.199	74(B)
St022	.297	.167	.230	74(A')	St059	.266	.308	.223	72(B')
St023	.244	.411	.380	72(A')	St060	.195	.145	-.131	76(A')
St024	.343	.191	.300	78(A')	St061	.366	.377	.317	70(B')
St025	.190	.239	-.045	68(B')	St062	.159	.227	-.126	72(B)
St026	.186	.055	-.010	72(B)	St063	.145	.259	.062	68(B')
St027	.186	.055	-.010	70(B)	St064	.202	.148	.198	66(C)
St028	.410	.472	.411	82(A)	St065	.164	.195	.117	70(B')
St029	.176	.183	.065	72(B)	St066	.299	.334	.185	68(B')
St030	.200	.170	.035	66(B')	St067	.284	.348	.335	66(C)
St031	.340	.384	.322	78(A)	St068	.402	.184	.287	78(A')
St032	.158	.359	.202	64(C')	St069	.307	.365	.180	70(B)
St033	.192	.328	.071	74(B)	St070	.280	.238	.219	68(B')
St034	.333	.369	.350	76(A')	St071	.309	.230	.312	72(B)
St035	.269	.261	.010	72(B)	St072	.333	.376	.345	76(A')
St036	.260	.206	.136	76(A')	St073	.217	.371	.284	74(B)
St037	.242	.261	.190	72(B)					

接著，開始以三種相似性指數為預測變項，期末成就測驗成績為效標變項，進行逐步回歸分析 (stepwise regression analysis)，以期找出最佳預測變項和預測模式來。茲將這三種指數與期末成就測驗成績間的相關係數矩陣與

逐步回歸分析結果，陳列於表 2 與表 3。

從表 2 資料分析結果顯示可知，受試者的三種相似性指數間，及其與期末成就測驗成績間，都具有顯著的相關，可見這三種相似性指數皆適合作為預測期末成就測驗成績之用。

表 2 三種相似性指數與期末成就測驗成績間的相關係數

	期末成績	PFC 指數	PRX 指數	GTD 指數
期末成績		.620 (.000)	.246 (.018)	.278 (.009)
PFC 指數			.502 (.000)	.522 (.000)
PRX 指數				.549 (.000)
GTD 指數				

註：括弧內的數字為顯著程度。

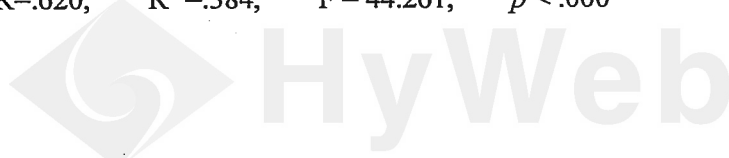
從表 3 資料分析結果顯示可知，當以受試者的三種相似性指數來預測其期末成就測驗成績時，真正具有預測力的變項僅有 PFC 指數，其標準化迴歸係數值為： $\beta = .620$ ($p < .000$)，達 $\alpha < .01$ 的顯著水準，但是，僅能解釋到期末成就測驗成績變異量的 38.40% 而已。這表示

根據「路徑蒐尋量尺化算則」所計算出的三種相似性指數中，僅有其中的 PFC 指數在本例中可以有效地預測受試者的期末成就測驗成績，其餘兩者則不行；因此，PFC 指數似乎可以被考量作為傳統紙筆式成就測驗分數外的另一種評量指標。

表 3 三種相似性指數預測期末成就測驗成績的結果

變項	B	SE(B)	β	t	p	R ² 改變量
PFC 指數	35.535	5.341	.620	6.653	.000	.384
截距	62.541	1.441		43.402	.000	

R = .620, R² = .384, F = 44.261, p < .000



接著，爲了進一步驗證上述三種相似性指數是否具有良好的區別力，特以 S-P 表分析法（余民寧，民 86a）所求得的不同學習類型學生做爲分類標準，其中，A 型爲學習穩定型學生，C 型爲學習異常型學生。由於過去的研究文獻發現（余民寧、陳嘉成，民 87，民 90），不

論就學業成就、考試成績、或與教師間知識結構的相似性指數而言，學習穩定型學生的表現均比學習異常型學生的表現爲優。因此，本研究擬進一步比較此二類型學生在這三種相似性指數上的差異，資料分析結果如表 4 所示。

表 4 學習穩定型與學習異常型學生的三種相似性指數之差異考驗

指數	組別	平均數	標準差	人數	自由度	t	p
PFC	學習穩定型(A)	.39733	.06385	6	8	6.126	.000
	學習異常型(C)	.15800	.05452	4			
PRX	學習穩定型(A)	.36400	.10142	6	8	2.037	.076
	學習異常型(C)	.17225	.19892	4			
GTD	學習穩定型(A)	.25983	.12239	6	8	2.307	.050
	學習異常型(C)	.03600	.18780	4			

由表 4 所示可知，上述三種相似性指數中，僅 PFC 指數達到顯著差異（即學習穩定型學生的 PFC 指數優於學習異常型學生的 PFC 指數），顯示 PFC 指數具有區別不同學習類型和高低學業成就的作用；從這項 PFC 指數的涵義中亦可顯示，學習穩定型學生的知識結構比學習異常型學生的知識結構更接近教師的知識結構，這兩者間並達到顯著的差異。其餘，PRX 指數與 GTD 指數則未達顯著差異，表示學習穩定型學生和學習異常型學生在此二指數上並沒有明顯

差異存在。由此可見，PFC 指數比較具有預測效用和區別力，其餘兩種指數則較缺乏。

爲了能夠進一步說明不同學習程度學生的知識結構概況，茲將教師（St000）的知識結構參照圖、表 1 中的 PFC 指數最高爲 .510 的學生（St046）知識結構圖、PFC 指數最低爲 .100 的學生（St049）知識結構圖，呈現於圖 1 至圖 3，以供比較；同時，亦將此三人的知識結構圖中，各節點（即附錄中的每一個概念）與節點間的聯結關係和權重，呈現於表 5，以

供參考判斷。

表 5 三種知識結構圖中節點與節點間的聯結關係及其權重

St000			St046			St049		
節點	節點	權重	節點	節點	權重	節點	節點	權重
1	19	1	1	17	1	1	11	1
1	18	1	2	12	1	1	10	1
1	17	1	3	19	1	1	8	1
1	15	1	4	12	1	2	20	1
1	14	1	1	19	1	2	17	1
1	10	1	2	7	1	2	14	1
1	7	1	6	17	1	2	7	1
1	2	1	7	19	1	3	18	1
2	18	1	7	11	1	3	17	1
2	15	1	7	10	1	3	16	1
2	3	1	8	10	1	3	15	1
3	19	1	9	17	1	3	13	1
3	18	1	9	16	1	4	18	1
3	16	1	9	14	1	4	11	1
3	15	1	9	10	1	5	7	1
3	13	1	10	14	1	6	19	1
3	11	1	10	11	1	6	18	1
3	10	1	11	19	1	6	13	1
3	9	1	11	14	1	7	12	1
3	8	1	12	18	1	9	16	1
3	7	1	12	15	1	10	12	1
3	5	1	12	14	1	11	17	1
4	10	1	13	15	1	11	12	1
4	7	1	13	14	1	12	19	1
5	20	2	14	20	1	13	15	1
5	11	1	14	19	1	13	14	1
5	10	1	14	18	1	14	19	1
5	7	1	14	17	1	16	19	1
6	17	1	14	16	1	16	18	1
7	16	1	14	15	1	19	20	1
7	11	1	15	20	1			
7	10	1	15	18	1			
8	20	2	16	19	1			
8	16	1	16	18	1			

9	20	2	16	17	1	
9	10	1	18	20	1	
10	20	2	18	19	1	
12	14	1				
13	16	1				
15	19	1				
15	18	1				
16	20	2				
17	20	2				
17	19	1				
18	20	2				
18	19	1				
19	20	2				
鏈結總數：47			鏈結總數：37			
			精確鏈結數：11 (.30)		精確鏈結數：7 (.23)	
			缺失鏈結數：36 (.97)		缺失鏈結數：40 (1.33)	
			多餘鏈結數：26 (.70)		多餘鏈結數：23 (.77)	

註：括號內的數字為百分比值。

由圖 1 至圖 3 及表 5 所示可知，依據點對點之間的集合理論方法 (set-theoretical) 所計算出的 PFC 指數，當然可以顯示出學生 St046 比 St049 的知識結構更接近教師 St000 的知識結構；因為，從表 5 中的百分比數據可知，學生 St046 比 St049 具有較多與教師 St000 相同的點對點之間精確的鏈結關係、較少缺失的鏈結關係 (亦即教師 St000 具有，但學生 St046 和 St049 不具有的聯結關係)、和較多多餘的鏈結關係 (亦即教師 St000 不具有，反而是學生 St046 和 St049 具有的聯結關係)。其中，精確的鏈結關係顯示出兩個結構之間的一致性程度，缺失的鏈結關係反應出學生的學習未達

精熟的程度，而多餘的鏈結關係則顯示出學生的學習產生錯誤的程度。然而，光是由鏈結次數來看，會因為缺乏比較的基準而有失客觀，因此，若改以百分比值來看，不論是精確鏈結、缺失鏈結、和多餘鏈結的百分比值，學生 St046 的該三項數值各為 .30、.97、和 .70，而學生 St049 的該三項數值則為 .23、1.33、和 .77；由此可見，相較之下，學生 St046 比 St049 更具有接近教師 St000 知識結構的一致性 (.30 > .23)、較少的不精熟程度 (.97 < 1.33)、和較少的錯誤程度 (.70 < .77)。這也就是為什麼看起來圖 2 會比圖 3 更接近圖 1 的理由所在，同時，這也可以說明學習穩定型學生 (如 St046)

會比學習異常型學生（如 St049），在一般期望結果的原因所在。
一般正常學習之後，會比較接近教師教學

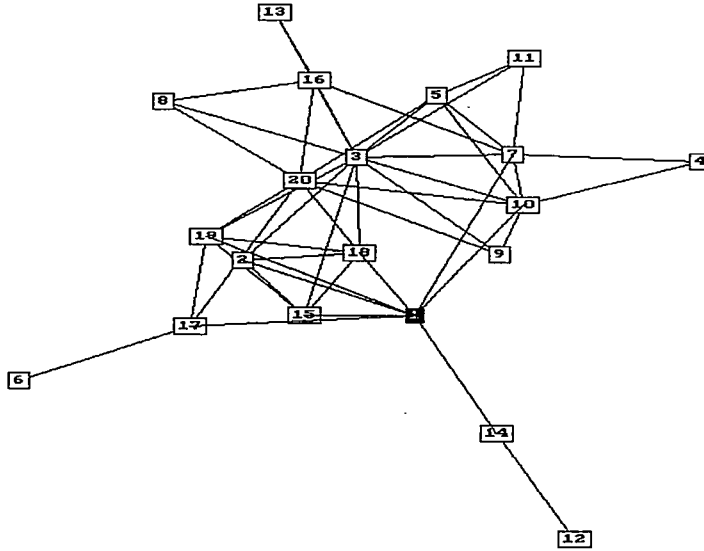


圖 1 教師(St000)的知識結構參照圖

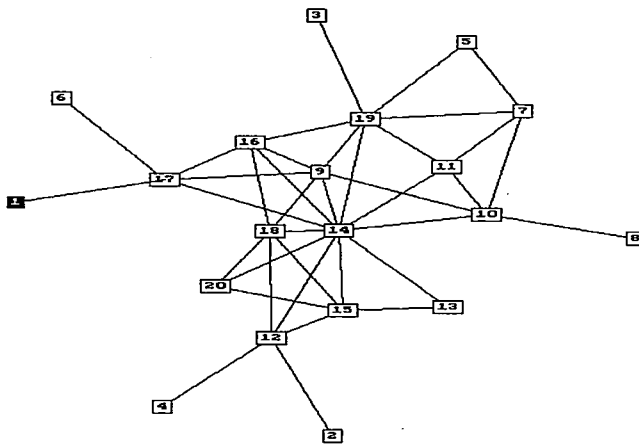


圖 2 學生(St046)的知識結構圖

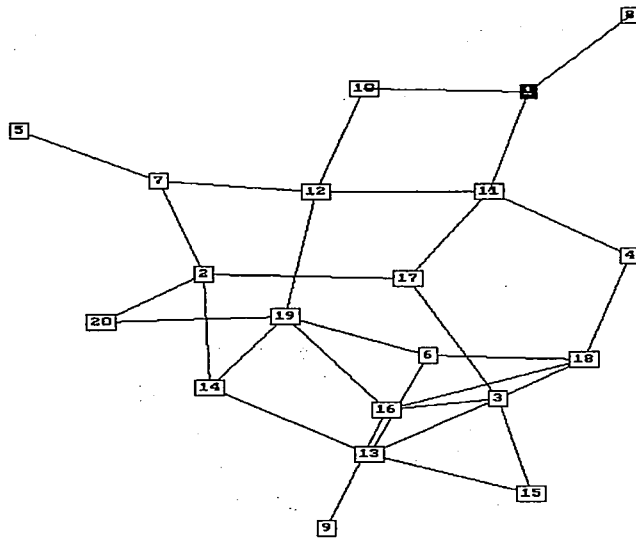


圖3 學生(St049)的知識結構圖

綜合上述的發現與說明，傳統式的紙筆測驗有它的方便性和優點，但是，本研究所用的知識結構評量模式亦具有發展潛力。若以譬喻來說，傳統式的紙筆測驗好比是一種「見樹不見林」的評量方法，而本研究所使用的知識結構評量模式則比較是屬於一種「見林不見樹」的評量方法，各有其優點及適用對象和時機。如果評量的目的只是在了解個別概念已經學會的總數量的話，傳統式的紙筆測驗就已經夠用了；如果評量的目的不僅止於此的話，那麼，搭配路徑蒐尋網路分析方法學的使用，以圖形表徵及網路分析技術所建構的知識結構評量模式，正可以用來彰顯看不見的個別概念間的結構關係，補充詮釋紙筆測驗結

果所得資訊不夠充分的缺失，它同時也是本研究努力發展的目標。

肆、結論與涵義

當今各級學校所慣用的評量方法，仍然是以傳統的紙筆測驗為主，尤其是使用客觀測驗題目（如：選擇題、是非題、填充題等），來作為評量學生在有關認知方面的學習成就。然而，紙筆式的客觀測驗，常被批評為多半只能評量到學生較低層次的認知能力（如記憶、理解層次）而已，即使是今年度首度實施的國中升高中之「基本學力測驗」，雖然號稱命題活潑、創新、彈性、千變萬化，但若仔細進行試題的命題分析和雙向細

目表的分析，將可發現該測驗的試題也多半只評量到理解、應用、分析層次的認知能力而已，對於更高層次的認知能力（如綜合、評鑑能力）的評量仍力歛未殆。

當然，學校的教學評量並不以此為滿足，學校的教學評量研究重點，也著重在開發各種不同的評量策略或模式，試圖去評量較高層次的認知目標，否則，光是強調使用評量較低層次認知能力的多數傳統紙筆測驗方式，不僅會招致「考試領導教學」之譏，更會誤導學生認為「學習只是在記憶」而已，而使學習變成一件多麼無趣又沒意義的活動，最終導致學生採取避之唯恐不及的行為。因此，在教學與評量上，我們必須嘗試做一些改變。

本研究即根據此理念，從「路徑蒐尋網路分析」(Pathfinder network analysis)的觀點，嘗試以圖形表徵及網路分析的技術和方法，將它應用到國內任何一門學科課程（本研究選定以「教育測驗與評量」課程為例）的教學評量上，以客觀的指數分析方式來彰顯這種評量模式的可行性，並探索這種知識結構的評量模式和技術做法是否可以作為傳統紙筆測驗方法外的輔助工具。

綜合上述分析結果與討論，本研究可以結論如下：以「路徑蒐尋網路分析」方法學所建構出來的知識結構評量模

式，確實可以應用到諸如「教育測驗與評量」課程的教學評量上，當成是另一種新式的評量方法，以評量較高層次的認知能力；它不但可以一種客觀的相似性指數（即 PFC 指數），來彰顯學生們在經過學習後，其知識結構與教師的知識結構之間的一致性程度，同時，亦可區別出不同學習成就學生間的學習差異，並解釋為什麼會有學習成就上的差異所在。這種評量方式，不僅讓教學評量活動變得較有意義，同時，因為可以圖形表徵方式來彰顯學生看不見的知識結構，因此使得評量結果更具有說服力；此外，因為知識結構圖可以直接指明知識結構的缺失及有待改進之處，因此，後續的輔導工作將可提出對症下藥的補救教學措施，讓教學、評量、與補救三者能夠結合成一體，以促進學生的學習成效，並提高學生的學業成就。

經由上述結論可知，本研究至少對教學及評量具有下列的啓示或涵義：

（一）學習應是一種概念聯結與概念組織的結果

認知性的學習，不僅是學習者不斷地作有意義的知識建構與重組的過程，更是學習者主動建構自己的概念聯結與概念組織的結果，如此，才能在有意義的學習過程中獲得持續性的效果。因此，評量方法的應用，必須要能夠彰顯

出這樣的學習涵義和結果，才足以稱得上是一種適當的評量方法；而本研究亦嘗試進行的知識結構評量模式，即是具有如此特性的一種新式評量方法。

(二)知識結構評量模式可以增進評量結果的說服力

以「路徑蒐尋網路分析」觀點所構築的知識結構評量模式，不僅可以彌補傳統教學評量方式的不足，更可以作為一種新式的評量策略。知識結構評量模式和知識結構圖的使用，不僅讓評量結果得以客觀指標來呈現而已，更可以利用圖形表徵方式，來呈現一般人看不見的知識結構，彰顯教師評量技巧的專業性，並增進評量結果的說服力。隨著教學與學習的進展，經由這種評量方式的多次實施，不僅可以洞察學生知識結構變化的發展趨勢，更可以即時掌握學生什麼時候產生錯誤概念、什麼地方產生錯誤聯結、以及什麼情況產生錯誤推理等訊息，以便利教師適時診斷學習的進步情形，對症下藥提出匡正與補救教學。

(三)研發本土化的知識結構評量工具是未來趨勢

知識結構評量模式雖經初步證實是一種有效的評量方式，但其應用價值仍有待各種推論效度的證據來支持。尤其是，目前本研究所使用的分析工具，是

來自國外的 KNOT 程式，欠缺本土性的分析工具可資使用。況且，本研究對 PFC 指數涵義所進行的分析和補充詮釋，是原本 KNOT 程式功能所沒有的；至於，要廣泛應用到成就測驗上慣用的二元化計分資料，亦是 KNOT 程式所無法提供的。因此，未來的研究方向，除了宜朝多種學科的應用推廣之外，更應該積極研發適用於本土的分析工具，才能讓這種新式評量模式的優點得以發揮。

參考書目

- 朱智賢(民 78)。心理學大辭典。北平：北京師範學院。
- 余民寧(民 84)。認知診斷測驗的發展趨勢。教育研究，45 期，14-22 頁。
- 余民寧(民 86a)。教育測驗與評量——成就測驗與教學評量。台北：心理。
- 余民寧(民 86b)。有意義的學習：概念構圖之研究。台北：商鼎。
- 余民寧、陳嘉成(民 85)。概念構圖：另一種評量法。政大學報，73 期(上冊)，161-200 頁。
- 余民寧、陳嘉成(民 87)。排序理論在概念結構評量上的應用。政大學報，76 期，17-48 頁。
- 余民寧、陳嘉成(民 90)。領域知識結構之評量研究——以「垃圾分類處理」領域知識為例。政大教育與心理

- 研究, 24 期(下冊), (出版中)。
- 余民寧、陳嘉成、潘雅芳(民 85)。概念構圖法在測驗教學上的應用。測驗年刊, 43 輯, 195-212 頁。
- 余民寧、潘雅芳、林偉文(民 85)。概念構圖法：合作學習抑個別學習。政大教育與心理研究, 19 期, 93-124 頁。
- 邱淵、王鋼、夏孝川、洪邦裕、龔偉民、李啞玲等譯(民 78)。教學評量。台北：五南。
- 邱上真(民 78)。知識結構的評量：概念構圖的發展與試用。特殊教育學報, 4 期, 215-254 頁。
- 胡志偉(民 78)。經常與不常從事活動的記憶表徵。中華心理學刊, 31 卷(2 期), 91-105 頁。
- 陳嘉成(民 87)。合作學習式概念構圖在國小自然科教學之成效研究。政大教育與心理研究, 21 期(上冊), 107-128 頁。
- 陳嘉成、余民寧(民 87)。以概念構圖為學習策略之教學對自然科學習的促進效果之研究。政大學報, 77 期, 201-235 頁。
- 郭姿媚(民 76)。知識結構的計量分析。國立臺灣大學心理研究所碩士論文。
- 張春興(民 77)。知之歷程與教之歷程：認知心理學的發展及其在教育上的應用。教育心理學報, 21 期, 17-38 頁。
- 張春興(民 78)。張氏心理學辭典。臺北：東華。
- 張春興、林青山(民 73)。教育心理學(五版)。台北：東華。
- 鄭昭明(民 83)。認知心理學：理論與實踐。台北：桂冠。
- Acton, W. H., Johnson, P. J., Goldsmith, T. E. (1994). Structural knowledge assessment : Comparison of referent structures. *Journal of Educational Psychology, 86*, 303-311.
- Airasian, P. W. (1994). *Classroom assessment* (2nd ed.). New York: McGraw-Hill.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart & Winston.
- Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational psychology: A cognitive view* (2nd ed.). New York: Holt, Rinehart, & Winston.
- Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407-428.

- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247.
- Cooke, N. J. (1992). Predicting judgment time from measures of psychological proximity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 640-653.
- Cooke, N. J., Durso, F. T., & Schvaneveldt, R. W. (1986). Recall and measures of memory organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 538-549.
- Cooke, N. J., Durso, F. T., & Schvaneveldt, R. W. (1994). Retention of skills after nine years. *Human Factors*, 36, 597-605.
- Cooke, N. J., Neville, K. J., & Rowe, A. L. (1996). Procedural network representations of sequential data. *Human Computer Interaction*, 11, 529-568.
- Cooke, N. J., & Schvaneveldt, R. W. (1988). Effects of computer programming experience on network representations of abstract programming concepts. *International Journal of Man-Machine Studies*, 29, 407-427.
- Cook, N. M., & McDonald, J. E. (1987). The application of psychological scaling techniques to knowledge elicitation for knowledge-based systems. *International Journal of Man-Machine Studies*, 26, 533-550.
- Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: Harper & Row.
- Galambos, J. A. (1983). Normative studies of six characteristics of our knowledge of common actives. *Behavior Research Methods & Instrumentation*, 15, 327-340.
- Galambos, J. A., & Rips, L. J. (1982). Memory for routines. *Journal of Verbal Learning and Behavior*, 21, 260-281.
- Glaser, R. (1962). Psychology and instructional technology. In R. Glaser (Ed.), *Training, research and education*. Pittsburgh: University of Pittsburgh Press.
- Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology*, 83, 88-96.
- Gomez, R. L., Hadfield, O. D., & Housner, L. D. (1996). Conceptual maps and

- simulated teaching episodes as indicators of competence in teaching elementary mathematics. *Journal of Educational Psychology*, 88, 572-585.
- Gonzalvo, P., Canas, J. J., & Bajo, M. T. (1994). Structure representations in knowledge acquisition. *Journal of Educational Psychology*, 86, 601-616.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
- Hart, D. (1994). *Authentic assessment: A handbook for educators*. Menlo Park, CA: Addison-Wesley.
- Housner, L. D., Gomez, R. L., & Griffey, D. C. (1993). Pedagogical knowledge structures in prospective teachers: Relationships to performance in teaching methodology course. *Research Quarterly for Exercise and Sport*, 64, 167-177.
- Johnson, P. J., Goldsmith, T. E., & Teague, K. W. (1994). Locus of the predictive advantage in pathfinder-based representations of classroom knowledge. *Journal of Educational Psychology*, 86, 617-626.
- Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Joyce, B., & Weil, M. (1996). *Models of teaching* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Mayer, R. E. (1981). *The promise of cognitive psychology*. New York: Freeman.
- Mayer, R. E. (1987). *Educational psychology: A cognitive approach*. New York: Harper Collins.
- Nichols, P. D. (1994). A framework for developing cognitively diagnostic assessment. *Review of Educational Research*, 64, 575-603.
- Nichols, P. D., Chipman, S. F., & Brennan, R. L. (1995). *Cognitively diagnostic assessment*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Novak, J. D. (1990a). Concept maps and vee diagrams: Two metacognitive tools to facilitate meaningful learning. *Instructional Science*, 19, 29-52.
- Novak, J. D. (1990b). Concept mapping: A useful tool for science education. *Journal of Research in Science*

- Teaching*, 27, 937-949.
- Novak, J. D. (1991). Clarify with concept maps. *Science Teacher*, 58, 44-49.
- Novak, J. D. (1993). Human constructivism: A unification of psychological and epistemological phenomena in meaningful making. *International Journal of Personal Construct Psychology*, 6, 167-193.
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge, London: Cambridge University Press.
- Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal*, 28, 117-153.
- Paivio, A. (1986). *Mental representation: A dual coding approach*. New York: Oxford University Press.
- Roid, G. H., & Haladyna, T. M. (1982). *A technology for test-item writing*. Orlando, FL: Academic Press.
- Rumelhart, D. E., McClelland, J. L., & the PDP research group (Eds.) (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: Bradford.
- Rumelhart, D. E., & Norman, D. A. (1985). Representation of knowledge. In A. M. Aitkenhead, & J. M. Slack (Eds.), *Issues in cognitive modeling* (pp. 15-62). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schvaneveldt, R.W. (Ed.) (1990). *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex.
- Schvaneveldt, R. W. (1994). *Knowledge network organizing tool* (PCKNOT Version 4.2). Las Cruces, Mn: Interlink.
- Schvaneveldt, R. W., Dearholt, D. W., & Durso, F. T. (1988). Graph theoretic foundations of Pathfinder networks. *Computers and Mathematics with Applications*, 15, 337-345.
- Schvaneveldt, R. W., Durso, F. T., & Dearholt, D. W. (1989). Network structures in proximity data. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol.24, pp.249-284). New York: Academic Press.
- Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., Cook, N. M., Tucker, R. G., & DeMaio, J.

- C. (1985). Measuring the structure of expertise. *Inter-national Journal of Man-Machine Studies*, 23, 699-728.
- Schvaneveldt, R. W., Durso, F. T., & Mukherji, B. R. (1982). Semantic distance effects in categorization tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 1-15.
- Shavelson, R. J. (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology*, 63, 225-234.
- Shavelson, R. J. (1974). Methods for examining representations of subject-matter structure in a student's memory. *Journal of Research in Science Teaching*, 11, 231-249.
- Solso, R. L. (1998). *Cognitive psychology* (4th ed.). Boston, MA: Allyn & Bacon.
- Stiggins, R. J. (1994). *Student-centered classroom assessment*. New York: Macmillan.
- Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom as-sessment*. New York: State University of New York Press.
- Wallace, J. D., & Mintzes, J. J. (1990). The concept as a research tool: Exploring conceptual change in biology. *Journal of Research in Science Teaching*, 27, 1033-1052.
- West, L., & Pine, L. (1985). *Cognitive structure and concept change*. Orlando, FL: Academic Press.
- Yoshino, R. (1991). A note on cognitive maps: An optimal spatial knowledge representation. *Journal of Mathematical Psychology*, 35, 371-393.

89/12/ 8 收件

90/ 3/24 修改

90/ 6/28 接受

附 錄

下列 20 個概念，請您依照它們概念間的關聯性，進行兩兩配對的判斷。判斷時，請用五點評量方式判斷之；即愈關聯者，給予 5 分，關聯程度依序下降者，則依序分別給予 4、3、2、1 分不等。在作答時，請依據您的直覺作判斷，不要思考太久。

- | | | | |
|------------|----------|-----------|------------|
| 1. 成就測驗 | 2. 雙向細目表 | 3. 選擇題 | 4. 效標關聯效度 |
| 5. 鑑別度 | 6. 常模 | 7. 試題分析 | 8. 注意係數 |
| 9. 再測信度 | 10. 測驗分析 | 11. 難度 | 12. 教學目標 |
| 13. 反應組型 | 14. 教學評量 | 15. 命題原則 | 16. 試題差異係數 |
| 17. 常模參照測驗 | 18. 測驗編製 | 19. 誘答力分析 | 20. 測驗理論 |

Assessment of Content Knowledge Structures: An Example for Educational Testing and Assessment

Min-Ning Yu
Department of Education
National Chengchi University

Abstract

The purpose of this paper is to investigate the possibilities of applying the Assessment Model of Knowledge Structures to the course of "Educational Testing and Assessment". 73 subjects participated in this study. Results showed that the model can be a potential tool to be used in instructional assessment with a quantitative index (i.e., PFC index) to represent the degrees of similarity of knowledge structures among instructors and students. In the meantime, the Assessment Model could also discriminate and interpret the differences of students' learning achievement. And because of using a graphic representation, the model can make students' content knowledge structures visible, identify wrong links and missing links in the structures, improve the convincing power of assessment results, and suggest the potential value of remedial instruction. Discussions of the results and suggestions for future instructional assessment and research projects are proposed.

Key words : knowledge structure, pathfinder network analysis, content knowledge