

第二章 文獻探討

本研究希望藉由文獻探討的歸納思索，建立理論基礎與發展系統架構時的理念基礎。在建立系統架構及研究架構前，必須先對網路教學系統、教材製作原則、建構主義、知識工程、本體論、網路測驗、測驗回饋、中文斷詞、測驗理論、電腦測驗、評分效力等主題之文獻進行概念性探討。

第一節 網路教學

隨著資訊科技的快速發展，各種應用於教育上的名詞不斷推陳出新，例如：電腦輔助教學(Computer Assisted Instruction, CAI)、智慧型電腦輔助教學(Intelligent Computer Assisted Instruction, ICAI)、電腦輔助學習(Computer Assisted Learning, CAL)、網路教學(Web-Based Instruction, WBI)、線上學習(On-line Learning)、隨選學習(Learning on Demand)、遠距教學及電子化學習(e-Learning)等。

Rosenberg(2001)指出網路教學是透過網際網路傳遞大量經過整理的解決方案，並藉以促進知識的獲得與提升學習者的績效表現。因此，網路教學是透過網際網路技術的教學方式，可即時更新、散佈、存取及分享教學內容或資訊。美國訓練發展發展協會(American Society for Training & Development, ASTD)認為網路教學包含網路化學習(Web-Based Learning)、電腦化學習(Computer-Based Learning)、虛擬教室(Virtual Classrooms)、數位式合作(Digital Collaboration)的廣泛應用及相關過程。

而依照教育部電算中心的「專科以上學校開辦遠距教學作業要點」第五條第四項規定，以電腦網路實施之非同步(網路)教學之基本需求規格，包括教學系統功能、教師教學方式及教材製作等三部分。陳年興等人(民 86)則指出網路教學環境依據教學者及課程軟體特徵，而將網路教學區分為三類：

- 一、線上沒有教學者，也沒有教學課程軟體，學習者是透過網路交換資訊、互動與討論來進行學習。例如：線上討論區、BBS、News 等。
- 二、線上沒有教學者，但有課程軟體的學習環境，例如：WebTitle。
- 三、线上有教學者，也有教學課程軟體的學習環境，此類學習環境又稱為遠距教學，並可依其教學模式細分為：
 - (一) 同步教學模式：強調一種即時的資訊傳輸，亦即所有參與學習的人員必須在同一時間透過傳輸系統產生互動，一般又稱之為即時群播視訊教學。
 - (二) 非同步教學模式：指教師和學生間的教學活動並非同時進行，而且通常會將學習內容長時間放置在教學網站上，使用者不受時間與地點限制，隨時可上網學習新知。實施的方式上也可分為課程隨選視訊及全球資訊網教學網站兩種，其中全球資訊網教學網站則屬於非同步網路教學管理系統。
 - (三) 整合式教學模式：將上述兩種教學模式整合而成的學習環境。

林敏慧等人(民 90)認為傳統教學是一個以教師為中心，學習者為被動角色的教學方式；網路學習則是以學習者為中心，學習者可以不受時空限制，並主動依據自己的方式抉擇欲學習之進度與內容。吳明隆(民 87)及 Bannan-Ritland et al.(1998)認為網路教學(WBI)並非只是將教材放置在網路上而已，因為許多課程在 Web-based 要素中，只是一種教學或非教學方式的「簡易並排」(Juxtaposition)呈現而已，並無法提供完整的網路線上教學

活動。因此如何利用 WWW 線上資源，有系統而完整的規劃設計教學活動以建構以學習者為中心的 WBI 學習環境，才是 WBI 整合式學習環境的核心所在，這也是 WBI 與 WWW 的主要差異。

目前網路教學系統及環境的建置、網路教學教材設計及網路教學績效評量等，也成為網路教學技術上經常被討論的議題。因此學者普遍認為教學網站應提供友善的學習介面，以降低學習障礙及負擔，系統也需要能記錄並蒐集學習者之學習歷程(李長峰等人, 民 91; Sugrue & Kobus; 1997, 藍月蓮, 民 92)。然而部分教學網路網站只是將傳統課程教材變成網路教材，並未有效利用網路互動特性，此外，由於互動性網路教學網站經常缺乏整合性及整體規劃，教學者反而需要投入大量時間精力去建構及維護教學環境(張紹勳、林騰蛟, 民 89)。

McGreal(1998)指出，一個理想的網路學習環境應具備線上課程、線上測驗、虛擬教室、教學管理、學習工具等五項要素。而林奇賢(民 86)也指出，一個理想的網路學習環境應該具備多媒體線上課程區、多媒體線上測驗區、公共討論與公告區、個人信箱、學習資源區及教材管理區。

藍月蓮(民 92)在網路上提供教學資料的方式相當多種，包括線上提供電子書、課程內容綱要、投影片、隨選視訊等，然而這些教材則需要教師事先製作完成，以提供學習者使用。而網路教學本研究中，則泛指透過網路進行教學或學習的方式。

Liu 等人(2000)認為對於網路學習環境中的學習者而言，線上課程是進行網路學習時的主要依據，所以課程安排及教材設計就顯得非常重要。而教材又分為教學教材及自學教材，楊家興(民 89)認為自學式教材因為沒有老師講解，所以內容必須比教學教材更加詳細清楚，此外，自學式教材通常都會提供明確學習目標、注意學習者可能的困難、依學習者需求決定教材、強調自我評量活動、使用簡單的語法詞句、較短的課程片段、需要學習者主動回應、使用大量的範例圖表、尋求學習者的評鑑回饋等。因此，自學式教材主要是以「學」的角度來規劃教材的設計，而非從「教」的角度來設計。

而當學習者進行網路學習時，課程教材必須經過規劃與設計。因此 Dick 與 Carey(1990)將系統化教學設計歸納為五個階段：

- 一、分析階段：確定教學目標、分析學習者特性、分析教學內容。
- 二、設計階段：訂定單元學習目標、設計教學活動。
- 三、發展階段：編寫教材、製作教學媒體。
- 四、執行階段：實施教學。
- 五、評鑑與修正階段：針對教材進行評鑑與修正。

此外，Gagne 等人(1992)認為一個有效的教學活動應該包含注意力集中、告知學習目標、追憶學前經驗、呈現教材、指導學習、實際演練、提供回饋、評量學習成果及強化學習保留和遷移等九項事項。然而，在線上學習環境中，學習者若使用自學式線上教材，則通常無人從旁給予學習指導並給予回饋，因此如何在線上學習系統中保留這些教學活動將是值得重視的議題。

第二節 建構主義

Piaget 認為知識不是客觀或主觀的東西，而是個體與環境交互作用的過程中逐漸建構的結果。所以學習者本是就是自我知識的建構者，其角色也由被動的聽眾，轉為積極的參與者(張史如，民 86)。

此外，建構主義論者認為學習是認知的建構過程，因此學習者需要藉由環境及過去經歷的記憶與理解方式去處理資訊。此外，由於每個人成長背景及學習經驗不同，所以個人知識建構的過程與途徑也不同，傳統教學中，教師很難提供不同的資訊滿足不同的學生需要，但全球資訊網以其特有的超連結能力，能將豐富的資料串連，讓學生可依照自己的興趣及學習方式，獲取相關的知識內容(Woolfolk, 1993)。

建構式學習理論強調知識是經由學習者主動探索和發現的方式所建立，因此建構式學習的成功關鍵，在於引起學習者對學習目標的興趣，如此才會有動機主動找尋所需的相關資料，以建構知識和解決問題，並達到良好的學習成效(張家倩、楊國德，民 87)。

Shih 與 Chen(2000)認為以建構主義為基礎的網路學習環境設計原則，應包括包括以下七個項目：

- 一、提供知識建構：學習者可以自行選擇所要學習的課程及學習順序。
- 二、具有學習社群的情境：學習者之間可以互動，並經由互動檢視自己的觀點和其他同儕是否有差異。
- 三、提供容易使用的環境：系統必須讓老師及學習者方便使用。
- 四、提供動態搜尋：由於知識是經由動態搜尋及發現而建立的，因此網路學習環境必須讓學習者可以動態搜尋資訊。
- 五、提供觀察其他同儕的作業狀況及學習活動。
- 六、提供可合作學習的環境及空間。
- 七、記錄學習者的學習歷程及學習狀況。

第三節 本體論

本體論最早是由 Bunge 於 1977 年所提出的哲學概念應用，Bunge 指出本體論是關於真實世界的基本特性哲學理論，其用來探討真實世界中存在的實體，並進行有系統的說明，而此說明並無特定的語言加以表示。近年來，本體論在人工智慧領域中被廣泛的使用，包括在自然語言處理(Natural Language Processing)、知識分享(Knowledge Sharing)及知識工程(Knowledge Engineering)中，去描述知識構成要素之間的關係(曾憲雄等，民 91)。此外，在人工智慧的領域中，本體論的發展主要是用來使知識分享和再用更為容易(Fensel, D., et al., 2001)。

本體論的研究大致上可為兩大方向，一是針對特定問題領域建立大量的本體論，例如：建立某領域詞彙的本體論，用於該領域自然語言的剖析；另一方面則是研究本體論的建構方法與表示方法，例如利用可延伸標記語言(eXtensible Markup Language, XML)或是資源描述格式(Resource Description Format, RDF)來描述本體論(曾憲雄等，民 91)。

Uschold and Gruninger(1996)認為本體論是一種被認同的概念性分享，所謂的概念性分享指的是模組化特定領域知識的概念架構、代理人內部操作的溝通協定規格及有關特定領域

的理論表示法。而 Guarino(1998)認為本體論是一個邏輯理論(Logic Theory)，用來說明一系列字彙的特定意涵(Intended Meaning)。為了達成描述一個概念化範疇(Conceptualization)的目的，使用一系列的邏輯語言來表達，然而這個語言基於其本體論行為(Ontological Commitment)的限制，必須從邏輯語言中找到適當的特定模型(Intended Models)來說明概念化範疇的特定意涵。

就知識的類型與概念而言，Winograd 與 Flores(1994)認為知識主要來自經驗、推論與社交等三個來源。「經驗」是指人類所具備的知識是由從前的經驗而來的；「推論」是指知識經由人類對現象的歸納或邏輯推理而產生；「社交」是指知識經由人際互動中的所見所聞中學習或模仿而來。

Holsapple 與 Whisten(1983)將知識分為經驗知識(Empirical knowledge)、公式知識(Formula Knowledge)、程序知識(Procedure Knowledge)、推論知識(Derived Knowledge)、類推知識(Assimilation Knowledge)、表現知識(Presentation Knowledge)等六種。Holsapple 與 Whisten 在所發展的知識管理系統中，讓知識庫內包含以上六種知識，再透過問題處理系統和使用者溝通。而問題處理系統則包含了分析語言(Linguistic Analysis)、控制環境(Environment Control)、資料管理(Data Management)、特定問題的詢問(Ad-hoc Inquiry)、統計分析(Statistical Analysis)、試算表(Spreadsheet Analysis)、報表管理(Printed Forms Management)、輸出管理(Screen I/O Management)及執行程序(Procedure Execution)等機制。

Nonaka and Takeuchi(1995)則將知識分成外顯(Explicit)知識與內隱(Tacit)知識兩部分，外顯知識是可以透過有系統的組織，且易於傳播與分享的知識，通常以手冊、流程、範本、表格、資料庫、知識庫等結構化方式來表現；而內顯知識則是屬於個人主觀經驗知識，通常具有不易編碼且難以表達的特性，因此不易進行知識分享，例如個人經驗、專業技能等。Nonaka & Takeuchi(1995)認為人類知識的建立是一系列內隱與外顯知識相互交織與不斷轉化的螺旋過程，即知識會經由社會化、外部化、組合化與內部化的四個不同階段逐次地累積、整合、創造與增長，其個各階段的具體涵義說明如下：

- 一、社會化：將內隱知識轉換為另一內隱知識的過程，是一種藉觀察、模仿、體會或學習來累積知識與分享經驗的過程。此一階段重視知識的累積、分享與傳遞。
- 二、外部化：將內隱知識轉化為外顯知識的過程，是一種透過類比或隱喻將內隱知識書面化與表達解釋的過程，此階段所重視的是集思廣益及知識的編碼、演繹與歸納。
- 三、內部化：將外顯知識轉化為內隱知識的過程，是一種透過評估與判斷新知識的價值，過濾並確認具附加價值的外顯知識，並將其概念化成為個人的知識。在這個階段中，所重視的是知識的過濾、篩選與判斷。
- 四、組合化：將外顯知識轉化為另一種外顯知識的過程，將跨領域外顯知識加以系統化形成另一種整合性的知識體系或工作模式，並加以驗證其可行性。在這個階段中，所重視的是知識的應用與整合。

Guarino(1998)將本體論分為上層本體論(Top-Level Ontology)、領域本體論(Domain Ontology)、作業本體論(Task Ontology)及應用本體論(Application Ontology)等四類。上層本體論描述一般的概念，例如 Space、Time、Matter、Object、Event 及 Action 等，這些概念與特定的問題領域並無關聯。領域本體論及作業本體論描述特定相關領域(例如醫學、汽車)、作業或活動相關的字彙。應用本體論描述與特定領域或作業相關的概念，這

些概念通常相當於領域實體(Domain Entity)在執行活動時所扮演的角色,例如 Replaceable 或 Spare Component 等。

Labrou 和 Finin(1999)的論文中提到諸如 Yahoo 這類型提供目錄服務的索引搜尋網站,其目錄結構就如同網路文件的本體論。因為藉由目錄的結構可描述文件與文件之間的連結關係,這些資訊也可以從問題領域的觀點去找到構成這個問題領域的相關文件與其結構關係,也可以看成這個問題的本體論。

IEEE 標準上層知識本體工作小組發展了一個建議上層共用知識本體(Suggested Upper Merged Ontology, SUMO),透過 SUMO 將能促進資料互通性、資訊搜尋和檢索、自動推理和自然語言處理。事實上知識本體(ontology)類似於字典或者術語表,但是能使電腦處理更多內容的細節和其結構。知識本體包括將人們有興趣的領域正規化為一套概念、關係和定理(axiom)。上層的知識本體被限制在 meta 的概念、一般、抽象或者哲學,因此足夠一般使用者提出一個涵蓋廣闊範圍的領域區域。特殊領域具體的概念不被包括在上層知識本體中,但是這樣的知識本體確可提供特殊領域的知識本體結構的建立。SUMO 希望藉由最高層次的知識本體,鼓勵其他特殊領域知識本體以其為基礎衍生出其他特殊領域的知識本體,並為一般多用途的術語提供定義。

黃居仁(民 92)從語意網及知識運籌的角度指出,知識本體論是對任何一個網頁、資源知識內容及資訊架構的描述與定義。從語意網(Semantic Web)的構想而言,期望知識本體論是用 RDF 的形式(或類似語言)寫成的文件,並清楚定義概念間的關係和推理的邏輯規則。

Russell and Norvig(2003)則從知識工程的觀點論及本體論在建構知識庫程序中的角色,Russell and Norvig 認為在發展知識庫時可使用五個步驟方法來協助知識工程師:

- 一、決定要討論的物件與論據:唯有對於領域有足夠的瞭解,才能知道有哪些物件和論據需要被論及,哪些需要忽略,如此方可避免在未完全瞭解領域之前就進行公式化。
- 二、決定述詞、函數及常數要用到的字彙:此步驟的目的是將重要的領域層次概念轉換至邏輯層次的名字。
- 三、此領域之一般性知識編碼:本體論是此領域概念的非正規之條列,藉由寫下在本體論中術語的邏輯命題或公設,可達成兩個目的。第一個目的是使這些術語更加精確,並使人們同意這樣的解釋,第二個目的是使執行推論程序而從知識庫自動推導出結論是可能的。
- 四、編碼問題特殊案例描述:如果本體部分已經考慮周全,它將只包含寫下已經是本體一部份的有關案例之簡單單元命題。
- 五、對推論程序提出查詢並得到解答:藉由推論程序在公理及問題明確的論據上操作,以得到我們有興趣想知道的論據。

第四節 中文動斷詞與中文結構

英文的文句結構與中文文句結構的組成方式相當不同,而且英文詞彙與詞彙間通常會以空格來區隔,而中文詞彙通常由一個或一個以上中文字組成,詞彙與詞彙間則是緊緊相鄰,並無任何標記區隔,僅有文句與文句間會採用不同的標點符號來區隔,因此知識萃取的相關研究若要能夠剖析中文文章的內容,就需要先進行文字內容的斷詞工作,

再透過文句內容的文法結構及詞彙語意剖析來萃取知識。

為了國內中文自然語言處理及其它相關研究提供基本的研究資料與知識架構，中研院詞庫小組(Chinese Knowledge Information Processing Group, CKIP)於 1995 年 11 月開始建構一個中文自然語言研究環境，此環境包括一個資訊處理用中文詞知識庫及一個代表當代白話文的語料庫和中文語句剖析系統。中央研究院所發展出的中文斷詞系統具有中文斷詞及中文詞類自動標記功能，使用者可以根據自己的需求，而選擇不同的辭典，以做為斷詞及標記系統的參考(中央研究院資訊科學研究所，民 93;中央研究院，民 93)。

CKIP 的中文自動斷詞系統中由於建構了中文結構樹的基本原則，因此對於中文詞句的斷詞工作能有相當不錯的能力。在 CKIP 的中文結構樹的基本原則中，有別於一般只包含語法訊息的結構分析的方法，CKIP 的中文結構樹它同時包含了語法及語意(semantic)訊息，依據「中心語主導原則」(Head-Driven Principle)，任何句子或片語結構都會有一個中心語，並由中心語決定各詞組所扮演的角色，不僅可標示出句子中詞與詞之間的語意限制關係，由於語法功能一目了然，也可使中文語意被樹狀結構化(Chen et al.,2004;Huang et al., 2000)。

一、詞庫剖析原則

整個詞庫的剖析原則共包括一個主要原則及三個輔助原則，主要原則是指中心語主導原則，三個輔助原則則包括詞類小而美、由左而右聯併及扁平原則(Chen et al.,2004; Huang et al., 2000)。

- (一) 中心語主導原則：此原則是指句子或詞組結構皆有一中心語(Head)，詞組結構是由中心語與其論元或附加成分組成。句子(S)和述詞詞組(VP)的中心語皆為述詞(V)，名詞詞組(NP)的中心語為名詞(N)，介詞詞組(PP)的中心語為介詞(P)，方位詞詞組(GP)的中心語為方位詞(Ng)，對等連接詞(XP) 的中心語為連接詞(C)，XP 的詞類由連接成分決定，連接成分為述詞詞組(VP)，則為述詞詞組(VP)，連接成分為名詞詞組，則為名詞詞組(NP)。
- (二) 詞類小而美：此原則是指如果下層結構屬於語義等重的詞組，則下層結構會先取詞組最小類，再構成詞組。此原則中還細分為三種情況，分別為最下層結構皆由單詞構成、最下層結構有一方是詞組及最下層結構皆由詞組構成。
- (三) 由左至右聯併：此原則是指在「對等連接詞詞組」或「並列結構」裡，組成成分的語義等重，由左至右聯併。例如「老虎、獅子和大象」是「對等連接詞詞組」的例子，而老虎獅子大象則是「並列結構詞組」的例子。通常在「對等連接詞詞組」和「並列結構詞組」中會配合「詞類小而美原則」，先將同是單詞的詞組以共同詞類標示。
- (四) 扁平原則：根據「中心語主導原則」，每個詞組結構皆由中心語與其論元或附加成分組成。而「扁平原則」則使得每個詞組結構中的論元或附加成分的最上層結構都與中心語位於同一層次。由於中文語義的重心大多位於最右邊，故 Parser 會由最右邊的詞組聯併到左邊，而在中心語左方的詞組則以語義為單位來組成中心語的論元或附加成分。

二、詞彙語意訊息：詞彙語意訊息部分包括詞條(Entry)及語意角色(Role)兩部分。詞條包括同形異義字的不同意面，其所整理的常用同形異義字包括「是」、「有」、「來」、「給」；而語意角色則處理該字的意義、語意角色，以及與各語意角色間的區別。然而，中

心語又分語法上的中心語及語義上的中心語，一般時候，語法上的中心語同時也是語義上的中心語。但在某些特定結構裡，語法上的中心語不等於語義上的中心語。因此其角色符號中使用 Head 表示語法上的中心語，使用 Head 表示語義上的中心語，其餘符號如上述中心語主導原則內容所描述。

三、語法訊息：語法部分包括連接詞(Conjunction)及特殊表義式中文句(Construction)。其中連接詞分為並列連接詞(Ca)及關聯連接詞(Cb)。特殊表義式中文句則列舉出三種語言現象，包括不以單詞表義而以整體結構表義的語言現象處理、同位語處理、沒有述詞的句子。

第五節 語意網與超文字文件結構

語意網(Semantic Web)是 Berners-Lee(2001)在 Scientific American 所提出的概念，Berners-Lee 同時也宣告語意網將成為下一代的網際網路，由於 Berners-Lee 是網際網路的發明人，因此他的宣告也受到各界的重視。在全球資訊網的環境中，其功能僅可被視為是一種交換文件的媒介，而文件中的資訊內容及語意卻無法讓電腦能自動瞭解或直接運用，因此語意網其實就是一種新的網路內容形式，它能提供電腦閱讀及理解網頁中的語意，並能有助於概念的溝通與知識體系的整合(黃居仁，民 92)。

黃居仁(民 92)與黃居仁、張如瑩、蔡柏生等人(民 92)的研究中指出，要讓電腦能閱讀語意，可利用資源描述架構(Resource Description Framework, RDF)與通用資源標誌碼(Universal Resource Identifier, URI)連結到相關的網頁/資源，在透過超連結找到關鍵詞後，則可再藉由知識本體(Ontology)來定義關鍵詞，並進行邏輯推理。雖然目前在電腦中應用最多的語言是英文，但黃居仁則預測十年內網上的人口中有四分之一是講中文的，因此語意網應該是靠知識溝通，而非靠語言溝通。對於語意網的內容標示而言通常會採用 XML，而在描述本體論時有些人則認為應採用 OWL。

為了使網頁內容能更容易檢索及使用，因此有學者認為可在超文字格式的資料架構中加強其結構化的段落定義，使之具有網際網路資源描述與意義內涵，以解決網路文件非結構化的問題。而其中較有名的研究包括由美國線上電子圖書館中心(OCLC)與美國高速電腦中心(NCSA)舉行的 Meta data Workshop。他們整理出十二項具有網際網路資源意義內涵的資料項目，遵循超文字文件目前架構，擴充超文字語言較具結構性的 HEAD 標記段落中的 META 標記子段落(Weibel,2004; Weibel, Godby and Miller, 2004; Raggett,2004; Quinn, 2004)，稱為 Dublin Metadata Core Elements Set，簡稱為 Dublin Core。這十二種資料項目包括主題(Subject)、標題(Title)、作者(Author)、出版者(Publisher)、其他相關人員(Other Agent)、文件出版日期(Date)、文件種類(Object Type)、文件格式(Form)、辨認字串(Identifier)、與其他文件的相關性(Relation)、語言(Language)、涵蓋範圍(Coverage)。

Desai(1997)則認為 Dublin Core 過於簡略與遷就現行架構，因而針對超文字語言架構提出修正，並徹底改寫超文字語言中較具結構性的 HEAD 標記段落，重新定義其子段落組成，稱為「語意標頭」(Semantic Header)。而語意標頭與 Dublin Core 的差異在於語意表頭是在一些重要的標頭項目中採取一些較具有資料保存意義的做法，包括增加標頭項目的資料層數(Levels)、增加歸屬領域(Domain)與對應值(Value)的組合表示法、提供具演化特性的標頭項目、針對與其他文件關係特性的標頭項目做更佳的定義與釐清、將一些

具有關連性的欄位做整合與擴充...等。

楊亨利與金士俊(民 90)的研究則指出不論是 Dublin Core 或是 Semantic Header, 不僅有助於在全球資訊網的資料中找出資訊, 其網際網路資源定義的資料項目也隱含一些網際網路與出版之知識, 因此具有資料發掘之價值。因此若是超文字文件能提供 Dublin Core 或 Desai 之 Semantic Header 則將能提供更多的資料挖掘的規則應用。而目前許多網頁製作工具已經提供簡便的介面, 方便超文字文件製作者填註 Meta Data, 許多搜尋引擎的註冊要求與公開的搜尋策略也有助於 Meta Data 建置的推廣。為使網頁超文字文件包含豐富的 Meta Data, 本研究將透過本體論、知識庫及詞義關係來自動產生動態的線上課程教材講義文件, 文件中亦將整合上述學者所提及之相關的標頭項目, 使得講義文件能提供更多的內容語意及描述知識主題間的關係。

第六節 詞彙網路

詞彙網路(WordNet)是普林斯頓大學的一個專案, 事實上他是一個英語詞彙資料庫系統, 此系統透過將同義詞的集合聚集到稱為同義詞集(SynSet)的群組中來描述和分類單字和概念(Fellbaum, 1998)。WordNet 目前的版本是 2.0, 它包括一個表示成千上萬名詞的同義詞集詞彙表。同義集中具有大量概念及其它概念類型的概念以及子類別, 此外, WordNet 還包括相似但不同義的概念之間的對應。

黃居仁(民 92)認為要著手知識本體的呈現, 可採用詞彙網路(WordNet)來進行。詞彙網路的構成元素包括某個語言內所有的詞彙(任一個詞形 Lemma 與詞義 Sense 的獨特配對定義為一個詞彙)及該語言表達的所有概念(即所有詞義 Sense)。對於詞彙網路的架構而言, 第一個準則是以詞義為基準, 把有相同詞義的所有詞彙放在一個同義詞集(SynSet), 而同義詞集則是用來表達相同概念的所有詞之集合。第二個準則是採用定義過的詞彙語意關係, 連結所有的同義詞集, 如此便可建立所有概念間的語意關係, 而對詞彙網路而言, 除了同義、反義、近義之外, 更重要的有上位、下位及功能等近二十種關係。綜合上述, 詞彙網路提供的知識基本架構採用了同義詞集 Synset 做為詞彙驅動的概念(知識)單位, 之後再透過語意關係 Semantic Relation 進行概念連結與建立知識衍生的基本關係。而在語意學理論中有九種最典型的詞義關係, 分述如下:

- 一、反義關係(Antonymy): 用以描述概念與另一概念間具有相反之意涵關係, 即此概念為另一概念之反義詞。例如:「胖」與「瘦」,「買」與「賣」等。
- 二、上位關係(Hypernymy/Hypernym): 用以描述某概念具有其他上位關係的「泛稱」概念, 即此概念有其他上位詞的概念。例如:「車」的上位詞是「交通工具」。
- 三、下位關係(Hyponymy): 用以描述某概念具有其他下位關係的概念, 即此概念有其他上位詞的概念。例如:「車」是「交通工具」的下位詞。
- 四、整體一部份關係(Holonymy/Holonym): 用以描述某概念有組成此概念的其他部分概念, 即此概念有其他可構成此概念的部份詞。例如: 整體詞「身體」的部分詞為「心臟」及「胃」等。
- 五、部份一整體關係(Meronymy/Meronym): 用以描述某概念是其他概念的一部分, 即此概念與其他概念組合後, 能生成另一個整體詞概念。例如:「胃」、「心臟」、「血管」等部分詞的整體詞為「身體」。此外, 在 Holonymy 及 Meronymy 關係中, 也可再區

分出區域、成分、單位、成員...等類型之概念。

六、轉指關係(Metonymy/Metonym): 利用一個組織、法人、地區等名稱來代替實際執行工作的人便稱為轉指關係。

七、近義關係(Near-Synonymy): 是指某概念有與其概念相似的其他概念, 例如「幽默」與「滑稽」在概念意義上相近, 但卻不完全相同。

八、同義關係(Synonymy): 是指某詞彙與其他詞彙在意義上有相同的概念, 例如「美麗」及「漂亮」在生活中被視為同意關係的詞彙。

九、方式關係(Troponymy): 主要用在動詞上, 早期理論將其視為上位、下位關係, 現在則稱為方式關係。在語意與概念的比對裡, 他主要是方式的不同, 例如「移動」的方式詞有走、跑、慢跑等, 這些事件都是移動, 只是方式上的不同。因此語言表達上, 方式關係是把同類的事件用方式區分。

除了上述的主要關係之外, 黃居仁還列出許多種類型的涉入(Involved)關係, 例如因果關係(CAUSES)、角色(Role)關係、工具(INSTRUMENT)關係等。本研究將透過上述這些詞義關係的觀念來連結知識單元與知識單元間的關係。

第七節 測驗形式之演化

一、傳統測驗

由於教育的目的在於引發學生行為的改變, 因此教師必須善用測驗來評量學生學習後行為改變的情形(林璟豐, 民 90)。就教學而言, 測驗的主要目的在於增進學習的效果, 所以在學習的各個階段以及擬定教學計畫時, 必須將測驗納入成為整個教學過程中的一部份(陳英豪等, 民 71)。

然而, 試題及題型的設計, 通常會考量教學的目標。在教育測驗理論中教育目標可分成認知領域、情意領域及動作技能等三種不同的層次(陳李綱, 民 86)。傳統的評量方式, 依照上述三大層次分類, 各有不同的測驗方法, 每一種學習結果的測量, 都有適合的試題類型。在認知領域之中, 傳統的測驗方式即是以紙筆測驗為主; 但是在情意與技能領域上, 相當重視實際情境中行為表現的測驗, 所以在教學結果之測量中經常被忽視(李大偉, 民 84)。

二、電腦輔助測驗

所謂的電腦輔助測驗(Computer-Based Testing, CBT)是將傳統的考試工具及考題轉移到電腦之中, 讓學生藉由電腦螢幕閱讀考題、利用鍵盤或滑鼠來移動游標並點選答案。Alessi 和 Trollip(1991)指出, 電腦輔助測驗在應用上, 具有選擇組合試題、易於產生試題及共享試題題庫等效用。何榮桂(民 79)及朱寧等人(民 92)也認為電腦輔助測驗, 具有提高測驗效率、減少測驗時間、減少評分誤差、提供多元豐富的試題呈現方式、提供無紙化測驗並降低印刷成本、電腦亂數選題、獲得測驗相關的其他資訊等優點。

依照電腦輔助測驗的發展現況而言, 主要採用古典測驗理論(Classical Test Theory, CTT)和試題反應理論(Item Response Theory, IRT), 來進行題庫的建構。這兩種理論的核心部分都是數學模型, 而且是基於不同的假設所提出的理論。古典測驗理論採用線性的定性模型, 試題反應理論則採用非線性的機率模型(朱寧等人, 民 92)。

CTT 經常使用的測量指標有平均分數和標準差，但 CTT 的施行存在一些限制，其中包括 CTT 用以評量試題品質的指標(如難度與鑑別度)嚴重依賴受測者的樣本，因此對同一問題，如果受測樣本的能力差異很大，則鑑別值會很高。此外，CTT 只適用於檢驗受測者相對能力水準(常模參照測驗)，不適用於檢驗實力能力水準(目標參照測驗)。而使用 CTT 來檢驗學生的某項能力或知識程度時，必須同時使用同一套測驗題，否則測驗結果無法直接進行比較(朱寧等人, 民 92)。

目前已有許多電腦輔助測驗輔以 IRT，來發展電腦化測驗系統，此類系統又稱為電腦化適性測驗(Computer-Adaptive Testing, CAT)，它屬於一種智慧型的測驗方法(Van der Linden & Glas,2000;Wainer & Dorans,2000)。IRT 包括試題反應模型、試題反應模型參數估計及試題與測驗信息函數等三個介面內容。雖然 IRT 可以克服 CTT 的一些缺點，但 IRT 的測驗依賴大量且預先準備的高品質試題，則是在現實教育領域中難以完全做到的。IRT 的應用，早在 1992 年美國著名的 GRE(Graduate Record Examination)測驗便已利用電腦進行考試，更於 1993 年實施電腦適性測驗的考試形式(周倩、簡榮宏,民 86)。美國管理會計師協會 IMA(Institute of Management Accountants)也於 1997 開始採用電腦測驗方式進行會計師認證考試(Peterson & Reider, 2002)。此外，台灣地區托福測驗(TOEFL)自 2000 年 10 月起由傳統的紙筆測驗改為電腦化測驗，托福電腦化測驗分為聽力、文法結構、閱讀及寫作四個項目(財團法人語言訓練中心，民 93)。

美國心理學會(American Psychological Associate, APA)在 1986 年，發表了以電腦為基礎的測驗在發展、使用以及解釋上的指導原則，以使電腦測驗能夠與傳統紙筆測驗具有同樣的效力。其中包括「在解釋電腦測驗之成績時，任何因為電腦造成的影響效果必須清除或重新計算」、「電腦測驗的發展者必須證明其測驗效度」、「決定效度的方法與紙筆測驗相同」、「個人在電腦測驗與紙筆測驗上的成績等級順序必須近乎相同」、「平均數散佈情形也必須近似相等」(Bugbee, 1996)。因此近年來有一些學者便針對電腦適性測驗的試題等化(Equation)等問題進行研究，並認為電腦測驗之評量應該與傳統測驗評量有著相同的評量效度(Tsai, et al., 2001;Han & Poglmann, 1997)。

三、網路測驗

電腦輔助測驗與全球資訊網結合就成了網路測驗，它不但具備有電腦輔助與電腦適性測驗的優點，更能兼具網際網路的各項特色(林璟豐,民 90)。以網路測驗所能夠達成的效果而言，它不僅只是施測的工具，同時也可以是教學的工具(周文正,民 87)。它能夠協助教師和學生得知學生對於課程的瞭解程度，如有不瞭解之處能夠即時予以適當、正確的回饋，並能夠協助學生獲得必要的資訊或是補救教材。因此網路測驗不僅是一個獨立的電腦輔助測驗的環境，也是網路教學環境中的一環。林璟豐(民 90)指出，目前國內外的網路測驗系統題型主要包括是非題、選擇題、填充題、配合題、問答題及模擬表現測驗。

McCormack 及 Jones(1997)認為網路測驗能夠改善評量的程序和方法，因為網路測驗具備有節省時間、即時回饋、減少資源、保存記錄、更加便利等特性。此外，Khan(1997)認為測驗及教學實際在網路上進行時能夠提供的特色包括互動性、方便性、全球一致性、全球化資訊、合作學習、多媒體特性、使用容易、容易發展與維護、跨越時空及作業平台、開放式系統、經濟效益、線上支援、線上評量、線上搜尋、線上資源等。周倩、簡榮宏(民 86)在其針對網路測驗的優缺點進行研究時指出，大部份的學生較信任網路測驗

的結果、認為網路測驗是比較具有效率的考試方式、大部份的受試者喜歡網路測驗的方式、認為利用網路測驗並不會影響考試實力。但是網路測驗的缺點是比傳統紙筆考試的速度慢、比較容易作弊、閱讀線上考試比閱讀紙張試題更容易感到疲倦、面對螢幕有礙考試時的思考、以及打字速度會影響答題表現等。

Mark(1997)及 Bennett et al.(1999)則認為，在網路測驗的試題呈現上，利用多媒體的方式來呈現更多元的測驗問題，以及利用人工智慧使電腦自動判斷如何去測驗及進行評量，是影響網路測驗未來發展的重要技術。

四、人工智慧與模糊理論在測驗系統的應用

認知診斷測驗的發展，是由當代測驗理論和認知心理學對學習理論的研究與發現，結合發展而來。為了使測驗系統能解決複雜的出題、評分、解釋及回饋等問題，並能支援教學、促進學生學習績效，許多研究結合人工智慧及模糊理論，進行測驗評分及診斷系統的發展(Langley et al.,1990; Marshall,1993; Sun,1999; Moundridou,2003; Hwang,2003; 林明達,民 87; 孫光天等,民 87)。而 Devedzic(2003)則指出「內容導向智慧」(Content-Oriented Intelligence)是下一代網站教育系統發展的重要議題。

模糊集合(Fuzzy Set)是指用來表示界限不明的概念集合，例如「很熱」的溫度集合，我們很難去界定「很熱」的溫度範圍。為了尋求新的表示方法，Zadeh(1965)提出模糊集合論，該理論將傳統集合的特徵函數從 0 到 1 之間絕對選擇，推展為 0 到 1 之間的任意值，此種新的特徵函數可稱之為歸屬函數(Membership Function)，常用之歸屬函數型態包括三角形、梯形、鐘形、高斯(常態)形等(Jang, et al., 1997)。在填充題測驗中，由於不同的填答者在填寫答案使用的詞彙並不相同，有些詞彙與標準答案的語意完全相同，有些則是部分相同，有些則是完全不同，因此若使用模糊集合的觀念將能使測驗系統不再只是具有非「對」即「錯」的二元評分能力，而能依照相似程度給予合理的評分。因此林明達(民 87)認為申論題及填充題的測驗題型使用電腦閱卷不易，其中填充題的答案常有不同的寫法或有許多相同意義的詞彙，使得無法明確的指定標準答案，這將是一個十分值得努力方向。

過去已有許多結合人工智慧所發展出來的診斷系統，例如 Marshall(1993)以「故事問題解決」(Story Problem Solve, SPS)、及 Langley 等人(1990)以「自動認知模型化」(Automated Cognitive Modeler, ACM)概念發展測驗系統，這些系統都是依據認知心理學理論，結合人工智慧技術建構及模擬人類解決算術問題時的認知程序與知識結構。這些系統不僅診斷受試者在解題時所使用的「綱要知識」(Schema Knowledge)，還能根據診斷結果辨認出受測者在認知失誤或隱含錯誤概念上的意義。

此外，Hwang(2003)使用概念地圖模型發展人工智慧教學及測驗系統，Hwang 認為測驗系統並非只給予學生一個測驗分數，而應分析學生測驗結果並給予一些學習建議與相關教材內容。Moundridou 等人(2003)也使用人工智慧技術，發展網站編輯工具以產生智慧型教學系統。

然而，目前國內外對於如何建立一套可以自行學習關鍵字詞、同義字詞、具有人工智慧的閱卷系統，並未見有特別的研究。Mark(1997)認為利用人工智慧使電腦能自動判斷如何去測驗及進行評量，是值得研究的方向，其對於網路測驗的未來也有明顯的影響。

五、線上測驗系統及智慧型測驗系統發展

隨著資訊科技及網路技術的發展，藉由電腦來進行測驗已成為不可避免之趨勢，而國內外對於線上測驗系統或智慧型測驗系統之發展應用也相當多，例如：美國 IBM 與 Arthur Anderson 公司所發展之電腦輔助測驗系統(何榮桂等人,民 85)、1992 年起美國 GRE 測驗採用電腦適性測驗、1997 年美國管理會計師協會 IMA 採用電腦測驗進行會計師認證考試(Peterson and Reider,2002)、2000 年 10 月台灣地區 TOEFL 測驗採用電腦化測驗、2003 年 9 月考試院宣布未來國家考試將採行電腦化測驗(財團法人語言訓練中心,民 93)，由此可見採用電腦及網路技術來輔助測驗，將成為未來測驗的必然趨勢。

目前國內外對於使用電腦進行測驗的研究，以 Lord(1980)所介紹的試題反應理論(IRT)及 Weiss(1980)所介紹的電腦適性測驗(CAT)為主，其中較為著名的研究包括何榮桂等人(民 85)提出一個採用遠距適性測驗的系統架構及其後續一連串 CAT 之研究(何榮桂,民 89a; 何榮桂,89b)，以及游寶達(民 87)對於智慧型電腦適性測驗進行研究，並利用模糊理論及智慧選題法，建構一個 ICAT 系統(<http://ical.cs.ccu.edu.tw/>)，以提升 CAT 對受測者能力的評估精確度和穩定性。

此外，在目前遠距教學環境中，許多教學系統已將測驗系統納入其中，例如 ClassNet 系統支援教師在線上編寫試題、解答，測驗題型包括是非、選擇及申論，其中是非題及選擇題由系統批改，申論題則以電子郵件寄給教師批閱(Van Grop & Boysen,1997)。Mallard 系統提供選擇、計算、布林運算等題型，系統會適時進行提示、回饋及計分(Swafford & Brown, 1996)。黃國禎等人(民 91)在國科會科學教育處推動的整合計畫 ITED(Intelligent Tutoring, Evaluation and Diagnosis)中，針對試題進行智慧型的線上測驗題型分析與改進研究，其中採用題目語意分析，來解決相似題目重複出現的問題，本研究則是針對填充題型的測驗答案語意進行分析，以解決線上填充題測驗的效力等化問題。

六、國內知名線上測驗系統功能與本研究系統預期功能之比較

本研究針對國內知名的線上測驗系統，從功能、測驗內容領域、測驗題型方面進行比較，並做為本研究系統發展之參考。由表 1 可知，大多數的線上測驗系統都會提供選擇題的測驗題型，而少部分系統則提供是非題及填充題等題型。值得注意的是，這些支援填充題型的測驗系統，僅能設計一個填充格題目，不能支援一個以上格子的題目，本研究規劃之系統則能支援最多 20 個格子的題目。此外，在答案的分析辨識上，現有系統大多採用二元評量的字元樣式比對，只有輸入的答案與標準答案完全相同才判定為「答對」，如果標準答案是「雙十節」，受測者輸入「國慶日」、「雙十國慶」則都算「答錯」。在表 1 中只有「國中生活科技線上測驗系統」提供了相似詞的答案分析功能，該系統允許老師預先將可以接受的答案一併輸入在同一個答案欄位中(所有答案間用半形逗號分開)，例如老師在標準答案的欄位中輸入「國慶日,雙十國慶,雙十節」，則當學生輸入「國慶日」、「雙十國慶」或「雙十節」時，則可判定為「答對」。但這些詞彙由於未採用相似語意詞庫來設計，無法被其他題目或試題重複採用，而且無法解決相同語意答案之擴充、相似語意之智慧模糊辨識等問題。本研究之系統則透過相似語意詞庫，將特定知識領域中具有相似語意的詞彙知識，建立到相似語意詞庫中，以支援教師們在同一個測驗系統環境中使用，並能讓不同的試題重複利用。而在回饋部分，大多數的現有系統只針對測驗結果給予成績結果的查詢，並未保留當初測驗之電子試卷及回饋內容，本研究將針對

這些問題進行加強。此外，本研究規劃之測驗系統，亦將提供其他測驗系統未提供之功能，諸如線上申訴、相似語意詞彙管理、同音異字管理、語意排除管理、教師個人化評分風格設定管理、11種回饋時機方式、系統自動產生具有高品質及共識的相似語意詞彙內容，以及同音異字、漏字、相似語意、全半形、英文大小寫、集合順序等智慧評分機制。

表 1：國內知名線上測驗系統功能比較表

網站名稱	功能	測驗內容領域	測驗類型	其他說明
e-School 電子校園 http://eschool.com.tw	線上測驗、考題製作、線上評分、答案詳解、成績排名、時間計時	兒童部、國中部、高中部、大學、社會大學等各級教育多元內容。	單選題、複選題、填充題	1. 填充題無語意辨識能力 2. 填充題僅可設計一個空格的題目
長春藤線上測驗中心 http://www.ivytest.com.tw	線上測驗、線上考卷製作、線上評分	大學、四技二專、高中、五專、托福、高普考、證券業務、駕照考試等試題測驗	單選題、複選題	
K-bar 考試網 http://www.k-bar.com.tw	線上測驗、線上考卷製作、線上評分、時間計時	升學模擬測驗(大學及高中)、語文能力測驗(中文、日文、托福、GEPT 全民英檢)、高普考、證照等	單選題、複選題、填充題	1. 填充題無語意辨識能力 2. 填充題僅可設計一個空格的題目
狀元考試資訊網站 http://elec.taivs.tp.edu.tw/kitty	線上測驗、線上考卷製作、線上評分、記錄及評估題目答題難易度、隨機出題	四技二專、二技、技能檢定、預官、高普考測驗	選擇題	
線上測驗服務中心 http://www.onlinetest.org	線上測驗、線上評分、成績查詢、時間計時、題目支援圖形內容	學科能力測驗(國文、英文、數學、社會、自然、歷史)	單選題、複選題	
階梯數位學院評量測驗 http://new.ladder100.com/	線上測驗、線上評分	九年一貫、高中、英文能力檢定等領域	單選題	
國中生活科技線上測驗系統 http://home.kghs.kh.edu.tw/olt	線上測驗、線上考卷製作、線上評分、成績查詢、時間計時、題目支援圖形內容、試題轉換	國中生活科技領域	單選題、是非題、填充題	1. 填充題僅可設計一個空格的題目 2. 填充題提供相似詞答案處理(需教師自行建立，系統未支援語意詞庫)
本研究之線上測驗子系統 http://140.126.115.230/OLT	線上測驗、線上考卷製作、線上模糊評分、電子試卷作答內容保存、成績查詢、線上申訴、相似語意詞彙管理、同音異字管理、語意排除管理、教師個人化評分風格設定管理、11種回饋時機方式、答案回饋...等	學科能力測驗(電子商務概論、管理資訊系統、商用程式語言、資料庫管理系統)	單選題、複選題、是非題、填充題	1. 填充題可設計 20 個空格的題目 2. 系統提供相似語意詞庫辨識填充題答案語意 3. 填充題提供同音異字、漏字、相似語意、全半形、英文大小寫、集合順序等智慧評分分析

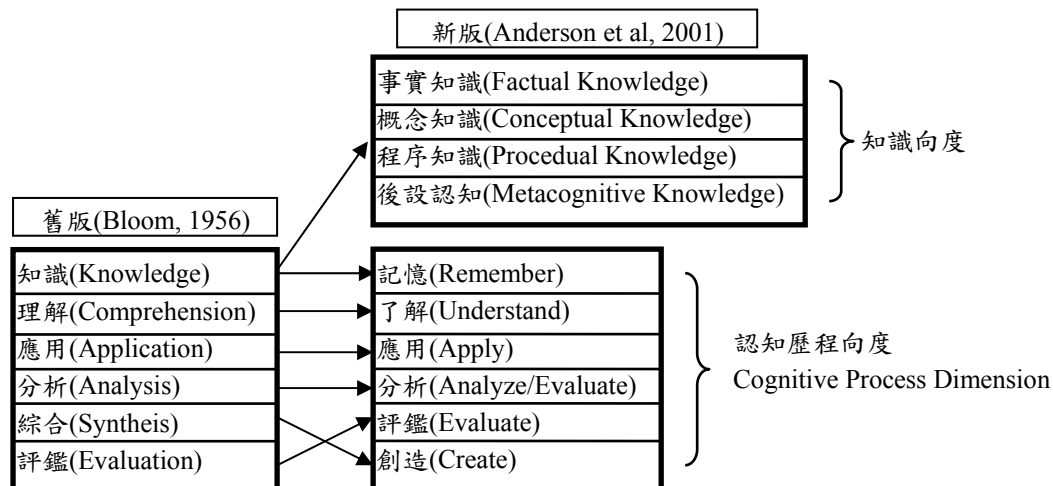
第八節 Bloom 教育目標分類理論

West 等人(1991)則認為教學所要的評量的知識包括敘述性(declarative)、程序性(procedural)和情境性(conditional)等三類。敘述性知識是有關於「了解是什麼(knowing that)」或「有關什麼(knowing about)」的知識，包括事實、事件、概念、現象、物體等陳述;程序性知識是有關於「了解如何(knowing how)」的知識，包括原則、程序、方法等提示操作執行參考的陳述;情境性知識是有關於「了解何時和為何(knowing when and why)使用」的知識，或稱策略性知識(strategic knowledge)。

何榮桂等人(民 88)指出，紙筆測驗是目前教育評鑑上最重要的一種方法，尤其是屬於認知領域的各種學習結果。Bloom 等人(1956)認為認知領域的學習，是以心智能力發展為核心，著重在理智、學識及問題解決能力之培養，因此將認知領域區分為知識、理解、應用、分析、綜合、評鑑等六個層次。何榮桂等人(民 88)指出認知、技能領域及情意的低層次能力，皆可使用紙筆測驗來評量，但紙筆測驗在編撰試題、施測、閱卷、評分過程中，相當耗費人力及物力；此外，受試者也必須在定時、定點接受測驗，缺乏測驗彈性。

表 2:新版 Bloom 認知領域教育目標分類表

知識向度(學習)	認知歷程向度(思考)					
	1.記憶	2.了解	3.應用	4.分析	5.評鑑	6.創造
A.事實知識						
B.概念知識						
C.程序知識						
D.後設認知知識						



資料來源：Anderson et al (2001), p.268.

圖 1：Bloom 教育目標分類系統新舊版本對照圖

Bloom 的認知領域教育目標分類系統不僅是一套測驗的工具，也是教學者在撰寫學習目標的共通語言，此外也可用以制定依照國家層級教育目標而設定下一層課程目標的基礎，甚至能促進課程中的教育目標、教學活動與評量具備一致性(Krathwohl, 2002)。因此，Bloom 等人(1956)的認知領域教育目標分類系統廣被國內外教育界採用，但隨著許多研究的檢驗及認知歷程的研究發展，一些學者從哲學思辨、實證研究方面對此分類提出

若干的批評，例如：高層次目標(綜合與評鑑)的階層性是否合理(Bloom, 1994; Furst, 1994; Kreitzer & Madaus, 1994; Postlethwaite, 1994; Krathwohl, 2002; 李宜玫等人, 民 93)，最後 Anderson 與 Krathwohl(2001)針對 Bloom 的分類提出了新的版本，如表 2。新舊版之間的主要差異，如圖 1 所示，其中原先舊版的名詞被改為動詞，原先的單向度則區分為「知識向度」與「認知歷程向度」雙向度。新版中的知識向度是專指知識的分類，並將知識分成事實(Factual)、概念(Conceptual)、程序(Procedural)及後設認知(Metacognitive)，各類別還包含若干次類別；而認知歷程向度則區分為記憶、了解、應用、分析、評鑑與創造，各分類另外也包含若干次分類(葉連祺與林淑萍, 民 92)。知識向度著重從「學習」的角度來區分學習者學習的知識類型；認知向度則著重從「思考」的角度來區分學習者思考的認知層次。

在知識向度的四類知識中，前三項知識類型源自於先前的目標分類系統架構，並加以調整命名而來；而後設認知知識則是將認知心理學中關於後設認知的觀點納入。這些知識類型依照具體到抽象的程度排列，雖然有部分的重疊，但是大抵仍維持此原則，而且各類知識之下又可區分為若干次類型知識，關於各類知識的內容說明如下(Anderson, 2001; Krathwhol, 2002; 李宜玫等人, 民 93; 葉連祺等人, 民 94)。事實知識：是指學習科目後和解決問題時應知的基本元素，此類別的知識另外包含二種次類別知識。

A. 事實知識：是指學習科目後和解決問題時應知的基本元素，此類別的知識另外包含二種次類別知識。

Aa.術語的知識(Knowledge of Terminology)：指特定語文與非語文的標題與符號知識；
例如：資料流程圖中的圖形符號、顧客關係管理(CRM)。

Ab.特定細節和元素的知識(Knowledge of Specific Details and Elements)：指事件、位置、人物、資料、資訊等，可包括精確、特定的資訊或約略的資訊；例如：千禧危機、決策支援系統特徵、實體關係模式組成元素。

B. 概念知識：是指基本要素與較大的結構共同發揮功能的互動關係，此類別的知識另外包含三種次類別知識。

Ba.分類和類別的知識(Knowledge of Classifications and Categories)：意指對不同事物的類別、等級、區分和排列的知識；例如：電子商務經營模式的類型，組織活動的層級。

Bb.原理和通則的知識(Knowledge of Principles and Generalizatons)：指透過觀察現象所摘述的摘要，且可用於描述、預測、解釋或決定行動與採取的方向；例如：全球電子商務發展五大原則。

Bc.理論/模式/結構的知識(Knowledge of Theories, Models, and Structures)：意指對複雜的現象、問題和事物提出清楚、完全與系統性的觀點，以理論和模式呈現，亦可用來描述、了解、解釋和預測現象，是事實與概念中最抽象的層次；例如：實體關係模式、三層式系統架構、科技接受模式。

C. 程序知識：指知道如何做某事的知識，通常是一系列或有步驟的流程，以及決定何時運用不同程序的規準，此類別的知識另外包含三種次類別知識。

Ca.特定學科的技能 and 演算知識(Knowledge of Subject-Specific Skills and Algorithms)：指有固定最終結果，或具固定順序或步驟的知識；例如：計算漢明碼、中序轉後序的資料結構。

- Cb. 特定學科的技術和方法知識(Knowledge of Subject-Specific Techniques and Methods)：對結果具共識或學科規範的知識，反應專家思考和解決問題的方式；例如：正規化方法、發展知識庫的步驟。
- Cc. 運用規準的知識(Knowledge of Criteria for Determining When to Use Appropriate Procedures)：指知道何時使用程序和過去使用該程序的知識；例如：選用資訊委外或自製方法的時機與程序。
- D. 後設認知知識：指一般認知以及自我知識的認知和覺察；含認知知識、監控、控制、調整認知(Monitoring, Control and Regulation of Cognition)，此類別的知識另外包含三種次類別知識。
- Da. 策略知識(strategic knowledges)：指用於學習、思考和問題解決的策略知識，可跨領域應用；例如：複誦、記憶術、概念圖等認知策略。
- Db. 認知任務知識(包括脈絡和情境的知識) (Knowledge About Cognitive Tasks, Including Appropriate Contextual and Conditional Knowledge)：指學生了解可以運用後設認知的知識，意即「何時」與「為何」能適當運用策略的知識；例如：知道看地圖比純背文字更易記憶方位，知道 How, When, Why 的運用。
- Dc. 自我知識(Self Knowledge)：可分四部分，包括對自己認知與學習能力優缺點的知識、偏好的策略，個人動機、信念、興趣、目標的察覺等。

認知歷程向度分為記憶、了解、應用、分析、評鑑、創造等六類，而各類之下又可再細分成若干次類別，在第二層的次類別中共包括了 19 個次類別項目。此六類的認知層次彼此間也有階層(Hierarchy)的概念，並且是以「漸增複雜性階層」(Increasing Complity Hierarchy)的關係來表示，也就是指類別層次是顯示出複雜性漸增的概念(Krathwhol,2002)。以下將各類認知層次的內容說明彙整如下(Anderson, 2001; Krathwhol, 2002; 李宜玫等人, 民 93; 葉連祺等人, 民 94)。

1. 記憶：是指從長期記憶中提取重要知識的能力，此類別中又可再細分出再認及回憶二種次類別項目。
 - (1) 再認(Recognizing)：是指從長期記憶中提取重要知識和所呈現的訊息加以比較的能力。
 - (2) 回憶(Recalling)：是指當給予提示或線索時，可以從長期記憶中提取出重要訊息的能力。
2. 了解：是指學生能從教學訊息中建構有意義知識的能力，其中也包含新舊知識的連結，此類別中又可再細分出七種次類別項目。
 - (1) 詮釋(Interpreting)：是指以能進行不同表徵間的轉換能力，例如：轉述、讀圖、依文畫簡圖。
 - (2) 舉例(Exemplifying)：是指可以根據概念或原則定義特徵，舉出例子的能力。
 - (3) 分類(Classifying)：是指根據概念或原則將某些事情歸屬於某類別中的能力。
 - (4) 總結/摘要(Summarizing)：是指能簡述所呈現的訊息或是摘要出一個要旨的能力。
 - (5) 推論(Infering)：是指藉由一系列的例子或項目間找到一組模式，並能夠解釋每個例子的特點和彼此關係的能力。
 - (6) 比較(Comparing)：是指檢視兩個(或以上)的事物的相似和相異之處(類比的推論)能力。

- (7) 解釋(Explaining)：是指能建構和運用一個系統的「原因—效果」模式的能力，通常評量「解釋」的方式有推究原因(Reasoning)、偵錯(Troubling)、再設計(Redesigning)和預測(Predicting)等三種方式
3. 應用：是指運用程序以表現動作或解決問題的能力，此類別中又可再細分出二種次類別項目。
- (1) 執行(Executing)：是指執行例行程序於熟悉工作(程序性知識)上的能力，如：運算，解方程式等。
- (2) 實行(Implementing)：是指將概念化的知識(理論模式或結構)應用到情境或事例的能力。
4. 分析：是指將材料打散成不同的組織成分，再由學生提出成分間與整體結構間關聯性的能力，此類別中又可再細分出三種次類別項目。
- (1) 辨別(Differentiating)：是指能從整體結構中區分出重要和不重要部分之能力。
- (2) 組織(Organizing)：是指建立系統性和連貫性的要素連結能力。
- (3) 歸因(Attributing)：是指確定溝通(文章、資料)材料中，隱含的觀點、偏見、價值觀和意圖之能力。
5. 評鑑：是指學生根據準則或標準進行判斷的能力，這些準則通常是品質、效能、效率和一致性，此類別中又可再細分出二種次類別項目。
- (1) 檢查(Checking)：是指檢測出測試操作過程或產品與內在標準不一致性或錯誤的能力。
- (2) 評論(Critiquing)：是指基於外在的準則或標準判斷一個產品或操作過程的能力。
6. 創造：是指將要素聚在一起形成一個協調或具功能整體的能力。此類別中又可再細分出三種次類別項目。
- (1) 產生(Generating)：是指能呈現問題並達成符合某些準則的多種選擇方向或假設能力。
- (2) 規劃(Planning)：是指能規劃一個符合問題準則的解決方法能力。
- (3) 製作(Producing)：是指能執行一個計劃或創造一個產品，而得以解決(符合某些特徵的)問題的能力。

陳筱菁(民 93)針對民國 90 年至 93 年四技二專統一入學測驗中的商業類計算機概論試題，以新版的 Bloom 分類進行試題分析，試題分佈如表 3 所示。其研究結果指出在以選擇題為主的四技二專統一入學測驗中，計算機概論的試題所包含的認知歷程無法達到評鑑與創造等二種層次，而在知識類型部分則涵蓋了事實知識、概念知識與程序知識等三類。就認知歷程而言，「記憶」層次通常與「事實知識」相關，「了解」層次通常與「概念知識」相關，而「應用」層次通常與「程序知識」相關。此外，陳筱菁也指出 Bloom 教育目標分類理論是用來檢視試題的知識類型與認知歷程之實用工具，但卻無法從分類結果中得知試題的難易度。

Bloom 教育目標分類概念已經被廣泛應用在線上測驗試題的發展研究上，近年也有不少研究以 Bloom 新版的分類理論為基礎，進行線上測驗試題的轉換、選題、評量策略等研究(沈晏仕等人,民 94a; 沈晏仕等人,民 94b; 翁瑞鋒, 民 94; 沈晏仕等人,民 94c)。

表 3:民國 90 年~93 年四技二專商業類計算機概論試題分佈

知識向度(學習)	認知歷程向度(思考)					
	1.記憶	2.了解	3.應用	4.分析	5.評鑑	6.創造
A.事實知識	56%	2%				
B.概念知識		17%	2%	4%		
C.程序知識		10%	7%	4%		
D.後設認知知識						

資料來源：陳筱菁(民 93)

註：因 93 年有 2 個試題跨 2 個類型，因此總和為 102%

第九節 測驗題型與試題發展原則

Gronlund(1998)依據題型的表現方式，將測驗題型歸納成選擇反應(Selected Response)、供應反應(Supply Response)、限制性實作(Restricted Performance)及延伸性實作(Extended Performance)等四大類。本研究以認知領域為教學目標進行智慧型線上測驗系統設計，因此僅探討認知領域中的選擇反應及供應反應測驗題型，對於情意領域、技能領域的實作性測驗則不進行探討。在認知領域之中，對於學生的學習測驗，大致上可以區分成認識型題目(Recognition-Type Items)及建構型題目(Constructed-Response Items)(鼎茂, 民 89)。認識型題目即 Gronlund(1998)分類中的選擇反應型題目，此類題目通常會包含一段文字敘述的刺激，以及一些選擇或選項以提供學生根據刺激來辨認各個選項，並選擇出適當的答案，例如是非題、選擇題及配合題。而建構型題目即 Gronlund(1998)分類中的供應反應型題目，此類題目包含了問題或狀況的刺激，再要求受試者根據題目的要求，將自己對於內容的瞭解，以文字或是語言的方式表達出來。例如名詞解釋、填充題及問答題等。

認識型題目的特點是它的答案不會因為人、時、地而有所改變，亦即其答案都是固定的(Rigid)，不會因為評分者的不同或情境時間不同，而產生重閱評分結果的差異。然而，建構型題目之答案卻經常會因評分者的不同或情境時間不同，而產生重閱評分差異。填充題主要是要求學生在一個未完成的敘述句之中，填上「字」或是「句子」，而使之成為一個正確且完整的敘述。雖然填充題在施測上，可以更加精確地測量學生各方面的知識，並降低猜答的機率，但因為同義字彙太多，容易造成計分不客觀的情形(陳英豪等, 民 71；李大偉, 民 84)，因此如何使填充題型的測驗有更客觀合理的計分方式，並使測驗系統能辨識出與答案相關的語意詞彙，並給予正確客觀的評分機制，將有其研究的重要性，本研究將針對此問題進行探究。理論上，我們可以在網路環境中實施任何型態的測驗。但是由於客觀環境限制及技術困難，目前最常應用在網路上的測驗題型，仍以選擇反應題型為主；因為電腦評分不易，電腦難以自動辨識相同及相似之語意詞彙，因此供應反應題型較為少見。若能使電腦在供應反應題型評分上，擁有與老師相同的專業評分準則及專業知識，那麼對於未來線上測驗系統的發展及應用，將能更加成熟。因此，本研究便針對建構型之填充題型進行測驗系統之研究。

葉連祺(民 89)認為各種測驗題型彼此間可相互轉換，例如是非題去掉關鍵詞，可變成填充題或申論題；增加多個答案項，可變為選擇題。而各種題型試題適合評量的認知目標層次也並不相同，是非和選擇等題型適合評量知識、理解、應用和分析等層次，申論題適合評量綜合、評鑑和創造能力(Bloom, et. al, 1956)。

McKenna 與 Bull(1999)認為利用電腦輔助評量的試題題型主要包括以下十一類：

- 一、單選題：此類題型要求學生從題目表列的可能答案中，選出一個正確的答案。
- 二、是非題：此類題型要求學生直接評估題目敘述的內容是否正確。
- 三、斷言題(Assert-Reason Questions)：此類題型屬於新一代的是非題型，它必須在選定是非題答案外，額外說明認為該題「是」或「否」的理由。
- 四、複選題：此類題型與單選題類似，但是此類題型則要求學生從題目表列的可能答案中，選出一個以上的答案。
- 五、圖形熱點題(Graphical Hotspot Questions)：此類題型要求學生從螢幕圖形顯示的區域選項中選出正確的選項及位置。
- 六、文數字題(Text/Numerical Items)：此類題型要求學生根據題目內容的問題，直接從鍵盤鍵入文字或數字。此類試題一般包含測驗試題中熟知的填充題型試題，填充題則要求學生在題目中的空格處，寫出適當的答案。
- 七、配合題：此類題型要求學生根據一個表列的清單項目內容，依據項目間的相關性，連結到另一個表列中的清單項目。
- 八、關鍵點指示題(Sore Finger Questions)：此類題型被使用在語言教學和電腦程式撰寫上，其要求學生依據題目中所缺少的字、程式碼或片語，利用鍵盤輸入文數字或利用關鍵點指示題方式完成回答。
- 九、評比題(Ranking Questions)：此類題型要求學生將一系列的相關項目，根據題目內容，在另一處以知識順序、事件順序或層次等級概念進行選評。
- 十、排序題(Sequencing Questions)：此類題型要求學生去將文字或圖片物件依照題目所指示概念，放置出正確的順序，此類試題對於評量方法(Methodology)知識特別有效。
- 十一、領域模擬題(Field Simulation Questions)：此類題型提供真實問題的模擬檢驗。

由於不同的試題題型都有不同的評量目的，許多研究者也紛紛指出教師在建構試題時，應該遵守每一種題型的試題建構原則，以下針對線上測驗系統中最常見的主要題型試題建構原則進行歸納說明(歐滄和，民 91; 李坤崇，民 88; ; NAVEDTRA, 1992; Borich, 1997; Clay, 2001; Thomas, 2004)。

- 一、選擇題：發展單選題或複選題的試題所需遵守的原則主要包括以下 15 項原則。
 - (一) 題幹僅包含單一的定義或概念，避免內容過於冗長複雜。
 - (二) 題幹中避免不必要或與問題無關的內容敘述。
 - (三) 題幹內容避免使用過於艱澀或不常見的詞彙，盡可能使用受測者能容易瞭解的詞彙文字，以使受測者不會因文字的可讀性而對於題目產生誤解。
 - (四) 若題幹使用否定問項，需要將否定詞以大寫字母、加底線、加粗或特別強調出來，以提醒學生注意。
 - (五) 每個選項中避免出現重複的敘述內容。
 - (六) 確保只有一個備選答案是正確的。
 - (七) 同一試題內的備選答案內容，其描述方式、長度及複雜度盡可能具有一致性。

- (八) 使用近似合理、近乎真實，且具有吸引力的備選答案。
- (九) 避免提供選擇正確答案或刪除不正確答案之線索。
- (十) 盡可能避免使用「以上皆是」或「以上皆否」的備選答案。
- (十一) 題幹中避免使用過度極端或模糊的字眼(例如：「必然」、「不曾」、「也許」、「經常」、「可能」...等)，以避免出現其他亦可接受的例外情形。
- (十二) 備選答案項目數量應該要充足，通常應有 3~4 個選項，較多的備選選項，可以降低學生猜題的機率。
- (十三) 備選答案項目數量需保持一致。
- (十四) 盡可能採用完整問句描述，題幹後應加註問號，避免使用不完整的題目描述。
- (十五) 試卷題目及試題的備選答案內容間需要彼此獨立且互斥，避免不同選項間具有重疊概念。

二、是非題：發展是非題的試題所需遵守的原則主要包括以下 9 項原則。

- (一) 每個題目都應該是絕對的「對」或「錯」，不可有模稜兩可或部分例外可接受的情形。
- (二) 試題內容僅涵蓋單一知識概念相法，避免由多個不同的知識概念來組成。
- (三) 題目內容需提供充分及完整的背景資訊描述，以確保試題不會產生不同情境產生不同答案的情形。
- (四) 試題描述需簡潔及清楚。
- (五) 避免直接使用簡單否定的描述出題。
- (六) 答案為錯的試題內如，需以相關且有意義的詞彙概念來取代，以達到觀念釐清及提升評量層次的目的。
- (七) 題幹中避免使用過度極端或模糊的字眼(例如：「必然」、「不曾」、「也許」、「經常」、「可能」...等)，以避免出現其他亦可接受的例外情形。
- (八) 避免使用否定詞來陳述問題。
- (九) 題幹內容避免使用過於艱澀或不常見的詞彙。

三、填充題：發展填充題的試題所需遵守的原則主要包括以下 5 項原則。

- (一) 除了與該答案相同的可替換詞彙外，每個填充答案的內容只有一種可能的答案。
- (二) 避免設計太多空格，而造成題目意涵遭受破壞的問題。
- (三) 需讓學生回答的空格避免設計在句首，並盡可能將空格調整製試題的句尾。
- (四) 試題內容若出現數個連續空格，且具有順序性時，題目應特別標示出來。
- (五) 計分標準要採「可接受」答案，而非「正確」答案

四、配合題：發展配合題的試題所需遵守的原則主要包括以下 8 項原則。

- (一) 各題幹應具同質性。
- (二) 選項反應數目應多於題幹數目，且最好告知反應被選次數。
- (三) 作答指導語必須明確規定和說明。
- (四) 題幹與選項反應盡量簡短，較短選項宜條列在題幹右方。
- (五) 選項反應項目宜依邏輯順序排列。
- (六) 題幹與選項反應序號不應相同。
- (七) 配合題中一個完整試題應印在同一頁。
- (八) 配對題目以不超過 10 項為原則(以 5~8 項為宜)。

由於每一種題型均有其優缺點及適用時機，國內外許多文獻也針對不同題型的效益進行比較，各種測驗題型的優缺點簡述如下(歐滄和, 民 91; 李坤崇, 民 88; Rowntree, 1977; McBeath, 1992; NAVEDTRA, 1992; Borich, 1997; Clay, 2001; Thomas, 2004)：

一、是非題：

- (一) 優點：命題容易、可適用於各種教材、計分客觀、作答方法簡單迅速、短時間可產製較多此類試題、可容納題數較多，且內容取樣更具代表性。
- (二) 缺點：鑑別力較差、信度低、試題內容可評量層次較低、知其然未必知其所以然、猜對答案機率偏高、無法涵蓋二分法以外之現象、教師無法從學生的錯誤答案中進行分析以獲得診斷性訊息。
- (三) 適合評量層次：通常用以評量 Bloom 舊版之知識層次，更高層次不易使用此類型達成。

二、選擇題(單選題/複選題)：

- (一) 優點：可適用於 Bloom 舊版的各種不同層次學習結果測量、可避免答題常出現的題意不清現象、編製較配合題容易、信度優於是非題、可以避免學生依照習慣的反應心態作答、具有診斷評量的效果、計分容易、客觀可靠、作答時間短、題目可容納長、內容取樣可較廣、可提高內容效度、相較於是非題有較低的猜對答案機率。
- (二) 缺點：通常只能以語文問題呈現、較不適用某些數學或科學問題的測量、不適合於組織和提出觀念能力的測量、誘答的編製不易、有學生猜測之困擾。
- (三) 適合評量層次：可用以評量 Bloom 舊版之知識、理解、應用、綜合、分析、評鑑等六個層次的的能力。

三、配合題 (match items)

- (一) 優點：可在極短的時間內測量大量相關聯的事實資料、編製容易、計分客觀容易、信度高、格式比是非題及選擇題更濃縮精簡、較不易猜測、可節省許多作答時間、可增進學生對比的觀念。
- (二) 缺點：僅能測量機械記憶的事實資料、容易提供額外的線索、很難找到一些符合教育目標和學習結果的同質材料、同一題材題數太多造成內容取樣口的偏差。
- (三) 適合評量層次：Bloom 舊版之知識層次，若能精心設計便可發展評量理解層次的試題。

四、填充題

- (一) 優點：大量產製此類試題容易、減少猜題機率、鼓勵學生更專注於知識內容的研讀、受測者可自長期記憶中將知識讀出或寫出，用以評量學生對於與 who、what、where、when 相關知識內容的學習成就特別有效。
- (二) 缺點：計分較主觀困難、可能出現許多也可被接受的其他詞彙答案、更高層次認知能力不易被評量。
- (三) 適合評量層次：適用於評量 Bloom 舊版之知識層次，特別是事實的回憶與記憶。

第十節 測驗回饋

回饋是一種提供給個人的資訊，其包括實際績效及被期望績效間的比較資訊。在早期研究中，回饋應用於績效領域，被視為是增加績效精確度的一種刺激物，而且亦能增加學習者的學習效果(Devedzic, 2003)。因此，Alderson(2000)針對回饋對態度及績效等構面的影響進行研究，並主張未來測驗方法的研究，應對其績效及態度等影響進行評估。回饋主要的兩個目的分別是：1.當學生對問題作答時，讓學生知道他的作答是對的；2.當他答錯時，讓他知道他的作答是錯的。因此，回饋不僅僅是簡單的對或錯，它還可能包含各種回應資訊。洪榮昭(民81)也認為回饋對學生的學習成效會有明顯的影響，而且也可以減少作答的時間。

自1970年起已有一些回饋方法被提出來，洪榮昭(民81)則認為，回饋通常有兩種方式，一是強調知識的正確回應(Knowledge of the Correct Response, KCR)，另一種是強調知識結果(Knowledge of Result, KR)。KCR是以提供學生正確的答案為重點，亦即提供學生增強正確作答的能力。KR則常被引用以選擇題作答時用來暗示，假如答案是錯的，另外的則可能是對的。

Dempsey 等人(1993)，則根據情境及回饋的資訊內容，而將回饋分成以下五類：

- 一、無回饋：未針對問題和需要提供回饋，而且也未指出學習者的回答是否正確。
- 二、知識結果回饋(Knowledge of Result, KR)：此回饋會針對學習者的回答內容，給與一個簡單的「正確」或「錯誤」的結果通知，因此這個回饋方法僅提供簡單的檢驗回饋結果。在實際的應用上，通常是對學生的回答給與「錯誤」或「正確」的訊息，這些訊息包括了一般常用的字意、熟悉的圖形聲音，其基本功能是取悅或引導學生的學習。
- 三、正確知識回應回饋 (Knowledge of Correct Response, KCR)：又稱正確回應回饋(Correct Response Feedback)，此回饋會通知學習者，關於該問題的正確答案內容應該是什麼。因此KCR是以提供學生正確的答案為重點，在設計時要考慮及預期學生的錯誤範圍，當學生作答錯誤時，給與學生具思考性的答案指示。
- 四、詳盡式回饋(Elaborated Feedback)：此回饋會提供一些解釋詳盡的回饋說明，以描述為何學習者的回答是正確或不正確，或者允許學習者去回顧參考與正確答案相關的教材資料。
- 五、再次嘗試回饋(Try-Again Feedback)：此回饋適用於當學習者已經完成答案回應，而且所回應的答案也是不正確的狀況。此時，系統允許學習者在嘗試選擇或回答其他的答案，以期學習者最終能正確的回答出答案。

上述各種回饋類型中，並沒有任何一種回饋方式被認為是最好或最差的回饋方式。Gilman(1969)在回饋類型應用在對電腦基礎教學的研究上發現，以就讀科學相關系所的大學生而言，接收到 KCR 回饋的學生其學習績效顯著優於沒有接收到 KCR 回饋的學生。

另外，Dempsey 與 Wager(1988)則認為依照回饋時間分類，可將回饋區分成立即回饋及延遲回饋。然而 Kulik 與 Kulik(1988)對於立即回饋及延遲回饋進行研究發現：立刻回饋比延遲回饋更有效；延遲回饋僅僅在重複的語詞刺激或測驗內容的實驗情境上，才會明顯有效。以下是 Dempsey 與 Wager(1988)所提及的立即回饋及延遲回饋類型：

一、立即回饋(Immediate Feedback)：立即回饋是允許在教學或測驗期間，透過硬體及軟體即使給予學習者或受測者一些正確的資訊回饋。立即回饋的類型包括：逐題(Item-by-Item)、學習者控制(Learner-Controlled)、邏輯的內容中斷(Logical Content Break)、末端模組回饋(End-of Module)、學習者中斷(Break by Learner)、時間控制(Time-Controlled)等。

二、延遲回饋(Delayed Feedback)：延遲回饋是學習者或受測者在教學或測驗期間，經過一個程式特定的延遲時間後，才給予學習者或受測者一些正確的資訊回饋。延遲回饋的類型包括：逐題(Item-by-Item)、邏輯的內容中斷(Logical Content Break)、少於一小時(學習或測驗結束後)、1-24 小時(學習或測驗結束後)、1-7 天、擴展延遲(Extended Delay)、下一次參與學習或測驗前(Before the Next Session)等。

此外，立即回饋及延遲回饋二種方式，對學生的學習效果皆有意義，立即回饋對那些課程生疏的學生較為有效，而延遲回饋則有助於學習知識的保留。對於理解性的教材，學生在作答時，應給予相關知識的訊息，以讓學生更正其錯誤，獲得良好的學習效果。