

網路評量系統融入師資培育之探究

王子華

仁德醫護管理專科學校

王國華

彰化師範大學科學教育研究所

王瑋龍、黃世傑

彰化師範大學生物學系

摘 要

本研究擬實際應用具有「試題組成、考試管理、試後評估」功能之 Triple-A (Assembling、Administrating、Appraising)架構的網路評量系統-「網際網路評量與試後分析系統 WATA(Web-based Assessment and Test Analysis System)」來輔助生物科職前教師之師資訓練，用以評估網路模擬教學策略對師資培育的輔助效益與可行性。在為期四個月的研究過程中，已經修過「教育測驗與評量」課程的 30 位職前教師，實際利用 WATA 系統在網路上執行與接受評量後發現：(1) 受試者認為小 WATA 系統(形成性評量策略)的「重複做題」、「查詢他人成績」等功能對於學科知識的學習有良好的輔助效果、(2)受試者將更能掌握評量流程，以及願意運用正確的評量策略來掌握學習者的迷思概念以改善教學，除此之外，也更願意主動了解自己的命題缺失，以提升自己的命題水準；(3)受試者能跳脫出過去只依賴參考書或考古題等現有資源的窠臼，以較積極的態度來面對評量。(4)影響受試者未來實際於教學環境中應用網路評量的主因除了軟硬體資源的充足與否外，其他教師對教育的看法也是影響的主要因素。在未來，應該針對各項師資培育的需要，開發各種可「模擬教學」的網路工具(Web Tools)，讓職前或實習教師親身使用，對同儕進行親身的模擬教學，如此對於師資培育訓練將有事半功倍的效果。

關鍵字：WATA (Web-based Assessment and Test Analysis System)、網路評量、試後分析、師資培育、學科教學能力(PCK)。

Improving Preservice Teachers' Assessment Knowledge with A Web-based Assessment System

Tzu-Hua Wang

Jen-Teh Junior College of Medicine, Nursing and Management

Kuo-Hua Wang

Graduate Institute of Science Education, National Chunghua University of Education

Wei-Lung Wang, Shih-Chieh Huang

Department of Biology, National Chunghua University of Education

Abstract

This research intended to use WATA (Web-based Assessment and test Analysis System), equipped with Triple-A (Assembling, Administrating, and Appraising) module which meant the functions of 'assembling test', 'administrating test' and 'test analysis', practically in assistance to the teacher education of pre-service biology teachers, to evaluate the effectiveness and practicalities of the Web-Based Instruction strategy of being a simulated teacher on the internet to teacher education. In the 4-month research, 30 of the pre-service teachers, who have taken the course of "Educational Test and Assessment," continuously used WATA system to administrate tests to peers and take peers' test on the Internet. The findings are as follows. First, the participants pointed out those functions, such as "answering questions repeatedly" and "querying others' grades," in WATA formative assessment module have positive effects to the absorption of biology knowledge. Second, the participants will be able to get hold of the standard assessment process more, and more willing to use correct assessment strategies to know learners' misconceptions and to improve teaching. Besides, the participants are more active in understanding the disadvantages of their own test items, and thus they can improve their item-construction abilities. Third, the participants can face the assessment more positively with wider ranges of selecting test items, not simply relying on reference books or the test papers in the past. Last but not least, the major reasons, which influence the subjects on the application of WATA to the actual instruction environment, are not only the adequate software and hardware resources but also the viewpoints of other teachers about education. In the future, to meet the needs of PCK training of teacher education, various online-instruction web tools of simulating teachers on the internet should be designed for pre-service teachers to actually practice alone or among peers. Thus, the training of PCK abilities will be greatly improved.

Keywords: WATA (Web-based Assessment and Test Analysis System) 、 Web-Based Assessment 、 Test Analysis 、 Teacher Education 、 Pedagogical Content Knowledge(PCK).

壹、緒論

隨著網路通訊與資訊技術的成熟與普及，有不少團體研發各種線上評量系統，試圖將傳統的紙筆測驗，轉移到電腦網路上進行，建立一個具有龐大題庫的評量系統，讓很多人一起「連線會考」[1]。這種評量模式對參與考試的學生而言，可以讓大家都知道來自各地考生的實力，分享認識各種不同評量題型，在無形中也可以縮短城鄉差距；對教育研究者而言，這套系統所蒐集的資料將有助於分析教學成效以及命題好壞，甚至可以分析出學生普遍的迷失概念，所以這種利用電腦網路突破時空限制的優點，進而達成讓各地考生一起進行評量的做法，在未來將有其發展性。

藉由電腦進行評量測驗中最著名的例子就是 GRE(Graduate Record Examination)，該測驗在 1992 年已用電腦版本進行考試，1993 年更以 CAT(Computerized Adaptive Test)的形式，利用人工智慧來自動調整試題，施行適性化評量，並預計未來將全面停止紙筆型式的考試，而採行電腦考試。除此之外，LPI(Linux Professional Institute)、MCSE(Microsoft Certificate Engineer)、CCNA(Cisco Certificate Network Administrator)等國際性的工程師執照考試也都採用網路評量的方式，由各地的考試中心(Testing Center)連線到特定的考試服務單位(Testing Service Provider)(例如:VUE, <http://www.vue.com>)，只要通過網路評量考試，就立刻發給證照，由此足見，網路評量的效益已經逐漸受到重視以及肯定。

本研究擬根據 Magnusson, Krajcik & Borko [2] 所建議之學科教學能力(Pedagogical Content Knowledge, PCK)師資培育課程中所應包含的成分(components)中：「對瞭解的知識(knowledge of students' understanding)」與「科學評量的知識(knowledge of science assessment)」二項，在網路上建立一個可以讓職前教師親自模擬與執行完整評量過程的環境，並可以讓職前教師有效率的掌握學生學習狀況與自己的命題良莠狀況。因此本研究擬利用網際網路評量系統作為輔助師資培育的工具，在此環境中，除了讓職前教師可以依照自己的方便性，異時異地的登入並且進行主動學習外，更要讓職前教師能夠藉著在網路上交替的扮演教師與學生的角色的機會，親自利用該網路評量系統實際執行(administrate)與參與(apply)網路評量，同時體會學生與教師在教學環境中的角色，而在不斷的角色轉變與扮演中獲得 PCK 能力的提升。

貳、文獻探討

一、電腦化評量(CBT)與網路化評量(WBT)的發展

以傳統教學的角度而言，「評量」與「教學」之間有相當密切的關係，有一些學者提出獨到的見解，國外學者 Gronlund & Linn [3]指出評量對於教師教學的意義在於：

(1)事先了解學習者的需求、(2)監控學習的過程、(3)診斷學習者學習的困難、(4)確定學習目標的達成狀況...(p.109)

國內學者林清山[4]亦指出，「評量對於教學過程而言是一個相當重要的步驟，由於教與學是循環的歷程，評量回饋是經驗成長的重要因素，評量提供教師掌握學生學習效果的訊息，藉著評量的回饋，教師能了解教學目標達到的情形，並可根據評量的資料修正目標，學生透過評量的回饋來了解學習的重點並且掌握自己學習的優點及弱點」。由此足見，教學與評量的實在是密不可分的，也因此，無論在過去的電腦輔助教學或是近幾年的網路教學相關研究中，應用最新的資訊技術輔助電腦評量(CBT, Computer-Based Test)或網路評量(WBT, Web-Based Test)的進行，就成為另一個重要的議題。

「電腦評量」，基本上就是將傳統紙筆測驗(如：選擇題)轉移到電腦上，讓學生透過電腦螢幕閱讀考題，利用鍵盤移動游標或用滑鼠點選答案，在評量系統通常允許學生重複檢查、修正答案，在確定完成之後送出試卷批改並離開測驗系統，另一些特殊考題形式如配合題、填充題或問答题，也可以透過程式撰寫，讓考試可以在區域網路(Local Area Network, LAN)上的個人電腦或工作stations上完成，換句話說就是採用電腦作為學生應試、教師命題與試卷傳遞的媒介[5]。近年來由於網際網路的互動與資料庫技術成熟，這些本來只能運作在於個人電腦或是區域網路(LAN)上的電腦評量系統(CBT)，逐漸轉移到網際網路(Internet)運作，而以 Web 為介面，使用者經由瀏覽器(Browser)就可以由世界各地隨時隨地的上來使用這評量系統，比起早期只建構在於個人電腦或是區域網路的電腦評量，例如：IBM [6]與 Arthur Andersen [7]公司所開發的 CAT 系統，在使用上更有彈性，對於教育研究上也將更有意義，而將這類電腦評量方式特稱為網路評量，用來區分只可以單機執行或是於區域網路上執行的封閉型電腦評量。

CBT 與 WBT 隨著使用的時機以及根據的立論不同共有兩大類，分別為常規(regular)與適性(adaptive)兩種形式的評量系統。「常規性評量系統」：只是建立一個電腦化的紙筆測驗，學生評量題目(assessment items)是相同的(identical)。「適性化評量系統」：其評量題目(assessment items)會隨著應考者上一題的答題狀況而有所改變，整份考卷是根據應考者的能力(abilities)與技能(skills)而特製的[8]。「適性化評量系統」又稱為 CAT(Computer Adaptive Test)，該系統是依據 IRT(Item-Response Theory) [9]所設計的，IRT 的基本假設是：「應試者應試是根據自己的能力進行的，而有較高能力的應試者會給予正確的答案[8]」，CAT 則是根據應試者不同能力(根據應試者對於前一個題目的答題狀況而斷定能力

不同)，由系統給予不同的考試題目，其主要優點在於正確與有效的評估確認應試者的認知能力(cognitive ability)或是能力限制(ability limit) [5]。由上述的定義可以發現，CAT 由於可以建置一個適性化的電腦評量環境，對於目前重視個別差異以及因材施教的教育環境下，無疑是相當好的一種電腦評量方式，尤其適合作為「診斷性評量」的工具，雖然立意很好，但是許多研究者仍舊反對將 CAT 應用於所有種類的評量上[10]，因為 CAT 的立論根據是 IRT，而 IRT 中基本要求是「每一個題目必須施測於大量的樣本上以確認每一個題目是有效的，如此才可以用來每一個應試者的能力分數(ability score) [5]」，換句話說，每一個題目都必須經過考驗，證實題目的代表性以及有效性，很顯然的，要建立一個可以進行 CAT 的電腦題庫難度高且成本也很高，較無實用與推廣性，因此 Mason et al. [5]認為，目前這種 CAT 並不適合於小型的評量或是教室評量，而一般的 CBT 或 WBT 將相對的比較有推廣使用價值。

近幾年，由於網路通訊與資料庫技術的提昇，新開發出來的評量系統均屬於可運作於網際網路上，以瀏覽器(Browser)作為介面的 WBT。這些 WBT 本身除了提供線上即時考試功能外，也做到線上出題、改考卷及登錄成績，另外因為這些線上評量系統可呈現 Web 格式的畫面，所以能取代傳統的紙筆考試，並可加入音效(Audio)和影像(Video)於試卷當中，甚至可加入 JAVA 設計的虛擬或動態畫面，以增加試題的靈活度。WBT 比較有名的有，學術團體開發的系統，例如：University of Nebraska at Lincoln(Dept. of Mathematics and Statistics) 針對微積分教學所開發的 Gateway Testing System 系統(<http://www.math.unl.edu/webtests/>)與 University of Illinois(Department of Electrical and Computer Engineering) 開發的 Mallard(<http://www.cen.uiuc.edu/Mallard/Demo/>)以及商業團體所開發的系統，例如：Question Mark Computing 開發的 Question Mark(<http://www.qmark.com>)與 Computer PREP 所開發的 Web@ssessor (<http://testdrive.webassessor.com/default.asp>)等[1]。

針對於評量系統的發展，國內外均有學者提出建議，Bonham et al.[1]指出，一個好的評量系統應該具備的特色與功能包括有：「能夠經由一般標準的瀏覽器進行網路連線、利用密碼判別使用者身分、自動評分、收集紀錄相關學生成績等四種功能的達成」；Alessi & Trollip[11]指出，應該能夠具備「題庫建置、形成測驗、舉行考試」；Kumar, Helgeson & White[12]指出，應該要能「被動收集考試評量歷程資料以及主動回饋引導學習與偵測學習迷思」。

Wang et al.[13] 依據上述建議，再整理 Gronlund & Linn [3, pp.109-141, p.228] 歸納出「完整評量流程」所應包含步驟：「確定測驗目的」、「建立雙向細目分析表」、「雙向細目表命題」、「選取適當的題型與適切的題目」、「組成試卷」、「執行考試」與「試卷之評定與結果的應用」以及實際針對在職教師進行問卷調查與訪談，指出一個完整的評量系統應該具備有 Triple-A Module (Assembling, Administrating and Appraising)，這 Triple-A module 分別為：

- 試題組成(Assembling)：教學者可以自建題庫並自編雙向細目分析表以及依據此表進行命題。
- 考試管理(Administrating)：教學者管理執行多重考試行程。
- 試後評估(Appraising)：教學者可以分析出各種試後分析與試題分析的資訊。

Wang et al.並針對其所提出的 Triple-A 架構，發展出了一套網路評量系統-WATA 系統(Web-based Assessment and Test Analysis System) [13,14,15,16,17,18,19,20,21,22,23]。

本研究是要探討網路評量系統應用於職前教師之師資培育的效益，因此擬採用具備 Gronlund & Linn 所指出之「完整評量流程」精神的網路評量系統，下表則是直接只以 Gronlund & Linn 的角度來比較上述文獻整理出的五套網路評量系統，如表 1 所示：

表 1：五套網路評量系統的比較

網路評量系統	Assembling				Administrating	Appraising
	確定測驗目的&建立雙向細目分析表(註一)	雙向細目表命題(註二)	選適當的題型與適切的題目(註三)	組成試卷(註四)	執行考試(註五)	試卷之評定與結果的應用(註六)
Gateway Testing System	X	X	◎	◎	◎	△
Mallard	X	X	◎	◎	◎	△
Question Mark	X	X	◎	◎	◎	△
Web@ssessor	X	X	◎	◎	◎	△
WATA	◎	◎	◎	◎	◎	◎

X：代表完全不具備此功能

◎：具備此功能

△：具備部分功能

註一：具備「以建立雙向細目分析表的方式確定測驗的目的」之功能

註二：具備「根據雙向細目分析表搜尋題庫命題」之功能

註三：具備「建立題庫」以及「勾選、搜尋適當的題庫題目」之功能

註四：具備有「形成試卷」之功能

註五：具備有「執行考試」、「設定考試執行方式」以及「考試行程管理」功能

註六：除了學生成績分析(原始分數、T 分數、名次、成績分布圖等)之外，尚須包含「試卷分析(標準差、變異數、KR20 等)」與「試題分析(鑑別度、難易度、誘答力分析等)」。

由上表可以發現，除了 WATA 系統之外，均沒有提供「雙向細目分析表」編輯以及針對該表命題的功能，而且在「試卷之評定與結果的應用」方面也有很多差異，例如：Gateway Testing System、Web@ssessor、Mallard 在本部分的功能只是簡單的原始成績呈現、統計圖表的呈現與試題對錯狀況，而沒有進行「試題分析」；Question Mark 則除了上述之外，會紀錄所有的學生作答過程資訊，而且會進行試題分析，但是仍無法更詳細分析學生成績(T 分數)與試卷分析(KR20)。基於以上的分析可知，以師資培育的角度來看，WATA 系統將較符合本研究的需求，因此擬採用 WATA 系統為本研究中的網路評量系統。

二、網路評量系統融入師資培育

近幾年來，國內一直持續有一些學者投身於將網路資訊技術融入師資培育用以提升師資培育的成效研究，並且都獲得具體的成果[24,25,26]，另外，亦有學者針對網路技術融入師資培育設計國科會專案計畫進行長期研究，例如：國立台灣師範大學的林陳涌、林金盾等教授所主持之國科會計畫：「生物科實習教師遠距輔導系統之研究—教學輔導系統的內涵與其效果考驗(NSC88-2511-S-003-052)」與國立彰化師範大學黃世傑、王國華等教授所主持的國科會計畫：「輔導中小學數理教師專業成長之遠距進修研究(NSC 90-2511-S-018-017)」、「運用網路資源系統提昇生物教師學科教學知能之研究—運用網路資源系統提昇生物教師學科教學知能之研究(NSC 89-2511-S-018-025)」也都對於網路技術輔助師資培育提出正面的研究結果。除了上述應用網路教材與網路輔助資源系統輔助師資培育之外，也有一些師資培育者嘗試將網路評量系統用以輔助師資培育的進行，例如：Ahern, Jamison & Olivarez[27]指出，在 Web 環境中開發師資培育工具是一個可行的方式，他們設計網路評量系統來協助職前教師的評量工作，而且發現受試之職前教師喜愛這樣的一個做法，也確實可行。王子華、黃世傑[14]應用「網際網路評量與試後分析系統(WATA)」做為進行實習教師的學科知識(Content Knowledge, CK)的評量工具，用此代替傳統的紙筆測驗，結果發現實習教師普遍認同(98%)認為以網路評量取代傳統紙筆測驗以輔助師資培育是可行的，且全部的實習教師都表示如果有機會可以接觸該系統，願意在未來執教學校採用網路評量。王子華、王瑋龍、黃世傑[15]應用 WATA 系統之形成性評量模組於「生物科教材教法」課程上，對生物科職前教師進行為期二個月 PCK(Pedagogical Content Knowledge)訓練，研究發現，職前教師願意主動利用這套系統輔助學習，而且在學習效益上有顯著的輔助效果。王子華、王瑋龍、王國華、黃世傑[16]更進一步應用 WATA 系統針對職前教師的評量能力進行訓練，在為期兩個月的研究後發現，職前教師除了對評量的觀念、理論精神與實施流程更能掌握之外，更願意主動利用評量來糾正自己命題的缺失、了解學習者迷思概念之所在以及改善教學。由此足見，網路評量系統無論在教學或是師資培育上都有其實用價值。

由上述可以發現，過去對於 WATA 系統應用於師資培育的研究，主要在於：由授課教授利用 WATA 系統在其網路教學環境中執行網路評量，用以探討該教學策略對於生物科職前教師在網路學習環境中的學科知識(Content Knowledge)學習效益[14,15]，以及將 WATA 系統的教師命題介面作為師資培育工具，用以評估 WATA 系統在培養職前教師的評量知能的效益[16]。而本研究，主要將過去的研究更進一步的整合以及延伸，直接應用網路評量系統做為師資培育的主要工具，讓職前教師能夠藉著這套系統在網路上交互的扮演教師與學生的角色，換言之，職前教師將可以應用 WATA 系統的教師介面模擬教師出題，並在網路上對同儕執行網路考試，也可以參與其他同儕所執行的網路考試，而本研究將探討在這樣交互的角色互換下，對於職前教師的評量知能的影響為何？除此之外，將更深入探討 WATA 系統中的特殊應考策略對於職前教師的學科知識的學習效益以及探討影響職前教師在未來執教時是否採用網路評量系統輔助教學活動的因素。

參、待答問題

本研究共有三個待答問題

1. WATA 系統中形成性評量模組(小 WATA)的四項功能設計(重複作題、不提供答案的策略、發問功能與查詢成績)，應用在職前教師之學科知識(Content Knowledge, CK)之培育的效益為何？
2. 採用 WATA 系統融入師資培育，對於職前教師的評量觀念之影響為何？
3. 影響職前教師在未來是否採用網路評量系統輔助教學活動進行的因素為何？

肆、研究方法

一、研究樣本

研究對象是某國立師範大學生物學系 90 學年度大三的職前教師(pre-service teacher)共 30 人(男 16 人與女 14 人)，他(她)們均擁有自己個人電腦者且可以利用宿舍網路環境上網。受試者在文書資料處理以及電腦打字方面能力都沒有問題。除此之外，也在大二時都已經修習過「教育測驗與評量」課程，因此對於評量理論已經有基本認識。

二、研究工具

(一) WATA (Web-based Assessment and Test Analysis System)系統

本研究擬採用「WATA 系統(Web-based Assessment and Test Analysis System)」，作為本研究中師資訓練的工具。WATA 系統屬於常規性的 WBT 系統，主要是應用於「總結性評量」與「形成性評量」，該系統具有 Triple-A(Assembling、Administrating、Appraising)架構，擁有高效率之「試題組成引擎」可以讓教學者管理考試行程與題庫與試卷的編製，「考試管理引擎」可以讓教學者掌握應考者的考試，「試後分析引擎」可以在幾秒鐘之內分析出所有試後分析的資訊。以下為個人化之 WATA 系統的畫面如圖 1：

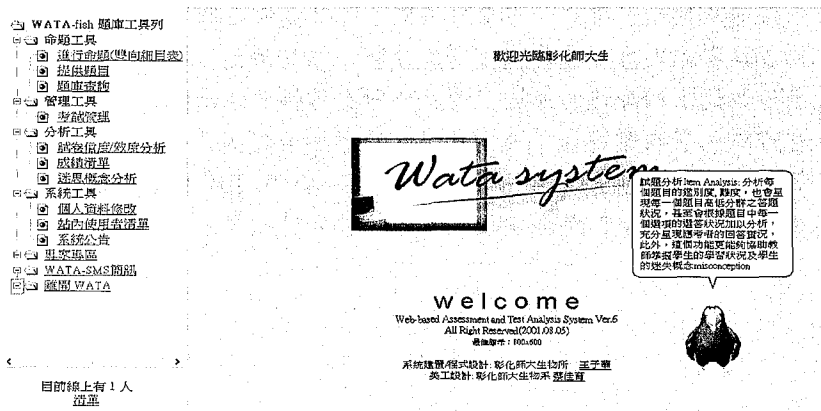


圖 1：WATA 系統個人化介面

WATA 的 Triple-A 架構簡介如下：

1. 試題組成引擎(Assembling Engine)

WATA 系統中 Assembling Engine 的功能執行模式如圖 2 所示：

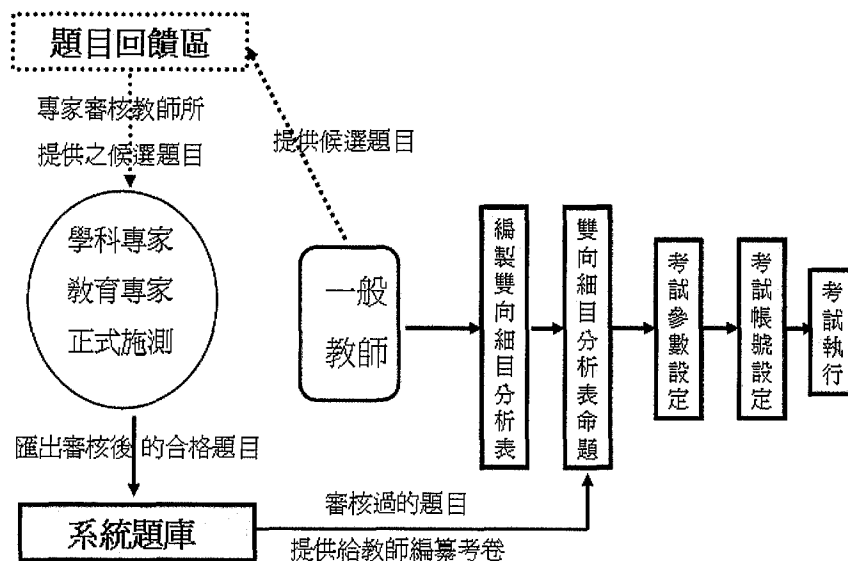


圖 2：Assembling Engine 執行模式圖

Assembling Engine 的功能細部運作介紹如下：

(1). 「屬性編號導向多媒體系統題庫建置」(圖 3)

WATA 題庫具備有「專家題庫」與「屬性編號」的概念。「專家題庫」的意義在於，系統題庫中的題目均來自於使用題庫的教師，教師將題目提供到「題目回饋區」之後，這些「候選題目」會經過學科與教育專家的編修、正式施測與修改，以建立題目的難易度與鑑別度。「候選題目」一但審核通過之後，編會匯入系統題庫，提供給予教師使用。「屬性編號」的意義在於，每一個系統題庫的題目都有一串屬性編號，用以代表這個题目的屬性，此描述的屬性包括有鑑別度、難易度、布倫分類、概念編號(九年一貫能力編號)，這串屬性編號巨細靡遺的描述了题目的所有特性。除此之外，題庫中每一個題目都可以是多媒體題目，因為系統支援考題呈現圖片、影像、聲音與 Flash.. 等任何可在瀏覽器上呈現的多媒體效果，並且可建立每一個题目的簡單回饋與說明，讓學生於應考完後，立即給予即時回饋。

新增題庫題目	
題目	
附圖	☐ 無附圖 ☐ 圖解: _____
章節	☐ 無附圖 ☐ 圖解: _____
鑑別度	☐ 影片 ☐ Flash
難易度	☐ 易 ☐ 中 ☐ 難
布倫分類	☐ 知識 ☐ 理解 ☐ 應用 ☐ 分析 ☐ 綜合 ☐ 評鑑
概念編號	_____
選項	_____ _____ _____ _____ _____
答案	_____
解答說明	_____
提交 重新設定	

圖 3：屬性編號導向多媒體題庫的建立畫面

(2). 「雙向細目表試卷編撰」(圖 4、圖 5)

在 WATA 系統內執行考試，必須遵守標準的命題流程，首先必須決定雙向細目分析表，以決定試卷的命題內容，才可以命題。

符合章節範圍的概念編號以及題庫資訊	
可選擇概念編號	概念敘述
號	(總題數/知識-理解-應用-分析-綜合-評鑑)
☐ 0301	植物取得養分的形式 (0/0-0-0-0-0-0)
☐ 0302	動物取得養分的形式 (1/0-1-0-0-0-0)
☐ 0303	提供能量的養分 (2/1-0-0-1-0-0)
☒ 0304	不具能量的養分 (5/4-1-0-0-0-0)
☒ 0305	氧化作用 (2/2-0-0-0-0-0)
☒ 0306	能量測定 (2/0-1-1-0-0-0)
☒ 0307	能量變化 (2/0-2-0-0-0-0)
☐ 0308	代謝作用 (3/1-0-2-0-0-0)
☐ 0309	酵素的性質 (4/4-0-0-0-0-0)
☐ 0310	酵素的活性 (4/1-0-2-1-0-0)
☐ 0311	行光合作用的要素 (3/2-1-0-0-0-0)
☐ 0312	葉綠體是行光合作用的場所 (2/2-0-0-0-0-0)
☐ 0313	光合作用的器官 (1/0-0-0-1-0-0)

圖 4：指定要出題的概念，由系統搜尋符合的題目

CK-03 - 第 3 --> 3 章 雙向細目分析表

*部分請務必輸入

概念/布倫	知識	理解	應用	分析	綜合	評鑑
0304	4	1	0	0	0	0
0305	2	0	0	0	0	0
0306	0	1	1	0	0	0
0307	0	2	0	0	0	0

提交 | 重新設定

圖 5：依據雙向細目表命題

(3). 「考試參數的設定」(圖 6)

依據網路考試的需求，教師可以設定考試的執行方式，包含「考試說明」、「考試範圍」，除此之外，也可以設定「是否顯示標準答案」、「是否寄送考卷成績單」、「是否允許學生查詢他人成績」、「是否允許重複做題」、「每一次提領試卷的題數」與「作答的時間限制」，另外，也可以設定教師與助教的 email 信箱，以便利師生之間的非同步互動，另外也提供是否啟動或關閉選項與題序亂數的功能。

參數名稱	參數內容
設定方式	○ 新建試卷(完全歸零, 含已有的成績資料) ◎ 只修改參數
考試說明	國中自然與生活科技小考
考試路徑	NI07
考試範圍	Ch7 - Ch7
是否顯示標準答案	◎ 是 ○ 否
是否寄送考卷成績單	◎ 是 ○ 否
是否允許學生查詢他人成績	○ 是 ◎ 否
是否允許學生重複作題	○ 是 ◎ 否
學生考試/答題數目(該考卷共有 50 題)	50 ▾
學生成績單/問題信件 正本教師信箱	
學生成績單/問題信件 副本教師信箱	
作答的時間限制(分鐘)	
關閉選項與題序亂數功能? ☐	
確定設定 重新設定	

圖 6：考試參數的設定

(4). 「多重考試行程的建立與管理」(圖 7)

可同時在 WATA 系統中舉辦多場考試，也可以隨意取消、終止或暫停；除此之外，教師對於每一個考試行程，都可設定其網路評量的實施方式，包含有：應考時間設定與限制、應考密碼設定、考生名單建立、是否顯示標準答案、是否自動以 E-mail 寄發成績單、是否提供考生查詢他人成績、設定學生試後發問的問題傳送 E-mail 位址..等。

fish 的考試行程

更新考試行程			
命題者	說明	範圍	時間
bio01 刪除 fish	> 設定考試參數 > (N/A) > 雙向分析表再命題 > 列出試卷內容 > 列印紙筆測驗 > 處理問題帳號成績 > 題目送分與答案變更 > 複製本行程	0 -- 1	2001 年 12 月 20 日 09 時 10 分 06 小時內可登入 登入密碼 30 修改 設定應考名單
bio02 刪除 fish	> 設定考試參數 > (N/A) > 雙向分析表再命題 > 列出試卷內容 > 列印紙筆測驗 > 處理問題帳號成績 > 題目送分與答案變更 > 複製本行程	2 -- 3	2001 年 12 月 12 日 07 時 20 分 1 小時內可登入 登入密碼 genbioeco 修改 設定應考名單
bio03 刪除 fish	> 設定考試參數 > (N/A) > 雙向分析表再命題 > 列出試卷內容 > 列印紙筆測驗 > 處理問題帳號成績 > 題目送分與答案變更 > 複製本行程	1 -- 6	2001 年 09 月 12 日 06 時 20 分 200 小時內可登入 登入密碼 fish 修改 設定應考名單

圖 7：多重考試行程管理畫面

2. 考試管理引擎(Administrating Engine)

WATA 的 Administrating Engine 架構中，特別注意到網路會考的安全性問題，因為網際網路是一個開放的連線環境，為了同時保留網際網路的便利性及安全性，WATA 設計了三層安全架構，所有應考者必須先輸入舉辦考試老師的帳號，並且擁有這個老師授與的特有應考密碼，再加上輸入正確的應考者帳號，才可以登入領取試卷，如此便可更有效地避免考試進行中未知的侵擾。在批改考卷方面，除了一般評量系統的登錄成績外，更會紀錄應考者答題資訊以供試後分析之用，除此之外，也會在學生送出答案批改之後，由系統立即告知成績及各題的對錯資訊並且以 E-mail 方式寄發學生成績單給學生個人與老師，另外，如果題庫中已經預先建立了該題簡單的回饋與說明，也會於此時列出，給予學生作為即時回饋的參考，藉以強化學習，此外，也提供非同步互動回饋的功能，讓應試者可藉由該功能以 E-mail 與教師互動討論問題。其詳細運作模式如圖 8 所示：

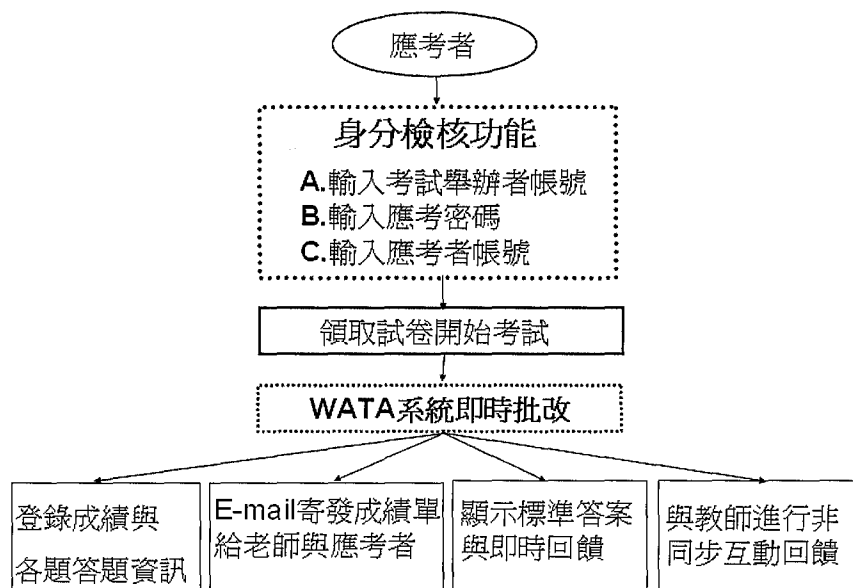


圖 8：Administrating Engine 運作模式圖

考試模式分為兩種，分別為「小 WATA 形成性評量」與「大 WATA 總結性評量」兩種模式，兩種模式的運作方式分別為：

(1). 小 WATA 形成性評量：

教師可以建立一個小型的題庫，而由 WATA 系統在每一次應考者領取試卷時只亂數選取一部分題目形成試卷顯示出來，而不會將所有題庫中的題目都同時顯示出來，系統也提供線上發問功能，教師也可以決定是否顯示標準答案或是以 E-mail 寄發成績單。系統運作的特色是，每一個題目的是否顯示、顯示的題序與選項順序都採用亂數處理。應考者可以重複作題，直到所有題目都做完為止，因為同一題目只要連續答對三次，系統就認定為真正答對而非猜對(因為在選項亂數的前提下，要三次都猜同一個答案的機會很低)，而該題將不在顯示，否則就歸零重新累計。

(2). 大 WATA 總結性評量：

乃一般總結性評量的模式，應考者不可以重複作題，應考者每一次提領試卷，WATA 系統都是顯示所有題目，但是每一份試卷的題續予選項順序都不同，除此之外，系統也提供線上發問功能，教師也可以決定是否顯示標準答案或是以 E-mail 寄發成績單。

3. 試後評估引擎(Appraising Engine)

WATA 系統除了於命題之前，有「雙向細目分析表命題」的功能來建立評量的效度外，也具備有 Appraising Engine 來進行「試後分析(Test Analysis)」的功能。WATA 系統分析應考者參與評量後所紀錄的資訊，分析出各種有用的資料，讓一份評量真的能夠發揮最大的功效，並有「診斷學生學習情形以提供教師教學參考的功能」，以下就是 WATA 系統所具備的「試後分析」，其具體做法包括試題分析(Item

Analysis)與試卷分析(Test Analysis)，WATA 系統採用的分析模式介紹如下圖 9 所示：

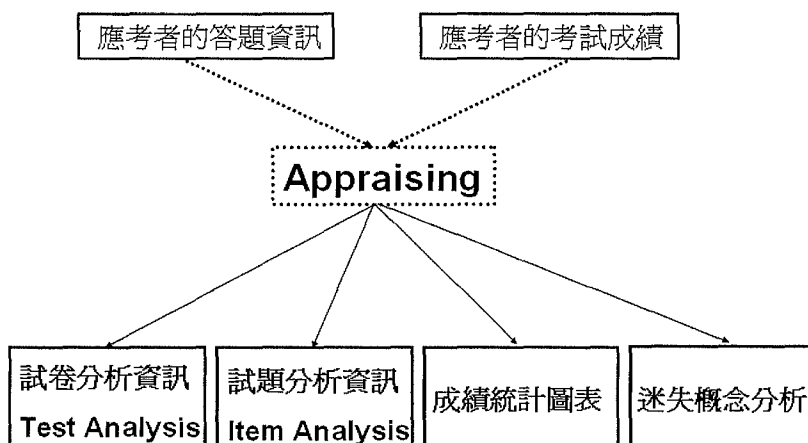


圖 9：Appraising Engine 運作模式圖

WATA 系統將紀錄於資料庫中應考者的答題過程資訊進行 Test Analysis 和 Item Analysis(圖 10-圖 11)，分析結果形成 Web 格式的試後分析報告清單，可以分別呈現出「全體考生成績的變異數、標準差、總平均」、「試卷的難度及 KR20 信度」、「每一題試題的鑑別度與難度」、「每一題試題高分群及低分群答題狀況」、「每一題學生答錯的答案」、「每一個题目的選項分析(誘答力分析)」...等資訊，提供教師參考，以利於教師掌握學生學習情況與自己教學的狀況，也可獲悉自己的命題是否恰當。相關分析資訊的定義如下：

- (1). 難度分析: 採用「通過率」的方式，也就是計算每一題平均的答對率，作為難度分析的指數[28]。
- (2). 鑑別度分析: 採用「鑑別度公式」，也就是該題「高分群答對比率減去低分群答對比率」，而高分群與低分群的認定方式，是以全班「常態化之 T 分數」前 27% 為高分群，而後面 27% 為低分群[28]。
- (3). z 分數與 Z 分數: 乃直接由原始分數直線轉換，z 分數為「原始分數的離均差除以全班原始成績標準差」，而 Z 分數由公式: 「 $50 + 10z$ 」，求得，所以該處的 z 與 Z 分數都屬於由原始分數線性轉化而來的成績，在教育上的利用價值不高，所以另有「常態化的 T 分數與 z 分數」，如下一項所介紹。[29]

$$z = \frac{X - \bar{X}}{S_x} \quad \begin{array}{l} X - \bar{X}: \text{離均差} \\ S_x: \text{全體成績的標準差} \end{array}$$

- (4). 常態化 z 分數與 T 分數: 是先將學生的成績排名，計算每一個考生勝過多少比率的人，然後經由常態分配表查表[29]，求出該考生成績於常態分配表上的 z

分數，然後以「 $50+10z$ 」求出 T 分數。

(5).信度的計算: 採用 Kuder & Richardson 於 1937 年提出的第 20 號公式-KR20 計算[31]，公式如下:

$$KR_{20} = \frac{n}{n-1} \left[1 - \frac{\sum pq}{Sx^2} \right]$$

n : 測驗的題數
 p : 答對某一題的人數比率
 q : 答錯某一題的人數比率
 $\sum pq$: 全部題目答對與答錯的百分比總乘積
 Sx^2 : 測驗總分的變異量

除此之外，也依據考生成績狀況與分布情形，自動繪出柱形圖與圓餅圖，提供教師圖形化的統計資訊，讓教師能更輕易掌握應考者成績分布與表現狀況。

試卷分析

人數	34
變異數	134.18
標準差	11.58
總平均	54.38
KR20	0.95
難度	48.92%

班級個人成績分析

學號/姓名	原始分數	z 分數[未常態]	Z 分數[未常態]
01	57.78	0.29	52.9
02	62.22	0.68	56.8
03	35.56	-1.61	33.9
04	48.89	-0.46	45.4
05	55.56	0.1	51
06	48.89	-0.46	45.4

班級個人標準化成績分析

名次	學號/姓名	原始分數	z 分數[常態化]	T 分數[常態化]
1	20	73.33	1.35	63.5
1	24	73.33	1.35	63.5
2	29	71.11	1.18	61.8
3	33	68.89	0.93	59.3
3	28	68.89	0.93	59.3
4	09	66.67	0.82	58.2
5	23	64.44	0.63	56.3
5	08	64.44	0.63	56.3
6	02	62.22	0.45	54.5
6	27	62.22	0.45	54.5
7	25	60	0.30	53

圖 10：試卷分析

試題/誘答力分析簡表

題目	A	B	C	D	E	N/A	Num	DP	ID
小英設計了下图裝置，乙探討唾液所含的酵素特性。40分鐘後，甲、乙、丙、丁四支試管都加入本氏液，並隔水加熱，結果只有甲試管變色成正反應。根據此一實驗，請問小英想探討酵素的何種特性(! fish! 357! 3! *! 普! 應用! i0336-i0334----+!) 圖解: (A)本氏液對唾液所含酵素有何作用 (B)溫度對唾液所含酵素有何作用 (C)鹼度對唾液所含酵素有何作用 (D)澱粉對唾液所含酵素有何作用 (E) 如果小孩穿白色褲子坐在草地上，褲子因為與草摩擦而沾染了綠色汁液，下列哪一種方法最容易把綠色汁液洗去(! fish! 365! 3! *! 普! 應用! i0337----+!) 圖解: (A)用肥皂水洗 (B)用自來水洗 (C)用沸水洗 (D)用酒精洗 (E)	25.21% H:0% L:21.43%	33.61% H:54.55% L:7.13%	4.2% H:9.09% L:14.29%	25.21% H:27.27% L:7.14%	0% H:0% L:0%	11.76% H:9.09% L:50%	119	0.47	34%
	16.81% H:0% L:14.29%	3.36% H:0% L:7.14%	14.29% H:0% L:7.14%	57.14% H:100% L:21.43%	0% H:0% L:0%	8.4% H:0% L:50%	119	0.79	57%

圖 11：試題分析、誘答力分析

(二) WATA 系統於職前教師評量理論與實務訓練成效評估問卷

該問卷共包含四種調查表，分別為「小 WATA 系統(形成性評量測驗模式)功能成效評估」、「大 WATA 系統對師資評量訓練的成效評估」、「大 WATA 系統對評量理論(想法)的改變評估」、「大小 WATA 系統融入師資評量訓練對評量活動實施之流程觀念改變評估」，四種調查表之內容依照需求不同，前兩種採用 Likert 五點格式，後二種為類別性測量(categorical measurement)格式。

該問卷之效度方面，經過專家效度與表面效度考驗，題目內容編製過程如下：

- 「小 WATA 系統(形成性評量測驗模式)功能成效評估」：
- 根據過去小 WATA 系統相關研究[21,23,32]中的調查表以及研究結果加以整理以及改編，再依據三位生物教育專家的意見加以修改，以建立專家效度與表面效度。
- 「大 WATA 系統對師資評量訓練的成效評估」、「大 WATA 系統對評量理論(想法)的改變評估」、「大小 WATA 系統融入師資評量訓練對評量活動實施之流程觀念改變評估」：
- 主要根據 Gronlund & Linn [3, pp.228-261]針對於評量的執行與精神的介紹設計題目內容，再經由三位生物教育專家的意見加以修改，以建立專家效度與表面效度。

該問卷之信度部份，分別針對使用過 WATA 系統的者實施預試：

- 「小 WATA 系統(形成性評量測驗模式)功能成效評估」：

該調查表共分為四個子調查表，本調查表針對 35 位某師範大學生物學系大一採用小 WATA 系統輔助網路教學的學生進行預試，其內部一致性 α 值分別

為.85、.91、.91、.95。

- 「大 WATA 系統對師資評量訓練的成效評估」：

該調查表共分為四個子調查表，本調查表針對 20 位使用過 WATA 系統的某師範大學生物學系大四學生進行預試，其內部一致性 α 值分別為.88、.93、.91、.95。

- 「大 WATA 系統對評量理論(想法)的改變評估」、「大小 WATA 系統融入師資評量訓練對評量活動實施之流程觀念改變評估」：

該調查表所得的資料均為二元性(binary)的類別資料，此兩調查表針對 20 位使用過 WATA 系統的某師範大學生物學系大四學生進行預試，其內部一致性 α 值分別為.71、.78。

三、研究設計與步驟

本研究採用個案研究的方式，所測試的教學課程是大三之「生物科教材教法」，這門課對於中等學校生物科職前教師而言是一門重要的必修課，該課程目的是用以訓練中等學校生物科職前教師的生物科學科教學能力(Biology Pedagogical Content Knowledge, PCK)。本研究對象為國立某師範大學生物學系對於該課程的教學方式是採用 WBI(Web-Based Instruction)的教學模式，但是實施方式迥異於一般的 WBI，因為整個 WBI 的教學內容與設計是由全體職前教師合力建構而成的，步驟如下：

- (一) 在進行本研究之前，先以「WATA 系統於職前教師評量理論與實務訓練成效評估問卷」中的「大 WATA 系統對評量理論(想法)的改變評估」與「大小 WATA 系統融入師資評量訓練對評量活動實施之流程觀念改變評估」兩子調查表進行前測，用以了解職前教師在研究進行前對於評量的基本態度與想法。接著，由研究者對職前教師施以為期一週的 WATA 系統操作訓練，使其了解 WATA 系統操作方式。
- (二) 授課教授首先將中等學校的生物課程教學內容分為 30 個主題(topics)，如：酵素、細胞分裂等，然後分配給三十位職前教師負責，每位職前教師需收集資料並且建立屬於所負責主題的 WBI 環境，並根據其網站教材內容命五十題單選題，作為評量網路學習效益之用。所有的評量試題，都必須附上九年一貫編號，該編號包含教育部所頒定的九年一貫自然與生活科技領域之「能力指標」，以及針對自己負責範圍的教材內容進行教材分析所設計出來之次編號。這些具有九年一貫編號的題目與 WBI 教材，必須經過任課教授修改與審核，才可以上網與建入 WATA 系統中。題目建入 WATA 系統後，則依照標準的命題程序形成與總結性評量試卷，作為評鑑 WBI 的學習效益之用。
- (三) 職前教師利用自己的 WBI 環境對其他二十九位同儕實施為期二週的教學，並安排一次的面授。二十九位同儕被要求在面授之前，必須先利用閒暇時間進入該主題負責人的教學網站進行學習，換言之，主題負責人儼然就是該主題的專家以及老師，必須扮演該主題之網路教學的教師角色。各主題的 WATA 總結性評量會在各個主題為期二週的 WBI 結束之後進行，而考試的方式採用異時異地的方式進行。
- (四) 主題負責人完成為期兩週的 WBI 教學之後，會公佈其所負責單元的 WATA 總結性評量實施時間，要求其他 29 位職前教師在一個期限內由任意地點上線完成考

試。在總結性評量完成後，利用 WATA 系統進行試後分析，而利用這些試後分析的資料，寫出一份報告，用以說明其他 29 位職前教師在其 WBI 中的學習效益與該如何改善教學，並且指出哪些題目品質不良以及該如何改善。

(五) 在為期四個月(一個學期)的課程與完成三十個單元的學習之後，以完整的「WATA 系統於職前教師評量理論與實務訓練成效評估問卷」實施調查。

四、資料收集與分析

本研究所收集的資料以量化資料為主，資料來自於「WATA 系統於職前教師評量理論與實務訓練成效評估問卷」，該問卷共四子調查表，在實施之後，將問卷中「小 WATA 系統(形成性評量測驗模式)功能成效評估」、「大 WATA 系統對師資評量訓練的成效評估」二子調查表加以整理，由於均為 Likert 五點格式(5 分代表非常同意、1 分代表非常不同意)，因此將其中平均得分高於 3.0 (四捨五入後)的題項判定為傾向同意，然後加以歸納整理出調查的結論。除此之外，會另外計算各題項不同性別者的平均分數，並以 SPSS 8.0 英文版施以獨立樣本 T 檢定，進行顯著性考驗分析，以了解性別變項是否影響結果。

其餘的兩子調查表-「大 WATA 系統對評量理論(想法)的改變評估」、「大小 WATA 系統融入師資評量訓練對評量活動實施之流程觀念改變評估」所得的前後測資料屬於二元類別資料，針對同意與不同意分別給予 1 與 0 分，之後以 SPSS 8.0 英文版施以無母數 Cochran 檢定，由於樣本僅有 30 人，所以檢定過程以 Monte Carlo 法在 Confidence Level 為.99 與 10000 個樣本下執行 Exact Test，以進行顯著性考驗分析。

伍、研究結果與討論

一、小 WATA 系統(形成性評量測驗模式)功能成效評估

在「小 WATA 系統(形成性評量測驗模式)功能成效評估」這部分的調查表是採用 Likert 五點量表格式，有效樣本為 30 人，該部分的內部一致性 $\alpha=.92$ ，且該部分所有的子調查表的內部一致性 α 也都大於.70，代表所有子調查表均具備有相當好的信度，由所有調查表整理出來的結果發現：

(一)「重複作題」功能

如表 2 所示，針對小 WATA 形成性評量「重複作題」功能，參與的職前教師認為可以協助找出學習的缺失，並且引發想把不懂觀念弄懂的動機，進而精熟課程內容。除此之外，也可以讓參與者找出哪些學習方向需要加強並使其對所學習到的觀念更有自信，此與 Buchanan[30,31]的研究一致。小 WATA 形成性評量所具備之獨特「連續答對三次才不會再度命題」的功能，使參與的職前教師養成更謹慎作答的態度，並對自己的學習狀況有所警覺，此外，如果能完成整個形成性評量題庫，會認為是一件很有成就感的事，此與王子華等[32]針對於應用小 WATA 形成性評量於大一普通生物學網路教學的研究結果一致。除此之外，各題項的作答狀況在性別變項的考驗上，均無顯著差異，因此，性

別差異並不影響該研究結果。

表 2：職前教師對小 WATA 的「重複作題」功能之反應與性別差異考驗(n=30)

題項大意	平均	SD	性別差異考驗		
			男 平均	女 平均	T 值
該功能可以幫我找出學習的缺失	3.5	.6	3.4	3.6	-.63
該功能可以幫我精熟課程內容	3.6	.6	3.8	3.4	1.98
該功能會引起我想把不瞭解懂的觀念弄清楚懂的動機	3.4	.8	3.3	3.5	-.89
該功能可以讓我哪些學習方向需要加強	3.7	.6	3.5	3.9	-1.51
該功能讓我對我所學習到的觀念更有自信	3.3	.6	3.3	3.4	-.18
該功能養成我更謹慎作答的態度	3.0	.7	3.2	2.9	1.37
當發現整個形成性評量題目都已經做完時，會很有成就感	3.5	.6	3.4	3.6	-.57
該功能會讓我對自己的學習狀況有所警覺	3.3	.6	3.3	3.4	-.80

內部一致性 $\alpha=.75$

(二)「不提供答案的策略」功能

如表 3 所示，參與的職前教師傾向於不認同小 WATA 形成性評量「不提供答案策略」的功能，他們不認為這樣的設計可以引發其釐清概念的動機，反而認為，如果直接給予答案更能夠讓參與者有機會去主動思考、主動搜尋課外資料來輔助學習、主動反覆閱讀教材，因而加強題目所附與的概念。此研究結果可以由 Buchanan[30]與王子華[23]的研究獲得說明，Buchanan 針對其所發展的網路形成性評量系統-PsyCAL[31]加以研究，指出

...當學生作答錯誤的時候，系統提供學生一個參考資料(reference)(摘自一本或更多本課本中恰當的片段)，學生將可以依據此找到正確的答案，而這樣的一個設計將可以使學生更主動融入於課程當中...[30, p195]

而王子華針對普通生物學生態學單元設計一個 GenBio-ECO WBI 進行網路教學，也採用 WATA 系統進行形成性評量，並且以「重複作題」與「不提供答案」作為後設認知策略，研究發現，受試者不認為「不提供答案」具備有引發後設認知能力的功能，並且在質性資料中發現：「不提供答案的設計會使學習者於尋找答案時沒有方向而成效不彰，並降低學習者的學習效益...沒有答案較不易為學習者所接受，因為這會增加學習者的不確定感」，除此之外，研究者也提出該功能具體改善的建議：

...「有考試沒有答案、但有提示」會比「有考試沒有答案」跟「有考試就有答案」來的好。因為「考試必有答案」沒有一個讓學生加深印象的延宕過程來強化學習者吸收到的正確知識。故提升評量考試效益的最佳方式，是在試後給予學生提示而非答案，使有思考的空間並將所學內化之...[23, p74-75]

因此在未來的設計中，應該在不直接提供答案的狀況下，可提供答題提示，才可以

讓這個功能發揮更大的輔助學習效果。除此之外，各題項的作答狀況在性別變項的考驗上，均無顯著差異，因此，性別差異並不影響該研究結果。

表 3：職前教師對小 WATA 的「不提供答案的策略」功能之反應與性別差異考驗(n=30)

題項大意	平均	SD	性別差異考驗		
			男 平均	女 平均	T 值
該功能比直接給答案更能引起我想把不懂的觀念弄懂的動機	2.6	.9	2.3	2.9	-1.64
該功能比直接給答案讓我更有機會去主動思考	2.9	.7	2.8	3.1	-1.16
該功能比直接給答案讓我更想搜尋課外資料來輔助學習	2.9	.9	2.8	3.1	-1.19
該功能比直接給予答案讓我對題目印象更為深刻	2.7	.9	2.6	2.8	-1.85
該功能比直接給予答案會讓我有更多機會主動反覆閱讀教材	2.9	.9	2.8	3.2	-1.23

內部一致性: $\alpha=.87$

(三)「發問功能」

如表 4 所示，參與的職前教師認為小 WATA 形成性評量「發問」功能可以提供一個更有效率的途徑來破除迷思概念，並可以使他們養成更主動的學習態度，而更願意主動向老師發問並解決學習上的問題，而獲得更多的延伸學習的機會。除此之外，他們認為寫作式發問功能，可以讓他們在發問的過程中將題目的內涵更具體地與自己其他的學習經驗發生連結並產生新的想法，也可以使他們在發問的時候對題目觀念更為清楚，並更能具體地了解自己學習缺失之所在，此與王子華[23]的研究結果一致，王子華指出，「小 WATA 形成性評量發問功能」對於網路學習者的後設認知能力有良好的輔助效果，其效果在研究者所設計的 23 種後設認知策略當中排名第六，而且對於四個向度的後設認知能力(目標設定、自我監控、自我修正與自我評鑑)的「自我修正」部份最好。除了上述的正向結果外，也有一些負面的結果，因為他們認為這樣的設計並不能使其獲得更好與老師互動的機會，而讓他們更願意上來參與測驗，可能的原因為，本研究中的老師角色，並不是真的由授課教授擔任，而是由單元負責的同學擔任，換言之，這些老師其實是同班同學，所以，如果真的有問題需要詢問，平常見面詢問或用電話會更加直接與便利，因此喪失它的效果。針對無法引起參與者更願意上來進行測驗的現象，其原因應該是在於，小 WATA 形成性評量本身就是一個考試環境，因此具備有嚴肅的特質，故參與者會較不喜歡上來作答。除此之外，各題項的作答狀況在性別變項的考驗上，均無顯著差異，因此，性別差異並不影響該研究結果。

表 4：職前教師對小 WATA 的「發問功能」之反應與性別差異考驗(n=30)

題項大意	平均	SD	性別差異考驗		
			男 平均	女 平均	T 值
該功能可以得到比傳統教學還要多與老師互動的機會	2.8	.7	2.9	2.8	.32
該功能可以提供我一個更有效率的途徑來破除迷思概念	3.1	.7	3.1	3.2	-.56
該功能讓我輕易針對特定問題向老師請益而獲得更多的學習	3.3	.8	3.3	3.2	.38
該功能讓我覺得老師就在我身旁，我隨時可以向他請教	3.0	.8	3.3	2.8	1.61
該功能讓我成績單除了有成績外還能讓我有延伸學習的機會	3.4	.5	3.4	3.4	.10
該功能讓我更願意上來進行測驗學習	2.7	.8	2.6	2.9	-.76
該功能讓我在決定發問問題的過程中，對題目觀念更為清楚	3.1	.7	3.0	3.2	-.77
藉由發問與老師互動的功能會讓我養成更主動學習的態度	3.0	.6	2.9	3.1	-1.26
該功能讓我在發問時更能具體地了解自己學習缺失之所在	3.3	.7	3.2	3.4	-.94
該功能讓我更願意主動向老師發問並解決學習上的問題	3.1	.8	2.9	3.3	-1.19
寫作式發問功能讓我在發問的過程中將題目的內涵更具體地與我其他的學習經驗發生連結並產生新的想法	3.3	.7	3.4	3.2	.85

內部一致性 $\alpha=.87$

(四)「查詢成績」功能

如表 5 所示，小 WATA 形成性評量「查詢成績」功能，參與的職前教師認為，可以讓他們更清楚自己在班上的排名、更清楚他人的學習狀況、引起學習動機與激起見賢思齊之感。除此之外，也認為會引起班上同學產生良性競爭、適時地掌握自己的學習狀況與主動修正自己的讀書方法或態度，此與王子華等[32]針對於應用小 WATA 形成性評量於大一普通生物學網路教學的研究結果一致。除此之外，各題項的作答狀況在性別變項的考驗上，均無顯著差異，因此，性別差異並不影響該研究結果。

表 5：職前教師對小 WATA 的「查詢成績」之反應與性別差異比較(n=30)

題項大意	平均	S D	性別差異考驗		T 值
			男 平均	女 平均	
查詢他人的成績會讓我更清楚自己在班上的排名	3.7	.7	3.7	3.6	.17
查詢他人的成績會讓我更清楚他人的學習狀況	3.6	.8	3.6	3.5	.41
藉由查詢了解他人的成績會引起我的學習動機	3.2	.9	3.4	2.9	1.42
查詢他人的成績會激起我有見賢思齊之感	3.4	.8	3.6	3.1	1.44
我認為查詢他人成績的功能會引起班上同學產生良性競爭	3.4	.8	3.4	3.3	.54
查詢自己的成績可以讓我適時地掌握自己的學習狀況	3.7	.7	3.7	3.7	-.10
查詢自己或他人的成績會引起我修正自己的讀書方法或態度	3.1	.8	3.3	3.0	.88

內部一致性 $\alpha=.92$

二、大 WATA 系統對師資評量訓練的成效評估

在「大 WATA 系統對師資評量訓練的成效評估」這部分的調查表是採用 Likert 五點量表格式，有效樣本為 30 人，該部分的內部一致性 $\alpha=.96$ ，且該部分所有子調查表的內部一致性 α 也都大於 .70，代表所有子調查表均具備有相當好的信度，由所有調查表整理出來的結果發現：

(一)「雙向細目分析表」的功能

如表 6 所示，WATA 系統中的「雙向細目分析表」功能應用於師資培育上，參與的職前教師普遍認為可訓練自己在未來命題時主動採用「雙向細目分析表」進行命題。此外，該功能可以使他們對實施測驗的教材更加熟悉，並對這些教材概念適當的分配題數，如此可以強化他們在這部分的概念學習，並能針對不同程度的測驗學生，選擇不同難度的考題形成試卷，這可以由 Gronlund & Linn[3, pp.112-117]的論點獲得解釋，Gronlund & Linn 指出，教師命題前需要先針對於教材「準備好其教學目標(preparing a list of instructional objectives)」、「摘要出教材內容(outline the course content)」，才可以進行雙向細目分析表的編製，換言之，也就是必須對於授課以及教材內容與架構有一定程度的熟悉，才可以編製好一份雙向細目分析表。除此之外，這樣的功能，可以訓練他掌握試卷中的命題資訊，更進一步地讓他們養成試後進行詳細的迷思概念診斷與分析之習慣。除此之外，各題項的作答狀況在性別變項的考驗上，均無顯著差異，因此，性別差異並不影響該研究結果。

表 6：「雙向細目分析表功能」對職前教師評量訓練的成效評估與性別差異考驗(n=30)

題項大意	平均	SD	性別差異考驗		
			男 平均	女 平均	T 值
訓練自己對測驗教材的熟悉度並選出適當的考題以形成試卷	3.2	.6	3.2	3.3	-.47
訓練自己對不同程度的學生，選擇不同難度的考題形成試卷	3.4	.5	3.4	3.4	.43
訓練自己依據考試範圍內的教材概念，適當分配題數，強化概念的學習	3.4	.7	3.5	3.4	.53
訓練自己詳掌握命題資訊，並可隨時更動試題	3.3	.6	3.6	3.3	1.22
可以訓練自己試後進行詳細的迷思概念的診斷與分析	3.3	.6	3.3	3.3	-.15
可以訓練自己養成未來命題時主動採用此法進行命題	3.2	.7	3.3	3.2	.13

內部一致性 $\alpha=.81$

(二)「教材概念編號設計」的功能

如表 7 所示，WATA 系統中強迫每一道題目都必須具備有自己專屬的「教材概念編號」，這種代表題目屬性的編號，對於職前教師而言，具備有輔助其命題的功能，也可以讓題庫中的題目容易搜尋，此與 Gronlund & Linn[3, p259]的看法一致。除此之外，這樣一個設計應用於師資培育上，也可以提升他們對教材的熟悉度，訓練其整理出學習內容主概念與次概念的能力，與訓練他們依據不同的學習階段(如：國小、國中)，設定不同的學習概念。而且，該設計也可以讓他們練習教材概念編號設計，並具備獨立命題的能力，而主動掌握題目的命題方向，他們也認為這些編號與雙向細目分析表搭配起來，可以讓他們於試後分析時，作為迷思概念偵測及診斷的依據來源，這是由於 WATA 系統的試後分析功能，是以這些編號作為依據進行分析。除此之外，各題項的作答狀況在性別變項的考驗上，均無顯著差異，因此，性別差異並不影響該研究結果。

表 7：「教材概念編號設計」對職前教師評量訓練的成效評估與性別差異考驗(n=30)

題項大意	平均	SD	性別差異考驗		
			男 平均	女 平均	T 值
可增加自己對測驗教材的熟悉度與擬定出學習主概念、學習次概念	3.2	.8	3.2	3.2	-.09
可訓練自己依據不同的學習階段(如：國小、國中)，設定不同的學習概念	3.3	.7	3.3	3.3	-.14
可建立自己獨立命題的學習概念藍本，掌握題目的命題方向	3.3	.7	3.3	3.3	-.14
可訓練自己利用雙向細目分析表進行命題	3.3	.6	3.3	3.4	-.44
可以訓練自己試後分析，並作為迷思概念偵測的依據來源	3.3	.7	3.3	3.4	-.68

內部一致性 $\alpha=.90$

(三)「個人題庫建立與分享」的功能

如表 8 所示，在採用 WATA 系統進行師資培育的過程中，要求職前教師建立自己的個人題庫並與他人分享，這樣一個環境設計可以訓練他們主動建立一套完整的個人題庫以供日後線上命題使用，並養成隨時保存自我命題的題目，以隨時作適當修正的習慣，另外，他們也認為，這樣的設計除了能隨時與他人分享題目，更可以使自己接受並樂於在網路上與同好合作發展題庫。除此之外，參與的實習教師認為，個人題庫的建立與分享，使自己能隨時與他人分享題目，如此讓自己在觀摩他人題目的同時激發命題的創意，這樣的結果可以由 Falchikov[33]、Freeman[34]與林珊如[35]等學者針對於「同儕互評(peer coaching)」的教學策略之相關研究獲得解釋，Falchikov 與 Freeman 等學者建議把「同儕互評」融入各種專業領域的訓練之中，因為這樣的設計將可以學生們在相互觀摩、學習過程中，達到較佳的學習效果；國內學者林珊如[35]也指出，觀摩(cognitive modeling)別人的學習成果或模範作品，是促使自己進步與反思的不二法門，而本研究中的「題庫分享」就有這個特質，因為在題庫分享的狀況下，建立題庫的教師們會互相觀摩與參考到同儕的命題，可以藉此交流與提升命題技巧或是命題創意。除此之外，各題項的作答狀況在性別變項的考驗上，均無顯著差異，因此，性別差異並不影響該研究結果。

表 8：「個人題庫建立與分享」對職前教師評量訓練的成效評估與性別差異考驗(n=30)

題項大意	平均	SD	性別差異考驗		
			男 平均	女 平均	T 值
可以訓練自己建立一套完整個人題庫，供日後線上命題使用	3.7	.6	3.6	3.8	-.67
可以訓練自己隨時保存自我命題的題目，並作適當的修正	3.8	.5	3.8	3.8	.13
可以訓練自己隨時與他人分享題目，使自己在觀摩他人題目的同時主動激發命題的創意	3.7	.6	3.8	3.7	.16
可以訓練自己隨時與他人分享題目，使自己接受並樂於在網路上與同好合作發展題庫	3.7	.7	3.6	3.8	-.58

內部一致性 $\alpha=.87$

(四)「WATA 試後分析」的功能

如表 9 所示，參與的職前教師普遍(平均大於 3.0)認為，WATA 系統中的「試後分析工具」應用於師資培育時，可以讓自己對於「教育測驗與評量」中的評量理論與實際應用更為了解，也更願意在考試完後進行試後分析，以掌握學習狀況與命題品質。他們認為，由於系統可以輕易進行試題鑑別度、難易度與誘答力指數等分析，因而可以讓他們實際地應用評量理論來定義自己命題的良莠，並進一步練習修改題目作為未來命題參考，此研究結果與王子華、黃世傑[14]針對 44 位民國八十九年畢業之生物科實習教師之研究結果一致。除此之外，他們也認為，這樣的功能可訓練自己依據試卷雙向細目分析表中所涵蓋的學習概念，進行學生學習成效評估、教學成效的評估，以作為改進教學的參考。除上述外，他們也認為這樣的功能可以訓練自己偵測出學生易發生混淆之迷思概

念的能力，以協助自己進行試後複習與個別化補救教學；而誘答力指數的分析，因可呈現出每一個選項高低分群的選答狀況，不但可訓練監控與改善每一個選項的品質外，更可訓練自己進一步掌握高、低分群之迷失概念所在，讓自己的教學作更有效的改善，此研究結果與 Gronlund & Linn[3, pp.244-247]所提出的看法一致。除此之外，各題項的作答狀況在性別變項的考驗上，均無顯著差異，因此，性別差異並不影響該研究結果。

表 9：「WATA 試後分析工具」對職前教師評量訓練的成效評估與性別差異考驗(n=30)

題項大意	平均	SD	性別差異考驗		
			男 平均	女 平均	T 值
訓練依據試卷雙向細目分析表進行學生學習成效的評估	3.4	.6	3.4	3.5	-.54
訓練依據試卷雙向細目分析表進行教師教學成效的評估，作為改進教學的參考	3.5	.7	3.5	3.4	.28
訓練依據試卷雙向細目分析表偵測出學生易發生學習混淆的概念，進行試後的複習及補救教學	3.4	.6	3.5	3.4	.62
訓練依據試卷雙向細目分析表偵測不同學生所產生的不同迷失概念，進行個別化補救教學	3.6	.6	3.6	3.5	.54
試後分析工具，可訓練進行兩次考試的概念分析比較，評估自己兩次教學中學生進步的成效	3.6	.7	3.5	3.7	-.80
試後分析工具的鑑別度指數，可訓練找出鑑別度不良的題目，加以修改提昇命題題目的品質	3.4	.7	3.5	3.4	.57
試後分析工具的難易度指數，可訓練找出難易度不良的題目，加以修改提昇命題題目的品質	3.5	.7	3.6	3.4	.76
試後分析工具的誘答力指數，可訓練找出誘答力不良的題目，加以修改，提昇命題題目的品質	3.6	.6	3.7	3.6	.23
試後分析工具的誘答力指數，可訓練找出學生主要錯誤的選項，了解學習迷失以改善教學	3.6	.7	3.6	3.6	-.08
同上，可訓練找出學生主要錯誤的選項，了解高低學習成就學生學習迷思之所在，用以改善教學	3.7	.7	3.7	3.6	.44
試後分析工具自動估算試題鑑別度與難易度，可以訓練判斷題目的品質，以供未來命題的修正	3.6	.7	3.6	3.6	.23
試後分析工具紀錄每題試題高低分群學生答錯狀況，訓練找出教學過程中遺漏的重要學習概念	3.5	.7	3.5	3.6	-.28
試後分析工具分析全體考生成績，有助訓練掌握各班學習狀況，並比較各班差異	3.6	.7	3.6	3.6	-.07
試後分析工具分析全體考生成績變異數及標準差，有助訓練了解整體學生程度分布狀況	3.4	.7	3.4	3.4	.03
試後分析工具自動估算考生 T 分數與 Z 分數，有助訓練掌握考生成績的進退步情形	3.4	.8	3.4	3.4	.03
試後分析工具自動估算考生 T 分數與 Z 分數，有助訓練比較兩個不同班級的成績與表現狀況	3.4	.8	3.4	3.4	-.18
試後分析工具分析整份試卷難度，可訓練比較不同班級間整體學生程度差異度	3.5	.6	3.4	3.6	-.89
試後分析工具分析整份試卷 KR20 信度，能訓練分析試卷評量成績的可信程度	3.4	.8	3.4	3.5	-.43
試後分析工具自動繪出成績分布長條圖與圓餅圖，能訓練了解成績分布狀況再次命題的參考值	3.5	.8	3.6	3.5	.22
試後分析的功能讓我對於「教育測驗與評量」中的評量的理論與其實際應用更為了解	3.3	.7	3.2	3.4	-.88
試後分析的功能讓我更願意在進行完考試之後進行試後分析，以掌握學習狀況與命題品質	3.5	.6	3.4	3.6	-.85

內部一致性 $\alpha=.96$

三、大 WATA 系統對評量理論(想法)的改變評估

本項研究結果主要來自於分析研究對象在「大 WATA 系統對評量理論(想法)的改變評估」問卷之反應而得，該部分資料屬於二元類別資料，認同題項敘述者給予 1 分，不認同者給予 0 分，有效樣本為 30 人，估算出每一個題項之認同百分率，並採用 Cochran 檢定併以 Monte Carlo 法執行 Exact Test，以進行顯著性考驗分析。資料分析後發現，參與的職前教師在經過 WATA 訓練後，在「認為評量所能達成的目標」方面有發生變化(如表 10)，尤其在變得更能認同評量可以「分析個別學生所產生的不同迷思概念、分析及改善教師的教學策略、了解題目命題的優劣、了解題目中各選項命題的優劣與了解整份試卷命題的優劣」五個方面最為顯著，這種現象無論男性或是女性皆是如此。

表 10：職前教師在 WATA 訓練前後認為評量所能達成的目標調查與性別差異考驗(n=30)

認為評量能達成的目標	WATA 訓練前 (%)			WATA 訓練後 (%)			CochranQ 檢定統計量		
	男	女	全體	男	女	全體	男	女	全體
鑑別學生的學習成效	81.3	64.3	73.3	93.8	78.6	86.7	2.00	0.67	2.00
瞭解學生容易產生學習混淆教材概念	75.0	64.3	70.0	75.0	78.6	76.7	0.00	0.67	0.40
分析個別學生所產生的不同迷思概念	31.3	14.3	23.3	81.3	57.1	70.0	6.40*	6.00*	10.89**
分析及改善教師教學策略	6.3	14.3	10.0	68.8	57.1	63.3	10.00**	4.5*	14.22**
了解題目命題的優劣	25.0	14.3	20.0	75.0	78.6	76.7	5.33*	9.00**	13.76**
了解題目中各選項命題的優劣	12.5	21.4	16.7	87.5	92.9	90.0	12.0**	10.00**	22.00**
了解整份試卷命題的優劣	25.0	28.6	26.7	75.0	85.7	80.0	5.33*	8.00**	12.80**

* p<.05

** p<.01

前測內部一致性: $\alpha=.65$, 後測內部一致性: $\alpha=.89$

另外，針對在未來擔任實習教師時會不會使用 WATA 系統協助實施評量進行調查，在研究對象 30 人中有 19 人(63%)表示願意採用，而在表示願意在實習時使用 WATA 系統的 19 人當中，有 13 人(68%)表示會排除萬難使用 WATA 系統，而其餘 6 人(32%)表示不會排除萬難使用 WATA 系統。另外，預期在國中高中環境中利用 WATA 系統進行網路評量會遭遇的問題如表 11：

表 11：職前教師預期採用 WATA 系統會遭遇的問題(人數百分比)

認為會遭遇問題	願意使用 WATA 系統者 (n=19)	不願意使用 WATA 系統者 (n=11)	願意使用但不會排除萬難使用 WATA 者(n=6)	全體 (n=30)
電腦教室的安排	19(100%)	10(91%)	6(100%)	29(97%)
學校教師異樣眼光	6(32%)	3(27%)	4(67%)	11(37%)
學生操作不熟練	8(42%)	9(82%)	4(67%)	17(57%)
網路連線速度太慢	13(68%)	6(55%)	5(83%)	20(67%)

由此可見，大多數願意在實習時採用 WATA 系統者，認為「電腦硬體資源」(連線狀況與電腦教室)是影響職前教師使用意願的最重要因素，其次才是學生的操作和同事的眼光，此與王子華等[14]針對民國 89 年畢業進行實習的 47 位生物科實習教師所做的研究結果一致。另外，對於不願意使用者而言，問題主要還是在於「電腦硬體資源」與「學生操作的熟練度」，而願意使用 WATA 系統但「不會」排除萬難使用者，問題除了「電腦硬體資源」與「學生操作的熟練度」外，也在乎「學校教師的異樣眼光」，此與王子華等[14]研究結果一致。針對於「學校教師的異樣眼光」的影響部份，可以由 Zeichner, Tabachnick & Densmore [36]針對影響初任教師之教學價值觀的研究結果獲得解釋，Zeichner et. al [36]指出「初任教師會將有經驗的同儕視為對他們學習教學影響最大的人…同儕中的常模也是一個重要的影響因素」，因此，學校中其他資深或是有經驗教師的看法，是影響實習教師是否採用 WATA 系統來進行網路評量的重要因素之一。上述這些因素都是未來在推廣 WATA 系統的普及或是網路評量時要注意到的問題，也是在推廣使用網路評量者應該注意到的議題。

四、大小 WATA 系統融入師資評量訓練對評量活動實施之流程觀念改變評估

該部分主要用以針對「評量活動實施流程之觀念改變」方面進行調查，該部分資料屬於二元類別資料，認同題項敘述者給予 1 分，不認同者給予 0 分，有效樣本為 30 人，估算出每一個題項之認同百分率，並採用 Cochran 檢定併以 Monte Carlo 法執行 Exact Test，以進行顯著性考驗分析。資料分析後發現，分析後發現，參與的職前教師在經過 WATA 訓練後在評量活動實施流程方面有改變。參與的職前教師經過在網路上以 WATA 系統實際模擬標準的命題流程，使得他們對於整個流程有更具體的概念，而使其未來在進行評量時遵循更完整的步驟。其中除了標準命題流程的「組成試卷、執行考試與試卷成績評定」三項步驟，由於一般的評量也一定會執行此步驟而沒有顯著改變外，其餘皆有顯著變化，這樣的結果無論是男性或市女性皆相同。如下表 12 所示：

表 12：職前教師在 WATA 訓練前後對評量活動實施流程之觀念改變調查與性別差異考驗
(n=30)

會進行之流程或步驟	WATA 訓練前 (%)			WATA 訓練後 (%)			CochranQ 檢定統計量		
	男	女	全體	男	女	全體	男	女	全體
確定測驗目的	56.3	64.3	60.0	87.5	100	93.3	5.00*	6.00*	11.00**
建立命題的雙向細目分析表	25	21.4	23.3	93.8	100	96.7	11.00**	11.00**	21.00**
依照雙向細目表命題	25	21.4	23.3	93.8	100	93.3	11.00**	11.00**	22.00**
選取適當的題型與適切的題目	50	71.4	60.0	93.8	92.9	100	7.00*	4.00*	11.00**
組成試卷	100	100	100	100	100	100	N/A	N/A	N/A
執行考試	100	100	100	100	100	100	N/A	N/A	N/A
試卷成績的評定	100	100	100	100	100	100	N/A	N/A	N/A
學習概念分析與結果的應用	25	21.4	23.3	100	100	100	12.00**	11.00**	23.00**

* p<.05

** p<.01

N/A 由於 WATA 訓練前後的填答狀況均無顯著差異，所以無法進行 CochranQ 檢定
前測信度: $\alpha=.73$, 後測信度: $\alpha=.92$

陸、結論與建議

一、結論

由於電腦網路技術的發展，利用網路於教育的相關研究有越來越多的趨勢，尤其是 e-Learning 的相關研究更是受到重視，但是直接利用網路來輔助師資培育的相關研究目前還不是很多，本研究旨在 Web 環境中開發師資培育工具，設計一個網路評量系統 (WATA) 來協助職前教師發展學科教學知能，結果發現：

(一) 小 WATA 形成性評量所具備的四個策略-「重複作題」、「不提供答案的策略」、「發問功能」與「查詢成績」能夠增進職前教師對學科知識的瞭解，而此四個策略各有不同的功能，以「重複作題」策略而言，可以協助職前教師找出學習的缺失、指引學習方向、對所學習到的觀念更有自信，進而精熟課程內容，而其中的「連續答對三次才不會再度出題」的設計，使參與的職前教師養成更謹慎作答的態度，並對自己的學習狀況有所警覺。「發問的功能」可以提供機會讓參與的職前教師養成更主動的學習態度，並且可使職前教師輕易針對特定的問題向老師請益而獲得更多的延伸學習的機會。除此之外，這種功能可以讓他們在發問的過程中將題目的內涵更具體地與自己其他的學習經驗發生連結並產生新的想法，更能具體地了解自己學習缺失之所在。「查詢成績」功能，可以讓他們更清楚自己在班上的排名、更清楚他人的學習狀況、引起學習動機與激起見賢思齊之感，並引起班上同學產生良性競爭，促使自

已掌握自己的學習狀況與主動修正自己的讀書方法。

- (二)參與研究的職前教師在使用 WATA 評量系統四個月後，無論男性或女性職前教師，均對評量功能的認知有顯著的改變。在未接觸 WATA 系統前(只有修過一般的教育測驗與評量課程)，他們認為評量的功能主要為：「鑑別學生的學習成效」、「瞭解學生容易產生學習混淆的教材概念」與「了解題目命題的優劣」，而接觸 WATA 系統之後，除了上述功能之外，也認為評量的功能具備有「分析個別學生所產生的不同迷思概念」、「分析及改善教師的教學策略」、「了解題目中各選項命題的優劣」與「了解整份試卷命題的優劣」的功能，由這樣一個改變，代表著參與的職前教師可以更有效的運用評量策略來改善教學、提升學習者學習效益與命題品質。
- (三)WATA 系統於師資培育過程中，會要求職前教師建立專屬的個人題庫並與他人分享，這樣的設計不但可訓練他們主動建立一套完整的個人題庫以供日後線上命題使用外，於觀摩他人題庫時亦能激發其命題的創意，讓整個職前教師間的創意在這個互為流通的題庫園地裡汨汨而生。
- (四)WATA 評量系統中的「依據雙向細目分析表命題」與「教材概念編號設計」設計，應用於參與調查之職前教師的評量訓練上，無論對男性或女性職前教師，均不但能促其精熟教材內容，更能使其精確地掌握命題方向進而提升其命題能力；此外，這兩項設計於試後分析時，有助於養成職前教師對學習者迷思概念診斷與分析之習慣，而此正向之態度及習慣更可作為補救教學與改進教學的指南。
- (五)WATA 評量系統中的「試後分析工具」，讓參與的男性或是女性職前教師，不但對評量理論與實際應用有更清楚的了解外，亦使其更願意於考試完後進行試後分析，以掌握自己的命題品質與學生的學習狀況。而系統中針對試題「鑑別度」、「難易度」與「誘答力指數」等分析，讓職前教師透過評量理論的定義更精準地掌握自己命題的良莠以為日後命題的參考，並了解到學習者的迷失概念所在，以運用於輔助日後教學。
- (六)大多數的職前教師表示，在實習時有意願應用 WATA 系統進行評量。另外針對於在國中高中教學環境使用 WATA 系統，預期會遭遇到的困難進行分析發現，會遭遇到的問題包括：在硬體方面，受限於電腦硬體資源的與網路的流暢度，在軟體方面，為週遭同儕教師的眼光與學生操作的熟練度，而其中以硬體方面的影響最為嚴重。另外也發現，對不願意使用者而言，其預期遭遇的問題主要在於「電腦硬體資源」與「學生操作的熟練度」，而願意使用 WATA 系統但「不會」排除萬難使用者，其預期遭遇的問題除了「電腦硬體資源」與「學生操作的熟練度」外，也很在乎「學校教師的異樣眼光」，由此可知，學校同儕或是資深教師會是影響本來就有意願在實習期間使用 WATA 系統的職前教師是否堅持在實習期間採用 WATA 的重要因素。

二、建議

- (一)小 WATA 形成性評量「不提供答案策略」的功能不被參與的職前教師認同可以引發其釐清概念的動機，反而認為，如果直接給予答案將更能夠讓參與者有機會去主動進行學習與思考找出答案，因此在未來的設計中，應該把「提供答題提示」與「不

提供答案策略」搭配在一起，才可以讓「不提供答案策略」這個功能發揮更大的輔助學習效果。

- (二) 由於傳統的評量理論對於「評量實施的流程」有嚴謹的要求，在這些要求中，尤其以「依據雙向細目分析表命題」與「試後分析與應用」這兩部分的難度較高且較為繁複，這也是造成一般的老師即使知道必須要做，但仍就是望而怯步的主因，利用資訊技術來節省進行這些步驟所需要耗費的人力物力，將可以使更多教師願意根據標準的評量流程來執行評量，而使得傳統評量能夠發揮最好的效果，並且成功的用來改善教學與學生的學習效益。除此之外，將開發出來的評量工具應用於師資培育，也可以讓接受師資訓練的職前教師對於評量能夠有更為具體的概念，而對於評量的學習不再只是紙上談兵與築空中樓閣。
- (三) 未來在「資訊技術融入師資培育」這個議題上，應該採取更為主動與開放的態度，研究者應該針對各項教師 PCK 能力訓練的需要，開發各種「線上模擬教學」的網路工具(Web Tools)，讓職前或是實習教師親身使用，並在網路上對同儕進行親身的模擬教學，而不是一味的把這些接受師資訓練的人視為一般的學生，讓他們在傳統或是網路教學(WBI)的環境中，只擔任知識接受者的角色，如此對於他們的 PCK 能力的訓練有更好的幫助。

本研究採用的樣本均為生物背景的師範體系學生，且因為資源限制無法獲得大量的樣本參與，因此本研究結果在應用與推估上需要謹慎小心。在未來，有興趣於這個議題的研究者將可以在現有的基礎上更進一步針對於受試者學科知識、學科概念、評量知識與評量概念的改變進行更細膩的質性與量化的研究。

誌謝

感謝國科會科教處專題研究計畫 NSC90-2511-S-018-015 與 NSC91-2511-S-018-024 經費補助，得以完成本研究結果報告。另外，也感謝僑光商業技術學院金融與風險管理學系 楊志強博士，對於本文的統計部分的指導。

參考文獻

1. Bonham, S.W., Beichner, R.J., Titus, A. & Martin, L. "Education research using Web-based assessment systems," *Journal of Research on Computing in Education* (33), 2000, pp:28-45.
2. Magnusson, S., Krajcik, J., & Borko, H. "Nature, sources, and development of pedagogical content knowledge for science teaching," In J. Gess-Newsome & N. Lederman (Eds.) *Examining pedagogical content knowledge: The construct and its implications for science education*, Dordrecht: Kluwer Academic Publishers, 1999, pp:125-126.
3. Gronlund, N.E. & Linn, R.L. *Measurement and Evaluation in Teaching*. (6th ed.), Macmillan, New York, 1990.

4. 林清山, 2000,『國小數學多元思考教材與整合型評量工具模組化之研究—子計畫二：整合型評量工具模組化』, 行政院國家科學委員會研究報告, NSC 89-2511-S-003-070。
5. Mason, B.J., Patry, M., & Bernstein, D.J. "An Examination of the Equivalence between Non-adaptive Computer-Based Test and Traditional Testing," *Journal of Educational Computing Research*, (24:1), 2001, pp:29-39.
6. Oblinger, D. "Teaching and Learning with Computers, An LAT Technical Primer," North Carolina University, Chapel Hill, NC: Institute for Academic Technology, 1992.
7. Dennis, V.E. & Gruner, D. "Computer Managed Instruction at Arthur Anderson & Company: Astatus Report," *Educational Technology* (March), 1996, pp:7-16.
8. Green, B.F. (1991). "Computer-Based Adaptive Testing in 1991," *Psychology & Marketing* (8:4), 1991, 243-257.
9. Hambleton, R.K. & Swaminathan, H. "Item Response Theory," Kluwer-Nijhoff Publishing, Boston, Massachusetts, 1985.
10. Bugbee, A.C. "The Equivalence of Paper-and-Pencil and Computer-Based Testing," *Journal of Reseach on Computing in Education* (28:3), 1996, pp.282-299.
11. Alessi, S.M. & Trollip, S.R. "Computer-based Instruction: Methods and Development," Englewood Cliffs, NJ: Prentice-Hall, 1991.
12. Kumar, D.D., Helgeson, S.L. & White, A.L. "Computer Technology-Cognitive Psychology Interface and Science Performance Assessment," *Educational Technology Research and Development* (42:4), 1994, pp:6-16.
13. Wang, Tzu-Hua, Huang, Shih-Chieh, Wang, Wei-Lung & Kuo, Li-Yu. "Development of a Web-based Assessment System-A Case Study," *Proceeding of the 6th Global Chinese Conference on Computers in Education/National Education Informatization Forum(GCCCE2002)*, 2002, pp385-388.
14. 王子華、黃世傑, 2001,『網際網路評量與試後分析系統(WATA)之探討-以生物科實習教師學科評測為例』, 九十學年度師範學院教育學術論文發表會, 國立台中師範學院主辦。
15. 王子華、王瑋龍、黃世傑, 2002,『應用網際網路評量與試後分析系統(WATA)於師資培育以生物科教材教法課程為例』, 九十一學年度師範學院教育學術論文發表會, 國立嘉義大學主辦。
16. 王子華、王瑋龍、王國華、黃世傑, 2002,『應用網際網路評量與試後分析系統(WATA)於師資培育：提升生物科職前教師之評量能力為例』, 收錄於 TANet2002 台灣區網際網路研討會論文集編 II, 397-402, 新竹市：國立清華大學。
17. 王子華、王瑋龍、王國華、黃世傑, 2001,『網際網路教學環境後設認知策略設計對於大學學生學習之影響』, 收錄於 TANet2001 台灣區網際網路研討會暨網路學習與繼續專業教育國際會議 (ELCPE, Conference on E-Learning and Continuing Professional Education) 論文集編 I, 53-58, 嘉義市：國立中正大學。
18. Wang, Tzu-Hua, Wang, Kuo-Hua, Wang, Wei-Lung & Huang, Shih-Chieh "Design and

- Evaluation of a Web-based Assessment and Test Analysis System (WATA),” Proceeding of the 6th Global Chinese Conference on Computers in Education/National Education Informatization Forum(GCCCE2002), 2002, pp.441-447.
19. 王子華、黃世傑，2002，『網際網路評量與試後分析系統(WATA)之設計』，資訊與教育雜誌，2001年資訊教育特刊：276-286頁。
 20. 王子華、王瑋龍、王國華、黃世傑，2002，『進階型多功能網路評量與試後分析系統(WATA)的發展與設計』，視聽教育，第43卷.第4期：21-45頁。
 21. 王子華、楊凱悌、王瑋龍、黃世傑，2002，『運用網路評量與試後分析系統(WATA)之形成性評量模組對學習者自我學習之成效分析』，收錄於TANet2002台灣區網際網路研討會論文集編II，415-420，新竹市：國立清華大學。
 22. 王子華、王瑋龍、郭禮瑜、黃世傑，2001，『多功能網路評量與試後分析系統(WATA)的發展與設計』，收錄於中華民國第十七屆科學教育學術研討會論文集編，57，高雄市：國立高雄師範大學，Available: <http://www.nknu.edu.tw/~gise/17years/B33.doc>。
 23. 王子華，2002，網際網路教學環境之後設認知策略設計對於大學學生學習效益之影響，國立彰化師範大學生物學研究所碩士論文。
 24. Wang, Tzu-Hua, Wang, Wei-Lung & Huang, Shih-Chieh “Utilizing FFS (Frontpage Feedback System) Sign-And-Feedback Teaching Module as Meta-cognition Strategy for Pre-service Biology Teachers Training: A Case Study of PCK Training Program,” Proceeding of the 2002 International Conference on Chinese Language Computing (ICCLC2002), 2002, pp105-111.
 25. 張基成、童宜慧，2001，『網路化學習歷程於師資培育課程之實施經驗及探討』，資訊與教育，第86期：88-101頁。
 26. 詹雅婷、張基成，2001，『網路化專題導向學習於師資培育課程之應用』，視聽教育，第42卷.第6期：26-39頁。
 27. Ahern, T. C., Jamison, M. & Olivarez, A. “Designing the Next Generation of Tools Using an Open Systems Approach: A Usability Study,” ERIC Document Reproduction Service, No. ED444585, 2000.
 28. 郭生玉，2000，心理與教育測驗，台北：精華書局。
 29. 林清山，1999，心理與教育統計學，台北：東華書局。
 30. Buchanan, T. “The efficacy of a World-Wide Web mediated formative assessment,” Journal of Computer Assisted Learning, 16, 2000, pp.193-200.
 31. Buchanan, T. “Using the World Wide Web for formative assessment,” Journal of Educational Technology Systems, 27, 1998, pp.71-79.
 32. 王子華、王國華、王瑋龍、黃世傑，2002，『網路形成性評量策略對於網路教學效益之影響--以 WATA 網際網路評量與試後分析系統形成性評量模組為例』，中華民國第十八屆科學教育學術研討會，彰化市：國立彰化師範大學。
 33. Falchikov, N., “Peer feedback marking-Developing peer assessment,” Innovation and Training International, 32, 1995, pp.175-187.

34. Freeman, M., "Peer assessment by groups of group work," *Assessment and Evaluation in Higher Education*, 20, 1995, pp.289-301.
35. 林珊如、劉旨峰、袁賢銘，2001，『以大學生學習動機與學習策略預測網路同儕互評之學習成就』，*教學科技與媒體*，第 57 卷：2-11 頁。
36. Zeichner, K., Tabachnick, B., & Densmore, S. "Individual, institutional and cultural influences on the development of teachers' craft knowledge," In J. Calderhead (Ed.), *Exploring teachers' thinking*. London: Cassell, 1987.