

2004天津

國際教育考試論壇

側記

編者按：2004年8月14至16三日，天津市教育招生考試院在天津舉辦「國際教育考試論壇」，邀請各地專家學者100多人與會，本中心由夏蕙蘭和管美蓉兩位研究員出席會議，以下是該會議報告側記及與會心得分享。

■研究發展處研究員／
管美蓉

由於天津市教育招生考試院(以下簡稱「考試院」)喬麗娟院長的邀請，夏蕙蘭和我在8月13日上午9:00由中正機場起程前往北京，途經香港轉機，下午4:00抵達北京，再由「考試院」派人接至天津。抵達之後，歡迎晚宴已開始，我們錯過了舞龍舞獅的熱鬧開場，十分可惜。整個晚宴就在東道主盛情款待的氣氛下展開，酒過三巡，來自四面八方的與會者，一掃先前的陌生，紛紛把酒言歡，不論是工作經驗或風俗民情，大家就這麼熱切地交流起來，為往後三天研討會奠下充分對話的基礎。

本次大會主題有三，一是各國大學入學考試制度改革的現況與未來發展；二是教育考試技術及命題管理等相關議題；三是回顧與展望。就第一個主題而言，會中有來自日

本、韓國、台灣、澳洲及美國等地的專家學者，對該國大學入學考試的現況發展及未來改革動向進行報告，對於來自中國各地、從事招生工作的學者專家而言，這些他山之石十分珍貴。第二個主題是針對各種測驗統計技術在各類考試中的應用，包括上海、天津教育招生考試院、美國ETS及美國Harcourt公司等測驗機構，均分別提出測驗評量的最新技術，對從事測驗評量專業的考試專責機構及相關研究人員而言，十分具有參考價值。第三個主題則是總結了美國及中國大學入學考試的經驗，由兩位學養深湛的學者進行講演，也為三天的會議畫下完美的句點。

以下各場報告重點，主要是節錄自「國際教育考試論壇簡報」及各場演講的摘要，詳細內容應會刊登於近期《考試研究》(由

「考試院」和天津人民出版社編輯出版)，有興趣者可自行參閱。

為大會揭開序幕的是中國教育部高校學生司司長林蕙菁博士，她以「構建有中國特色的考試制度」為題進行演講，主要提出三個問題：(一)中國的高考面臨著新的形勢和新的挑戰。林博士認為，近年來高考(即普通高等學校入學考試)改革的內、外部環境都發生著深刻的變化，因此必須體認到改革的勢在必行。(二)下一步高考改革的目標和方向：建立以國家統一考試錄取為主，與多元化考試、多樣化選才相結合，高校更加自主、政府加強宏觀管理、社會有效監督的具有中國特色的現代化考試制度。(三)未來高考改革的幾個方面。一是進一步加大全國統一考試分省命題的力度，二是繼續探索並不斷完善多元化考試、多樣化選拔的考試評價制度，三是積極探索多層次考試方式，四是繼續推進高考內容改革，五是大力推進考試形式現代化的進程。

日本大學入學考試中心大津起夫博士以「NCUEE(日本大學入學考試中心)組織的統一考試在日本大學入學中的作用」為題，他首先介紹了日本的教育體系，現行的6-3-3-4體系是在1947年建立。第二部分是回顧人口變動和大學註冊人數變化的趨勢。伴隨著日本經濟高速發展，在1960年代申請進入大學的人數快速增長，此導致了1970至80年代大學入學競爭的不斷加劇。為了適應這種狀況，統一考試乃應運而生。第三部分是介紹日本大學入學招生的程序，對申請程序、NCUEE考試的設計及形式及其使用進行說明。第四部分是對NCUEE進行歷史回顧，並描繪日本大學招生的前景。

韓國教育課程評價院金京堦博士以「韓國大學入學考試制度」為題，他指出：韓國大學入學考試制度的一大特點是政府提供基本的指導性原則，而學校能夠制定和使用自己的入學條件。然儘管有著招生制度的一般原則，入學考試的分數卻一直是大學招生的主要條件，因此，高中教育指向在入學考試中取得高分。為解決高中教育的問題，韓國教育界建議在下列方面進行改革：(一)高中教育從大學入學考試的羈絆中解脫出來，回歸教育的本質。(二)大學在選拔學生時應該更具專業和自主。(三)為提高國家競爭，應尊重大學教育的多樣性和質量。(四)與大學入學相關的數據資料應相互補充，而非專取其中之一使用。

本中心夏蕙蘭博士以「台灣大學入學考試改革的看法和今後發展趨勢」為題，主要內容分為：(一)背景；(二)對台灣大學入學考試改革的看法；(三)台灣大學入學考試發展的重要歷程；(四)台灣大學入學考試命題研究的未來發展。而筆者則是以「大學入學考試對高中教育的影響——以1991年和2004年台灣地區問卷調查為例」，以13年間兩次相同內容的問卷調查結果，來說明十年教改下的大學多元入學方案到底改善了什麼？

美國大學理事會高級研究員王湘波博士以「『新』SAT1介紹：其在美國大學入考試中的新特點、研究和影響」為題，主要內容是：(一)介紹大學理事會及其承辦的大學招生考試項目；(二)介紹新SAT：從其基本原理、推理測驗的目的、批判性閱讀的目的、數學問題解決的目的及寫作測驗等項目來說明；(三)介紹新SAT的作文評分步驟；(四)說明新SAT的研究成果；(五)說明新SAT在美國大學招生中的意義；(六)對中國大學入學考

試提出個人建議。

澳大利亞ETC中心主任Peter Knapp博士以「大學入學測驗未來的發展方向」為題，他首先介紹了澳大利亞的教育概況，並對新南威爾斯州高中證書(HSC)評價、成績報告、級別、原始分數到UAI的轉換等問題逐一進行說明，最後就一般成就測驗(General Achievement Test)進行了分析，提出未來的GAT類型的測驗模型。

澳大利亞維多利亞省課程與評價局考試部主任Bill Perrin以「高利害大學入學考試」為題進行演講。內容主要是針對GAT——一般成就測驗加以介紹，包括此測驗的使用、構建、寫作測驗、以英語為第二語言的學生和以英語為母語學生得分的比率、測驗統計在測驗建構中的應用、分數分布以及為何還需要其他的考試等內容進行說明。

就測驗評量與命題管理的主題而言，以下演講大抵圍繞著此一議題：

上海教育考試院雷新勇博士以「多元概化理論在大規模教育考試中的應用」為題進行演講。他介紹了概化理論的基本概念，並將之應用於五個方面：(一)研究考試的測量效果和相對測量誤差；(二)研究影響試卷測量效果的主要因素；(三)研究改變語文評分方法的效果；(四)研究如何設計更好的試卷；(五)確定最佳試卷以及各變量部分的最佳題量。最後並對應用多元概化理論研究大規模教育考試的主要優勢和侷限提出看法。

「考試院」趙彤璐女士以「考試數據統計分析系統和英語口語測試系統」為題進行演講。首先介紹了考試數據統計分析系統開發的理論與現實背景、系統程序、系統應用說明和使用示例，然後對英語口語測試系統

的開發背景、考試系統程序、閱卷系統程序以及系統應用情況進行了詳細的講解。

美國Harcourt公司副總裁John. F. Olson博士做了題為「Harcourt公司在大規模測驗管理中的命題和質量保證」的演講。他首先介紹Harcourt公司的概況，然後從命題的幾個關鍵步驟、Harcourt公司的命題方法、大規模考試項目的管理、心理計量學方面的研究與應用、以計算機測驗和在?評價為基礎的科學技術在評價中的應用、命題質量評價政策等方面進行說明。

Harcourt研究員王蜀東博士、焦虹博士以「美國中小學成就和診斷考試的發展和應用」為題進行演講。演講從常模參照考試的一般特徵分析入手，對美國中小學成就和診斷考試中的SAT10(斯坦福成就測驗)常模參照考試的發展流程進行講解，並結合常模參照考試和標準參照考試的特徵，對美國中小學成就和診斷考試的設計、發展研究加以說明。

Harcourt研究員易青博士做了題為「機遇與挑戰：在線考試在中小學教育中的應用」的演講。她從在線測驗的界定、種類、優點、州立考試中電腦測驗的應用情況，電腦測驗的效度問題、在線測驗需要考慮的問題等幾個方面做了說明，並從技術策略、技術投資、技術綜合實力和互動式服務平台等方面介紹Harcourt公司的發展狀況。

Harcourt高級研究員劉志明博士做了題為「美國中小學成就測驗中的原始分數向縱向量表報告分數的轉換」的演講。他從標準化成就考試、原始分數、量表報告分數、等值(化)、量度化等背景材料的分析入手，詳細介紹了縱向量表的定義、與等值(化)的異

同、運用以及對教育改革的啟示，並從基本假設、數據收集設計、共同試題設計、共同組設計、結合方法論層面說明了縱向量表化的方法，最後介紹標準設定的定義、設定、制定以及年級組標準設定的方法。

ETS高級研究員錢家和博士做了為題「電子閱卷在美國大規模英文考試中的運用、研究和發展前景」的演講。主要內容是：自動作文評分系統的定義、目標、階段、原理、基於計算機(電腦)的文本分析、評分標準的機器呈現、作文特徵的種類，作文特徵的因素分析和維度分析，根據直接估計和若干導出估計所得的真分數最佳線性估計值。同時他對中文作文評分系統發表了個人見解。

ETS高級研究員張金明博士做了題為「美國教育進程國家評估(NAEP)總體設計、抽樣和數據分析」的演講。他介紹了NEAP的概念、管理內涵、各責任部門的職責、組成部分，對為什麼需要NEAP進行了分析，並對NEAP的設計要求和目標、設計特點、統計分析進行了詳細的解說。

ETS高級研究員顏端麗博士以「認知診斷的幾種模型及其應用」為題，介紹了幾種心理測量模型，包括貝葉斯測量模型、能力特徵評分的共同方法、Arpeggio和StatShop模型，並提出應用於下列三種測驗的情形：PSAT/NMSQT、新TOEFL、高級教育ICT評價。

第三部分是以回顧與展望為主題，以下二篇演講：

德克薩斯大學奧斯汀分校張華華博士以「回顧與展望：計算機考試(以下簡稱「機考」)在美國研究和應用的三十年」為題進行

演講。他首先是回顧了「機考」的發展歷史；其次說明「機考」在大規模應用中出現的問題、發生問題的根源、成績可比性的問題以及如何解決大規模偷題的問題，並對「機考」及電腦適性測驗提出了批判與建議。

最後的壓軸演說，是由前教育部考試中心主任楊學為以「關於高考考能力的問題」進行演講，主要從三個方向來探討：一是科舉失敗的教訓；二是何謂能力；三是高考的侷限性。最後並提出對高考改革的建議：(一)本科生及專科生分開考試；(二)筆試達一定分數的考生加試口試、實作等；(三)考試機構的體質應產業化，與福利結合，大學應自主錄取。楊主任並指出，這次來自海外的多位學者專家，應該都是1977年恢復高考之後進入大學就讀的一代，因此可以稱得上是高考受益者。這批出生中國、在海外開疆闢土的測驗專家，也確實為中國教育考試的這片土壤提供許多養分，恢復高考的重要性及歷史意義由此亦可彰顯。

三天會議的內容十分充實，除了白天的知性交流之外，在「考試院」精心規畫下，每晚會議結束之後，與會人員還有機會認識感性的天津。第一天是夜訪天塔、360°地欣賞天津夜景，第二天到天津戲劇博物館欣賞表演的經驗則令人印象深刻。看戲地點是以前的廣東會館，建於1903年，現改建為茶館。四人一桌，熱茶及零食均置於桌上，斟茶的夥計穿梭其間，隨時為客人補充熱茶。表演舞台位於正前方，大夥一邊喝茶，一邊聊天，一邊欣賞演出，比起在音樂廳的正襟危坐，茶館瀟灑著輕鬆自在的氣氛，看到精彩處可以大聲叫好，文化就這麼在身旁流動著。當日演出的京劇及民族舞蹈等精彩節

目，均是由天津藝術學校的師生擔綱，此舉一方面讓天津的民俗文化活動得以延續，一方面也讓藝校學生有磨練的機會，應是不錯的合作方式。

身在天津，讓人感受到天津的特殊韻味，回顧天津歷史，它是近代中國最早開港的北方大港之一，因此早在1860年代初期就與西方遭遇，隨著天朝勢力的日趨衰微，天津受到列強駐兵、設使館、立租界等外力侵擾，在新舊勢力衝擊之下，它也逐漸發展出文化的特殊性。相對於北京集政治、軍事、經濟於一身，冠蓋雲集，動見觀瞻，天津就像是北京的後花園，沒有北京那樣濃烈的政治性格，反倒是透露著一種閒適與庶民氣息的津派文化。天津也是許多政治人物退隱之處，他們習於在宦海浮沈，下台之後，往往在此等待再度出山的機會，在此可以見到北洋時期的大總統徐世昌、奉系張學良等人的故居。

天津緊臨著北京，車程只有二小時，然北京像個大磁鐵似的，吸走了許多的資源、人才及世人的目光，因此天津在現代中國的發展史上，並不像廣州、深圳等地的生氣勃勃，甚至還因而有些停滯不前。不過，換個角度來看，這未始不是天津幸運之處，在其他大城市的許多歷史古蹟紛紛受到開發所需而被迫拆遷，天津卻因為進展的腳步較為緩慢，反而保留了大片而完整的各國使館租界區，使得天津在邁向現代化的同時，仍能保有自身的歷史景深，形成了一種獨特的城市性格與風采。在微涼的夏夜，乘著馬車，漫步於舊租界以及其所散發的歷史氛圍之中，有著一種既遠又近、既真實又虛幻的迷離之感。

這次大會在許多細節的安排上十分用心，令人激賞，也值得我們學習。例如，充分運用電腦數位影像處理的功能，在與會代表報到時，即照相立即將影像輸出到示別證上，同時，在會前已透過電子郵件的聯繫，將個人通訊資料存檔，因此，在會議結束以前，與會者均獲得一本載有全體人員名單、彩色照片及相關聯絡資訊的通訊錄，使會後的交流更加便利。此其一也。大會對於會議進行的情形是以快報方式加以報導，出刊的時間是在隔天下午，快報以中、英文摘記了前一天各場報告的內容，並輔以演講者的照片，快速提供會議的動態報導。此其二也。此次大會是採「論壇」方式進行，由每位演講人簡報二十分鐘，並於會前印發綱報告大綱，演講人均輔以power-point進行講演，由於演說內容精彩豐富，大會擬將各場報告內容燒錄成光碟，提供與會人員參考。此其三也。此外，大會致贈演講人紀念圓盤，並書上演講人姓名，相當具有紀念價值，也讓演講人留下深刻印象。此其四也。

三天的會議共有十九場演講，台上是傾囊相授，台下是頻繁交流，也讓我們此行滿載而歸。中國有七百萬考生參加，高考的重要性不言而喻，這也使得考試研究的課題益發迫切。不論是2003年10月在武漢召開的「兩岸四地考試與社會發展學術研討會」，抑或是此次的「國際教育考試論壇」，都顯示出「考試學」逐漸成為一個嶄新的研究領域。這是一門既需豐富實務經驗，又需進行深入理論探究的課題，它既能與現代社會接軌，又能對過去歷史探微，是一個跨學門的研究範疇，若能有不同專業領域的學者專家一起投入，必能發掘出更多有意義的課題。