

國立政治大學「教育與心理研究」
2008年3月，31卷1期，頁65-96

結構化教師甄試口試之初步調查

胡悅倫*

摘要

本研究旨在瞭解目前臺灣教師甄選口試結構化的程度，以期增進臺灣目前教師甄選之效度。

本研究受試者為2004年中等教師聯合甄選之口試委員，以臺北縣為主，有效樣本215人，其中男性131人，女性84人。研究工具為自編之「中等學校教師甄選口試結構化調查問卷」。在信度方面，Person相關係數約在0.5左右，達顯著水準。在效度方面，經過焦點團體得專家效度。

結構化面談分為15個向度，如：由「工作分析」建立口試題目、定錨量表的使用、面談訓練等。研究結果發現：目前教師甄試（以臺北縣為例）仍處於「低結構化」的狀態。而最困擾口試委員的問題是「如何擬定良好問題」及「如何評分」。目前，低結構化口試狀態會嚴重的影響臺灣教師甄試口試的信度與效度。

關鍵詞：教師甄試、口試、結構化

* 胡悅倫：政治大學教育學系教授

誌謝：本文為國科會補助之研究計畫成果之一（NSC93-2413-H-004-006），謹謝兩位匿名審查委員的修改意見與支持。整個研究過程中助理李仁豪、袁明玉、陳世芬（依姓名筆劃順序）的協助。本研究已被2006年夏威夷國際教育年會所接受。

電子郵件：joyhu@nccu.edu.tw

收件日期：2006.05.11；修改日期：2007.04.18；接受日期：2007.06.26

The Study of Structured Teacher Interview

Yueh-Luen Hu*

Abstract

The objective of this study is to understand the structured level of Taiwan's teacher selection system, thus promoting the validity of Taiwan's teacher selection system.

The subjects were interviewees of the Taipei County District 2004 joint high school teacher selection. A total of 215 effectiveness questionnaires are collected, consists of 131 males and 84 females. Survey instrument is the self-designed "High School Teacher Selection Interview Survey Questionnaire". The test-retest reliability coefficient of the questionnaire is about .50 with significance level of .05. The validity of the questionnaire was based on the expert validity, which was satisfying.

The general definition of a structured interview can be organized into fifteen dimensions. For example, to establish interview questions based on job analysis, to use of anchored rating scales and interview training. Results showed that the teacher selection interviews, in Taipei County, are in a low structured state.

The most overwhelming confusion among the interviewers is the lack of ability to propose the appropriate questions and to evaluate responses. Taiwan is still quite unfamiliar to the issue of structured interviews. Such a problem will damage the reliability and validity of Taiwan's teacher selection.

Keywords: teacher selection, interview, structured

* Yueh-Luen Hu: Professor, Department of Education, National Chengchi University

E-mail: joyhu@nccu.edu.tw

Manuscript received: 2006.05.11; Revised: 2007.04.18; Accepted: 2007.06.26

壹、緒論

「眼見為憑」是多數主管在人事甄選上的中心信念，因此，主管們仍堅持使用口試／面談做為錄用人員的重要依據。人事心理學多年來的研究成果亦認同結構化面談的重要性，究其原因為多數的研究者（Arvey & Campion, 1982; Harris, 1989; Mayfield, 1964; Schmitt, 1976; Ulrich & Trumbo, 1965; Wagner, 1949; Wright, 1969）皆認為自測驗學的觀點視之，結構化面談比非結構化面談具有更多的優勢。而後設分析的研究更證實，非結構面談的效度大約在0.14到0.30，但是，一旦面談結構化後，其效度則可提升至0.35與0.62之間（Campion, Palmer, & Campion, 1997; Huffcutt & Arthur, 1994; Hunter & Hunter, 1984; Marchese & Muchinsky, 1993; McDaniel, Whetzel, Schmidt, & Maurer, 1994; Wiesner & Cronshaw, 1988; Wright, Lichtenfels, & Pursell, 1989）。

結構化面談，英文中有人稱為「structured interview」，但也有人用標準化（standardized）、指導性（guided）、系統性（systematic）或是組型式的（patterned）等字詞代替結構化（structured）。但一般而言，在討論口試信度與效度的文章中，最常用的仍是結構化面談一詞。

就狹義而言，結構性面談研究都將焦點放在面談的問題及評分標準上。Huffcutt與Arthur（1994）將面談結構性依「面談問題」及「評分標準」區分為四個水準。面談的結構性有層次上的分別，其嚴謹程度由水準一至水準四逐漸提升，故水準四的結構性要求是最高的。而愈嚴謹的面談，乃是在「面談問題」與「評分標準」的處理上愈有一定的規格。後續的研究者，尤其有許多個重要的後設分析均採用Huffcutt與Arthur的分類標準，並獲得寶貴的結論（Campion et al., 1997; Huffcutt & Woehr, 1999）。

就廣義而言，Campion、Pursell與Brown（1988）認為結構性的面談，除了一套詳細的評分量表外，其他條件亦有助於提升面談效度。如評分員的先期訓練、評分員在面談期間作筆記、以數人組成的評分小組來進行面談，或者以行為評量為主的評定量表等。上述這些輔助條件若能一致性地被使用，則稱之為廣義的結構性面談。而所謂的「結構性」可視為任何使面談過程呈現公平一致的控制方式。這也說明了為什麼口試一旦結構化之後，的確有助於提升口試的效度。

結構性面談之所以獲得重視，其原因是結構性面談的效度是非結構性面談的二倍（Campion et al., 1997; Wiesner & Cronshaw, 1988）。結構性面

談的效度之所以能夠優於非結構性面談，其原因若由實驗設計或測驗學理論來說，實在不難理解。因為當控制無關干擾因素進入面談，使所有應試者接受相同一致的待遇，且確保每位面談委員皆能接收到正確的訊息，將使得口試的各個向度都得到一定的控制。是故，在信度及效度方面，結構性的面談當然優於非結構性的面談。

目前，國內師資培育多元化之後，造成待業的準教師大量增加，再加上「少子化」的社會趨勢，以及教師超額壓力的影響，各縣市教師甄試釋出的缺額愈來愈少，粥多僧少的情況之下，使得教職難求，教師甄試的競爭亦愈來愈激烈。以2005年臺北市的教師甄試為例，報名國中教師甄試的人數達4,000人，錄取率約2%；而國小教師甄報名人數高達9,595人，錄取率只有1%。另外，臺中市國立高中教師甄試也較往年競爭激烈，錄取率約1%（林宜靜，2005；林麗雪，2005）。而應徵者中有八成以上具有碩士學歷，半數以上在私立學校任教超過五年，競爭之激烈，歷年少見（喻文玟，2005）。由上可知，各縣市教師甄試釋出的缺額較往年減少許多，但相對地報名人數卻盛況空前，讓許多報考者大嘆老師難為，要成為正式教師實非易事。

在教師甄試中，有關應試者的反應，胡悅倫（2006）針對臺北縣1,995

名應試者的調查研究指出：近八成的應試者認為須「訂定標準透明化的口試流程」，方能有效地杜絕教師甄試口試的不公平。此外，近五成的應試者認為宣導公平公正的觀念、對面談委員予以標準化訓練，以及強制增聘外校面談委員，皆能有效杜絕教師甄試口試的不公平。故針對教師甄試的口試進行研究，的確有其迫切性。

以上研究在某種程度已經碰觸教師甄試問題核心。但是，在口試中，到底什麼需要標準化？是口試問題內容抑或評分標準化？而標準化程度又是什麼？像紙筆測驗一樣嗎？理當不是，否則就不像口試了！但又是什麼？故口試所引發的種種問題，著實需要完整的理論架構支持。

所謂結構化口試，必須在內容、程序及評分上有一套標準化的藍圖。在內容上，口試的問題必須是事先設定的，而且題目必須是具有鑑別力，可使應試者能力展現不同的能力配分。在程序上，涉及到由口試前，口試進行中，到口試後種種相關的流程細節。在評分上，面談委員則需根據應試者對問題回答的結果與預先設定的評分標準做比較後，給予評分。

口試的結構性議題並不是一個全然或無的問題。基本上，由非結構性到結構性是一連續性的向度。換言之，口試的結構性乃是一種程度上的差別。最

狹義的觀點是僅侷限在建立口試一套評分工具與標準。而廣義的觀點則是在內容、程序及評分上，均能有良好的控制。當然結構化程度愈高，口試效度愈佳；反之，則效度愈差。但是，結構化程度愈高者，在時間、人力，以及物力的花費成本相對地愈高，並不是每種口試都能負擔得起。因此，如何在這其中取得最佳的平衡，乃是教師甄試口試所最須面對的問題。

瞭解口試相關議題的確有其必要性，口試是那麼關鍵地影響許多人的未來。而在眾多的口試中，吾人覺得以教師甄試口試最為重要，因為它涉及不僅是教師個人生涯的發展，更是牽涉到國內教師素質良莠，一位教師的錄用，就所謂的「請神容易送神難」，豈可等閒視之。教師的品質，牽動著千萬學子的身心與未來的發展，故教師甄試口試之重要性不言而喻。據此，本研究目的歸納如下：

一、以國外結構化面談的理論為基礎，利用自編問卷分析現階段教師甄試口試的結構化程度。

二、依據研究結果，針對教師甄試口試提出未來的改進方向，以提升教師甄試口試之品質。

貳、文獻探討

過去近九十年來，人事心理學在面談甄試的研究成果豐富，近二十年來

更出現許多的後設分析，然而這些研究的焦點大都集中在少數幾個向度上，例如：面談內容的問題類型（如Janz, 1982; McDaniel et al., 1994; Motowidlo et al., 1992）、面談結構性高低程度（如Huffcutt & Woehr, 1999; McDanielet al., 1994; Wiesner & Cronshaw, 1988）、面談委員的組成形式（如Huffcutt & Woehr, 1999; McDaniel et al., 1994）、面談的效標（如Wright et al., 1989; McDaniel et al., 1994）、面談委員訓練（如Huffcutt & Woehr, 1999; Motowidlo et al., 1992）、面談筆記（如Burnett, Fan, Motowidlo, & Degroot, 1998; Huffcutt & Woehr, 1999）及面談委員的標準化（如Huffcutt & Woehr, 1999; Pulakos, Schmitt, Whitmey, & Smith, 1996）等幾個較大的向度上。但對於其他較細節的部分，並非所有研究都清楚說明他們如何控制其它可能會影響結構性的因素，例如：面談使用時間、面談問題題數、面談委員是否獨立評分等等，這些較少受重視的因素也會對結構性面談造成干擾，但卻未被有系統性地討論。

本研究發現Campion等人（1997）提出的15項影響結構性面談的向度整理的最詳細完善，除了上述較大向度外，亦針對細節提出討論，因此，本研究將參考其所提出的15項影響結構性面談的向度，並融合其他文獻及本研究的主

題，在此依序提出說明，但這不表示影響結構性面談的因素僅有這15項。15項影響結構性面談因素，茲說明如下：

一、由「工作分析」建立口試題目

結構性面談的口試題目不能憑空出現，必須奠基於重要事件的工作分析。所謂工作分析即是分析特定工作的重要細節，使能對工作的內容有一提綱挈領的瞭解。工作分析必須由該工作領域的專家學者來進行，通常使用問卷調查或專家座談方式達成，以確保所發展出的口試問題具有效度。

工作分析可以強化口試內容對工作關聯性，部分是因為它允許面談委員獲得應試者行為有關的工作樣本的資訊（Dipboye & Gaugler, 1993）。工作分析應盡量強化與工作有關之訊息數量，使重要訊息均能被納入考量，同時，藉由聚焦面談在工作有關的內容上，可減少無關訊息的污染。在面談之前，若無透過工作分析以提供面談委員一個共同的參考骨架，面談委員可能將面談奠基於個人獨特對工作需求的信念上（Dipboye, 1994），使整個口試內容流於面談委員各自發揮的現象，導致面談結構性降低，因而降低了面談的效度。

是故，經由適當完整的工作分析，方使面談內容更能符合實際工作的要求，也因而提升口試的內容效度或稱

專家效度。研究發現：擁有較高內容效度的面談問題，將會具有較高的效標關聯效度（Carrier, Dalessio, & Brown, 1990）。而結構式的面談必須與工作有關（job-related）的問題作為面談基礎，方能增加面談的效度（Wiesner & Cronshaw, 1988）。由此可知，工作分析可幫助面談委員擬出具有內容效度的題目，並藉此提升口試的品質。

二、每一位應試者均被問相同問題

從測驗學的角度來看，對每一位受測者使用相同的考題，是提供測驗結果具相互比較性的最基本要求。從面談的角度來看，使用相同的口試問題可以提高面談的結構性，使得面談委員的評分建立在相同問題的基礎上，如此評分結果的比較才有意義。

使用相同的問題可能是將面談從「討論」走向「科學性測量」最基本的通路（Campion et al., 1997）。如果能進一步以相同的問題順序詢問每一位應試者，將使面談的結構性有更進一步的提升，因為這樣做可以增加面談的一致性（Schriesheim, Solomon, & Kopelman, 1989）。

這一點看起來很理想，但實際上的操作，卻很困難。因為應試者眾，前面考完的考生出去以後，難道不會說出考什麼嗎？題目一旦洩漏，則效度盡

失。因此，實務上該如何處理，實有待考量。是以題庫或是其他方式取代之，則應再討論。

三、即興問題或後續追問的處理

在面談進行的過程中，面談委員有時會不按既定的問題發問，或者會針對應試者先前的回答，額外或臨時提問一些與工作內容無關的問題，目的可能是為了獲得更多有用的訊息，以利決策評分。但這種後續追問的動作必須謹慎小心，原則有二：(一)追問的問題必須扣緊相關工作主題；(二)必須給予所有應試者相同的機會獲得面談委員的追問 (Campion et al., 1997)。

對許多人來說，面談 (interview) 的基本特徵是兩個人之間的動態交互作用，然而這種方式可能會使訊息的蒐集導致偏誤 (Dipboye, 1994)。為了避免不必要的訊息進入面談中，並提升面談的結構性，因此，對於即興問題應該避免，而後續追問也應該扣緊主題，同時每一位應試者都能獲得相同的追問與澄清的機會。

四、問題形式與內容

經由工作分析後所發展出來的問題形式，是指問題是如何被問或是問題的內容。問題可能要求應試者做自我描述、對假設情境的反應、或者回答描述

過去行為、一般背景、特定知識或技巧、動機、或其他問題。由過去的研究 (如Janz, 1982; McDaniel et al., 1994; Motowidlo et al., 1992; Motowidlo, Dunnette, & Carter, 1990)，面談問題的形式大致可歸納為六種：情境式問題、過去行為問題、背景式問題、與工作知識有關的問題、問題是真實工作的模擬及心理特質問題等。

研究發現：以「情境式問題」及「過去行為問題」的面談效度最佳，而以「心理特質問題」的面談效度最低 (如 Janz, 1982; McDaniel et al., 1994)。情境問題的效度建立在目標或意圖與未來行為的關係上 (Locke & Latham, 1984)，而過去行為問題的效度通常是假定，過去行為是未來行為良好的預測因子，若能以具體的資料做為佐證更佳 (Mumford & Stokes, 1992)；心理特質問題十分模稜兩可，足以讓應試者表現出令人喜愛的態度或避免表現出缺點 (Campion et al., 1997)，至於，若欲使「意願」(willing) 問題具面談效度則需以應試者能提供證據為宜 (Wanous, 1980)。是故，未來口試問題應朝向「情境式問題」及「過去行為問題」發展，不過，基於教師工作所需的重要情意特質，因此，心理特質問題仍須加以考慮。

五、面談的時間與題數

一如測驗長度對測驗信度的影響，當測驗長度愈長時，信度愈好；而信度愈佳，效度也會隨之提高。同理，面談的時間與題數的增加時，面談的效度也會增加。面談長度是一個基本，但又常是被忽視的因素，在合理的時間下，「面談時間」與「面談題數」的增加導致結構性較佳，因為較長的面談可獲得有關應試者較多的資訊（Campion et al., 1997）。

過去的研究並未特別重視面談長度對面談結構性的影響，Campion等人（1997）所蒐集到的近200篇有關面談研究的文獻中，只有38個研究報告面談的時間，其範圍在3至120分鐘間，而平均面談時間是38.75分鐘；並且只有14個研究報告面談的問題題數，其範圍在4至34題間，而平均題數是16.50題。但沒有一篇研究對所選擇的面談長度做解釋。至於，到底要多少時間及多少題目，尚無定論，有待學者進一步研究。目前，國內教師甄試面談時間都不超過15分鐘，若加上試教時間，也不過30分鐘。在面談的時間與議題上，的確值得商議。

六、輔助訊息的控制

所謂的輔助訊息，是指除了正式的面談外，可用來幫助面談委員做決策的資訊，包括申請表、履歷、測驗分數、推薦信、早先的面談、成績單等

等。面談過程若能一致且客觀地使用輔助訊息，將增加面談的結構性，反之，若未有效控制輔助訊息將對面談結構造成重要威脅（Campion et al., 1997）。

就評分的一致性而言，應該讓所有面談委員都能獲得這些輔助訊息，才能避免評分不一致的情形發生；而就應試者的公平性而言，必須是在所有應試者都能提供這些輔助訊息的條件下，才能避免對某些應試者不公平。輔助訊息在一致且公平審查下，才能維持面談的結構性（Latham, Sarri, Pursell, & Campion, 1980; Mayfield, Brown, & Hamstra, 1980）。

七、應試者的發問時機

在面談結構性的考量下，應試者在面談期間的隨意發問可能會使面談結構性受到破壞，因此，事先規定應試者的發問時機是必須的，在面談期間與面談內容無關的發問都必須禁止，或留待面談結束後才統一給予回答（Campion et al., 1997; Campion et al., 1988; Tullar, 1989）。

禁止應試者的發問，可以減少面談委員與應試者間互動所產生的干擾，面談委員可能會因應試者的發問而產生主觀印象，進而影響評分的客觀性，或者應試者可能藉由與面談委員的對話中，漸漸迎合面談委員想要的答案（Beatty, 1986），導致面談效度的降

低。過去的研究對應試者發問時機並未特別重視，因此，有關的討論並不多，但基於結構性的考量，對應試者發問的時機應該謹慎處理。

八、立即且多次的評量

一如測驗學理論所說的，進行多次的評量其信度會較單一次評量來得高，評量的次數愈多，所涵蓋的評量範圍愈廣，可獲得的資訊愈多，則效度也會愈高；同理，面談委員對每一個問題的答案都能立即評分，這樣累積評分的次數就愈多，評分結果也就愈可靠，面談效度也會提升。同時，就結構性而言，對每一個應試者回答的答案進行評分，可以使答案與評價的連結更緊密，因而提升面談的結構性（Campion et al., 1997）。

因應多次評量所設計出的詳細評量表，提供了更多行為被評量的機會，同時，也可以使面談委員更仔細地辨認出所要評量的行為與無關行為，因而減少無關行為干擾相關行為的評價，使得面談的效度可以提升（Campion et al., 1997; Delery, Wright, McArthur, & Anderson, 1994; Latham & Saari, 1984）。

九、定錨量表的使用

定錨量表是指對評分量尺上的每一個量尺點所代表的程度或意義加以描

述，描述的程度可以是十分詳細地指出每個量尺點所應具備的能力或表現，也可以簡單地只點出如非常好、好、普通、不好、非常不好等等。發展定錨量表的目的是為了減少面談委員間評分的不一致性，使面談的結構性升高。常見的評分者效應，如寬容、嚴格的評分風格，都會使面談出現系統性的評分誤差，導致面談委員評分的信效度降低（Graves, 1993）。

實務的研究指出，並沒有強烈的證據可以支持愈詳細的定錨量表愈有助於面談效度的提升（Landy & Farr, 1980）。但從減少面談委員評分的模糊性來說，定錨量表的使用是一個有利的因素，並且被一些學者所重視（如 Green, Alter, & Carr, 1993; Nevo & Berman, 1994）。

十、筆記的運用

藉由作筆記可以減少記憶負擔，確認重要的相關資訊被記錄，並避免初始及新近效應（regency & primacy effects），因而增加面談的結構性（Campion et al., 1988; Schmitt & Ostroff, 1986）。

筆記的詳實度愈高，面談委員的評價將愈精確，因此，愈有助於效度的提升。同時，作筆記具有鼓勵面談委員去注意答案並組織思考的功能，因而增加面談委員評價的精確性（Campion et

al., 1997)。一致性地要求面談委員作筆記可以增加面談的結構性，有助面談效度的提升，但部分研究指出，對面談效度愈關心的面談委員，即使不主動做筆記，面談效度也不會較低（如Burnett et al., 1998）。

十一、多重面談委員

使用多位面談委員可以減少單一的面談委員對應試者的衝擊，因此，聚集較多面談委員可以避免單一主觀的判斷或評分偏好，並且消除隨機誤差，其效果就如同使用一個較長的測驗，據此所結合的分數較具信度（Dipboye, 1992; Hakel, 1982）。

在非結構式面談中，小組面談比個別面談更加有效（效度為0.37 vs.0.20），但在結構式面談中，兩者沒有顯著差異（效度為0.60vs.0.63）（Wiesner & Cronshaw, 1988）。另外的研究發現，在非結構式面談中，兩者沒有差異（效度為0.33vs.0.34），但在結構面談下，個別面談稍具優勢（效度為0.38vs.0.46）（McDaniel et al., 1994）。是故，未來的研究有必要進一步研究兩種組成方式的優劣。

十二、應試者均能接受相同面談委員的口試

使用相同的面談委員對面談的結構性是十分重要的，尤其是當其他影響

面談的因素未能有效控制其結構性時（Campion et al., 1997）。因為對不同的應試者使用不同的面談委員，增加了面談委員變異的干擾，不僅降低面談結構性，對效度也是一大影響。

幾個不同的研究都發現面談委員間效度的不同（Dipboye, Gaugler, & Hayes, 1990; Gehrlein, Dipboye, & Shahani, 1993; Green et al., 1993）。然而，最近的研究暗示面談委員效度的不同是由於抽樣誤差（Pulakos et al., 1996），但該證據是奠基於十分結構性的面談上，因此，可能減少了面談委員的不同所造成的影響。

十三、面談期間，面談委員的問題討論

在進行面談時，有些面談委員偏好與其他面談委員討論應試者的答案，可能是為了使自己的評分更客觀，但這樣做卻違反了獨立評分的精神，因而降低了面談的結構性。經由面談委員討論具有共識性後才進行獨立評分，可能會使無關訊息進入面談中，使面談委員對不同的應試者產生不同的評價標準（Cook & Campbell, 1979; Campion et al., 1997）。

然而，有後設分析研究指出，在結構式小組面談的條件下，共識性評分的效度（0.64）顯著高於各面談委員的平均評分效度（平均0.41）（Wiesner &

Cronshaw, 1988)。另外研究則發現，在結構性面談的情境下，共識性評分與獨立評分後的平均評分在面談效度上差異不顯著（Pulakos et al., 1996）。是故，共識性評分對面談效度的影響還需進一步澄清。

十四、面談訓練

如同在測驗施測上，施測問卷者必須受過特別的訓練，才能使施測的過程依標準化的程序進行。同理，訓練面談委員也必須接受一些先期訓練，才能使評分更加一致。面談訓練常從面談的目的及背景的描述開始，接著討論面談本身，訓練可能包括如何寫出面談問題，或者更典型地如何使用已寫出的問題（Campion et al., 1997）。訓練計畫時間從數個小時到一週都有，但大部分是一、二天，許多訓練計畫使用行為示範、角色扮演及實務面談，然後由訓練者及同學提供回饋（如，Motowidlo et al., 1992）。一般而言，廣泛的訓練內容，包括提供回饋的實務面談等，對面談效度的提升較有幫助（Dougherty, Ebert, Callender, 1986; Pulakos, Nee, & Kolmstetter, 1995）。但是若訓練的內容範圍太小，或訓練的時間太短，對面談效度是沒有幫助的（Vance, Kuhnert, & Farr, 1978）。

十五、計分方式

事先規定一套客觀的計分方式，有助於面談結構性的提升，若是在面談結束後才決定如何計分，可能會產生一些偏失，而有利於某些特定人士。常用的計分方式是對所有問題或向度使用簡單平均或加總，有時會根據題目的重要性或與效標的關係來做不同的加權，或者藉由特定問題的數目來反應出其重要性（Campion et al., 1997）。

相較之下，共識性評分似乎也有事後才決定評分的意味，儘管仍保有獨立評分的程序，但其客觀性仍不如純粹將各位面談委員的獨立評分加總或平均。然而，許多研究發現共識性評分的面談效度較高或至少不會較低（如 Pulakos et al., 1996; Wiesner & Cronshaw, 1988），因此，有必要進一步研究。

小結：綜合上述15項影響結構性面談因素，假如在時間、物力及財力的許可下，當然掌握各種能影響效度的因素是最理想的狀況，也就是說，口試的主辦單位應由廣義的結構著手。但是，在有限的時空環境下，本研究希望能對上述15項因素的重要性做一個輕重緩急的衡量，同時考慮簡單、便利的原則，為日後國內教師甄試提出方便可行且不失效度的可行性程序。

參、研究方法

一、研究樣本

受試者為九十三學年度臺北縣教育局辦理中等教師聯合甄選之面談委員，有效樣本215人，其中男性131人，約占61%，女性84人，約占39%。

在服務學校與職稱方面，今年度臺北縣教育局辦理之中等教師聯合甄選中，有99%的面談委員任職於臺北縣，僅有2位（0.9%）任職於臺北市。且大部分的面談委員是學校中的校長、主任、組長，如64.5%的面談委員是學校主任，校長與組長約占29%，僅有少部分的面談委員（約4%）是學校教師。

在教學資歷與行政資歷方面，今年度擔任教師甄試的面談委員，99%的人具有教學與行政資歷，其中更有35%的人，教學年資已達21年以上，屬於資深教師。另外，行政資歷在10年以上的人亦達50%以上，可見面談委員大部分屬於教學與行政資歷豐富的人。

在擔任面談委員之經驗方面，除了教學與行政資歷之外，此次聘請的面談委員中，不曾擔任校內自辦教師甄試面談委員者僅有19.4%，大部分的人（約80%）皆曾經擔任過校內自辦教師甄試的面談委員，其中67.2%的人更擔任過面談委員1~5次。

二、研究工具

研究工具為自編之「現階段中等學校教師甄試口試結構化現況調查問卷」。在工具發展過程中，以立意抽樣

方式選取曾經擔任教師甄試面談委員12人，再以滾雪球抽樣方式，請每位被選取者推薦10位曾擔任面談委員者，並詢問其填答意願，共發出106份問卷，回收有效問卷共計100份，其中每位受試者均需填答兩次相同的問卷。

在信度方面，面談委員版問卷信度以再測信度為主，施測間隔14天，問卷回收率達100%。由於本問卷為現況調查表，並非態度或成就量表。而且問卷題型包含單選、複選、多重排序、問答等，因此，各題得分無法加總，故再測信度方面，刪除問答題、多重選擇題、排序題型後，選取19題單選題進行Kappa一致性係數檢定，Kappa一致性係數在.031至.419之間，達.05或.01之顯著水準，顯示問卷單選題在時間方面的穩定性令人滿意。值得注意的是，單選題未達顯著之題目，依測驗學理角度原應刪除，但考量問卷以詢問最近五年內的口試情況，就實務角度而言，口試情況每年皆有變動，因此，有多數受試者在填答時常不知如何選擇，或將之視為複選題勾選，造成統計運算時視為遺漏值處理，進一步使結果未達顯著。研究者再度徵詢專家學者，並考量問卷題目的內容效度後，決定保留未顯著題目，使問卷更能符應結構化口試學理。

在效度方面，以專家效度為主。在面談委員版部分，從口試效度的文獻分析中所歸納出廣義式結構化口試15項

議題，依專家建議可歸納成四個層面，分別為口試內容、口試流程、口試計分、口試滿意度，且問卷內容大部分屬於口試實務經驗的反應題型。經由三次專家小組的焦點團體晤談，徵詢目前任教於中學之資深教師、主任及曾擔任學校教評會委員教師、校長、大學教授等15人，請其針對預試問卷加以修正，以求符合研究目的，並能確實反應教師甄試口試的現況，以及口試結構化的程度。經由多次修改，面談委員版內容已能展現理論與實務的效益，可謂具專家效度。

三、資料分析

由於問卷共計35題，每題均請面談委員根據自身參與口試經驗與主觀感受進行勾選，題目類型包含單選題與複選題。由於單選與複選題非如李克特量表有強度等級之分，亦無反向計分題，因此，無法將各題得分加總。在本研究中，僅以SPSS for Windows套裝軟體程式之百分比統計，呈現所有受試者在各題項上的資料分布情形，以瞭解現階段教師甄試口試的結構化程度概況。

此外，由於問卷中夾雜了單選題與複選題，以至於有些受試者會誤將某些單選題視為複選題，導致研究者在計算該題時，無法正確地歸類。然而，要蒐集到面談委員的資料並非易事，若因為受試者誤將某題的單選與複選弄錯，

而刪除該份問卷資料，恐將遺漏更多有效的訊息，因此，研究者決定保留該筆資料，並將受試者填答不清、重複勾選或遺漏填答的該題選項資料定義為遺漏值。

肆、研究結果

為瞭解目前國內教師甄試結構化的程度，研究者歸納整理結構化口試每一個向度在理論上的定義，並與目前的現況互相比對，使讀者對結果能一目瞭然，其研究結果參見表1。有關各向度之選項百分比，羅列於附錄，以免繁複。

一、教師甄試結構化程度

目前，教師甄試（以臺北縣為例），仍處於「低結構化」的狀態。由工作手冊觀之，研究者認為其他縣市大致也是大同小異，因為國內對於「結構化口試」的定義及運用，仍是十分生疏的。

結構化口試的15個細項中，以「定錨評分」的問題最為嚴重，一般而言，大家對此一無所知。就狹義的觀點，所謂結構化口試，僅局限在建立一套口試評分工具及標準。因此，定錨評分的疏漏為降低口試結構化的要害。口試的結構化程度，攸關口試效度的高低，結構化程度低，效度也就低，這是最令人堪慮的事情。而在實際研究過程

表 1 教師結構化口試議題分析

議題	理論	現況
1. 由「工作分析」建立口試題目	所謂工作分析即是分析特定工作的重要細節，使能對工作的內容有一提綱挈領的瞭解。工作分析必須由該工作領域的專家學者來進行，以確保所發展出的口試問題具有效度。	口試問題幾乎是由面談委員自由命題（93.9%），沒有經過一套事先嚴密的標準化產出程序（附錄附表1）。只有少部分是由主辦單位提供題庫，考試委員從中選擇發問（5.6%）
2. 每一位應試者均被問及相同的問題	從面談的角度來看，使用相同的口試問題可以提高面談的結構性，使得面談委員的評分建立在相同問題的基礎上，如此評分結果比較才有意義。	在教師甄試口試中，每位應試者被問及相同問題的機率極少（6.1%）（附錄附表2），而且也未能保持發問的順序（附錄附表3），以維持標準化的程序（93.9%）。
3. 即興問題或後續追問的處理	在標準化的程序上，應該要有一致的追問方式，倘若對某些特定應試者感興趣就追問的更仔細，而對不感興趣的應試者則匆匆帶過，就樣並不符合口試的標準化程序。	僅有少數（一成七）的人有自己特定的提問與追問方式，而約七成的面談委員並無固定的提問與後續追問方式，通常會視當時情況與應試者的反應而定（附錄附表4）。
4. 問題形式與內容	根據國外的研究發現：以「情境式問題」及「過去行為問題」的面談效度最佳，而以「心理特質問題」的面談效度最低。	國內的口試問題有83.3%是屬於「情境式問題」，故應有一定的效度存在，但是在效度最低的「心理特質」問題上也高達63.9%（附錄附表5），故有必要降低此種問題的比例。
5. 面談的時間與題數	面談的時間，範圍在3至120分鐘間，而平均面談時間是38.75分鐘；面談的問題題數，其範圍在4至34題間，而平均題數是16.50題。	五成八的委員認為6~10分鐘為最理想的口試時間，然亦有三成四的委員認為需要多於10分鐘的時間方足夠（附錄附表6），但是在此次教師甄試中，所有口試時間均少於10分鐘，其中更有一成五的口試時間是少於5分鐘的（附錄附表7）。這麼短的時間是否能真正瞭解應試者的想法，或讓應試者充分表達自己的想法。
6. 輔助訊息的控制	所謂的輔助訊息，是指除了正式的面談外，可用來幫助面談委員做決策的資訊。面談過程若能一致且客觀地使用輔助訊息將增加面談的結構性，反之，若未有效控制輔助訊息將對面談結構造成重要威脅。	就應試者所提供的輔助訊息部分，委員們認為傑出表現、獎狀或證明是最重要的，其次是教學經歷證明，第三重要的是教學經歷，第四是個人基本資料。而第二專長證明及推薦信則被視為不重要（附錄附表8）。就輔助訊息的計分而言，六成二的委員將其融入口試過程中計分，二成一的委員自行決定，僅少於一成的委員不列入計分（附錄附表9）。
7. 應試者發問的處理	在面談結構性的考量下，應試者在面談期間的隨意發問可能會使面談結構性受到破壞，因此，事先規定應試者的發問時機是必須的。	在口試過程中，若有應試者主動發問，有五成以上的面談委員會視問題而有不同的作法，有近三成的面談委員指出，未曾發生應試者主動提問的情形，僅有少數面談委員會制止應試者發問或迴避應試者的問題（附錄附表10）。
8. 立即且多次的評量	就結構性而言，對每位應試者回答的答案進行評分，面談委員對每個問題的答案都能立即評分，可以使答案與評價的連結更	僅有少數的委員（4.6%）會逐題循序計分，有六成的委員對應試者的評分方式為綜合口試整體表現而評定一個總分，然

表 1 教師結構化口試議題分析（續）

議題	理論	現況
	緊密，因而提升面談的結構性。	而，亦有三成的委員會針對應試者的表現先評出一個等級，最後再視所有應試者的表現情況調整分數（附錄附表11）。
9. 定錨量表的使用	定錨量表是指對評分量尺上的每一個量尺點所代表的程度或意義加以描述，描述的程度可以是十分詳細地指出每個量尺點所應具備的能力或表現，也可以簡單地只點出如非常好、好、普通、不好、非常不好等等。	臺北縣教師聯合甄選委員會已明定一套評分範圍，亦即每位面談委員的評分範圍必須在70~90分之間。然而，設定評分範圍並不同於定錨評分，定錨評分必須先產生題庫，再針對每個問題的不同類型的答案設定一套明確的評分標準。由本次調查發現，並沒有真正的定錨評分機制來幫助面談委員作評分的判斷，這對不同評分風格的口委來講，將形成對口試標準化的一大偏誤來源。
10. 筆記的運用	藉由作筆記可以減少記憶負擔，確認重要的相關資訊被記錄，並避免初始及新近效應（regency & primacy effects），因而增加面談的結構性。筆記的詳實度愈高，面談委員的評價將愈精確，因此，愈有助於效度的提升。	有八成四的面談委員會主動針對應試者在口試過程中的表現予以簡短地記錄，僅極少數（2.3%）的委員會詳細地記錄（2.3%），但有一成五的面談委員並未對應試者的口試表現予以記錄（附錄附表12、13）。就記錄內容而言，近七成的面談委員會記錄應試者表現好壞的評價，其次，六成的面談委員會記錄應試者的態度、情緒、外表等，有八成七的面談委員也會記錄應試者的人格特質，較少的面談委員（24.6%）會記錄應試者於口試過程中表現出的行為（附錄附表14）。
11. 多重面談委員	使用多位面談委員可以減少單一面談委員對應試者的衝擊，因此，聚集較多面談委員可以避免單一主觀的判斷或評分偏好，並且消除隨機誤差。	此次教師甄試每個試場只有兩位面談委員，很難保證評分不會受極端分數的影響。
12. 應試者均能接受相同面談委員的口試	對不同的應試者使用不同的面談委員，增加了面談委員變異的干擾，不僅降低面談結構性，對效度也是一大影響。	就本次調查結果發現，由於應試者太多，所以不是所有應試者都能接受相同面談委員的口試，這也是大型招募場合所無法避免的問題。本次考試共有217人，因此，更應該在其他面向上維持標準化，才不至於導致口試效度的降低太多。
13. 面談期間，面談委員的問題討論	在標準化條件下，自當以獨立評分為佳，然而，在過去文獻的分析中，獨立評分或共識性評分似乎都有其優點。	本次調查結果發現近五成六的面談委員未能保持一致，僅針對特定應試者加以討論（附錄附表15），此種不一致容易衍生口試效度的爭議。
14. 面談訓練	一般而言，廣泛的訓練內容，包括提供回饋的實務面談等，對面談效度的提升較有幫助。	僅有一成六的面談委員主觀認為主辦單位有舉辦過先期訓練，且其中有近九成的面談委員表示訓練時間在半天內就結束（附錄附表16、17）。研究者以為，面談委員誤把口試前的說明會誤認為訓練。

表 1 教師結構化口試議題分析（續）

議題	理論	現況
15.計分方式	事先規定一套客觀的計分方式，有助於面談結構性的提升，若是在面談結束後才決定如何計分，可能會產生一些偏失，而有利於某些特定人士。	大部分（六成四）面談委員的評分標準均符合主辦單位規定的範圍，而幾乎可以說是沒有在口試後，方進行口試評分標準的討論（附錄附表18、19）。

中，研究者赫然發現從預試到正式施測，參與此次研究又具多年主辦口試的校長們，告訴研究者：他們從來不知道會有這麼多因素影響學生的口試成績，這真是一件嚴重的事情。

二、教師甄試之現況

除了口試結構化的議題外，本研究另外有關於現況的報導。茲分述如下：

(一)近八成七以上的面談委員覺得口試能挑選出優秀的老師，認為口試無法挑選出優秀的教師者僅占3.1%（附錄附表20）。另外，不同的行政資歷、教學資歷、職稱的面談委員，在口試是否能選出優秀教師方面的統計考驗均未達.05顯著水準。亦即對不同的性別、行政資歷、教學資歷、職稱的面談委員而言，其共識為口試可以挑選出優秀的教師。

(二)有六成五的面談委員認為，經由口試結果甄選上的教師，其日後表現大致符合在口試時的預期結果，持反對意見的僅有3.1%。值得注意的是，有三成的面談委員表達無意見或不知道

（附錄附表21），可能是因為許多面談委員對於甄選上的教師，在日後並無接觸或瞭解，因此，無從瞭解教師的日後表現。

(三)對不同的性別、教學資歷、職稱的面談委員而言，其共識為：經由口試結果甄選上的教師，其日後表現大致符合在口試時的預期結果。另外，對於不同行政資歷、教學資歷、職稱之面談委員，在「對甄選教師之預期」上之統計考驗均未達.05顯著水準。

(四)不同行政資歷的面談委員在此項的意見並不一致，達統計顯著水準（ $F=2.59, p<.05$ ）（附錄附表22），進行事後比較發現，無行政資歷與5年以下行政資歷者，比具有11年以上行政資歷者較「不同意經由口試結果甄選上的教師，其日後表現符合在口試時的預期結果」。

(五)就口試成績占總成績而言，近七成的面談委員認為21~40%較為適當，但有二成的人認為，口試成績占總成績20%以下較為適當，而一成的人認為口試成績占總成績41~60%較為適當（附錄附表23），可見全部的面談委員

皆認為口試並非決定一個人錄取與否的關鍵，口試僅是總成績的一部分。

(六)最適合擔任面談委員的是主任與校長，約有九成左右的面談委員認為主任適合擔任面談委員，有七成五的面談委員認為校長適合擔任面談委員。值得注意的是，並未有許多面談委員支持教育局人員（6.6%）與家長代表擔任面談委員（7.6%）（附錄附表24）。

(七)有五成左右的面談委員人選是主辦單位決定的，另外，亦有四成左右的面談委員是由教評會或教務處提供名單，再由校長圈選決定。僅有少數的面談委員分別是由校長自行決定、各校依比例推薦、開會決定的（附錄附表25）。

(八)五成的人都是口試前三天才被告知需擔任面談委員，其次，有三成五的人是口試前七天被告知需擔任面談委員。值得注意的是，有少數的人是在口試前一天或臨時被告知需擔任面談委員（附錄附表26），此種情形是否會影響口試工作的品質，或影響應試者的權益，值得關注。

(九)有八成八的面談委員認為，大部分的面談委員詢問的問題是適合的，僅有1.6%的人認為大部分的面談委員詢問的問題是不適合的（附錄附表27）。

(十)有七成的面談委員認為，口試工作最困難的部分為「口試問題的應答

評分」，其次，有近四成的面談委員認為「口試題目的擬定」與「輔助資料的評分」（如自傳履歷、教學檔案、成長檔案等）也是口試工作中最困難的部分。至於，最為應試者關注的人情壓力，僅有3.6%的面談委員提及此項（附錄附表28）。

(十一)有五成以上的面談委員認為，口試工作最需加強的部分為口試問題的擬定，其次，也有四成五的面談委員認為必須加強口試的計分標準，至於面談委員的訓練也有近四成的面談委員認為有待加強（附錄附表29）。這些資料亦反應了結構化的口試首要關鍵「由工作分析建立題庫」的必要性與重要性。

三、小結

理論與實務上最大的共識如下：在檢視國中教師結構化的程度中，最缺乏的部分是「由工作分析建立口試問題」及「定錨評分」，而面談委員最感困擾的部分也是「口試問題的擬定」及「計分標準的訂定」。因此，教育部應該聚集人力，戮力從事對教師職業進行工作分析，建立專業教師甄試口試題庫，並具此建立定錨評分的機制。

伍、討論與建議

結構化口試議題的重要性乃在於國內對此議題相當生疏，而且在實務上

有其迫切性，而口試一旦被結構化並且被正確地使用後，方能有效提高口試的效度（如，Campion et al., 1997; Huffcutt & Arthur, 1994）。事實上，只要經過精細的設計，並正確的使用，面談可以是很有效度的（Gatewood & Field, 2001: 521）。亦即結構化對口試而言，可以經由標準化的程序，提升口試在心理測驗上的屬性，或者說讓面談委員知道「要問什麼題目」及「要如何評分」。經由本研究發現，至少在教師甄試口試這個部分，口試結構化程度是甚低，在理論上，此將對口試的效度，造成很不利的影響，進而影響人員甄選的品質。研究者以為，臺灣未來各種口試在結構化的程度，至少應朝狹義性定義，即是由良好工作分析擬定口試題目，以及建立定錨評分系統。至於廣義式的結構化口試，則需視各個市場成本及時間的配合。

Huffcutt與Woehr（1999）企圖找出如何降低公司人事面談成本的有效方式，也就是希望將實務上的問題與狹義結構性面談作一番比較。Huffcutt與Woehr將120個有關面談效度的研究，使用後設分析的統計方法來處理。將「狹義的面談結構性」（依「面談問題」及「評分標準」區分為四個水準）、「面談委員訓練」（分有、無兩類）、「面談委員組成」（分為小組式或個別式兩類）、「面談委員標準化」（分

有、無兩類）及「面談筆記」（分有、無兩類），這五個因素作為研究的自變項，而將效度係數作為依變項。Huffcutt與Woehr發現：一旦面談依「面談問題」及「評分標準」加以控制後，面談的效度便能達到一定的水準，在有經費、人力限制下，最簡單的方式就是將面談結構性的焦點鎖定在「面談問題」及「評分標準」的建立與安排上。

目前，假如欲在國內實施結構化口試，可能會遭遇下列的困難，這些現象在國外亦有類似的狀況，以下是相關的研究。首先，實務工作者可能對結構化口試理論的相關知識，知之甚少。即使實務工作者對結構化口試有所瞭解，但也可能懷疑其信度（credibility）、重要性（relevance）及實用價值（Oviatte & Miller, 1989; Terpstra & Rozell, 1997）。其次，結構化口試有這麼多規則要遵守，使得面談委員的權力及自由度大大受到限制，也會降低面談委員遵守的意願（Dipboye, 1997），委員會覺得口試有點無聊，缺乏挑戰性。再次，結構化口試可能會減低口試的意願（Dipboye, 1997; Zee, Bakker, & Bakker, 2002），理由是，應試者會比較喜歡活潑、溫暖、和善的面談委員，而這類的委員比較容易出現在非結構式的晤談中，因此，也有另類的聲音支持非結構式晤談口試，是較占優勢的（Latham

& Finnegan, 1993; Zehlein, 1985)。最後，結構式面談的時間及經費的成本，較非結構式面談高，假如碰到主管單位對結構式面談的好處採懷疑的態度，那麼結構式面談將會窒礙難行（Terpstra & Rozell, 1997）。

另外，最近Zee等人（2002）以「為什麼結構式面談甚少被用於人事甄選」為題，研究人事甄選的主管（N=99）較傾向使用結構式或非結構式的面談？結果發現這些主管比較傾向於使用非結構式的面談。而Zee等人乃以廣義式結構化面談來檢驗人事甄選主管的實際甄選行為所得到的結果。然而，這篇文章是在過去九十年的人事心理學在結構化口試中，極少數實證研究支持非結構口試的文章。

研究者以為，上述支持非結構化口試的研究，比較側重在實務工作者（亦或稱面談委員）的需求與試務工作的方便性。其實，它與結構化口試的真諦是有所距離的。口試結構化主要的目的乃是在提升口試的效度，以期提升人員進用的品質。各位讀者，請試想在教育的領域，是試務工作繁瑣麻煩大，抑或是選擇一位不適任教師麻煩大？相信這個道理是不言而喻的。一位不適任教師對學校的負擔，對學生學習的負面影響，其嚴重性真是罄竹難書。另外，一個結構化的口試，在心理測驗理論上較完整，對於目前競爭激烈的教師甄試錄

取率只有1%~2%的現況，口試的主辦單位或是教育主管機構，必須有強而有力的說帖，說明教師甄試口試可以是公平與公正的，以召信於社會大眾。基於上述理由，研究者以為，結構化口試的研究與推廣有其必要性。

一、理論上的建議

建立教師甄試口試結構化的歷程，非一蹴可幾的，而必須循序漸進。所謂的結構化程度亦非0與1的分野，而是程度上的差別；某程度可大致分為非結構、狹義性結構，最後到廣義性的結構。非結構化口試即是傳統上目前國內一般教師甄試，而狹義性的結構化口試則必須側重「口試內容擬定」與「定錨評分的建立」；廣義性的結構化口試則側重全面性的結構化，包括15項議題。從非結構至結構化其要考慮的因素很多，例如：時間、人力、物力及主辦單位的想法及意願。而本計畫以臺北縣教育局所主辦的教師聯合甄試為對象，研究者本身的主導性不強，是故，研究者以為可先瞭解目前國中教師甄試的現況，接著著手進行口試题目的定錨評量，以建立狹義的結構化口試。最後，再經由面談訓練達到促進廣義結構化口試的目的。期待此一過程能開始建立口試／面談結構化的里程碑。

研究者以為目前教育的學術界，需要做三件事：(1)建立專業口試題

庫；(2)建立定錨評量系統；(3)面談訓練。尤其是前兩項，必須由教育學術界自己發展，方能有效地解決目前面臨的問題。其流程可以借鏡於工商業界或人事心理學，但內容必須自己發展。此有待測驗學界各位先進一起努力。

面談訓練是最常被用來改進面談／口試的方法 (Dipboye, 1992)，主要目的在保證一種程序，使各種影響因素能在面談中被正確地實施。事實上，面談訓練就好像訓練測驗主試者一樣重要 (Society for Industrial and Organizational Psychology, 1987)。更重要的是，面談訓練是一種很容易被教得會的學問 (Howard, Dailey, & Gulanick, 1979)。

二、實務上的建議

從事教師甄試的研究工作，充滿著許多的挑戰，研究者願意記錄下來，以饗讀者，並提供給未來相關研究作參考。

(一)時間及樣本選取的困難度

由於樣本散布在全國各縣市，且各縣市教育局舉辦之教師聯合甄選或學校自辦的教師甄試都於7月10日至20日之間舉行，甚至有六個縣市的甄選集中於7月17、18兩日，如臺北縣、高雄縣、臺南縣等。在同一時間內欲蒐集到所有樣本的資料，實非易事，且需耗費極大的人力、物力，因此在最短的時

間、最精簡的人力、物力及金錢下，並兼顧南北樣本的平衡性，本研究選取臺北縣、臺南縣、高雄縣教育局舉辦之中等教師聯合甄選作為本研究的施測樣本。目前臺北縣的資料蒐集最完整，全部的面談委員及應試者，均予以施測。而因天候、場地及人員配合度問題，臺南縣完全沒有資料，高雄縣有部分資料。目前研究報告結果是以臺北縣為主，也因為如此，本研究的推論上，也甚受限制。

(二)場地及天候問題

在臺北縣部分，因報名人數眾多，所以考場分散於五所學校舉行，如中和國中、錦和高中、永平高中、福和國中、永和國中。然而，考試當日上午參加筆試者約1,500人，缺考率達65%，因此，各校於午餐會報時緊急更動下午口試會場的人員安排與口試場地，連帶影響本研究面談委員的心情。

在高雄縣部分，報名人數為1,808人，考試場地分散在鳳山高中、青年國中、中山國小，但考試當日，高雄下起滂沱大雨，有些場地進水，連帶影響應試者填答問卷的意願，有些應試者在考完之後便匆忙地離去。在臺南縣部分，如同高雄縣，一場午後大雨使得部分場地嚴重進水，協助人員忙於尋找新的場地與搬運口試資料，無暇顧及問卷施測事宜，以致於問卷淋濕，且亦忘了發放

問卷。故高雄與臺南的資料就沒辦法使用。

(三)人員配合的困難度

在本次問卷施測過程中，與相關人員聯繫，並請求協助是本研究最大的困難點。因為若由研究者聘請工讀生進行施測，恐會干擾口試人員的試務工作，且成效不佳。然而，近年來參加教師甄試人數眾多，試務工作相當繁重且敏感，若請求相關人員在試務當中協助問卷施測，試務人員則需付出更多的時間精力去規劃填答問卷的路線與方式。某些縣市在接洽的過程中，就直接告訴研究人員，不方便進行研究。因此，此次能徵得臺北縣聯合辦理教師甄試的同意，實屬難能可貴。

是故，若欲從事與此類相關的研究有志之士，必須要克服上述的議題，方能完成研究。

參考文獻

- 林宜靜。北市教師甄試9千人搶96缺額。聯合報。2005年6月13日，取自 <http://www.taotc.org/enews/e-news70.htm>
- 林麗雪。教職難求5萬人搶2千個缺。民生報。2005年6月7日，取自 <http://career11.mac.nthu.edu.tw/edu/road/940607a.html>
- 胡悅倫(2006)。教師甄試口試——應試者的觀點。教育心理與研究，送審中。
- 喻文玟(2005, 6月7日)。臺中女中甄選老師493人搶7缺額新聞。聯合報，C1版中市文教。
- Arvey, R. D., & Campion, J. E. (1982). The employment interview: A summary and review of recent research. *Personnel Psychology, 35*, 281-322.
- Beatty, R. N. (1986). *The five-minute interview*. New York: Wiley.
- Burnett, J. R., Fan, C., Motowidlo, S. J., & Degroot, T. (1998). Interview notes and validity. *Personnel Psychology, 51*, 375-396.
- Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology, 50*, 655-702.
- Campion, M. A., Pursell, E. D., & Brown, B. K. (1988). Structured interviewing: Raising the psychometric properties of the employment interview. *Personnel Psychology, 41*, 25-42.
- Carrier, C. A., & Titus, A. (1979). The effects of note-taking: A review of studies. *Contemporary Educational Psychology, 4*, 299-314.
- Carrier, M. R., Dalessio, A. T., & Brown, S. H. (1990). Correspondence between estimates of content and criterion-related validity values. *Personnel Psychology, 43*, 85-100.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi experimentations: Designs and analysis for field setting*. Boston: Houghton Mifflin.
- Delery, J. E., Wright, P. M., McArthur, K., & Anderson, D. C. (1994). Cognitive ability tests and the situational interview: A test of incremental validity. *International Journal of Selection and Assessment, 2*, 53-58.
- Dipboye, R. L., & Gaugler, B. B. (1993).

- Cognitive and behavioral processes in the selection interview. In N. Schmitt & W. C. Borman Associates (Eds.), *Personal selection in organizations* (pp. 135-170). San Francisco: Jossey-Bass.
- Dipboye, R. L. (1992). *Selection interviews: Process perspectives*. Cincinnati, OH: South-Western.
- Dipboye, R. L. (1994). Structured and unstructured selection interviews: Beyond the job-fit model. In G. R. Ferris (Ed.), *Research in personnel and human resources management, 12* (pp. 79-123). Greenwich, CT: JAI Press.
- Dipboye, R. L. (1997). Selection interviews: Why do they work? Why are they underutilized? In N. Anderson & P. Herriot (Eds.), *International Handbook of Selection and Assessment* (pp.455-473). New York: Wiley.
- Dipboye, R. L., Gaugler, B. B., & Hayes, T. (1990). *Individual differences among interviewers in the incremental validity of their judgements*. Paper presented at the Fifth Annual Conference of the Society for Industrial and Organizational Psychology, Inc., Miami.
- Dougherty, T. W., Ebert, R. J., & Callender, J. C. (1986). Policy capturing in the employment interview. *Journal of Applied Psychology, 71*, 9-15.
- Gatewood, R. D., & Field, H. S. (2001). *Human Resource Selection*. Orlando, FL: Harcourt College Publication.
- Gehrlein, T. M., Dipboye, R. L., & Shahani, C. (1993). Nontraditional validity calculations and differential interviewer experience: Implications for selection interviews. *Educational and Psychological Measurement, 53*, 457-469.
- Graves, L. M. (1993). Sources of individual differences in interviewer effectiveness: A model and implications for future research. *Journal of Organizational Behavior, 14*(4), 349-370.
- Green, P. C, Alter, P., & Carr, A. E. (1993). Development of standard anchors for scoring generic past-behaviour questions in structured interviews. *International Journal of Selection and Assessment, 1*, 203-212.
- Hakel, M. D. (1982). Employment interviewing. In K. M. Rowland & G. R. Ferris (Eds.), *Personnel management* (pp. 129-155). Boston: Allyn and Bacon.
- Harris, M. M. (1989). Reconsidering the employment interview: A review of recent literature and suggestions for future research. *Personnel Psychology, 42*, 691-726.
- Howard, G. S., Dailey, P. R., & Gulanick, N. A. (1979). The feasibility of informed pretests in attenuating response-shift bias. *Applied Psychological Measurement, 3*, 481-494.
- Huffcutt, A. I., & Woehr, D. J. (1999). Further analysis of employment interview validity: A quantitative evaluation of interviewer-related structuring methods. *Journal of Organizational Behavior, 20*, 549-560.
- Huffcutt, A., & Arthur, W. Jr. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology, 79*, 184-190.
- Hunter, J. E., & Hunter, R. F. (1984). The validity and utility of alternative predictors of job performance. *Psycho-*

- logical Bulletin*, 96, 72-98.
- Janz, T. (1982). Initial comparisons of patterned behavior description interviews versus unstructured interviews. *Journal of Applied Psychology*, 67, 577-580.
- Landy, F. R., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin*, 87, 72-107.
- Latham, G. P., & Finnegan, B. J. (1993). Perceived practicality of unstructured, patterned, and situational interviews. In H. Schuler, J. L. Farr, & M. Smith (Eds.), *Personnel selection and assessment: Individual and organizational perspectives* (pp. 41-55). Hillsdale, NJ: Erlbaum.
- Latham, G. P., & Saari, L. M. (1984). Do people do what they say? Further studies on the situational interview. *Journal of Applied Psychology*, 69, 569-573.
- Latham, G. P., Saari, L. M., Pursell, E. D., & Campion, M. A. (1980). The situational interview. *Journal of Applied Psychology*, 65, 422-427.
- Locke, E. A., & Latham, G. P. (1984). *Goal-setting: A motivational technique that works*. Englewood Cliffs, NJ: Prentice-Hall.
- Marchese, M. C., & Muchinsky, P. M. (1993). The validity of the employment interview: A meta-analysis. *International Journal of Selection and Assessment*, 1, 18-26.
- Mayfield, E. C. (1964). The selection interview: A reevaluation of published research. *Personnel Psychology*, 17, 239-260.
- Mayfield, E. C., Brown, S. H., & Hamstra, B. W. (1980). Selection interviewing in the life insurance industry: An update of research and practice. *Personnel Psychology*, 33, 725-739.
- McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599-616.
- Motowidlo, S. J., Carter, G. W., Dunnette, M. D., Tippins, N., Werner, S., Burnett, J. R., & Vaughan M. J. (1992). Studies of the structured behavioral interview. *Journal of Applied Psychology*, 77, 571-587.
- Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology*, 75, 640-647.
- Mumford, M. D., & Stokes, G. S. (1992). Developmental determinants of individual action: Theory and practice in applying background measures. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology: Vol.3* (2nd ed., pp. 61-138). Palo Alto, CA: Consulting Psychologists Press.
- Nevo, B., & Berman, J. A. (1994). The two-step selection interview: Combining standardization with depth. *Research & Practice in Human Resource Management*, 2, 89-96.
- Oviatte, B. M., & Miller, W. D. (1989). Irrelevance, intransigence and business professors. *Academy of Management Executive*, 3, 304-312.
- Pulakos, E. D., Nee, M. T., & Kolmstetter, E.

- B. (1995). Effects of training and individual differences on interviewer rating accuracy. In E. B. Kolmstetter (Chair), *Interviewer and contextual factors that make a difference in interview validity*. Symposium presented at the Tenth Annual Conference of the Society for industrial and Organizational Psychology, Inc., Orlando.
- Pulakos, E. D., Schmitt, N., Whitmey, D., & Smith, M. (1996). Individual differences in interviewer ratings: The impact of standardization, consensus discussion, and sampling error on the validity of a structured interview. *Personnel Psychology, 49*, 85-102.
- Schmitt, N. (1976). Social and situational determinants of interview decisions: Implications for the employment interview. *Personnel Psychology, 29*, 79-101.
- Schmitt, N., & Ostroff, C. (1986). Operationalizing the "behavioral consistency" approach: Selection test development based on a content-oriented strategy. *Personnel Psychology, 39*, 91-108.
- Schriesheim, C. A., Solomon, E., & Kopelman, R. E. (1989). Grouped versus randomized format: An investigation of scale convergent and discriminant validity using LISREL confirmatory factor analysis. *Applied Psychological Measurement, 13*, 19-32.
- Society for Industrial and Organizational Psychology, Inc. (1987). *Principles for the validation and use of personnel selection procedures* (3rd ed.). College Park, MD: Author
- Terpstra, D. E., & Rozell, E. J. (1997). Why some potentially effective staffing practices are seldom used. *Public Personnel Management, 26*, 483-495.
- Tullar, W. L. (1989). Relational control in the employment interview. *Journal of Applied Psychology, 74*, 971-977.
- Ulrich, L., & Trumbo, D. (1965). The selection interview since 1949. *Psychological Bulletin, 63*, 100-116.
- Van der Zee, K. I., Bakker, P., & Bakker, A. B. (2002). Why are structured interviews so rarely used in personnel selection? *Journal of Applied Psychology, 87*(1), 176-184.
- Vance, R. J., Kuhnert, K. W., & Farr, J. L. (1978). Interview judgments: Using external criteria to compare behavioral and graphic scale rating. *Organizational Behavior and Human Performance, 22*, 279-294.
- Wagner, R. (1949). The employment interview: A critical summary. *Personnel Psychology, 2*, 17-46.
- Wanous, J. P. (1980). *Organizational entry*. Reading, MA: Addison-Wesley.
- Wiesner, W. H., & Cronshaw, S. F. (1988). The moderating impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology, 61*, 275-290.
- Wright, O. R. Jr. (1969). Summary of research on the selection interview since 1964. *Personnel Psychology, 22*, 391-413.
- Wright, P. M., Lichtenfels, P. A., & Pursell, E. D. (1989). The structured interview: Additional studies and a meta-analysis. *Journal of Occupational Psychology, 62*, 191-199.

Zehlein, A. (1985). *Students' judgment of interviewers: The influence of gender and communication style*. Unpublished

master's thesis, Universität Erlangen-Nürnberg, Erlangen, Germany.

附錄

附表 1 口試問題的產生方式

	人數	百分比 (%)
依主辦單位提供的題目發問	1	0.5
主辦單位提供題庫，面談委員從中選擇發問	12	5.6
面談委員自由發問	200	93.9
總計	213	100.0

附表 2 口試問題詢問方式

	人數	百分比 (%)
詢問相同的問題	13	6.1
詢問不同的問題	26	12.1
不一定，視情形而定	175	81.8
總計	214	100.0

附表 3 面談委員的發問順序

	人數	百分比 (%)
所有應試者由固定委員發問	4	1.9
對每位應試者，由委員輪流發問	143	68.8
隨興發問，不特別規定委員發問次序	61	29.3
總計	208	100.0

附表 4 口試問題的提問方式

	人數	百分比 (%)
不給任何提示，只簡短複述問題	29	13.7
給予有限的提示	24	11.3
對答案深入追問	12	5.7
任何形式皆有，對有些應試者追問，有些則否	147	69.3
總計	212	100.0

附表 5 口試問題的類型

	次數	反應百分比 (%)	人數百分比 (%)
擔任特定職務的意願	83	11.5	39.5
問題情境的解決方式	175	24.3	83.3
背景問題，如工作或教學經驗或其他資格	146	20.2	69.5
專業理念與知識	179	24.8	85.2
自我描述、自我評價及期望等問題	132	18.3	63.9
其它	6	0.8	2.9
次數總計	721	100.0	343.3
人數總計	210		

附表 6 理想口試時間

	人數	百分比 (%)
5分鐘以下	9	4.2
6-10分鐘	125	58.7
11-15分鐘	64	30.0
16-20分鐘	6	2.8
不一定	9	4.2
總計	213	100.0

附表 7 實際口試時間

	人數	百分比 (%)
5分鐘以下	32	15.1
6-10分鐘	178	84.0
不一定	2	0.9
總計	212	100.0

附表 8 輔助訊息的重要性與排名

輔助訊息的類型	不重要	重要			小計
		排名 1	排名 2	排名 3	
個人基本資料 (自傳履歷、學校成績單)	81 41.5%	71 36.4%	13 6.7%	30 15.4%	114 58.5%
教學檔案	65 33.5%	62 32.0%	34 17.5%	33 17.0%	129 66.5%
教學經歷證明	66 33.8%	49 25.1%	53 27.2%	27 13.8%	129 66.2%
傑出表現、獎狀或證明	49 25.1%	76 39.0%	39 20.0%	31 15.9%	146 74.9%
修習第二專長證明	173 88.7%	6 3.1%	4 2.1%	12 6.2%	22 11.3%
推薦信	191 97.9%	0	1 0.5%	3 1.5%	4 2.0%
其它	184 94.4%	5 2.6%	1 0.5%	5 2.6%	11 5.7%

附表 9 輔助訊息計分方式

	人數	百分比 (%)
依甄選簡章規定比例計分	15	8.2
融入口試過程計分	114	62.0
自行決定	39	21.2
不列入計分	16	8.7
總計	184	100.0

附表 10 應試者發問處理

	人數	百分比 (%)
制止應試者發問	7	3.4
直接回答應試者的問題	17	8.2
迴避應試者的問題	6	2.9
視問題而定	117	56.5
未曾發生	60	29.0
總計	207	100.0

附表 11 面談委員評分方式

	人數	百分比 (%)
逐題循序計分	10	4.7
綜合整體評分	131	62.1
先評定等級再視情況調整分數	69	32.7
其他	1	0.5
總計	211	100.0

附表 12 面談委員是否做筆記

	人數	百分比 (%)
主動記錄表現	177	83.9
沒有主動記錄表現	34	16.1
總計	211	100.0

附表 13 面談委員記錄的內容 (複選題)

	次數	反應百分比 (%)	人數百分比 (%)
簡短地記錄應試者的回答	70	31.4	40.2
詳細地記錄應試者的回答	4	1.8	2.3
簡短地記錄評語與等級	147	65.9	84.5
其它	2	0.9	1.1
次數總計	223	100	
人數總計	174		128

附表 14 面談委員記錄類型（複選題）

	次數	反應百分比 (%)	人數百分比 (%)
試者的行爲	41	11.2	23.8
應試者的人格特質	98	26.7	57.0
描述應試者的態度、情緒、外表等	104	28.3	60.5
應試者表現好壞的評價	119	32.4	69.2
其它	5	1.4	2.9
次數總計	367	100.0	
人數總計	172		213.4

附表 15 面談委員對應試者表現的討論

	人數	百分比 (%)
會討論每位應試者的表現	47	24.1
僅討論部分應試者的表現	109	55.9
從不討論，各自評分	39	20.0
總計	195	100.0

附表 16 主辦單位是否曾邀集所有的面談委員舉辦口試先期訓練

	人數	百分比 (%)
有先期訓練	33	16.1
無先期訓練	172	83.9
總計	205	100.0

附表 17 口試先期訓練的時間

	人數	百分比 (%)
半天以內	34	89.5
一天	3	7.9
一天以上	1	2.6
總計	38	100.0

附表 18 口試評分標準的形成

	人數	百分比 (%)
主辦單位已預設評分範圍	137	63.7
面談委員討論出評分共識	36	16.7
面談委員各自決定	42	19.5
總計	215	100.0

附表 19 口試評分標準的討論確定

	人數	百分比 (%)
口試前	166	79.0
口試後	20	9.5
兩者皆有	24	11.4
總計	210	100.0

附表 20 口試是否能挑選出優秀的教師

	人數	百分比 (%)
非常同意	16	8.2
同意	155	79.5
無意見/不知道	18	9.2
不同意	6	3.1
總計	195	100.0

附表 21 甄選上的教師是否符合口試時的預期結果

	人數	百分比 (%)
非常符合	9	4.7
符合	117	60.6
無意見/不知道	61	31.6
不符合	5	2.6
非常不符合	1	0.5
總計	193	100.0

附表 22 不同行政資歷面談委員在 35 題之變異數分析摘要

變異來源	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>Sig</i>	事後比較
組間	3.74	4	0.93	2.59	.038	1>3,4,5 1：無或5年以下
組內	62.74	174	0.36			2：6-10年 3：11-15年
總計	66.48	178				4：16-20年 5：21年以上

附表 23 口試占總成績百分比

	人數	百分比 (%)
20%以下	40	20.4
21~40%	136	69.4
41~60%	20	10.2
61~80%	0	0
81~100%	0	0
總計	196	100.0

附表 24 適合擔任面談委員的人選

	次數	反應百分比 (%)	人數百分比 (%)
校長	150	27.7	76.1
主任	181	33.5	91.9
組長	53	9.8	26.9
教師	43	7.9	21.8
教育局人員	13	2.4	6.6
專家學者	48	8.9	24.4
家長代表	15	2.8	7.6
教師會代表	38	7.0	19.3
次數總計	541	100.0	
人數總計	197		274.6

附表 25 面談委員如何決定

	人數	百分比 (%)
開會決定	2	1.0
教評會或教務處提供名單，校長圈選	80	41.9
校長自行決定	6	3.1
主辦單位決定	102	53.4
各校依比例推薦	1	0.5
總計	191	100.0

附表 26 面談委員何時被告知需擔任面談委員

	人數	百分比 (%)
口試前一天	5	2.5
口試前三天	96	48.7
口試前七天	68	34.5
口試前一個月	25	12.7
臨時被告知	3	1.5
總計	197	100.0

附表 27 面談委員詢問的題目是否適合

	人數	百分比 (%)
非常適合	54	28.9
適合	111	59.4
無意見／不知道	19	10.2
部分不適合	3	1.6
全部不適合	0	0
總計	187	100.0

附表 28 口試工作最困難的部分

	次數	反應百分比 (%)	人數百分比 (%)
人情壓力	7	2.0	3.6
口試题目的擬定	76	21.8	39.2
應答的評分	136	39.1	70.1
輔助資料的評分 (如自傳履歷、教學檔案、 成長檔案等)	75	21.6	38.7
時間的掌握	44	12.6	22.7
其他	10	2.9	5.2
次數總計	348	100.0	
人數總計	194		179.4

附表 29 口試工作最需加強的部分

	次數	反應百分比 (%)	人數百分比 (%)
面談委員的挑選	54	14.3	27.8
口試問題的擬定	104	27.5	53.6
計分標準	88	23.3	45.4
面談委員的訓練	74	19.6	38.1
口試程序標準化	58	15.3	29.9
次數總計	378	100.0	
人數總計	194		194.8