

擴大用人機關參與公務人員考試之研究： 以司法官特考口試為例

黃一峯*

摘要

本文研究以司法官特考口試為例，探討如何擴大用人機關參與公務人員考試。研究透過焦點團體訪談，邀請司法院法官分享其參與司法官特考口試的經驗與看法。訪談結果顯示，現行口試制度仍有多項議題，包括評分標準與實務需求契合、透過考生分組提升公平性、口試題目設計充分反映司法官核心職能等。本研究建議提高用人機關代表在口試委員中的比例、建立口試試題檢討與回饋機制、增加情境模擬測驗等，以期精進口試的效度與公平性，確保考選出真正符合實務需求的司法官人才。作者並建議未來除應擴大研究對象，並須將檢察系統意見納入，方能對司法官口試做全面、整體的探討，尋求周延的精進方策。

關鍵詞：司法官特考、口試

*淡江大學公共行政學系專任副教授。作者感謝鞏宜萱女士的協助。

Expanding the Participation of Hiring Agencies in Civil Service Examinations: The Case of the Oral Examination in the Special Examination for Judges and Prosecutors

Irving Yi-Feng Huang, PhD*

Abstract

This paper examines how to expand the participation of hiring agencies in civil service examinations, using the oral examination of the Special Examination for Judges and Prosecutors as a case study. Through focus group interviews, judges from the Judicial Yuan were invited to share their experiences and perspectives on participating in the oral examination process. The findings indicate several issues with the current oral examination system, including the alignment of scoring criteria with practical needs, enhancing fairness through candidate grouping, and ensuring that oral exam questions fully reflect the core competencies of judicial officers. To improve the validity and fairness of the oral examination, this study recommends increasing the proportion of hiring agency representatives on the oral examination committee, establishing a review and feedback mechanism for oral exam questions, and incorporating situational simulation tests (or questions). These will help to ensure that the selection process identifies judicial officers who truly meet practical demands. Additionally, the study suggests expanding the scope of future research and incorporating input from the prosecutorial system to conduct a more comprehensive analysis of the oral examination process, ultimately seeking well-rounded improvements.

Keywords: Special Examination for Judges and Prosecutors, oral examination

* Associate Professor, Department of Public Administration, Tamkang University. The author appreciates the assistance of Ms. Yixuan Gong.

壹、前言

考試院係依孫中山先生的構想而設，其根本理念在於落實五權分立，讓考試院得以獨立辦理各類國家考試與人才選拔，避免行政機關在用人上出現「任人唯親」或關說等情形，杜絕歐美「分贓制」(spoil system) 的流弊，進而維持政府機構的公平與專業。此制度源於孫中山先生在研究西方三權分立的同時，亦保留了傳統科舉的獨立性，將「考試權」自「行政權」抽離，以對應人民對公共服務領域的多元需求。

在我國憲政體制下，考試院透過定期舉辦公務人員考試、專門職業及技術人員考試等方式，為國家選拔所需的各類專才，同時也確保透明的選才流程，避免外力干預或人事腐敗現象。尤其在過往威權時代後轉型至民主社會的歷程中，考試院透過整套獨立的命題與評分機制，成功提升政府施政效能及公信力。

在當今多元開放的社會氛圍下，考試制度面臨諸多新挑戰：其一，現行考試範圍能否隨社會變遷而調整，進而有效衡量公務人員的實務能力與創新思維；其次，應如何在嚴謹考評與有效篩選間取得平衡，同時兼顧公平與彈性；再者，面對資訊科技日新月異的衝擊，又該如何防範舞弊並維持考試品質。儘管如此，五權分立下的考試院，作為文官制度的核心機制，仍在維護公職人員專業素養與社會信任方面，具有無可取代的地位與功能，也是我國政治體制民主與自由的重要展現。

然而，近年來隨著社會事務日益繁複，公務人員工作內涵不斷擴增，僅以「確保公平」為中心的考試設計，往往無法充分因應外在環境對公務人員的多重要求，除了要鑑別基本知識與專業素養，也需能判斷考生是否具備充分的職能，尤其在面對高度機動且情境多變的業務時，能否展現快速應對、有效溝通與跨域整合的能力 (Grindle & Hilderbrand, 1995)。這些要素往往是筆試難以窺見的「隱性素質」，唯有用人機關積極參與並結合其對職能需求的了解，才可能在口試或其他考選方法中辨別出真正符合職務要求的人選 (李俊達, 2021)。若用人機關未能適度介入設計與評估流程，就可能造成考試結果與實際職務需求脫節，不利於有效選才。因此，如何在維持考試院獨立性與制度公平的前提下，積極傾聽並融入用人機關的聲音，已成為精進考試制度時的重要方向。

從實務觀點來看，用人機關參與考試流程不僅有助於加強監督考試品質、維護考生的權益，也能為整體選才帶來更深刻的意義。

第一，若用人機關能適度介入考試設計與執行，便能於考試過程中提供意見，協助考選單位釐清題目取向、篩檢方式以及評分準則，確保整個考試機制更貼近實務需求。

第二，用人機關最瞭解職務職能與工作內容，自然更能針對各項能力需求與考試內涵進行溝通和調整。當前公部門因應社會議題與國際變局，公務人員的工作往往需要快速應變、多元思考以及跨域協作，若能將這些能力具體反映在試題與評分標準中，將顯著提升考選的有效性。挑選出的公務人員更能契合實際工作所需，便能在職場上展現更高度的專業與效能。

第三，強化用人機關在考試制度中的角色，也有助於形塑符合時代需求的公務員團隊。透過機關與考選單位的持續互動，能讓考生的專業知識、政策分析與領導管理能力獲得更全面的評估，避免紙上談兵的考試形式。這些共同努力，最終目的在提昇政府部門人才素質，使公務體系得以更加適切地回應社會期盼，維護民主與法治價值。

本文以司法官特考口試為例，邀請司法院曾參與司法官特考口試的法官進行焦點座談，瞭解其參與口試的態度與意見。座談除探討法官對考選制度中題目設計、評分規準與實務需求的觀點，也關注法官們如何評估用人機關在題目擬定、考場觀察或後續訓練程序中所能發揮的功能與價值。

為更清楚闡述口試設計的基礎，本文首先對「職能」概念進行界定。職能係指個人執行特定工作角色時所需具備的知識、技能、能力與行為特質的綜合表現，涵蓋專業知識、實務技能及人格特質等多面向能力。在司法官特考的脈絡中，核心職能應反映法官與檢察官在執行審判、偵查、公訴等職務時所需的專業素養與實務能力。根據司法院與法務部的職能參考規範，司法官的核心職能包括：(1) 邏輯思辨與法律適用能力，透過理性分析與推理，適切認定事實並正確適用法律；(2) 證據評價與調查能力，依據經驗法則與論理法則，客觀評估卷證與證據；(3) 程序掌控與訴訟指揮能力，熟悉程序法規，確保公平訴訟程序；(4) 溝通與說理能力，以清晰、具說服力的方式表達判斷與決策；(5) 自律自省與廉潔操守，秉持公正、中立與高道德標準；(6) 敬業精神與責任感，忠於職守，勤勉辦案。這些核心職能不僅是口試題目設計與評分規準的基礎，亦是用人機關（司法院、法務部）對司法官專業能力與實務表現的期待。

目前的司法官特考在二試與三試之間大約有 83% 左右的錄取率（表 1），但考生對名次順位的判斷可能影響了三試的參與程度。由於錄取成績是採計二試

成績（占比 90%）與口試成績（占比 10%），因此，如何透過口試有效評估上述核心職能，以確保選拔出的司法官符合專業、公平與實務需求，成為考選制度的重要課題。

本研究蒐集並歸納法官的觀察與建議，期能進一步釐清在司法官考試情境下，如何有效納入用人機關的經驗與需求並精進考選過程，確保符合專業、公平與實務需求的三重需求。

表 1

司法官考試錄取人數統計（二、三試）

年度	試別	到考人數	錄取人數	錄取率
2024 年	三試	214	182	85.05%
	二試	2,774	217	7.82%
2023 年	三試	241	201	83.40%
	二試	2,776	245	8.83%
2022 年	三試	191	158	82.72%
	二試	2,523	192	7.61%
2021 年	三試	166	140	84.34%
	二試	2,649	168	6.34%
2020 年	三試	209	175	83.73%
	二試	2,724	212	7.78%
2019 年	三試	128	106	82.81%
	二試	2,531	128	5.06%
2018 年	三試	92	76	82.61%
	二試	2,045	92	4.50%
2017 年	三試	116	100	86.21%
	二試	2,226	116	5.21%

貳、文獻探討

一、公務人員考選中的口試

口試是一種能夠透過直接溝通評估候選人能力、人格、態度及適應性的評估工具 (Chen et al, 2013)。在審查公務人員職缺應徵者時，若該職位涉及行政管理、

監督職責或需要參與公共關係工作，通常會將口試納入考試的一環，以評估候選人的個人條件與職務適配性 (Powell & Levine, 1942)。口試內容通常包括對專業知識的考核、邏輯思維能力的測試以及語言表達能力的評估 (江宗正, 2024)。口試的主要功能有三：一是評估候選人是否具備勝任職位所需的知識與能力；二是評估候選人的行為能力與未來工作表現的潛力；三是了解候選人的動機、價值觀及與組織文化的適配程度 (Salgado & Moscoso, 2008)。口試的有效性並非自動形成，必須取決於口試的設計與執行方式 (Schmidt & Hunter, 1998)。

在英國，1854 年 Northcote-Trevelyan Report 即提出以公開競爭考試選拔文官，並逐步確立以筆試為主的錄取制度。隨著選拔方式的發展，口試逐漸成為考核的一部分，主要用於評估候選人的綜合素質與溝通能力。時至今日，英國文官口試內容豐富多樣，包括即席演說、外語閱讀、文件處理、會議組織等，並可能安排考生互評。口試的引入致使文官選拔由單一筆試向綜合評估發展，更全面地衡量考生的能力與潛質。

日本高度重視口試，將其視為評估考生能力與特質的核心工具，不僅在國家級考試中普遍採用，且地方政府亦多次口試，甚至輔以小組討論與心理測驗 (考選部, 2023; 吳新興, 2023)。日本公務員考試通過第一階段後，即進入「集團面接」和「集團討論」。前者是集體口試，由多位考官對應多名考生 (2-5 名考官對應 3-6 名考生)，持續約 20 至 50 分鐘，旨在提高口試效率並便於考官橫向比較申請人。口試官多由現職公務員擔任，以確保專業性與符合用人機關需求。口試問題主要圍繞申請動機、職業願景、相關政策及其實施挑戰，較少深入追問，考生間亦不互相辯論，但可能被問及他人回答相關內容，因此必須認真傾聽。與考官距離較遠，需注意音量、語速及儀態，確保表達清楚。口試時，考官更關注申請人的應對能力，而非深入瞭解其個性，因此展現熱情、自信、團隊合作意識至關重要，應確保回答緊扣自身經歷和公務員職責，避免冗長或模糊。良好的禮儀和清晰的表達可增強競爭力。日本的「官廳訪問」機制顯示實務機關有實質決定權 (江宗正, 2024)，地方政府可依職位需求調整口試題型與機制 (吳新興等, 2023)，可作為臺灣未來制度設計的啟發。

個別口試與集體口試各有不同的觀察重點。個別口試由考生單獨面對考官，主要評估其專業知識、邏輯思維、語言表達及應變能力，並可能涉及個人背景與生涯規劃。相較之下，集體口試則由多名考生共同參與，透過小組討論、情境模擬等方式，考察團隊合作、溝通協調、問題解決能力及臨場反應 (余民寧, 2013)。

個別口試重在個人綜合能力的深度評估，而集體口試則更強調互動能力與群體中的表現，考生需針對不同形式調整策略，以展現最佳實力。

口試不只是測試口語流暢度，而是對考生思維縝密度、分析力及創新力的全面檢視（江宗正，2024）。一個看似簡單的問題，可能暗藏多層含義，要求考生具備靈活應變的能力、邏輯推演的能力，以及從不同角度審視問題的格局。優秀的作答應該兼具「高度」——展現理論素養，「深度」——挖掘核心本質，「廣度」——涵蓋多面向思維，「溫度」——能貼近現實、引發共鳴。唯有透過批判性思考與創新表達，才能在競爭激烈的口試中突圍（馮文濤，2024）。

二、不同口試方法及其特點

不同口試方法具有不同的優缺點，選擇適合特定情境的口試方法有助於提高招聘決策的準確性。以下將進一步探討不同口試方法及其特點，分析其對於司法官口試改進的意義和適用性，詳如表 2。

表 2

口試方法分類與比較

口試方法	主要特徵	效度	公平性	候選人體驗	成本	適用性
非結構化口試	口試官自由提問，無統一標準	較低 (0.14)	較低 (受口試官主觀影響大)	高 (自然交流)	較低	了解個性、價值觀
結構化口試	問題與評分標準固定	中等 (0.63)	較高 (標準化評估)	中等	較高	司法官、公務員考試
結構化行為口試	依據過去行為預測未來表現	較高 (0.63)	較高 (問題標準化)	中等	較高	適用於專業、高責任職位
情境口試	假設情境測試應變能力	中等 (0.34)	較高	中等	中等	主管職位、危機管理
小組口試	多名口試官同時評估	中等	較高 (多維度評估)	較低 (壓力大)	較高	高階管理職位
壓力口試	施加壓力測試抗壓能力	較低 (0.14)	較低	較低 (體驗不佳)	較低	高壓環境的職位
AI 口試	AI 分析語音、表情等	未知(算法決定)	可能存在偏見	中等	較低	初步篩選、大規模考選
遠距口試	透過視訊或錄影口試	中等	較高	較低 (技術依賴)	較低	跨地區考選

（一）非結構化口試 (unstructured oral examination)

非結構化口試即傳統口試，口試官自由提問，沒有固定的問題結構與評估標準 (Salgado & Moscoso, 2005)。這種方法在公共管理領域廣泛使用，尤其在司法與行政官員招聘過程中常見。其優點在於：一，候選人能更自由地表達自己的看法，展現真實個性，有助於評估其溝通能力及文化適應性；二，候選人一般對非結構化口試有較好的主觀感受，因為較不拘束，氣氛更輕鬆自然。然而，其主要問題包括效度較低、易受到口試官主觀偏見影響，導致公平性不足。研究指出，非結構化口試的評估一致性差，不同口試官之間評分的差異較大 (Salgado & Moscoso, 2002)。Schmidt 和 Hunter (1998) 更指出，非結構化口試的效度僅約為 0.14，明顯低於其他方法。

（二）結構化口試 (structured oral examination)

結構化口試是公部門日益重視的形式，其特點為口試問題事先經由工作分析明確界定，所有候選人回答同樣的問題，評分標準一致 (Salgado & Moscoso, 2005)。

結構化口試的優勢包括較高的預測效度和較佳的公平性 (余民寧，2013)。Moscoso (2000) 的研究指出，結構化口試的預測效度明顯高於非結構化口試，以小組進行的結構化口試的效度達 0.63 (Sáez, 2007; Salgado et al., 2007)。結構化口試可有效降低口試官的主觀偏見，確保公平和透明性 (顧愛華、黃一峯，2014)。但其缺點也不可忽視：一是成本較高，需要設計標準化問題與評分準則，並需對面口試官進行專業培訓；二是候選人在高度結構化環境中可能感到壓力較大，難以自然表達真實能力 (Salgado et al., 2007；顧愛華、黃一峯，2014)。

（三）結構化行為口試 (structured behavioral oral examination)

結構化行為口試強調候選人過去行為對未來表現的預測效度，其要求候選人描述過往特定情境下的實際行為 (Salgado & Moscoso, 2005)。例如「過去你如何處理團隊內的衝突？」

結構化行為口試的預測效度相當高，主要來自於問題的情境性與回答的具體性，使得候選人的行為更容易評估且更加真實。研究顯示，結構化行為口試不僅具備高度的預測效度，還能提升評分者間的一致性與面試過程的公平性

(Motowidlo et al., 1992)，進一步強化其在公共與私人組織中作為標準化甄選工具的實用性。Turner (2004) 也提供了具體操作指南，建議以能力為基礎設計問題，並使用 STAR 架構提升面試的結構與效率。

然而，該方法的缺點包括成本較高，必須根據具體職位進行詳盡的工作分析設計問題；候選人必須回憶過去的具體行為，可能存在記憶誤差或刻意修飾 (Ellis et al., 200; Salgado, 2003)。此外，即便結構化面試能提升資訊真實性，但求職者可能仍會使用印象管理策略，其中包括誠實與欺瞞兩類型，前者能正向影響評價，後者則可能誤導面試官並在被識破後導致評價反彈 (Roulin et al., 2015)。進一步來看，Van Iddekinge 等 (2023) 指出，結構化行為面試在效度與候選人接受度間表現良好，但仍需注意對多元性與公平性的潛在影響，尤其在設計內容與語言時應避免文化偏見，以促進選拔制度的整體包容性。

(四) 情境口試 (situational oral examination)

情境口試透過提出假設的未來工作情境，評估候選人在特定情境中的決策與反應，效度約為 0.34 (Lievens, 2007)。例如「若你發現同事違規，你會如何應對？」，在面對這種涉及人際關係和真實情境的題目中，考生應保持積極正向的心態，友善待人、主動反思，並正確歸因。以與同事合作發生錯誤為例，應承擔應有責任，理性面對批評，並從心態與行動層面調整改進，以展現成熟穩健的處事能力。情境口試的優點包括較高的效度和良好的公平性，由於問題標準化，能有效減少主觀性 (Moscato, 2000)。但候選人可能給出理想化的答案，並非其真實的行動方式，這在某種程度上影響了其預測效度 (Salgado, 2003)。

隨著音訊、影像處理技術、電腦與通訊科技的成熟與硬體設備普及，多媒體技術已逐步應用於口試，如錄音題、影片題等，用來測試時事分析與應變處理能力。透過圖片、動畫、影片等方式，能豐富命題素材，提升考生的資訊處理與臨場反應能力。導入動態資訊能更真實地模擬工作情境，降低考生背誦、樣板作答或刻意包裝的影響，更準確測試其實際能力。雖然多媒體技術能提升考官與考生的互動體驗，但需增設投影幕、音響、麥克風、電腦等設備，並在考場安排技術人員支援。

(五) 小組口試 (panel oral examination)

小組口試由多名口試官同時評估候選人，能夠提供多維度的評價，具有良好

的公平性與客觀性 (Huffcutt & Arthur, 1994)。相較於單一口試官模式，多名評估者能針對候選人的能力、性格特質及組織適配度進行交叉驗證，從而降低個人偏見或認知盲點對決策的影響 (Lievens & De Paepe, 2004)。Barrick 等 (2010) 的實證研究指出，當口試官團隊涵蓋不同職能背景（如直屬主管、HR 專員與跨部門代表）時，其對候選人工作勝任力的預測效度可提高 18%-22%，此結果凸顯小組口試在提升評價客觀性上的潛力。但該方法可能增加候選人的心理壓力，在面對多名口試官的壓力情境下，可能因社交焦慮 (social anxiety) 影響其表現的真實性 (Salgado et al., 2007)；而且行政成本較高，需協調多方參與者的時間、設計標準化評分表，並進行口試官培訓以確保評分一致性 (Chapman & Zweig, 2005)。

（六）壓力口試 (stress oral examination)

壓力口試為刻意創造高壓環境測試候選人抗壓性，但其效度低，僅約 0.14，且候選人感受極差，可能影響機構的形象 (Hunter & Schmidt, 2004)。Chen 等 (2017) 發現，壓力口試能提高口試官對求職者情緒調節能力的評估準確度，但同時也會降低求職者對口試官友善度與組織吸引力的感知。不過，在司法高壓場景下適用於測試候選人的抗壓能力，對司法官考選可能具有參考意義。

（七）AI 與遠距口試 (AI and remote oral examination)

人工智慧 (AI) 口試利用演算法評估候選人，能提高考選效率，透過技術手段減少人為偏見 (Andrews & Bucher, 2022)。搭配遠距口試，能使遴選降低成本，克服地理限制。然而，使用 AI 口試存在潛在的算法偏見問題，可能影響公平性 (Salgado, 2007)。遠距口試雖降低了成本，適合大規模招聘，然而技術限制可能降低候選人真實的表達，影響互動質量。

以上對口試方法的綜合討論，對司法官特考口試制度的發展具有啟發。現行的司法官特考口試制度採用集體口試，但該模式可能仍需加強效度、增進公平性及強化多樣性。因此，結合多種方法和工具的優勢，或能改善目前的集體口試模式，提升選拔機制的科學性與合理性。同時，若能適度擴大用人機關的參與，讓其在題目設計與評分機制上提供意見並協同實施，也能使口試更貼近司法實務需求，增強評估真實度與公信力。可透過採用結構化口試與結構化行為口試為核心，輔以情境口試來提升效度與公平性，並適度引入 AI 口試與遠距口試降低行

政成本。如此可在確保選拔質量的同時，提高選拔過程的公平性與多樣性，實現更為優化的司法官遴選制度。

三、理想的口試建立與實施過程

在以上討論的基礎上，要建立理想的口試制度，應確保其效度、公平性、候選人體驗及成本效益四大核心要素 (Salgado & Moscoso, 2005; Schmidt & Hunter, 1998)，可遵循以下步驟來設計與實施。

(一) 口試設計階段

在口試設計階段，首先需要進行職位分析與能力需求界定。透過專家會議，綜合司法實務與學術領域的專業意見，明確司法官職位所需的核心能力，包括法律專業素養、邏輯推理、道德判斷以及壓力管理等關鍵要素。這一階段的目標是確保口試的設計能夠準確評估候選人的適任性，使評選標準符合司法官的實際職責 (Turner, 2004)。確定能力需求後，需進一步選擇適當的評估方法，確保評分標準具備高效度和公平性 (胡悅倫等，2009)。

同時，若能在此階段即邀請用人機關參與設計流程，便可依據實務現場對司法官的需求與痛點，釐訂出更契合崗位要求的題目範疇與評估指標 (Kristof-Brown et al., 2002)。透過用人機關的視角，可避免考題內容與真實職責脫節，並更有效檢驗考生是否具備所需之專業能力與實務應變力。

在選擇口試方法時，應建立多層次評估機制，以綜合不同測試方式的優勢，提高選拔的全面性和公正性。核心評估工具應以結構化行為口試為主，透過候選人對過去經歷的描述，評估其應對法律挑戰的能力。這種方式能有效降低主觀偏見，並提高對未來表現的預測準確性。此外，應結合情境口試，模擬司法案件或道德兩難情境，測試候選人在複雜環境下的臨場反應與決策能力，確保選拔出的司法官能夠在實務操作中勝任職責。

在初步篩選階段，應善用人工智慧口試與遠距口試技術，提高評估效率並降低行政負擔。人工智慧口試可快速分析候選人的語言表達、邏輯思維及法律知識，協助篩選不符合標準的申請者，而遠距口試則能讓來自不同地區的優秀候選人參與選拔，避免地域限制帶來的人才流失。

在正式口試階段，應採取小組口試的方式，由多位評審共同評估候選人，確保評估過程更具多元性和公正性 (Avery & McKay, 2006)。小組口試能夠融合不

同評審的專業視角，減少個別評審的偏見影響，使最終遴選結果更具代表性和說服力。此外，這一方式也能讓候選人在不同提問角度下展現多方面能力，提供更全面的表現機會。

（二）口試實施階段

在口試實施階段，需要確保評估過程的標準化、客觀性與公平性，以提升選拔的準確性和公正性。這一階段涵蓋口試官培訓、評估標準制定以及完整的口試流程設計，確保所有候選人都能在相同標準下接受評估，並展現其真實能力與潛力。同時，若在整體實施過程中加強用人機關的參與，便能將第一線的實務需求直接融入考題及評分標準，協助口試官更準確評估考生與司法工作的匹配度，進而避免與實際崗位落差過大。

首先，應加強口試官培訓與評估標準制定，確保評估過程的一致性與專業性。口試官需接受系統化的專業培訓，熟悉司法官的核心職責與考選標準，並學習如何識別候選人真正具備的能力，減少個人偏見對決策的影響。同時，應制定明確的評分標準，如採用行為標準評分量表 (BARS)，將各項評估指標量化，確保評分過程可比較且具備可操作性。這種標準化方式不僅能提高評估的一致性，也能提供具體依據，避免因主觀因素影響考選結果。

其次，應建立分階段的口試流程與評估方式，確保考選過程層層遞進，由基礎篩選到綜合評估，完整考察候選人的專業素養與適任性。在第一階段，應運用人工智慧口試或遠距口試進行初步篩選，確保候選人具備基本法律知識與語言表達能力，提升考選效率並降低行政負擔。這一階段主要篩選出符合基本資格的申請者，為後續深入評估奠定基礎。

進入第二階段，候選人需接受結構化口試與情境測試，進一步評估其法律推理能力與實務應用能力。結構化口試能確保所有候選人接受相同的問題挑戰，提供公平的競爭環境，而情境測試則能模擬司法實務中的決策場景，測試候選人在壓力與不確定情境下的應對能力，確保未來能夠勝任司法官的職責。

在最終的第三階段，應實施小組口試與綜合評估，透過多位口試官共同討論，確保候選人的價值觀與道德標準符合司法官的核心要求。小組口試能夠綜合不同口試官的觀點，降低個別偏見對評估結果的影響，並透過更深入的交流，確保候選人不僅具備法律專業能力，也展現出穩健的司法素養與公正性。綜合評估則應考量所有階段的表現，透過量化數據與專家討論，確保選拔出的司法官候選人能夠符合司法機關的專業與道德標準。

（三）口試後評估與制度改進

在口試後評估與優化階段，應建立完整的數據分析與成效評估機制，確保考選制度的持續改進與適應性。透過對不同方法的效度、公平性與候選人體驗的深入分析，可以確保司法官口試制度隨著法律環境與社會需求的變遷而不斷優化，進一步提升其公正性與科學性。

首先，應透過數據分析與成效評估來驗證口試方法的準確性與公平性。定期檢視不同評估方式的效度，分析各階段的口試結果與候選人後續任職表現的關聯性，確保所選拔的司法官具備優秀的專業能力與決策素養。為進一步提升口試制度的預測效度，應建立口試任用人員之績效追蹤制度，將任用人員的年度績效表現納入評估指標，透過長期追蹤其司法實務表現，分析口試評估與實際任職能力的關聯性，找出潛在的評估偏差或改進空間。此外，應建立長期數據跟蹤機制，對歷年候選人的評估數據進行交叉比對，找出可能的偏差或不足之處，進一步提升考選準確性。除了數據分析，也應透過問卷調查與候選人反饋來評估口試流程的體驗品質。透過收集候選人在口試過程中的感受與建議，可進一步調整口試設計，例如改善題目表達方式、優化評估標準或提升遠距口試的互動性，以確保考選過程既能維持高標準，又能提供公平、合理的競爭環境。

其次，應建立賡續優化與制度改進機制，確保口試內容與標準能夠適應法律與社會環境的變化。司法制度與社會需求不斷發展，考選司法官的標準也應隨之調整，因此應建立動態調整機制，定期檢視考核項目與評估方式，使其與最新的法律改革、社會價值趨勢及司法實務需求保持一致。為強化與用人機關的聯繫，可邀請司法機關參與口試制度改進的論證，結合其對任用人員績效的觀察，確保評估標準與實務需求一致。例如，若司法實務對道德判斷與人權保護的要求提升，口試評估應增加相關情境測試，以確保考選出的司法官能夠適應新時代的司法挑戰。

此外，應積極引入新技術來提升考選效能。例如，利用自然語言處理技術對候選人的回答進行語言分析，檢測其法律邏輯、語意準確性與專業知識表達能力，降低口試官主觀判斷的影響。亦可嘗試引入心理測試，評估候選人的壓力管理能力與道德穩定性，以確保司法官能夠在高壓環境下維持公正與理性決策。透過科技的輔助，考選機制將更趨科學化，減少人為誤差，進一步提升考選標準的客觀性與全面性。

參、研究方法

本研究採用焦點團體訪談 (focus group interview) 方法，蒐集受訪者參與口試的經歷與感想。焦點訪談在公共行政領域是一種富有彈性且能激發多元對話的定性研究法，透過一組具相關經驗和知識的受訪者進行有目的的討論，研究者得以深入了解訪談參與者對制度運作、政策規劃以及實務需求的多元觀點。焦點訪談特別適用於探討制度改進或政策制訂時面臨的複雜議題，能協助研究者掌握不同利益相關者的看法與利益衝突，並進一步形成具體的行動建議。

焦點訪談常被運用於探索性研究、理解多重觀點或評估現行政策成效。本研究聚焦的議題——用人機關如何更有效地參與公務人員考試，涉及人才考選標準、行政效率以及考試公平性等層面，正是一項需要多方深入對話才能釐清的複雜問題。為了貼近司法實務需求，作者邀請司法院法官進行焦點訪談，並針對司法官特考口試制度提出觀察與建議。一方面，法官長期處於法律與實務操作的最前線，能夠對試題設計、評分機制及考生真實能力需求等層面提出專業而實用的意見；另一方面，他們也能回饋對現行口試制度可能產生的盲點，如評量方式過度偏重形式或忽略核心職能等問題。

焦點團體訪談於 2025 年 3 月 14 日舉行，共有四位法官出席，由作者擔任主持人並記錄。

肆、訪談資料分析

根據訪談逐字稿內容，整理分析如下：

一、考慮以核心職能為口試評分標準

在司法官考選制度的探討中，評分標準的設置至關重要。法官 A 指出，現行口試主要聚焦於「應變與溝通能力」等一般性評價項目，較少深入探測司法官的實務核心職能，如法律素養與道德判斷。他認為，口試題目與時間的限制，使得考生很難在短時間內充分展示其專業內涵與法律思維的深度。法官 C 亦強調，司法官作為「說法」的職業，需具備以口語方式讓民眾理解法律精神的能力，這顯示出在評分標準中強調法律專業及論辯技術等核心職能的重要性。

然而，法官 B 的觀察顯示，在實務口試題目未能明確對應這些核心能力時，

考生往往僅展現「補習班教的模組化回答」，這種現象使得口試難以真正評估考生的潛在實力與專業素養。法官 D 更進一步指出，目前的口試分數在整體考試中占比極低，其「汰劣」效果遠大於「擇優」，導致筆試成績成為決定性因素，這種情況下，考生的綜合素質與核心職能難以得到充分評量。

各位法官普遍質疑現行評分標準的有效性，認為其無法真正檢測司法官最重要的職能，如邏輯推理、價值判斷、法律論述與溝通協調等能力。若僅以簡單的「儀態」、「溝通」等項目評分，不僅無法全面顯示考生在司法實務中的整體表現，更難以區分出優秀與普通考生。因此，他們一致呼籲在口試設計時，必須重新檢視並明確化核心職能，使評分標準更加具體且貼合實務需求。

具體而言，應透過設計針對性題目與制定行為指標，突顯出個別考生的特點與優勢，減少評分的主觀性與隨意性，避免流於形式化的「制式分數」。法官 C 強調，若不強調法律思維與口語闡述並重的評價方式，便難以真正發揮口試在挑選人才上的功效，使遴選出的司法官真正符合職業要求與社會期待。

法官 A 與 B 皆提到，補習班的影響使得考生在普遍性項目上的表現幾近一致，反而忽略了更深層的法律或道德素養的展現。因此，將評分重心轉向司法官應具備之核心職能，是目前口試改革的重要課題，這不僅有助於提升遴選制度的科學性與公正性，更能確保司法官隊伍的專業素質與職業能力，以應對日益複雜的司法實務需求與社會法治建設的挑戰。

二、考生分組問題

在司法官口試的實施過程中，考生分組方式的合理性與公平性是確保考選品質的關鍵因素之一。法官 B 明確指出，現行的分組方式存在明顯疑慮，容易導致組內考生之間的「互動與拉抬」現象。法官 B 觀察到，筆試成績較低、處於危險區的考生往往在口試現場搶先發言，試圖以積極表現彌補筆試分數的不足；而筆試成績較高、落於「安全區」的考生則傾向於低調應對，以減少因口或過度表現而失分的風險。這種策略性行為使得口試現場的表現並不能真實反映考生的實際能力與素質。

再者，考生在得知同組名單後，會私下建立社交媒體群組討論應對策略，這種現象在補習班的介入下尤為突顯。由於補習班掌握了筆試排名資訊，同組考生之間更容易形成默契，協力營造出整體友善的氛圍，以確保每個人的分數都不至於過低。這種策略性合作不僅使得口試的競爭環境變得可能影響公平，也減損了

口試作為人才考選工具的有效性。

法官 A 也注意到分組方式的問題，他指出分組並非完全隨機，這可能導致組內考生的口試行為模式趨於同質化。例如，若一個組內大多是高分考生，他們可能採取保守策略，避免過度表現；而若組內大多是低分考生，則可能選擇積極搶答，以爭取更多得分機會。這種組內動態使得口試官所觀察到的發言積極度並不一定與考生的真實實力成正比，從而影響了評分的準確性與公正性。

法官 D 進一步強調，由於考生在分組前已知曉各自的筆試成績，他們能夠更精準地進行策略合作，這種情況下，口試本應具有的鑑別功能被大大削弱。原本旨在區分考生能力差異的口試，成了考生間的策略博弈，這可能不利於考選出真正優秀的司法官人才。

在探討可能的改進方向時，法官 C 建議採用「完全隨機分組」或隱藏考生的筆試成績順位，以減少考生之間的資訊交流與策略合作。這種方式有助於讓口試更真實地呈現每個考生的個人發揮，避免因組內默契而產生的分數虛高或壓低現象。法官 B 則強調，一旦同組考生互相知曉分數高低，就會透過配合和分工來「拉抬」彼此，使口試分數差距受到壓縮。因此，應儘可能避免考生在分組前獲取彼此的筆試成績資訊。

法官 D 提出，唯有分組方式更加透明且隨機，才能杜絕考生「刻意表現」或「刻意隱藏」的情況，確保口試過程的公平與公正。法官 A 雖懷疑在現行制度下要完全杜絕資訊洩漏並不容易，但也建議至少讓組內人數更加多元，避免太多同分數區間的考生集中在一起，以減少組內行為的同質化傾向。

綜合各位法官的看法，分組方式確實是可能影響口試公平的關鍵因素。現行的分組方式容易導致考生的策略性行為與不公平競爭，從而影響口試的鑑別功能。因此，亟需對分組方式進行改革，採用更加科學、公正的分組方法，以減少策略式操作的空間，維持口試的鑑別作用。

三、口試題目應區分業務類型，避免過於專業

在司法官口試的實施過程中，口試題目的設計與分類對於準確評估考生的專業素養與綜合能力至關重要。法官 A 指出，目前的口試題目數量有限，且多偏向一般事務或基本法律情境，較少能深入測試考生的法律專業素養。他認為，由於口試時間短暫且考生處於高度緊張狀態下，「過於專業」的法律題目無法實際考出深層思維；然而若題目過度淺顯，又不足以區分考生真實的法律功底，這

使得現行題目在深度與廣度上存在一定的局限性。

法官 C 則強調，出身專業法官、檢察官背景的口試委員如果命題過度專精，會讓無相關經驗的考生無從回答，欠缺公平性；但如果題目都是大眾化的法律常識，又容易被補習班「套路化」準備，難以檢驗真實能力。這種情況下，口試的區分度與有效性都受到影響，無法全面、準確地評估考生的專業素養與職業潛力。

法官 B 在參與命題工作的過程中發現，補習班對口試題目的熟悉度很高，若題目與時事或一般性法律常識過於相似，考生便能以「標準化回答」順利應對，失去檢驗實力的意義。她還提到，題目若偏向刑事偵查領域，法院背景不熟悉此領域的委員在提問時難以掌握重點；考生若剛好具備該經驗，就能大幅領先其他人，這進一步凸顯了題目設計與考生背景之間的匹配問題。

法官 D 認為，現行口試題目的制式化不足，讓考生在「應變與溝通」之外，很難展現其他關鍵能力。例如，富有行政法院經驗的法官，如果看到題目只是普通刑事案例，提問時也不容易深入，導致雙方都只流於形式，無法充分展現專業素養與職業能力。

針對這些問題，法官 C 建議應將題目區分為基礎法律常識、情境案例、道德判斷與司法官素養等不同類型，依考生抽題情況進行多面向評估，如此才能兼顧公平與專業深度。法官 A 也呼應此建議，指出現行口試若有題庫輪替機制，並要求命題委員在事先明確列出「評估要點」，或能減少「考生猜題—補習班預演」的模式，提升題目設計的科學性與有效性。

綜合各位法官的看法，現行口試題目可能會單一化或高度專業化，這不僅影響了對司法官核心素養的辨識度，也限制了不同經歷背景考生的公平競爭機會。因此，題目設計宜避免過於偏向某一領域或難度過高，以提升口試對司法官專業素養與綜合能力的全面評量，確保挑選出的錄取人真正具備所需的職業素養與能力，同時兼顧考試的公平性與公正性。

四、評分標準差可能會限制了鑑別力

在司法官口試的評分過程中，評分「標準差」的問題成為影響人才考選品質的重要因素。法官 B 在談及「標準差」問題時指出，口試官為了維持評分的「統計合理範圍」，往往不會給出過高或過低的分數，這種「自我約束」可能導致優秀考生難以脫穎而出。他描述了自己看見某位考生表現非常突出，但因擔心分數

若拉太開有違規定，只能「忍痛」將分數壓在相對安全的區間，這種情況下，考生的真實實力無法得到充分顯現。

法官 D 認同這種狀況，認為現行制度採「口試占整體分數十分之一」的模式，導致口試官普遍傾向於將分數集中在某個區間，以免變成決定錄取與否的關鍵。這種作法使得那些表現中庸的考生和真正具潛力的考生之間的差距難以反映，削弱了口試的鑑別功能。

法官 C 強調，司法官職業需要強調突出能力，尤其在「說法」與「道德堅持」方面，若標準差被嚴格管控，就難以凸顯傑出或非常不適任者。他認為，評分標準差雖然在某種程度上可確保不同組口試官的平均分布一致，但反而壓縮了分數範圍，可能降低鑑別度。

法官 B 還提到，他甚至觀察到考生中「一直搶答者」最後也拿不到高分，因為標準差的考量使口試委員不會給分過高。法官 C 與 D 皆表示，若能將口試配分比例適度提高，或是放寬標準差管控，也許能更真實地呈現考生的差異，避免「高分不高、低分不低」的現象。

法官 A 補充，可以考量在制度上明確賦予口試官「評分自由度」，並同時加強評分理由的書面化，以防範任意打分與無理爭議；如此便能兼顧公平與鑑別功能，真正讓口試分數彰顯出優秀與普通之分。

綜合各位法官的觀點，現行評分標準差的限制在一定程度上壓制了口試官的評分自由度，影響了對考生真實能力的準確評估與區分。為了提升遴選制度的公正性與有效性，有必要對評分標準差的管控進行適度調整，賦予口試官更大的評分自由度，並通過制度設計確保評分過程的透明與公正，讓優秀考生能夠在口試中充分展現其實力，從而提升司法官隊伍的整體素質與專業水平。

五、應讓口委有較多時間閱讀考生基本資料

法官 D 先前提到，口試官往往在到達口試現場時才獲得考生自傳與履歷，時間緊迫導致他們只能快速翻閱，難以深入掌握每位考生的經歷與特質。法官認為，口試官對考生的背景「所知有限」，必須在候選人自我介紹的同時「邊看邊問」。他認為，若要判斷考生是否偽裝，或要評估考生的真實狀態，必須讓口試官事先充裕時間閱讀，並分析所有可得訊息，包括自傳、履歷與可能的其他背景資料。法官 B 則舉例，若是一般企業招聘，往往會讓口試官提早獲得應徵者的社群媒體資訊或職業紀錄，以便綜合觀察應徵者的行為模式。然而，司法官口試

的口試官在短時間內只能檢視制式資料，恐怕無法發掘考生是否「刻意包裝」或「深度準備」。

法官 C 則在這方面提出一項有趣的觀察：補習班為考生「建議」了包含儀態與回答方式的統一訓練，例如女性考生「不可穿裙子」以及坐姿、進場流程、發言句型等，都以一種「軍事化」的模式呈現。這些外在行為若無法對照考生的真實背景，口試官就難以區分誰是真正具備司法官所需素養、誰只是短期訓練後的制式化表現。倘若口試官能在口試前詳閱考生的相關經歷與興趣取向，也許能藉由深層對話，探查到考生更真實的想法與專長，而非僅從外在的儀態或標準化語句評分。

針對上述觀察，法官 A 與法官 D 建議，若能在口試前一週或更早就將考生資料提供給委員，並要求口委簽署保密協議，以確保資料不外流，這將使口委在正式口試時更能「有的放矢」。法官 B 強調，若僅在口試當天匆匆翻閱，口試官不但難以洞悉考生真實能力，甚至可能被「一致化」的外在表現所誤導。此外，提前閱覽資料也能引導口試官提問更具針對性，包括追問考生為何選擇司法官職涯、如何理解司法官倫理與公共利益等深層議題，增進評分的客觀性與差異性。

因此，在強調「核心職能」評分的當下，「口委有更多時間閱讀考生資料」是維持評分公平與準確的重要手段。唯有先充分掌握考生背景，口試官才能在短暫口試中對照其回答與履歷，判斷表現是否屬虛飾或真才實學，並衡量其是否具備擔任司法官所需的人格特質與專業素養。

六、口委評分一致性與獨立性

在司法官口試的評分過程中，確保口試官評分的一致性與獨立性是保障考選公正性的關鍵要素。法官 B 在談到實際評分時提及，口試官之間常有「彼此觀察」的心理，擔心給分過高或過低與其他委員落差太大，被認為不符常規。這種心理狀態導致口試官在評分時趨於保守，不敢大膽地根據考生的真實表現給出極端分數。

法官 A 指出，計算每組口試官的平均分差，若誰分數給得太「出格」，就可能受到質疑。這種制度設計間接影響了口試官評分的獨立性，無法完全依個人判斷給分，而是要考慮到整體的平均分差，這在一定程度上限制了口試官的評分自由度。

法官 C 表示，若每位口試官都採取「安全區」打分策略，不但限縮了表現優秀者的空間，讓真正有實力的考生無法獲得應有的高分，也讓表現不足者獲得過高分數，從而削弱了口試的區分度，使得遴選結果無法準確反映考生的真實能力水平。

對此，法官 D 建議，在制度設計上應給予口試官更大的「評分自由度」。例如，只要口試官能提供合理解釋，就可適度擴大高低分的區間，讓評分更貼近實際表現，這樣才能真正體現考生之間的差異，提升考選的公正性與有效性。

法官 B 強調，若要維持一致性，仍需保有一定的標準化參考，例如採用「行為標準評分量表」。一旦各口試官對行為具體指標有明確理解，彼此給分自然不會差異過大，又能保留個別判斷，這樣既確保了一致性，又維護了獨立性。

至於如何確保口試官獨立性，法官 A 建議，可以先讓每位口試官在看完考生回答後「獨立記分」，再經由系統統計出平均與分差，而不需進行「現場會商」，則可避免口試官間的相互影響，確保評分的獨立性。

法官 C 則認為，若口試官之間能在事前充分討論「司法官核心能力」與「評估邏輯」，即使最後給分不同，也能比較容易理解彼此的依據，不會產生不必要的懷疑或壓力，從而兼顧公平與彈性。

綜合各位法官的看法，一致性與獨立性並非互斥，唯有在標準化的基礎上保留口試官個人判斷，才可同時實現口試的公平與彈性，確保挑選出的錄取人真正具備所需的專業素養與職業能力，維護司法公正與法治精神。

七、集體口試與個別口試

與會法官也就集體口試與個別口試提出討論。法官 B 指出，在現行司法官口試中，集體口試往往呈現出「同組考生互相觀察與配合」的現象，筆試分數較危險的考生積極搶答，而分數較安全的則選擇沉默，藉此維持安全排名。這種策略性行為使得小組口試不易有效區分真實能力與策略性表現，影響了考選的公正性與有效性。

法官 A 認為，若組內同質性高，或皆知彼此分數差距，集體口試就會傾向「禮讓」或「團體作戰」，使得最終評分可能與個人實力不全相符。這種情況下，集體口試的優勢無法充分發揮，反而可能掩蓋考生的真實水準。

相較之下，法官 D 分享了司法院「律師轉任法官口試」的「個別口試」經驗。若有充裕時間檢視考生資料，再逐一進行訪談，可深入挖掘考生的動機與應

變能力，不至於受其他考生干擾。個別口試能提供更多針對性問題，讓具備特殊背景的考生有機會充分發揮，從而更全面地評估考生的專業素養與綜合能力。

然而，法官 C 提醒，個別口試雖能深入，但若口試官團隊過大、時間壓力過高，也可能造成評分不一致或資訊過載；而集體口試若運作恰當，也能觀察考生在群體互動下的態度，對未來在院檢實習的合作精神有指標意涵。因此，個別口試與集體口試各有利弊，關鍵在於如何合理整合與運用。

基於此，法官 B 建議，可結合兩種形式：先在第一階段透過集體口試，觀察考生的互動與臨場反應，再在第二階段進行個別口試，以精準針對每位考生的特質與背景。法官 A 也贊同此作法，認為如此能兼顧「群體互動」和「個人深度」兩方面優勢，避免只看到集體行為而忽略個人能力。法官 C 則補充，若能適度調整考試時程，將集體與個別口試做合理分配，就能在不增加太多成本的情況下，更全面地評估考生。

整體而言，各位法官對小組口試與個別口試各有所見。唯有整合多種方式並兼顧配分與時間安排，才能更有效檢驗司法官所需的複合能力，確保遴選出的司法官真正具備所需的專業素養與職業能力，維護司法公正與法治精神。

八、考生對口試公平性的感受

與會法官指出，現行口試因補習班高度介入，考生往往出現「制式化」的進退儀態與回答模式，難以看出個人真實特質。法官 B 強調「若口試官對考生背景了解不足」會放大這種問題：當所有人外在行為幾乎一致、語句也雷同，儀態評分就幾乎拉不開差距。更甚者，若口委想藉由臨場互動區分誰在偽裝，卻因準備時間與資訊不足，無從深究考生的動機與表達背後的含義。在這樣的情況下，考生可能會質疑口試能否真正區辨優秀與平庸，認為只要按補習班教的進行就不至於失分。

法官 C 則表示，偽裝在高度競爭的考試環境下「很難完全避免」。他提到，一些制式化的動作其實根源於過往軍事化訓練或補習班錯誤的引導，而這些舉止並非司法官工作的真實 DNA。對考生而言，受到補習班的引導是「安全」策略，畢竟口試占整體考試比例有限，能在短時間內以標準化表現獲得不錯的分數。若考生因此感到制度「默許」這種形式，便很難說服他們相信口試具有深層鑑別力。法官 C 進一步提到，自己在提問時會儘量「回歸核心職能」，例如要求考生解釋法律決策的邏輯或闡述面對各種案件的應對態度，以便拉開考生間的

真實差異；但若沒有明確機制鼓勵考委都這麼做，考生恐怕仍覺得口試流於形式。

此外，法官 B 指出，評分時雖有「標準差不宜太大」的建議，但實務上未必每位委員都嚴守規定，可能造成某些組分數的離散度與其他組不一致的情形。法官 D 則從「汰劣」與「擇優」的辯證觀點提醒，當考生發現口試主要是淘汰低分者，而非真正用來區分菁英，他們自然更可能懷疑口試的公平與價值。

法官 B 與法官 C 認為，若能加強口試官前置準備、增進對考生背景的了解，並鼓勵考委在提問中著重司法官核心職能，將能降低偽裝空間，增強考生信任。而法官 A 與法官 D 皆建議提高口試在總分中的比重，同時放寬標準差限制，使優秀考生能透過表現脫穎而出，才更能讓考生相信「口試真的會改變錄取命運」，最終也能促成司法官口試的實質改革與公信力提升。

九、擴大用人機關參與的具體作法

以下針對司法官特考口試提出具體建議，期能提升用人機關參與的深度與廣度，確保口試更能反映實務需求與選才目標。

（一）提高用人機關代表的口委人數比例

現行作法已根據修訂後的《典試法施行細則》，將用人機關代表在口試委員中的比例由不超過三分之一提升至二分之一，顯示考選部對用人機關意見已加強重視（江宗正，2024）。

為使口試更具實務對應性，可考慮對特定專業類科（如刑事法律、行政訴訟、家事少年法庭等）進一步增加用人機關代表，甚至可占三分之二以上委員席次。如此一來，第一線司法實務者有更多發言空間，能在甄選過程中提升命題、評分和監督品質，強化與職務需求的契合度。

（二）考慮由用人機關主導口試評分權重

目前司法官特考口試評分，用人機關的實質影響相對有限。若能推動「雙層評分機制」，由考選部和學者專家先針對通用能力進行第一層結構化評分；接著，用人機關可針對與職務高度相關的核心要素（如司法實務能力、道德判斷）進行第二層加權評分，能凸顯考生在面對複雜案件或壓力時的應對能力。如此一來，不僅維持考試的基礎一致性，也能讓用人機關更有效發揮篩選功能。

（三）建立口試試題檢討與回饋機制

現行制度已有嚴謹的命題與審題機制，仍建議擴大用人機關在試題方面的參與機制。例如：

- 召開「口試試題檢討工作坊」：由考選部、用人機關代表、命題委員（口試委員）共同檢視擬定題目的重點是否呼應核心能力與當前司法業務動態；並可進而討論試題的（預測）效度，做為未來命題參考。
- 強化口試官培訓：透過講習或工作坊，增進口試委員對提問方式、行為觀察與紀錄、評分標準、評分技巧等方面的瞭解，提高評分信效度。

（四）考慮增加「情境模擬測驗」

由於司法官工作常需面對複雜案例與突發情境，故建議在傳統口試以外，加入更動態的「情境模擬」內容，並讓用人機關實務代表直接觀察與評分。情境可包含：模擬法庭辯論、審判程序應對、刑事搜查現場等，考核考生如何在真實壓力下展現法律邏輯與應變能力。用人機關透過此環節能更精準判斷考生專業度，避免僅憑口語流暢度或背誦式知識便高分錄取的情況。

伍、結論

本研究以司法官特考為例，探討了用人機關如何在口試環節中扮演更積極且有效的角色。透過針對司法院法官所進行的焦點小組座談，瞭解現行司法官特考在口試制度與命題設計上猶可繼續精進的諸多議題，包括效度提升、提高公平性、口試方法多樣性，以及如何補救業者對考生表現的影響等。針對這些議題加以改進有助於讓考生的實力充分展現，使得用人機關順利選拔到真正符合實務需求的人才。

在歸納相關文獻與訪談資料後，研究進一步針對「擴大用人機關參與司法官特考口試」提出了四項建議，分別為：（1）提高用人機關代表的口委人數比例；（2）考慮由用人機關主導口試評分權重；（3）建立口試試題檢討與回饋機制；（4）考慮增加「情境模擬測驗」。

上述四項具體建議能讓口試內容與評分機制更貼近實際職務需求，並兼顧考試的公平與科學性。其中，包括在評分層面增加用人機關對特定核心職能的加權，或是在題目檢討與回饋階段納入實務者的深度意見，可加強命題與實務的連結；同時，若能在傳統口試之外導入「情境模擬測驗」，也有助於讓考生面對複

雜案件與突發情境的臨場應變能力。對於專業度要求高、工作情境多變的司法官職位而言，這些改革措施有助提昇人才考選的精準度，讓真正具備核心能力與實務應變力的考生能在口試環節中脫穎而出。若能透過這四項作法持續推動並滾動檢討，不僅可確保口試隨司法環境與社會變動而動態調整，也能加強考選部與用人機關間的溝通與合作，進一步提升整體司法人員考選的效率與品質。

最後，鑒於司法官特考在於考選未來的法官與檢察官，惟本研究受限於時間與資源，僅以法官為研究對象；未來除應擴大研究對象，並須將檢察系統意見納入，方能對司法官口試做全面、整體的探討，尋求周延的精進方策。未來研究應將考生納入訪談對象，了解其在口試過程中的體驗與建議，以進一步優化口試設計與公平性。同時，應深入探討補教業者對口試表現的影響。例如，Fletcher (2013) 的理論解釋了補習班訓練如何透過標準化的印象管理策略（如預先準備的應答模板或刻意展現特定特質）導致「Faking 效果」，這可能削弱口試對考生真實能力的評估 (Fletcher, 2013)。進一步地，Roulin 等 (2015) 指出，補習班可能助長欺騙性印象管理行為（如誇大能力或虛構經歷），而此類行為因考生準備充分而難以被面試官察覺，從而在評分中獲得不當優勢。為降低此類影響，未來可採用結構化面試、行為導向問題或情境模擬測驗，確保評估聚焦於與司法官核心職能相關的真實能力，而非印象管理技巧。此外，應善用現有資料，採用更多元的研究方法，例如進行口試考題的內容分析、口試成績的統計分析，以及口試成績與筆試成績的比較分析，以更全面地評估口試制度的效度與改進方向。

參考文獻

江宗正（2024）。我國國家考試口試實施現況及未來精進策略。**國家菁英**，17(1)，23-43。

考選部（2023，6 月 15 日）。日本公務人員考試基本能力測驗編製與面試考察報告。考試院第 13 屆第 142 次會議考選部重要業務報告。

余民寧（2013）。口試在國家考試應用之再檢討與改進。**國家菁英季刊**，9（2），87-107。

吳新興（2023）。從日本重視面試機能思考台灣考選改革〔Rethinking Taiwan's examination reform from Japan's emphasis on interview mechanisms〕。**國家考試院**。

https://www.exam.gov.tw/NHRF/News_EpaperContent.aspx?n=3778&s=46726

&type=3F111E57648FF61F

吳新興、劉建忻、董郢（2023）。考試院 111 年度日本中央及地方政府公務人員考選制度考察報告。考試院。

李俊達（2021）。英國一般文官考選制度：機關分權、公開競爭、測驗效度。國家人力資源論壇，12。

https://www.exam.gov.tw/NHRF/News_EpaperContent.aspx?n=3778&s=44694&type=5654361B474B8EFF

胡悅倫、陳皎眉、洪光宗（2009）。國家考試口試之命題與評分。國家菁英季刊，5（4），35-56。

馮文濤（2024）。談公務員面試作答的三種“層次”。應用寫作，06，63-64。

顧愛華、黃一峯（2014）。論大陸公務員面試制度。國家菁英季刊，8（4），139-148。

Andrews, L., & Bucher, H. (2022). Automating discrimination: AI hiring practices and gender inequality. *Cardozo L. Rev.*, 44, 145.

Avery, D. R., & McKay, P. F. (2006). Target practice: An organizational impression management approach to attracting minority and female job applicants. *Personnel Psychology*, 59(1), 157-187.

Barrick, M. R., Swider, B. W., & Stewart, G. L. (2010). Initial evaluations in the interview: Relationships with subsequent interviewer evaluations and employment offers. *Journal of Applied Psychology*, 95(6), 1163-1172.

Chapman, D. S., & Zweig, D. I. (2005). Developing a nomological network for interview structure: Antecedents and consequences of the structured selection interview. *Personnel Psychology*, 58(3), 673-702.

Chen, C. C., Lee, Y. H., Huang, T. C., & Ko, S. F. (2017). Effects of stress interviews on selection/recruitment function of employment interviews. *Asia Pacific Journal of Human Resources*, 57(1), 40-56.

Chen, J. C. C., Chiang, M. H., & Chen, T. M. (2013). Reinventing strategies of oral examination in Taiwan civil service recruitment. *Public Personnel Management*, 42(2), 293-309.

Ellis, A. P., West, B. J., Ryan, A. M., & DeShon, R. P. (2002). The use of impression management tactics in structured interviews: A function of question type? *Journal of Applied Psychology*, 87(6), 1200.

- Fletcher, C. (2013). Impression management in the interview. In *Impression management in the organization* (pp. 269-282). Psychology Press.
- Grindle, M. S., & Hilderbrand, M. E. (1995). Building sustainable capacity in the public sector: What can be done?. *Public Administration and Development*, 15(5), 441-463.
- Huffcutt, A. I., & Arthur, W. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, 79(2), 184-190.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis* (2nd ed.). Newbury Park, CA: Sage.
- Kristof-Brown, A., Barrick, M. R., & Franke, M. (2002). Applicant impression management: Dispositional influences and consequences for recruiter perceptions of fit and similarity. *Journal of Management*, 28(1), 27-46.
- Lievens, F. (2007). Tests de juicio situacional: Introducción y revisión de la investigación. *Revista de Psicología del Trabajo y las Organizaciones*, 23, 93-110.
- Lievens, F., & De Paepe, A. (2004). An empirical investigation of interviewer-related factors that discourage the use of high structure interviews. *Journal of Organizational Behavior*, 25(1), 29-46.
- Moscato, S. (2000). Selection interview: A review of validity evidence, adverse impact and applicant reactions. *International Journal of Selection and Assessment*, 8, 237-247.
- 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(5), 571.
- Powell, N. J., & Levine, H. (1942). Reliability of the Civil Service Oral Examination. *The American Journal of Psychology*, 55(3), 385-393.
- Roulin, N., Bangerter, A., & Levashina, J. (2015). Honest and deceptive impression management in the employment interview: Can it be detected and how does it impact evaluations?. *Personnel Psychology*, 68(2), 395-444.
- Sáez, J. (2007). Diseño y validación de una Entrevista Conductual Estructurada para la selección de agentes de policía local. *Revista de Psicología del Trabajo y las*

Organizaciones, 23, 57-74.

- Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, 76, 323-346.
- Salgado, J. F. (2007). La utilidad económica de la entrevista conductual estructurada en la selección de personal de la administración general del País Vasco. *Revista de Psicología del Trabajo y las Organizaciones*, 23, 139-154.
- Salgado, J. F., & Moscoso, S. (2002). Comprehensive meta-analysis of the construct validity of the employment interview. *European Journal of Work and Organizational Psychology*, 11, 299-324.
- Salgado, J. F., & Moscoso, S. (2005). *La entrevista conductual estructurada de selección de personal*. Madrid: Pirámide.
- Salgado, J. F., & Moscoso, S. (2008). Personnel selection in industry and public administration: From the traditional view to the strategic view. *Papeles del Psicólogo*, 29(1), 16-24.
- Salgado, J. F., Gorriti, M., & Moscoso, S. (2007). La entrevista conductual estructurada y el desempeño laboral en la administración pública española: Propiedades psicométricas y reacciones de justicia. *Revista de Psicología del Trabajo y las Organizaciones*, 23, 39-55.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Turner, T. S. (2004). *Behavioral interview guide: A practical, structured approach for conducting effective selection interviews*. Bloomington: Trafford Publishing.
- Van Iddekinge, C. H., Lievens, F., & Sackett, P. R. (2023). Personnel selection: A review of ways to maximize validity, diversity, and the applicant experience. *Personnel Psychology*, 76(2), 651-686.