

## 公務人員考試加入性向測驗的可行性評估

陳柏熹\*

### 摘要

本文主要是以多元觀點來探討公務人員考試的選才方式，並分析在公務人員考試中若加入性向測驗作為選才參考依據之一，需要考量哪些因素。測量多元潛能的性向測驗，早在第一次世界大戰前就普遍被用在軍事及工業中的人才甄選與職涯諮商，不論是統計方法的演進，還是智力理論的取向轉變，乃至於學校的入學選才，都為多元潛能在公務人員考試的應用奠定了良好的基礎，而性向測驗的主要目的正是用來評估應試者的多元潛能。考選部近二十年來陸續對於將不同考試方式納入公務人員考試的選才中作了法規制度的改變、評估研究案，以及試題範例的研發，但真正要將不同測驗方式落實公務人員考試中，還需要從人、事、時、地、物等五個層面作仔細的考量與及規劃。本文從這些層面進行了現況分析，以作為未來公務人員考試要將性向測驗納入公務人員選才時的政策擬訂及實務規劃參考。

關鍵詞：性向測驗、公務人員考試、可行性評估

---

\*國立臺灣師範大學教育心理與輔導學系教授

## Feasibility Evaluation of Aptitude Testing in Civil Service Examinations

Po-Hsi Chen\*

### Abstract

This article investigates the civil service examination from multiple perspectives. The author analyzes related factors and considerations when applying the aptitude testing in civil service examination. Aptitude testing has been used in military and industry for personnel and employee selection since the First World War. Progress in statistical methods, the developing trend in intelligence theories and the tendency of the college entrance system lay the foundation for using the multiple latent traits in civil service examinations. Aptitude testing is the best tool for evaluating multiple talent. The Ministry of Examination (MOEX) has carried out several projects applying different psychological tests in civil service examinations. To realize the ideal of using aptitude testing in civil service examinations, the author suggests five factors for practical consideration: human, administrative, time, space, and tools. Evaluation of these factors can be for reference in policy making and practical planning for using aptitude testing in civil service examinations.

Keywords: aptitude testing, civil service examination, feasibility evaluation

---

\*Professor, Department of Educational Psychology and Counseling, National Taiwan Normal University

## 壹、緒論

自第一套標準化心理測驗比西智力量表 (Binet & Simon, 1905) 發展至今，已有超過百年歷史，雖然此測驗提出了測量一般心智能力的標準程序，但由於它主要是測量語文與數字方面能力，因此心理學家們很快就發現其侷限性，進而發展出特殊性向測驗 (special aptitude test) 來彌補一般智力測驗的不足，以評估受測者的多元潛能，並符合工業及軍事用途中大規模招募不同類型人才的需求。

### 一、多元潛能的發展趨勢與多元性向測驗

性向測驗屬於最大表現型測驗 (maximum performance test)，是一種有標準答案或對錯之分的能力測驗，是用來了解個人未來發展潛能的測驗工具 (Anastasi & Urbina, 1997)。測量多元潛能的性向測驗是指在一份測驗中包含許多測量不同特殊性向的試題，又可稱為多元性向測驗 (multiple aptitude batteries)。多元性向測驗的發展，主要與下列幾項趨勢有關：

#### (一) 智力 (intelligence) 理論的演變

智力的定義與測量在心理學上一直引起熱烈的討論，早在 1921 年教育心理學期刊 (Journal of Educational Psychology) 就發行專刊邀請多位學者對智力與智力的測量進行探討，雖然史比智力量表 (Stanford-Binet Intelligence Scale) 提出智商 (Terman, 1916) 的概念，以心智年齡與生理年齡的比值來簡化大家對智力的觀點，但仍有許多心理學家對於只用單一指標來代表人類的潛能不以為然，並提出不同的觀點。例如：Spearman (1927) 的二因素理論就認為各種心智活動所需的能力不只有一般因素 (general factor)，還有特殊因素 (specific factor)。人類在這些測驗或任務 (task) 上的表現好壞，是同時受到這兩類因素所影響，題目類型或作業活動的差異愈大，一般因素的影響就愈小，特殊因素的影響就愈大。

由於 Spearman 並沒有對特殊因素的類型做更細緻的說明，產生了多因素理論的需求，Thurstone (1938) 正是多因素理論的代表，以 Thurstone 與 Thurstone (1941) 提出的智力理論為例，他們認為人類的心智能力包含語文理解、字彙流暢、數字、空間、聯想記憶、知覺速度、歸納或推理等。後續更有學者認為應該從作業活動的內容、認知運作方式與產物三個向度的結構，來看待人類的心智能力，也就是 Guilford 的智力結構模型 (Guilford, 1988)。

當代較著名的多元潛能觀點首推 Gardner (1983) 的多元智慧理論，他基於生物演化與適應社會功能的觀點，認為每個人應該都具備了八種不同的天生能力，只是某些能力較具優勢，有些能力則較不具優勢。這八項能力分別是：語文智慧、視覺空間智慧、音樂節奏智慧、肢體動覺智慧、邏輯數學智慧、自然觀察者智慧、人際智慧、內省智慧等。至今，我們仍可在臺灣的多元入學制度中看到 Gardner 理論的影子，而 Gardner 目前也在政府與企業支助下發展了零計畫 (Project Zero, Gardner et al., 1988)，以評鑑中心法 (assessment center) 觀察及記錄多元智慧的發

展，並開發培訓課程。

上述這些智力理論的演變，都支持著一個觀點，就是單一分數或單一能力不足以解釋人類在不同作業活動或測驗結果上的表現，應該要以多元潛能、多元的評量方式來看待人類的能力。

## （二）因素分析 (factor analysis) 的發展與應用

因素分析的概念最早由 Spearman (1927) 所提出，初期是以尋找共同因素為目的，後來由 Thurstone (1931) 提出多元因素的分析程序。因素分析對於瞭解測驗建構或作業活動所需的能力有相當重要的貢獻，也因此產生了上述智力理論中所提到的許多不同類型的多元潛能。

多元性向測驗就是用來評量受測者在各種不同方面的潛能，讓受測者的測驗結果不再只以單一智商的分數來呈現，而是改用剖面圖 (profile diagram) 的形式來呈現受測者在語文、數字、空間、知覺速度……等多個方面的能力。這些以因素分析為基礎的測驗研究，當時多集中在軍事用途中，尤其是在第二次世界大戰期間，為了招募具備不同特殊才能的軍人，因此大量使用了多元性向測驗，以彌補原本使用單一能力測驗（例如：Army Alpha, Army Beta）的不足，此趨勢造就出著名的軍職性向測驗 (Armed Services Vocational Aptitude Battery, ASVAB)，ASVAB 至今仍是美國軍隊招募人才最主要的測驗，該測驗目前可以做出五十種以上的量尺組合，以因應各種不同軍種的招募需求 (ASVAB, 2023)。

## （三）學校入學選才的趨勢

最早的比西智力測驗在法國是被用在學校入學前篩選心智遲緩兒童之用，並非用來評估學習成效，被用在評估學生學習成效的標準化測驗是 1923 年出版的史丹佛成就測驗 (Stanford Achievement Test)，該測驗使用單一常模群體來呈現出不同學校受測者的學習成就。隨著後來美國大學入學委員會 (College Entrance Examination Board, CEEB) 以及教育測驗服務社 (Educational Testing Service, ETS) 的成立，成就測驗也漸漸朝向以測量學生學習潛能的學術性向測驗去發展，著名的學術性向測驗 (Scholastic Aptitude Test, SAT) 就是在這樣的潮流下發展出來的。而 SAT 的演變也與上述智力理論的轉變趨勢一樣，從剛開始 1926 年只有單一分數，到 1931 年轉變成提供語文、數學兩項分數，再到 1993 年轉變成以推理能力為主的 SAT I 以及以學科為主的 SAT II，而一般所指的 SAT 就是 SAT I，因為它才是測量未來學習潛能的學術性向測驗，而非成就測驗。

性向測驗與成就測驗最大的不同點在於，性向測驗是在學習前測量，主要是評估受測者具備哪些學習能力或執行作業活動的潛能，以作為未來選擇學術生涯或職業生涯的參考，而成就測驗是在學生學習之後才進行測量，主要是了解受測者對於各項課程內容的學習成效。性向測驗又可分為學術性向測驗與職業性向測驗兩大類，學術性向測驗偏重在語文、數字與邏輯推理等方面，這些能力與大部分的課程學習有關，職業性向測驗則範圍較廣，除了測量前述特質外，通常也包含與實作技能有關的知覺速度、機械性向，視覺空間……等與職業活動較有關連的潛能。

臺灣的大學入學制度自 1954 年起也是採用單一入學管道單一分數，也就是

大學聯考制度，到 2001 年起轉變成目前多項管道採納多種表現的多元入學制度，而且各校可以針對每一種管道及採計項目設定權重。時至今日，大學入學管道主要包含繁星推薦（約 5~10%）、甄選申請入學（約 40~50%）、考試分發入學（約 40~50%），這三個管道除了考試分發入學只採計考試成績外，其餘兩個管道都會採計綜合學習表現 (portfolio)，包含學習歷程檔案（課程學習成果、多元學習表現……等）及自辦甄試項目（面試、實作，筆試……等）（李大偉等人，2012；葉玉玲、丁一顧，2021）。

學生入學後，各大學欲培養的學生核心能力或基本素養也是多元潛能取向，陳柏熹等人（2015）曾彙整臺灣十多所公私立大學公告的學生核心能力及基本素養，並參考 21 世紀技能評量與教學計畫 (Assessment & Teaching of 21st Century Skills [ATC21S], 2010) 提出的 21 世紀學生必備十項基本素養，彙整出共同的九項大學生基本素養，分別為：溝通合作、美感素養、科學思辨、資訊素養、終身學習、創新領導、問題解決、公民社會、生涯發展，並發展出大學生基本素養測驗，以不同題型測量此九項素養的認知能力與情意態度表現。從上述的轉變也可以看出，單一測驗總分的概念已經不符合適性發展的教育理念了，大學各科系還是希望能了解學生是否具備進入該科系學習所需的多元能力，並培養學生未來進入社會就業所需的核心理念。

#### （四）公務機構與企業的選才趨勢

世界各國將性向測驗列為公務人員選才項目的例子不少，例如：在亞洲方面，新加坡的公共服務署或金融管理局皆將心理測驗（包含性向測驗）列為甄選公務人員的評估項目之一，日本公務人員考試常用的內田－克雷佩林心理測驗 (Uchida-Kraepelin psychodiagnostic test) 也是一種測量運算能力與專注力的性向測驗。在歐美地區，美國聯邦政府的人事管理局 (Office of Personnel Management [OPM], 2019) 使用智力測驗與性向測驗來協助甄選聯邦文官，各州政府的警察考試也採用人格測驗、性向測驗與智力測驗作為甄選、淘汰或升遷的評量工具，英國、德國等歐洲國家也是如此（余民寧等人，2016；施能傑等人，2009；彭錦鵬等人，2017）。

在臺灣的公務人員考試中，民航特考之飛航管制人員科別已將性向測驗納入飛航管制人員體格複檢項目中，內容包含「動作判斷能力」、「認知及推理能力」及「空間能力」等（考選部，2020），也有許多學者專家建議應逐步將性向測驗與性格測驗納入部分公務人員考試的項目中，如「外交部外交領事人員」或「內政部警察人員」考試等（伍姿蓉，2012；余民寧，2015；李登科，2012）。

企業界也有使用性向測驗作為人力資源甄選或派任的例子（段美玉，2006）。例如：證券交易基金會辦理的證券交易所人員招募或集保結算所人員招募，都會使用職務適性評量，包含性向測驗及性格測驗兩大類試題。又例如臺北捷運局的人才招募，除了部分類科在第一試的專業科目就會測量數理邏輯的性向外，第二試也以口試或心理測驗來評估應試者的多元潛能與人格特質。

總和上述各項發展趨勢，多元潛能的評估幾乎已經是各領域人才甄選的主要考量，此趨勢包含了三個特色：其一，評估的面向從單一面向到多元面向，使用

的分數類型也從加總的單一分數轉變成有多個面向的剖面圖。其二，測驗的重點從強調過去學習表現的成就測驗，擴展到強調未來發展潛能的性向測驗。其三，評量的方式從過去較著重在紙筆測驗，漸漸加入非紙筆的實作評量、口試、電腦化測驗等不同類型的評量方式。這些趨勢對於我國採用性向測驗到公務人員考試中有相當大的參考價值。

## 貳、公務人員考試加入性向測驗的可行性評估

### 一、法規及評估研究的分析

將多元潛能的觀點納入公務機構的人才甄選上已有相當久的歷史，從 1988 年制定典試法時，就已經在第 19 條中將「審查著作或發明或論文，或口試，或實地考試」納入可能的測驗方式中，但較少實際採行。到了 2015 年修訂時，在第 14 條正式將「考試得採筆試、口試、心理測驗、體能測驗、實地測驗、審查著作或發明……」列入考試方式中，至於心理測驗或其他測驗方式的實施細則，考選部則於其他規則中另訂。

除了法規面的調整，政府部門對於將不同測驗方式納入公務人才選拔或評核也進行過研究評估，例如：公務人力發展中心曾於 2016 年委託學者專家來進行公部門運用心理測驗及其法律背景之研究(余民寧等人, 2016)，考選部也在 2017 年執行了公務人員考試採心理測驗之可行性研究(彭錦鵬等人, 2017)，這兩份近期的研究都建議應逐步針對臺灣的部分公務人員甄選或安置使用心理測驗(含性向測驗)。而在 2019 年，考選部也委託了心理測驗專家實際研發了公務人員性向測驗範例，並提出分數使用的評估與模擬研究(陳柏熹等人, 2020a)。然而，目前除了民航特考之飛航管制人員科別以及外交部駐外人員的升遷考試外，其他公務人員皆尚未採用性向測驗，殊為可惜。以下摘要這些法規及評估研究的重點。

#### (一) 法規制度的改變

法規制度是公務人員依法行政的最重要依據，若要將性向測驗納入公務人員的甄選中，修法是第一要務。如前所述，考試院於 2015 年已修改典試法將心理測驗列入考試方式中，而公務人員考試法第 10 條也將心理測驗納入考試方式中。至於心理測驗的實施方式，考選部於 2015 年 12 月 14 日另外發布了心理測驗規則，其中第 3 條明訂「各種考試得視需要採行心理測驗，以文字、數字、符號、圖形或操作等方式，評量應考人之智力、性向、人格、態度及興趣等心理特質。心理測驗種類如下：一、智力測驗。二、性向測驗。三、人格測驗。四、興趣測驗。五、其他心理測驗。前項心理測驗種類，得視考試性質選定其中一類或數類舉行。」因此，性向測驗納入公務人員考試中早有法源依據。從上述文字中亦可看出，現行的考試法規是將成就測驗與心理測驗區分開來，這樣的觀點與心理學領域中將成就測驗視為心理測驗的一種有所不同(Anastasi & Urbina, 1997)。

該規則第 4 條中又提到「採行心理測驗時，其種類、評分、參考標準或與其他考試方式分別占總成績之比例或其他相關事宜，於各該考試規則中訂定。」也

就是說，只要事先將性向測驗分數的計分參考標準及分數使用方式在考試前訂定清楚，並事先公告，實施性向測驗是符合法律規範的。也因此，有多位考試委員曾建議要將心理測驗納入公務人員考試中，例如：陳皎眉等人（2011）曾建議在公務人員考試的甄選上，應加入具有預測效用的心理測驗，以達預測工作績效的人才甄選目的。而 2015 年 8 月 13 日通過的「公務人力再造策略方案」，也將心理測驗列入策略議題「參、強化心理測驗與多元考試機制」之「研究發展心理測驗等多元考試方式」案，並配合規劃研議。

## （二）可行性評估研究

在評估研究方面，余民寧等人（2016）曾受公務人力發展中心委託進行公部門運用心理測驗及其法律背景之研究，主要目的是分析當時各國公務機構使用心理測驗的情形、法律背景，以及使用後的成效。該研究在瀏覽歐美及亞洲先進國家公務部門使用心理測驗的概況及其法律規範後，提出心理測驗導入公務機構之自我探索模式、職務分派模式與多階段複合功能模式，並進行焦點團體座談。該研究也針對心理測驗實施結果的運用、心理測驗導入實施時機、心理測驗實施法制化，以及心理測驗之資訊公告四個面向進行大規模問卷調查。結果顯示，受訪者與學者專家對於公務機構使用心理測驗普遍認同度高，但認為尚缺乏配套，以致存在些許待解議題。該研究建議未來在公務機構中可以先使用在低風險用途上、公開評量工具的內容品質，以及分階段漸進式實施心理測驗。

此外，彭錦鵬等人（2017）受考選部委託也在 2017 年執行了公務人員考試採心理測驗之可行性研究，內容包含蒐集並分析國內外公務人員考試併採心理測驗之類型與辦理情形、採行心理測驗的目的、適合採行的方式，以及如何規劃相關配套措施（例如：需對外公開的資訊以昭公信、心理測驗的專業品質要求、維持考試方式的持續運作機制）等，該研究從文獻資料分析、焦點團體座談及深入訪談等方式進行評估研究。研究結果建議，公務人員考試應採用二階段的篩選機制，考慮加入能測量高層次認知能力的測驗，例如：邏輯判斷、資料分析等性向測驗，並以職務工作分析來作為選擇公務人才特質的依據，而現行成就測驗之選擇題或申論題應盡量朝測量高層次認知能力或素養來設計，避免都是測量知識的記憶與理解。該研究並建議成立「公務人員考試採心理測驗推動計畫」的專責小組。

## （三）測驗內容與試題範例研究

除了可行性的評估，在考選部委託陳柏熹等人（2020a）進行公務人員性向測驗研發計畫中，規劃了適用於公務人員考試的性向測驗內容向度、研發試題範例、進行信度及效度評估，並研擬常模建置方式與分數使用之建議。該研究根據國內外公務人員核心職能相關文獻之內涵進行整理後（江明修，2002；考選部，2013；施能傑，2010；高永光，2013；陳玉貞，2013，2016；Cofsky, 1993；McClelland, 1973；Spencer & Spencer, 1993），大致將公務人員的職能區分為專業技能、性格、性向三大類職能。其中，專業技能較偏「成就測驗」，這是測量某特定學科的學習成就，與學習環境及個體努力程度有密切關係；而性格方面主要為測量受測者習慣的反應型態，並無對錯之分，屬於「人格測驗」；至於性向方面的核心職能，

則包含溝通表達、判斷思考、創新求變、數理統計、語文理解、問題解決、運算思維等七種，如表1所示。該研究並以此核心職能結合現行出版之學業及職業性向測驗重要向度，發展出公務人員性向測驗內容向度與試題範例，如表2及圖1所示（陳柏熹等人，2020b）。

表 1

## 公務人員核心職能中有關性向之分類整理

- (一) 溝通表達：媒體溝通、領導／指導他人、溝通協調／聯繫、言談表達／答辯／談判、行銷與顧客服務、公共關係、傳達訊息、協商技巧……等。
- (二) 判斷思考：思考／判斷／決策力、政策分析、環境變化與趨勢分析、邏輯演繹／邏輯分析、成本／效益分析、考核／評估、批判反思、觀察力……等。
- (三) 創新求變：思考創造、創意思考、創造力、創造革新、創新改進……等。
- (四) 數理統計：數理推理、統計……等。
- (五) 語文理解：資訊文書、應用文書／公文寫作、理解／閱讀理解力……。
- (六) 問題解決：策略思考與規劃、策略執行、問題解決、應變、方案設計、願景規劃、應變／危機處理、複雜問題／分析及解決、調查／研究……等。
- (七) 運算思維：資訊運用、電腦應用、資訊技術……等。

註：引自「公務人員性向測驗的發展與應用」，陳柏熹、黃宏宇、蘇少祖、丁姿云，2020b，國家菁英季刊，13（2），頁 144。

表 2

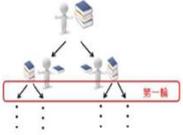
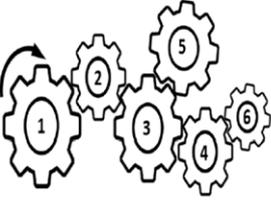
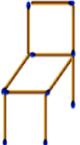
## 公務人員性向測驗之內容向度及內涵定義

向度	次向度	內涵說明
數理邏輯／ 運算思維	數理邏輯	對數目關係的理解和推理，以及處理數量材料的能力。
	運算思維	運用抽象化、演算法設計、模式辨識等運算及思維能力解決問題的能力。
空間概念／ 機械推理	空間概念	在腦中思考及運作圖形或三度空間的能力，依據圖解或圖畫來想像物體外觀及位置的能力。
	機械推理	應用基本的機械原理、工具配件與物理力學原理進行推理之能力。
問題解決／ 批判思考	問題解決	面對問題時能釐清問題重點，判斷有用資訊與提出解決方案，並評估方案優劣勢的能力。
	批判思考	辨認證據、假設、歸納、演繹、解釋、評鑑的批判思考能力。
語文推理／ 溝通表達	語文推理	語意了解、推斷兩字詞所含意義間關係並進行抽樣或概括推理的能力。
	溝通表達	正確理解訊息，並運用適當的表達與溝通技巧闡述個人見解的能力。
	創造思維	能以不同觀點提出新穎的想法或思維的能力。

註：引自「公務人員性向測驗的發展與應用」，陳柏熹、黃宏宇、蘇少祖、丁姿云，2020b，國家菁英季刊，13（2），頁 145。

圖 1

公務人員性向測驗範例試題

<h3>試題設計-數理邏輯</h3> <p>✓ 請你依照下方九宮格直欄及橫列的數字規則，判斷出？為何數字。</p> <table border="1" style="margin-left: 20px;"> <tr><td>4</td><td>6</td><td>2</td></tr> <tr><td>7</td><td>9</td><td>5</td></tr> <tr><td>15</td><td>17</td><td>?</td></tr> </table> <p>(A)12 (B)8 (C)19 (D)13</p> <p>✓ 行政院農委會某秘書參加了台北 101 垂直馬拉松競賽，秘書從 1 樓走到 8 樓需要花 56 秒的時間，請問若以同樣的速度往上走樓梯，秘書還需要幾秒的時間才能走到第 20 層樓？</p> <p>(A)84 (B)91 (C)96 (D)104</p> 	4	6	2	7	9	5	15	17	?	<h3>試題設計-運算思維</h3> <p>✓ 國民健康署某科長想發會議資料給不同部門的每一位同事，因每一份會議資料皆印有同事的名字，因此只能專人傳送。為了節省時間，他將會議資料分成兩大疊，並分別送到該疊會議資料第一份名字的同事手中(如下圖)，再請這兩位同事收下自己的會議資料後，將剩餘的會議資料平均分成兩份(若無法平分，兩疊的數量最多差一張)，分別送到該疊會議資料的第一位名字同事的手中(第一輪)，以此類推。而若會議資料總共有 40 份，請問最多需要經過幾輪的傳送，才可讓每個人拿到自己的會議資料？</p> <p>(A)2 (B)3 (C)4 (D)5</p> 
4	6	2								
7	9	5								
15	17	?								
<h3>試題設計-空間概念</h3> <p>✓ 公共工程甲委員前往新竹高鐵站進行相關工程場勘，他先在高鐵站外拍了一張照片，請你依據照片及衛星圖的相對位置，判斷出甲委員在何處拍照的。</p> <p>(A) A 處 (B) B 處 (C) C 處 (D) D 處</p> 	<h3>試題設計-機械推理</h3> <p>✓ 交通部甲科員正在進行船舶的維修工程，若齒輪 1 為順時針方向轉動，請問下列齒輪中何者方向與齒輪 1 的轉動方向相同？</p> <p>(A) 齒輪 2 (B) 齒輪 4 (C) 齒輪 5 (D) 齒輪 6</p> 									
<h3>試題設計-問題解決與批判思考</h3> <p>✓ 大鵬是話劇社團社長，最近在籌畫社團成果展由九位大四畢業生擔綱演出話劇，現在距離成果展剩下一周，阿強臨時告知大鵬因為家中父親急病過世，無法參加社團成果展，請他另作安排，大鵬如果希望社團成果展能夠順利演出的話，哪些不是他現在應該考慮的問題？</p> <p>(A) 阿強的戲份刪掉對整體的影響 (B) 勸阿強看看是否有辦法參加演出 (C) 由較熟悉劇本的大三生頂替阿強演出的可行性 (D) 阿強角色對整體戲的重要性</p> <p>✓ 行政院農業委員於農作物轉播指導時，在農田裡發現了一大塊礫石，此時甲科員說：「這不是礫，也不是銅」；乙科員說：「這不是鐵，是錫」；丙科員說：「這不是錫，是鐵」。此時隨行的地理老師則說：「你們三人有一人全對，一人全錯，一人半對半錯，請你依據上述的資訊，判斷出此礫石為何物？」</p> <p>(A) 鐵 (B) 錫 (C) 銅 (D) 其他礦物</p>	<h3>試題設計-語文推理與溝通表達</h3> <p>✓ 「歌舞蹈：演員」兩者間的關係與下列哪個選項的關係配對相同？</p> <p>(A) 操場：運動選手 (B) 演唱會：歌手 (C) 電腦：打字員 (D) 辦公室：職員</p> <p>✓ 阿強與阿達今天約好帶家人一起去看爬山，兩家人預計在山腳下會合，但因為阿達家中小孩臨時生病無法赴約，不得已得跟阿達取消原本的爬山計畫，如果阿達希望不要破壞雙方友情，且站在彼此的角度進行思考的話，他該如何跟阿達表達比較好呢？</p> <p>(A) 「我家小孩昨天生病，今天體力還沒恢復，可能爬不上山頂。」 (B) 「我家小孩昨晚臨時發高燒，身體不適，怕傳染給你家小孩。」 (C) 「我家裡有事情，臨時無法出席，你帶著你的家人去吧。」 (D) 「我家今天無法出席，跟你改約下周再去爬山，好不好？」</p>									
<h3>試題設計-創造思維</h3> <p>✓ 請由題目中所提供的三個詞彙進行發想，試著從下列選項中找到一個與題目中三個詞彙各自有相關的字詞。(此處的相關可以是語意關聯、或可以組成成語或用語)</p> <p>開始、太陽、鬧鐘</p> <p>(A) 黃昏 (B) 早晨 (C) 公園 (D) 手錶</p> <p>✓ 甲科員在中午休息時間用 10 根火柴棒排成了一把椅子，請問甲科員最少移動幾根火柴棒的位置，便可使這張椅子變成顛倒的呢？</p> <p>(A) 2 根 (B) 3 根 (C) 4 根 (D) 5 根</p> 										

註：引自「公務人員性向測驗的發展與應用」，陳柏熹、黃宏宇、蘇少祖、丁姿云，2020b，國家菁英季刊，13 (2)，頁 148。

## 二、性向測驗納入公務人員考試中的實務考量及現況分析

為了瞭解性向測驗落實於公務人員考試中可能會面臨到的問題，本文參考陳柏熹（2019）提出之擬定測驗實施計畫架構，結合前一節的幾項研究成果，對未來若將性向測驗納入公務人員考試中需要考量的人、事、時、地、物等因素進行現況分析，供政策擬定與實務規劃之參考：

### （一）人員相關因素

#### 1. 命題者培訓

如同其他公務人員考試的成就測驗一樣，性向測驗也需要招募一批命題委員，這些命題委員最好能具備表 2 所欲測量性向之專長，是熟悉性向測驗向度內涵的心理學家或教育學者。待正式聘任並簽訂試題保密合約後，利用陳柏熹等人（2020a）的內容向度與範例試題編製命題手冊，訓練性向測驗命題委員，每年定期產出所需的試題數量，並協助進行修題與審題工作。

#### 2. 應考者的宣導與公開練習試測

彭錦鵬等人（2017）以及余民寧等人（2016）的研究結果都指出，對外公開性向測驗的內容資訊及品質以昭公信是公務人員考試實施性向測驗的重要配套工作，而考選部於 2015 年公布之心理測驗規則的第 4 條也指出：採行心理測驗時，其種類、評分、參考標準或與其他考試方式分別占總成績之比例或其他相關事宜，於各該考試規則中訂定。其中測驗內容部分可以參考陳柏熹等人（2020b）的內容向度與範例試題編製應考者宣導文宣，未來也可定期公開部分試題讓考生作為練習題，以降低考生的焦慮。至於測驗分數的計算與分數使用方式則應該在實施前先制定好，並撰寫宣導文宣，於正式實行前在考選部網站上公告。

公開部分試題讓考生練習試測除了可以讓考生熟悉試題類型、降低應試焦慮外，還可以作為預試以蒐集考生的作答反應，評估難易度、鑑別度等試題品質資訊，只要都在考選部控制的環境下進行預試（最好是在電腦化測驗試場中進行），並在預試流程中做好試題保密措施及干擾記憶的程序，應能讓公開練習試測成為未來考選部控管試題品質、維持測驗難度與穩定性的常態活動。

### （二）試務相關因素

試務相關因素的考量可以從考試前、考試中及考試後三個時間點來進行規劃，根據考選部長期執行公務人員考試的專業經驗，要在公務人員考試中實施性向測驗應該不成問題。在考試前，需準備好測驗題本及答案卡（若以電腦化測驗的方式實施則是要將性向測驗建置在測驗系統中），並訓練監場人員如何監考性向測驗。此外，也應針對實施性向測驗過程中可能遇到的考生提問或突發狀況，訓練監場人員在考試過程中如何作回應及進行試場紀錄。這些由於跟其他考試的試務流程都相同，因此不再贅述。

測驗結束後的試題保密性、計分方式與分數使用將是公務人員考試實施性向

測驗需要額外規劃的重要工作。以下分別敘述：

### 1. 試題保密性

性向測驗並不像其他考科的成就測驗一樣，有大範圍的學習內容可以快速設計題目。由於性向測驗的題目需仔細考量抽象的潛在特質所表現出來的外顯行為，有些試題還需要設計公務情境讓應試者能展現出與核心職能相關的外顯行為，而且還需進行預試來了解試題難易度與鑑別度、分析測驗的信度與效度以確保測驗的效能，因此不適合在考試結束後就全面公開試題。若現有紙筆測驗的法規要求性向測驗考完後需全面公告試題，則可能需要進行修法，或考慮採用可以不公開試題的電腦化測驗來進行施測。

即使法規允許性向測驗施測後可以不公開試題，性向測驗在考試結束後的試題保密性也比其他成就測驗更為重要，因為這些試題是有標準答案的能力類型測驗，而且未來還會在其他類科考試中被重複使用。試題的保密性需要從施測過程中就開始進行，監場人員需確保應試者在考試時沒有抄錄試題或背誦試題的行為，測驗結束時沒有攜帶任何紙張資料離開，必要時也可以執行統一程序來干擾記憶。

### 2. 測驗計分與解釋

測驗結果的計分與解釋方式主要可以分為常模參照 (norm reference) 與標準參照 (criterion reference) 兩種，常模參照是將應試者的作答結果與群體作比較，以呈現應試者在群體中的相對位置，標準參照則是將應試者的作答結果與事先訂出的通過標準或切截分數 (cut-off score) 作比較，以了解受測者是否具備事先訂定的能力水準。從這兩項測驗結果呈現方式的特徵來看，常模參照可能較適合用在公務人員考試中，而標準參照則是較適合用在專技人員考試中。

性向測驗與其他常模參照的成就測驗較為不同之處在於，性向測驗的常模是整合所有類科的應試群體為常模，而不是以某特定類科的應試群體為常模，就好像國文、英文等普通科目一樣。未來在計分及解釋測驗結果時，應了解性向測驗的常模參照特性，在使用測驗分數時也應考量其參照群體，針對不同類科訂定不同的篩選標準或加權方式，才能符合各類科的選才需求。

測驗計分除了可用古典測驗理論的原始分數作常模參照的轉換外，也可以使用試題反應理論 (item response theory, IRT) 的能力量尺進行常模參照的測驗分數轉換。由於公務人員考試每年舉行數次，考量試題安全性的前提下，可能需要為不同類科考試組合不同的性向測驗試題，因此需考量測驗等化及常模可用性的問題，而 IRT 可根據受測者在各個題目上的答題反應來估計受測者的能力以及題目的難易度，由於 IRT 具有能力估計的不變性與試題參數的不變性，可以將測量相同能力但不同試題的測驗結果放在同一量尺上進行比較，因此很適合用在一年多次且使用不同試題但測量相同能力的性向測驗中，進行分數等化與常模參照轉換。

### 3. 測驗分數的使用

余民寧等人 (2016) 的研究中對心理測驗分數使用的評估主要是著重在公務

人員的晉升、職務分派或儲訓等現職公務員的不同風險用途中，因此本節有關心理測驗用於公務人員甄選的測驗分數使用方式主要參考彭錦鵬等人（2017）及陳柏熹等人（2020a）的研究結果。根據彭錦鵬等人（2017）的評估研究建議，心理測驗在公務人員考試中宜配合分階段考試的方式來實施，並先以試辦形式推展，待蒐集測驗結果的證據確定測驗分數可以預測未來表現後，再正式採計。

陳柏熹等人（2020a）曾邀請協助過考選部考試業務或對於公務人員篩選有相關研究的專家學者召開公務人員性向測驗分數使用諮詢會議，會中建議可以採用門檻法與加權法兩種方式來採計性向測驗分數。其中門檻法是將錄取程序分成兩階段，第一階段以現行公務人員考試的科目來計算總分，第二階段是判斷性向測驗分數是否達到預先訂定的最低門檻，總分高且通過性向測驗最低門檻者才錄取。加權法是將性向測驗的分數與公務人員考試的其他考科加權採計到總分中，作為最後錄取與否的依據。

為了瞭解這兩種分數使用方式在選擇不同權重或不同最低門檻時，對錄取結果之影響程度，陳柏熹等人（2020a）進行了模擬研究，以該研究 434 名樣本在公務人員考試的成就測驗各向度與性向測驗各向度的相關係數，模擬產生八向度多變量常態分布之 1,000 位考生資料，分別代表他們在國文、英文、專業科目、問題解決／批判思考、語文推理／溝通表達、數理邏輯／運算思維、空間概念／機械推理、創造思維的成績，而成就測驗與性向測驗之間的相關係數則設定為低相關（0.1~0.3）與中相關（0.4~0.6）兩套，以模擬當兩者間相關程度不同時使用性向測驗分數的影響，兩種測驗分數使用方式的錄取率都設定為 10%，結果如下：

### (1) 以門檻法使用性向測驗分數

由於性向測驗有五個向度，因此該研究的每一種設定又分為非補償性（性向測驗各向度都要達到某個標準）與補償性（性向測驗各向度的平均值達到某個標準，也就是各向度間可以優劣互補）兩種做法。結果顯示，若性向測驗門檻設為負一個標準差（-1.0 SD，PR 值 $\geq$ 17），只有當成就測驗與性向測驗之間為低相關且性向測驗門檻為非補償式時，對錄取與否的影響比例將達 26%（原本錄取者變成不錄取的比例），其餘情況下對錄取結果的影響比例皆在 5% 以下，如表 3 所示：

表 3

### 性向測驗門檻設定為負一個標準差（-1.0 SD）對錄取結果的影響

	非補償式門檻	性向測驗錄取	性向測驗未錄取	影響比例
成就測驗 與性向測 驗低相關	成就測驗錄取者	74	26	26%
	補償式門檻	性向測驗錄取	性向測驗未錄取	影響比例
	成就測驗錄取者	99	1	1%
成就測驗 與性向測 驗中相關	非補償式門檻	性向測驗錄取	性向測驗未錄取	影響比例
	成就測驗錄取者	96	4	4%
	補償式門檻	性向測驗錄取	性向測驗未錄取	影響比例
	成就測驗錄取者	100	0	0%

註：簡化自公務人員性向測驗研發計畫期末報告（頁 87-88），陳柏熹、黃宏宇、蘇少祖、丁姿云，2020a，考選部。

## (2)以加權法使用性向測驗分數

當性向測驗以 10%、20%、30% 三種不同權重納入總分採計時，原本僅採計成就測驗的錄取者在重新採計性向測驗後的錄取結果如表 4 所示。當成就測驗與性向測驗為中度相關時，若性向測驗加權為 10%，則原本錄取的 100 人中會有 6 人變為不錄取，影響程度僅 6%，若性向測驗加權為 20%，影響程度為 13%，若性向測驗加權為 30%，則影響程度提高到 19%。若成就測驗與性向測驗為低相關，錄取程度變化將提升至 13%~36%。整體而言，性向測驗的加權愈高，對原錄取與否的影響程度越高，特別是當成就測驗與性向測驗相關較低時，其影響程度更大。

表 4

性向測驗權重不同時對錄取結果的影響

成就測驗 與性向測 驗中相關	性向測驗權重 10%	加權後錄取	加權後未錄取	影響比例
	成就測驗錄取者	94	6	6%
	性向測驗權重 20%	加權後錄取	加權後未錄取	影響比例
	成就測驗錄取者	87	13	13%
	性向測驗權重 30%	加權後錄取	加權後未錄取	影響比例
	成就測驗錄取者	81	19	19%
成就測驗 與性向測 驗低相關	性向測驗權重 10%	加權後錄取	加權後未錄取	影響比例
	成就測驗錄取者	87	13	13%
	性向測驗權重 20%	加權後錄取	加權後未錄取	影響比例
	成就測驗錄取者	73	27	27%
	性向測驗權重 30%	加權後錄取	加權後未錄取	影響比例
	成就測驗錄取者	64	36	36%

註：簡化自公務人員性向測驗研發計畫期末報告（頁 81-82），陳柏熹、黃宏宇、蘇少祖、丁姿云，2020a，考選部。

根據上述研究建議，若要在公務人員考試的第二階段中採用門檻法來設定性向測驗最低門檻時，當性向測驗與成就測驗為低相關，最好設定性向測驗為-0.5 個標準差（PR 值 $\geq$ 32）以下的補償式門檻，對錄取結果的影響不會超過 10%，當性向測驗與成就測驗為中相關，則設定性向測驗為-1 個標準差（PR 值 $\geq$ 17）以下的補償式門檻對錄取結果的影響都能在 5% 以內，若設定性向測驗為各向度都需要通過的非補償性門檻，則最好將門檻控制在-0.5 個標準差（PR 值 $\geq$ 32）以下，對錄取結果的影響才能控制在 15% 以內。若以加權法將性向測驗併入公務人員考試總分計算時，當性向測驗與成就測驗之間為低相關，則性向測驗權重最好不要超過 10%，當性向測驗與成就測驗為中相關，則性向測驗權重最好低於 20%，這兩種情況對錄取結果的影響都不會超過 15%。

整體來說，在進行性向測驗分數使用之決策時，可先以試辦的形式蒐集某年度應考者之性向測驗及成就測驗的分數，以實際資料的相關係數進行不同加權與

不同門檻值的影響力分析，以確保決策之合理性與精準度。至於未來是否需要依不同類科來訂定向測驗的分數使用方式，抑或是要不要採用統一的篩選門檻或加權方式，最好能蒐集具有代表性的樣本來檢視各類科在性向測驗上的表現是否真有明顯區別，以及對未來工作表現的預測力，再來決定較適合的性向測驗分數使用方式。

### （三）時程相關因素

在實施性向測驗的時間點方面，根據彭錦鵬等人（2017）的研究建議，性向測驗的施測時間最好不要與現行的成就測驗在同一階段施測，而是在公務人員考試的第二階段中以電腦化測驗方式進行施測，並且在測驗後不要公告試題。陳柏熹等人（2020a）對性向測驗分數使用的專家諮詢會議，也建議可以採用門檻法在公務人員考試第二階段實施性向測驗。即使採用加權法，由於不建議在性向測驗考完試後公開全部試題，與成就測驗作法不同，因此最好也能安排在不同階段實施。

有關性向測驗的考試時間，則應根據未來性向測驗的向度數量、題數及試題難度實際進行評估。根據陳柏熹等人（2020a）的研究資料，當性向測驗為五個向度共 55 題時，若施測時間為 70 分鐘有 90% 以上的受測者皆能全部作答完成。其中，該研究的語文推理／溝通表達試題較簡單（平均答對率約 0.80），創造思維的試題略偏困難（平均答對率約 0.47）。未來性向測驗的考試時間需要多久則要根據向度數及題數再做試驗。

### （四）場地與設備相關因素

性向測驗的施測場地與環境主要與測驗的方式有關，若採用紙筆測驗施測，則施測場地與目前公務人員考試的成就測驗考場無異。若考量前述第二項試務相關因素的分析，為了做到試題保密性，採用電腦化測驗來進行施測，則需要考慮目前可使用的電腦化試場所容納的人數。根據考選部的考選統計年報，每年報考公務人員考試的人數在 2011 年約有 51 萬人，到了 2021 年已降至約 24 萬人（江宗正，2022）。即使每年分兩次舉行電腦化性向測驗，每次一天可以實施 4 梯次，每梯次也需要能容納將近 30,000 人的電腦考場。

面對如此大量人數的電腦化測驗，除了電腦考場的容納人數外，如何確保電腦設備的穩定性及資料傳輸的速度與安全性也是重要問題。在電腦設備穩定性方面，除了要定期檢測與考前壓力測試，讓各考場電腦都能維持在最佳的待測狀態外，也應該準備好面對單一電腦設備異常或整個考場停電等不同程度突發狀況的緊急應變措施。在資料傳輸速度與安全方面，可以在各考場架設伺服器（server）以區內網路（local intra-net）的方式來進行施測，分散每台伺服器需要承載的用戶端（client）電腦數量，較能確保資料傳輸的速度與安全性。由於考選部過去已有執行大量人數電腦化測驗的經驗，2023 年起電腦考場可以容納的人數也將突破萬人，相信未來要實施電腦化性向測驗應該不成問題。

## （五）測驗工具相關因素

目前公務人員考試尚未發展出正式版本的性向測驗，只有陳柏熹等人（2020a）承接考選部研究計畫所發展出來的試題範例及示範性測驗題本兩套。未來若要在公務人員考試中正式實施性向測驗，可能需要儘早評估是要採用市面上現有的性向測驗，還是自行發展適用於公務人員的性向測驗，並規劃性向測驗題目與常模汰舊換新事宜。

### 1. 使用現有測驗或自行研發測驗

若要使用現有的性向測驗，目前國內較常使用的性向測驗有學術性向測驗與職業性向測驗兩種，學術性向測驗的內容向度大多與學校的學習活動有關，職業性向測驗則是與生涯選擇或職業活動有關。早期的性向測驗有多因素性向測驗（路君約等人，1994）與通用性向測驗（內政部職業訓練局，1985）。前者適用對象為中學生，測驗內容為語文推理、數學推理、抽象推理、機械推理、空間關係、文法與修辭、知覺速度與確度等，主要用在幫助學生作為未來學習與職業選擇的參考，效度分析也是以學生的國文、英文、數學等成績進行相關研究。後者原為美國勞工局編製，由臺灣內政部職業訓練局翻譯並建置國內常模，測驗結果可算出一般智能、語文、數字、空間、圖形知覺、文書知覺、動作協調、手指靈巧與手部靈巧等九項分數，效度分析是以高職學生的共同科或職業類科成績進行相關研究。

近期的性向測驗有新編多元性向測驗（吳明雄等人，2011）與適性化生涯性向測驗（宋曜廷，2015）。新編多元性向測驗的內容向度有語文推理、數字推理、圖形推理、機械推理、空間關係、中文詞語、英文詞語、知覺速度與確度等，雖然可作為學生學習輔導或是企業篩選新進員工之參考，但測量的內容向度較偏重在學校的學習情境，效度分析方面也是以學生的入學考試成績進行相關研究。適性化生涯性向測驗的內容向度有語文、數學、空間、邏輯推理、觀察、科學推理、美感、創意八個分測驗，測驗對象主要是國中學生，提供學生瞭解自身優勢，選擇未來求學進路的參考，效度分析也是以學生的入學考試成績進行相關研究。

從國內現有的性向測驗來看，測量的內容向度與功能都是以輔導學生進行生涯適性發展為主，較不適合用於公務人員的甄選，可能需要開發較適合用在公務人員甄選的性向測驗。若要自行發展公務人員性向測驗，可以參考陳柏熹等人（2020a）的公務人員性向測驗研發計畫，該研究參考了過去性向測驗的常見向度，以及公務人員核心職能分析的結果，整理出公務人員性向測驗的五大向度及範例試題，見表 2 與圖 1 所示。為了能與公務人員的未來工作情境結合，該研究發展出一般選擇題與情境式選擇題兩種範例，其中情境式選擇題主要是將公務人員可能會遇到的情境結合所欲測量的能力向度，作為題目的問題情境，如圖 2 右方所示。情境式性向測驗題的情境不能太特殊化，否則該情境測量到的能力無法類化到其他類似的公務情境中，但也不能太廣泛性，否則只能測量到與一般選擇題相同的能力，無法測出應試者能否將在這些能力應用在實務情境中。

圖 2

## 公務人員性向測驗的選擇題範例

<p><b>試題設計-語文推理與溝通表達</b></p> <p>✓ 「歌舞劇：演員」兩者間的關係與下列哪個選項的關係配對相同？</p> <p>(A) 操場：運動選手 (B) 演唱會：歌手 (C) 電腦：打字員 (D) 辦公室：職員</p>	<p><b>試題設計-空間概念</b></p> <p>✓ 公共工程甲委員前往新竹高鐵站進行相關工程勘測，他先在高鐵站外拍了一張照片，請你依據照片及衛星圖的相對位置，判斷出甲委員在何處拍照的。</p> <p>(A) A 處 (B) B 處 (C) C 處 (D) D 處</p> 
<p>一般選擇題</p>	<p>情境式選擇題</p>

註：引自公務人員性向測驗研發計畫期末報告(頁 21)，陳柏熹、黃宏宇、蘇少祖、丁姿云，2020a，考選部。

雖然陳柏熹等人（2020a）的研究已發展出兩套示範性題本，但現有題本在部分向度的難易度及測驗信度方面仍有待提升，不論是古典測驗理論或試題反應理論的分析結果都發現，該測驗的語文推理／溝通表達的試題偏易，試題答對率偏高，可能因為這是一般人日常生活中的基本技能；創造思維的試題則偏難，可能是由於創造思維的試題是日常生活中較不常見的思考活動，不論是現職公務人員或大學生都是長期被要求須有正確的概念及答案，或要以既定的規則來辦事，因此面對這種需要從不同角度來思考的問題時較為困難。

信度方面，該研究以內部一致性係數與 IRT 的測驗訊息量所轉換的區分信度 (separation reliability) 作為測量精準度的指標。由於公務人員性向測驗中有部分試題是多元計分題，因此該研究採用 Cronbach's  $\alpha$  係數分別對題本 A 及題本 B 進行信度分析。其中創造思維向度由於較不適用單向度的概念，因此測驗信度較低，其他向度在題數僅為 8~14 題的條件下，信度約在 0.4~0.75 左右。由於該研究兩個題本之間有共同題，可利用 IRT 同時估計法將兩題本合在一起進行試題連結與受測者能力等化，再根據能力值變異數與能力估計誤差推估兩題本合併後測驗的區分信度，其結果顯示，除了創造思維向度外，其餘四個向度在 12~23 題條件下，能達到受測者區分信度介於 0.67~0.72，試題區分信度高於 0.95 的水準（陳柏熹等人，2020a）。

效度方面，該研究採用不同群體差異比較來呈現區辨效度。從不同群體在公務人員性向測驗上的測驗結果可以發現，現職公務人員在性向測驗各向度上的表現略高於非公務人員與一般大學生，尤其是現職公務人員與非公務人員的差異更為明顯，如表 5 所示。

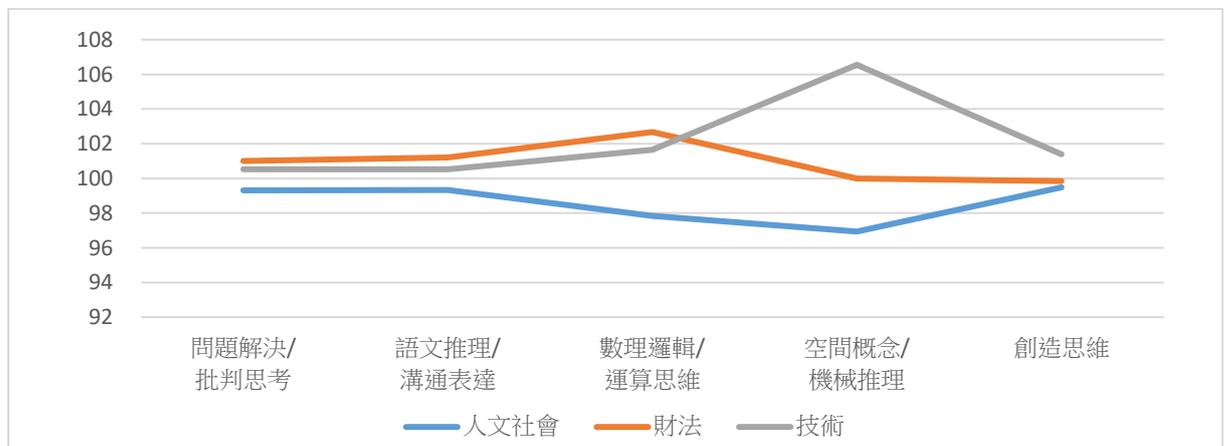
表 5  
不同群體在公務人員性向測驗的描述性統計

現職	人數	向度				
		問題解決/ 批判思考	語文推理/ 溝通表達	數理邏輯/ 運算思維	空間概念/ 機械推理	創造思維
		平均數 (標準差)	平均數 (標準差)	平均數 (標準差)	平均數 (標準差)	平均數 (標準差)
非公務員	92	97.7 (6.8)	97.6 (7.0)	97.2 (10.1)	92.9 (17.8)	98.1 (4.9)
公務員	235	100.9 (6.3)	101.0 (7.0)	100.9 (10.1)	103.0 (13.4)	100.9 (3.8)
學生	107	100.2 (7.0)	100.1 (7.6)	100.7 (10.5)	100.4 (16.6)	100.0 (4.4)

註：表中能力量尺平均數 100，標準差 15。引自公務人員性向測驗研發計畫期末報告（頁 64），陳柏熹、黃宏宇、蘇少祖、丁姿云，2020a，考選部。

根據陳柏熹等人（2020a）的研究結果，報考不同類科公務人員考試之應考者在性向測驗各向度上的表現也有所不同，如圖 3 所示，在「數理邏輯／運算思維」上，財法類與技術類皆高於人文社會類；在「空間概念／機械推理」上，以技術類表現最佳，人文社會類表現較弱；在「問題解決／批判思考」、「語文推理／溝通表達」上，三類科的表現差異並不大；在「創造思維」上，技術類表現略高於人文社會類與財法類。圖 3 之測驗結果也可看出人文社會類與財法類雖同屬於行政類科，但兩者在公務人員性向測驗各向度上的表現有所不同，頗符合該研究專家會議中學者們對這兩大類科的核心職能有所不同應區分開來討論的論點。此外，除了技術類科的現職公務員與非公務員在某些向度的表現差異不大之外，人文社會及財法類的現職公務員在性向測驗各向度上的表現均優於非公務員，表示測驗各向度對於人文社會類科與財法類科的公務人員篩選都能發揮良好的區辨作用，對於技術類科公務人員的篩選，則是「空間概念／機械推理」向度可以發揮較高的區辨作用。

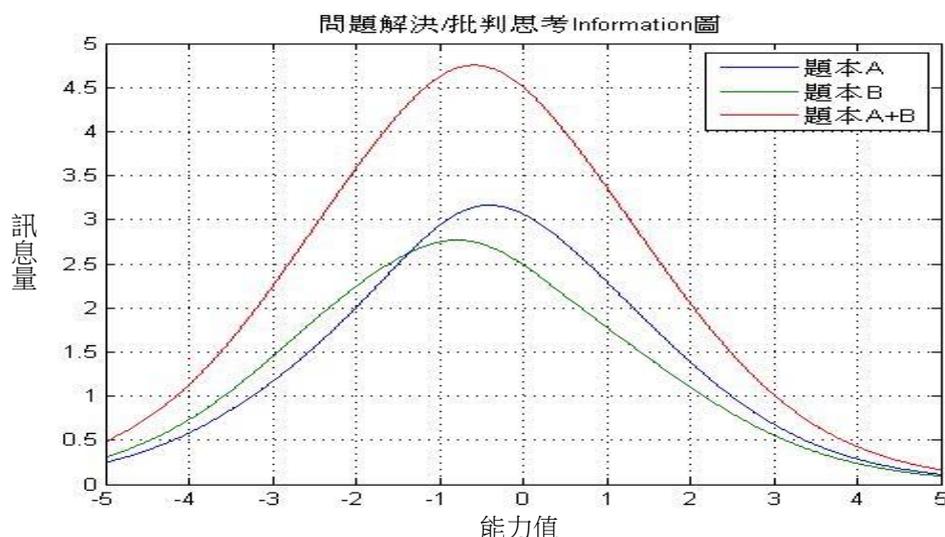
圖 3  
不同群體受測者在公務人員性向測驗的結果



註：修改自公務人員性向測驗研發計畫期末報告（頁 65），陳柏熹、黃宏宇、蘇少祖、丁姿云，2020a，考選部。

綜合上述，未來若要使用公務人員性向測驗，可以使用陳柏熹等人(2020a)所發展的題本為基礎，增加作答時間並提高題數至每向度 20 題左右，使各向度的內部一致性信度能達到 0.80 以上，從該測驗試題反應理論的訊息量圖可以看出，大部分內容向度都能在中等能力範圍 (-2.0 ~ +1.0) 提供較高的測量精準度 (圖 4 以問題解決/批判思考向度為例)，大致能符合前述第(二)項試務相關因素以門檻法以低於平均數 0.5 或 1 個標準差為篩選標準進行公務人員初步篩檢的需求。

圖 4  
公務人員性向測驗的訊息量 (問題解決/批判思考向度)



註：修改自「公務人員性向測驗的發展與應用」，陳柏熹、黃宏宇、蘇少祖、丁姿云，2020b，*國家菁英季刊*，13 (2)，頁 154。

## 2. 題庫建置與常模更新

公務人員性向測驗與一般性向測驗不同，由於測驗分數使用可能涉及工作機會，試題風險性比一般性向測驗高，而且每年需要在不同時段舉行大規模施測，因此需要有多套不同題本交替使用，以免應試者背誦試題影響公務人員甄選的公平性。較佳的測驗建置方式是比照部分類科的考試科目一樣建置題庫。為了符合每年能組合多套平行題本的需求，本文建議未來可以根據每年考試的次數來規劃題庫建置的目標題數，若每年考 4 次性向測驗，則題庫中的題數建議維持在正式題本題數的 4 倍以上，若考試完能不公開試題，則在組合不同題本時可以有適當比例的重疊試題，並以試題反應理論來協助進行分析，讓不同題本的組題及測驗分數等化更有心理計量的學理依據。

常模是公務人員性向測驗另一項重要考量，由於此測驗要用來瞭解應試者在所有報考者中的相對位置，甚至可能要了解在該報考類科群體中的相對位置，因此需要以全體考生與所報考類科考生兩種不同群體來建置不同的常模對照表，未來才能在測驗結果中展現出考生在不同常模群體中的多元潛能表現，提高分數解釋與預測的效度。由於有些考科的報考人數可能很少，即使建置常模也不太具有

代表性，因此可以考慮先將不同類科考試整併成類群（例如：人文社會群、財政法律群、技術類群），先以類群為單位來建置常模，待未來蒐集較多年數據且人數較多時，若發現某類科與所屬類群的其他類科在性向測驗各向度上的表現有明顯差異，再考慮是否獨立出該類群另外建置常模。

## 參、結論與建議

多元潛能是公務人員選才與適性發展的重要觀念，尤其在人口大量減少，年齡結構逐漸老化的趨勢下，願意投入公務服務的人才也逐漸減少，從近十年來報名公務人員考試的人數銳減的情況就能看出此趨勢。隨著每年公務人力的退休，在公務人力需求數量不變，但人口逐漸減少與老化、企業積極搶人才的趨勢下，未來恐會面臨公務人力供不應求的情形。面對這樣的危機，評估多元潛能並給予公務人員適才適所的安置，將是未來選拔公務人員的重要措施。單純只以知識概念學習成效的成就測驗來評估一個人是否適合擔任公務員，恐不足以因應未來公務人力的整體發展需求，若能加入性向測驗，不僅可以讓應試者了解自己各方面的多元潛能，也能讓用人機關知道如何將不同潛能的公務員安置在較適合的位置，讓公務人力發揮最大的效用，以因應未來公務人才不足的潛在危機。

要將性向測驗真正落在公務人員的選拔中，需要從人、事、時、地、物等多方面作仔細規劃，才能讓性向測驗的推動發揮真正效用。根據過去幾份相關的評估研究報告結果，以及本文對此五大面向的現況分析，若政府有意將性向測驗納入公務人員選拔的評測項目，作者建議近期可以朝三個方向來進行努力：

### 一、規劃實施方式與分數使用方式

根據本文的分析，性向測驗最好與一般科目成就測驗的施測日期與時間不同，可以在第二階段以電腦化測驗方式來執行，每次從題庫中抽組不同題本，讓不同題本試題重疊率控制在一定程度以下（例如 20%），考完不公布全部試題，只公布範例試題或定期公布曝光率（使用率）較高的試題供考生練習，可以同時兼顧試題保密性及測驗結果的效度。

分數使用可以採用門檻法或加權法，但初期建議是以不列入計分的方式試測 1~2 年，蒐集性向測驗與一般科目成就測驗的相關性，並分析性向測驗在各類科較適合的門檻值或加權值，以及對未來工作表現的預測效果後，再來決定適當的分數使用方式。

### 二、建置題庫與分析常模

目前尚無適用於公務人員且品質令人滿意的性向測驗，需要儘快籌組團隊來設計試題及各類科常模。可使用考選部委託的公務人員性向測驗研發計畫為基

礎，根據五大內容面向來設計試題，並試測 1~2 年蒐集考生的作答反應以分析試題品質與測驗效能，去蕪存菁，並建置常模，分析全體常模與類科常模參照表的篩選效能與行為預測力，以符合未來每年多套題本的施測需求及測驗常模解釋力。

隨著時代的演變愈來愈快，近十年來科技發展、人口年齡結構，以及執行公務的類型都有所改變，執行公務所需要的核心職能也會逐漸改變，未來公務人員性向測驗的題庫及常模也需要隨著時代的變化做適當的調整，才能符合國家的需要拔擢優秀的公務人員，造福社會。

### 三、成立公務人員考試心理測驗推動小組

由於上述各項工作涉及了考選部規劃司、高普考試司、題庫管理處、資訊管理處……等多處室的業務協調，雖然不必大費周章成立財團法人或獨立機構來推動，但建議可以由部長或次長擔任召集人，成立公務人員考試心理測驗推動小組，邀請外部測驗專家擔任諮詢顧問，各相關處室主管擔任推動小組委員，聘任 2~3 位約用人員來規劃與執行上述各項業務，每月定期開會報告推動進度，並適度協調各處室的人力兼辦相關業務。初期可先以發展性向測驗之題庫、常模及分數使用方式為推動目標，未來再視情況加入性格測驗或價值觀測驗等其他心理測驗，以優化公務人員考試人才選拔的效能。

### 參考文獻

- 內政部職業訓練局（1985）。**通用性向測驗指導手冊**。內政部職業訓練局。
- 伍姿蓉（2012）。論警察特考納入心理測驗施測之內容與施測之時機。**國家菁英季刊**，8（1），117-151。
- 江宗正（2022）。公務人員考試不再熱門了嗎？影響報考人數的原因分析。**國家人力資源論壇**，18。  
[https://www.exam.gov.tw/NHRE/News\\_EpaperContent.aspx?n=3778&s=45441&type=DD102593FDB1A032](https://www.exam.gov.tw/NHRE/News_EpaperContent.aspx?n=3778&s=45441&type=DD102593FDB1A032)
- 江明修（2002）。公務人員各官等核心能力與訓練體系建立之研究。公務人員保障暨培訓委員會委託研究。
- 考選部（2013）。**國家考試職能分析推動工作第一及第二梯次職能評估成果**。考選部。
- 考選部（2020）。**民航人員考試飛航管制人員體格複檢標準表**。  
[https://wwc.moex.gov.tw/main/controls/wHandEditorExtend\\_File.ashx?Fun=Property&menu\\_id=335&item\\_id=4856&file\\_id=9808](https://wwc.moex.gov.tw/main/controls/wHandEditorExtend_File.ashx?Fun=Property&menu_id=335&item_id=4856&file_id=9808)
- 余民寧（2015）。典試法修正後的幾個題庫建置與測驗編製問題。**T&D 飛訊**，211，1-19。
- 余民寧、謝進昌、張藝馨、黎佩欣、彭英錡（2016）。公部門運用心理測驗及其

- 法律背景之研究**。行政院人事行政總處公務人力發展中心。
- 吳明雄、陳榮華、陳心怡 (2011)。新編多元性向測驗指導手冊。中國行為科學社。
- 宋曜廷 (2015)。適性化職涯性向測驗國中版指導手冊第五版。國立臺灣師範大學心測中心。
- 李大偉、李建興、胡茹萍、黃嘉莉 (2012)。我國大學多元入學制度之評估研究。行政院研究發展考核委員會委託研究報告。[https://ws.ndc.gov.tw/001/administrator/10/refile/5644/3267/0058949\\_1.pdf](https://ws.ndc.gov.tw/001/administrator/10/refile/5644/3267/0058949_1.pdf)
- 李登科 (2012)。外交領事人員考試與核心職能之研究。國家菁英季刊，8 (1)，87-115。
- 施能傑 (2010)。職能理論對國家考試制度設計的啟示。國家菁英季刊，6 (3)，17-35。
- 施能傑、曾瑞泰、蔡秀涓 (2009)。美國、英國和日本中央政府初任文官的甄補制度介紹。國家菁英季刊，5 (1)，13-34。
- 段美玉 (2006)。臺灣市面販售用於人事甄選之心理測驗的比較研究〔未出版之碩士論文〕。國立中山大學。
- 高永光 (2013)。公務人員考試與核心職能測驗分析。國家菁英季刊，9 (2)，1-22。
- 陳玉貞 (2013)。職能分析導入國家考試口試制度之探討。考選論壇季刊，3 (4)，16-28。
- 陳玉貞 (2016)。從職能觀點探討國家考試應試類科、科目簡併。考選通訊，68。
- 陳柏熹 (2019)。心理與教育測驗：測驗編製理論與實務。精策教育。
- 陳柏熹、黃宏宇、蘇少祖、丁姿云 (2020a)。公務人員性向測驗研發計畫期末報告。考選部。
- 陳柏熹、黃宏宇、蘇少祖、丁姿云 (2020b)。公務人員性向測驗的發展與應用。國家菁英季刊，13 (2)，139-162。
- 陳柏熹、黃馨瑩、陳郁欣、葉泰廷、蘇少祖、李柏瑋 (2015)。大學生基本素養測驗的發展及信度效度分析。教育科學研究期刊，60 (3)，95-125。
- 陳皎眉、黃富源、孫旻暉、李睿杰 (2011)。論心理測驗與國家考試。國家菁英季刊，7 (1)，15-31。
- 彭錦鵬、鄭夙珍、李俊達 (2017)。公務人員考試採心理測驗之可行性。考選部。
- 葉玉玲、丁一顧 (2021)。我國大學多元入學方案之回顧與省思。臺灣教育評論月刊，10 (4)，1-9。
- 路君約、盧欽銘、歐滄和 (1994)。多因素性向測驗。中國行為科學社。
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Prentice Hall.
- Armed Services Vocational Aptitude Battery. (2023). *History of military Test*.  
<https://www.officialasvab.com/researchers/history-of-military-testing/>
- Assessment & Teaching of 21st Century Skills. (2010). *21st century skills white*

- paper. <http://atc21s.org/index.php/resources/white-papers/>
- Binet, A., & Simon, T. (1905). New methods for the diagnosis of the intellectual level of subnormals. In H. H. Goddard (Ed.), *Development of intelligence in children (the Binet-Simon Scale)*. Williams & Wilkins.
- Cofsky, K. M. (1993). Critical Keys to Competency-based Pay. *Compensation and Benefits Review*, 25(6), 46-52.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books.
- Gardner, H., Feldman, D. H., & Krechevsky, M. (1988). *Project spectrum: Early learning activities, Project Zero frameworks for early childhood education* (Vol. 2). Teachers College Press.
- Guilford, J. P. (1988). Some changes in the structure-of-intellect model. *Educational and Psychological Measurement*, 48(1), 1-4.
- McClelland, D. C. (1973). Testing for competence rather than for “intelligence.” *American Psychologist*, 28(1), 1-14.
- Office of Personnel Management. (2019). *Delegated examining operating handbook: A guide for federal agency examining offices*. [https://www.opm.gov/policy-data-oversight/hiring-information/competitive-hiring/deo\\_handbook.pdf](https://www.opm.gov/policy-data-oversight/hiring-information/competitive-hiring/deo_handbook.pdf)
- Spearman, C. E. (1927). *The abilities of man, their nature and measurement*. Macmillan.
- Spencer, L. M., & Spencer, S. M. (1993). *Competence at work: Models for superior performance*. John Wiley & Sons.
- Terman, L. M. (1916). The measurement of intelligence. In W. Dennis (Ed.), *Readings in the history of psychology* (pp. 485-496). Appleton-Century-Crofts.
- Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review*, 38(5), 406-427.
- Thurstone, L. L. (1938). *Primary mental abilities*. University of Chicago Press.
- Thurstone, L. L., & Thurstone, T. G. (1941). Factorial studies of intelligence. *Psychometric Monographs*, 2, 94.