

交通警察取締闖紅燈執法能力感知- RASCH 模式之應用

張新立¹ 施昌谷² 陳少旭³

摘要

由於闖紅燈照相設備價格昂貴，且無法普遍設置，為防制闖紅燈肇事行為，警察執行取締作為是一種經常性又符合經濟原則的手段。本文針對台灣358位外勤員警隨機抽樣調查，應用Rasch模式分析交通警察執法能力問卷（Traffic Police Enforcement Ability Questionnaire, TPEAQ），並據以評估警察取締闖紅燈行為的執法能力。分析結果發現，整體而言，尋找取締闖紅燈合適地點、發現闖紅燈行為，以及避免申訴的能力，是員警認為較容易把握的項目。另一方面，員警普遍認為追緝逃逸的闖紅燈違規行為車輛，是較不易執行的任務。此外，對於執行取締闖紅燈的執法行為，女警自我感知其能力高於男警，而經DIF檢測，主管則自覺其在避免民眾抱怨申訴處置的能力上高於員警，再年輕警察比年長警察執法能力為佳。進一步而言，在體能的負荷上，員警對於2個小時的取締勤務已顯疲倦，甚至會有腰酸背痛現象，因此，編排取締勤務以不超過二小時為宜。在特殊天候的表現上，員警認為雨天及交通尖峰時間執行取締勤務較為困難。因此為顧及安全，遇雨天時，應避免執行取締勤務，在交通尖峰時間應以疏導交通為主。根據本研究的結果，我們建議為普遍提高警察執法能力，警察訓練應加強安全駕駛訓練、強化體能負荷能力、及訓練員警在天候不良狀況下的執法技巧。在取締能力應加強部分，女警首應加強的能力為心智（Mental）與心理（psychological）需求能力，而男警應加強訓練其心理與生理（Physical）需求能力；最後本文以平方和指標（DI, Sum of Squares Indicator）測量，在全部358個受測員警中，有75個員警（20.95%），執行取締闖紅燈行為的能力顯著低於平均值，針對這些員警應優先列為教育與訓練的對象。

Keywords: red light running, enforcement ability, Rasch model

壹、前言

闖紅燈是經常發生的交通違規行為，Retting 與 Williams（1996）研究發現路口每小時平均可發現2個闖紅燈交通違規行為[1]，Porter 與 England（2000）觀察交岔路口車流行為，進一步指出，每小時平均可發現10個闖紅燈交通違規行為[2]。Retting等人（1998）研究美國1992年至1996交通事故指出，有3%的死亡交通事故與闖紅燈行為有關[3]。美國

¹交通大學運輸與科技管理研究所教授

²交通大學運輸與科技管理研究所博士班研究生

³台北市警察局交通大隊執法組組長

都市道路交通事故有22%與闖紅燈駕駛行為涉入有關[4]，這些現象均顯示闖紅燈行為對交通事故確有極大影響，也是交通管理重點之一。但由於駕駛人普遍存在闖紅燈時不一定會有警察取締，且警察取締也不一定能攔下闖紅燈駕駛人的心態，致某些狀況下取締闖紅燈行為不易執法，Porter 與 Berry (2001) 研究即指出，在有警察執法路口，對於10件闖紅燈違規，警察有辦法攔下取締者，少於2件[5]。因此，評估警察執法工作能力，並進而研擬提升執法技能，為有效遏制闖紅燈行為的方法之一。

依據我國內政部警政署統計，台灣地區機動車輛92年至96年A1及A2類(死亡及受傷)交通事故原因分析，因闖紅燈駕駛行為(包括違反號誌管制)涉入者，平均每年有11,258件，其中因而死亡人數，每年平均有136人；以A2+A1件數而言，居肇事原因的3位，僅次於「未依規定讓車」與「未注意車前狀態」等二項肇事原因；以A1類肇事原因而言，僅次於「酒醉駕駛失控」、「超速」、「未依規定讓車」與「未注意車前狀態」等4項歸責駕駛人因素(human factor)，居肇事死亡原因的5位，可見台灣闖紅燈違規行為的確是肇事的主因之一，有必要加強取締，以維護交通安全。台灣已將闖紅燈行為與酒醉駕車及超速並列為三項重大嚴重違規案件，然而每一路口的設置費用在一百萬元以上，只能重點佈置，無法在每個路口、每一方向普遍設置，因此，闖紅燈違規行為仍有賴警察以巡邏及路口站崗取締模式執法，其中又以路口站崗模式為警察普遍執法模式，因此，員警在交岔路口站崗取締闖紅燈違規行為執法能力，即執法是否成功的關鍵原素之一。因此，提升員警在交岔口執行取締闖紅燈能力，對於降低闖紅燈違規行頻率並進而降低交通事故，是一既有效又經濟的方法。為提升員警執法能力，則必需先對目前員警執法能力及盲點有所了解，並據以發展一套有效提升員警執法能力的訓練計畫，本文目的即在發展一可以有效評估員警取締闖紅燈執法能力問卷，並進而評估員警執行取締闖紅燈違規行為執法能力，以為未來教育與訓練計畫實施方向，並做為相關法規及規範訂定的參考。

本研究係研究交通警察執法的能力，而績效(Performance)是能力(Ability)與動機(Motivation)的函數，即績效 = 能力 × 動機[6]。Schmidt等人(1986)研究工作績效因果模式(Causal model of job performance)時，提出影響工作績效最重要原因為感知能力(cognitive ability)，且能力與績效隨著時間遞移，呈現穩定的關係[7]。Libby與Luft(1993)認為能力系經由績效表現出來[8]。Hunter(1986)回顧諸多學人研究後指出，以一般感知能力(general cognitive ability, GCA)預測工作績效可以獲致良好效果[9]，並表示，一般感知能力之所以對工作績效有高度的預測效果，系因其可以預測學習(Learning)與工作優勢(job mastery)，且能力與工作知識高度相關，而工作知識又與工作績效高度相關[9, p. 354]。Aron等人(1998)對176個製造工人進行研究，得到能力經驗(experience)是預測工作績效的重要指標[10]。Campbell(1993)對工作能力研究進行後均指出，能力與績效間呈現正向相關關係[11]。綜合上述研究，本文評估警察取締闖紅燈行為的感知能力，

除了執法績效與執法能力具有高度相關外，研究目前警察執法能力，得做為未來警察訓練、教育及相關規範擬定方向的參考。

在評估方法上，估計受試者潛在構面的方法很多，如因素分析（factor analysis）、路徑分析（path analysis）、結構方程模型（structural equations model）...等。然而這些方法只應用原始得分（raw scores）計分並以為效度推論的論據，這些傳統方法最大的限制為每一個試題賦予相同的權重（equal weight），也就是受測者對一個非常簡單的問題回答「非常不滿意」，與回答一個困難試題的得分，都計同分1分，然而這兩試題難度，相對於同一受測者能力而言並不相等。對於困難的試題，得分1分表示：「對於困難的事情難以執行」；對於簡單的試題，得分1分表示：「對於簡單的事情難以執行」，這是不合理的現象。此外，試題對不同的團體有極大差距的反應，代表試題存有潛在偏誤（latent bias），也就是試題可能具有不同的意涵，或試題並非針對與欲測量的能力構面而設計[12]，因此，傳統的測量方法在測量能力時並非最佳的方法。為改善以上問題，Rasch模型[13]就廣為應用於發展更具信度的問卷，Rasch分析是樣本獨立及試題獨立，受試者能力與試題難度的估計是互相獨立，也就是所謂的特定客觀性（specific objectivity），Rasch模式將順序分數轉換為洛基（logit）的等距分數，使量尺具有等距及簡單充份性（simple sufficiency）[14]，可用以檢測問卷是否符合單向度的特性，另外，Rasch模式中試題差異功能（differential item functioning，DIF），可用以評估受測者能力跨越不同群體的影響程度。因此，本研究的目的為應用Rasch模式（1）評估試題潛在的偏誤，（2）發展一適合評估警察執法能力的問卷，（3）進而測量警察執法的能力，以為教育、訓練實施方向的參考。

貳、警察取締闖紅燈違規行為能力感知概念性架構

工作能力（work ability）以單一面向無法得到客觀的測量，而需由各種不同面向來衡量，在早期的研究中，工作能力指標（Work Ability Index）被廣泛應用於測量工作能力[15-19]，工作能力指標內容最重要的指標為有關心智工作負荷（mental job demands）、心理資源（psychological resources）、及生理負荷（physical demands）等[20]。Markku（2002）[21]等人視工作能力為有關生理、心理、環境及其他因素間複雜的交互作用。

依據荷蘭1988年實施之醫療檢驗法（Medical Examination Act），對於特別工作負荷量（specific job demands）定義為：必需測量目前人類工程（state-of-the-art or ergonomic）及其可能逾越暴露安全水準（exposure safety levels）或人類平均能力之工作負荷量。高度負荷工作可定義為：「在工作過程中對公眾安全與健康有潛在風險影響者」[22]。在特別工作負荷中，有關高度負荷工作的特性可依據生理、心智與心理等工作特質分類[23]。生理負荷包括精力負荷量（energetic demands）、與生物力學負荷量（biomechanical demands），其中精力負荷包括有氧（aerobic）及無氧（anaerobic）負荷；生物力學負荷量包括肌肉與

骨格系統靜態 (static) 及動態 (dynamic) 負荷；心智負荷即與專注力 (concentration)、注意力 (attention)、記憶力 (memory)、與決策力 (decision-making) 有關；心理負荷則由情緒負荷 (emotional demands)、人際關係 (interpersonal relations)、時間壓力 (time-pressure) 及安全 (safety) 等因素決定[23]。

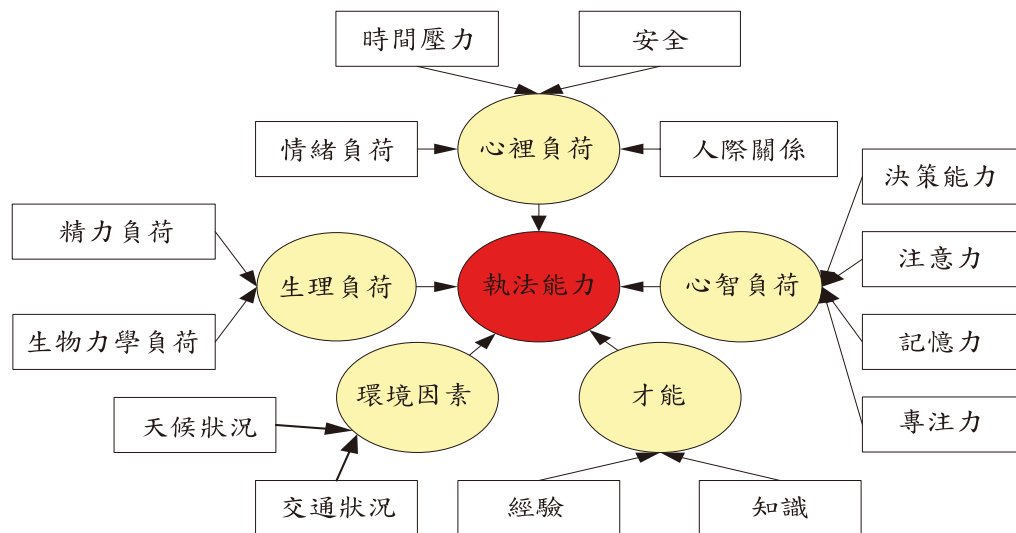


圖 1 交通警察執法能力感知模型

就警察於路口交通執法而言，交通警察任務必需執行交通疏導及取締交通違規，他們必須站立於路口、交通手勢指揮車輛行進與停止、攔停違規車輛、開製違規舉發單、處理違規行為人抱怨與申訴、甚至遇突發狀況時對違規行為人做必要的強制作為，等等作為與執法者本身、違規行為人及其他用路者安全與健康息息相關，是一兼具生理、心理及心智的高度負荷量工作。

才能 (competence) 為影響能力的一個因素，知識 (knowledge) 與經驗又為才能的重要元素。而教育程度是才能的重要元素[24]，而經驗也會影響工作能力[25-26]，Robert 與 David (1990) [27]也指出：「審計決策會因審計人員在訓練與經驗的差異而不同」。目前台灣外警察階級區別，有接受警察大學教育者始得任巡官 (分隊長、派出所主管)，而巡佐 (或小隊長) 係經任職警員多年而提升，因此，本研究以警察階級替代警察知識。在交通環境方面，由於交通狀況隨時間動態變化，故天候狀況 (weather condition) 對於交通執法績效有一定程度的影響。而擁擠程度 (perceived congestion) 對闖紅燈行為也有顯著影響[28]，因此在環境因素方面，本文以天候狀況及交通狀況 (traffic condition) 為影響環境因素的元素。

依據上述文獻，警察執法工作自我感知能力應包括心智負荷 (mental demands)、生理負荷 (physical demands)、心理負荷 (psychological demands)、才能 (competence) 及環境因素 (environment fsactors)，圖一即交通警察執法工作能力感知模型[29]。

參、方法論

3.1 問卷設計

為發展一有效測量警察取締闖紅燈執法問卷，參酌內政部警政署彙編一般交通違規程序，並依本研究測量能力內容，將警察取締闖紅燈執法流程，分解如圖2所示。

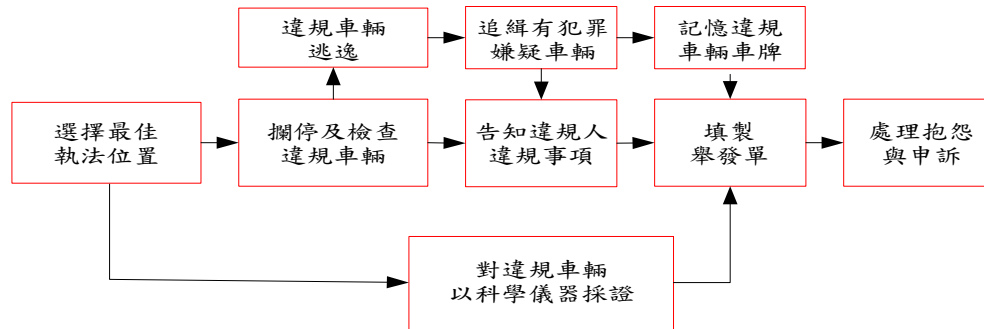


圖 2 交通執法流程示意圖

為能有效率地取締闖紅燈行為，執勤警察首先應具備選擇最佳執法位置的能力，接下來就是能有效率地攔檢違規車輛，檢查人、車，攔檢車輛時並應注意安全。在告知違規人違規事實後，填製違規單舉發，在攔停至填製舉發單的連續作為中，應注意程序合法、及執法態度，以避免受民眾抱怨與事後申訴。若違規車輛經攔查不停逃逸時，除有犯罪嫌疑重大情況得追緝逃逸車輛，一般以記錄車牌（並有採證為佳）後，以逕行舉發方式處理。

本研究依據圖2執法流程，設計警察取締闖紅燈違規「執法能力感知」問卷，問卷試題計20題，20題試題，均以圖1「交通警察執法能力感知模型」中各項元素設計，並採用李克特7尺度（7-point Likert scale），以為衡量、計算參與者的反應。

20題試題中有5題（第1-5題）係依據心智負荷能力設計，這些題目用以測量警察執法時在注意力、專注力、記憶力、或決策的能力；第6-13題等8題，係依據心理負荷能力設計，用以測量執法安全性、情緒負荷、人際關係處理、時間壓迫等的處理能力；第14-16等3題，係依據生理負荷能力設計，用以測量執法時體能負荷與抗疲勞度的能力；最後4題（17-20題）為依據警察執法時不同交通狀況設計，包括單警執勤、雨天、嚴寒天氣、交通尖峰等交通狀況的執法能力。

表 1 TPEAQ 試題設計內容

心智負荷能力	
Item01:	我可以找到取締闖紅燈的最佳位置。
Item02:	一有闖紅燈違規行為，我很容易就可以立即發覺。
Item03:	我很容易可以查覺闖紅燈違規人是否有駕車逃逸意圖。
Item04:	我有辦法記錄下逃逸車輛的車牌號碼。
Item05:	只要我想要攔截逃逸車輛時，我都可以攔下來。
心理負荷能力	

-
- Item06:我可以安全的指揮闖紅燈車輛靠邊停車。
Item07:當違規人要棄車逃逸時，我可以快速將其攔下。
Item08:我可以迅速的上巡邏車（機車），追緝逃逸車輛。
Item09:我有能力駕車緊跟違規逃逸車輛之後。
Item10:我的EQ很好，不會被違規人激怒。
Item11:當違規人現場不認同我的處置，我可以和言悅色的對待以避免紛爭。
Item12:闖紅燈違規人不會在現場執疑我的取締過程。
Item13:當我開出舉發單時，違規人很少提出申訴。

生理負荷能力

- Item14:執行取締勤務2小時，我不會腰酸、背痛。
Item15:執行取締勤務4小時，我不會腰酸、背痛。
Item16:執行取締闖紅燈勤務2小時，我不會覺累。

交通環境處理能力

- Item17:我可以獨自執行取締闖紅燈勤務，不須其他同事幫忙。
Item18:兩天我一樣可以有效率地執行取締闖紅燈勤務。
Item19:在嚴寒天候我仍可以有效執行取締闖紅燈勤務。
Item20:在尖峰時間我也可以有效取締闖紅燈違規。
-

3.2 實證研究

本研究設計TPEAQ問卷計20題試題，於95年9月進行問卷抽樣調查，於台灣北、中、南三區隨機抽樣，對358位警察進行調查，其中125份問卷抽樣自北台灣的警察、120份問卷抽樣自中台灣的警察、113份問卷抽樣自南台灣的警察。警察年齡分佈自22歲至55歲，平均年齡 37.45 ± 9.81 歲。樣本中有220位（61.4%）為警員、87位（24.3%）為巡佐（包括小隊長）、51位（14.3%）為主管（包括分隊長與分駐、派出所所長）；55%為交通警察，45%為行政警察；333位（90.2%）為男警，25位（9.8%）為女警。

3.3 Rasch 模式

心理計量學是一門研究心理測驗與評斷的科學，是一門包括量化心理學（quantitative psychology）、個別差異（individual differences）、和心理測驗理論（mental test theories）等研究範圍的學問。測驗理論（test theory）是一種解釋測驗資料間實證關係（empirical relationships）的有系統的理論學說。目前並行流通於測驗學界中，測驗理論主要劃分成二大學派：一為古典測驗理論（classical test theory），主要是以真實分數模式（true score model）為骨幹；另一為當代測驗理論（modern test theory），主要是以試題反應理論（item response theory）為架構。

模式化的心理計量方法以試題反應理論（item response theory）為基礎[13,30-31]，與傳統以發展試題間相關性與基本構面的因素分析模式具有互補的作用。而由於古典測驗理論忽視受試者的試題反應組型（item response pattern），認為原始得分相同的受試者，其能力必定一樣；其實不然，即使原始得分相同的受試者，其反應組型亦不見得會完全一致，因此，其能力估計值應該會有所不同。

試題反應理論建立在兩個基本概念上：(1)考生在某一測驗試題上的表現情形，可由一組因素來加以預測或解釋，這組因素叫作潛在特質 (latent traits) 或能力 (abilities)；(2)考生的表現情形與這組潛在特質間的關係，可透過一條連續性遞增的函數來加以詮釋，這個函數便叫作試題特徵曲線 (item characteristic curve, 簡寫為 ICC)。我們把能力不同的考生得分點連接起來所構成的曲線，便是能力不同的考生在某一測驗試題上的試題特徵曲線，把各試題的試題特徵曲線加總起來，便構成所謂的試卷特徵曲線 (test characteristic curve, 簡寫為 TCC)。試題特徵曲線即是一條試題得分對能力因素所作的迴歸線，這條迴歸線在基本上是非直線的，但直線的試題特徵曲線也是有可能的，端視所選用的試題反應模式 (item response model) 而定。因此，試題特徵曲線即是一條試題得分對能力因素所作的迴歸線，這條迴歸線在基本上是非直線的，但直線的試題特徵曲線也是有可能的，端視所選用的試題反應模式 (item response model) 而定。試題特徵曲線所表示的涵義，即是某種潛在特質的程度與其在某一試題上正確反應的機率，二者之間的關係；這種潛在特質的程度愈高 (或愈強)，其在某一試題上的正確反應機率便愈大。在試題反應理論中，每一種試題反應模式就有其相對應的一條試題特徵曲線，此一曲線通常包含一個或多個參數來描述試題。

在IRT 的理論中，又以將模型設定為僅探討個人能力 (ability) 與題目難度 (difficulty) 之差異，造成每個人在每個題目測度上差異之 Rasch 模型為最廣為使用之探討架構。正如 Strong 等人 (2005) 指出：「諸多試題反應模式 (item-response models, IRM) 已經成型，而全部 IRM 均發展統計模式，描述受者在橫跨多層潛在構面之試題的反應機率」[32]，而 Rasch 模式即聚焦在測量單一向度試題，並且將原始分數轉換為線性單位 (linear unites) 以定義潛在變數或理論構面。Rasch 模式以等距洛基 (logit) 尺度估計每一位受測者的期望水準 (如能力或特性)，並估計每一個試題的困難度，再據以評估每一受測者填答試題得分的機率 [33]。此模型為 Rasch 所提出，其理論最主要精神乃透過勝算比 (Logodds) 之觀念，將每個題目 (item) 之每個選項 (category) 上，由受試者整體之回應，校估出該題目之難度 (difficulty)。換言之，當多人未答對該題 (或多人選擇該題相對較低之選項)，即代表該題本身設計之難度較高。當每個題目皆以此方式校估出難度，則可藉由每個受試者在各個難易不同之題目上表現，得到受試者在此份試題上之評量分數。如果資料符合 Rasch 或其延展的模式，受試者能力與試題難度的估計是互相獨立，其所生的量尺。就具有等距 (interval) 及簡單充份性 (simple sufficiency) [34]，如評定量表模式 (rating scale model) [35] 及部份給分模式 (Partial Credit Model) [36]。如果資料不符合羅吉特模式，即表示試題未建置為量化變項。以二元計分之 Rasch 模型而言，假設第 n 位受測者能力為 θ_n ， b_i 為試題 i 的困難度，則其答對第 i 題的機率可表示為：

$$P_i(\theta_n) = \frac{e^{(\theta_n - b_i)}}{1 + e^{(\theta_n - b_i)}}$$

在取勝算比並取自然對數後可得

$$\ln \left\{ \frac{P_i(\theta_n)}{1 - P_i(\theta_n)} \right\} = \theta_n - b_i$$

以多元選項之Rasch 模型而言，部份給分模式可表示為：

$$\ln \left(\frac{P_{nix}}{P_{ni(x-1)}} \right) = \theta_n - b_{ix}$$

其中， b_{ix} 為在試題 i 第 x 點（rating category）困難度， P_{nix} 及 $P_{ni(x-1)}$ 分別示計分點 x 和 $x-1$ 的機率。

Rasch 模型之評分具有許多尺度變數評分的優點，包括：

- 一、單一維度性：所有分數可以在同一個尺度座標上做比較。
- 二、局部獨立性：當影響測驗表現的能力被固定不變時，受試者對任一試題上的反應，在統計學上而言是獨立的；簡單地說，這意謂著涵蓋在試題反應模式裡的能力因素，才是唯一影響考生在測驗試題上做反應的因素。
- 三、可再製性：其操作架構具有一定數學理論基礎，排除研究者之主觀給分造成之差異。
- 四、可驗證性：其模型可針對模式之配適進行統計探討，並提供對應之指標。

就Rasch 模型之參數而言，均為相對量尺所估計而得之結果。一般在教育與心理計量領域中，多透過平均難度之定錨，也就是將試題反應理論之試題平均難度設定為固定值，常見的預設值通常設定為0 logit；個別試題之難度與受測者的能力則按該預設值進行校估。難度與能力校估值亦以logit 為單位，利用羅吉特分析將順序分數轉換為洛基(logit)的類等距分數，使具可加性等良好性質，可提供更多訊息與更精確的數據表現。

3.3.1 Rasch 模式適合度 (Fit) 評估

Infit與Outfit統計值為Rasch模式用以檢測資料是否符合模式單一向度的期望。Infit均方 (MNSQ) 是表示接近受試者能力測量反應的非預期敏感度，Outfit均方是表示對試題太易或太難之非期望反應的敏感性指標 (Linacre, 1994) [37]。均方可轉置為t統計，以標準化Z值 (ZSTD) 表示出來，如符合模式的適合度，即呈t或標準常態分配。Infit是加權變異數 (σ^2) 後的標準化殘差平方加總，以公式 $\sum(z^2\sigma^2)/\sum(\sigma^2) = \sum(X - E)^2/\sum(\sigma^2)$ 表示，Outfit是傳統標準化平方的加總值，由標準化殘差平方 $z^2 = (X - E)^2/\sigma^2$ ，則Outfit是 $\sum(z^2)/N$ ，此處 X 為觀察值， E 為期望標準， σ^2 是模式期望變異變，而 N 為觀察值樣本個數。

當資料符合Rasch模式時，Infit (Zstd) 與Outfit (Zstd) 統計值均應介於可接受的範圍[38]，當題少及受試者較少時，可以觀察Infit與Outfit的ZSTD，當介於±2間時，可以認定符合Rasch的假設[39]，當樣本數大於500以上時，可以Infit MNSQ及Outfit MNSQ介於

0.6–1.4間為標準[40-42]，不符者，遂予除列。

3.3.2 單一向度與局部獨立性評估

當資料符合模式時，觀察分數減期望分數的殘差值 (residuals) 應是隨機分配[43]，主成份因素分析是驗證殘值中是否有主成份，而單一向度表示殘值中沒有主成份。單一向度 (unidimensionality) 在Rasch模式分析時，係經由主成份因素分析中，使用標準化後的殘差離差值 (dispersion of standardized residuals) 來估計，因素分析在Rasch轉換後分數殘值，顯示沒有主宰因素 (dominant factors) 存在。Smith & Miao (1994)經由模擬研究指出：「殘差因素 (residual factors) 的Eigenvalues小於1.4可認為具有隨機效果」[44]，Linacre (2004)更進一步指出：「第一因素解釋殘值變量小於3 logits者，表示具有單一向度」[45]。Rasch分析中對於鑑別度過高或過低的題目會被視為不好的題目，因為他們和其他試題相較下，鑑別度過高較不符局部獨立性假定，過低就不符測量同一個特徵的單向度。Infit太大表示該試題與試卷特徵關聯性較低，有違單一向度假設；Infit太小表示其關聯性較高，有悖局部獨立性 (local independence) 假設。

3.3.3 試題差異功能 (differential item functioning, DIF)

良好的測驗對所有的團體 (如男女、不同種族等) 應是公平的測驗，也就是說程度一樣的受測者對每道試題的答對率或通過率應該是一樣的，否則就表示題目對不同的團體有不同的功能(Holland & Wainer, 1993) [12]，亦即試題差異功能。也就是說，該試題對不同的團體，測量了不同的建構，同樣的分數在不同的團體內，代表了不同的意義，因此無法進行跨團體的比較，此類題目應與剔除，以確保測驗的公平性。一般對測驗公平性問題的看法認為：「在某個試題上，如果多數族群和少數族群的平均表現有所不同的話，該試題顯現出有DIF現象。」[46]。本研究意圖發展一個適用於男警與女警、警察各年齡層、及警員與警官的量表，因此題目在性別、年齡與階級上不可有試題差異功能。T檢定為檢定團體間是否具有試題差異功能[39,47]，本研究以團體間試題整體參數達到統計上0.05顯著差異，且其差異值在0.5 logit以上[37,48-49]，認定其具有試題差異功能，此類DIF試題刪除。

肆、結果分析

4.1 效度分析

表 2 Rasch 初測試題困難度與效度分析表

試題	困難度 (logit)	Infit MNSQ	Infit Zstd	Outfit MNSQ	Outfit Zstd
18	0.28	1.09	1.3	1.27	3.4
15	0.38	1.23	3.0	1.25	3.1
03	0.18	1.13	1.7	1.15	1.9
01	-0.35	1.07	1.0	1.02	0.2
02	-0.85	1.07	0.9	1.01	0.1

09	0.59	1.06	0.9	1.05	0.7
08	0.38	1.06	0.8	1.04	0.6
20	0.42	0.99	-0.1	1.04	0.5
14	0.45	1.03	0.5	0.99	-0.1
12	-0.49	0.99	-0.2	1.01	0.2
11	-0.46	1.01	-0.2	1.00	0.0
13	-0.6	1.01	0.1	0.98	-0.2
16	0.31	0.99	-0.1	0.98	-0.2
07	-0.84	0.96	-0.6	0.97	-0.4
05	0.33	0.92	-1.1	0.94	-0.9
04	0.17	0.94	-0.9	0.88	-1.5
19	0.54	0.90	-1.5	0.91	-1.3
10	0.05	0.90	-1.4	0.84	-1.9
06	-0.08	0.89	-1.4	0.87	-1.6
17	-0.39	0.84	-2.3	0.85	-1.9

本研究使用WINSTEP軟體[50]進行Rasch模式分析，其問卷信度，亦即樣本信度(Person reliability)為0.89(相似於Cronbach's α)。IRT假定測驗需符合單一向度及局部獨立性，表2為進行初始Rasch分析，各試題困難度結果，顯示除了第15、17、及18題等3題外，其餘15題(85%)試題的infit及outfit均落於 ± 2 可接受範圍內，而infit及outfit的MNAQ均落於 1 ± 0.4 範圍內。

經刪除適合度不良3題後，再進行第二次Rasch分析，各試題的infit及outfit均落於 ± 2 可接受範圍內，而infit及outfit的MNAQ也均落於 1 ± 0.4 範圍內。因素分析顯示第一因素解釋2.7%的殘值變量，可謂此18個指標，構建出一有效單一向度資料模式。

4.2 問卷試題差異功能(DIF)分析

表3 DIF的t檢定

試題	t-tests			困難度參數 (parameter estimate)		
	1 vs. 2	1 vs. 3	2 vs. 3	group 1	group 2	group 3
Panel A: 職位 (警員、巡佐、主管)						
11	0.85	2.11*	1.23	-0.35	-0.47	-0.71
19	0.61	-2.32*	-2.49*	0.56	0.5	0.87
自由度: 1 vs. 2 tests=307; 1 vs. 3 tests=271; 2 vs. 3 tests=138.						
group 1: 警員; groups 2: 巡佐; groups 3: 主管; *p-value < 0.05						
Panel B: 工作經驗 (年資)						
11	-1.38	0.63	2.00*	-0.23	-0.05	-0.31
自由度: 1 vs. 2 tests=241; 1 vs. 3 tests=233; 2 vs. 3 tests=236.						
group 1: 0-10年; groups 2: 11-20年; group 3: 大於20年; *p-value < 0.05						

經DIF分析不同年齡層間，並未顯示DIF，而在不同職位，即主管(包括分隊長)、

巡佐（包括小隊長）、及警員間，經t檢定後，顯著具有DIF題目為第11及19題等2題，第11題是「當違規人現場不認同我的處置，我可以和言悅色的對待以避免進一步的紛爭」，DIF顯示，警員認為處理現場紛爭的難度高於主管，但其差異值小於0.5 logit，因此第11題仍予以保留。第19題是「在嚴寒天候下，我仍能夠有效的執行取締闖紅燈勤務」，DIF顯示，巡佐及警員在嚴寒冷的天氣比主管更能有效的執行取締勤務，但其差異值小於0.5 logit，因此第19題仍予以保留。在工作年資比較方面，經t檢定後，顯著具有DIF題目為第11，但其差異值小於0.5 logit，因此第11題仍予以保留。

男警與女警間，在第11題顯示有DIF，女警的困難度（-0.05 logits）顯著（t-statistic=2.10, p-value < 0.05）高於男警（-0.56 logits），因此處理現場糾紛對兩男、女警察而言，有其分別的意涵，因此第11題予以刪除，如此所得的量表才能分別適用於男警與女警，以便進行性別的比較。

4.3 試題適合度檢測

經刪除第11、15、17、及18等4題目，由表4整體問卷信度（person reliability index）為0.89（類似Cronbach's α 值）[51]，試題信度為0.99（item reliability index），整體能力參數infit及outfit分別為-0.1及-0.2，MNAQ分別為1.01及0.99；題目難度參數infit及outfit分別為0及-0.1，MNAQ分別為1及0.99，結果符合Rasch模式的基本假設。

由表5針對16題試題檢測，各題目infit及outfit均落於 ± 2 可接受範圍內，而infit及outfit的MNAQ均落於 1 ± 0.4 範圍內，以上結果顯示，TPEAQ整體問卷、能力參數、及試題均已符合Rasch模式的單一向度、局部獨立性的假設，各試題適合度檢測也符合模式的基本假設，而圖3即為受試者能力分佈與題目困難度分佈線性圖。

表 4 Rasch 模式整體參數估計與適合度統計

		樣本數：358		試題數：16			
能力參數	測量值 (logit)	標準誤 (Standard Error)	Infit		Outfit		
平均值	0.14	0.29	IMNSQ	ZSTD	IMNSQ	ZSTD	
			1.01	-0.1	0.99	-0.2	
問卷信度 (Person Reliability) : 0.89				試題信度 (Item Reliability) : 0.99			
試題參數	困難度	標準誤 (Standard Error)	Infit		Outfit		
平均值	0.00	0.05	IMNSQ	ZSTD	IMNSQ	ZSTD	
			1.00	0.00	0.99	-0.1	

表 5 問卷指標難易度分析（由難至易順序排列）

試題	變數 / 問題	困難度	Infit Zstd	Outfit Zstd
09	我有能力駕車緊跟違規逃逸車輛之後	0.62	1.2	1.1
19	在嚴寒天候我仍可以有效執行取締闖紅燈勤務	0.54	0.5	0.3
14	執行取締勤務 2 小時，我不會腰酸、背痛。	0.45	1.4	0.8

20	在尖峰時間我也可以有效取締闖紅燈違規	0.42	0.5	0.8
08	我可以迅速的上巡邏車(機車),追緝逃逸車輛	0.39	0.4	0.3
05	只要我想要攔截逃逸車輛時,我都可以攔下來	0.35	-0.9	-0.6
16	執行取締闖紅燈勤務2小時,我不會覺累	0.30	1.5	1.2
03	我很容易可以查覺闖紅燈違規人是否有駕車逃逸意圖。	0.17	1.2	1.7
04	我有辦法記錄下逃逸車輛的車牌號碼。	0.16	-1.1	-1.8
10	我的EQ很好,不會被違規人激怒。	0.03	-1.6	-1.9
06	我可以安全的指揮闖紅燈車輛靠邊停車	-0.1	-2.0	-1.7
01	我可以找到取締闖紅燈的最佳位置	-0.38	0.1	-0.7
12	甚少闖紅燈違規人在現場執疑我的取締過程	-0.52	0.1	0.6
13	當我開出舉發單時,違規人很少提出申訴。	-0.64	0.2	-0.3
7	當違規人要棄車逃逸時,我可以快速將其攔下	-0.90	0.1	-0.8
2	一有闖紅燈違規行為,我很容易就可以立即發覺。	-0.90	-1.2	-0.5

伍、討論

本研究經由單一向度、局部獨立分析、模式配合度檢測及探討DIF,計刪除4題(20%)題目,餘留16題題目,經Rasch模式分析後,樣本能力參數在16題問卷與原20題試題問卷間相關性為0.997 ($p < 0.01$),表示刪除不適宜的變項後,與原始欲測度的工作能力,仍具有極高度相關性,也就是不會產生有意義的變化,仍保持原20題問卷欲測量執法能力的主要架構與構面。雖然最後的問卷變項較少,但卻能測量出潛在構面變項間真正的相關性,且各變項經DIF檢測後,在不同的團體間呈現最小變異。由於RASCH模型,將原始分數經由logit轉後,具有等距尺度的特性,除了具備分類和分級的特性外,以等距層次量度出來相同數量的差異,其相差的程度也是一樣,因此困難度參數與能力參數可以做有效的運算。

20個試題,刪除適合度不良的3個試題,17個變項再經DIF檢測後,第11、及第19題顯示,在不同階級的DIF表現上,雖因其差異未大於0.5 logit而予以保留,但仍具有顯著性,值得進一步探討。第11題:「當違規人現場不認同我的處置,我可以和言悅色的對待以避免進一步的紛爭」,DIF顯示,警員認為處理現場紛爭的難度高於主管,表示對員警應加強EQ情緒管理,並應施以處理民眾報怨應對進退的教育訓練,才能提升執法品質。第19題:「在嚴寒天候下,我仍能夠有效的執行取締闖紅燈勤務」,DIF顯示,巡佐及警員在嚴寒的天氣比主管更能有效的執行取締勤務,這與事實符合,因員警經常在各種天候下執行取締的工作,其抗寒執法的能力會比主管還好。不同性別間,第11題顯示有DIF並予刪除,女警的困難度(-0.05 logits)顯著高於男警(-0.56 logits),這有二種不同涵意,值進一步研究,首先是女警與男警對於處理現場紛爭難度的感受不同,而女警自我期許

的分數較高，也就是男警回答時給分較為保守，但實際上女警並未比男警能力還高；第二，女警處理紛爭的能力真的比男警還好，則女警當可與男警搭配共同執行取締闖紅燈違規勤務，以彌補男警EQ情緒管理不佳的缺點。

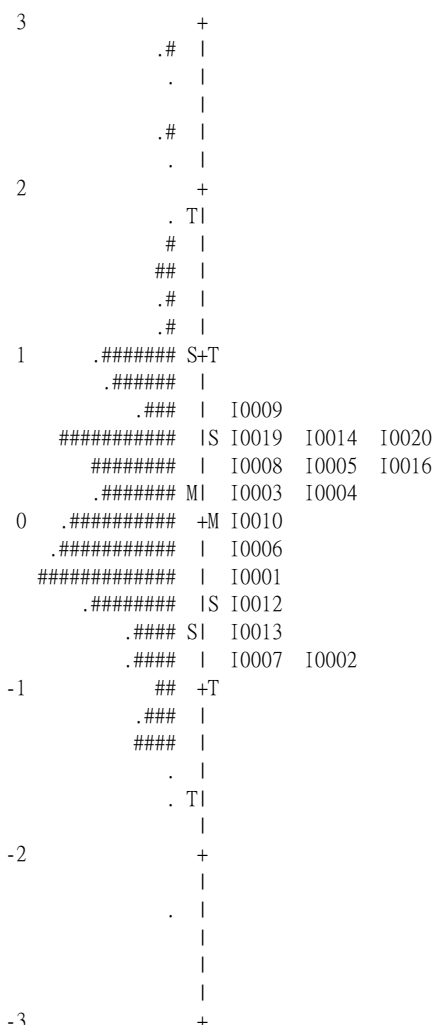


圖 3 受試者能力分佈與題目困難度分佈線性示意圖

本研究以警察階級替代警察的知識（包括教育與訓練），應用單因子變異數分析（one-way analysis of variance, ANOVA），進一步探討才能（competence）與工作年資（group 1: 0-10 年; group 2: 11-20 年; group 3: 大於20年）對於執法能力是否有顯著差異存在。分析結果，以不同階級（主管、巡佐、警員）而言，未達顯著水準，並無顯著差異（ $F(2, 355)=0.238, p=0.788$ ），表示不同階級並不會影響取締闖紅燈執法的能力，進一步而言，「才能」並不是響取締績的主因；但在不同工作年資上，顯示三種不同年資間存在顯著差異（ $F(2, 355)=3.969, p=0.02$ ）。經再採用Dunnett T3檢定，結果顯示，警察年資在「10年以下」者的執法能力，顯著高於「20年以上」的警察（ $t=2.807, p<0.01$ ），表示年齡是影響警察的執法能力的重要因素。由於工作年資越長，則警察年紀越大，警察年資大於20年以上者，年齡至少有40歲以上，而年資少於10年者，一般而言，年紀在35歲以下，這表示警察超過

40歲時，在路口執行取締闖紅燈的能力會有衰退的現象，顯示出，闖紅燈執法勤務是一項具生理負荷（physical load）的工作，越年輕的警察越能勝任，因此搭配年輕警察共同執法有其相互輔助的意義。

在男警與女警的差異上，女警的平均能力參數為0.38 logit，男警為0.12 logit，差異雖未達顯著，但對各別題目而言，仍值得進一步討論。男警認為第2題（一有闖紅燈違規行為，可以立即發覺）、與第7題（當違規人要棄車逃逸時，可以快速將其攔下）為最容易的項目，而女警認為第6題（我可以安全的引導闖紅燈車輛靠邊停車）、及第2題為最容易執行的項目。可見，女警與男警在查覺闖紅燈違規車輛方面，均足能適任。

另一方面，女警認為第10題（我的EQ很好，不會被違規人激怒）、第5題（只要我想要攔截逃逸車輛時，我都可以攔下來）、及第3題（我很容易可以查覺闖紅燈違規人是否有駕車逃逸意圖）為最難的題目，男警認為第9題（駕車緊跟違規逃逸車輛之後的能）、第19題（在嚴寒天候我仍可以有效執行取締闖紅燈勤務）、及第14題（執行取締勤務2小時，我不會腰酸、背痛）為最難執行的項目。可見女警最應加強EQ情緒管理及現場處理糾紛能力，而男警應加強體的訓練，另外由於男警認為跟車並非容易的執勤項目，而大部份駕駛巡邏車者以男警居多，因此，應針對男警加強安全駕駛訓練。綜合而言，女警應加強心智負荷能力（mental demands ability）與心理負荷（psychosocial demands ability），而男警則應加強心理負荷能力與生理負能力（physical demands ability）。

Rasch模式將平均試題困難度參數設為0 logit，以為等距尺度比較的基礎，由表4可得樣本平均能力參數為0.14 logit，表示大多數的警察對於闖紅燈執法勤務是能勝任的。由表5可知，平均而言，在TPEAQ試題中，超過一半（9題）的闖紅燈執法工作（即困難度參數大於0.14logit者），警察認為是較難執行的。警察認為第9題「緊追逃逸闖紅燈違規車輛，並成功的攔停」、19題「在極嚴寒的天氣執法」及第14題「執法2小時不會腰酸背痛」為最難達成的三項任務。第3「查覺闖紅燈違規人是否有駕車逃逸意圖」、4「記錄逃逸車輛車牌號碼」、5「攔截逃逸車輛的成功率」、7「快速反應阻止想要棄車逃逸的違規人」、8「迅速的上巡邏車（機車），追緝逃逸車輛」、9「駕車緊跟違規逃逸車輛之後」等6題為有關追緝逃逸違規人、車題目，除第7題外，其餘5題均為追緝逃逸違規車輛變項，其困難度參數均大於平均能力參數。此乃因執行闖紅燈時，警察一般站立於路口處，當違規車輛拒絕攔停而逃逸時，執勤警察若要追緝，必需先快速跑到巡邏車（機車），再發動引擎，執行追緝，因此反應與操作時間過長，違規車輛已有足夠時間逃逸，因此，對於闖紅燈違規逃逸車輛，應以記錄車牌，而予以逕行舉發為宜。由上述可知，為提高警察取締闖紅燈的執法能力，首應研擬針對逃逸車輛現場應變與處置的標準作業程序，以為執員警遵循，其次應加強體能的訓練，再者警察應加強適應嚴寒天候的取締勤務及執勤技巧。

由表4可知，在環境因素影響方面，天候嚴寒及交通尖峰時執法的困難度分別0.54 logit 與 0.42 logit，均大於平均能力參數0.14 logit，可見環境因素確為影響執法能力的重要因素。進一步而言，交通尖峰時間，警察在交通勤務的執行，應以路口交通疏導為原則，其執法的效果事倍功半。

為計算受測者的能力表現，本研究利用平方和指標(Derek & Wilson, 2003) DI (sum of squares indicator) 值[52]，以解釋整體的能力表現，DI如公式 (10) 所示：

$$DI_n = \sum_1^D (\bar{\theta} - \theta_{nd})^2 \quad (10)$$

其中D為構面數； $\bar{\theta}$ 為全體受測者所有構面能力值的平均數； θ_{nd} 為受測者n在構面d的能力值。當DI值低時，表示受測者各構面的能力值與平均數差異不大，當DI值高表示受測者至少有一個構面能力值與平均數差異很大。若將DI值定於 $DI_n = 0.5$ [52]，經檢測，在全部358個受測員警中，有75個員警（20.95%），執行取締闖紅燈行為的能力，顯著低於平均值，為提高執法績效，應針對這些員警應優先列為教育與訓練的對象。

陸、結論與建議

本研究應用Rasch模式發展一「警察執法能力問卷」(TPEAQ)，並據以評估警察取締闖紅燈執法能力感知，結果可做為日後研究交通警察執法能力的參考，並做為警察教育、訓練方向參考、與執法取擇。本研究結果可歸納如下：首先，女警自覺其取締闖紅燈能力比男警稍高；第二，除第11題「避免取締紛爭的能力」外，其餘變項均沒有顯著DIF；再者，女警最應加強的能力為心智與生理的負荷，男警首要加強的能力為生理負荷能力；在不同職務區分上，主管處理違規人抱怨的能力較強，而男警也較女警較容易處理違規人抱怨的情況；最後，為提升警察取締闖紅燈執法能力，未來交通警察的教育、訓練，應針對能力顯著不足的員警實施，並著重於安全駕駛訓練、增加體能負荷能力及強化適應不良天候的執法任務。

參考文獻

1. Retting, R. A., and Williams, A. F. (1996) Characteristics of red light violators: results of a field investigation, *Journal of Safety Research*, Vol. 27, 9-15.
2. Porter, B. E. and England, K. J. (2000) Predicting red-light running behavior: a traffic safety study in three urban settings, *Journal of Safety Research*, Vol. 31, 1-8.
3. Retting, R. A., Ulmer, R. G., and Williams, A. F. (1998) Prevalence and Characteristics of Red light Running Crashes in the United States. Insurance Institute for Highway Safety,
4. Retting, R. A., Williams, A. F., Preusser, D. F. and Weinstein, H.B. (1995) Classifying urban crashes for countermeasure development, *Accident Analysis & Prevent*, Vol. 27, 283- 94. Arlington, VA.
5. Porter, B. E., and Berry, T. D. (2001) A nationwide self-report survey of red-light running behavior: measuring prevalence, predictors, and perceived consequences, *Accident Analysis & Prevent*, Vol. 33: 6, 735-741.
6. Holland, B. K. (1989). "Motivation and Job Satisfaction." *Journal of Property Management*, 54(5) , pp.32-36.
7. Schmidt, FL; Hunter, JE & Outerbridge, AN (1986) The impact job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology* 71, 432-439
8. Libby, R. and Luft, July. (1993), Determinants of Judgment performance in accounting settings: Ability, Knowledge, Motivation, and Environments, *Organization and Sociality*, pp. 425-450.
9. Hunter, John E., Cognitive Ability, Cognitive Aptitudes, Job Knowledge, and Job Performance, *Journal of Vocational Behavior*, v29 n3 p340-362 Dec 1986
10. Aron R. KOLZ, Lynna. McFARLAND, Stanley B. SILVERMAN, Cognitive Ability and Job Experience as Predictors of Work Performance, *Journal of Psychology*, Vol. 132, 1998, page 539.
11. Campbell, J. (1993). *Creative art in groupwork*, Oxford. U. K.: Winslow Press. Rubin(ED.), *Approachs to art therapy – theory and techniques*, New York: Brouner/. pp. 92-112.
12. Holland, P. W., and Wainer, H. (1993) *Differential Item Functioning*. Hillsdale, NJ7 Lawrence Erlbaum.
13. Rasch, G. (1960) *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Denmarks Paedagogiske Institut.

14. Andersen, E.B. (1973). Conditional inference and models for measuring. Copenhagen ; Mentalhygiejnisk Forlag.
15. Tuomi K., Toikkanen J., Eskelinen L., Backman A-L., Ilmarinen J. and Järvinen E. (1991) Mortality, disability and changes in occupation among aging municipal employees. Scand J. Work Environ Health, Vol. 17(suppl, 1), 58-66.
16. Seitsamo, J., Ilmarinen, J. (1997) Life-style, aging and work ability among active Finnish workers in 1981-1992, Occupational : Psychology and Behaviour, Elsevier Science and Industrial Medicine, Vol. 37, No. 4 , 192.
17. Liira, J., Matikainen, E., Leino-Arjas, P., Malmivaara, A., Mutanen, P., Rytönen, H. and Juntunen, J. (2000) Work ability of middle-aged Finnish construction workers – a follow-up study in 1991–1995, International Journal of Industrial Ergonomics, Vol. 25, NO. 5, May, 477-481.
18. Ilmarinen, J., and Tuomi, K. (2004) Past, present and future of work ability, People and Work., Finnish Institute of Occupational Health, Research Reports, Vol. 65, 1-25.
19. Bugajska, J., and AAastowiecka, E. (2005) Life style, work environment factors and work ability in different occupations, International Congress Series, 1280, Complete, June, 247-252.
20. Tuomi K., Toikkanen J., Eskelinen L., Backman A-L., Ilmarinen J. and Järvinen E. (1998) Work ability index, 2nd revised edition, Finnish Institute of Occupational Health, Helsinki.
21. Markku, T., et al. (2002) Physical and psychosocial prerequisites of functioning in relation to working ability and general subjective well-being among office workers ,Tuulikki Sjögren-Rönkä, M.Sc., Department of Health Sciences, University of Jyväskylä, Finland.
22. Sluiter, J. K., (2006) High-demand jobs: Age-related diversity in work ability? Applied Ergonomics, Vol. 37, No. 4, July, 429-440.
23. Dijk, F. J., Meijman, T. F. and Ulenbelt, P. (1988) Towards a dynamic model of exposure, susceptibility and effect, International Archives of Occupational and Environmental Health, Suppl 1,144-150.
24. Ylikoski M, Ilmarinen J, Huuhtanen P, Punnonen O (2005). Age Managment Training -tool.European Network of Workplace Health Promotion, Essen, www.enwhp.com/toolbox
25. Ilmarinen J, Tuomi K, Seitsamo J. (2005). New dimensions of work ability. In Costa G, Goedhard W, Ilmarinen J (eds.) Assessment and Promotion of Work Ability, Health and Well-being of Ageing Workers. Amsterdam: Elsevier, pp. 3–7.
26. Nicholls, J. G. 1984. Achievement motivation: Conceptions of ability, subjective experience, task

- choice, and performance. *Psychological Review*, 91(3): 328-346.
27. Groeger, J.A. and Brown, I.D. (1989). Assessing one's own and others driving ability: Influences of sex, age, and experience. *Accident Analysis and Prevention* 21:155-168.
 28. Robert S. Kaplan., & David P. Norton 著，「策略核心組織：以平衡計分卡有效執行企業策略」，臉譜文化出版，台北，2001。
 29. Shinar, D., (1998) Aggressive driving: the contribution of the drivers and the situation, *Transport. Research: Part F : Psychology and Behaviour*, Vol. 1: 2, 137-160.
 30. Chang, H. L., and Shih, C. K. (2007). How do the traffic police perceive their ability for red light running enforcement ? -An application of the Rasch measure, *Journal of the Eastern Asia Society for Transportation Studies*, Vol.7, 2007
 31. Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Lawrence Erlbaum.
 32. Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores* (chapter 17-20) Reading, MA: Addison-Wesley.
 33. Strong, D. R., Kahler, C. W., Green, R. L. and Schinka, J. (2005) Isolating a primary dimension within the Cook-Medley hostility scale: a Rasch analysis, *Personality and Individual Differences*, Vol. 39, 21-33.
 34. Wright, B. D., and Masters, G. N. (1982) *Rating Scale Analysis*. Chicago: MESA.
 35. Andersen, E.B. (1973). Conditional inference for multiple-choice questionnaires. *British Journal of Mathematical and Statistical Psychology*, 26, 31-44.
 36. Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 357-74.
 37. Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.
 38. Linacre, J. M. (1994). *Many-Facet Rasch Measurement*. Chicago: MESA Press.
 39. Bond, T. G., and Fox, C. M. (2001) *Applying the Rasch Model: Fundamental Measurement in The Human Sciences*, Mahway, NJ: Erlbaum.
 40. Smith, R. M. (2004) Detecting item bias with the Rasch model, *Journal of Applied Measurement*, Vol. 5, 430-449.
 41. Linacre, J. M. & Wright, B. D. (1999). *A user's guide to Bigsteps/Winsteps*. Chicago: Mesa.

42. Wright, B. D., Linacre, J. M., Gustafson, J. E. and Martin-Loeffel, P. (1994) Reasonable mean-square fit values. *Rasch Measurement Transaction*, Vol. 8, NO. 3, 370.
43. Wang, W. C., & Chen, C. T. (2005). Item parameter recovery, standard error estimate and fit statistics of the WINSTEPS program for the family of Rasch models. *Educational and Psychological Measurement*, 65, 376-404.
44. Hsueh, I.P., Wang, W.-C., Sheu, C.F., Hsieh, C.L.(2004). Rasch analysis of combining two indices to assess comprehensive ADL function in stroke patients. *Stroke*, 35(3), 721-726.
45. Smith, R. M., and Miao, C. Y. (1994) Assessing unidimensionality for Rasch measurement. In M. Wilson (Ed.), *Objective Measurement: Theory into Practice*, Vol. 2. Norwood, NJ: Ablex, 314-327.
46. Linacre, J. M. (2004) WINSTEPS: Rasch Measurement Computer Program. Chicago: Winsteps.com.
47. 王文中, 錢才瑋、張文信、林宏榮、陳承德、劉歐, Rasch 分析在醫療界之應用, 聞道出版社, 2006年2月25日。
48. Draba, R. E. (1977) The identification and interpretation of item bias. MESA Memo, No. 25, [On-line] Available URL [http:// www.rasch.org](http://www.rasch.org).
49. Cohen, J. (1977) *Statistical power analyses for the behavioral sciences (Revised Edition)*. New York: Academic Press.
50. Wright, B. D., and Douglas, G. (1975) Best test design and self-tailored testing. MESA Memorandum No. 19, Department of Education, University of Chicago.
51. Linacre, J. M., and Wright, D. M. (1998) *A user's Guide to Winsteps, and Ministeps: Rasch Model Computer Programs*. Chicago, IL: Mesa Press.
52. Wright, B. D. (1996) Comparing Rasch measurement and factor analysis, *Structural Equation Modeling*, Vol. 3, 3-24.
53. Derek C. B., Wilson M. An introduction to multidimensional measurement using Rasch models. *Journal of Applied Measurement*: 2003; 4(1):87-100.

