發展適用於思覺失調症病人之多向度多途徑情緒辨識測驗

成冠緯

1. **研究計畫中英文摘要：**請就本計畫要點作一概述，並依本計畫性質自訂關鍵詞。
2. 計畫中文摘要。（五百字以內）

**背景與目的：**思覺失調症病人常有中至重度之情緒辨識損傷（難以正確辨識快樂、悲傷、生氣、厭惡、害怕、驚訝與平靜），限制其人際社交互動、社會參與以及生活品質。情緒辨識能力為極具潛力之介入目標，因其損傷可藉由心理社會訓練改善，並增進病人之人際社交表現。介入或追蹤病人情緒辨識能力的前提之一為需具備良好的評估工具。然而，目前常用評估工具卻有諸多限制，包含未完整評估7種基本情緒、未涵蓋3種情緒刺激物之感覺傳遞途徑（視覺、聽覺與整合）、施測費時、以及缺乏華人情緒刺激物等，嚴重影響可用性。電腦適性測驗能達成精準且快速之評估，故為理想之施測方式。本研究之目的為發展適用於思覺失調症病人之多向度多途徑情緒辨識測驗 (a multi-dimensional-multi-channel emotion recognition computerized adaptive testing for patients with schizophrenia, MER-CAT)，並驗證其應用於目標族群之心理計量特性。

**方法：**本研究將以3年五階段，完成MER-CAT之發展與驗證。

**階段一、發展MER題庫：**主持人將於「台灣地區華人情緒與相關心理生理資料庫」挑選合適情緒刺激物（含視覺、聽覺與整合途徑），並經專家會議以確認題庫之內容效度。**階段二、臨床測試：**測試MER題庫於500位思覺失調症病人與200位健康成人。**階段三、題庫驗證與發展MER-CAT算則：**以羅序模式分析 (Rasch model analysis) 檢驗題目適配度以及不同性別於特定項目是否具有「差異試題功能 (differential item functioning)」，再以驗證性因素分析確認建構效度，並藉模擬分析挑選快速且精準之項目組合，以最佳化MER-CAT之施測效能。**階段四、開發MER-CAT施測介面：**建立網頁為基礎之施測平台，以利後續應用與推廣。**階段五、驗證MER-CAT之心理計量特性：**測試MER-CAT於160位思覺失調症患者（急、慢性各半，共評2次，慢性間隔2週、急性4週），以驗證MER-CAT之心理計量特性。

**預期結果：** MER-CAT可能為國際間第一個多向度多途徑情緒辨識測驗，能克服常用工具之不足，提供完整（7種基本情緒）、詳盡（個別情緒分數）、周全（視覺、聽覺與整合途徑）、快速（<10分鐘）、精準（信度 ≥0.90）、有效（內容效度與建構效度等）、無性別差異試題功能，且能偵測病人情緒辨識功能改變之電腦測驗。因此，MER-CAT應可大幅改善情緒辨識之評估之效能，增進臨床與學術人員對患者情緒辨識能力之掌握與介入。

**關鍵字：**情緒辨識、文化差異、電腦化測驗、思覺失調症、羅序分析、社會認知

**三、研究計畫內容（以中文或英文撰寫）：**

1. 研究計畫之背景。請詳述本研究計畫所要探討或解決的問題、研究原創性、重要性、預期影響性及國內外有關本計畫之研究情況、重要參考文獻之評述等。如為連續性計畫應說明上年度研究進度。

**情緒辨識之定義、重要性、情緒類別與傳遞途徑**

情緒辨識 (emotion recognition) 是個體以互動對方之外顯行為（如臉部情緒或情緒聲調），以推測對方情緒狀態之能力。正確辨識他人情緒，有助於個體正確解讀他人之行為，進而做出合宜之行為反應，故為和諧人際互動之基礎。1,2 此外，情緒辨識為情緒管理歷程 (emotion processing) 之初始步驟，3 代表人們區辨他人有、無情緒之產生，以及該情緒之類別/分類的過程，故影響人們對於情緒之感知、學習與應用。綜上所述，良好的情緒辨識能力，是有效之人際關係與情緒管理/學習之基礎。

文獻探討之情緒種類繁多，但目前學界最常用之情緒類別為Ekman等人提倡之基本情緒 (basic emotion)。4 依據其跨文化之研究，不同文化間普遍存在7種基本情緒（快樂、悲傷、生氣、厭惡、害怕、驚訝與平靜）。4-6 這些基本情緒具有相似之表現方式（快樂時常嘴角上揚且語調輕快；生氣時則怒目瞪視且語氣不善）、觸發原因（快樂通常發生於有利於生存之事物，如享用美食或充分休息；生氣通常產生於生命或利益遭受侵害之情形，如面對攻擊者或搶奪），以及後續行為反應（快樂則增加行為頻率；生氣則勇於抵抗或消除危害），故有其獨特之演化意義（如評價周遭事務之好壞，以利於趨吉避凶）與社交價值（如快速且不受語言隔閡地解讀他人之行為）。4-6 有鑑於基本情緒具備跨文化且意義通用之特性，目前情緒辨識之研究大多以此7種基本情緒為主要架構。3

此外，情緒辨識亦有傳導途徑之別，通常分為3種辨識途徑（視覺途徑、聽覺途徑與整合途徑），7 以下詳述之。視覺途徑可能是人們最仰賴之情緒辨識方式，因為人們在社交過程中，絕大多數的時間皆在觀看他人之臉部表情，8 且於失去表情資訊時，人們較難以正確判讀他人之情緒狀態。9 視覺途徑常以靜態照片或無聲影片為情緒刺激物，適合用於環境嘈雜/聲音難以有效傳遞（如在市場或派對場合）時，或個體沒有直接參與互動的觀察情境（如觀察其他客人的表情，以判斷該餐點好吃與否）。聽覺途徑情緒辨識即傳統之情緒聲調辨識能力，通常以不帶有情緒意涵之錄音檔為情緒刺激處。其主導性雖不如視覺途徑強烈，但由於聽覺途徑情緒辨識對特定情緒（如害怕）較為敏銳，9 且部分情境難以使用視覺途徑情緒辨識（如使用電話溝通時，或他人在相鄰房間之呼喚），故仍具有獨特價值。整合途徑之情緒辨識可能是最生活化的情緒辨識能力，與日常人際互動等皆相關。整合途徑是綜合視覺與聽覺途徑之應用，理論上因其涵蓋最多資訊，故能提供最穩定有效之情緒辨識，通常以影音片段做為刺激物。此外，整合途徑之情緒辨識亦有助於人們察覺複雜的社交情境，如視覺與聽覺途徑提供的資訊不一致時（如表情開心但語帶哽咽），此差異即可能提供額外資訊，進而幫助人們進一步探查他人潛在的想法/感受。反之，若整合途徑之情緒辨識能力不佳，則人們可能難以協調視覺與聽覺資訊不一致之情形，進而誤導對他人行為之解讀。綜上所述，視覺、聽覺與整合途徑情緒辨識皆有其獨特之使用情境，故為情緒辨識不可或缺之面向。

**思覺失調症患者之情緒辨識損傷與影響**

思覺失調症 (schizophrenia) 患者常有情緒辨識之困難，影響其人際社交、社會參與以及生活品質。儘管目前尚無研究顯示此損傷具體之人數比率，但系統回顧與統合分析結果顯示，絕大多數的思覺失調症患者皆有中至重度之情緒辨識損傷，10-12 且其損傷受到情緒類別（如平靜與正向情緒通常較輕微，約為中度；負向情緒如生氣、厭惡及害怕通常較嚴重，相當於重度）、10 患者性別（如男性損傷程度通常較高，女性損傷程度較低），13,14 以及症狀嚴重度（急性期損傷程度較高，於症狀穩定後則減輕）等因素影響。1,2,15-17 此外，即便於症狀穩定之患者，其情緒辨識能力仍明顯低於健康成人，10,11,18-24 顯示情緒辨識損傷為患者長期難解的問題之一。再者，思覺失調症患者之情緒辨識損傷，是影響患者人際關係與社會功能之重要因素之一。23,25-31 具體而言，患者對他人情緒狀態之誤判，可能造成患者的社交行為失當（社交技巧不佳），進而造成患者難以維持人際關係；更甚者，若患者長時間處於上述狀態，可能造成患者缺乏參與社交活動之信心，影響其社會功能。32 因此，情緒辨識之損傷可被視為思覺失調症患者重要損傷之一，並對患者之人際互動與社會功能有顯著影響。

**情緒辨識評估對臨床與研究之價值**

思覺失調症患者之情緒辨識功能為精神復健主要成效指標之一，主要理由有二：其一、情緒辨識等社會認知功能是神經認知（neurocognition，如注意力及執行功能等）及功能性表現的中介因子 (mediator)，31,33 顯示神經認知對功能性表現之影響有部分是透過社會認知達成，支持評估與介入社會認知功能之重要性。此外，同時考慮神經認知與社會認知向度之模型，可大幅提升對功能性表現之解釋力。34,35 因此，評估並介入患者的情緒辨識損傷，是改善患者社會功能障礙的重點目標之一。其二、心理社會介入有助於改善患者之情緒辨識能力（如教導患者辨識各情緒之臉部表情與聲調特徵），且此進步能轉移至其人際社交功能（如減少對他人之敵意，以及更願意參與人際互動）。36 綜上所述，情緒辨識功能是極具潛力之介入目標，有助於改善思覺失調症患者長期難解之人際社交問題。

評估思覺失調症患者之情緒辨識能力，對臨床與研究人員之價值至少有三：其一、提供個別化治療或職場安置：臨床人員需瞭解患者之功能損傷情形與嚴重度（如患者不善於判讀哪些情緒，以及不擅長辨識之訊息途徑，如視覺或聽覺途徑），才能擬定合宜之介入或職場安置目標與計畫、提供個別化的復健治療、或選擇患者情緒辨識能力可勝任的職務進行訓練與安置，進而最佳化精神復健成效。其二、追蹤治療成效：臨床人員需良好之成效指標，以檢視所擬定之治療目標與計畫的成效，如此方能於治療成效不彰時，及時察覺並調整介入目標，以改善/維持療效。於長期而言，亦有助於臨床人員選擇治療成效較佳的方法，進而系統性地改善臨床介入之品質與效能。其三、探討影響因子：臨床與研究人員皆需完整、精準且有效的評估工具，以量化患者於各種情境下的情緒辨識能力（如辨識臉部情緒之能力、辨識情緒聲調之能力，或整合視覺與聽覺途徑之資訊，以辨識影音或現實情境中他人的情緒狀態），方能調查/驗證與其情緒辨識功能相關之影響因素，進而拓展對潛在機制之瞭解，或研發/改良目前的治療方法。綜上所述，臨床與研究人員皆需完整、精準、有效，且對反應敏銳之評估工具，方能有效掌握患者之情緒辨識能力與進展。此外，由於臨床業務繁忙，故除上述特性外，兼具快速且有效特性之測驗將格外適用於常態評估，以瞭解患者情緒功能損傷程度與進步情形。

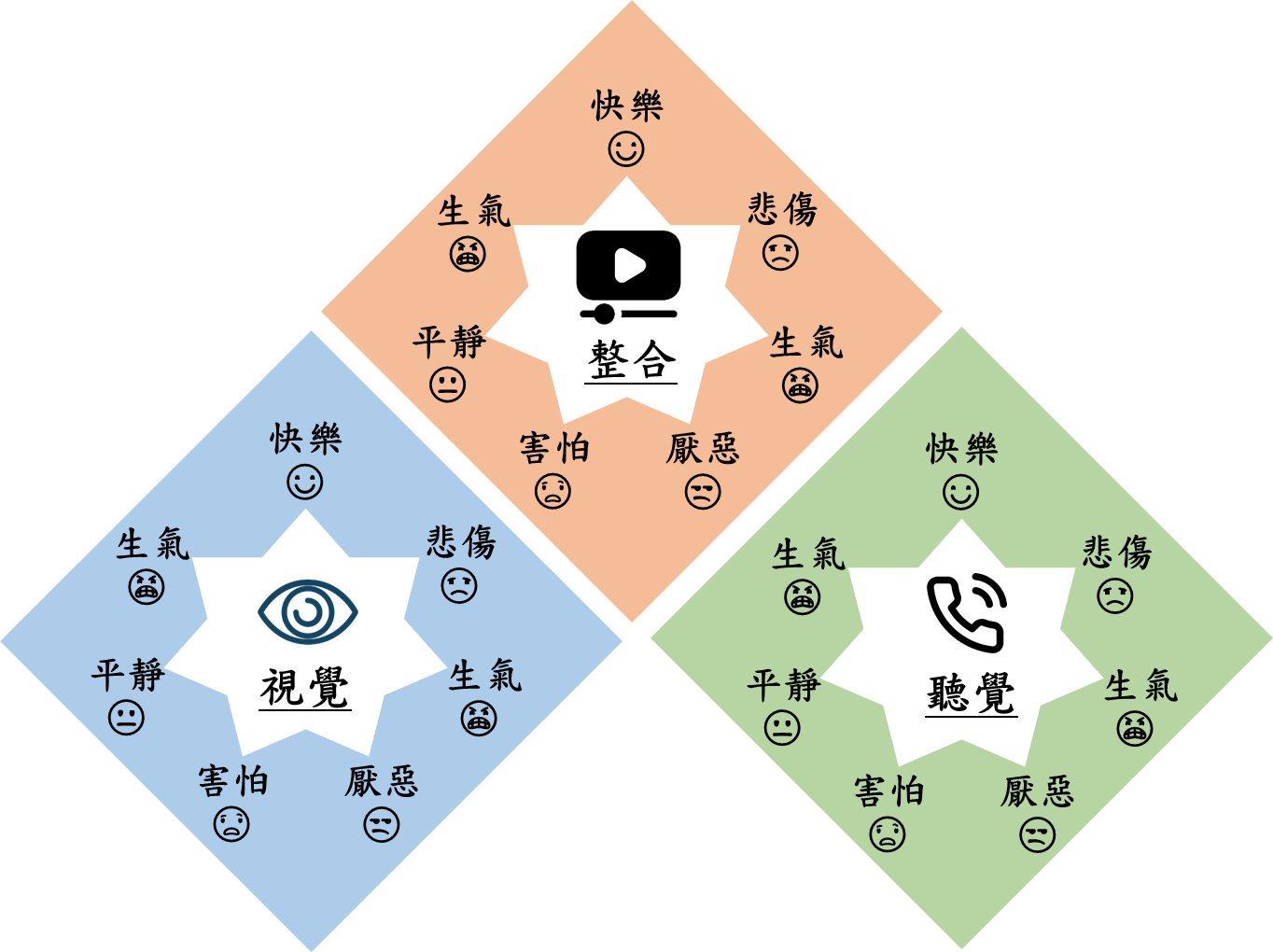
**理想情緒辨識測驗之架構與特性**

理想之情緒辨識測驗應包含7種基本情緒之辨識能力，並涵蓋3種主要訊息傳遞途徑（即視覺、聽覺與視聽整合；圖一）。此外，亦需使用華人之臉部表情或情緒聲調，並具備良好之可用性（題數少且易於操作）與心理計量特性（即信度、效度、反應性與無性別偏差），以下詳述之。

於基本情緒方面，由於各情緒皆有其獨特演化/社交意義，且思覺失調症患者辨識各種情緒的損傷程度不一（患者辨識厭惡、害怕跟驚訝達嚴重損傷；辨識悲傷與快樂則為中度損傷）40，故完整的情緒辨識測驗應完整涵蓋7種基本情緒，並呈現受試者於判讀個別基本情緒之能力，以幫助使用者周全瞭解患者的情緒辨識能力，從而避免低估其功能損傷情形。

於傳導途徑方面，由於3種辨識途徑皆有其獨特價值（如視覺途徑易辨識厭惡，聽覺途徑對害怕較敏銳，而整合途徑則因資訊豐富故最穩健）與適用情境（視覺途徑用於觀察、聽覺途徑用於通話，而整合途徑則用於日常人際社交），故完整之情緒辨識測驗應涵蓋3種傳導途徑，以利使用者掌握患者於各種途徑情緒辨識之優劣，進而瞭解其社交障礙之可能原因與預後，並規劃最適當之介入策略（如善於聽覺辨識者，可優先強化聽覺能力；如難以整合視聽資訊者，則加強其整合應用之能力）。

之前計畫「發展適用於思覺失調個案之臉部情緒辨識電腦適性測驗」主要為使用視覺途徑之臉部表情辨識，未包含聽覺途徑及整合途徑，因此無法全面評估個案之情緒辨識能力。例如：單純評量視覺辨識，可能無法正確評量現實社交情境中的弦外之音（皮笑肉不笑，或板起臉孔講冷笑話）或語音接觸（講電話）之情緒辨識能力。此外，因為日常生活大多數的社交情境（如朋友聚餐、與同事面談及視訊會議等）皆須同時使用此三種途徑之情緒辨識（會同時接收視覺及聽覺之情緒刺激物） 41,42，故須同時評估三種途徑之情緒辨識才最符合日常生活之情境，以提升測驗之生態效度。



圖一、理想情緒辨識測驗之架構（含7種基本情緒與3種訊息傳遞途徑）

於華人之臉部情緒與情緒聲調方面，儘管基本情緒普遍存在於不同文化間，但昔日研究仍指出人們的情緒辨識能力具備他族效應 (other race effect)，43 意旨人們於判讀與自己不同文化/族裔者之情緒時，其正確率較判讀與自己相同文化/族裔者低。44-46 此結果可能與不同文化間對情緒表達的方式略有不同有關。由於他族效應可能造成受試者情緒辨識測驗之表現偏差，進而錯估其情緒辨識能力，故理想之情緒辨識測驗，應使用與受試者相同文化/族裔之刺激物（臉部表情或情緒聲調）。

於可用性方面，由於臨床業務繁忙，題目多/施測時間長之測驗易造成施測負擔（如疲勞與缺乏動機等），進而不利於常態評估使用。因此，理想之評估工具應具備題目少、快速且易於操作之特性，最好能於10分鐘內完成，最多不超過15分鐘，以免造成過重之施測負擔。

於心理計量特性方面，由於精準、有效、能偵測功能改變，且不受患者性別影響，是成效評估工具之必備條件，亦為正確解讀評估結果之先決條件。因此，理想之評估工具設計應充分驗證前述心理計量特性，至少應以下條件：良好之羅序/再測信度（係數 ≥0.90）、微小之隨機測量誤差（估計標準誤 <平均分數之10%）、微小之練習效應（效應值*d* <0.20）、良好之因素效度（符合羅序模型，或因素分析之架構）、良好之收斂/發散效度（與相關效標呈現符合文獻所載之相關性）、良好之已知族群效度（know-group validity, 代表能區分健康人與思覺失調症患者之程度，各向度差異符合文獻所載數值，約為中至重度損傷）、良好之反應性（於團體層級能呈現情緒辨識功能之進步；於個別層級能呈現個別分數達顯著差異），以及不受患者性別影響之題目特性（即不具有試題性別差異功能函數，對男、女性患者難度相符，才能提供公平之比較）。

**常用情緒辨識測驗之簡介**

文獻中常見之11種情緒辨識測驗：Facial Expressions of Emotion–Stimuli and Tests (FEEST),47 Face Emotion Identification Test (FEIT),48 Penn Emotion Recognition Test (ER40),49 Voice Emotion Recognition Test (VERT),45 Voice Emotion Identification Test (VEIT),48 Florida Affect Battery (FAB),50 The Diagnostic Analysis of Nonverbal Accuracy 2 (DANVA2),51 Multimodal Emotion Recognition Test (MERT),9 Bell Lysaker Emotion Recognition Task (BLERT),52 Geneva Emotion Recognition Test (GERT)。53 另有國人於近期發展之臉部情緒辨識電腦適性測驗 (Computerized Adaptive Test of facial emotion recognition, CAT-FER)。54 各項工具之簡介與優劣分析詳述如下。

FEEST（視覺）

FEEST為Young等人所發展，47 共有60張黑白之白人照片，用以評估受試者辨識6種基本情緒（快樂、悲傷、生氣、厭惡、害怕及驚訝）之能力。FEEST之分數計算以答對題數為指標，並將結果分為整體分數（代表整體情緒辨識之能力，總分介於0至60分間）與個別情緒分數（代表辨識特定情緒之能力，滿分介於0至10分間）二部分，分數越高代表受試者情緒辨識能力越好。就計畫主持人所知，目前尚無文獻記錄FEEST之施測時間及其應用於思覺失調症患者之心理計量特性。

FEIT（視覺）

FEIT為Kerr等人所發展，55 共使用19張黑白之白人照片，以評估受試者辨識5種基本情緒（快樂、悲傷、生氣、害怕與驚訝）與1種特殊情緒（羞恥）之能力。FEIT之分數以答對題數計算，總分介於0至19分間，分數越高代表其情緒辨識之能力越好。一般而言，作答FEIT所需時間約為8分鐘。目前尚無證據顯示FEIT應用於思覺失調症患者之心理計量特性。

ER40（視覺）

ER40是由Kohler等人所發展，49 共納入40張性別、年齡與族裔均衡之彩色照片，以評估受試者判讀5種基本情緒（快樂、悲傷、生氣、害怕與中性/無表情）之能力。於ER40中，每種情緒之照片皆有情緒強度之分（高、低強度各4張），故能反應受試者對不同強度之辨識能力。ER40採用答對題數為指標，其總分介於0至40分間，分數越高代表受試者辨識情緒的能力越好。昔日研究顯示，完成ER40所需時間約為9分鐘。

ER40應用於思覺失調症患者之心理計量特性實證有限，主要來自單篇工具驗證與比較之文獻。56 整體而言，ER40具可接受之內部一致性 (Cronbach’s *α* = 0.75)、可接受之再測信度 (Pearson’s *r* = 0.71)、可忽略之練習效應 (Cohen’s *d* = 0.05)，以及可接受之收斂效度（與社會功能效標呈現低至中度相關，Pearson’s *r* = 0.17–0.41）。56

VERT（聽覺）

VERT為Kucharska等人所發展，57 包含35段由同一位白人男性專業表演者所錄製之錄音（內容語意不涉及情緒相關詞彙，故屬中性語意），用以評估受試者判讀7種基本情緒（快樂、悲傷、生氣、厭惡、害怕、驚訝與平靜）之能力。VERT之計分以答對題數為指標，其總分介於0至35分間，分數越高代表情緒辨識能力越好。完成VERT約需20分鐘。目前尚無證據顯示VERT應用於思覺失調症患者之心理計量特性。

VEIT（聽覺）

VEIT為Kerr等人所發展，43 內含21段由6位大學生（男女各半）所錄製之錄音（內容為中性語意），以評估受試者辨識5種基本情緒（快樂、悲傷、生氣、害怕與驚訝）與1種特殊情緒（羞恥）之能力。VEIT亦採用答對題數為指標，其總分介於0至21分間，分數越高代表受試者之情緒辨識能力越好。作答VEIT所需時間約為7分鐘。但目前尚無學者驗證VEIT應用於思覺失調症狀患者之心理計量特性。

FAB（視覺與聽覺）

FAB是由Bowers等人所發展，50 涵蓋多種途徑之情緒辨識測驗，如臉部情緒辨識、情緒聲調辨識與二者結合（即以觀看表情挑選對應聲調，或聆聽聲調配對對應表情）。於臉部情緒與情緒聲調辨識向度中，FAB包含20張黑白照片與20段中性語意之錄音，以分別評估受試者於視覺途徑與聽覺途徑辨識5種基本情緒（快樂、悲傷、生氣、害怕與中性）之能力；於二者結合之向度則分別有20題，以評估受試者觀看表情挑選對應聲調，或聆聽聲調配對對應表情之能力。FAB之計分皆以答對題數計算，各向度總分介於0至20分間，分數越高代表受試者辨識情緒之能力越好。作答FAB之情緒相關向度約需20分鐘。目前尚無證據顯示FAB應用於思覺失調症患者之心理計量特性。

DANVA2（視覺與聽覺）

DANVA2為Nowicki等人所發展，58 涵蓋多種非語言向度，包含臉部情緒辨識（視覺途徑）與情緒聲調辨識（聽覺途徑）。於情緒辨識相關向度中，DANVA2包含24張彩色照片與24段中性語意之錄音，以分別評估受試者於視覺途徑與聽覺途徑辨識4種基本情緒（快樂、悲傷、生氣與害怕）之能力。DANVA2採用答對題數計分，故臉部情緒辨識與情緒聲調辨識向度總分皆介於0至24分間，分數越高代表受試者於特定途徑（視覺或聽覺途徑）情緒辨識之能力越好。完成DANVA2相關向度所需時間約為16分鐘。但目前尚無學者驗證DANVA2應用於思覺失調症患者之心理計量特性。

MERT（視覺、聽覺與整合）

MERT為Ba¨nziger等人所發展，9 涵蓋多種途徑之情緒辨識測驗，包含視覺途徑、聽覺途徑，與整合途徑。具體而言，MERT包含30張黑白照片（視覺-靜態圖片）、30段無聲音之臉部表情影片（視覺-動態影像）、30段中性語意之錄音（聽覺），以及30段包含語音之表情影片（整合），以檢驗受試者判讀4種基本情緒（快樂、悲傷、生氣與厭惡）之能力。MERT以答對題目計分，並分為整體分數（即120題分數加總）與個別向度分數（即個別情緒之答對題數），總分分別介於0至120分與0至30分間，分數越高代表受試者之情緒辨識能力越好。完成MERT約需45分鐘。目前尚無證據顯示其應用於思覺失調症患者之心理計量特性。

BLERT（整合）

BLERT是由Bryson等人所發展，其包含21部白人男性演員上半身且帶有聲音之短片（每片約10秒），並要求受試者於觀看短片後，判斷該演員所表現之最接近於7種基本情緒中的哪一種。BLERT之單題以答對/答錯計分，總分介於0至21分間，分數越高代表受試者情緒辨識之能力越好。一般而言，BLERT之施測時間約為10分鐘。

就計畫主持人所知，BLERT應用於思覺失調症患者之心理計量特性實證有限，僅有一篇研究顯示其具備可接受之內部一致性 (Cronbach’s α = 0.78)、可接受之再測信度 (Pearson’s *r* = 0.81)、微小至中等的練習效應 (Cohen's *d* = 0.40)，以及大致可接受之收斂效度 （與社會功能相關效標呈現適度相關，Pearson’s *r* = 0.21–0.42）。56

GERT（整合）

GERT為Schlegel等人所發展之整合途徑情緒辨識測驗，53 其包含83段包含聲音與表情之影片，以評估受試者判讀6種基本情緒（快樂、悲傷、生氣、厭惡、害怕與驚訝）之能力。由於GERT是以羅序分析發展而成（有別於前述測驗皆使用傳統測驗理論），故受試者於各題目之作答表現會經羅序分析（依題目難度與受試者答對/答錯之情形），轉換成羅序分數（一種標準分數，數值介於正負無限大間，數值越大代表能力越好），並以分數越高代表受試者情緒辨識之能力越好。完整作答GERT約需20分鐘。但目前尚無學者驗證GERT應用於思覺失調症患者之心理計量特性。

CAT-FER（視覺）

CAT-FER是國人所自行研發之情緒辨識測驗，其總題庫包含165張華人之臉部表情辨識，涵蓋7種基本情緒（快樂、悲傷、生氣、厭惡、害怕、驚訝，以及平靜）。測驗以隨機順序且每次呈現一張之方式，呈現表演者之表情照片及情緒選項，並記錄其作答反應與時間。此外，因CAT-FER為電腦適性測驗，可依據受測者作答表現，挑選符合其能力水準之題目。CAT-FER為經由羅序模型發展驗證而成，故受測者於各題目之作答表現會經羅序分析轉換成羅序分數，以分數越高代表受測者情緒辨識之能力越好。CAT-FER依據其不同模式施測時間約為11–20分鐘。CAT-FER目前應用於思覺失調症患者之心理計量特性實證僅有一篇，54 研究顯示其總題庫具備可接受之羅序信度 (average Rasch reliability = 0.73–0.84)。此外，CAT-FER之精確模式和快速模式亦具有可接受之羅序信度 (average Rasch reliability = 0.73–0.81 vs. 0.69–0.72)。

**常用情緒辨識測驗之評論**

儘管目前已有多樣評估工具，但這些工具主要缺點有七（表一），嚴重影響其評估效能與結果解讀，以下詳述之。

其一、評估情緒向度不完整：絕大多數評估工具僅包含4至6種基本情緒（常為快樂、悲傷、生氣與害怕），僅有VERT、BLERT與CAT-FER完整涵蓋7種基本情緒。然而，由於思覺失調症患者判別不同情緒之損傷程度有別（如中性與正向情緒較輕微；負向情緒較嚴重），10 若使用評估情緒不完整之工具，可能錯估/遺漏患者辨識特定情緒（如厭惡及害怕）之損傷與其嚴重度，進而誤導使用者對於評估結果之解讀。

其二、計分指標不詳盡：大多數之測驗採用單一總分（如總答對題數）為指標，僅FEEST、FAB、MERT、GERT與CAT-FER可呈現7種基本情緒之個別分數。然而，單一總分雖能代表受試者之整體情緒辨識能力，卻無法呈現其辨識特定情緒表現之優劣（此資訊可為臨床介入之指引與成效評估，如發現患者不善判讀厭惡，可加強其捕捉厭惡特徵之能力，並以此追蹤介入成效），故能提供之資訊有限。此外，受試者辨識不同情緒之表現可能相互抵銷（如善於辨識快樂，但不善辨識厭惡者，其總分應靠近中間值、無明顯好壞之別），進而掩蓋患者於情緒辨識之優勢/損傷。

其三、評估感覺傳遞途徑不周全：絕大多數工具僅評估情緒刺激物「特定感覺傳遞途徑」之情緒辨識能力（如純視覺或純聽覺），除MERT外，無其他工具同時評估受試者於3種途徑（視覺、聽覺與整合）情緒辨識之表現。然而，人們於不同途徑之情緒辨識表現有別（如於整合途徑之表現通常優於純視覺或純聽覺途徑），且個別傳遞途徑皆有其獨特之應用價值/情境（如講電話之情境僅能仰賴聽覺途徑），若使用未周全評估3種感覺傳遞途徑之工具，可能導致評估結果過於偏頗，進而侷限對受試者情緒辨識能力之瞭解。

其四、缺乏華人情緒刺激物：絕大多數之測驗皆以白人之表情或聲調為情緒刺激物，僅有ER40與CAT-FER包含華人之表情照片。然而，情緒辨識受文化/族裔因素影響（亦即辨識者與被辨識者若屬於相同文化/族裔，其情緒辨識之正確率高於二者屬於不同文化/族裔者）。若使用不同文化/族裔之表情或聲調為刺激物，可能低估受試者之情緒辨識能力，且其結果亦可能與受試者於日常生活之實情不符（評估判讀白人情緒之能力，但國人大多非白人），故生態效度不佳。

其五、心理計量特性未知或不佳：絕大多數之評估工具缺乏應用於思覺失調症患者之心理計量特性實證，僅有ER-40、BLERT與CAT-FER曾被驗證並具備可接受之信度（信度係數約為0.70至0.80，CAT-FER平均羅序信度為0.73–0.84）與效度（與社會功能呈現中度相關，CAT-FER以羅序分析驗證建構效度）。由於心理計量特性是確認評估工具品質之指標，未知之心理計量特性，意即缺乏實證證實評估結果之穩定性（信度）、效度（正確性）與偵測病患之變化（反應性），可能誤判受試者情緒辨識之能力。

其六、未校正性別「差異試題功能 (differential item functioning, DIF)」偏差：絕大多數之工具並未驗證其試題有無性別DIF（即受試者性別對於評分結果之影響），僅CAT-FER驗證與校正性別DIF。由於情緒辨識可能受性別因素影響（如生活經驗與情緒表達習慣等），可能導致相同題目受到表現方式或情境熟悉度等因素影響，對男或女性受試者特別容易/困難，進而導致相同能力水準之男、女性受試者卻有不同的評估結果/分數，影響測驗之公平性。此外，試題之性別DIF可能是造成昔日研究結果不一致之原因，例如過往研究發現女性之情緒辨識能力通常優於男性，但結果不一致，即可能是所用工具具有不同程度之試題性別偏差（如部分對男/女性有利）所致。然而，缺乏試題性別DIF之檢驗，則難以排除其對評估結果之影響，進而侷限跨性別情緒辨識之研究。

其七、施測效率不佳：多數工具之題目多/施測時間長（>15分鐘），嚴重限制這些工具之可行性。於前述工具中，僅有FEIT, ER40, VEIT與BLERT因題數少，能於10分鐘內完成，但由於題數與信度為權衡關係（即題數多則測量資訊足，故評估結果能趨於穩定；題數少則資訊缺，評估結果較不穩定，故信度偏低），故這些測驗之信度表現普遍不佳。此外，前述信度與題數之權衡問題隨著評估向度數增加而加劇，故多向度測驗常需更多題數，但導致更長的施測時間，故不利於常態應用。

**情緒辨識測驗之文獻回顧總結**

綜上所述，目前文獻中常用之情緒辨識工具難以提供情緒向度完整、計分指標詳盡、評估感覺傳遞途徑周全、適用於華人、精準、有效、能偵測差異、不受性別偏差影響且快速之多向度多途徑情緒辨識測驗，嚴重影響其評估效能。於臨床方面，缺乏良好之評估工具，將難以辨識患者之問題與損傷程度，進而不利於制訂個別化治療目標，並定期追蹤以確保治療成效。於研究方面，缺乏良好之評估工具，則不易探究情緒辨識損傷之影響因素，進而不利於驗證其潛在機制，以及發展/改良介入模式。因此，發展兼具前述理想特性之情緒辨識測驗，將有助於提升臨床與學術評估之效能。

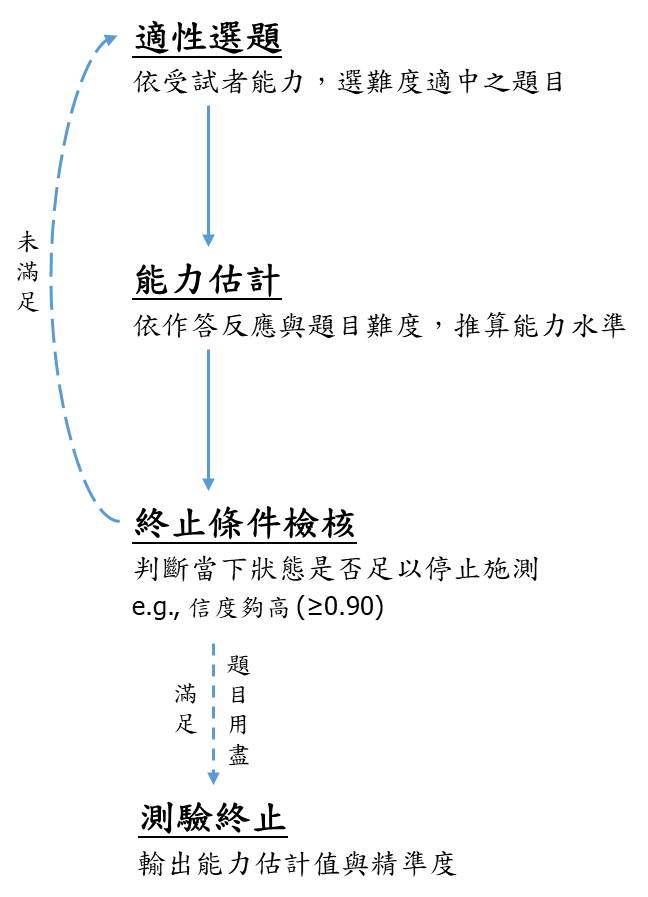
**電腦化適性測驗之簡介與優勢**

電腦適性測驗 (computerized adaptive testing, CAT) 為理想之測驗形式，因其具備適性選題之特性，能挑選並僅施測與受試者能力最相符之題目（個別化選題/施測，對每位受試者不過於簡單/困難），故能克服信度與施測效率權衡關係的侷限，可提供快速且精準之評估。59,60

簡言之，CAT之施測流程大致可分為三階段（圖二）：一、適性選題：CAT依據受試者之能力估計值（此數值受到受試者作答各題之表現以及這些題目之難度影響而變化），自題庫中挑選能提供最多測量訊息（可最大幅度提升測驗之信度/精準度）之題目（通常為難度與受試者能力估計值最接近的題目），再由受試者作答該挑選題目。二、能力估計：CAT依據受試者於測驗中作答之所有題目，以及這些題目之難度參數，彙整判斷受試者最可能的能力估計值，以及該能力估計值之精準度（即估計標準誤或個別化信度）。三、終止條件檢核：CAT依據受試者目前的能力估計值與精準度是否滿足既定條件（通常於發展CAT時即會預設測驗將於何時結束，常見標準如信度夠高，如 ≥0.90；或多施測對於能力估計之精準度無益，如信度增加量低於0.001），若滿足該條件，則停止測驗，並輸出評估結果（包含分數與信度）；反之，則回到第一階段，繼續挑選下一題，直到終止條件被滿足，或所有題目皆被作答為止。因此具備良好題庫CAT之施測效能高（快速且精準）。

CAT尤其適用於多向度測驗，因其接受向度間之資訊彼此共享（亦即受試者作答某一向度題目之表現，除能用以估計其於該向度之能力外，亦可透過向度間之相關性，以估計受試者於另一向度之能力水準），故能大幅減少評估所需之題目。如昔日研究即顯示相較於單向度之CAT僅能減少20%至50%之題數，若發展為多向度CAT則可減少50%至80%之題目。60 因此，CAT尤其適用發展情緒辨識等向度數較多之測驗，以更節省評估時間，但不犧牲施測結果之精準度。

圖二、CAT之架構/流程圖



**研究目的**

本研究之目的有二：其一、發展適用於思覺失調症患者之多向度多途徑情緒辨識電腦適性測驗 (A multi-dimension-multi-channel emotion recognition computerized adaptive testing for patients with schizophrenia, MER-CAT)，以提供完整、周全、資訊充足、快速、準確、有效、能呈現功能改變、適用於國人且無試題性別偏差之情緒辨識測驗。其二、完整驗證MER-CAT之心理計量特性，包含再測信度、練習效應、建構效度、收斂/發散效度、已知族群效度，以及反應性。

**創新與價值**

MER-CAT可能為國際間第一個多向度多途徑情緒辨識之電腦適性測驗。

MER-CAT可提供完整（7種基本情緒）、詳盡（個別情緒分數）、周全（視覺、聽覺與整合途徑）、快速（<10分鐘）、精準（信度 ≥0.90）、有效（內容效度與建構效度等）、無性別差異試題功能，且能偵測病人情緒辨識功能改變之電腦測驗。

MER-CAT可大幅改善情緒辨識之評估之效能，增進臨床與學術人員對患者情緒辨識能力之掌握與介入。

表一、文獻中常用的情緒辨識工具之限制

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 評估工具 |  | 情緒完整性 | | | | | | |  | 分數有效性 |  | 途徑完整性 | | |  | 文化適用性 |  | 心理計量特性a | | |  | 性別偏差 |  | 效率b |
|  | 快樂 | 悲傷 | 生氣 | 厭惡 | 害怕 | 驚訝 | 平靜 |  | 個別情緒 |  | 純視覺 | 純聽覺 | 整合 |  | 華人表情/聲調 |  | 信度 | 效度 | 反應性 |  |  |
| FEEST |  | **** | **** | **** | **** | **** | **** |  |  | **** |  | **** |  |  |  |  |  | ? | ? | ? |  | ? |  | ? |
| FEIT |  | **** | **** | **** |  | **** | **** |  |  |  |  | **** |  |  |  |  |  | ? | ? | ? |  | ? |  | **** |
| ER40 |  | **** | **** | **** |  | **** | **** |  |  |  |  | **** |  |  |  | **** |  | **** | **** | ? |  | ? |  | **** |
| VERT |  | **** | **** | **** | **** | **** | **** | **** |  |  |  |  | **** |  |  |  |  | ? | ? | ? |  | ? |  | **** |
| VEIT |  | **** | **** | **** |  | **** | **** |  |  |  |  |  | **** |  |  |  |  | ? | ? | ? |  | ? |  | **** |
| FAB |  | **** | **** | **** |  | **** |  | **** |  | **** |  | **** | **** |  |  |  |  | ? | ? | ? |  | ? |  | **** |
| DANVA2 |  | **** | **** | **** |  | **** |  |  |  |  |  | **** | **** |  |  |  |  | ? | ? | ? |  | ? |  | **** |
| MERT |  | **** | **** | **** | **** |  |  |  |  | **** |  | **** | **** | **** |  |  |  | ? | ? | ? |  | **** |  | **** |
| BLERT |  | **** | **** | **** | **** | **** | **** | **** |  |  |  |  |  | **** |  |  |  | **** | **** | ? |  | ? |  | **** |
| GERT |  | **** | **** | **** | **** | **** | **** |  |  | **** |  |  |  | **** |  |  |  | ? | ? | ? |  | ? |  | **** |
| CAT-FER |  | **** | **** | **** | **** | **** | **** | **** |  | **** |  | **** |  |  |  | **** |  |  |  |  |  | **** |  | **** |

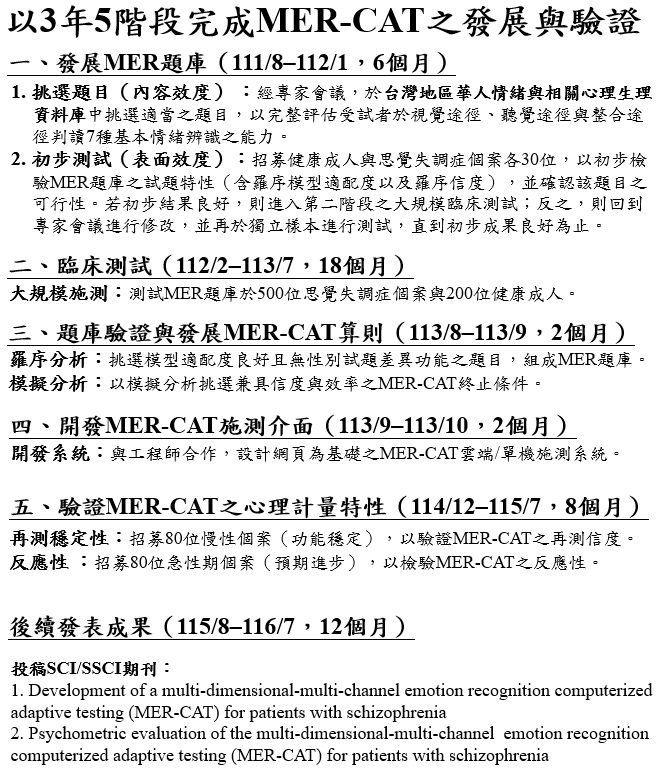
FEEST, Facial Expressions of Emotion–Stimuli and Tests; FEIT, Emotion Identification Test; ER-40, Penn Emotion Recognition Test; VERT, Voice Emotion Recognition Test; VEIT, Voice Emotion Identification Test; FAB, Florida Affect Battery; DANVA2, The Diagnostic Analysis of Nonverbal Accuracy 2; MERT, Multimodal Emotion Recognition Test; BLERT, Bell Lysaker Emotion Recognition Task; GERT, Geneva Emotion Recognition Test; CAT-FER, Computerized Adaptive Test of facial emotion recognition.

a於心理計量特性欄位中，代表良好，代表中等至可接受，代表不佳，?代表未知

b於效率欄位中，代表好（<10分鐘），代表中等至可接受（<20分鐘），代表不佳（≥20分鐘），?代表未知

1. **研究方法、進行步驟及執行進度。請分年列述：1.本計畫採用之研究方法與原因及其創新性。2.預計可能遭遇之困難及解決途徑。3.重要儀器之配合使用情形。4.如為須赴國外或大陸地區研究，請詳述其必要性以及預期效益等。**

本研究預計以3年5階段以發展與驗證MER-CAT（圖三），以下詳述之。



圖三、MER-CAT之發展、驗證與發表進度圖

1. **發展MER題庫（6個月）**

目的

本階段之目的為挑選適當之華人情緒刺激物（臉部表情、情緒聲調與溝通影片），並初步驗證其影音資料之品質，以做為MER-CAT之施測題目/題庫。

對象

本階段將招募健康成人與症狀穩定之思覺失調症患者各30人，以初步測試MER題庫之品質。

思覺失調症患者之納入條件有三：(1) 依據DSM-5或ICD-10診斷為思覺失調症；(2) 年滿20歲；(3) 可遵守3步驟指令。然而，若受試者曾有其他精神診斷（如憂鬱症與強迫症等）、曾有腦傷病史、曾有物質濫用（酒精與藥物）之問題、於近半年內調整藥物者，或無法完成所有測驗者，則予以排除。

健康成人之納入條件主要有二：(1) 年滿20歲；(2) 有意願參與本研究。然而，若受試者曾有任何精神診斷（如思覺失調症、憂鬱症或強迫症等）、曾有腦傷病史、曾有物質濫用（酒精與藥物）之問題，或無法完成所有測驗者，則予以排除。

流程

組成專家小組並確認情緒刺激素材來源：計畫主持人將召開專家會議（預計徵詢7位專家，並至少包含1位以上之精神科醫師、精神科職能治療師、臨床心理師與測驗發展專家），以挑選合適題目，做為MER-CAT之候選題目。計畫主持人預計自「**台灣地區華人情緒與相關心理生理資料庫」**中挑選適當素材，12 以做為MER-CAT之題目來源。該資料庫之發展是由梁庚辰與廖瑞銘等人為正副主持人，集合大量心理學門學者之努力與貢獻，並於國科會人文處之經費支持下，所發展之適用於本土文化的情緒刺激與測量模式資料庫。該資料庫共分為9個子資料集（專業表演者臉部表情、大學生臉部表情、情緒圖片、情緒詞、情緒相關隱喻或比方、引發情緒之敘述—笑話、情緒聲調、電影短片，以及音樂情緒特性），12 內容包含及大量之影/音與整合類型之情緒相關刺激物，並經過我國大學生初步評量/測試，以支持其刺激物之品質。故此資料庫之素材具備實證，極具潛力可用以發展適用於我國思覺失調症患者之情緒辨識測驗。

初步挑選情緒刺激物：於專家會議之前，計畫主持人將依據理想之情緒辨識測驗架構，先挑選性別（男女各半）、年齡（盡可能涵蓋不同年齡層）、情緒（涵蓋7種基本情緒）與難度（綜合考量情緒強度、特徵強弱、動/靜態媒介之分別，以及昔日大學生作答之正確率）均衡之華人刺激物，以做為MER題庫之候選。這些題目將挑選自前述資料庫中，包含專業表演者臉部表情及大學生基本情緒臉部表情常模資料庫中168張照片、基本情緒聲調中192筆純聲音及基本情緒聲調中246段動態影音。各傳遞途徑與資料庫之配對關係呈現如表二。研究者預期此資料庫之素材數量大，應可充分挑挑選出各種性別、年齡、情緒與難度配對之刺激物。

專家確認情緒刺激物/候選題庫：計畫主持人以電子郵件，將預先挑選之題目寄給與會專家，並請專家逐題/項目判斷這些初選題與理想架構之合適性，以做為題目挑選之依據。各題目需經過8成專家（亦即6位專家）同意，始納入為候選題；未達8成但超過6成專家同意者（亦即超過5人但未達6人），計畫主持人將再參考專家意見，並重新挑選或發展合適替代項目，並於後續會議經專家再次確認（標準與此階段相同）；未達6成同意度者則直接刪除。前述程序將重複直到所有題目皆已被納入或刪除為止。由於此階段已經過專家確認其題目挑選之合適性，故亦完成MER候選題目之內容/專家效度。

預試MER候選題庫：再來，計畫主持人將測試這些題目測試於思覺失調症患者與健康成人各30人，並於測試後進行簡單之認知訪談，以確認這些題目之測量品質、臨床可用性，以及表面效度。如有測量品質不佳之題目，則會回到專家會議重新討論，以確認題目之合適性；如有題目內容不當，需重新設計，則計畫主持人將另外邀請不重複之思覺失調症與健康成人受試者，以再確認這些題目之品質，直到所有候選題目皆表現良好，再進入後續大規模驗證。

題目呈現方式：MER題庫之電腦化施測介面委由一位電腦軟體程式設計師開發而成。該軟體以隨機順序且每次呈現一段情緒刺激物之方式，於電腦上呈現表演者之情緒刺激物，爾後畫面即會跳出7種情緒選項，由受測者點選所呈現的情緒類別為何。軟體自動記錄受測者作答反應與時間，並於結束測驗時呈現該物患者在各種情緒與各種感覺傳遞途徑的作答正確率。為避免受測者疲勞，該施測軟體於受測者每作答30題後，強制受測者休息至少10秒（實際休息時間由受測者自行決定，休息時間亦納入完整施測之時間內）。

表二、各途徑之刺激物與預計挑選之資料庫

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **途徑** | **刺激物類型** | **資料庫** | **發展階段題數** | **範例** |
| 視覺 | 靜態照片、  無聲音之表情影像 | 大學生基本情緒臉部表情資料庫61  專業表演者臉部表情常模資料62 | 168 | 純影像（靜態圖片） |
| 聽覺 | 情緒聲調錄音 | 基本情緒聲調63 | 192 | 純語音：「我早就知道了…」 |
| 整合 | 含聲音之影片 | 台灣地區華人情緒與相關心理生理資料庫—基本情緒聲調）64 | 246 | 動態影音（影像及聲音） |

工具

MER題庫，由計畫主持人初步挑選，經專家會議同意之情緒辨識刺激物。

1. **臨床測試（18個月）**

目的

測試MER候選題庫於大量思覺失調症患者與健康成人，以做為後續羅序分析與CAT模擬分析之數據。

對象

本階段預計招募500位思覺失調症患者（急性患者歧異性預計大，所需人數較多，預計300人；慢性患者症狀較為穩定，預計200人）與200位健康成人，以測試MER題庫之品質。

思覺失調症患者之納入條件有三：(1) 依據DSM-5或ICD-10診斷為思覺失調症；(2) 年滿20歲；(3) 可遵守3步驟指令。然而，若受試者曾有其他精神診斷（如憂鬱症與強迫症等）、曾有腦傷病史、曾有物質濫用（酒精與藥物）之問題，或無法完成所有測驗者，則予以排除。

健康成人之納入條件主要有二：(1) 年滿20歲；(2) 有意願參與本研究。然而，若受試者曾有任何精神診斷（如思覺失調症、憂鬱症或強迫症等）、曾有腦傷病史、曾有物質濫用（酒精與藥物）之問題，或無法完成所有測驗者，則予以排除。

流程

於簽署施測同意書後，同意參與本研究之受試者將被安排至安靜的評估室內，以進行後續測試。依據實際挑選題目之數量，受試者將以多次施測方式（為免疲勞，每次作答以不超過30分鐘為限制；並於過程中受試者可視需求予以休息），於2天內，完成所有MER題目之評量。

計畫主持人將記錄前述施測過程中各階段所花費的時間（精確至每題作答之秒數），但不會限制受試者之作答時間，亦不會要求患者越快越好。

於完成前述MER後，主持人將再收集受試者之基本資料（如年齡、性別、教育程度、職業別）與臨床資料（僅限於病人組，包含其認知功能、症狀嚴重度、發病年齡以及發病時間）。

工具

MER題庫與前一階段相同，此不贅述。

蒙特利爾認知評估 (Montreal Cognitive Assessment, MoCA) 為研究常用之認知篩檢工具，可粗略地呈現受試者之認知功能程度。62 MoCA共有30題，涵蓋8種常見之認知向度（專注力與集中力、執行功能、記憶力、語言能力、視覺空間建構、抽象概念、及計算與定向感），並以答對題數為計分依據。MoCA總分介於0至30分間，分數越高通常反應受試者沒有認知相關之問題。MoCA應用於思覺失調症患者具備良好之心理計量特性，包含良好之再測信度 (intraclass correlation coefficient = 0.81)與良好之同時效度（與Mini-Mental State Examination具中高度相關，Pearson’s *r* = 0.62)。55,63

臨床整體印象量表-嚴重度 (Clinical Global Impressions-Severity, CGI-S) 為研究常用之症狀嚴重度評估工具。65 CGI-S僅有1題，以7點量尺呈現受試者近2週內受精神症狀影響之程度，並以7分代表非常嚴重（為評估者看過最嚴重者），1分代表輕微（幾乎沒有症狀）。CGI-S應用於思覺失調症狀患者具備良好之施測者間信度 (intraclass correlation coefficient > 0.70)，以及良好之同時效度（與Positive and Negative Syndrome Scale具備高度相關，Pearson’s *r* > 0.75）。66

1. **題庫驗證與發展MER-CAT算則（2個月）**

目的

藉羅序分析挑選合適之候選題（符合羅序模型且無性別偏差），以組成最終版MER題庫，並驗證其已知族群效度。

流程

計畫主持人將使用前一階段所收集之大量臨床資料（預計500位思覺失調症患者與200位健康成人），以羅序分析檢驗MER題庫之模型適配度羅序模型適配度以及不同性別於特定題目是否具有「差異試題功能(differential item functioning, DIF)」，刪除適配度不佳以及具有性別DIF之題目後，組成最終版MER題庫。此外，主持人亦將檢視最終版MER題庫之建構效度、羅序信度以及已知族群效度，以確認評估結果之信度與效度。

於確認MER最終版題庫後，主持人將以模擬分析方式，檢視在不同終止條件下（如信度 ≥0.90），可達成之平均信度、個別化信度（信度 ≥0.90, ≥0.80, ≥0.70之人數比率），以及所需之施測題數，並挑選兼具信度與施測效率之最佳終止條件，以完成MER-CAT之發展。

資料分析

羅序模型適配度將以均根值 (mean square, MnSq) 為指標。67 MnSq有二種類型，其一、遠方均根值 (outfit MnSq)，亦稱為加權之適配度指標，代表各種能力水準受試者與羅序模型預期之符合程度；其二、近方均根值 (infit MnSq)，亦稱為加權性之適配度指標，特別著重於能力與題目難度相符者之作答反應的適配度。此指標之數值介於正負無窮大間，並以數值1為期望值：數值大於1者代表不適配，即作答反應不符預期；反之，則為過份適配，代表受試者行為完全符合預期，故難以提供額外有效之評估資訊。67 計畫主持人將採用Wright及Linacre之建議，68 以二種適配度指標皆不超過1.4為有效適配之標準；本研究不刪除MnSq低於0.6（過度適配）之題目，乃因為這些題目仍可提供臨床上有意義之資訊，或有助於避免練習效應；此外，本研究將以CAT方式施測，故系統能自動挑選並排除資訊量有限之題目，不至於造成測驗過長之問題。羅序分析預計以ConQuest version 2軟體進行，並以最大概似估計法為參數估計方式。

良好之羅序模型適配度，亦代表一群題目符合個別單向度 (individual unidimensionality) 之特性。67 由於單向度是計分與分數解釋之必要條件（一群題目需測量相同構念，則這群題目之分數才得以加總/平均，以反應受試者於該構念或潛在特質 (latent trait) 之程度，否則該分數將無意義）。為進一步確認個別單向度之假設，計畫主持人再以驗證性因素分析 (confirmatory factor analysis, CFA) 檢驗各題目是否符合單向度結構。由於CFA之模型評估尚無單一黃金標準，故本研究將同時考慮4種模型適配度指標，以評估單向度是否被滿足：一、：為卡方值與模型自由度 (degree of freedom, *df*) 之比值，以反應於校正模型複雜度後，其整體適配度之情形。此比值越低，代表觀察模型與假設模型越相符，進而支持假設模型之有效性。一般而言，此指標通常以數值 <2為可接受之標準。69 二、Comparative Fit Index (CFI)：為假設模型與虛無模型（與假設模型相同，唯其題目與構念間毫無關聯）卡方適配度之比值，用以比較與隨機/最低模型適配度數值之差異。CFI越高，代表模型距離最差情形越遠，進而支持假設模型的有效性。CFI通常以 ≥0.95為良好，並以 ≥0.90為可接受。69 三、Standard root-mean-square residual (SRMR)：為標準化之殘差方均根值，乃以殘差角度檢視理論模型與觀察資料間的差異程度。SRMR數值越小，代表二者差異越小，進而支持模型之有效性。通常以SRMR < 0.08為良好之判斷標準。69 四、Root-mean-square estimation approximation (RMSEA)：其概念類似推論統計，乃以假設模型推論真實模型的差異程度。RMSEA越小，代表假設模型與真實模型之間差距越小，進而支持假設模型之有效性。一般而言，以RMSEA < 0.06為良好，並以 <0.08為可接受。69 唯有4項指標皆符合者才代表適配度良好，否則皆暗示該模型可能有某一方面之瑕疵，故為適配度不佳。CFA之分析將以EQS軟體進行。70

試題性別偏差將以羅序分析之DIF分析進行，以檢驗男、女性受試者於作答特徵是否有別。由於羅序模型中，受試者之作答反應僅受到題目難度與受試者能力之影響，故作答特徵之差別，即代表題目難度之差別，故可用以偵測各題目是否對於男、女性受試者較為困難或簡單，進而予以必要之調整（如刪除或予以參數校正）。主持人將使用Zieky之建議，71 以DIF值與其差異顯著性做為試題是否具有DIF之判定標準：DIF值代表特定試題對男性與女性受試者之平均難度差異，如此差異程度超過0.38，則代表二者具有明顯的差別，將誤導評估結果之解讀。

已知族群效度將以獨立樣本*t*檢定 (independent *t* test) 與效應值 (Cohen’s *d*) 為指標。*t*檢定主要用以檢驗思覺失調症患者之平均得分是否顯著地低於對照組（健康成人）；效應值則用以呈現前述平均數之差異程度，並以*d* ≥ 0.20為小，*d* ≥ 0.50為中，*d* ≥ 0.80為大。72

CAT之模擬分析，將納入近期CAT研究常用之終止條件，以最佳化MER-CAT之評估效能。目前研究中最常使用之終止條件有二：73 其一、最大所需信度值 (minimal reliability required, MRR)，代表測驗結束時所需之最低信度，常見標準包含MRR ≥0.90, MRR ≥0.80, MRR ≥0.70；其二、最低信度改變量 (limited reliability increase, LRI)，代表多施測一題後，信度之增加量，通常以LRI <0.001, LRI <0.005, LRI <0.010為指標。由於MRR與LRI個別有其適用情境，如MRR適合用於題庫數量充沛，但信度已滿足時，可及早結束測驗以維持效率；而LRI則用於多增加施測題目已難以有效提升測驗信度時，以及早結束測驗。本研究將綜合考慮此二種終止條件（即單獨使用MRR、單獨使用LRI，以及MRR與LRI任一條件滿足即停止CAT），以挑選最佳組合。

效能最佳之終止條件，應能使各向度之信度 ≥0.90，且施測時間應不超過10分鐘，並不超過15分鐘。如最佳終止條件難以達成，主持人將因應不同使用需求（如臨床人員與研究人員），分別挑選信度最佳（適用於臨床評估，需要精準評估結果，以進行個別治療之情形），以及效率最佳（如研究情境，通常使用平均數等團體層級之指標，故對精準度要求較低），以最佳化MER-CAT之實用價值。

1. **開發MER-CAT施測介面（2個月）**

目的

發展MER-CAT之施測介面與資料儲存系統，以利後續使用者之施測與資料保存。

流程

計畫主持人將依據前二階段之結果（即最終版MER題庫之題目難度，以及所挑選之最佳終止條件），與專業程式設計師合作，以開發MER-CAT之前台使用者介面（含使用者登入系統、施測介面、患者資料管理介面、評估結果輸出介面與規格等），以及其後台資料管理系統（數據庫與使用者權限管理等）。

主持人預計使用JavaScript語法為主，並開發以網頁為基礎之施測平台，以最佳化其實用價值（使用者可以不同載具，如智慧型手機、平版電腦或桌上電腦等，登入系統以近行MER-CAT之施測）；此外，MER-CAT施測平台將兼具連結網路施測（需連網，資料上傳並保存於雲端伺服器）與單機施測（不需連網，系統建置於單機電腦，且評估結果保存於該電腦硬碟內）二種施測模式，以兼顧方便性與資訊安全之考量，更貼近臨床情境之需求。

1. **驗證MET-CAT之心理計量特性（8個月）**

目的

測試MER-CAT於慢性與急性思覺失調症患者，以檢驗其重要之心理計量特性，包含再測信度、練習效應、收斂與發散效度，以及反應性。

對象

本階段預計招募160位思覺失調症患者，並依其病程分為再測穩定性組（慢性患者80人），以及反應性組（急性患者80人），以分別檢驗MER-CAT之重要心理計量特性。

再測穩定性組之患者納入條件有四：(1) 依據DSM-5或ICD-10診斷為思覺失調症；(2) 最近一次發病時間達半年以上；(3) 年滿20歲；(4) 可遵守3步驟指令。然而，若受試者曾有其他精神診斷（如憂鬱症與強迫症等）、曾有腦傷病史、曾有物質濫用（酒精與藥物）之問題、於近半年內調整藥物者，或無法完成所有測驗者，則予以排除。

反應性組之患者納入條件有四：(1) 依據DSM-5或ICD-10診斷為思覺失調症；(2) 最近一次發病時間未超過3個月；(3) 年滿20歲；(4) 可遵守3步驟指令。然而，若受試者曾有其他精神診斷（如憂鬱症與強迫症等）、曾有腦傷病史、曾有物質濫用（酒精與藥物）之問題，或無法完成所有測驗者，則予以排除。

流程

參與本研究之患者將接受二次評估，二次評估之間隔時間分別為2週（再測穩定性組）與1個月（反應性組）。於簽署受試者同意書後，參與本研究之患者將被安排至安靜的評估室內，以完成後續測驗。參與二組之患者，每次皆需完成MER-CAT、MoCA、CGI-S與生態效度驗證之效標「個人與社會功能滿意度評估量表-簡易圖示版 (self-reported version of the graphic personal and social performance scale, SRG-PSP)」74 之評估。

工具

MER-CAT為前幾階段之成果，理論上可於10分鐘內完成所有評估。

MoCA與CGI-S已於第二階段詳述，於此不再重述。

SRG-PSP為國內學者由PSP改編而成，以簡易圖示呈現，藉以提升個案自行填寫之可行性。74 PSP評估個案過去一個月的社會功能表現狀況。此量表參考義大利所發展的精神復健介入計畫，並依據DSM-IV對思覺失調症的評估標準，包括四個面向的評估範疇：(1) 社會性活動 (socially useful activities); (2) 個人與社會關係 (personal and social relationships)；(3) 自我照顧 (self-care)；(4) 擾亂與攻擊行為 (disturbing and aggressiveness behaviors)。各面向以「個案功能表現的困難程度」分成6個等級，等級越高表示該範疇功能越差。分數範圍為1-100，並分成10個區間。評分時須先綜合四個面向的障礙程度，先選擇個案功能表現的分數區間(如40-50)，再考慮其它相關資訊，如社交網絡、興趣、使用交通工具與電話等，選擇10分區間內的分數(如46)做為最後的總分。PSP具有良好的再測信度75-77與施測者間信度。77,78 PSP具有可接受的建構效度、79收斂效度。75-77,79 SRG-PSP 於國內思覺失調個案之內在一致性與建構效度良好。74

資料分析

再測信度將以組內相關係數 (intraclass correlation coefficient, ICC) 為指標。ICC可視為測量變異數分析的一種應用，其主要檢測患者內變異（重複測量造成之分數差異）與患者間變異（個體間差異，可視為誤差）之比率：如變異主要來自於患者內變異，則代表評估結果不甚穩定，故再測一致性不佳；反之，如變異主要來自於患者間變異，則代表其重複測量分數的一致性高，故再測穩定性好。ICC之數值介於0至1間，數值越高即代表重複測量結果越穩定、一致。本研究採用ICC ≥ 0.90為極佳，適用於個別層級之評估；並以ICC ≥ 0.70為可接受，屬團體層級比較之最低標準。80

練習效應將以獨立樣本*t*檢定 (independent *t* test) 與效應值 (Cohen’s *d*) 為指標。練習效應代表在受試者能力穩定不變的前提下，重複操作測驗所造成之分數提升之偏差。理論上，練習效應易導致使用者高估受試者之能力（因能力實際上不變，僅受試者更熟練如何操作此測驗），故為重複測量情境下必須確認之心理計量特性。本研究使用*t*檢定以檢驗重複測量結果間是否存在系統性偏差；而Cohen’s *d*則用以量化該偏差之程度，並以*d* ≥ 0.20為小、*d* ≥ 0.50為中，並以*d* ≥ 0.80為嚴重。72

收斂與發散效度將以皮爾森相關係數 (Pearson’s *r*) 為指標。本研究將以二面向檢視MER-CAT之收斂與發散效度：其一、向度間之相性：本研究將檢視各情緒刺激之感覺傳遞途徑下，受試者判讀7種基本情緒之分數間的相關性是否與昔日文獻相符，以判斷各向度分數之收斂/發散效度；其二、與效標之相關性：本研究將比較各向度分數與MoCA及CGI-S分數間之相關性是否與昔日文獻相符，以檢驗各向度分數之有效性。

反應性之驗證將分為團體層級與個別層級進行。73 團體層級反應性 (group-level responsiveness) 代表以團體性指標（如平均數等）時，評估工具偵測功能改變之能力，本研究採用相依樣本*t*檢定 (paired *t* test) 與標準化反應平均 (standard response mean, SRM) 為指標，以分別檢驗在已知患者功能有改變的前提下，其平均分數是否顯著地不同，以及其差異程度之大小。SRM之標準與Cohen’s *d*相同，於此不再贅述。於個別層級反應性 (individual-level responsiveness) 方面，本研究將採用分數改變量超過隨機測量誤差之人數比率為指標，以檢驗若以MER-CAT施測，有多少比率的受試者其進步可為MER-CAT所掌握/呈現。

生態效度：以 Pearson *r* 驗證MER-CAT與SRG-PSP分數之關連程度。研究者預期MER-CAT與SRG-PSP之關連程度中等。

1. **預期完成之工作項目及成果。請分年列述：1.預期完成之工作項目。2.對於參與之工作人員，預期可獲之訓練。3.預期完成之研究成果（如實務應用績效、期刊論文、研討會論文、專書、技術報告、專利或技術移轉等質與量之預期成果）。4.學術研究、國家發展及其他應用方面預期之貢獻。**

**預期完成之工作項目**

第一年（民國111/8–112/7）

1. **挑選題目：**經專家會議與初步測試，完成MER候選題庫之挑選。
2. **大規模測試（部分）：**測試MER候選題庫於100位思覺失調症患者與與100位健康成人。
3. **計畫/成果分享：**於國內學術研討會發表研究成果，如 “Development of a multi-dimensional-multi-channel emotion recognition computerized adaptive testing for patients with schizophrenia: Study Protocol”。

第二年（民國112/8–113/7）

1. **大規模測試：**測試MER候選題庫於400位思覺失調症患者與與100位健康成人，至此將完成大規模測試之所有收案（共納入思覺失調症患者500人與健康成人200人）。
2. **計畫/成果分享：**於學術研討會發表研究成果，如 “Development of a multi-dimensional-multi-channel emotion recognition computerized adaptive testing for patients with schizophrenia: Preliminary results”。

第三年（民國113/8–114/7）

1. **資料分析：**完成MER-CAT所需之資料分析，包含藉羅序分析（挑選適配度良好且無性別偏差之題目，以組成最終版MER題庫）、CFA分析（確認最終版MER題庫之個別單向度），以及CAT模擬分析（挑選兼具信度與施測效率之終止條件組，以最佳化MER-CAT之施測效能）。
2. **系統開發：**與工程師合作完成MER-CAT施測系統之設置。
3. **驗證心理計量特性：**完成160位（急性與慢性各半）思覺失調症患者之資料收集（共2次，急性間隔1個月、慢性間隔2週），以確認MER-CAT之再測信度與反應性等。
4. **計畫/成果分享：**
   1. 撰寫稿件 “Development of a multi-dimensional-multi-channel emotion recognition computerized adaptive testing (MER-CAT) for patients with schizophrenia for patients with schizophrenia”，預期投稿至SCI期刊如 Schizophrenia Bulletin。
   2. 於學術研討會發表研究成果，如 “Validation of the multi-dimensional-multi-channel emotion recognition computerized adaptive testing for patients with schizophrenia:”。

第四年（114/8–115/7）

1. **計畫/成果分享：**撰寫稿件 “Psychometric evaluation of the multi-dimensional-multi-channel emotion recognition computerized adaptive testing (MER-CAT) for patients with schizophrenia
2. **成果推廣：**於國內各醫院精神部與研討會/工作坊推廣MER-CAT。

**預期貢獻**

本研究可能發展出國際間第一個適用於思覺失調患者之多向度多途徑情緒辨識測驗 (MER-CAT)。主持人預期MER-CAT可提供情緒向度完整、計分指標詳盡、傳遞途徑周全、施測快速、結果精準/有效/無性別偏差，且適用於我國患者之情緒辨識測驗。參與本研究之人員，預期可獲得以下訓練/經驗：

1. 學習情緒辨識相關之理論與評估方法。

2. 學習羅序模型與DIF分析之方法、結果解讀與實務經驗。

3. 發展CAT之概念、資料分析方法與實務操作經驗。

4. 學習心理計量特性之研究設計、分析方法、結果解讀與實務經驗。

5. 學習研究結果的呈現及論文報告之書寫。

**預計可能遭遇之困難及解決途徑**

1. **CAT發展不易：**由於CAT涉及羅序分析與模擬分析等，計算過程繁複，為確保計畫可行，計畫主持人擬延攬相關專家（如臺大職能治療學系謝清麟教授，已發展多套CAT系統）擔任共同主持人。
2. **樣本需求大（題庫測試需700人 [病人500位；健康人200位]；心理計量特性驗證則需思覺失調症患者160人）：**為確保可行性，除主持人所屬院區（高雄市立凱旋醫院等地）外，我們亦考慮與其他地區之醫院/精神機構合作（如台大醫院精神部、衛生福利部八里療養院與桃園市立療養院等），以確保受試者人數充足。其次，為確保收案人力，主持人乃編列經費招募研究助理（預期1至2人），以聯繫、資料收集、校對語彙整等任務。再者，為提升受試者之參與意願，本研究亦提供受試者費（以現金或禮券形式，相當於每小時200元計），以鼓勵/酬謝受試者對研究之貢獻。

**參考文獻**

1. Morris RW, Weickert CS, Loughland CM. Emotional face processing in schizophrenia. *Curr Opin Psychiatry*. 2009;22:140-146. doi: 10.1097/YCO.0b013e328324f895

2. Behere RV. Facial emotion recognition deficits: The new face of schizophrenia. *Indian J Psychiatry*. 2015;57:229-235. doi: 10.4103/0019-5545.166641

3. Salovey PE, Sluyter DJ. *Emotional development and emotional intelligence: Educational implications*. Basic Books; 1997.

4. 梁庚辰（2013）。心理學門熱門及前瞻學術研究議題調查簡介。《人文與社會科學簡訊》，第15卷第1期，頁116-123. doi:

5. Ekman P, Cordaro D. What is meant by calling emotions basic. *Emot Rev*. 2011;3:364-370. doi: 10.1177/1754073911410740

6. Ekman P. An argument for basic emotions. *Cogn Emot*. 2008;6:169-200. doi: 10.1080/02699939208411068

7. Ekman P. Facial expression and emotion. *The American psychologist*. 1993;48:384-392. doi: 10.1037//0003-066x.48.4.384

8. Bortolon C, Capdevielle D, Raffard S. Face recognition in schizophrenia disorder: A comprehensive review of behavioral, neuroimaging and neurophysiological studies. *Neurosci Biobehav Rev*. 2015;53:79-107. doi: 10.1016/j.neubiorev.2015.03.006

9. Banziger T, Grandjean D, Scherer KR. Emotion Recognition From Expressions in Face, Voice, and Body: The Multimodal Emotion Recognition Test (MERT). *Emotion*. 2009;9:691-704. doi: 10.1037/a0017088

10. Barkl SJ, Lah S, Harris AW, Williams LM. Facial emotion identification in early-onset and first-episode psychosis: A systematic review with meta-analysis. *Schizophr Res*. 2014;159:62-69. doi: 10.1016/j.schres.2014.07.049

11. Savla GN, Vella L, Armstrong CC, Penn DL, Twamley EW. Deficits in domains of social cognition in schizophrenia: A meta-analysis of the empirical evidence. *Schizophr Bull*. 2013;39:979-992. doi: 10.1093/schbul/sbs080

12. Cotter J, Granger K, Backx R, Hobbs M, Looi CY, Barnett JH. Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neurosci Biobehav Rev*. 2018;84:92-99. doi: 10.1016/j.neubiorev.2017.11.014

13. Andric Petrovic S, Jerotic S, Mihaljevic M, Pavlovic Z, Ristic I, Soldatovic I, Maric NP. Sex differences in facial emotion recognition in health and psychotic disorders. *Cognitive neuropsychiatry*. 2019;24:108-122. doi: 10.1080/13546805.2019.1582411

14. Waaramaa T. Gender differences in identifying emotions from auditory and visual stimuli. *Logoped Phoniatr Vocol*. 2017;42:160-166. doi: 10.1080/14015439.2016.1243725

15. Green MF, Penn DL, Bentall R, Carpenter WT, Gaebel W, Gur RC, Kring AM, Park S, Silverstein SM, Heinssen R. Social cognition in schizophrenia: An NIMH workshop on definitions, assessment, and research opportunities. *Schizophr Bull*. 2008;34:1211-1220. doi: 10.1093/schbul/sbm145

16. Buck BE, Healey KM, Gagen EC, Roberts DL, Penn DL. Social cognition in schizophrenia: Factor structure, clinical and functional correlates. *J Ment Health*. 2016;25:330-337. doi: 10.3109/09638237.2015.1124397

17. Balogh N, Egerhazi A, Berecz R, Csukly G. Investigating the state-like and trait-like characters of social cognition in schizophrenia: A short term follow-up study. *Schizophr Res*. 2014;159:499-505. doi: 10.1016/j.schres.2014.08.027

18. Kerr SL, Neale JM. Emotion perception in schizophrenia: Specific deficit or further evidence of generalized poor performance? *J Abnorm Psychol*. 1993;102:312. doi:

19. Leentjens AF, Wielaert SM, van Harskamp F, Wilmink FW. Disturbances of affective prosody in patients with schizophrenia: A cross sectional study. *J Neurol Neurosurg Psychiatry*. 1998;64:375-378. doi: 10.1136/jnnp.64.3.375

20. Shaw RJ, Dong M, Lim KO, Faustman WO, Pouget ER, Alpert M. The relationship between affect expression and affect recognition in schizophrenia. *Schizophr Res*. 1999;37:245-250. doi:

21. Edwards J, Pattison PE, Jackson HJ, Wales RJ. Facial affect and affective prosody recognition in first-episode schizophrenia. *Schizophr Res*. 2001;48:235-253. doi: 10.1016/s0920-9964(00)00099-2

22. Ross E, Orbelo D, Cartwright J, Hansel S, Burgard M, Testa J, Buck R. Affective-prosodic deficits in schizophrenia: Profiles of patients with brain damage and comparison with relation to schizophrenic symptoms. *J Neurol Neurosurg Psychiatry*. 2001;70:597-604. doi:

23. Hooker C, Park S. Emotion processing and its relationship to social functioning in schizophrenia patients. *Psychiatry Res*. 2002;112:41-50. doi: 10.1016/s0165-1781(02)00177-4

24. Cotter J, Granger K, Backx R, Hobbs M, Looi CY, Barnett JH. Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neurosci Biobehav Rev*. 2017. doi:

25. Brekke J, Kay DD, Lee KS, Green MF. Biosocial pathways to functional outcome in schizophrenia. *Schizophr Res*. 2005;80:213-225. doi: 10.1016/j.schres.2005.07.008

26. Kee KS, Kern RS, Green MF. Perception of emotion and neurocognitive functioning in schizophrenia: What's the link? *Psychiatry Res*. 1998;81:57-65. doi: 10.1016/s0165-1781(98)00083-3

27. Pinkham AE, Penn DL. Neurocognitive and social cognitive predictors of interpersonal skill in schizophrenia. *Psychiatry Res*. 2006;143:167-178. doi: 10.1016/j.psychres.2005.09.005

28. Poole JH, Tobias FC, Vinogradov S. The functional relevance of affect recognition errors in schizophrenia. *J Int Neuropsychol Soc*. 2000;6:649-658. doi: 10.1017/s135561770066602x

29. Couture SM, Penn DL, Roberts DL. The functional significance of social cognition in schizophrenia: A review. *Schizophr Bull*. 2006;32 Suppl 1:S44-63. doi: 10.1093/schbul/sbl029

30. Mancuso F, Horan WP, Kern RS, Green MF. Social cognition in psychosis: Multidimensional structure, clinical correlates, and relationship with functional outcome. *Schizophr Res*. 2011;125:143-151. doi: 10.1016/j.schres.2010.11.007

31. Schmidt SJ, Mueller DR, Roder V. Social cognition as a mediator variable between neurocognition and functional outcome in schizophrenia: Empirical review and new results by structural equation modeling. *Schizophr Bull*. 2011;37 Suppl 2:S41-54. doi: 10.1093/schbul/sbr079

32. Green MF. Impact of cognitive and social cognitive impairment on functional outcomes in patients with schizophrenia. *J Clin Psychiatry*. 2016;77 Suppl 2:8-11. doi: 10.4088/JCP.14074su1c.02

33. Addington J, Saeedi H, Addington D. Facial affect recognition: a mediator between cognitive and social functioning in psychosis? *Schizophr Res*. 2006;85:142-150. doi:

34. Pinkham AE. Social cognition in schizophrenia. *J Clin Psychiatry*. 2014;75 Suppl 2:14-19. doi: 10.4088/JCP.13065su1.04

35. Schmidt SJ, Mueller DR, Roder V. Social Cognition as a Mediator Variable Between Neurocognition and Functional Outcome in Schizophrenia: Empirical Review and New Results by Structural Equation Modeling. *Schizophrenia Bulletin*. 2011;37:S41-S54. doi: 10.1093/schbul/sbr079

36. Kurtz MM, Richardson CL. Social cognitive training for schizophrenia: A meta-analytic investigation of controlled research. *Schizophr Bull*. 2012;38:1092-1104. doi: 10.1093/schbul/sbr036

37. Pinkham AE. Social cognition in schizophrenia. *J Clin Psychiatry*. 2014;75 Suppl 2:14-19. doi: 10.4088/JCP.13065su1.04

38. Schmidt SJ, Mueller DR, Roder V. Social cognition as a mediator variable between neurocognition and functional outcome in schizophrenia: Empirical review and new results by structural equation modeling. *Schizophr Bull*. 2011;37 Suppl 2:S41-54. doi: 10.1093/schbul/sbr079

39. Addington J, Saeedi H, Addington D. Facial affect recognition: A mediator between cognitive and social functioning in psychosis? *Schizophr Res*. 2006;85:142-150. doi: 10.1016/j.schres.2006.03.028

40. Barkl SJ, Lah S, Harris AW, Williams LM. Facial emotion identification in early-onset and first-episode psychosis: a systematic review with meta-analysis. *Schizophr Res*. 2014;159:62-69. doi: 10.1016/j.schres.2014.07.049

41. Phillips LH, Slessor G. Moving beyond basic emotions in aging research. *Journal of Nonverbal Behavior*. 2011;35:279. doi:

42. Hall JA. Gender effects in decoding nonverbal cues. *Psychol Bull*. 1978;85:845. doi:

43. O’toole AJ, Deffenbacher KA, Valentin D, Abdi H. Structural aspects of face recognition and the other-race effect. *Memory & Cognition*. 1994;22:208-224. doi:

44. Elfenbein HA, Ambady N. On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychol Bull*. 2002;128:203-235. doi: 10.1037/0033-2909.128.2.203

45. Albas DC, McCluskey KW, Albas CA. Perception of the emotional content of speech: A comparison of two Canadian groups. *Journal of Cross-Cultural Psychology*. 1976;7:481-490. doi:

46. Laukka P. *Vocal expression of emotion: Discrete-emotions and dimensional accounts*. Acta Universitatis Upsaliensis; 2004.

47. Young AW, Perrett D, Calder A, Sprengelmeyer R, Ekman P. *Facial expressions of emotion: Stimuli and tests (FEEST)*. Thames Valley Test Company; 2002.

48. Kerr SL, Neale JM. Emotion perception in schizophrenia: Specific deficit or further evidence of generalized poor performance? *J Abnorm Psychol*. 1993;102:312-318. doi: 10.1037/0021-843x.102.2.312

49. Kohler CG, Turner TH, Bilker WB, Brensinger CM, Siegel SJ, Kanes SJ, Gur RE, Gur RC. Facial emotion recognition in schizophrenia: Intensity effects and error pattern. *Am J Psychiatry*. 2003;160:1768-1774. doi: 10.1176/appi.ajp.160.10.1768

50. Bowers D, Blonder LX, Heilman KM. *Florida affect battery*. Center for Neuropsychological Studies, Department of Neurology Florida, USA; 1998.

51. Nowicki S, Duke M. *Manual for the receptive tests of the diagnostic analysis of nonverbal accuracy 2 (DANVA2)*. Department of Psychology, Emory University; 2008.

52. Bryson G, Bell M, Lysaker P. Affect recognition in schizophrenia: A function of global impairment or a specific cognitive deficit. *Psychiatry Res*. 1997;71:105-113. doi: 10.1016/s0165-1781(97)00050-4

53. Schlegel K, Grandjean D, Scherer KR. Introducing the Geneva emotion recognition test: An example of Rasch-based test development. *Psychol Assess*. 2014;26:666-672. doi: 10.1037/a0035246

54. Lee SC, Lin GH, Liu CC, Chiu EC, Hsieh CL. Development of the CAT–FER: A Computerized Adaptive Test of Facial Emotion Recognition for adults with schizophrenia. *Am J Occup Ther*. 2021;75:1-11. doi:

55. Ramirez LRB, Saracco-Alvarez R, Escamilla-Orozco R, Orellana AF. Validity of the Montreal Cognitive Assessment Scale (MoCA) to detect cognitive impairment in schizophrenia. *Salud Ment*. 2014;37:517-522. doi: 10.17711/SM.0185-3325.2014.062

56. Pinkham AE, Harvey PD, Penn DL. Social Cognition Psychometric Evaluation: Results of the final validation study. *Schizophr Bull*. 2018;44:737-748. doi: 10.1093/schbul/sbx117

57. Kucharska-Pietura K, Klimkowski M. Perception of facial affect in chronic schizophrenia and right brain damage. *Acta Neurobiol Exp (Wars)*. 2002;62:33-43. doi:

58. Nowicki S, Duke MP. Individual-differences in the Nonverbal-Communication of Affect - the Diagnostic-Analysis of Nonverbal Accuracy Scale. *J Nonverbal Behav*. 1994;18:9-35. doi: Doi 10.1007/Bf02169077

59. Segall DO. Multidimensional adaptive testing. *Psychometrika*. 1996;61:331-354. doi: Doi 10.1007/Bf02294343

60. Wang WC, Chen PH. Implementation and measurement efficiency of multidimensional computerized adaptive testing. *Appl Psych Meas*. 2004;28:295-316. doi: 10.1177/0146621604265938

61. 黃世琤、李明純、李麗雯、詹雅婷、蔡鑫廷。台灣地區華人情緒與相關心理生理資料庫-基本情緒聲調。中華心理學刊。2014；56：437-452. doi:

62. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53:695-699. doi: 10.1111/j.1532-5415.2005.53221.x

63. Wu C, Dagg P, Molgat C. Measuring stability of cognitive impairment in inpatients with schizophrenia with alternate forms of the Montreal Cognitive Assessment during acute hospitalization. *Psychiatry Res*. 2017;258:299-304. doi: 10.1016/j.psychres.2017.08.065

64. 黃世琤, 李明純, 李麗雯, 詹雅婷, 蔡鑫廷. 台灣地區華人情緒與相關心理生理資料庫-基本情緒聲調. *中華心理學刊*. 2014;56:437-452. doi:

65. Haro JM, Kamath S, Ochoa S, Novick D, Rele K, Fargas A, Rodriguez M, Rele R, Orta J, Kharbeng A. The Clinical Global Impression–Schizophrenia scale: A simple instrument to measure the diversity of symptoms present in schizophrenia. *Acta Psychiatrica Scandinavica*. 2003;107:16-23. doi:

66. Haro JM, Kamath SA, Ochoa S, Novick D, Rele K, Fargas A, Rodriguez MJ, Rele R, Orta J, Kharbeng A, et al. The Clinical Global Impression–Schizophrenia scale: A simple instrument to measure the diversity of symptoms present in schizophrenia. *Acta Psychiatr Scand Suppl*. 2003;107:16-23. doi: 10.1034/j.1600-0447.107.s416.5.x

67. Linacre J. What do infit and outfit, mean-square and standardized mean? . *Ras Meas Trans*. 2002;16:878. doi:

68. Wright BD, Linacre M. Reasonable mean-square fit values. *Ras Meas Trans*. 1994;8:370. doi:

69. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. *J Educ Res*. 2006;99:323-337. doi: Doi 10.3200/Joer.99.6.323-338

70. Bentler PM, Wu EJC. *EQS 6.1 for Windows*. Multivariate Software INC; 2005.

71. Zieky M. *A DIF primer*. Educational Testing Service; 2003.

72. Cohen J. A power primer. *Psychol Bull*. 1992;112:155-159. doi: 10.1037//0033-2909.112.1.155

73. Huang YJ, Lin GH, Lee SC, Chen YM, Huang SL, Hsieh CL. Group- and individual-level responsiveness of the 3-point Berg Balance Scale and 3-point Postural Assessment Scale for Stroke Patients. *Arch Phys Med Rehabil*. 2018;99:529-533. doi: 10.1016/j.apmr.2017.08.472

74. Bai YM, Hsiao CY, Chen KC, Huang KL, Lee IH, Hsu JW, Chen PS, Yang YK. The development of a self-reported scale for measuring functionality in patients with schizophrenia--self-reported version of the graphic Personal and Social Performance (SRG-PSP) scale. *Schizophrenia research*. 2014;159:546-551. doi: 10.1016/j.schres.2014.08.024

75. Juckel G, Schaub D, Fuchs N, Naumann U, Uhl I, Witthaus H, Hargarter L, Bierhoff HW, Brune M. Validation of the Personal and Social Performance (PSP) Scale in a German sample of acutely ill patients with schizophrenia. *Schizophrenia research*. 2008;104:287-293. doi: 10.1016/j.schres.2008.04.037

76. Nasrallah H, Morosini P, Gagnon DD. Reliability, validity and ability to detect change of the Personal and Social Performance scale in patients with stable schizophrenia. *Psychiatry research*. 2008;161:213-224. doi: 10.1016/j.psychres.2007.11.012

77. Patrick DL, Burns T, Morosini P, Rothman M, Gagnon DD, Wild D, Adriaenssen I. Reliability, validity and ability to detect change of the clinician-rated Personal and Social Performance scale in patients with acute symptoms of schizophrenia. *Current medical research and opinion*. 2009;25:325-338. doi: 10.1185/03007990802611919

78. Morosini PL, Magliano L, Brambilla L, Ugolini S, Pioli R. Development, reliability and acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale (SOFAS) to assess routine social functioning. *Acta psychiatrica Scandinavica*. 2000;101:323-329. doi:

79. Kawata AK, Revicki DA. Psychometric properties of the Personal and Social Performance scale (PSP) among individuals with schizophrenia living in the community. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 2008;17:1247-1256. doi: 10.1007/s11136-008-9400-z

80. Linacre JM. *A user’s guide to Winsteps. Rasch-model computer programs*. Winsteps.com; 2007.