

「軌跡標示測驗平板版」的信、效度研究暨本土常模建置

黃詩媛¹、黃博聖^{2,3}、陳志洪⁴、陳珮臻¹、陳學志^{1,5,6,7}

軌跡標示測驗 (Trail Making Test) 能評估個體高層次的認知執行功能及基本認知能力, 廣泛用於健康樣本、腦損傷及組件功能缺陷之患者。然在測驗工具使用上多以人工閱卷, 以致有計分相對耗時費力之劣勢。在臺灣亦尚無以數位化平板電腦操作之軌跡標示測驗, 實為研究缺口, 發展中文版平板操作軌跡標示測驗暨本土常模建置有其必要性。本研究發展一套電子數位化版本, 即「軌跡標示測驗平板版」(Tablet-Based Trail Making Test, TB-TMT)。本測驗涵蓋五個情境: 情境一視覺掃描、情境二簡單圖形數字序列、情境三選擇圖形數字序列、情境四轉換圖形數字序列、情境五動作速度, 藉由情境間的分數轉換, 可得四項指標: 視覺掃描時間、動作時間、抑制情境時間、轉換情境時間。依北、中、南、東, 採分層隨機取樣 1,046 位健康研究參與者, 年齡介於 8 至 29 歲, 進行信效度檢驗。結果顯示, TB-TMT 再測信度具有可接受的穩定性。各指標成分能代表其特性, 具有適當的內部相關性。效度考驗中, 以彩色路徑描繪測驗、西蒙作業、塔測驗、反應時間作業、中文版執行功能指標為效標, 發現 TB-TMT 具有良好之效標關聯效度。總結而言, TB-TMT 具有良好的信度與效度, 深具實務應用價值, 適用於未來學術研究、教學實務, 以及檢視臨床功能的良好工具, 以提供有效之參考依據。

關鍵詞: 執行功能、軌跡標示測驗、常模、平板電腦

¹ 國立臺灣師範大學教育心理與輔導學系

² 國立臺灣科技大學數位學習與教育研究所

³ 國立臺灣科技大學技職賦能研究中心

⁴ 國立臺灣師範大學資訊教育研究所

⁵ 國立臺灣師範大學學習科學跨國頂尖研究中心

⁶ 國立臺灣師範大學華語文與科技研究中心

⁷ 國立臺灣師範大學社會情緒教育與發展研究中心

通訊作者: 陳學志, 國立臺灣師範大學教育心理與輔導學系, chcjyh@ntnu.edu.tw。
本研究獲教育部核定的國立臺灣師範大學高等教育深耕計畫「學習科學跨國頂尖研究中心」、「華語文與科技研究中心」, 以及「社會情緒教育與發展研究中心」之支持與經費補助, 特此感謝。本研究作者群感謝審查委員耐心指導與寶貴建議。

執行功能 (executive function) 指的是一系列具相互關聯性的高層次認知過程, 能夠對個體思考、情緒與行動, 進行規劃與執行目標導向之控制 (Diamond, 2013), 係一組廣泛的認知處理能力, 使個體能夠靈活的適應不斷變化的環境需求 (Raver et al., 2011), 為多模態處理 (multi-modal processing) 與高層次認知技能的核心 (Anderson, 2002; Stuss & Alexander, 2000)。Miyake 等人 (2000) 和 Miyake 與 Friedman (2012) 提出執行功能最典型的三個核心能力: 抑制 (inhibition)、轉換 (switching)、更新 (updating)。執行功能是認知過程的集合與總稱, 用來描述各種相關但不同的認知功能, 對於個體成長階段的歷程至關重要 (張之捷等人, 2023; Rodríguez-Nieto et al., 2022; Sutin et al., 2022)。許多研究指出, 執行功能對學習表現、心理健康之預測性得到證實 (簡馨瑩, 2020; Best et al., 2011; Liao & Chen, 2022; Malone et al., 2022)。

常見使用於執行功能的神經心理學測驗為軌跡標示測驗 (Trail Making Test, 簡稱 TMT), TMT 起初的使用目的是作為智力測驗, 至今主要應用於神經心理學, 及其相關認知病症的診斷 (Hagen et al., 2014)。TMT 是神經心理學中, 最常被使用的認知評估測驗之一 (Butler et al., 1991), TMT 係一項視覺消除任務與一系列的圓圈連結任務, 涉及有關視覺掃描、動作速度、抑制、轉換之執行功能相關能力 (Delis et al., 2001)。在臨床神經心理學中, TMT 具備區分正常人與患者的認知功能表現之有效性 (Ashendorf et al., 2008; Ozcan et al., 2016), 因其具備測驗簡易性、實用性與敏感性 (Zeng et al., 2017), 過去也經常被使用在評估認知障礙與神經損傷之診斷工具 (Heilbronner et al., 1991; Periañez et al., 2007; Stanczak & Triplett, 2003)。可見 TMT 對於評估個體執行功能發揮著重要的作用。

典型的 TMT 有 A、B 兩種測驗情境, A 情境 (TMT-A) 係請研究參與者依序將隨機散放的數字 1 至 25 編號的圓圈進行連接, 例如: 1—2—3—4—5, 以此類推; B 情境 (TMT-B) 則是將 25 個圓圈進行連接, 並需依序數字 1 至 13, 與英文字母 A 至 L 之間進行交替連接, 例如: 1—A—2—B—3—C, 以此類推 (Reitan, 1992; Reitan & Wolfson, 1985)。後來 Delis 等人 (2001) 發展 Delis-Kaplan 執行功能系統 (Delis-Kaplan Executive Function System, D-KEFS), 係一組高等層次的認知執行功能測驗, 適用於正常人與臨床個案 (額葉受損程度) 之評估工具, 此測驗系統的 TMT 也被廣泛使用, 其將典型的 TMT 進行擴充以外, 並根據認知歷程取向提供五種情境測驗, 包括視覺掃描、數字序列、文字序列、數字文字轉換與動作速度, 藉此能讓研究參與者確認情境的缺陷分數是與否與認知彈性或一種或多種潛在組件技能的損害有關。

近期許多國家以 TMT 建立本土常模之相關研究中, 在研究對象部分, 大多以成年人至高齡者 (Indorewalla et al., 2022; Magnusdottir et al., 2021; Siciliano et al., 2019)、中高齡者 (St-Hilaire et al., 2018) 與高齡者 (Fröhlich et al., 2024) 為主, 針對涵蓋學齡兒童、中學生、高中生至成年人之建立本土常模較少。其次, 在收案人數部分, 大約範圍介於 200 至 1,000 人左右。在臺灣的研究學者由郭曉燕與花茂琴 (2015) 根據 D' Elia 等人發展的 Color Trails Test (D' Elia, 1996) 進行編譯, 修訂為中文版彩色路徑描繪測驗 (Color Trails Test, 簡稱 CTT)。然而, 此測驗係根據軌跡標示測驗的原理所設計之, 其建立常模係以臺灣 50 歲以上的健康成人為研究對象, 提供臺灣中高齡者常模資料, 對於執行功能、額葉功能的神經心理測驗提供重要的基礎, 由此可見軌跡標示測驗的重要性。不過, 此測驗並未針對不同年齡層進行廣泛蒐集, 且係以紙本測驗形式進行測量, 耗時又費力。此外, 在測驗內容僅有兩個測驗, 涉及的認知功能與基本能力有限。有鑑於此, 本研究嘗試發展一套數位化平板操作之軌跡標示測驗, 以打破人工閱卷的限制, 改良人工計分的劣勢, 建立中文版平板操作軌跡標示測驗之本土常模。研究對象橫跨四個不同年齡層, 學齡兒童、國中生、高中生、大學生 (含研究所以上)。其測驗涉及抑制情境時間、轉換情境時間等執行功能指標, 以及視覺掃描時間、動作時間等基本認知能力表現。

在認知心理研究領域, 回顧臺灣軌跡標示之相關測驗而言, 就研究者所知, 目前國內尚無發展標準化之相關測驗工具。據此, 本研究欲發展標準化之軌跡標示測驗平板版, 建立認知發展至穩定階段, 研究參與者年齡 8 至 29 歲之臺灣本土常模。藉此瞭解研究參與者的發展變化, 能夠更清楚地看到執行功能歷程的差異。並加以深化臺灣研究領域, 對於認知層次之執行功能與軌跡標示之探究, 提供數位化、創新性的標準化評估工具做為認知心理、臨床神經診斷、學習者評量之工具, 有其研究與實務之必要性。

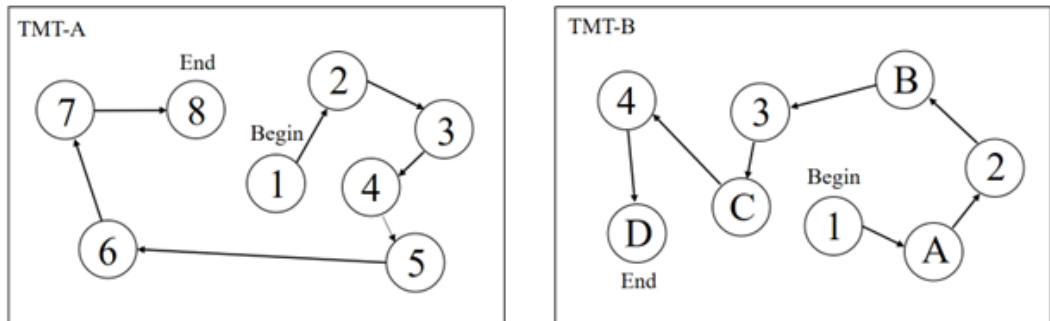
文獻探討

(一) 軌跡標示測驗之定義與發展

軌跡標示測驗 (TMT) 又稱為路徑描繪測驗，係一項依照指定任務將圓圈進行序列連結及視覺消除之任務。最初是在 1938 年由 Partington 所發展之，作為一般智力的衡量標準，直到 1942 年，使用成年男性群體進行標準化，並與斯比智力量表 (Revised Stanford-Binet Scale) 具顯著正相關 (Partington & Leiter, 1949)。TMT 被廣泛收錄在認知測驗組合 (Army Individual Test Battery, 1944)、認知評估工具－蒙特利爾認知評估 (Montreal Cognitive Assessment, MoCA) (Nasreddine et al., 2005)，以及牛津認知篩查 (Oxford Cognitive Screen, OCS; Oxford Cognitive Screen-Plus, OCS-Plus) (Demeyere et al., 2015; Demeyere et al., 2021) 之中。

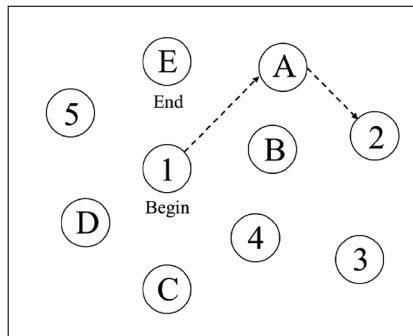
此測驗最典型的是 Reitan 與 Wolfson (1985) 發展的 TMT，透過紙本測驗的方式，檢測個體是否存在腦損傷，是注意力、動作速度和認知靈活度的衡量標準。測驗包括 TMT-A 和 TMT-B 兩部分，TMT-A 部分的指定任務是畫線連結 1 至 25 個圓圈數字，其隨機分散在測驗紙上，由數字 1 開始依序串聯，例如：1—2—3—4—5，以測量視覺掃描、數字序列與動作速度；TMT-B 部分的指定任務是 13 個數字 (1 至 13) 和 12 個英文字母 (A 至 L) 之間交替連結，例如：1—A—2—B—3—C，以測量視覺運動、視覺空間能力與認知靈活度，較 TMT-A 部分而言，TMT-B 會需要更長的時間完成 (如圖 1 所示)。接著，MoCA 是由 Nasreddine 在 1995 年發展認知評估測驗，同樣也是紙本測驗的形式，檢測健康專業人員的輕度認知障礙，以及評估認知衰退症狀或記憶喪失的病患，並能識別阿茲海默症及其他癡呆症風險的個體 (Nasreddine et al., 2005)。MoCA 包括 30 個題目，TMT 在該測驗為其中一題，採用的是 Trails B 測驗進行檢查，其指定任務為 5 個數字 (1 至 5) 和 5 個英文字母 (A 至 E) 之間交替串聯，例如：1—A—2—B—3—C—4—D—5—E (如圖 2 所示)。

圖 1
TMT-A 與 TMT-B



Note. Adapted from *The halstead-Reitan neuropsychological test battery: Theory and clinical interpretation* (p. 95), by R. M. Reitan and D. Wolfson, 1985, Neuropsychology Press.

圖 2
MoCA-Trails

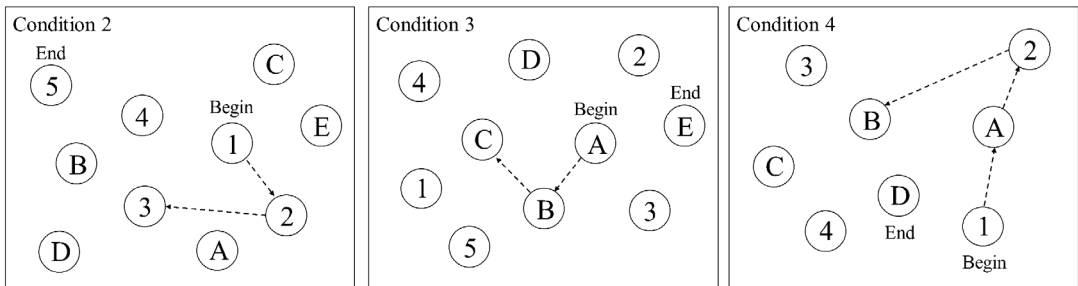


Note. Adapted from “The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment,” by Z. Nasreddine, N. A. Phillips, V. Bédirian, and S. Charbonneau, 2005, *Journal of the American Geriatrics Society*, 53(4), p. 697 (<https://doi.org/10.1111/j.1532-5415.2005.53221.x>). Copyright 2024 by the John Wiley & Sons.

關於上述提及的 Reitan 與 Wolfson (1985) 發展的 TMT，以及 Nasreddine 等人 (2005) 發展的 Trails B，施測對象以認知相關疾患為主，且此二個測驗雖有所差異，但實際測驗之相似性高，涉及檢測的認知能力較為侷限。後來 Delis 等人 (2001) 發展 Delis-Kaplan 執行功能系統 (D-KEFS) 共有九個高層次認知測驗，使用對象涵蓋 8 歲至 89 歲，適用於一般兒童至高齡者，以及患有嚴重神經或精神疾病患者。Delis 等人結合 Reitan 的形式，發展 D-KEFS 軌跡標示測驗 (Delis-Kaplan Executive Function System-Trail Making Test)，涵蓋五個執行功能任務情境：視覺掃描 (visual scanning)、數字序列 (number sequencing)、文字序列 (letter sequencing)、數字文字轉換 (number-letter switching) 以及動作速度 (motor speed)，能以評估多種情境測驗辨識認知神經機制的受損，評估個體在執行視覺動作序列任務時的思維靈活性、視覺注意力、運動功能、動作速度 (如圖 3 所示)。

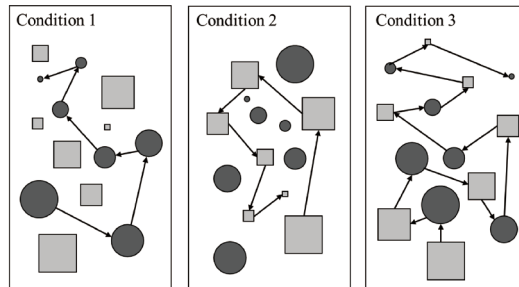
近期 Demeyere 等人 (2015) 開發針對中風症狀患者的認知篩檢工具 OCS，評估與中風相關的不同認知領域，包括注意力、實踐 (praxis)、記憶、數字和語言，並在 2021 年開發平板電腦的數位化工具 OCS-Plus，用於檢測更細微的認知變化。然而，在上述的 TMT 測驗版本皆涉及文字辨識之問題，Demeyere 等人認為軌跡標示測驗係評估非語言的認知功能，但在很大的程度上卻依賴於語言功能，若有語言缺陷的患者在測驗上可能會受到評估限制，便無法確切檢測認知障礙能力，對於語言障礙特徵的患者群體而言會是個需要正視的問題。因此，Demeyere 團隊解決此問題，發展 OCS-Plus 包括 10 個子測驗，其中一項 Trails 測驗包含 3 個情境，將原本涉及到數字與英文字母的軌跡路徑形式，調整為形狀之圓形和正方形，以及由大至小的方式進行檢測。例如：情境一為具有正方形的干擾物的情況下，按照圓圈由大至小進行連接；情境二為具有圓形干擾物的情況下，按照正方形由大至小進行連線；情境三為圓形和正方形之間進行交替的軌跡，並依照正方形和圓形由大至小進行串聯銜接 (如圖 4 所示)，則 OCS 所涉及的執行功能包括：注意力 (attention)、記憶 (memory)、空間 (spatial)、覺察 (awareness)、失用症 (apraxia)，以及知覺 (perception) 等能力。其中，OCS-Plus 為數位化平板電腦測驗工具，具備執行功能指標記錄、自動評分與報告功能，但與本研究開發的 TB-TMT 有所差異，且 OCS-Plus 仍存在一些研究限制：其一，OCS-Plus 主要評估中風患者的認知缺陷，相關研究大多使用在中風患者、失智症患者、健康高齡者樣本、認知障礙高齡者樣本 (Demeyere et al., 2021; Humphreys et al., 2017; Kobayashi et al., 2020; Webb et al., 2022)；其二，Demeyere 等人編製 OCS 測驗時，研究對象以成年人至高齡者為主，年齡介於 18 歲至 98 歲之間，研究對象採便利取樣方式，並未將年齡和教育程度進行匹配與控制，大多係接受高等教育的研究參與者，故對教育程度較低的研究參與者能力表現解釋之代表性與可推論性產生限制 (Iosa et al., 2022)。

圖 3
D-KEFS-TMT



Note. Adapted from *Delis-Kaplan Function System: Examiner's Manual* (pp. 41–43), by D. C. Delis, E. Kaplan, and J. H. Kramer, 2001, The Psychological Corporation.

圖 4
OCS-Trails



Note. Adapted from “The oxford cognitive screen (OCS): Validation of a stroke-specific short cognitive screening tool,” by N. Demeyere, M. J. Riddoch, E. D. Slavkova, W.-L. Bickerton, and G. W. Humphreys, 2015, *Psychological Assessment*, 27(3), p. 884 (<https://doi.org/10.1037/pas0000082>). Copyright 2024 by the American Psychological Association.

在測驗編製的考量中，本研究依據前述文獻探討的測驗設計，針對各測驗的指標特性、測驗程序進行評估，D-KEFS 軌跡標示測驗較其他測驗版本具有優勢。首先，Reitan 與 MoCA-Trails 雖為典型路徑描繪測驗，涉及的認知指標為執行功能基礎能力，較無明確能測得個體高層次認知能力。其次，OCS-Trails 雖克服文字辨識之問題，以形狀與大小作為測驗間的變化，能評估到不同的認知領域指標，但在形體大小之間的區辨，需要更有系統化的設計與組織；而 D-KEFS 軌跡標示測驗包括五個情境，能提供不同認知神經機制之執行功能指標與基礎能力的訊息，涵蓋視覺掃描、數字序列、文字序列、數字文字轉換及動作速度。其中，在 D-KEFS 第三個情境與第四個情境涉及語言文字層次，對於後續發展測驗工具時，可進行測驗設計之調整與精進。D-KEFS 整體而言，適合作為評估個體執行功能指標與軌跡標示識別的測量工具，且其適用於測量認知神經心理層次之高齡者、患者，也廣泛應用在健康樣本中 (Delis et al., 2001)。於此，本研究將參考 D-KEFS 軌跡標示測驗的設計編製原理，發展數位電子化軌跡標示測驗平板版。

(二) 發展平板電腦操作之軌跡標示測驗之編製基礎與必要性

本研究參考 D-KEFS-TMT 的任務情境與測量程序為基礎，D-KEFS-TMT (Delis et al., 2001) 共有五個情境。情境一視覺掃描，研究參與者需在任務頁面中，找到所有的阿拉伯數字 3，並劃一條斜線做記號，該情境涉及視覺掃描與視覺注意力之基本能力。情境二數字序列，研究參與者需從數字 1 劃一條線到數字 2，數字 2 到數字 3，數字 3 到數字 4，以此類推，按照順序畫到最後一個數字 16，該情境除了基本數值處理之外，還需要足夠的視覺掃描／注意力和運動功能。情境三文字序列，研究參與者需從英文字母 A 劃一條線到英文字母 B，英文字母 B 到英文字母 C，以此類推，按

照順序畫到最後一個英文字母 P，該情境以視覺、動作的任務形式對字母進行排序，是評估數字與字母切換技能的重要基線任務。情境四數字文字轉換，研究參與者需從數字與字母之間不斷轉換，從數字 1 開始劃一條線到字母 A，再從字母 A 到數字 2，數字 2 到字母 B，以此類推，此情境為 D-KEFS-TMT 的核心，涉及經典執行功能認知能力之認知彈性，是諸如多工處理、同時處理和分散注意力等高層次能力必不可少的技能。情境五動作速度，研究參與者需要從「開始」處沿著虛線劃線，持續劃到終點，此情境用以評估認知彈性的基礎為研究參與者的動作反應。透過一系列的視覺消除任務與圓圈數字、字母的連接任務，評估個體在執行視覺動作序列任務時的思維靈活性。

上述提及參考 D-KEFS-TMT 較其他軌跡標示測驗的優勢之處，但在相關測驗設計層面上需要進行調整與精進。其一，D-KEFS-TMT 情境三文字序列，使用英文字母作為刺激材料，即對於研究參與者而言，可能會受限於教育程度及識字能力，在英文字母辨識上造成施測適用範圍之限制，因此本研究將 D-KEFS 原版之情境三文字序列的測驗設計，改良為非語言層次之辨識；其二，D-KEFS 原版為紙本測驗，在測驗計分上較為繁瑣，人工計分耗時又耗費人力；其三，OCS-Plus 係採取數位化工具平板電腦的形式進行施測 (Demeyere et al., 2021)，足見數位化趨勢成為國內外測驗所重視之議題。本研究參考 D-KEFS-TMT 的情境與程序，克服認知測驗上的限制與侷限，將原本數字與文字序列，以圓形及六邊形數字序列的形式呈現，並且解決在測驗人工計分上耗時、耗力的窘境。至今在臺灣尚無以平板電腦操作的軌跡標示測驗進行開發，為建立平板電腦作為媒材的可靠性與有效性，以及擴充操作型執行功能測驗之創新性，實有發展之必要。故本研究將發展數位電子化之軌跡標示測驗平板版 (Tablet-Based Trail Making Test, 以下簡稱 TB-TMT)。

(三) 軌跡標示測驗之效度證據的研究

軌跡標示測驗平板版 (TB-TMT) 涉及的執行功能與認知基本能力，包括抑制情境時間、轉換情境時間等執行功能指標，以及視覺掃描時間、動作時間等基本認知能力。TB-TMT 是快速執行視覺消除任務，以及一系列的數字與圖形之序列連結任務。當個體執行測驗的視覺掃描及動作速度的反應時間愈短，表示其任務處理速度愈快 (Woods et al., 2015)。

本測驗 TB-TMT 主要測量執行功能與基本認知能力。TB-TMT 涉及的執行功能，包括抑制情境時間、轉換情境時間，選擇對神經心理功能具敏感度的彩色路徑描繪測驗 (CTT) (郭曉燕、花茂琴, 2015) 作為同時效度之檢驗，以及西蒙作業 (Simon Task, 簡稱 SmT) (Simon, 1990)、塔測驗 (Tower Task, 簡稱 TwT) (Delis et al., 2001) 檢驗 TB-TMT 之抑制情境時間及轉換情境時間的效標關聯效度。TB-TMT 涉及的基本認知能力，包括視覺掃描時間、動作時間，則以反應時間作業 (Reaction Time Task, 簡稱 RTT) (Deary et al., 2011) 檢驗 TB-TMT 之基本認知能力的效標關聯效度。

彩色路徑描繪測驗 (CTT) 包括兩個測驗，簡單知覺排序時間 (CTT-1) 與概念轉換時間 (CTT-2)，係被廣泛用來評估執行功能或額葉功能的神經心理測驗之一 (Shibuya-Tayoshi et al., 2007)。CTT 涉及的功能為排序、知覺追蹤速度、分散注意力、注意力及心智控制、概念轉換，其反應時間愈短，意即其認知功能愈佳 (郭曉燕、花茂琴, 2015)。其中，CTT 簡單知覺排序時間與 TB-TMT 動作時間，以及 CTT 概念轉換時間與 TB-TMT 轉換情境時間涉及相似歷程與測驗型態，其差別在於 CTT 為紙本測驗形式進行測量，故以「簡單知覺排序時間」作為 TB-TMT 「動作時間」之效標工具，則「概念轉換時間」作為 TB-TMT 「轉換情境時間」之效標工具。西蒙測驗 (SmT) 係一項測量執行功能表現的認知作業，根據刺激指示的方向與刺激位置之間的特徵材料進行作業反應 (Simon & Rudell, 1967)。SmT 在評估個體在抑制與轉換之刺激反應的有效性，在做出刺激與反應的過程中，同時也會受到視覺空間訊息及干擾之影響，其反應時間愈短，同時正確率高，意即其抑制、轉換能力愈佳 (Verbruggen et al., 2005)，本研究將以 SmT 的「抑制效果」與「轉換效果」作為 TB-TMT 「抑制情境時間」與「轉換情境時間」之效標工具。塔測驗 (TwT) 是一項涉及計畫 (Gnys & Willis, 1991)、抑制和認知靈活性 (Welsh & Huizinga, 2001)、空間規劃、規則學習並持續執行任務的測驗 (Delis et al., 2001)，因此除了被視為執行功能測驗工具之一外，也常作為神經心理學的衡鑑工具。在 D-KEFS-TwT 版本中，包括移動時間、違反規則等測驗反應的紀錄指標，違反規則即涉及反應執行功能之抑制能力，違反規則次數愈少，表示其抑制能力愈佳，故以 TwT 的「違反規則」作為 TB-TMT 「抑制情境時間」之效標工具。反應時間作業 (RTT) 係測量對特定刺激所

做出的認知處理反應時間 (Deary et al., 2010)。研究發現，處理速度測驗與軌跡標示測驗具有顯著相關性 (MacPherson et al., 2017; Salthouse, 2011)。針對認知障礙的患者，會使用軌跡標示 TMT-A 和 TMT-B 評估處理速度與注意力 (Christidi et al., 2015; Hu et al., 2023; Wang et al., 2023)。此外，軌跡標示測驗得分與視覺空間能力、處理速度、記憶力等認知領域測驗也具關聯性 (MacPherson et al., 2019)，本研究將以 RTT 的「選擇反應時間」與「選擇移動時間」作為 TB-TMT「視覺掃描間」及「動作時間」之效標工具。

執行功能的測驗方式包含認知實作作業與自陳行為評量，上述相關測驗與作業屬認知實作作業之型式，而自陳行為評量則為評估日常生活的執行功能行為及表現。過去研究指出，此二種不同型式測量執行功能的方法，其測量的結果可解釋個體不同的認知表現面向 (Dekker et al., 2017)，表示認知實作作業與自陳行為評量之執行功能指標或成份，可能係涉及於不同認知歷程。本研究將以中文版執行功能指標 (CEFI-C, CEFI-J) (游勝翔、林緯倫, 2018)，與 TB-TMT 之指標與基本認知能力相關測量向度進行效度檢驗。

綜上可知，CTT 的簡單知覺排序時間愈短，TB-TMT 的動作時間也愈短，概念轉換所需時間愈短，TB-TMT 的轉換情境時間也愈短。SmT 的抑制、轉換時間愈快，TB-TMT 的抑制、轉換情境時間也愈短。TwT 的違反規則次數愈少，TB-TMT 的抑制情境時間也愈短。RTT 的選擇反應與移動時間愈短，TB-TMT 的視覺掃描時間也愈短，TB-TMT 的動作時間也愈短。CEFI 的策略行動、組織規劃與衝動控制表現愈佳，TB-TMT 的抑制與轉換情境時間則愈短。

藉此，以 CTT 的「簡單知覺排序時間」和「概念轉換時間」，SmT 的「抑制效果」和「轉換效果」之所需時間，TwT 的「抑制能力」，RTT 的「選擇反應／移動時間」、「簡單反應／移動時間」，以及 CEFI 的「策略行動」、「組織規劃」及「衝動控制」作為效標關聯效度。再者，本研究發展的 TB-TMT，主要涉及的執行功能任務為抑制情境時間、轉換情境時間等兩項，以及基本認知能力為視覺掃描時間和動作時間兩項。因此，本研究以再測信度作為信度分析，依據各指標特性進行內部作業相關作為效度證據，以彩色路徑描繪測驗 (CTT)、西蒙作業 (SmT)、塔測驗 (TwT)、反應時間作業 (RTT) 與中文版執行功能指標 (CEFI) 五者作為重要的參照效標，檢驗 TB-TMT 的效標關聯效度，同時檢驗 TB-TMT 在性別及教育階段上的差異。

(四) 研究目的

本研究目的在發展中文版「軌跡標示測驗平板版」(TB-TMT)，並進一步對此操作式測驗進行建立常模、信度與效度研究。將採取數位電子化的平板電腦操作方式，擴展認知神經心理學測驗在年齡層及族群的適用範圍，並能降低人力、解決傳統使用上的相關限制。國內教育與認知神經心理領域目前亦缺乏一套適用於兒童至成年人且兼具數位電子化、本土常模的軌跡標示測驗，因此為有效且精確地探究正常研究參與者及認知需求的疾患之執行功能，促使能尋求相關適合的治療或訓練方案，本研究所編製兼具本土常模的數位電子化軌跡標示評量工具實有其重要性與意義。

方法

首先，茲就研究參與者、研究工具、研究流程三部分，說明量表之編製過程。其次，以測驗實施研究樣本進行信效度分析，說明驗證信效度之程序。

(一) 研究參與者

1. 測驗實施研究樣本

本研究之測驗實施對象橫跨四個不同年齡層，共有 1,046 位研究參與者，分別為：國小生（三年級至六年級）、國中生、高中生、大學生（含研究所以上）。為顧及樣本結構特徵具母體代表性，分別從年齡、地區、性別等變項採取分層隨機取樣。平均年齡為 14.17（標準差 = 4.05；範圍：8 ~ 29 歲）。女性有 515 位，占總數的 49%，男性有 531 位，占總數的 51%。約 37% 來自北部，26%

來自中部，24%來自南部，13%來自東部。有關教育年分成四組，包括國小中高年級（ $N = 354$ ）、國中（ $N = 280$ ）、高中（ $N = 217$ ）、大學（含研究所以上）（ $N = 195$ ），如表 1 所示。本研究參與者中文閱讀能力無明顯障礙，另排除未矯正的聽力障礙、未矯正的視覺障礙、色盲／色弱、上肢運動障礙／上半身運動障礙、神經疾病或精神疾病（例如：腦血管疾病、腦炎、青光眼、藥物濫用、癲癇、思覺失調症、憂鬱或躁鬱症）、正在服用抗憂鬱、抗焦慮或抗精神疾病之藥物、一週內超過兩天飲用三份以上的酒精飲品、曾無端失去意識 5 分鐘以上、曾因頭部受傷而住院，以及有認知或記憶障礙病史，或曾被診斷為學習障礙等症狀。研究參與者分布見表 1，以考驗 TB-TMT 的信效度並建立常模。本研究已得到「國立臺灣師範大學研究倫理中心之研究倫理審查委員會」（Institution Review Board, IRB）審查通過，在研究開始之前，所有參與者均已獲悉研究程序並獲得知情同意，完成測驗後提供參與者禮品。

表 1
測驗實施研究參與者分布（ $N = 1,046$ ）

區域	性別	國小	國中	高中	大學	全體
		N (%)	N (%)	N (%)	N (%)	N (%)
北部	男	48 (13%)	74 (19%)	25 (7%)	46 (12%)	193 (51%)
	女	51 (13%)	76 (20%)	31 (8%)	31 (8%)	189 (49%)
	小計	99 (26%)	150 (39%)	56 (15%)	77 (20%)	382 (100%)
中部	男	60 (22%)	25 (9%)	39 (14%)	20 (7%)	144 (52%)
	女	59 (21%)	27 (10%)	26 (10%)	19 (7%)	131 (48%)
	小計	119 (43%)	52 (19%)	65 (24%)	39 (14%)	275 (100%)
南部	男	46 (19%)	19 (7%)	42 (17%)	18 (7%)	125 (50%)
	女	40 (16%)	19 (8%)	27 (11%)	38 (15%)	124 (50%)
	小計	86 (35%)	38 (15%)	69 (28%)	56 (22%)	249 (100%)
東部	男	31 (22%)	17 (12%)	13 (9%)	9 (7%)	70 (50%)
	女	19 (14%)	23 (16%)	14 (10%)	14 (10%)	70 (50%)
	小計	50 (36%)	40 (28%)	27 (19%)	23 (17%)	140 (100%)
總計	男	185 (18%)	135 (13%)	119 (11%)	93 (9%)	532 (51%)
	女	169 (16%)	145 (14%)	98 (9%)	102 (10%)	514 (49%)
	總計	354 (34%)	280 (27%)	217 (21%)	195 (18%)	1,046 (100%)

2. 常模樣本

本研究透過嚴謹的程序進行抽樣以建立常模，依照比例選取 500 位進行各條件的控制與取樣。本研究參與者 1,046 位，因各教育階段（國小中年級、國小高年級、國中、高中、大學以上）的人口數不同，為兼顧樣本代表性，加以控制研究樣本的年齡、家長教育程度，以及依據全臺四個區域進行人口比例抽樣，刪除不符合條件的樣本。因此，本測驗建立常模，為符合人口比例，保留教育階段能夠符合代表性的各 100 位，共計 500 位研究參與者樣本建置常模。

依據內政部戶政司的人口數按年齡及教育程度分配資料（內政部戶政司，2023），以各年齡組及教育程度的人口比例進行抽樣，以每組 100 位為例，見表 2。其計算方法為搜尋該年齡組在出生時的平均生父、生母年紀，再計算其現今年紀，最後經查具配偶人口年齡與教育程度表，以此估之控制研究參與者之父母教育程度。

表 2
各年齡組及研究參與者父母教育程度的人口比例抽樣表

各年齡組		國小	國中	高中	大學
8 ~ 9 歲	比例分配	1.51	8.28	28.57	61.65
	抽樣人數	2	8	29	61

（續下頁）

表 2
各年齡組及研究參與者父母教育程度的人口比例抽樣表（續）

各年齡組		國小	國中	高中	大學
10 ~ 11 歲	比例分配	1.52	8.42	30.42	59.65
	抽樣人數	2	8	30	60
12 ~ 14 歲	比例分配	1.53	10.00	33.07	55.39
	抽樣人數	2	10	33	55
15 ~ 17 歲	比例分配	1.53	10.00	33.07	55.39
	抽樣人數	2	10	33	55
18 ~ 36 歲	比例分配	0.07	1.45	18.29	80.19
	抽樣人數	0	2	18	80

將全臺灣地區分為北部、中部、南部、東部等四個區域，依據內政部戶政司的各縣市人口比例（內政部戶政司，2019），各區域抽樣人數比例約為 4：3：3：1，人數依次為：220 位、130 位、130 位、20 位。本研究建立本土常模，共計 500 位參與者樣本的測驗資料，常模樣本分布如表 3 所示。平均年齡為 16.35（標準差 = 4.78；範圍：8 ~ 29 歲）。女性有 250 位，占總數的 50%，男性有 250 位，占總數的 50%。約 44% 來自北部，26% 來自中部，26% 來自南部，4% 來自東部。有關教育年分成四組，包括國小中年級（ $N = 100$ ）、國小高年級（ $N = 100$ ）、國中（ $N = 100$ ）、高中（ $N = 100$ ）、大學（含研究所以上）（ $N = 100$ ）。

表 3
常模樣本分布（ $N = 500$ ）

區域	性別	國小	國中	高中	大學	全體
		N (%)	N (%)	N (%)	N (%)	N (%)
北部	男	22 (20%)	22 (20%)	22 (20%)	22 (20%)	110 (100%)
	女	22 (20%)	22 (20%)	22 (20%)	22 (20%)	110 (100%)
	小計	44 (20%)	44 (20%)	44 (20%)	44 (20%)	220 (100%)
中部	男	13 (20%)	13 (20%)	13 (20%)	13 (20%)	65 (100%)
	女	13 (20%)	13 (20%)	13 (20%)	13 (20%)	65 (100%)
	小計	119 (43%)	26 (20%)	26 (20%)	26 (20%)	130 (100%)
南部	男	13 (20%)	13 (20%)	13 (20%)	13 (20%)	65 (100%)
	女	13 (20%)	13 (20%)	13 (20%)	13 (20%)	65 (100%)
	小計	119 (43%)	26 (20%)	26 (20%)	26 (20%)	130 (100%)
東部	男	2 (20%)	2 (20%)	2 (20%)	2 (20%)	10 (100%)
	女	2 (20%)	2 (20%)	2 (20%)	2 (20%)	10 (100%)
	小計	4 (20%)	4 (20%)	4 (20%)	4 (20%)	20 (100%)
總計	男	50 (20%)	50 (20%)	50 (20%)	50 (20%)	250 (100%)
	女	50 (20%)	50 (20%)	50 (20%)	50 (20%)	250 (100%)
	總計	100 (20%)	100 (20%)	100 (20%)	100 (20%)	500 (100%)

（二）研究工具

本研究的主要研究工具類型涵蓋認知實作作業與自陳行為評量，其中包括彩色路徑描繪測驗（CTT）、西蒙作業（SmT）、塔測驗（TwT）、反應時間作業（RTT），以及中文版執行功能指標（CEFI），作為檢驗軌跡標示測驗平板版（TB-TMT）的效標關聯效度之用。

1. 本研究編譯及發展之「軌跡標示測驗平板版」（TB-TMT）

（1）TB-TMT 之編製程序。本研究針對 Delis 等人（2001）所編製的 D-KEFS 軌跡標示測驗（Delis-Kaplan Executive Function System）Trail Making Test（TMT）進行中文版之修訂，並將傳統

的紙本測驗，開發為平板操作式的測驗型態，以發展軌跡標示測驗平板版（TB-TMT）。首先，研究者取得原作者的同意後，再由研究者與三位心理學專家討論，將原始測驗之操作方式、指導語、計分指標、分數解釋等，完整翻譯成中文後，為能表達適切的測驗題意、通順的任務語句，再由一位精通中英文的教育心理學專家檢驗測驗內容，以確認中英文翻譯之精確性。在構建 TB-TMT 的數位化測驗中，首先，將已完成中文版軌跡標示測驗，由 APP 軟體工程師進行數位化設計，以開發 Android 設備作業系統之應用程式。在實施 TB-TMT 時，研究參與者則使用平板電腦，以手指頭觸控的方式進行測驗。

（2）**TB-TMT 之測驗內容與目的。**TB-TMT 係一項視覺消除任務與一系列的數字與圖形的連結任務。主要的執行功能任務是情境 4 圓形六邊形數字轉換，用以評估在執行視覺動作序列任務時的思維靈活性，涉及轉換、視覺掃描、數字排序、運作速度之能力。情境 1 視覺掃描、情境 2 圓形數字序列、情境 3 六邊形數字序列、情境 5 動作速度等四個情境，係幫助研究參與者能夠量化並產生執行轉換任務時所需之關鍵組成技能的規範資料。利用這些方法，研究參與者可以確認轉換情境的缺陷分數與其認知彈性，或與一種或多種潛在組件技能的損害有關。此五個情境需在指定的時間停止，測驗時間以秒計算，依序為：150 秒、150 秒、150 秒、240 秒、150 秒，測驗時間約為 15 分鐘。

（3）**TB-TMT 施測程序。**研究參與者在接受測驗之前，先由施測人員解說測驗目的與內容，接著填寫健康調查表與基本資料問卷，評估其認知功能，以確保排除參與者可能之認知功能缺損。各情境皆有時間限制，若研究參與者在 150 秒鐘內仍未完成情境 1、2、3、5，以及在 240 秒內仍未完成情境 4，即終止測驗，並記錄每位研究參與者完成各情境所需的時間。本測驗採團體施測的方式，由施測人員進行測驗說明與注意事項，使研究參與者能理解與投入於測驗之操作。

（4）**各情境與指標所涉及的認知能力、計分指標。**本測驗共有五個情境，每個情境將可算得各自的完成時間，若研究參與者未能在時限內完成某情境的題目，其完成時間為該情境的最長時限，該題會被標註為錯誤反應。例如：研究參與者未完成情境 4，其完成時間為 240 秒，且該次嘗試為錯誤反應。後續研究者則檢視其各情境測驗原始資料之正確率，以判斷該樣本的測驗結果。情境 1 至 5 的原始分數以秒計算。以下針對各情境表現反應個體認知能力之說明。接著，以測量指標解釋說明分數反應個體的認知功能。

情境 1 視覺掃描：提供對於視覺掃描與視覺注意力的快速測試。情境 2 簡單圖形數字序列：需要研究參與者在視覺動作任務的形式下依序連數字，除了基本數值處理之外，還需要足夠的視覺掃描／注意力和運動功能。情境 3 選擇圖形數字序列：研究參與者被要求以視覺一動作的任務形式對圖形數字進行排序，此為評估圖形與數字切換技能的重要基線任務。情境 4 轉換圖形數字序列：此為軌跡標示測驗之核心，任務要求研究參與者按照順序在連接不同圖形與數字之間來回轉換。是諸如多工處理、同時處理和分散注意力等高層次能力必不可少的技能，係為經典的認知能力－執行功能中的認知彈性。情境 5 動作速度：用以評估認知彈性的基礎，為研究參與者的動作反應，如果研究參與者因為在中樞或周邊的動作功能障礙而造成在繪製線條時的速度減慢的話，在這項測驗的大多情境中都會獲得低分，包含情境 4。

以下根據各項比較測量之分數進行解釋。第一項指標：視覺掃描時間（情境 1）花費時間愈短表示視覺掃描與注意力表現良好，時間愈長代表研究參與者的視覺掃描、注意力不佳。第二項指標：動作時間（情境 5）時間愈短代表研究參與者的速度反應愈佳，反之亦然。第三項指標：抑制情境時間〔（情境 2 加情境 3）／2 減（情境 1 加情境 5）／2〕時間愈短代表研究參與者的抑制能力愈佳，反之亦然。第四項指標：轉換情境時間〔情境 4 減（情境 2 加情境 3）／2〕時間愈短表示轉換能力良好，反之亦然。

2. 效度考驗之研究工具

（1）**彩色路徑描繪測驗（中文版）（CTT）。**本研究所使用的 CTT，係郭曉燕與花茂琴（2015）根據 D'Elia 等人發展之 Color Trails Test（D'Elia, 1996）進行編譯與修訂。CTT 包括兩個部分，CTT-1 簡單知覺排序與 CTT-2 概念轉換，前者測驗包含數字 1 至 25，奇數出現在粉紅色圓圈裡，偶數出現在黃色圓圈裡，依序數字 1 至 25 進行連線排序，其測驗涉及的功能包括：手部動作技巧、

知覺追蹤、排序及持續注意力；後者測驗包含數字 1 至 25，除了數字 1 在粉紅色圓圈裡，其他數字各會出現粉紅色與黃色，依序數字 1 至 25，以及切換兩個顏色的轉換，其測驗需要同時在數字和顏色序列之間做輪替，涉及分散注意力、認知彈性、轉換等能力。CTT 的常模收集對象為臺灣 50 歲以上的健康成人，提供臺灣中老年人常模資料，有效進行執行功能評估。施測時間約 10 分鐘。

(2) 西蒙作業 (SmT)。本研究所使用的 SmT，主要參考 Simon (1990) 的研究，主要測量抑制 (inhibition) 和轉換 (switching) 之能力，以修改呈現刺激以及題目數量，並增加情境一的基礎測量與情境三的規則轉換。該測驗以平板電腦進行施測，本測量共計 136 題，每題呈現時間限制 2 秒，參與者被要求根據給予的規則進行判斷色塊的測驗，選擇按左鍵或右鍵，並被要求盡快且準確地做出判斷。本測驗共有三個情境，情境一中，色塊會出現螢幕正中間，參與者須在 2 秒內根據顏色點選左鍵或右鍵，回答速度越快越好，一共有 32 題。情境二的規則與情境一相同，但色塊會出現在螢幕左側或右側，當色塊顏色與回答的方向一致，反應時間會有促進的效果；當色塊顏色與回答的方向不一致，則會有抑制的效果，此情境共有 40 題。情境三中色塊會出現在左右兩側，判斷規則與先前不同，參與者須先判斷色塊的形狀，再根據形狀對應的顏色規則點選左鍵或右鍵，因此，可以根據反應時間測量到參與者在規則轉換上的能力，情境三共有 64 題。抑制分數為情境二不一致的反應時間減去情境一中的反應時間，而轉換分數為情境三中兩種規則轉換的反應時間減去情境二不一致的反應時間。施測時間約 10 分鐘。

(3) 塔測驗 (TwT)。本研究所使用的 TwT，係參考 Delis 等人 (2001) 發展之 D-KEFS 執行功能之塔測驗，主要測量抑制能力。測驗目標是盡可能用最少的移動步數將三個圓柱上不同大小的圓盤移動成指定的塔型，並需遵循兩個規則：一次只能移動一個圓盤，以及較小的圓盤只能擺放在大的圓盤之上。當研究參與者移動圓盤建構目標塔時，測驗計分記錄該題完成的移動次數、違規次數、移動時間與完成時間，以評估個體空間規劃、規則學習、抑制衝動和持續反應力。而塔測驗中的違規次數與抑制能力存在關聯性，且抑制可能對於減少違規行為有所幫助 (Baughman & Cooper, 2007)。施測時間約 10 分鐘。

(4) 反應時間作業 (RTT)。反應時間係用於測量個體對特定刺激做出反應所需的時間，涉及認知處理速度的能力，而反應時間與視覺刺激、動作速度皆存在關聯性 (Deary et al., 2010)，且選擇反應時間與抑制之間也存在相關性 (Burle et al., 2004)，本研究以「反應時間作業」作為效標考驗。本研究所使用的 RTT，主要參考 Deary-Liewald 之 RTT 程序 (Deary et al., 2011) 之選擇反應時間 (Choice Reaction Time, CRT)。其作業情境中，研究參與者針對四個正確目標項進行按鍵反應，共 32 個嘗試次。每個嘗試次的時間距範圍在 1 至 3 秒之間，研究參與者對目標項的反應時間愈快，表示反應時間表現愈佳，並記錄作業情境在按鍵反應的移動時間。該作業的內部一致性係數為 .97，再測信度係數為 .83。效標關聯效度顯示，CRT 與數字反應時間方框作業之評估反應時間 (numbers reaction time box) 達相關 ($r = .68, p < .01$)。施測時間約 5 分鐘。

(5) 中文版執行功能指標 (CEFI, CEFI-J)。本研究採用游勝翔與林緯倫 (2018) 「中文版執行功能指標」中的大學生版 (Chinese Executive Function Index-for College Students, CEFI-C) 與國中生版 (for Junior High School Students, CEFI-J) 作為研究工具，可分為「策略行動」、「組織規劃」及「衝動控制」等三因素，各版本題數各為 15 題。以李克特氏五點量表的方式進行回答，1 分代表「完全不符合」，5 分代表「完全符合」，得分愈高，表示其執行功能程度愈佳。CEFI-C 具有良好的內部一致性 (Cronbach' $\alpha = .80$)，CEFI-J 也具有良好的內部一致性 (Cronbach' $\alpha = .77$)，在建構效度部分，CEFI-C 與 CEFI-J 三因素模型之整體模式適配度良好，顯示兩個版本之量表的三因素模式與觀察資料之適配程度高，量表因素結構穩定。表示此兩個版本皆具有良好的信、效度。施測時間約 10 分鐘。

(三) TB-TMT 研究流程

本測驗以 1,046 位學齡兒童至成年階段（8 ~ 29 歲）健康樣本為研究參與者（女性 514 位、男性 532 位），並依據全臺灣地區（北部、中部、南部、東部）四個區域的人數比例，採取分層隨機抽樣，以顧及樣本代表性。本研究採團體施測的方式，由施測人員說明測驗須知及注意事項，根據施測人員與研究參與者之人數分配比例，以一位施測人員對應 15 位研究參與者為上限。進行施測之前，施測人員事先將平板電腦設備設定為飛航模式與靜音狀態，避免測驗時受到干擾而影響施測環境。在正式施測時，提供研究參與者平板電腦，引導其點擊打開測驗作業之應用程式。首先，研究參與者進行填答基本資料，包括姓名、編碼、性別、慣用手、教育階段，以及生日。基本資料填答完成後，開始進行 TB-TMT 平板版測驗。TB-TMT 包含五個情境，每個情境皆包括練習階段和正式階段，並附有指導語及操作範例影片。研究參與者在練習階段中，需正確無誤完成限時的練習題，系統再將轉至正式階段進行測驗。研究參與者依序完成所有情境測驗後，平板畫面即出現測驗結束，表示該測驗資料已儲存。在信、效度考驗部分，以信度分析確認再測信度；在效度驗證階段，依指標特性進行內部相關性之分析，同時施測 TB-TMT 與 CTT、SmT、TwT、RTT 及 CEFI 具有良好信效度之效標工具，以進行本測驗之效度檢驗。

結果

(一) 描述統計與分析

本研究在 TB-TMT 各情境與指標之完成時間（秒）的平均數：情境一為 37 秒（標準差 11.07）、情境二為 60 秒（標準差 18.62）、情境三為 52 秒（標準差 18.87）、情境四為 66 秒（標準差 21.00）、情境五為 39 秒（標準差 14.16）。透過比較測量發現，情境三的完成時間顯著少於情境二（ $F(1, 1,045) = 176.01, p < .001$ ），顯示出情境二簡單圖形數字序列與情境三選擇圖形數字序列之間的練習效果；情境四的完成時間顯著多於情境二與情境三之完成時間（ $F(2, 2,090) = 224.95, p < .001$ ），顯示情境四的轉換圖形序列之歷程。接著，TB-TMT 中四個指標：視覺掃描時間、動作時間、抑制情境時間、轉換情境時間，四者的平均數各為 37 秒、39 秒、18 秒、10 秒，標準差各為 11.07、14.16、14.83、19.68。本研究四個指標中，抑制情境時間與轉換情境時間指標之計算公式與分數解釋，詳見研究工具之說明，而視覺掃描時間及動作時間為基本認知能力。以上結果橫跨四個不同年齡層之健康樣本在 TB-TMT 的整體表現。

再者，透過偏態與峰度係數瞭解資料之常態化，本量表之偏態係數介於 0.49 至 2.49 之間，屬於正偏態。峰度係數介於 1.74 至 12.92 之間，表示為高峽峰，分配呈現高聳型態。Curran 等人（1996）根據以機率統計理論為基礎的蒙地卡羅法（Monte Carlo method），透過數值計算提出多元常態分布，其指出偏態係數介於 2.0 至 3.0 之間，峰度係數介於 7.0 至 21.0 之間，其分布判定為中等程度偏離常態分配。余民寧（2006）指出當偏態係數大於 3.0、峰度係數大於 21.0 時，此資料被判定為嚴重偏離常態分配。根據上述可知，本研究資料的分配屬於中等偏離常態分配，判斷資料常態性的標準而言，即稍微偏離常態。在運用變數分配的偏態和峰度而言，有些微偏離常態之情況仍可正常的估計（陳寬裕，2023）。

承上述，針對使用於資料分析的統計方法對常態分配的重要性而言，相關文獻提及常態性不是特別的重要。Mishra 等人（2019）指出，在資料分析中的相關性、迴歸、 t 檢定和變異數分析，當樣本數達 100 個或以上，違反常態分配已不是主要關注的問題。再者，英國格拉斯哥大學（University of Glasgow）之心理與神經科學院的 PsyTeachR 團隊（2021）探討常態性檢定的本質，指出常態性在小樣本（ $N = 10$ ）中，易於呈現常態性的偏差，而在大樣本（ $N = 5,000$ ）中，易於因檢測與理論常態性呈現較微小的偏差，但實際上常態性檢定沒有實際用途。而來自義大利比薩大學（University of Pisa）的統計學教授 Paolo Frumento（2023）提出反對常態分配的原因，他指出現實中的數據很少會完全符合常態分配，測試常態性的假設大多為多餘的，不應成為判斷數據可靠性的標準。

(二) 信度分析—再測信度

為探討 TB-TMT 測驗分數在各個情境中所展現的再測穩定性，以效度研究樣本抽取 40 位研究參與者進行間隔六個月的再測。再測信度的樣本係採取立意取樣的方式，為同意參與測驗、聯繫順利，以及適合實施本測驗的研究參與者。其樣本分布北、中、南、東各地區，人數各 10 位。再測分析結果發現，情境 1 至情境 5 之再測係數呈顯著正相關 ($r = .40 \sim .55, ps < .05 \sim .001$)，顯示 TB-TMT 具有可接受的穩定性。

(三) 效度證據

1. 內部相關性

本測驗為操作式平板測量，根據標準化樣本進行相關分析，依各指標特性進行內部作業相關以求各指標測驗完成時間（秒）得分之相關係數，作為本測驗之內部相關性。以 Pearson 相關來考驗各指標之間的相關性，由表 4 可知，原始各情境間的相關皆達 .001 的顯著水準。經轉換為執行功能指標，以求之間的相關性，結果發現視覺掃描時間與動作時間、轉換時間達正相關 ($r = .07 \sim .21, ps < .05 \sim .001$)；動作時間與抑制時間達負相關 ($r = -.28, ps < .001$)；抑制時間與轉換時間為負相關 ($r = -.36, ps < .001$)。可知，透過各情境轉換為指標後，各指標成分能代表其特性。

表 4

TB-TMT 各情境間及指標間之相關係數摘要表 ($N = 1,046$)

	1	2	3	4	5
1. 情境一視覺掃描	—				
2. 情境二簡單圖形數字序列	.34***	—			
3. 情境三選擇圖形數字序列	.32***	.42***	—		
4. 情境四轉換圖形數字序列	.36***	.38***	.39***	—	
5. 情境五動作速度	.21***	.22***	.22***	.20***	—
1. 指標一視覺掃描時間	—				
2. 指標二動作時間	.21***	—			
3. 指標三抑制時間	-.06	-.28***	—		
4. 指標四轉換時間	.07*	-.01	-.36***	—	

* $p < .05$. ** $p < .01$. *** $p < .001$.

2. 效標關聯效度

在效標關聯效度方面，以整體健康樣本進行 TB-TMT 之效度檢驗，採用平板操作測驗之認知實作作業及自陳行為評量，包括：彩色路徑描繪測驗作業、西蒙測驗、塔測驗、反應時間，以及中文版執行功能指標。為使測驗進行效度檢驗以及比較測量，本研究將五個情境進行分數轉換為四個指標，指標 1 視覺掃描時間、指標 2 動作時間、指標 3 抑制情境時間、指標 4 轉換情境時間。

表 5 為 TB-TMT 與 RTT、SmT、TwT、CTT 及 CEFI 的相關係數。結果發現，簡單知覺排序與動作時間具有顯著正相關 ($r = .27, ps < .05$)，表示彩色路徑描繪測驗的簡單知覺排序所需時間愈短，其動作時間也愈短，意即知覺追蹤速度能力愈好。概念轉換時間與轉換情境時間具有顯著正相關 ($r = .31, ps < .05$)，顯示概念轉換時間愈快，轉換情境時間也愈短，意即認知彈性與轉換能力表現愈好，符合過去研究發現。抑制效果所需時間、違反規則次數與抑制情境時間具有顯著正相關 ($r = .26 \sim .38, ps < .05 \sim .001$)，表示西蒙測驗的抑制效果所需時間愈短，軌跡標示的抑制情境時間也愈短，意即抑制能力愈佳，以及塔測驗的違反規則次數愈少，抑制情境時間愈短，意即抑制能力愈好。轉換效果所需時間與轉換情境時間達顯著正相關 ($r = .16, ps < .05$)，表示西蒙測驗的轉換效果所需時間愈短，軌跡標示的轉換情境時間也愈短，轉換能力愈好，皆符合過去研究發現。選擇反應時間與視覺掃描時間有顯著正相關 ($r = .26, ps < .001$)，顯示選擇反應時間愈快，視覺掃描時間也愈短，意即視覺掃描表現愈佳，以及選擇移動時間與動作時間有顯著正相關 ($r = .17, ps < .001$)，顯

示選擇移動時間愈快，動作時間也愈短，意即動作移動能力愈佳，符合過去研究發現。策略行動、組織規劃及衝動控制與視覺掃描時間及抑制情境時間達顯著正相關 ($r = .17 \sim .32, ps < .001$)，表示執行功能指標能力愈佳，視覺掃描時間、抑制情境時間則愈短，執行功能指標具有良好的生態效度，其各別指標能力涵蓋多種執行功能，符合過去研究發現。再者，進一步檢驗四個計分指標的效度。結果發現，視覺掃描時間與選擇反應時間、策略行動、組織規劃、衝動控制有顯著正相關 ($r = .19 \sim .32, ps < .001$)。動作時間與簡單知覺排序時間、選擇移動時間有顯著正相關 ($r = .17 \sim .27, ps < .05 \sim .001$)。抑制情境時間與抑制效果、違反規則、策略行動、組織規劃、衝動控制有顯著正相關 ($r = .17 \sim .38, ps < .05 \sim .001$)。轉換情境時間與概念轉換時間、轉換效果有顯著正相關 ($r = .16 \sim .31, ps < .05$)。結果顯示，TB-TMT 具有效標關聯效度。

表 5
TB-TMT 與各效標關聯效度指標之相關分析表

	TB-TMT 執行功能指標與基本認知能力			
	視覺掃描 時間	動作時間	抑制情境 時間	轉換情境 時間
CTT ($N = 54$)				
簡單知覺排序時間	.20	.27*	-.02	.20
概念轉換時間	.23	.13	.13	.31*
SmT ($N = 625$)				
抑制效果	-.02	3.05	.26*	.11
轉換效果	-.09	-.01	-.18	.16*
TwT ($N = 545$)				
違反規則	.33***	.17**	.38***	.13*
RTT ($N = 771$)				
選擇反應時間	.26***	.14**	.30***	.05
選擇移動時間	.23***	.17***	.31***	.01
CEFI ($N = 461$)				
策略行動	.32***	-.01	.24***	-.05
組織規劃	.30***	-.04	.24***	-.06
衝動控制	.19***	-.08	.17***	-.02

註：CTT 為 Color Trails Test「彩色路徑描繪測驗」；SmT 為 Simon Task「西蒙作業」；TwT 為 Tower Task「塔測驗」；RTT 為 Reaction Time Task「反應時間作業」；CEFI 為「中文版執行功能指標」。

* $p < .05$. ** $p < .01$. *** $p < .001$.

(四) 性別及教育階段差異檢定

針對 TMT 樣本特性之差異檢定，大多研究係以認知障礙之高齡者樣本進行分析 (Fröhlich et al., 2024; Humphreys et al., 2017; Kobayashi et al., 2020)，較少採取健康樣本進行探究。過去研究顯示，軌跡標示 TMT-A 與 TMT-B 存在性別及教育程度之差異 (Cangoz et al., 2009; Tamayo et al., 2012)，亦有研究顯示二者經指標換算後皆無差異 (Cangoz et al., 2009; MacPherson et al., 2019)。本研究檢驗健康樣本之性別及教育階段之差異表現。

1. 性別考驗

在性別考驗中，男性與女性在情境四的表現具有顯著差異 ($t = 2.54, p < .05$)，女性 ($M = 64.08$) 在情境四所需時間低於男性 ($M = 67.35$)，在其他情境中，男性與女性的表現未達顯著差異。整體而言，女性在情境四之轉換圖形數字序列之能力表現較優於男生，而在其他情境中，不同性別之間並沒有差異，與過去研究結果相似 (如 Cangoz et al., 2009; Tamayo et al., 2012)。接著，分析性別與指標之差異表現。結果發現，各指標中男性與女性的表現不具有顯著差異，符合過去研究發現 (Cangoz et al., 2009; MacPherson et al., 2019; Tamayo et al., 2012)。表示性別在視覺掃描時間、動作時間、抑制情境時間與轉換情境時間等能力表現皆無差異 (如表 6)。

表 6
性別在 TB-TMT 各情境與情境完成時間分數之差異檢定

	男性 (N = 532)	女性 (N = 514)	t
各情境			
情境一-視覺掃描	37.37 (11.23)	36.40 (10.88)	1.42
情境二-簡單圖形數字序列	60.61 (19.68)	59.36 (17.46)	1.08
情境三-選擇圖形數字序列	52.64 (20.60)	50.80 (16.85)	1.59
情境四-轉換圖形數字序列	67.35 (23.54)	64.08 (17.88)	2.54*
情境五-動作速度	39.37 (13.79)	39.08 (14.54)	0.91
各指標			
指標一-視覺掃描時間	37.37 (11.23)	36.40 (10.88)	1.42
指標二-動作時間	39.37 (13.79)	39.08 (14.54)	0.32
指標三-抑制情境時間	18.26 (15.53)	17.34 (14.08)	1.00
指標四-轉換情境時間	10.73 (21.78)	9.00 (17.22)	1.43

* $p < .05$.

2. 教育階段組考驗

檢驗四組不同教育階段學生（國小中高年級、國中生、高中生、大學生含研究所以上）之間與各情境及指標是否有差異，結果如表 7 所示。結果顯示，各情境的完成時間分數在各年齡之間皆達顯著差異（ $F = 25.22 \sim 114.85, p < .001$ ）；事後比較大致趨勢顯示，教育年分增加，其完成測驗之時間則愈短，顯示 TB-TMT 在教育階段組之間具有顯著差異，符合過去研究發現（Cangoz et al., 2009）。接著，分析教育階段與各指標之差異性。結果發現，四個指標的完成時間分數在各教育階段之間皆達顯著差異，指標一（ $F(3, 1045) = 42.07, p < .001$ ）、指標二（ $F(3, 1,045) = 25.22, p < .001$ ）、指標三（ $F(3, 1,045) = 55.32, p < .001$ ）指標四（ $F(3, 1,045) = 4.60, p < .01$ ）。特別的是，國小生（中高年級）的轉換能力顯著優於高中生，此結果與相關研究相似，Ólafsdóttir 等人（2020）研究發現，學齡兒童到 12 歲時，其轉換能力已達成人水平。同樣類似研究發現，8 至 13 歲的學齡兒童表現出更高的轉換能力，則抑制和更新能力沒有差異，其推測此階段的學習者在動態技能與執行功能指標具有密切關聯性（Möhrling et al., 2022）。承上結果可知，就五個情境與四個指標而言，不同教育階段在各個情境與指標具有顯著差異。

表 7
教育階段在 TB-TMT 各情境與情境完成時間分數之差異檢定

	1 國小生 (N = 354)	2 國中生 (N = 280)	3 高中生 (N = 217)	4 大學生 (N = 195)	F	事後比較
各情境						
情境一	41.22	37.01	34.75	31.25	42.07***	1 > 2、3 > 4
視覺掃描	(11.83)	(11.51)	(9.25)	(6.90)		
情境二	71.05	57.48	52.47	51.92	83.22***	1 > 2 > 3、4
簡單圖形數字序列	(20.16)	(14.54)	(16.02)	(13.33)		
情境三	64.24	49.41	43.99	40.98	114.85***	1 > 2 > 3、4
選擇圖形數字序列	(21.54)	(13.43)	(14.18)	(10.66)		
情境四	74.54	63.84	60.68	58.13	37.86***	1 > 2、3、4
轉換圖形數字序列	(21.57)	(15.75)	(23.33)	(18.25)		
情境五	43.86	37.10	34.32	39.29	25.22***	1 > 2、3、4 4 > 3
動作速度	(16.10)	(12.30)	(12.00)	(12.56)		
各指標						
指標一	41.22	37.01	34.75	31.25	42.07***	1 > 2、3 > 4
視覺掃描時間	(11.83)	(11.51)	(9.25)	(6.90)		
指標二	43.89	37.10	34.32	39.29	25.22***	1 > 2、3、4 3 > 4
動作時間	(16.10)	(12.30)	(12.00)	(12.56)		

（續下頁）

表 7
教育階段在 TB-TMT 各情境與情境完成時間分數之差異檢定 (續)

	1 國小生 (N = 354)	2 國中生 (N = 280)	3 高中生 (N = 217)	4 大學生 (N = 195)	F	事後比較
指標三 抑制情境時間	25.09 (16.82)	16.39 (12.21)	13.69 (12.00)	11.18 (11.54)	55.32***	1 > 2、3、4
指標四 轉換情境時間	6.90 (23.72)	10.39 (14.64)	12.45 (19.77)	11.68 (17.09)	4.60**	3 > 1

** $p < .01$. *** $p < .001$.

(五) 常模

本研究分別建立指標一視覺掃描時間、指標二動作時間、指標三抑制情境時間及指標四轉換情境時間等四個指標的常模分數對照表，分別為國小中年級、國小高年級、國中、高中、大學以上的研究參與者編製四個指標的原始分數與量尺分數（平均數 10、標準差 3）之常模分數對照表，以符合測驗使用者的需求。透過對照表分數與常模，能提供評估者判斷與進一步釐清與確認研究參與者的指標表現。因全部樣本的常模資料較多，由於篇幅關係無法逐一贅述，詳細內容故呈現於指導手冊中，並以教育階段在指標一視覺掃描時間之量尺分數對照表作為範例提供參酌，見表 8。

表 8
教育階段在指標一視覺掃描時間之原始時間分數與量尺分數對照表 (N = 500)

教育階段	國小中年級	國小高年級	國中	高中	大學以上
量尺分數	原始分數				
1		> 73.91	> 72.51	> 80.28	> 48.92
2	> 71.61	73.91~69.24	72.51~66.49	80.28~49.24	48.92~44.43
3	71.61~67.12	69.24~66.08	66.49~61.99	49.24~45.59	44.43~40.47
4	67.12~61.41	66.08~66.04	61.99~58.29	45.59~44.79	40.47~40.24
5	61.41~55.72	66.04~56.39	58.29~54.35	44.79~41.32	40.24~39.10
6	55.72~52.22	56.39~48.84	54.35~44.87	41.32~38.19	39.10~35.49
7	52.22~48.60	48.84~46.13	44.87~41.56	38.19~34.07	35.49~33.24
8	48.60~43.48	46.13~40.94	41.56~39.16	34.07~31.81	33.24~30.45
9	43.48~40.14	40.94~37.14	39.16~36.25	31.81~30.17	30.45~27.84
10	40.14~38.10	37.14~34.15	36.25~31.51	30.17~26.45	27.84~24.92
11	38.10~35.31	34.15~31.80	31.51~29.13	26.45~24.37	24.92~22.33
12	35.31~31.45	31.80~28.96	29.13~25.91	24.37~21.14	22.33~20.90
13	31.45~28.10	28.96~25.39	25.91~24.65	21.14~19.19	20.90~19.14
14	28.10~25.41	25.39~24.66	24.65~22.05	19.19~18.33	19.14~18.11
15	25.41~22.85	24.66~22.09	22.05~20.32	18.33~15.00	18.11~16.76
16	22.85~19.51	22.09~18.87	20.32~19.48	15.00~14.16	16.76~16.23
17	≤ 19.51	18.87~18.55	19.48~18.60	14.16~11.67	16.23~14.50
18		≤ 18.55	≤ 18.60	≤ 11.67	≤ 14.50

討論

(一) 研究意涵

「軌跡標示測驗」(TMT) 主要評估個體運用執行視覺動作之認知彈性。本研究參考「D-KEFS 軌跡標示測驗」為基礎，發展以平板電腦操作之中文版電子化軌跡標示測驗 (TB-TMT)，並建立本土常模，歸納出五個情境，分別為情境一視覺掃描、情境二簡單圖形數字序列、情境三選擇圖形數字序列、情境四轉換圖形數字序列、情境五動作速度；再藉由情境之間的分數計算，得出較為純

粹的執行功能指標，以避免干擾物所影響之其他基本認知能力（Bowie & Harvey, 2006; Drane et al., 2002; Kao et al., 2023; Hashimoto et al., 2006），轉換得分後，計算得出兩項執行功能指標：抑制情境時間和轉換情境時間，以及兩項基本認知能力：視覺掃描時間和動作時間。本測驗建立過程經由理論架構之參考、標準化抽樣過程，建立國小至大學（含研究所）常模。在信度考驗中，再測信度分析有顯著正相關（ $r = .40 \sim .55, ps < .05 \sim .001$ ），表示具有良好的穩定性。在效度考驗部分，首先以內部相關性檢驗五個情境與四項執行功能指標之相關性，分析結果顯示，各情境之間皆具相關性；同時將各情境轉換為執行功能指標後，視覺掃描時間與動作時間、轉換情境時間呈現正相關，動作時間與抑制情境時間呈現負相關，抑制情境時間與轉換情境時間呈現負相關，表示各個指標成分能代表其特性，係能評估相對應能力之指標分數。接著，在效標關聯效度方面，本研究發現 TB-TMT 的視覺掃描時間與選擇反應時間、策略行動、組織規劃、衝動控制均達顯著正相關，此結果與過去國外研究一致（Brooks et al., 2017; Jahanshahi, 2003; MacPherson et al., 2019）。TB-TMT 的動作時間與簡單知覺排序時間、選擇移動時間達顯著正相關（Elkin-Frankston et al., 2007）。此外，TB-TMT 的執行功能指標（抑制情境時間、轉換情境時間）與不同效標關聯效度之測驗工具呈現低度至中度的顯著相關。在抑制時間指標的效度分析，塔測驗之違反規則指標，同樣涉及抑制能力（Delis et al., 2001）有中度正相關，抑制情境時間指標需要控制個體違反規則的能力，抑制能力愈好，其違反測驗規則的次數則愈低，且與涉及抑制效果的西蒙測驗也存在中度顯著相關，此結果支持抑制能力指標的獨立評分效果。在轉換情境時間指標的效度分析，轉換情境時間指標與同樣涉及轉換效果的西蒙測驗具有低度正相關（Delis et al., 2001），與概念轉換則有中度相關（Elkin-Frankston et al., 2007），也藉此驗證轉換情境時間指標的獨立評分效果。在自陳行為評量部分，策略行動涉及自我監控、策略擬定與執行，組織規劃涉及分配與歸納資源及維持注意力，衝動控制則涉及對他人不適的行為進行自我控制（游勝翔、林緯倫，2018），其執行功能指標屬於高階認知功能之集合，涵蓋多種混合的功能及能力，因此在視覺掃描時間、抑制情境時間皆具有低度相關，符合過去研究發現（Dekker et al., 2017）。結果呼應過去研究發現抑制情境時間指標與多項認知執行功能指標之關聯性，顯示 TB-TMT 具有良好的效標關聯效度。以上分析結果表示，TB-TMT 之四項基本認知與執行功能指標係能獨立評分之能力，適合評估個體執行功能與軌跡標示之數位化平板評估工具。

在性別與教育階段之考驗，本研究檢驗性別在 TB-TMT 各情境與指標之差異檢定，女性在情境四轉換圖形數字序列中表現顯著高於男性，則在其他情境的表現上，性別不存在差異性。此外，將各情境轉換為四項指標後，各指標的性別表現皆不具差異性，以上結果符合過去實徵研究發現（如 Bouattour, et al., 2017; Cangoz et al., 2009; Tamayo et al., 2012）。檢驗教育階段在 TB-TMT 各情境與指標之差異考驗，各情境的完成時間在各年齡區間具有差異性，表示隨著教育年之增加，其完成測驗的時間則愈快，符合過去研究發現（Cangoz et al., 2009）。再者，教育階段與各指標具有差異性，顯示出視覺掃描時間會隨著年齡增長，速度愈快；在動作時間與抑制情境時間表現上，自中學時期的青少年階段開始，動作時間及抑制情境時間較學齡兒童反應佳，在中學階段與成人之間無差異性，正如 Band 等人（2000）與 Davidson 等人（2006）研究指出，青少年階段和成人階段的研究參與者在動作時間與抑制情境時間之反應優於學齡兒童，他們確實發現，對於學齡兒童而言，抑制能力反應較為困難，需要持續更長的時間發展；亦即，抑制能力在中學階段開始趨於穩定。學齡兒童在 7 歲至 11 歲之間，是執行功能抑制和轉換能力發展的關鍵時期（Brydges et al., 2014; Diamond, 2002）。藉由認知經驗與大腦成熟變化的相互作用，促進學齡階段認知過程的發展（Richardson et al, 2018）。在轉換情境時間指標中，國小生的轉換能力顯著優於高中生，符合 Davidson 等人（2006）實徵研究分析的結果，發現在學齡前幼兒到成年人的受試者中，9 歲到 13 歲的學齡兒童較能準確地進行轉換能力。本研究也發現轉換能力在高中階段開始到成人階段趨於穩定，符合 Reimers 與 Maylor（2005）實徵研究結果，從 10 歲到成年中期，部分轉換能力則保持穩定。因此，本研究根據健康樣本之特性與結果，建立國小（中高年級）、國中、高中、大學（含研究所）以上之臺灣常模。

（二）研究貢獻

本研究所發展之「軌跡標示測驗平板版」（TB-TMT），為臺灣首創數位化操作型軌跡標示測驗。在研究貢獻與特色上，首先，在研究參與者部分，過去研究多以成年人至高齡者（Fröhlich et

al., 2024; Indorewalla et al., 2022; Magnusdottir et al., 2021; Siciliano et al., 2019; St-Hilaire et al., 2018) 為主, 本測驗適用對象涵蓋學齡兒童、中學生、高中生至成年人, 並建立本土常模, 可作為專業評量工作者、教師、家長或研究參與者本身瞭解執行功能表現。其次, 在測驗型態部分, 過去執行功能軌跡標示測驗多以紙本測驗 (Delis et al., 2001; Nasreddine et al., 2005; Reitan & Wolfson, 1985) 為主, 尚未有一套數位化平板操作之軌跡標示測驗, 本測驗操作方式以平板電腦進行測驗, 兼具數位電子化、建立本土常模之標準化測驗工具, 克服人工閱卷、計分之限制, 並設計改良為非語言文字層次之辨識, 改善測驗情境中涉及語言文字之限制, 研究參與者約可於 15 分鐘完成測驗。其三, 在測驗分析部分, 研究參與者透過本測驗, 可反映個體執行功能基礎能力以及高層次認知能力, 完成測驗後將原始資料置入報表分析系統中, 即可獲得該位研究參與者之測驗結果, 將此數據對照本土常模, 針對研究參與者之測驗表現與反應能力提供適當之回饋與評語, 藉以提供適其訓練方案, 提升能力表現。

(三) 研究限制與未來研究建議

首先, 本研究開發 TB-TMT 以建立認知功能發展至穩定階段之樣本, 年齡介於 8 至 29 歲, 此樣本集中在學齡兒童至成年人之健康樣本, 故其代表性與可推論範圍仍有受限。本研究團隊未來將擴大研究對象, 目前已著手設計發展適用於學齡前幼兒與中高齡者之 TB-TMT 版本, 檢驗 TB-TMT 在不同年齡樣本應用之普及性。同時建議未來研究包含適用本測驗的潛在認知缺陷族群樣本, 有待未來研究與測驗進一步開發與檢驗, 以探討不同條件樣本之發展階段的變化。其次, TB-TMT 係參考 D-KEFS-TMT 的設計編製原理, 以其任務情境與測量程序為基礎, 發展數位電子化 TMT 平板版, 本測驗涉及抑制情境時間、轉換情境時間等執行功能指標, 以及視覺掃描時間、動作時間等基本認知能力表現, 測驗結果無法推論至其他執行功能之特定指標。最後, 本研究開發 TB-TMT 是以數位化平板電腦作為資料蒐集之測驗工具, 在桌上型或筆記型電腦上安裝與設定應用程式模擬器, 即可開啟 TB-TMT, 但仍有其操作限制, 若以電腦作答則需以滑鼠操作進行, 與原測驗媒材平板電腦操作模式不同。若後續依原始資料所得測驗結果對照本土常模, 建議以平板電腦為主要測量媒材進行收案及施測。

以下提出四點實務建議。第一, 本平板版測驗在各情境之呈現畫面中, 皆有提供練習題與正式題的規則說明, 以及各情境任務的影片示範, 藉以使研究參與者清楚與理解測驗完整運行。因此, 建議在測驗開始前, 施測人員需進行測驗說明與作答方式; 在測驗過程中, 也需留意參與者實際執行任務的作答情況。其次, 本測驗能以個別及團體的方式進行施測與蒐集資料, 施測人員與參與者人數比例以一對十五為上限, 施測人員需隨時留意參與者在各情境的規則理解、作答反應, 適時提供個別的引導與說明。其三, 在臨床實務建議部分, 本測驗建立臨床實務之有效工具, 可用於篩檢認知功能受損與障礙之個體。藉由 TB-TMT 報表分析輸出測驗結果, 可初步評估個體執行功能指標表現推論大腦受損情況。其四, 在教學實務建議部分, 本研究建立學齡兒童至成年人階段之健康樣本常模, 以平板版認知評估為基礎的執行功能測驗, TB-TMT 具有可行性、可靠性與有效性。因此, 教學實務工作者可提供多樣性教育身分之學生 (例如一般生、特殊生與資優生) 進行測驗, 並針對 TB-TMT 的測驗結果與回饋評語, 瞭解學生執行功能指標中抑制能力、轉換能力, 以及基本認知能力中視覺掃描、動作速度等認知功能表現, 亦有助於檢測需要教學協助或學習輔導之學生, 以提供適合之執行功能訓練方案。

結論

TB-TMT 為我國首創發展平板版數位化操作之軌跡標示測驗, 適用於評估臺灣學齡兒童至成年人階段 (8 ~ 29 歲) 健康樣本之執行功能指標與基本認知能力。本測驗優勢在於施測便利、克服人工閱卷與計分之限制, 相較於自陳式量表而言, 較為客觀評估個體執行功能之認知表現, 並能有效蒐集執行功能指標中抑制情境時間與轉換情境時間, 以及基本認知能力中視覺掃描時間與動作時間等能力表現。再者, TB-TMT 為標準化測驗程序, 兼具計分一致性與自動化報表數據整合性, 並建立具有代表性的本土臺灣常模, 可作為測驗結果之分數解釋。本研究發展 TB-TMT 以建立一套完整

的數位化執行功能測驗暨分析系統，適用於個體在執行功能初步篩檢，有利於教學實務參考、學術研究認知功能成效分析，以及檢視臨床相關功能等多元情境，有助於提供學校端、研究界與臨床端重要研究實踐與實務應用價值。

參考文獻

- 內政部戶政司（2019）：〈縣市人口性比例及人口密度〉。 <https://www.ris.gov.tw/app/portal/346>
[Department of Household Registration, Ministry of the Interior. (2019). *Xianshi renkou xing bili ji renkou midu*. <https://www.ris.gov.tw/app/portal/346>]
- 內政部戶政司（2023）：〈十五歲以上人口教育程度〉。 <https://www.ris.gov.tw/app/portal/346>
[Department of Household Registration, Ministry of the Interior. (2023). *Shiwusui yishang renkou jiaoyu chengdu*. <https://www.ris.gov.tw/app/portal/346>]
- 余民寧（2006）：《潛在變項模式：SIMPLIS 的應用》。高等教育。[Yu, M.-N. (2006). *Latent variable models: The application of SIMPLIS*. Higher Education.]
- 張之捷、林緯倫、高詩姮（2023）：〈國英雙語平衡程度對臺灣大學生執行功能與兩類創造力表現的影響〉。《教育科學研究期刊》，68（4），191–225。 [https://doi.org/10.6209/JORIES.202312_68\(4\).0007](https://doi.org/10.6209/JORIES.202312_68(4).0007) [Chang, C.-C., Lin, W.-L., & Kao, S.-H. (2023). The influences of balanced Mandarin-English bilinguals on executive function and two types of reativity in Taiwanese University Students. *Journal of Research in Education Sciences*, 68(4), 191-225. [https://doi.org/10.6209/JORIES.202312_68\(4\).0007](https://doi.org/10.6209/JORIES.202312_68(4).0007)]
- 陳寬裕（2023）：《結構方程模型：運用 AMOS 分析（第二版）》。五南圖書。[Chen, K.-Y. (2023). *Jiegou fangcheng moxing: Yunyong AMOS fenxi* (2nd ed.). Wu-Nan Book.]
- 郭曉燕、花茂琴（2015）：《彩色路徑描繪測驗（中文版）指導手冊》。中國行為科學社。[Kuo, H.-Y., & Hua, M.-S. (2015). *Color Trails Test (Chinese Version) instruction manual*. Chinese Behavioral Science Corporation.]
- 游勝翔、林緯倫（2018）：〈「中文版執行功能指標」（大學生版與國中生版）的編製與其信效度報告〉。《測驗學刊》，65，29–67。[Yu, S.-H., & Lin, W.-L. (2018). The development and validation of Chinese executive function index-for college students (CEFI-C) and for junior high school students (CEFI-J). *Psychological Testing*, 65, 29–67.]
- 簡馨瑩（2020）：〈「記憶－抑制控制」活動融入語文教學對幼兒在執行功能與口語理解表現的效果研究〉。《教育科學研究期刊》，65（4），275–304。 [https://doi.org/10.6209/JORIES.202012_65\(4\).0009](https://doi.org/10.6209/JORIES.202012_65(4).0009) [Chien, H.-Y. (2020). Effects of memory-inhibitory control activity with embedded repeated read-aloud programs on executive function and oral comprehension ability of preschoolers. *Journal of Research in Education Sciences*, 65(4), 275-304. [https://doi.org/10.6209/JORIES.202012_65\(4\).0009](https://doi.org/10.6209/JORIES.202012_65(4).0009)]
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, 8(2), 71–82. <https://doi.org/10.1076/chin.8.2.71.8724>
- Army Individual Test Battery. (1944). *Manual of directions and scoring*. War Department, Adjutant

General's Office.

- Ashendorf, L., Jefferson, A. L., O'Connor, M. K., Chaisson, C., Green, R. C., & Stern, R. A. (2008). Trail making test errors in normal aging, mild cognitive impairment, and dementia. *Archives of Clinical Neuropsychology*, *23*(2), 129–137. <https://doi.org/10.1016/j.acn.2007.11.005>
- Band, G. P. H., van der Molen, M. W., Overtoom, C. C. E., & Verbaten, M. N. (2000). The ability to activate and inhibit speeded responses: Separate developmental trends. *Journal of Experimental Child Psychology*, *75*(4), 263–290. <https://doi.org/10.1006/jecp.1999.2538>
- Baughman, F. D., & Cooper, R. P. (2007). Inhibition and young children's performance on the Tower of London task. *Cognitive Systems Research*, *8*(3), 216–226. <https://doi.org/10.1016/j.cogsys.2007.06.004>
- Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327–336. <https://doi.org/10.1016/j.lindif.2011.01.007>
- Bouattour, N., Farhat, N., Hadjkacem, H., Hdiji, O., Dammak, M., & Mhiri, C. (2017). Trail making test (TMT): Tunisian normative values from 339 normal adult controls. *Journal of the Neurological Sciences*, *381*(Sup), 846. <https://doi.org/10.1016/j.jns.2017.08.2381>
- Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols*, *1*(5), 2277–2281. <https://doi.org/10.1038/nprot.2006.390>
- Brooks, J., Seeanner, J., Hennessy, S., Manganelli, J., Crisler, M., Rosopa, P., Jenkins, C., Anderson, M., Drouin, N., Belle, L., Truesdail, C., & Tanner, S. (2017). Interactive tools for measuring visual scanning performance and reaction time. *The American Journal of Occupational Therapy*, *71*(2), Article 7102350010. <https://doi.org/10.5014/ajot.2017.020461>
- Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. *Intelligence*, *47*, 34–43. <https://doi.org/10.1016/j.intell.2014.08.010>
- Burle, B., Vidal, F., Tandonnet, C., & Hasbroucq, T. (2004). Physiological evidence for response inhibition in choice reaction time tasks. *Brain and Cognition*, *56*(2), 153–164. <https://doi.org/10.1016/j.bandc.2004.06.004>
- Butler, M., Retzlaff, P. D., & Vanderploeg, R. (1991). Neuropsychological test usage. *Professional Psychology: Research and Practice*, *22*(6), 510–512. <https://doi.org/10.1037/0735-7028.22.6.510>
- Cangoz, B., Karakoc, E., & Seleklir, K. (2009). Trail making test: Normative data for Turkish elderly population by age, sex and education. *Journal of the Neurological Sciences*, *283*(1), 73–78. <https://doi.org/10.1016/j.jns.2009.02.313>
- Christidi, F., Kararizou, E., Triantafyllou, N., Anagnostouli, M., & Zalonis, I. (2015). Derived trail making test indices: Demographics and cognitive background variables across the adult life span. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *22*(6), 667–678. <https://doi.org/10.1080/13825585.2015.1027650>

- Crone, E. A., Bunge, S. A., van der Molen, M. W., & Ridderinkhof, K. R. (2006). Switching between tasks and responses: A developmental study. *Developmental Science*, *9*(3), 278–287.
<https://doi.org/10.1111/j.1467-7687.2006.00490.x>
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29.
<https://doi.org/10.1037/1082-989X.1.1.16>
- D'Elia, L. F., Satz, P., Uchiyama, C. L., & White, T. (1996). *Color trails test*. Psychological Assessment Resources.
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078.
<https://doi.org/10.1016/j.neuropsychologia.2006.02.006>
- Deary, I. J., Liewald, D., & Nissan, J. (2011). A free, easy-to-use, computer-based simple and four-choice reaction time programme: The Deary-Liewald reaction time task. *Behavior Research Methods*, *43*(1), 258–268. <https://doi.org/10.3758/s13428-010-0024-1>
- Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, *11*(3), 201–211. <https://doi.org/10.1038/nrn2793>
- Dekker, M. C., Ziermans, T. B., Spruijt, A. M., & Swaab, H. (2017). Cognitive, parent and teacher rating measures of executive functioning: Shared and unique influences on school achievement. *Frontiers in Psychology*, *8*, Article 48. <https://doi.org/10.3389/fpsyg.2017.00048>
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System: Examiner's Manual*. The Psychological Corporation.
- Demeyere, N., Haupt, M., Webb, S. S., Strobel, L., Milosevich, E. T., Moore, M. J., Wright, H., Finke, K., & Duta, M. D. (2021). Introducing the tablet-based Oxford cognitive screen-plus (OCS-Plus) as an assessment tool for subtle cognitive impairments. *Scientific Reports*, *11*(1), Article 8000.
<https://doi.org/10.1038/s41598-021-87287-8>
- Demeyere, N., Riddoch, M. J., Slavkova, E. D., Bickerton, W.-L., & Humphreys, G. W. (2015). The Oxford cognitive screen (OCS): Validation of a stroke-specific short cognitive screening tool. *Psychological Assessment*, *27*(3), 883–894. <https://doi.org/10.1037/pas0000082>
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780195134971.003.0029>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168.
<https://doi.org/10.1146/annurev-psych-113011-143750>
- Drane, D. L., Yuspeh, R. L., Huthwaite, J. S., & Klingler, L. K. (2002). Demographic characteristics and normative observations for derived-trail making test indices. *Neuropsychiatry, Neuropsychology,*

- and Behavioral Neurology*, 15(1), 39–43.
- Elkin-Frankston, S., Lebowitz, B. K., Kapust, L. R., Hollis, A. M., & O'Connor, M. G. (2007). The use of the color trails test in the assessment of driver competence: Preliminary report of a culture-fair instrument. *Archives of Clinical Neuropsychology*, 22(5), 631–635.
<https://doi.org/10.1016/j.acn.2007.04.004>
- Fisher, A. V. (2011). Automatic shifts of attention in the dimensional change card sort task: Subtle changes in task materials lead to flexible switching. *Journal of Experimental Child Psychology*, 108(1), 211–219. <https://doi.org/10.1016/j.jecp.2010.07.001>
- Fröhlich, S., Müller, K., & Voelcker-Rehage, C. (2024). Normative data for the CERAD-NP for healthy high-agers (80–84 years) and effects of age-typical visual impairment and hearing loss. *Journal of the International Neuropsychological Society*, 30(7), 697–709.
<https://doi.org/10.1017/S1355617721001284>
- Frumento, P. (2023). *Against normality testing*. Cornell University. <https://doi.org/10.48550/arXiv.2302.11536>
- Gnys, J. A., & Willis, W. G. (1991). Validation of executive function tasks with young children. *Developmental Neuropsychology*, 7(4), 487–501. <https://doi.org/10.1080/87565649109540507>
- Hagen, K., Ehrlis, A.-C., Haecussinger, F. B., Heinzl, S., Dresler, T., Mueller, L. D., Herrmann, M. J., Fallgatter, A. J., & Metzger, F. G. (2014). Activation during the trail making test measured with functional near-infrared spectroscopy in healthy elderly subjects. *NeuroImage*, 85(1), 583–591.
<https://doi.org/10.1016/j.neuroimage.2013.09.014>
- Hashimoto, R., Meguro, K., Lee, E., Kasai, M., Ishii, H., & Yamaguchi, S. (2006). Effect of age and education on the trail making test and determination of normative data for Japanese elderly people: The Tajiri Project. *Psychiatry and Clinical Neurosciences*, 60(4), 422–428.
<https://doi.org/10.1111/j.1440-1819.2006.01526.x>
- Heilbronner, R. L., Henry, G. K., Buck, P., Adams, R. L., & Fogle, T. (1991). Lateralized brain damage and performance on trail making A and B, digit span forward and backward, and TPT memory and location. *Archives of Clinical Neuropsychology*, 6(4), 251–258.
[https://doi.org/10.1016/0887-6177\(91\)90002-Q](https://doi.org/10.1016/0887-6177(91)90002-Q)
- Hu, Y., Li, J., Zhao, Y., Dong, Z., Qiu, P., Yang, S., Xu, H., Kuang, W., & Gong, Q. (2023). Memory and processing speed impairments in first-episode drug-naïve patients with major depressive disorder. *Journal of Affective Disorders*, 322, 99–107. <https://doi.org/10.1016/j.jad.2022.10.048>
- Humphreys, G. W., Duta, M. D., Montana, L., Demeyere, N., McCrory, C., Rohr, J., Kahn, K., Tollman, S., & Berkman, L. (2017). Cognitive function in low-income and low-literacy settings: Validation of the tablet-based Oxford cognitive screen in the health and aging in Africa: A longitudinal study of an INDEPTH community in South Africa (HAALSI). *The Journals of Gerontology: Series B*, 72(1), 38–50. <https://doi.org/10.1093/geronb/gbw139>
- Indorewalla, K. K., Osher, J., Lanca, M., Kartik, R., Vaidya, N., & Moncata, S. (2022). A normative study

- of the color trails test in the adult Indian population. *Applied Neuropsychology: Adult*, 29(5), 899–906. <https://doi.org/10.1080/23279095.2020.1819279>
- Iosa, M., Demeyere, N., Abbruzzese, L., Zoccolotti, P., & Mancuso, M. (2022). Principal component analysis of Oxford cognitive screen in patients with stroke. *Frontiers in Neurology*, 13, Article 779679. <https://doi.org/10.3389/fneur.2022.779679>
- Jahanshahi, M. (2003). Reaction time as an index of motor preparation/programming and speed of response initiation. In M. Hallett (Ed.), *Handbook of clinical neurophysiology: Movement disorders* (pp. 203–229). Elsevier. [https://doi.org/10.1016/s1567-4231\(09\)70163-7](https://doi.org/10.1016/s1567-4231(09)70163-7)
- Kao, S.-C., Baumgartner, N. W., Pritt, T., Wu, S., Schmitt, S., Ullrich-French, S., & Wang, C.-H. (2023). Acute effects of mindful interval exercise on cognitive performance in a higher education setting. *Psychology of Sport and Exercise*, 64, Article 102326. <https://doi.org/10.1016/j.psychsport.2022.102326>
- Kobayashi, L. C., Farrell, M. T., Payne, C. F., Mall, S., Montana, L., Wagner, R. G., Kahn, K., Tollman, S., & Berkman, L. F. (2020). Adverse childhood experiences and domain-specific cognitive function in a population-based study of older adults in rural South Africa. *Psychology and Aging*, 35(6), 818–830. <https://doi.org/10.1037/pag0000552>
- Liao, Y.-H., & Chen, H.-C. (2022). Happiness takes effort: Exploring the relationship among academic grit, executive functions and well-being. *Personality and Individual Differences*, 199, Article 111863. <https://doi.org/10.1016/j.paid.2022.111863>
- MacPherson, S. E., Allerhand, M., Cox, S. R., & Deary, I. J. (2019). Individual differences in cognitive processes underlying Trail Making Test-B performance in old age: The Lothian Birth Cohort 1936. *Intelligence*, 75, 23–32. <https://doi.org/10.1016/j.intell.2019.04.001>
- MacPherson, S. E., Cox, S. R., Dickie, D. A., Karama, S., Starr, J. M., Evans, A. C., Bastin, M. E., Wardlaw, J. M., & Deary, I. J. (2017). Processing speed and the relationship between Trail Making Test-B performance, cortical thinning and white matter microstructure in older adults. *Cortex*, 95, 92–103. <https://doi.org/10.1016/j.cortex.2017.07.021>
- Magnusdottir, B. B., Haraldsson, H. M., & Sigurdsson, E. (2021). Trail making test, stroop, and verbal fluency: Regression-based norms for the Icelandic population. *Archives of Clinical Neuropsychology*, 36(2), 253–266. <https://doi.org/10.1093/arclin/acz049>
- Malone, S. A., Pritchard, V. E., & Hulme, C. (2022). Domain-specific skills, but not fine-motor or executive function, predict later arithmetic and reading in children. *Learning and Individual Differences*, 95, Article 102141. <https://doi.org/10.1016/j.lindif.2022.102141>
- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, 22(1), 67–72. https://doi.org/10.4103/aca.ACA_157_18
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21(1), 8–14.

<https://doi.org/10.1177/0963721411429458>

- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>
- Möhring, W., Klupp, S., Ludyga, S., & Grob, A. (2022). Executive functions in children engaging in open- and closed-skilled sports. *Psychology of Sport and Exercise*, *61*, Article 102218. <https://doi.org/10.1016/j.psychsport.2022.102218>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Ólafsdóttir, I. M., Gestsdóttir, S., & Kristjánsson, Á. (2020). Age differences in foraging and executive functions: A cross-sectional study. *Journal of Experimental Child Psychology*, *198*, Article 104910. <https://doi.org/10.1016/j.jecp.2020.104910>
- Ozcan, H., Ozer, S., & Yagcioglu, S. (2016). Neuropsychological, electrophysiological and neurological impairments in patients with obsessive compulsive disorder, their healthy siblings and healthy controls: Identifying potential endophenotype(s). *Psychiatry Research*, *240*, 110–117. <https://doi.org/10.1016/j.psychres.2016.04.013>
- Partington, J. E., & Leiter, R. G. (1949). *Partington Pathways Test*. Psychological Services Center.
- Periáñez, J. A., Ríos-Lago, M., Rodríguez-Sánchez, J. M., Adrover-Roig, D., Sánchez-Cubillo, I., Crespo-Facorro, B., Quemada, J. I., & Barceló, F. (2007). Trail making test in traumatic brain injury, schizophrenia, and normal ageing: sample comparisons and normative data. *Archives of Clinical Neuropsychology*, *22*(4), 433–447. <https://doi.org/10.1016/j.acn.2007.01.022>
- PsyTeachR Team. (2021). *PsyTeachR tutorials*. <https://psyteachr.github.io/tutorials/>
- Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CRSP’s impact on low-income preschoolers’ preacademic skills: Self-regulation as a mediating mechanism. *Child Development*, *82*(1), 362–378. <https://doi.org/10.1111/j.1467-8624.2010.01561.x>
- Reimers, S., & Maylor, E. A. (2005). Task switching across the life span: Effects of age on general and specific switch costs. *Developmental Psychology*, *41*(4), 661–671. <https://doi.org/10.1037/0012-1649.41.4.661>
- Reitan, R. M. (1992). *Trail Making Test: Manual for administration and scoring*. Reiten Neuropsychological Laboratory.
- Reitan, R. M., & Wolfson, D. (1985). *The halsted-Reitan neuropsychological test battery: Theory and clinical interpretation*. Neuropsychology Press.
- Richardson, C., Anderson, M., Reid, C., & Fox, A. M. (2018). Development of inhibition and switching: A longitudinal study of the maturation of interference suppression and reversal processes during

- childhood. *Developmental Cognitive Neuroscience*, 34, 92–100.
<https://doi.org/10.1016/j.dcn.2018.03.002>
- Rodríguez-Nieto, G., Seer, C., Sidlauskaitė, J., Vleugels, L., Roy, A. V., Hardwick, R., & Swinnen, S. (2022). Inhibition, shifting and updating: Inter and intra-domain commonalities and differences from an executive functions activation likelihood estimation meta-analysis. *NeuroImage*, 264, Article 119665. <https://doi.org/10.1016/j.neuroimage.2022.119665>
- Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence*, 39(4), 222–232. <https://doi.org/10.1016/j.intell.2011.03.001>
- Shibuya-Tayoshi, S., Sumitani, S., Kikuchi, K., Tanaka, T., Tayoshi, S., Ueno, S.-I., & Ohmori, T. (2007). Activation of the prefrontal cortex during the Trail-Making Test detected with multichannel near-infrared spectroscopy. *Psychiatry and Clinical Neurosciences*, 61(6), 616–621.
<https://doi.org/10.1111/j.1440-1819.2007.01727.x>
- Siciliano, M., Chiorri, C., Battini, V., Sant'Elia, V., Altieri, M., Trojano, L., & Santangelo, G. (2019). Regression-based normative data and equivalent scores for trail making test (TMT): An updated Italian normative study. *Neurological Sciences*, 40, 469–477.
<https://doi.org/10.1007/s10072-018-3673-y>
- Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 31–86). Elsevier. [https://doi.org/10.1016/S0166-4115\(08\)61218-2](https://doi.org/10.1016/S0166-4115(08)61218-2)
- Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, 51(3), 300–304.
<https://doi.org/10.1037/h0020586>
- Stanczak, D. E., & Triplett, G. (2003). Psychometric properties of the mid-range expanded trail making test an examination of learning-disabled and non-learning-disabled children. *Archives of Clinical Neuropsychology*, 18(2), 107–120. [https://doi.org/10.1016/S0887-6177\(01\)00190-1](https://doi.org/10.1016/S0887-6177(01)00190-1)
- St-Hilaire, A., Parent, C., Potvin, O., Bherer, L., Gagnon, J.-F., Joubert, S., Belleville, S., Wilson, M. A., Koski, L., Rouleau, I., Hudon, C., & Macoir, J. (2018). Trail making tests A and B: Regression-based normative data for Quebec French-speaking mid and older aged adults. *The Clinical Neuropsychologist*, 32(sup1), 77–90. <https://doi.org/10.1080/13854046.2018.1470675>
- Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: A conceptual view. *Psychological Research*, 63(3-4), 289–298. <https://doi.org/10.1007/s004269900007>
- Sutin, A. R., Sesker, A. A., Stephan, Y., & Terracciano, A. (2022). Socioeconomic status, internalizing and externalizing behaviors, and executive function in adolescence: A longitudinal study with multiple informants. *Psychiatry Research*, 313, Article 114585. <https://doi.org/10.1016/j.psychres.2022.114585>
- Tamayo, F., Casals-Coll, M., Sánchez-Benavides, G., Quintana, M., Manero, R. M., Rognoni, T., Calvo, L., Palomo, R., Aranciva, F., & Peña-Casanova, J. (2012). Spanish normative studies in a young adult

- population (NEURONORMA young adults project): Guidelines for the span verbal, span visuo-spatial, Letter-Number Sequencing, Trail Making Test and Symbol Digit Modalities Test. *Neurologia (English Edition)*, 27(6), 319–329. <https://doi.org/10.1016/j.nrleng.2012.07.008>
- Verbruggen, F., Liefvooghe, B., Notebaert, W., & Vandierendonck, A. (2005). Effects of stimulus-stimulus compatibility and stimulus-response compatibility on response inhibition. *Acta Psychologica*, 120(3), 307–326. <https://doi.org/10.1016/j.actpsy.2005.05.003>
- Wang, H. F., Tian, S., Yan, R., Tang, H., Shi, J. B., Zhu, R., Chen, Y., Han, Y. L., Chen, Z. L., Zhou, H. L., Yao, Z., & Lu, Q. (2023). Convergent and divergent cognitive impairment of unipolar and bipolar depression: A magnetoencephalography resting-state study. *Journal of Affective Disorders*, 321, 8–15. <https://doi.org/10.1016/j.jad.2022.09.126>
- Webb, S. S., Hobden, G., Roberts, R., Chiu, E. G., King, S., & Demeyere, N. (2022). Validation of the UK english Oxford cognitive screen-plus in sub-acute and chronic stroke survivors. *European Stroke Journal*, 7(4), 476–486. <https://doi.org/10.1177/23969873221119940>
- Welsh, M. C., & Huizinga, M. (2001). The development and preliminary validation of the tower of hanoi-revised. *Assessment*, 8(2), 167–176. <https://doi.org/10.1177/107319110100800205>
- Woods, D. L., Wyma, J. M., Herron, T. J., & Yund, E. W. (2015). The effects of aging, malingering, and traumatic brain injury on computerized trail-making test performance. *PLoS ONE*, 10(6), Article e0124345. <https://doi.org/10.1371/journal.pone.0124345>
- Zeng, Z., Miao, C., Leung, C., & Shen, Z. (2017). Computerizing trail making test for long-term cognitive self-assessment. *International Journal of Crowd Science*, 1(1), 83–99. <https://doi.org/10.1108/IJCS-12-2016-0002>

收稿日期：2023 年 07 月 13 日

一稿修訂日期：2023 年 08 月 07 日

二稿修訂日期：2023 年 10 月 06 日

三稿修訂日期：2023 年 10 月 31 日

四稿修訂日期：2024 年 01 月 17 日

五稿修訂日期：2024 年 01 月 23 日

接受刊登日期：2024 年 01 月 23 日

Bulletin of Educational Psychology, 2024, 56(2), 231–260
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

Reliability and Validity of Tablet-Based Trail Making Test and Establishment of Normative Data

Shih-Yuan Huang¹, Po-Sheng Huang^{2,3}, Zhi-Hong Chen⁴,
Pei-Zhen Chen¹, and Hsueh-Chih Chen^{1,5,6,7}

The Trail Making Test (TMT) is widely used to evaluate higher-order executive cognitive functions and basic cognitive abilities. This test is suitable for both healthy individuals and patients with brain injuries or specific functional impairments. The TMT consists of a visual cancellation task and a series of exercises requiring the sequential connection of numbered and lettered circles. It evaluates several aspects of executive functioning, such as visual scanning, motor speed, inhibition, and switching. In clinical neuropsychology, the TMT is an effective measure for distinguishing the cognitive performance of healthy individuals from that of patients. Its ease of administration, practicality, and sensitivity make it a valuable tool for diagnosing cognitive disorders and neurological injuries. Overall, the TMT is a crucial instrument for evaluating executive function.

In recent years, researchers worldwide have utilized the TMT to establish local normative data. Most studies have focused on adults, particularly middle-aged and older individuals (Indorewalla et al., 2022; Magnusdottir et al., 2021; Siciliano et al., 2019; St-Hilaire et al., 2018) as well as seniors (Fröhlich et al., 2024). However, research providing local normative data for school-aged children, middle school students, high school students, and younger adults remains scarce. Additionally, most of these studies have sample sizes ranging from approximately 200 to 1000 participants. In Taiwan, Kuo and Hua (2015) translated and adapted the Color Trails Test (CTT), originally developed by D' Elia et al. (1996), into a Chinese-language version. This adaptation applied TMT principles to establish normative data for Taiwan's middle-aged and older populations. Their data were derived from a study conducted on healthy adults aged over 50 years. While this research provided a valuable foundation for neuropsychological assessments of executive and frontal lobe function, it did not include data from diverse age groups. Furthermore, the study employed a paper-based format, which is both time-consuming and labor-intensive, and included only two subtests, limiting its scope for assessing cognitive and basic abilities.

In Taiwan, a significant research gap exists due to the lack of a digital version of the TMT designed for tablet administration. Therefore, developing a Chinese-language TMT for tablet use is essential to establish local normative data. This study introduced the Tablet-Based TMT (TB-TMT), a digital adaptation of the TMT, comprising five conditions. Condition 1, Visual Scanning evaluates visual scanning and attention through a quick and focused task. Condition 2, Simple Symbol

¹ Department of Educational Psychology and Counseling, National Taiwan Normal University

² Graduate Institute of Digital, National Taiwan University of Science and Technology

³ Empower Vocational Education Research Center, National Taiwan University of Science and Technology

⁴ Graduate Institute of Information and Computer Education, National Taiwan Normal University

⁵ Institute for Research Excellence in Learning Sciences, National Taiwan Normal University

⁶ Chinese Language and Technology Center, National Taiwan Normal University

⁷ Social Emotional Education Development Center, National Taiwan Normal University

Corresponding author:

Hsueh-Chih Chen, Department of Educational Psychology and Counseling, National Taiwan Normal University. Email: hcjyh@ntnu.edu.tw

Number Sequencing requires participants to sequentially connect numbers in a visual-motion task. This condition assesses basic numerical processing, visual scanning, attention, and motion. Condition 3, Selective Symbol Number Sequencing involves sequencing symbol-number pairs in a visual-motion task. This condition provides a foundational measure of the ability to alternate between symbols and numbers. Condition 4, Switching Symbol Number Sequencing is a core component of the TMT. Test takers must alternate between connecting symbols and numbers in a specific sequence. This condition evaluates higher-order abilities, such as multitasking, simultaneous processing, and divided attention, serving as a key measure of cognitive flexibility within executive functioning. Condition 5, Motor Speed assesses the foundational aspects of cognitive flexibility by measuring motor reaction times. Individuals with motor function disorders involving the central nervous system or related areas may demonstrate poorer line-drawing performance, leading to lower scores across most conditions, particularly in Condition 4.

Through calculations based on scores for the different conditions, we derived four indicators. The first indicator—Visual Scanning Time, reflected in Condition 1—measures performance in visual scanning and attention, with shorter times indicating better performance. The second indicator—Motion Time, derived from Condition 5—examines reaction performance, where shorter times indicate better reaction performance. The third indicator, Inhibition Time, is obtained as follows: $(\text{Condition 2} + \text{Condition 3})/2 - (\text{Condition 1} + \text{Condition 5})/2$. Shorter times represent stronger inhibition skills. The fourth indicator, Switching Time, is obtained as follows: $\text{Condition 4} - (\text{Condition 2} + \text{Condition 3})/2$. Shorter times reflect higher switching skills.

This study included 1,046 participants aged 8–29 years, divided into four age groups: primary school students in Grades 3–6, middle school students, high school students, and college students (including graduate and postgraduate levels). To ensure the sample's structural features were representative of the population, stratified random sampling was conducted based on variables such as age, region, and gender. This approach aimed to evaluate the reliability and validity of the TB-TMT and establish normative data. The study was approved by the Research Ethics Committee of Taiwan Normal University. All participants were informed of the research procedures and provided written informed consent before participation. Upon completing the test, participants received a small gift as a token of appreciation.

This study employed practical cognitive tasks and self-report behavior assessment scales—including the CTT, Simon Task (SmT), Tower Task (TwT), Reaction Time Task (RTT), and Chinese Executive Function Index (CEFI)—to evaluate the criterion-related validity of the TB-TMT. A reliability analysis was conducted to assess test–retest reliability, and internal correlation analysis was performed to examine the relationships among the indicators derived from the TB-TMT. To further evaluate criterion validity, established tools with proven reliability and validity—such as the CTT, SmT, TwT, RTT, and CEFI—were utilized. The results demonstrated that the test–retest reliability of the TB-TMT ranged from .40 to .55, indicating acceptable stability. Additionally, each indicator effectively reflected its intended construct and exhibited appropriate internal correlations. Correlation coefficients ranging from .17 to .38, obtained using the CTT, SmT, TwT, RTT, and CEFI as validity criteria, confirmed that the TB-TMT possesses satisfactory criterion-related validity.

The TB-TMT developed in this study represents Taiwan's first digital adaptation of the Trail Making Test (TMT). It offers several unique contributions, including its emphasis on diverse participant demographics, its innovative digital test format, and its advanced methodology for analyzing test results.

First, unlike previous studies that primarily focused on adults and seniors (Fröhlich et al., 2024; Indorewalla et al., 2022; Magnusdottir et al., 2021; Siciliano et al., 2019; St-Hilaire et al., 2018), the present study includes school-aged children, middle school students, high school students, and adults. By establishing local normative data, this study provides a valuable resource for professionals, educators, parents, and researchers seeking to assess and understand executive functioning.

Second, regarding the test format, most existing TMTs used to assess executive function are paper-based (Delis et al., 2001; Nasreddine et al., 2005; Reitan & Wolfson, 1985), and no tablet-based digital TMT had been developed prior to this study. The tablet-based test developed in this study is a standardized digital tool designed to establish local normative data. It overcomes the limitations of manual scoring with an enhanced design that incorporates non-text identification. This feature eliminates text-related constraints during testing and allows participants to complete the test within 15 minutes.

Third, the test evaluates individuals' executive function and higher-level cognitive abilities. Upon completion, the raw data are transmitted to a report analysis system, which compares the participants' results against the established local normative data. The system then generates detailed feedback and comments on the participants' performance and response abilities, as well as

tailored recommendations for training to enhance their cognitive skills and improve future performance.

The TB-TMT is Taiwan's first tablet-based digital version of the TMT. It is designed to assess executive function indicators and basic cognitive abilities in healthy individuals aged 8–29 years, spanning school-aged children to adults. This digital test offers convenient administration and overcomes the limitations of manual scoring. Compared to self-report scales, the TB-TMT provides a more objective evaluation of an individual's executive function. It effectively gathers data on executive function indicators such as inhibition and switching time, as well as basic cognitive skills, including visual scanning time and motor speed. Furthermore, the standardized testing procedure of the TB-TMT ensures consistent scoring and integrates automated report generation. The TB-TMT establishes representative local normative data in Taiwan, providing a solid foundation for interpreting test results.

This study developed the TB-TMT as a comprehensive digital system for evaluating and analyzing executive function. The TB-TMT is suitable for preliminary screening of executive function and can be applied in various settings. It provides valuable insights for teaching practices, supports academic research on cognitive function, and facilitates the assessment of clinical functions. Its utility extends to schools, research institutions, and clinical environments, making it a valuable resource for both research and practical applications. In summary, the TB-TMT demonstrates strong reliability and validity, making it a highly practical tool for future academic research, educational practices, and clinical evaluations.

Keywords: executive function, Trail Making Test, normative data, tablet computer

