

「設計流暢性測驗平板版」的信、效度研究 暨本土常模建置

陳珮臻¹ 吳清麟² 陳志洪³ 廖育秀⁴ 陳學志⁵

摘要

「設計流暢性測驗」(Design Fluency Test)能評量大腦額葉的執行功能表現。著名的「Delis-Kaplan 執行功能系統」(Delis-Kaplan Executive Function System, D-KEFS)，其「設計流暢性測驗」已被廣泛應用，且適用對象較廣，從資優個體到嚴重腦損傷患者。然而，該測驗的計分相對費時且費力。本研究參考 D-KEFS 的「設計流暢性測驗」，欲發展一套電子化版本，即「設計流暢性測驗平板版」(Tablet-Based Design Fluency Test, TB-DFT)。本測驗流程：情境一（流暢）、情境二（流暢＋抑制）、情境三（流暢＋轉換），以「正確設計數量」反映該情境的得分；透過情境間的比較，可以獲得流暢、抑制與轉換的三個計分指標。本研究於北、中、南、東取樣 1,099 位健康參與者（年齡介於 8~29 歲），進行信、效度之檢驗。結果顯示，本測驗之三個情境具有適當的再測信度。再者，本測驗具有適當的內部相關性，並且三個計分指標也能反映相應的能力。在效度考驗中，「流暢」指標與更新能力、反應時間、圖形推理之間具有顯著正相關，顯示適當的效標關聯效度；「轉換」指標與涉及轉換能力的卡片分類作業也呈顯著正相關，顯示適當的區辨效度。這顯示本測驗具有適當的效度，「流暢」與「轉換」是能有效且獨立評估的計分指標。另外，在樣本特性差異考驗顯示，性別沒有顯著差異，但測驗表現會隨著年齡逐漸上升，直至高中趨於穩定；本研究進一步建立認知發展至認知穩定階段（8~29 歲）之常模參照表。本研究發展的 TB-DFT 能有效測量健康個體之設計流暢性，提供未來研究與教育場域使用，期能深化設計流暢性與執行功能的探究。

關鍵詞：平板電腦、執行功能、常模、設計流暢性測驗、圖形流暢性

1. 陳珮臻，國立臺灣師範大學教育心理與輔導學系博士候選人

2. 吳清麟，國立臺灣師範大學學習科學學士學位學程副教授

3. 陳志洪，國立臺灣師範大學資訊教育研究所教授

4. 廖育秀，國立臺灣師範大學教育心理與輔導學系博士

5. 陳學志，國立臺灣師範大學教育心理與輔導學系教授

收件日期：2022.06.06；完成修改：2022.10.19；正式接受：2022.10.21

通訊作者：陳學志；Email：chcjh@ntnu.edu.tw

地址：106308 臺北市大安區和平東路一段 162 號

國立臺灣師範大學教育心理與輔導學系

The Reliability, Validity and Normative Data of a Tablet-Based Design Fluency Test

Pei-Zhen Chen¹ Ching-Lin Wu² Zhi-Hong Chen³
Yu-Hsiu Liao⁴ Hsueh-Chih Chen⁵

Abstract

Design Fluency Test (DFT) is an assessment for evaluating the performance of the frontal lobes. The Design Fluency Test of the famous Delis-Kaplan Executive Function System (D-KEFS) has been widely used, and is applicable to a wide range of subjects, from gifted individuals to severely brain-damaged patients. However, the scoring method for this test is relatively laborious. Based on D-KEFS-DFT, the present study develops an electronic version — the Tablet-Based Design Fluency Test (TB-DFT) — and inspect its psychometric characteristics. TB-DFT is divided into three conditions: condition 1 (fluency), condition 2 (fluency + inhibition), condition 3 (fluency + switching); the number of “correct design” reflects the score of condition. Through the comparison between conditions, we calculate three scoring indicators of fluency, inhibition and switching. 1,099 participants in aged 8-29 were recruited from different areas in Taiwan to exam the reliability and validity of this test. In the reliability, the three conditions of TB-DFT have great test-retest reliability. Furthermore, this test has great internal correlation, and the three scoring indicators also reflect the relative ability. In the criterion-related validity, there is a significant positive correlation between the “fluency” indicator and updating ability, reaction time, and figural reasoning. In the discriminant validity, the “switching” was also significantly positively correlated with card sorting tasks involving conversion ability. This indicates that the test has good validity, and “fluency” and “switching” are scoring indicators that can be effectively and independently assessed. In addition, there was no significant difference between genders, but performance would gradually increase with age until it became stable in high school; this study further established the norm of cognitive development to cognitive stability stage (Aged 8-29). This study provides effective measurement. It is expected to deepen the exploration of design fluency and executive function.

Keywords: Design fluency test, executive function, figural fluency, normative data, tablet computer

1. Pei-Zhen Chen, Doctoral Candidate, Department of Educational Psychology and Counseling, National Taiwan Normal University
2. Ching-Lin Wu, Associate Professor, Program of Learning Sciences, National Taiwan Normal University
3. Zhi-Hong Chen, Professor, Graduate Institute of Information and Computer Education, National Taiwan Normal University
4. Yu-Hsiu Liao, Ph.D, Department of Educational Psychology and Counseling, National Taiwan Normal University
5. Hsueh-Chih Chen, Professor, Department of Educational Psychology and Counseling, National Taiwan Normal University

Received: 2022.06.06; Revised: 2022.10.19; Accepted: 2022.10.21

Corresponding Author: Hsueh-Chih Chen; Email: chcjyh@ntnu.edu.tw

Address: No. 162, Sec. 1, Heping E. Rd., Da'an Dist., Taipei City 106308, Taiwan

Department of Educational Psychology and Counseling, National Taiwan Normal University

壹、緒論

執行功能 (executive function) 是指個體有效率地執行具目標導向的活動，並在過程中隨時依外在環境的需求與變化進行調整 (Burgess & Alderman, 2018; Ni et al., 2011)，坐落於前額葉 (prefrontal cortex) 的大腦區位 (Alvarez & Emory, 2006)。近年來，許多學者投注時間和心力致力於執行功能，尤其在教育方面 (Diamond & Ling, 2016; Zelazo et al., 2016)，執行功能被發現是預測學生未來學業成績的重要預測指標 (Alloway & Alloway, 2010; Dumontheil & Klingberg, 2012; McClelland & Cameron, 2011; St Clair-Thompson & Gathercole, 2006)。在執行功能相關研究的快速累積之時，發展有效評估執行功能的工具是研究的重要基礎。執行功能的定義較為廣泛，最典型的是 Miyake 等 (2000) 及 Diamond (2013) 所提出之執行功能三大因素：抑制 (inhibition)、轉換 (switching)、更新 (updating)；然而，近年研究發現流暢性 (fluency) 也是重要因素之一 (Karr et al., 2019)，應重視流暢性在執行功能領域的探究。

流暢性是指個體快速產生符合規則語文或非語文反應的能力，過去經常被用來了解個體的額葉功能 (Phillips, 1997)，包括語文流暢性與設計流暢性 (又稱為圖形流暢性)。其中，設計流暢性 (design fluency) 被認為最適合用來檢視個體額葉變化的情形 (Bryan & Luszcz, 2000)，測驗具備非語文特性、且不易受到性別變項的影響 (Ross et al., 2019; Van der Elst et al., 2011; Woods et al., 2016)。許多國家也重視設計流暢性的發展趨勢，並紛紛建立常模 (如 Battista et al., 2021; Cattelani et al., 2011; Delis et al., 2001; Izaks et al., 2011 等)，且國外著名「Delis-Kaplan 執行功能系統」(Delis-Kaplan Executive Function System, D-KEFS; Delis et al., 2001) 的「設計流暢性測驗」也被廣泛使用，足見設計流暢性的重要性。然而，回顧臺灣的流暢性相關測驗，並沒有發展標準化的相關測量工具，且對於設計流暢性的研究也相對較少 (如楊宗仁, 2002)，顯示臺灣研究領域缺少對設計流暢性的探討。再者，設計流暢性雖然不受語文能力影響，但實證研究已表明東方與西方之間具有執行功能表現的差異性，例如：香港學童的執行功能表現比英國學童佳 (Xu et al., 2020)，且無論在學齡前、兒童、青少年時期，東方學生的表現都比較好，支持本研究建立臺灣常模之想法。據此，本研究欲發展標準化之「設計流暢

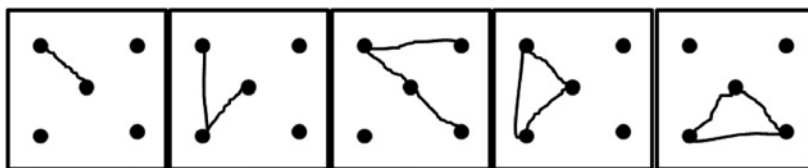
性測驗」並建立認知發展至穩定階段（8~29 歲）的臺灣常模，藉以深化臺灣研究領域對於執行功能與設計流暢性的探究，並提供創新的評估工具於臨床診斷及教學評估中。

貳、文獻探討

一、設計流暢性之測驗發展與比較

設計流暢性是在有限規則中繪製產出許多圖形的能力，為非語言流暢性表現。第一版的「設計流暢性測驗」是作為語文流暢性的非語文對比而開發（Jones-Gotman & Milner, 1977），作業要求受測者在 5 分鐘內繪製出愈多不同設計，對額葉受損有敏感性，尤其是右側額葉受損的個體（Jones-Gotman, 1991）。然而，最初的版本存在計分困難的問題，Regard 等（1982）為了解決此問題，發展了「五點測驗」（Five-Point Test），並以結構化的方式來評估設計流暢性的表現。在「五點測驗」中，受測者會看到許多方格，每個方格中有五個對稱放置的點（如圖 1 所示），受測者以連接點的方式來繪圖，每張圖最少繪製一條線，並且避免重複繪製相同的設計；計分方式則以繪製圖的數量反映設計流暢性的能力。隨著「五點測驗」的出現，學者相繼發展出不同形式的設計流暢性測驗，例如：「改版五點測驗」（Modified Five-Point Test; Cattelani et al., 2011）、「Ruff 圖形流暢性測驗」（Ruff Figural Fluency Test; Ruff et al., 1987）、「D-KEFS 設計流暢性測驗」（Delis-Kaplan Executive Function System-Design Fluency Test; Delis et al., 2001）。本研究為了發展標準化的設計流暢性測驗，進一步整理各種測驗的優勢與劣勢，說明如下。

圖 1 「五點測驗」的作答範例



註：取自 Regard 等（1982）。

首先，Regard 等（1982）的「五點測驗」有施測容易且快速的優勢，因為該測驗僅有一個情境且施測時間只有 5 分鐘，能在短時間內獲得測驗表現；然而，該測驗提供的訊息較少，僅有設計流暢性的表現。Cattelani 等（2011）修改了原本的「五點測驗」，將施測時間改為 3 分鐘，並增加設計策略的評分，這不僅減少時間花費，還獲得其他指標訊息。再者，Ruff 等（1987）參考「五點測驗」的程序，修改了正方形內的配置（例如：五個點的位置呈現不對稱、或增加讓人分心的干擾物），開發「Ruff 圖形流暢性測驗」（如圖 2 所示），測驗共有五個情境，每個情境皆呈現不同的刺激材料，且皆有 1 分鐘進行繪圖。雖然「Ruff 圖形流暢性測驗」經常被使用，但與「五點測驗」相比，施測程序相對費時且費力，卻沒有獲得其他指標訊息。第三，Delis 等（2001）結合 Regard 的形式（方格內有數個點）與繪製四條線的限制（維持受測者之間的设计複雜性與計分一致性），開發「D-KEFS 設計流暢性測驗」。該測驗包含三種情境（皆有 1 分鐘進行繪製），難度逐漸上升，能以多種情境的測量辨識認知神經機制（如設計流暢、抑制、轉換）的受損狀況。

圖 2 「Ruff 圖形流暢性測驗」



註：取自 Ruff 等（1987）。

在測驗編製基礎上，本研究考量較少花費時間與多種指標訊息的特性，「改版五點測驗」與「D-KEFS 設計流暢性測驗」皆具有的優勢，尤以後者更適合發展。「改版五點測驗」雖然可以獲得設計流暢性、設計策略的指標訊息，但「D-KEFS 設計流暢性測驗」卻能提供不同認知神經機制的訊息，更適合作為評估執行功能與設計流暢性的測量工具。「D-KEFS 設計流暢性測驗」具備施測時間短、且能獲得多項指標的特性，也是目前廣泛應用於健康樣本之測量工具（Delis et al., 2001），其多項指標也反映在計分的複雜性與時間花費上。是此，本研究將參考「D-KEFS 設計流暢性測驗」的編製原理，發展電子化版本。

二、「設計流暢性測驗平板版」之編製基礎與必要性

前已提及，「D-KEFS 設計流暢性測驗」相較其他設計流暢性測驗，係能花費較少時間且能獲得最多評估結果的測驗工具，但該測驗係以人工計分，於計分上過於繁瑣且耗時。本研究為改善計分冗長的劣勢，將發展數位化設計流暢性測驗。曾有研究發展「電腦版設計流暢性測驗」（Computerized Design Fluency; Woods et al., 2016），該測驗為五點的形式呈現，並參考「D-KEFS 設計流暢性測驗」的程序（僅繪製四條線），受測者可使用滑鼠進行繪圖。Woods 等（2016）提出「電腦版設計流暢性測驗」的幾個特性：第一，自動化施測，減少施測者對測驗結果的影響，具有標準化的施測程序；第二，評分是自動化且無錯誤的；第三，自動獲得整套測驗結果，包括反應的時間，此足見數位化測驗的優勢。「電腦版設計流暢性測驗」雖有效簡化計分方式，以數位方式輸出結果，但在媒介上可能存在限制。電腦版是以滑鼠進行測驗，在表現上恐受到對設備使用的熟悉程度而影響。若採用平板電腦進行操作，更接近在紙上進行繪製的行為型態。目前，還沒有以平板電腦操作的設計流暢性測驗被開發，為了檢驗平板電腦作為媒介的適用性，本研究將開發「設計流暢性測驗平板版」（Tablet-Based Design Fluency Test，以下簡稱 TB-DFT），拓展行為測驗之媒介的創新性。

TB-DFT 編製之基礎，將參考「D-KEFS 設計流暢性測驗」的程序進行編製。「D-KEFS 設計流暢性測驗」共有三個情境：情境一，黑點連線，與「五點測驗」相似，方格中有五個黑點刺激，受測者需繪製四條線且必須是連續線段（線與線之間相交於一個黑點），該情境涉及設計流暢性的表現。情境二，加入白色點刺激，方格中呈現五個黑點與五個白點，受測者必須抑制先前對黑點（先前目標）的行為反應，以白點為目標進行連線、繪製圖形，該情境涉及設計流暢性與抑制能力的參與。情境三則是轉換情境，方格中有五個黑點與五個白點，受測者需在黑點與白點之間互相轉換進行連線，該情境為設計流暢性與轉換能力的參與；其中，轉換能力與卡片分類（Heaton & Staff, 1993）涉及相似的轉換機制（Delis et al., 2001）。最終，透過情境間的比較將能解釋受測者的流暢、抑制與轉換的表現。該測驗難度係逐漸提升，不僅能避免天花板或地板效應，同時也能看出個體認知功能的興衰。

三、設計流暢性測驗之效度證據的研究

設計流暢性測驗與其他測驗的關聯性反映出效標關聯效度或區辨效度 (Ross et al., 2019; Tucha et al., 2012; Van der Elst et al., 2011; Woods et al., 2016)。如前所述，設計流暢性是快速產出符合特定規則圖形的能力，若個體的反應時間、處理速度愈佳，設計流暢性的表現也愈佳 (Ross et al., 2019; Tucha et al., 2012; Woods et al., 2016)。而且在進行繪圖時，為避免重複先前的設計，需要記得繪製過的圖形，與更新能力（工作記憶）也有關聯性 (Ross et al., 2019; Tucha et al., 2012; Woods et al., 2016)，表示個體需要有足夠的工作記憶容量記得先前繪製過的圖畫，進而持續將注意力放在繪圖上，產出更多反應。再者，本研究發展 TB-DFT 之情境二與情境三，分別涉及抑制能力與轉換能力，兩者與設計流暢性之間呈弱相關或無相關 (Ross et al., 2019; Tucha et al., 2012; Woods et al., 2016)，顯示三者之間的獨立性。此外，設計流暢性是執行功能的一環，研究指出智力與執行功能幾乎無相關 (Ardila et al., 2000)，且執行功能比智力更能預測國小到大學的成績 (Alloway & Alloway, 2010; Dumontheil & Klingberg, 2012; McClelland & Cameron, 2011; St Clair-Thompson & Gathercole, 2006)，智力測驗所得分數只能預測實際表現的 20 至 25% 左右 (Sternberg et al., 2001)，這表示兩者可能涉及不同歷程。由上可知，反應時間愈快、更新能力愈佳，個體的設計流暢性表現也愈好，本研究將以「反應時間」及「更新能力」作為效標關聯效度；而抑制能力、轉換能力及智力對個體設計流暢性的表現，顯示弱相關或無相關，本研究將以「抑制能力」、「轉換能力」、「智力」作為區辨效度，藉以檢驗 TB-DFT 的效度。

另一方面，設計流暢性測驗的表現不存在性別差異，且受到年齡與教育年的影響很大 (Chelune et al., 2006; Ross et al., 2019; Tucha et al., 2012; Van der Elst et al., 2011; Woods et al., 2016)，例如：Van der Elst 等 (2011) 曾考驗 6.5 至 15.5 歲兒童的設計流暢性表現，結果發現年齡與產出設計數量有中高等程度的正相關，且隨著年紀增加而產生更多的設計；又如：Tucha 等 (2012) 建立兒童 (7~15 歲) 與成人 (20~88 歲) 常模，發現兒童的表現隨年紀上升、成人表現卻隨年紀而下降，顯示認知成長與衰退之變化。然而，一項建立成人 (35~82) 常模的研究將參與者依高 (> 12)、低 (≤ 12) 教育年作區分，發現高教育年的設計數量顯著大於低教育年，顯示設計流暢性也受到

教育年的影響 (Izaks et al., 2011)。因此，便有研究將年齡與教育年視為設計流暢性的重要預測因子 (Battista et al., 2021; Cattelani et al., 2011)。由上可知，設計流暢性在樣本特性（如性別、教育年）的穩定結果，本研究將以樣本特性檢驗 TB-DFT 是否也具有相同的結果。

四、研究目的

綜上所述，Delis 等 (2001) 開發「D-KEFS 設計流暢性測驗」，具備施測時間短、施測容易、計分多樣性之優勢，唯獨計分方式只能採人工計分，會花費大量的時間與精力。為此，本研究將發展電子化版本，參考「D-KEFS 設計流暢性測驗」，改良操作（平板電腦）與計分（數位計分）的程序。本研究之目的即是開發 TB-DFT，並檢驗其信、效度，最後建立認知發展至穩定階段（8~29 歲）之臺灣常模，期能提供創新的研究工具，深化對設計流暢性與執行功能之探究。

參、研究方法

一、參與者與流程

依分層立意取樣的方式，取樣臺灣北、中、南、東四個地區的 8~29 歲 ($M = 14.24$ 、 $SD = 3.93$) 之健康參與者（男性 49.32%、女性 50.68%），共計 1,099 位（參與者分布如表 1 所示）。其中，常模樣本的教育年分組參考 D-KEFS (Delis et al., 2001)，將本研究參與者教育年組分成：國小中年級、國小高年級、國中、高中、大學（含）以上。本研究為建立健康樣本之常模，在開始正式施測之前，每位參與者皆填寫健康調查表，內容包含：聽力障礙、視覺障礙（含色盲／色弱）、上肢運動障礙／上半身運動障礙；神經疾病或精神疾病（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）；是否服用抗憂鬱、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；有認知或記憶障礙病史，或曾被診斷為學習障礙。確保無以上狀況後，則開始正式施測。本研究採團體施測的方式，由本研究安排 2~3 位施測人員進入班級教室進行施測，施測人員須注意參與者對測驗的理解情形。根據班級人數分配施測人員，以一位施測者對應十五位參與者為上限，

例如：一個班級有 28 位學生，將安排 2 位施測人員協助施測。

表 1 全體樣本之人口變項（地區、性別、教育年組）分布情形

	地區	北		中		南		東		總人數
	性別	男	女	男	女	男	女	男	女	
國小中年級		31	31	21	23	28	18	14	11	177
國小高年級		26	26	44	44	18	21	9	11	199
國中		79	80	25	27	21	23	16	24	295
高中		28	33	40	27	34	39	17	17	235
大學（含）以上		44	32	20	19	18	37	9	14	193
總計		208	202	150	140	119	138	65	77	1,099

另外，部分參與者進行了信度及效度考驗。本研究於全體樣本中抽取次樣本 50 人，進行 TB-DFT 再測信度的考驗；於全體樣本中另外再抽取次樣本，進行 TB-DFT 效標關聯效度與區辨效度的考驗。施測時間約需一節課的時間，先進行 TB-DFT，再進行 1~2 個效度考驗的研究工具。本研究考量到施測一節課可能會有練習疲勞的問題，在效標工具的施測順序上，依照注意力參與程度進行排序（由高到低：RST、SmT、RTT、SrT、CEFI），以減少練習疲勞的問題；性向測驗（推理能力）施測時間約 40 分鐘，則不會同時施測其他效標工具。針對信、效度的考驗，參與者的取樣需與課堂老師商量，考量教學現場的限制，本研究以三個教育階段（小學、中學、大學）進行信、效度的檢驗為原則（詳細人數請見表 2）。由於實體收案與教學現場的限制，逢 Covid-19 疫情期間，學校執行防疫措施，進行遠距教學，唯智力／推理能力（即性向測驗）僅有中學階段的樣本，後面將討論此限制。

表 2 信、效度考驗之樣本數分布

	教育階段			總計
	小學	中學（國高中）	大學（含）以上	
信度考驗的樣本數				
再測信度	11	26	13	50
效度考驗的樣本數				
RST	65	104	29	198
RTT	36	78	29	143
CEFI	66	83	67	216
SmT	77	105	3	185
SrT	33	80	29	142
性向測驗（推理能力）	—	71	—	71

註：RST = Running Span Task（記數廣度作業）；RTT = Reaction Time Task（反應時間作業）；CEFI = Chinese Executive Function Index（「中文版執行功能指標」）；SmT = Simon Task（「西蒙測驗」）；SrT = Sorting Task（「分類測驗」）。

二、研究工具

（一）本研究開發 TB-DFT

1. 平板電腦的內容設計

本研究參考 Delis 等（2001）的「設計流暢性測驗」，開發 TB-DFT。首先，本研究選取普遍使用的 10 吋平板電腦進行開發。接著，本測驗是呈現一系列內有數個點的方格，並要求受測者在每個方格中以四條直線連接這些點且每條線必須相交於一點。為符合此一程序，本研究將平板電腦視為畫紙，將方格以 1 比 1 的比例呈現在平板電腦上，每個情境皆有 4×9 個方格，受測者需在 60 秒內盡可能繪製出不同圖形。本測驗共有三個情境；情境一為「黑點連線」，方格內只有五個黑點，受測者需藉由連接這些點來繪製圖形；情境二為「白點連結」，方格內有五個黑點及五個白點，受測者只能連結白點，抑制之前以黑點連線的行為反應；情境三為「黑點與白點轉換」，方格中有五個黑點及五個白點，受測者需要交替連接黑點與白點來繪製圖形。另外，本測驗設置「橡皮擦」功能，當受測者需要修改時，可以啟用橡皮擦功能，將需要修改的方框清除，再重新進行繪圖。

2. TB-DFT 的施測流程

開始施測前，將平板電腦設備設定為：飛航模式、靜音且無連結網路的狀態，確保測驗進行時無任何干擾訊息。當受測者拿到平板電腦後，請受測者將其橫向平放於桌面上，並點開測驗 App 開始進行施測。進入測驗畫面後，受測者需填寫基本資料（如附錄圖 A）：姓名、編碼、性別（男／女）、慣用手（右手或左手）、教育年（從國小一年級開始算起之受教育時間）、生日。完成基本資料填寫後即開始測驗內容，每個情境皆有練習階段與正式階段，兩階段開始前都有指導語（如附錄圖 B），於畫面左側為指導語、右側為操作範例影片。其中，練習階段（如附錄圖 C）為確保參與者了解測驗進行方式，若有錯誤則會顯示錯誤原因（如附錄圖 D），受測者必須修改錯誤的圖，直到三題練習題完全正確（如附錄圖 E）。正式階段有 36 個方格（4×9），受測者需在 60 秒內繪製圖形，並依循由左至右、由上到下的順序；同時，能以右側的上卷鍵及下卷鍵滑動介面，檢視曾繪製過的圖形；計時結束後，將轉跳至下一個情境。最後，所有情境完成，將轉跳至測驗結束的畫面（如附錄圖 F），確保資料儲存。

3. TB-DFT 的計分與說明

(1) 原始資料

每個情境將可算得正確設計數（符合規則的設計數量），並以正確設計數反應各個情境之表現；情境一提供了設計流暢性的基本測量，情境二量測出設計流暢性與抑制能力，而情境三則評估了設計流暢性與認知轉換。本測驗將算得整體正確率，也就是在三個情境加總的總嘗試次數中正確設計數量之比例。另外，受測者在每個情境中繪製的圖畫皆會被記錄，可以檢視受測者所繪製的圖畫。

(2) 比較測量

首先，比較情境一與情境二的正確設計數之結果：當情境一的正確設計數小於情境二，顯示出練習效果；當情境一的正確設計數大於情境二，顯示出與持續性傾向有關，個體容易被影響而無法跳脫先前情境的規則；當正確設計數在情境一跟情境二皆低時，表示基本視覺注意力或動作速度受損，或是基本能力表現在正常範圍內，可能在涉及的高層次執行功能上有缺陷；當正確設計數在情境一跟情境二皆高時，表示視覺注意力、動作速度、非語言

創造力與同時處理的能力上具有優勢。再者，情境三的正确設計數大於情境一＋情境二，顯示出愈佳的轉換能力；反之，則愈差。

(3) 計分指標

本測驗共可算得三個指標：第一是流暢（fluency），指個體在特定規則下快速產出圖形設計的能力，指標計算由情境一的正确設計總數反映之；第二是抑制（inhibition），指個體能有效抑制先前的目標項，指標計算由情境二的正确設計總數減掉情境一的正确設計總數，再除以基準值（即情境一的正确設計總數），顯示個體抑制先前目標項（即黑點）後產出設計反應的比例，其數值愈大、抑制能力愈好；第三是轉換（switching），指個體能在兩種目標項中來回轉換的能力，指標計算由情境三的正确設計總數，減掉情境一與情境二的平均正确設計總數，再除以基準值（即情境一與情境二的平均正确設計總數），最後加上常數 1 換算成大於 0 的數值，顯示個體在轉換情境下所能產出設計反應的比例，其數值愈大、轉換能力愈好。

（二）效度考驗之工具與流程

1. 效標關聯效度

根據過去研究顯示，設計流暢性表現愈佳，個體的更新能力、反應時間、處理速度也愈好（如 Ross et al., 2019; Tucha et al., 2012; Woods et al., 2016），且「設計流暢性測驗」也被視為執行功能測驗（Swanson, 2005）。本研究以「記數廣度作業」（Running Span Task, RST）、「反應時間作業」（Reaction Time Task, RTT），以及「中文版執行功能指標」（Chinese Executive Function Index, CEFI；游勝翔、林緯倫，2018）作為效標關聯效度之研究工具。以下說明效標關聯效度研究工具之適宜性與介紹。

(1) 「記數廣度作業」（RST）

RST 是評估工作記憶容量的工具，但該測驗所需要的認知需求與工作記憶容量有所不同，需要動態且即時的監控連續出現之刺激，並根據刺激做出反應。因此在 RST 中，個體必須不斷更新對目標項的心理表徵，同時將不相關的項目排除，是最適合評估「更新能力」的測量工具。本研究參考 Moore 與 Ross（1963）的 RST 之程序，編製本研究效度考驗的工具。在 Moore 與

Ross 的測驗程序中，將呈現數字序列，一次呈現一個數字，受測者必須記住最後幾個數字，例如：畫面提示「請記住最後 3 個數字」，若數字序列為「2-5-4-1-...-7-8-4-3-」，正確反應是「8、4、3」。受測者記住的數字量被視為「廣度」(span)，廣度範圍為 1~8，一個廣度兩題。練習題(廣度「2」)階段，若答對兩題可進入下一個廣度(如廣度「3」)，答錯兩題則退回前一個廣度(如廣度「1」)，答對一題則維持在原廣度(如廣度「2」)，直到某一廣度進入第四次則完成測驗；測驗沒有時間限制，平均約 10 分鐘完成(含指導語說明)。計分方式如公式：
$$\sum \frac{\text{正確嘗試次數}}{\text{總計嘗試次數}} \times \text{廣度}$$
，分數愈高表示更新能力愈好。

(2)「反應時間作業」(RTT)

反應時間被視為處理速度的能力，是與年齡相關的高階認知功能下降之重要評估基礎(Madden, 2001)，而反應時間與設計流暢性之間也存在顯著相關(Woods et al., 2016)，本研究以「反應時間」作為效標進行考驗。本研究參考Deary-Liewald的RTT程序(Deary et al., 2011)，編製本研究效度考驗的工具。在Deary-Liewald的RTT中共有兩個情境：簡單反應時間(Sample Reaction Time, SRT)與選擇反應時間(Choice Reaction Time, CRT)。在SRT中，受測者只需針對一個目標項作出按鍵反應，共有 20 個嘗試次；在CRT中，有四個目標項，受測者必須針對正確的目標項作出按鍵反應，共有 32 個嘗試次。在畫面中，下方呈現一個準備鍵、上方則呈現一個／四個目標項，該測驗要求受測者先按著準備鍵，當目標項變色後，受測者必須快速地按下目標項來反應。每個嘗試次間距(inter-trial interval, ITI)範圍在 1~3 秒之間。施測時間約 10 分鐘內完成(含指導語說明)。計分方式將排除 2.5 個標準差後，計算在SRT與CRT中，正確嘗試次的平均反應時間，若時間愈快，表示「反應時間」表現愈佳。在信度上，內部一致性SRT達.94、CRT達.97，再測信度SRT達.64、CRT達.83；在效度上，SRT與CRT皆與數字反應時間方框作業(numbers reaction time box；評估反應時間)有正相關，具有效標關聯效度。

(3)「中文版執行功能指標」(CEFI)

CEFI(游勝翔、林緯倫, 2018)是評估個體在日常生活中運用執行功能

的表現。由於設計流暢性被視為執行功能的因素之一（Karr et al., 2019; Phillips, 1997），本研究將以日常生活的執行功能表現作為效標進行效度考驗。此測驗為評估一般國中生與大學生之執行功能表現，測驗分為「策略行動」、「組織規劃」、「衝動控制」三個分量表（共 15 題）。該測驗之內部一致性信度，全量表達 .77、各分量表介於 .69~.76，而全量表之再測信度達 .80，且策略行動（ $r = .55$ ）、組織規劃（ $r = .76$ ）、衝動控制（ $r = .75$ ）也達顯著相關；與衝動性、憂鬱症狀、情緒控制之間有顯著相關。該測驗具有良好的信、效度。以李克特式五點量表計分，依其版本（國中生版或大學生版）之因素結構計算三個因素的得分，得分愈高表示該因素表現愈好。

2. 區辨效度

根據過去研究發現執行功能與智力之間幾乎沒有相關（Ardila et al., 2000），本研究以「國民中學學業性向測驗七至九年級」（周台傑、葉瓊華，2016）評估智力表現。再者，TB-DFT 的情境二、情境三，分別涉及抑制能力及轉換能力，將分別以「西蒙測驗」（Simon Task, SmT）及「分類測驗」（Sorting Task, SrT）作為研究工具，區辨三個情境的評量指標。以下說明區辨效度研究工具之適宜性與介紹。

(1) 「西蒙測驗」（SmT）

SmT 是要求受測者根據刺激材料的特徵條件，按下特定方向的按鍵，例如：Simon（1990）的研究中有兩種刺激，刺激會隨機出現在左右邊，受測者要根據規則按下正確反應鍵。在這種實驗設計中，若刺激位置與按鍵位置一致，反應速度會比較快，相對的若刺激位置與按鍵位置不一致，則反應速度會比較慢且更容易出錯。這種效果被稱為西蒙效應（Simon effect），反映受測者能否成功抑制空間訊息的干擾。西蒙效應是經常用來反映「抑制能力」的指標，本研究參考 Simon 的測驗程序，編製本研究效度考驗的工具。此測驗由兩個情境組成：情境一為熟悉規則，色塊將出現在螢幕正中間，參與者被要求看到紅色色塊按左鍵、看到藍色色塊按右鍵，共有 32 次嘗試；情境二為抑制情境，色塊將隨機出現在螢幕的左側或右側，同樣地參與者被要求看到紅色按左鍵、看到藍色按右鍵，共有 40 次嘗試。每次嘗試之流程，先出現一個凝視點（500ms），接著出現色塊（5000ms），看到色塊後並做出反應，若在 5000 毫秒內按下反應鍵後進入下一次嘗試，若超過 5000 毫秒未

反應，將以不正確反應（太慢反應）紀錄，並進入下一次嘗試。計分方式以每次嘗試及與計分指標（流暢、抑制、轉換）的反應時間做計算，首先進行資料處理，排除錯誤反應（答錯）、太快或太慢的反應（反應時間小於 300ms、大於 5000ms）；再者，計算抑制能力，以情境二的不一致反應時間減去一致的反應時間（即 $RT_{\text{不一致}} - RT_{\text{一致}}$ ），算得西蒙效應為抑制能力的指標。

(2) 「分類測驗」(SrT)

本研究開發 TB-DFT 之情境三涉及認知轉換能力，而執行功能的經典測驗「威斯康辛卡片分類」(Heaton & Staff, 1993) 也要求受測者運用認知轉換能力。因版權問題，本研究另參考 D-KEFS 的「分類測驗」(SrT; Delis et al., 2001) 編譯中文版本。該測驗涉及認知轉換能力的參與。在測驗程序中，提供六張卡片，要求受測者將卡片分成兩組（一組三張），分在同一組的卡片具有概念相似。共有兩個卡片組，一個卡片組最多可以被分成八個目標分類（三種語文分類、五種視覺模式分類），分出愈多個目標分類則得分愈高，而重複分類則不重複計分。得分愈高顯示概念形成能力，且愈能抑制先前分類過的反應，具備認知轉換的思考與行為。施測時間為 5 分鐘。

(3) 「國民中學學業性向測驗七至九年級」

過去研究發現，執行功能與智力之間呈弱相關，甚至是無相關的結果 (Ardila et al., 2000)。本研究欲針對不同特定領域的智力與推理能力來檢驗效度，周台傑與葉瓊華 (2016) 發展的「國民中學學業性向測驗七至九年級」能評估個體的智力與推理能力，包含語文推理、數學推理、圖形推理，將作為本研究區辨效度考驗的研究工具。此測驗能評量受測者的推理能力（即智力），包含語文推理（類比、歸納、記憶）、數學推理（系列、算術推理）、圖形推理（類比、系列、歸納）三個分測驗（共 60 題），施測時間為 40 分鐘。再測信度達 .48~.79、內部一致性信度達 .66~.89、內部相關達 .48~.85，且與國文成績 ($r = .44 \sim .82$) 及數學成績 ($r = .59 \sim .81$) 皆達顯著相關。得分愈高表示該分量表之表現愈好。

三、資料整理與分析

首先，本研究對健康樣本的表現進行描述統計，計算 TB-DFT 在三個情境中的正確設計總數，並進一步作比較測量，以單因子相依變異數分析考驗

情境間的差異；同時，計算此測驗的三個計分指標（流暢、抑制、轉換）之平均數與標準差，藉以檢驗全體健康樣本在本測驗之整體表現。接著，進行信、效度的考驗。在信度分析中，將考驗再測信度，分析第一次施測與第二次施測在三個情境的正確設計總數及正確率之相關係數。在效度考驗中，本研究將分成三個部分呈現：其一，考驗內部相關性，分析三個情境間的正確設計總數之相關係數；其二，考驗效標關聯效度，分析三個計分指標（流暢、抑制、轉換）與效標關聯效度工具之間的相關性；其三，考驗區辨效度，分析三個計分指標（流暢、抑制、轉換）與區辨效度工具之間的相關性。最後，進行樣本特性考驗，以教育年〔國小中年級、國小高年級、國中、高中、大學（含）以上〕與性別進行差異考驗，依差異檢定結果建立 TB-DFT 的常模對照表。

肆、研究結果

一、描述統計

首先，本研究算得 TB-DFT 三個情境之正確設計數的描述統計：情境一為 6.49（標準差 2.76）、情境二為 8.02（標準差 3.01）、情境三為 5.41（標準差 2.97）。進一步「比較測量」發現，情境一的正確設計數顯著小於情境一（ $F = 378.25, p < .001$ ），顯示出練習效果；情境三的確設計數顯著小於情境一與情境二之平均正確設計數（ $F = 487.47, p < .001$ ），顯示情境三的轉換歷程。再者，TB-DFT 中的流暢、抑制、轉換計分指標，三者的平均數各為 6.48、0.39、0.78，標準差各為 2.76、0.88、0.51。「流暢」指標顯示所有參與者平均產出 6.48 個設計反應。「抑制」指標顯示所有參與者在抑制先前目標項（即黑點）後平均能多設計出 0.39 的反應比例；該數值大於 0，顯示健康樣本不易受到先前目標項影響，而能產出更多設計反應。「轉換」指標顯示所有參與者在轉換情境下所產出的設計反應，平均產出 0.78 的反應比例；此說明轉換情境下需要花費更多時間來產出設計反應。上述結果顯示，健康樣本在 TB-DFT 之整體表現。

二、信、效度考驗

（一）信度考驗：再測信度

以整體健康樣本中的 50 位參與者進行再測信度之考驗，男性 25 人、女性 25 人。第一次施測與第二次施測之時間間隔約六個月。結果發現（如表 3 所示），各情境之正確設計總數及正確率在兩次施測的表現中，具有顯著正相關（ $r = .39 \sim .68$ ），表示 TB-DFT 具有良好的再測信度。

表 3 TB-DFT 之再測信度（ $n = 50$ ）

		第二次施測			正確率
		情境一	情境二	情境三	
第一次施測	情境一	.39**	-	-	
	情境二	-	.45***	-	
	情境三	-	-	.48***	
	正確率				.68***

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

（二）效度考驗

由上可知，TB-DFT 三個情境評估的能力是不同且複雜的，本研究進一步計算出三個計分指標（流暢、抑制、轉換），藉以區分出各情境所評估的能力。以下檢驗 TB-DFT 三個計分指標（流暢、抑制、轉換）是否有效反映此測驗所要評估的能力。

1. 內部相關性

首先，以整體健康樣本考驗測驗內部相關性，將三個情境之正確設計數及三個計分指標（流暢、抑制、轉換）進行相關分析。結果顯示（如表 4 所示）三個情境之間有顯著正相關；且，抑制與情境二有顯著正相關（ $r = .18$, $p < .001$ ）、轉換與情境三有顯著正相關（ $r = .56$, $p < .001$ ），顯示計分指標與相應情境表現之間的關聯性，並與無關情境沒有相關（ $r = -.06$ 、 $.04$, $ps > .05$ ）。另外，三個計分指標之間有顯著相關，可能受到測驗規則（需繪製四條直線）及計算指標（皆以情境的正確設計數之計算來反映）的影響。

上述結果指出，TB-DFT 具有良好的內部相關性，且三個計分指標（流暢、抑制、轉換）能反映相對應的能力。

表 4 TB-DFT 之內部相關性

	1	2	3	4	5
1 流暢（情境一）	-				
2 情境二	.60***	-			
3 情境三	.43***	.47***	-		
4 抑制	-.48***	.18***	-.06	-	
5 轉換	-.20***	-.17***	.56***	.04	-

*** $p < .001$

2.效標關聯效度與區辨效度

透過內部相關性的考驗，可知三個計分指標（流暢、抑制、轉換）能反映相對應的能力表現。本研究進一步檢驗三個計分指標的效度，提供計分指標之效度證據。

在效標關聯效度中，以整體健康樣本中的部分參與者進行效標關聯效度之考驗（如表 5 所示）。結果發現，流暢與更新能力、簡單及選擇反應時間、執行功能皆有顯著相關（ $r = -.19 \sim .34$ ， $ps < .05$ ），符合過去研究發現；抑制與轉換幾乎與效標關聯效度指標無相關，顯示三者之間的獨立性；而轉換與簡單反應時間（ $r = -.22$ ， $p < .01$ ）及選擇反應時間（ $r = -.21$ ， $p < .05$ ）有顯著負相關，顯示在轉換情境中反應速度愈快、轉換表現愈佳。上述結果顯示，TB-DFT 具有效標關聯效度。

表 5 TB-DFT 之效標關聯效度考驗結果

	TB-DFT 之計分指標		
	流暢	抑制	轉換
RST ($n = 198$)			
更新能力	.32**	-.004	.09
RTT ($n = 143$)			
簡單反應時間	-.19*	-.003	-.22**
選擇反應時間	-.34**	.08	-.21*
CEFI ($n = 216$)			
策略行動	.28***	-.06	-.04
組織規劃	.26***	-.04	-.04
衝動控制	.16*	-.04	-.11

註：RST = Running Span Task（記數廣度作業）；RTT = Reaction Time Task（反應時間作業）；CEFI = Chinese Executive Function Index（中文版執行功能指標）。

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

在區辨效度中，為進一步檢驗三個計分指標（流暢、抑制、轉換）的獨立評分效果，將以 SmT 的西蒙效應作為抑制能力、另以 SrT 的分類得分反應轉換能力。結果指出（如表 6 所示），西蒙效應及分類得分與流暢、抑制沒有相關（ $r = -.12 \sim .12$ ， $ps > .05$ ），但與轉換有顯著相關（ $r = -.15$ 、 $.24$ ， $ps < .05$ ）。另外，本研究以「國民中學學業性向測驗七至九年級」中不同領域推理能力作為區辨效度，藉以區分三個計分指標，結果指出圖形推理與流暢有顯著正相關（ $r = .38$ ， $p < .001$ ），語文及數學推理與轉換有顯著正相關（ $r = .33$ 、 $.33$ ， $ps < .001$ ），而抑制則皆無相關（ $r = -.10 \sim -.07$ ， $ps > .05$ ）。上述結果顯示，TB-DFT 的轉換指標係能有效獨立評估之指標，而抑制指標的效果可能較弱。

表 6 TB-DFT 之區辨效度考驗結果

	TB-DFT 之計分指標		
	流暢	抑制	轉換
SmT ($n = 185$)			
Simon effect	-.03	-.12	-.15*
SrT ($n = 142$)			
分類得分	.12	-.09	.24**
性向測驗 ($n = 71$)			
語文推理	.06	-.10	.33***
數學推理	.15	-.07	.33***
圖形推理	.38***	-.09	-.01

註：SrT = Sorting Task (「分類測驗」)；SmT = Simon Task (「西蒙測驗」)；性向測驗 = 「國民中學學業性向測驗七至九年級」。

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

三、樣本特性

根據過去研究顯示，設計流暢性不存在性別差異，且會隨著教育程度而有所提升。本研究檢驗不同樣本特性之差異表現。在性別考驗（如表 7 所示）中，女性的正確率顯著高於男性（ $t = -2.52$ 、 $p < .05$ ），而在總正確設計數之平均表現中男性（ $M = 19.63$ 、 $SD = 7.50$ ）與女性（ $M = 20.18$ 、 $SD = 6.74$ ）無顯著差異（ $t = -1.28$ 、 $p > .05$ ）；另外，在各情境與指標的差異檢定上，三個情境之正確設計數及三個指標（流暢、抑制、轉換）皆無顯著差異（ $t = 1.04 \sim -1.43$ 、 $ps > .05$ ），符合過去研究發現（如 Ross et al., 2019; Van der Elst et al., 2011; Woods et al., 2016）。在教育年組的考驗中（如表 8 所示），分成國小中年級（三四年級）、國小高年級（五六年級）、國中、高中及大學（含）以上（共五組）。結果發現，正確率及正確設計總數皆具有顯著差異，事後比較顯示教育年愈高，其正確率愈高、正確設計數愈多，亦符合過去研究發現（如 Battista et al., 2021; Cattelani et al., 2011; Van der Elst et al., 2011）。在各情境與指標上，亦發現相同結果，除了抑制指標。本研究進一步根據教育年的分組，建立不同教育年的常模對照表，可以得知不同教育年段在三個情境之原始分數所對應的百分等級。

表 7 性別在 TB-DFT 的整體及各情境與指標之描述統計與差異檢定

	男性 (n = 542)	女性 (n = 557)	t
整體			
正確率	73%	76%	-2.52*
正確設計數總數	19.63 (7.50)	20.18 (6.74)	-1.28
各情境與指標			
流暢 (情境一)	6.42 (2.92)	6.55 (2.59)	-0.77
情境二	7.94 (3.13)	8.10 (2.90)	-0.91
情境三	5.27 (3.15)	5.53 (2.78)	-1.43
抑制	0.42 (1.00)	0.37 (0.76)	1.04
轉換	0.77 (0.53)	0.79 (0.49)	-0.68

* $p < .05$

表 8 教育年分組在 TB-DFT 的整體及各情境與指標之描述統計與差異檢定

	1 國小 中年級 (n = 177)	2 國小 高年級 (n = 199)	3 國中 (n = 295)	4 高中 (n = 235)	5 大學 (含) 以上 (n = 193)	F	事後比較
整體							
正確率	63 %	72 %	74 %	80 %	82 %	43.11***	1 < 2, 3 < 4 < 5
正確設計數總數	12.45 (4.31)	15.74 (4.63)	20.49 (6.24)	24.49 (5.88)	24.61 (5.67)	185.82***	1 < 2 < 3 < 4, 5
各情境與指標							
流暢 (情境一)	4.26 (2.13)	5.21 (2.12)	6.74 (2.53)	7.95 (2.64)	7.68 (2.44)	86.35***	1 < 2 < 3 < 4, 5
情境二	5.36 (2.11)	6.35 (2.32)	8.32 (2.84)	9.62 (2.55)	9.78 (2.49)	119.37***	1 < 2 < 3 < 4, 5
情境三	2.82 (1.86)	4.19 (2.25)	5.44 (2.76)	6.91 (2.71)	7.14 (2.79)	100.21***	1 < 2 < 3 < 4, 5
抑制	0.48 (1.11)	0.36 (0.79)	0.46 (1.17)	0.32 (0.54)	0.36 (0.50)	1.354	
轉換	0.62 (0.44)	0.75 (0.44)	0.76 (0.39)	0.83 (0.41)	0.85 (0.35)	9.283***	1 < 2 < 3 < 4 < 5

*** $p < .001$

伍、綜合討論

「設計流暢性測驗」係執行功能的評量工具之一，個體必須在有限的規則繪製圖形，同時避免重複繪製，能反應個體的設計流暢性與執行功能之表

現。過去研究已發展許多版本的設計流暢性測量工具，並進一步建立常模。然而，臺灣目前尚未建立標準化的「設計流暢性測驗」及其常模，實徵研究也指出東西方文化下的個體，其執行功能表現有差異性（Xu et al., 2020），即使非語文測驗工具不受到東西方文化的影響，但在常模對照上並不合適（如 D-KEFS 為美國常模）。因此，本研究之目的係發展電子化測驗並建立臺灣常模。本研究參考「D-KEFS 設計流暢性測驗」的情境，發展電子化、平板電腦操作的測驗，即 TB-DFT。本測驗共有三個情境，以正確設計數反映該情境的表現，並可透過情境間的計算，算得三個計分指標（流暢、抑制、轉換）。本研究取樣 1,099 位健康樣本，檢驗 TB-DFT 的信、效度，並建立 8~29 歲的臺灣常模。在信度考驗中，執行兩次施測（間隔不超過六個月），結果顯示兩次施測的表現有顯著正相關，具有良好的再測信度。在效度考驗中，先以內部相關性檢驗各情境表現與三個計分指標（流暢、抑制、轉換）之相關性，確定三個計分指標是能評估相對應能力的指標分數；接著，檢驗三個計分指標之效標關聯效度與區辨效度，由結果支持流暢與轉換是能獨立評分之能力。據此，TB-DFT 是適合評估個體設計流暢性與執行功能的評估工具。

進一步檢視效度考驗的資料，結果發現 TB-DFT 的計分指標（流暢、轉換）與不同效標關聯效度之研究工具間存在低度至中度的顯著相關。首先，在流暢指標的效度考驗上，與過去實徵研究得到相同的結果（如 Ross et al., 2019; Woods et al., 2016），表示「流暢」能力需要較好的更新能力，且個體訊息處理速度愈快、愈能快速產出更多圖形。再者，轉換指標的效度考驗，與同樣涉及認知轉換能力的分類測驗（Delis et al., 2001）有低度至中度的正相關，且與涉及抑制能力的西蒙效應也有低度負相關，支持轉換指標的獨立評分效果。第三，過去研究將「設計流暢性測驗」視為執行功能的評估工具，TB-DFT 與自陳式執行功能測驗（即 CEFI）有低度至中度的正相關，支持「設計流暢性測驗」能評估個體執行功能。第四，過去研究指出執行功能比智力更能預測未來的表現（Alloway & Alloway, 2010; Dumontheil & Klingberg, 2012; McClelland & Cameron, 2011），兩者之間相關性不高（Ardila et al., 2000）；本研究發現「流暢」與圖形推理有低度至中度的正相關，表示設計流暢性表現與更好的圖形推理有關，且「轉換」與特定領域推理能力（語文推理、數學推理）有關，此顯示「轉換」能力可能涉及其他認知歷程。第五，「抑制」指標與效標工具幾乎沒有相關，可能顯示情境二的抑制

效果不大，從情境一與情境二的比較測量亦可發現是練習效果，這與本研究之參與者（8~29 歲的健康樣本）有關；本研究建議測驗使用者，若要評估健康樣本，可以僅測量 TB-DFT 的情境一、情境三。

在性別考驗中，設計流暢性與執行功能的表現上不存在差異性，突破語文能力的限制（Herlitz & Yonker, 2002），能更單純的評量個體之流暢力表現。在教育年的差異考驗中，顯示設計流暢性與執行功能的表現隨著教育年之增加而有上升的趨勢，表示額葉功能隨著年齡上升而發展（Mills & Tammes, 2014），且高中與大學（含）以上的流暢性能力相當。本研究針對不同樣本特性的差異分析結果，與過去實徵研究得到相同的結果（如 Ross et al., 2019; Van der Elst et al., 2011; Woods et al., 2016）。因此，本研究根據樣本特性的結果，進一步建立臺灣常模：國小中年級、國小高年級、國中、高中、大學（含）以上。測驗使用者可以根據施測對象參照常模對照表，例如：給國小四到六年級施測，三四年級學生的表現請參照「國小中年級」的常模；而五六年級學生的表現請參照「國小高年級」的常模。

本研究之貢獻。在領域知識中，過去執行功能的測量工具多以紙本（Delis et al., 2000）或自陳式（陳心怡等，2020；游勝翔、林緯倫，2018）為主，本研究為第一個開發操作式執行功能（「設計流暢性測驗」），不僅解決了計分的時間花費，業以客觀評估、不受社會讚許影響。在實際場域的貢獻中，本研究開發標準化 TB-DFT 並建立報表產出系統，測驗使用者僅需要在平板電腦中下載「TB-DFT」App 程式，便能獲得受測者在設計流暢性表現的原始資料（各情境的正確設計數及正確率）。此外，將原始資料置入報表分析系統中，便可以獲得該位受測者的測驗結果，並且統一將結果數據對照常模，針對受測者的表現給予適當評語。舉例而言，若國中生在三個情境的原始分數分別為 9、10、8，報表分析系統會自動對應國中常模，獲得三個情境的 PR 值為 78、71、77，表示在 100 人中，該位國中生在情境一的表現贏過 78 人、情境二贏過 71 人、情境三贏過 77 人。測驗使用者及受測者可以根據結果數據與評語，安排適合的訓練方案，精進自身的能力表現。

本研究尚有幾點限制。首先，本研究的信、效度考驗是以小學、中學、大學三個階段進行考驗為原則。然而，於收案期間遇到 Covid-19 疫情，學校實施遠距教學、或不便安排校外人員入校。在實體收案過程遇到阻礙，導致部分效標工具的人數不多，如「智力／推理能力」效標僅有中學階段的樣本，此為本研究之限制。未來研究可以增加探討小學、大學的智力推理表現

與設計流暢性之間的關係；或是開發「線上版」，將更能排除實體收案的限制。再者，本研究開發 TB-DFT 僅建立認知發展至穩定階段（8~29 歲），致使適用對象的限制。雖然設計流暢性不受語文能力影響，但仍無法確定幼兒或中老年人是否適合使用 TB-DFT，有待未來研究進一步檢驗。另外，本研究開發 TB-DFT 是以平板電腦的硬體作為收案工具，若想在桌上型電腦／筆記型電腦上進行施測，只要具備 App 模擬器，也能開啟。但桌上型電腦／筆記型電腦的作答方式是以滑鼠進行操作，與平板電腦的作答方式略有不同。若測驗使用者要對照本研究建立的常模，建議仍要使用平板電腦進行施測。

未來研究建議。在執行功能理論中，最初提出的執行功能三因素（抑制、轉換、更新）並沒有包含流暢性，但近年 Karr 等（2019）發現流暢性亦是執行功能的因素之一。本研究認為流暢性可能與其他因素（抑制、轉換、更新）涉及不同的歷程。過去執行功能的測量工具多涉及邏輯式、分析式的思考歷程（如轉換、更新、抑制等）（Diamond, 2013; Miyake et al., 2000）。根據雙重歷程理論（dual processes theory）提及人類擁有兩種類型的訊息處理模式（Evans, 2007; Evans & Stanovich, 2013）：類型一（Type 1）是直覺的、連結的處理歷程，類型二（Type 2）則是分析的、邏輯式的處理歷程。此理論也應用於其他領域，例如：創造力（林緯倫，2019；Lin & Lien, 2013; Lin et al. 2012），有研究曾將創造力分成開放式創造力與封閉式創造力，前者係能發想愈多愈不尋常想法的能力，倚賴類型一的思考；後者則是需要跳脫現有框架、改變問題表徵以提出唯一正解，具類型一與類型二之間轉換的參與。將其應用於執行功能範疇中，抑制、轉換、更新等被視為類型二的思考歷程（Benedek et al., 2012; Chuderski & Jastrzębski, 2018; Lee & Theriault, 2013），而自由聯想（free associative）、流暢性作業等則屬於類型一的連結歷程（Benedek & Neubauer, 2013）。因此，本研究認為在執行功能的因素中，「流暢性」與典型三因素（抑制、轉換、更新）涉及不同歷程，建議未來研究可以採用 TB-DFT，探討設計流暢性在執行功能理論中所扮演的角色，進而深化執行功能理論之探討，及執行功能與兩種訊息處理模式（類型一、類型二的思考）的關聯性。

本研究提出四點實務建議：第一，以團體施測方式蒐集行為資料，需安排相當數量的施測人員，以一對十五為上限。因為在施測過程中施測者須留意參與者對測驗的理解狀況，適時提供個別指導與說明。第二，針對 8~12 歲受測者（國小階段學生），須注意他們對測驗的理解。本研究 TB-DFT 雖然附有文字與影片的指導語說明，但本研究在蒐集 8~29 歲樣本的行為資料

時，發現國小階段學生往往自行開始，沒有注意施測人員的指示，以至於不理解測驗操作方式，需另外再個別指導說明。其他對象較不會出現這種情形，只需要在練習階段有錯誤時提供指導說明。整體而言，參與者有注意聽施測人員的指導語說明，並參照測驗指導語，皆能理解本測驗的操作方式。第三，針對臨床實務建議，可用以評估額葉或認知功能受損的個體，過去研究指出設計流暢性與右前額葉的受損有關（Jones-Gotman, 1991），透過 TB-DFT 的評估，可以初步從行為指標推論大腦受損的情形。第四，針對教學實務的建議，本研究建立健康樣本的常模，教學實務工作者可以透過 TB-DFT 的測驗結果，了解學生在設計流暢性與執行功能的表現；此外，TB-DFT 亦可提供特殊生與資優生評估其表現，將能檢視不同群體的認知功能表現。

陸、結論

本研究首次發展平板操作之設計流暢性測驗（TB-DFT），不僅有客觀評估、容易施測、計分快速之優點，能有效蒐集 8~29 歲對象的設計流暢性與執行功能（抑制、轉換）之表現。同時，採用具備標準化評估之測驗程序、數位化計分之一致性、自動化獲得結果數據的優勢與特性，亦能進一步比較常模結果進行分數解釋。本研究開發 TB-DFT，提供未來設計流暢性與執行功能研究之探討與發展，也提供教育或臨床領域一個重要的測量基礎，期能深化應用 TB-DFT 探究設計流暢性與執行功能理論的研究發現，進而實踐。

謝誌

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」、「華語文科技研究中心」，以及「大專校院人文與社會科學領域標竿計畫（教育學領域）」支持，特此感謝。本研究作者群感謝審查委員耐心指導與寶貴意見。

參考文獻

中文部分

- 周台傑、葉瓊華（2016）。**國民中學學業性向測驗七至九年級**。中國行為科學社。
- 林緯倫（2019）。創造力的不同類型：以雙重歷程理論探討開放式與封閉式創造力的差異。載於鄭英耀（主編），**華人創造力：理論與實務**（頁 319-338）。遠流。
- 陳心怡、陳美芳、胡心慈（2020）。**青少年執行功能行為評量表**。中國行為科學社。
- 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**（1），29-67。
- 楊宗仁（2002）。自閉症青少年之「執行功能」研究：認知彈性。**特殊教育研究學刊**，**22**，49-73。

英文部分

- Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20-29. <https://doi.org/10.1016/j.jecp.2009.11.003>
- Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*, *16*(1), 17-42. <https://doi.org/10.1007/s11065-006-9002-x>
- Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, *15*(1), 31-36. [https://doi.org/10.1016/S0887-6177\(98\)00159-0](https://doi.org/10.1016/S0887-6177(98)00159-0)
- Battista, P., Griseta, C., Tortelli, R., Guida, P., Castellana, F., Rivolta, D., & Logroscino, G. (2021). The Modified Five-Point Test (MFPT): Normative data for a sample of Italian elderly. *Neurological Sciences*, *42*(6), 2431-2440. <https://doi.org/10.1007/s10072-020-04818-3>
- Benedek, M., & Neubauer, A. C. (2013). Revisiting Mednick's model on creativity-related


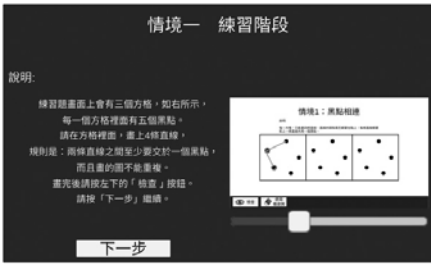
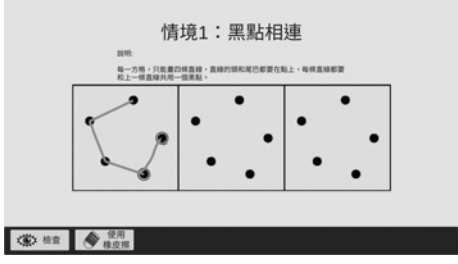
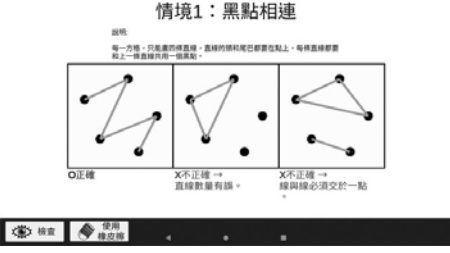
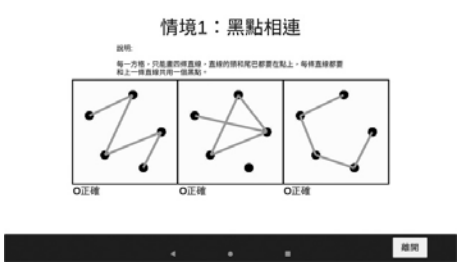
- differences in associative hierarchies. Evidence for a common path to uncommon thought. *The Journal of Creative Behavior*, 47(4), 273-289. <https://doi.org/10.1002/jocb.35>
- Benedek, M., Könen, T., & Neubauer, A. C. (2012). Associative abilities underlying creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 6(3), 273-281. <https://doi.org/10.1037/a0027059>
- Bryan, J., & Luszcz, M. A. (2000). Measurement of executive function: Considerations for detecting adult age differences. *Journal of Clinical and Experimental Neuropsychology*, 22(1), 40-55. [https://doi.org/10.1076/1380-3395\(200002\)22:1;1-8;FT040](https://doi.org/10.1076/1380-3395(200002)22:1;1-8;FT040)
- Burgess, P. W., & Alderman, N. (2018). Rehabilitation of dyscontrol syndromes following frontal lobe damage: A cognitive neuropsychological approach. In P. W. Burgess & N. Alderman (Eds.), *Cognitive rehabilitation in perspective* (pp. 183-203). Routledge.
- Cattalani, R., Dal Sasso, F., Corsini, D., & Posteraro, L. (2011). The Modified Five-Point Test: Normative data for a sample of Italian healthy adults aged 16-60. *Neurological Sciences*, 32(4), 595-601. <https://doi.org/10.1007/s10072-011-0489-4>
- Chelune, G. J., Holdnack, J., & Levy, J. (2006). Education effects and base-rate information for the Delis-Kaplan Executive Function System (DKEFS) fluency measures. *Archives of Clinical Neuropsychology*, 21, B45.
- Chuderski, A., & Jastrzębski, J. (2018). Much ado about aha!: Insight problem solving is strongly related to working memory capacity and reasoning ability. *Journal of Experimental Psychology: General*, 147(2), 257-281. <https://doi.org/10.1037/xge0000378>
- 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>
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system*. The Psychological Corporation. <https://doi.org/10.1037/t15082-000>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135-168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience*, 18, 34-48.
- Dumontheil, I., & Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cerebral Cortex*, 22(5), 1078-1085. <https://doi.org/10.1093/cercor/bhr175>

- Evans, J. S. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*. Psychology Press. <https://doi.org/10.4324/9780203947487>
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on Psychological Science*, 8(3), 223-241. <https://doi.org/10.1177/1745691612460685>
- Heaton, R. K., & Staff, P. (1993). *Wisconsin Card Sorting Test: Computer Version-2*. Psychological Assessment Resources.
- Herlitz, A., & Yonker, J. E. (2002). Sex differences in episodic memory: The influence of intelligence. *Journal of Clinical and Experimental Neuropsychology*, 24(1), 107-114. <https://doi.org/10.1076/jcen.24.1.107.970>
- Izaks, G. J., Joosten, H., Koerts, J., Gansevoort, R. T., & Slaets, J. P. (2011). Reference data for the Ruff Figural Fluency Test stratified by age and educational level. *PLoS ONE*, 6(2), e17045. <https://doi.org/10.1371/journal.pone.0017045>
- Jones-Gotman, M. (1991). Localization of lesions by neuropsychological testing. *Epilepsia*, 32, S41-52.
- Jones-Gotman, M., & Milner, B. (1977). Design fluency: The invention of nonsense drawings after focal cortical lesions. *Neuropsychologia*, 15(4-5), 653-674. [https://doi.org/10.1016/0028-3932\(77\)90070-7](https://doi.org/10.1016/0028-3932(77)90070-7)
- Karr, J. E., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2019). Examining the latent structure of the Delis-Kaplan executive function system. *Archives of Clinical Neuropsychology*, 34(3), 381-394. <https://doi.org/10.1093/arclin/acy043>
- Lee, C. S., & Theriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence*, 41(5), 306-320. <https://doi.org/10.1016/j.intell.2013.04.008>
- Lin, W.-L., & Lien, Y.-W. (2013). The different role of working memory in open-ended versus closed-ended creative problem solving: A dual-process theory account. *Creativity Research Journal*, 25(1), 85-96. <https://doi.org/10.1080/10400419.2013.752249>
- Lin, W.-L., Hsu, K.-Y., Chen, H.-C., & Wang, J.-W. (2012). The relations of gender and personality traits on different creativities: A dual-process theory account. *Psychology of Aesthetics, Creativity, and the Arts*, 6(2), 112-123. <https://doi.org/10.1037/a0026241>
- Madden, D. J. (2001). Speed and timing of behavioral processes. *Handbook of the Psychology of Aging*, 5, 288-312.
- McClelland, M. M., & Cameron, C. E. (2011). Self-regulation and academic achievement

- in elementary school children. *New Directions for Child and Adolescent Development*, 133, 29-44. <https://doi.org/10.1002/cd.302>
- Mills, K. L., & Tamnes, C. K. (2014). Methods and considerations for longitudinal structural brain imaging analysis across development. *Developmental Cognitive Neuroscience*, 9, 172-190. <https://doi.org/10.1016/j.dcn.2014.04.004>
- 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>
- Moore, M. E., & Ross, B. M. (1963). Context effects in running memory. *Psychological Reports*, 12(2), 451-465.
- Ni, T. L., Huang, C. C., & Guo, N. W. (2011). Executive function deficit in preschool children born very low birth weight with normal early development. *Early Human Development*, 87(2), 137-141. <https://doi.org/10.1016/j.earlhumdev.2010.11.013>
- Phillips, L. H. (1997). Do “frontal tests” measure executive function? Issues of assessment and evidence from fluency tests. In Rabbitt P. (Ed.), *Methodology of frontal and executive function* (pp. 191-213). Psychology Press.
- Regard, M., Strauss, E., & Knapp, P. (1982). Children’s production on verbal and non-verbal fluency tasks. *Perceptual and Motor Skills*, 55(3), 839-844. <https://doi.org/10.2466/pms.1982.55.3.839>
- Ross, T. P., O’Connor, S., Holmes, G., Fuller, B., & Henrich, M. (2019). The reliability and validity of the action fluency test in healthy college students. *Archives of Clinical Neuropsychology*, 34(7), 1175-1191. <https://doi.org/10.1093/arclin/acz016>
- Ruff, R. M., Light, R. H., & Evans, R. W. (1987). The Ruff Figural Fluency Test: A normative study with adults. *Developmental Neuropsychology*, 3(1), 37-51. <https://doi.org/10.1080/87565648709540362>
- Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. *Advances in Psychology*, 65, 31-86.
- St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology*, 59(4), 745-759. <https://doi.org/10.1080/17470210500162854>
- Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, 47(1), 1-41. <https://doi.org/10.1353/mpq.2001.0005>
- Swanson, J. (2005). The Delis-Kaplan executive function system: A review. *Canadian*

- Journal of School Psychology*, 20(1-2), 117-128. <https://doi.org/10.1177/0829573506295469>
- Tucha, L., Aschenbrenner, S., Koerts, J., & Lange, K. W. (2012). The five-point test: Reliability, validity and normative data for children and adults. *PLoS ONE*, 7(9), e46080. <https://doi.org/10.1371/journal.pone.0046080>
- Van der Elst, W., Hurks, P., Wassenberg, R., Meijs, C., & Jolles, J. (2011). Animal verbal fluency and design fluency in school-aged children: Effects of age, sex, and mean level of parental education, and regression-based normative data. *Journal of Clinical and Experimental Neuropsychology*, 33(9), 1005-1015. <https://doi.org/10.1080/13803395.2011.589509>
- Woods, D. L., Wyma, J. M., Herron, T. J., & Yund, E. W. (2016). A computerized test of design fluency. *PLoS ONE*, 11(5), e0153952. <https://doi.org/10.1371/journal.pone.0153952>
- Xu, C., Ellefson, M. R., Ng, F. F.-Y., Wang, Q., & Hughes, C. (2020). An East-West contrast in executive function: Measurement invariance of computerized tasks in school-aged children and adolescents. *Journal of Experimental Child Psychology*, 199, 104929. <https://doi.org/10.1016/j.jecp.2020.104929>
- Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2016). *Executive function: Implications for education*. NCER 2017-2000. National Center for Education Research.

附錄

<p>圖 A. 基本資料填寫</p> 	<p>圖 B. 指導語介面</p> 
<p>圖 C. 練習階段</p> 	<p>圖 D. 練習階段說明錯誤原因</p> 
<p>圖 E. 練習階段完全正確的回饋</p> 	<p>圖 F. 測驗結束介面</p> 