

# 更新能力測驗之編製及發展軌跡研究

郭郡羽<sup>1</sup>、陳珮臻<sup>2</sup>、陳志洪<sup>3</sup>、陳學志<sup>2, 4, 5, 6</sup>

更新 (updating) 是一種重要的執行功能，涉及在工作記憶中持續調整和更新信息，以適應環境的變化。本研究旨在發展一套標準化的評估工具，以便更準確地測量更新能力並進一步探討更新能力的發展軌跡。本研究首先回顧了現有的更新能力測量方法，並指出了這些方法在與其他執行功能區分上的侷限性。為了解決這一問題，本研究基於 **running memory task** 設計了一個平板版的更新測驗，該測驗可計算最高廣度指標和更新分數指標來反映個體之更新能力。為了驗證更新測驗的有效性，本研究採用了多種效標關聯效度分析，結果顯示更新測驗的表現與處理速度、軌跡標示測驗、分類測驗及學習成效均具有顯著相關性。此外，更新測驗在六個月的再測信度檢驗中也表現出良好的穩定性。本研究進一步探討了更新能力的發展軌跡，發現更新能力在 7 歲至成年早期之間隨年齡增長逐漸提高，並在成年後趨於穩定。本研究的結果不僅強調了更新能力在不同年齡段的發展特徵，也為未來在教育及臨床實務中對更新能力的評估提供了可靠的工具和理論支持。

**關鍵詞：**平板、更新能力、執行功能、發展軌跡

<sup>1</sup> 國立臺灣師範大學社會教育學系

<sup>2</sup> 國立臺灣師範大學教育心理與輔導學系

<sup>3</sup> 國立臺灣師範大學資訊教育研究所

<sup>4</sup> 國立臺灣師範大學學習科學跨國頂尖研究中心

<sup>5</sup> 國立臺灣師範大學華語文與科技研究中心

<sup>6</sup> 國立臺灣師範大學社會情緒發展研究中心

**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](mailto: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」、「華語文科技研究中心」及「社會情緒教育與發展研究中心」支持，特此感謝。

請想像一個日常生活中常經歷到的情境，你詢問朋友他的分機號碼，他回答「3184，喔不，我的分機剛改過，應該是3165。」在理想的情況下，我們應該要將腦海裡的末兩位數字84取代為65，並以3165作為最後需要撥號時的依據。這個認知運作歷程即為更新（updating）。更新能夠避免不再重要的訊息（例子中的84）對我們的記憶與行為產生影響（De Beni & Palladino, 2004; Oberauer, 2001）。

更新（updating）與抑制（inhibition）、轉換（shifting）同為重要的執行功能組成成分，彼此獨立且共同運作在工作記憶之中（Miyake et al., 2000）。其中更新係指：隨著外界刺激的變化，不斷地調整記憶中表徵的能力（Morris & Jones, 1990, p. 112）。這個定義中涵蓋兩個重要的歷程：用新表徵取代工作記憶中的舊有表徵、捨棄與當下作業不再相關的表徵。這兩個歷程對於日常生活極為重要，在面臨時時變動的環境時，我們須將容量有限的資源投注在與當下任務相關的訊息上，並捨棄不再重要的舊訊息。此外，更新能力也是許多高階認知能力的有效預測指標，例如：流體智力（Santarecchi et al., 2021; Yntema & Mueser, 1962）、學習成就（Gathercole et al., 2004; Gijsselaers et al., 2017）、推理（Barrouillet, 1996; Barrouillet & Lecas, 1999）、閱讀理解（Daneman & Merikle, 1996; Just & Carpenter, 1992; Palladino et al., 2001; Swanson & Berninger, 1995; Zaccoletti et al., 2020）、數學能力（Lee et al., 2012; Toll et al., 2011; Zheng et al., 2011）。甚或在Chen與Li（2007）、Friedman等人（2006）的研究中都發現更新是唯一能預測流體智力的執行功能成分。因此許多研究透過提升更新能力來改善學習表現，例如Zhang等人（2018）透過訓練更新能力來提升中學生的數學成績；Gao等人（2023）則成功透過更新能力的訓練提升閱讀理解。

### （一）精準評估更新能力的必要性

有鑑於更新能力與許多重要能力間的緊密連結，正確評估個體的更新能力成為一個重要的議題。研究者們多以工作記憶相關作業來測量更新能力，例如：複雜工作記憶作業（n-back task）、工作記憶廣度作業（forward or backward digit span task）（Wechsler, 1955）、閱讀廣度測驗（reading span test）（Daneman & Carpenter, 1980）。某些研究中同時使用此些作業作為更新能力的指標（Schmiedek et al., 2009）。然而，這些作業中不僅需要更新能力，往往也同時使用到了其他與工作記憶相關的能力（例如：維持、轉換等）。如此一來，更新難以與工作記憶中的其他歷程區分開來。因著過去慣常使用的測驗中往往同時涉及多項認知功能，因此其作業表現難以清楚反映個體之更新能力，而更新能力被視為執行功能中的獨立成分，因此更新能力在理論建構及概念上都具有被明確區別開來的必要性（Ecker et al., 2014）。

Ecker等人（2010）的研究中設計了八種不同情境來反映更新歷程中涉及的三種主要認知成分：提取（retrieval）、轉換（transformation）、抽換（substitution）。我們可以透過下面的例子來說明這三種認知成分，當你想以台幣在國外購物網站上購買東西時，你看到標價上寫10美元，此時若依照1:32的匯率轉換，你知道你需要付320元台幣。這個過程中，你需要提取10，並將其轉換為 $10 \times 32$ ，並在最後結帳時將商品金額更新為320元台幣。Ecker等人透過結構方程模式分析作業表現正確率，結果發現提取與轉換與工作記憶廣度相關，但抽換則無關。這表示抽換是唯一能夠獨立代表更新能力的歷程。

在上述文獻中，我們可以看出準確評估更新能力的必要性。然而，實證研究中使用的更新能力評估方法各有不同，且可能涉及其他與工作記憶相關的過程，這使得難以精確地提取出更新能力的表現。因此，本研究的主要目的是設計一個標準化的評估工具，用於測量更新能力，以便為基礎研究、應用及臨床工作者提供便捷的評估方法。

### （二）更新能力的評估方式

傳統實驗室情境中用以評估更新能力的方法主要包含：running memory task（Morris & Jones, 1990）、keeping track task（Yntema, 1963）、n-back task（Kirchner, 1958）等。running memory task將要求參與者在序列呈現的刺激中記得最後呈現的n個，並依序進行回報，其中常用的刺激屬性可能為數字、字母、物體、位置等不同版本。以字母版本為例，若當下任務難度為須回憶最後3個刺激，

而刺激序列為「G → E → K → N → Q → T → B」，則參與者須回答：Q、T、B。keeping track task 中將會依序呈現不同類別的刺激，例如：字母、顏色、數字、動物等，參與者需要記憶每種類別中最後出現的刺激，例如出現刺激序列為：「紅色→老虎→H→綠色→B」，最終參與者需要分別回憶顏色、動物及字母類別中的最末一個刺激：綠色、老虎、B。n-back task 是記憶負荷相對較高的更新測驗，參與者須針對序列呈現的刺激中的每一個進行反應，判斷其餘前 n 個是否相同。例如刺激序列為「3 → 7 → 4 → 8 → 4 → 9 → 2」並進行 2-back 任務難度時，參與者須將每個出現的刺激與前 2 個出現的刺激進行比較，並判斷兩個刺激是否相同。因此當第三個刺激 4 出現時，其與第一個出現的刺激 3 不同，因此須回答「不同」；而當第五個刺激 4 出現時，其與第三個刺激 4 相同，因此須回答「相同」。此些電腦化測驗的方式的優勢在於能夠將更新能力從工作記憶廣度中獨立地分離出來，更精準地評估個體移除舊訊息並登錄新表徵的能力，亦能更細緻區辨健康個體更新能力上的差異。缺點則在於評估及計分的要求較高，一方面需要硬體及程式資源的協助；另一方面也需要針對參與者的狀況進行難度調整，因此測驗時間往往較長。在計分時也需要對資料進行前處理後才能產生有意義的指標。因此在臨床及實務應用場域中較難廣泛使用。

在實務情境中對更新能力的測量則偏向鑑別診斷，因此往往採用難度較低、但更具生態效度的版本，例如：selective updating of sentences task (Fellman et al., 2017)。在此作業中主試者會呈現一個句子，裡面可能包含人物、地點、時間、物品的關鍵訊息，例如「小明昨天把他的書放在地上」，接著再持續更新句子中的關鍵訊息，例如再次呈現「現在，小明把他的書放到椅子上」，最後詢問參與者關於句子中的關鍵訊息，例如「現在書在哪裡？」此時須回答「椅子上」。神經心理學測驗中亦有些分測驗被用來作為更新能力評估的參考，例如：Wisconsin card sorting test (Grant & Berg, 1948)，在此測驗中參與者須依照規則將帶有不同特徵的卡片進行分類，例如：依照顏色分類；並在下一個階段中轉換分類的規則，例如依照形狀分類。其失敗多反映由於舊規則固著所導致的持續性錯誤上 (Nagahama et al., 2005)，因此得以藉由觀察是否能成功轉換規則來反映更新能力。軌跡標示測驗 (trail making test) (Reitan, 1955) 則需要參與者在不同向度的特徵中轉換並完成連續作業，測驗時會在空間中的隨機位置上呈現多個數字及英文字母 (3、D、K、7、2、A)，參與者須依照數字→字母→數字→字母的順序將數字及字母依序由小連線到大，在上述例子中，參與者須正確依照：「2 → A → 3 → D → 7 → K」的順序作答。digit span backward (Wechsler, 1955) 則須將聽到的數字依倒序的方式重新念出，例如聽到「3、5、1、8、5、6、2」，則須回應「2、6、5、8、1、5、3」。上述作業中所評估的能力多涉及多樣執行功能的組成，例如：轉換、工作記憶廣度等，因此其表現雖能涵蓋更新能力於其中，卻無法精準評估更新能力本身的狀態，因此較偏向在臨床的廣泛性認知能力鑑別診斷上。

基於對上述用以評估更新測驗的回顧與了解，在理論向度上，根據 Ecker 等人 (2010) 的研究，我們選用涉及抽換歷程最多的作業；在實務向度上，我們同時考量測驗是否能夠精準分離並評估更新能力及施測時長，選擇了 running memory task 作為更新測驗設計的基礎，以此版本進行改良並設計完整的背景計分模組，以便能夠在實務場域中達到準確評估及便利計分的目標。

### (三) 更新測驗之效標與理論基礎

本研究選用處理速度、軌跡標示測驗表現、分類測驗表現與學習成效作為效標以檢驗本研究所設計之更新測驗之有效性。以下將依序說明各效標選擇之原因與理論基礎：首先，根據處理速度理論 (Salthouse, 1996) 認為處理速度必然與更新測驗的表現相關。在實徵研究中亦可觀察到在工作記憶中登錄每個表徵的時間約為 500 ~ 600 毫秒 (Jolicoeur & Dell'Acqua, 1998; Vogel et al., 2006)，將之移除則需耗費更長的時間 (Oberauer, 2001)，這表示處理速度與更新能力間存在必然的關聯性 (Albinet et al., 2012; Chen & Li, 2007; Schubert et al., 2018)。因此本研究以處理速度作為效標。

第二，軌跡標示測驗常被用在臨床鑑別診斷上，其主要評估轉換的能力，在測驗中要求參與者須在兩個規則中轉換，並選擇符合當下規則的刺激連線，其過程同時涉及將工作記憶中的舊訊息抽換為新訊息。根據 Ecker 等人 (2010) 的研究，抽換是唯一能夠獨立代表更新能力的歷程，因此本研究選擇此作業表現作為效標之一。我們預期更新能力越佳者在軌跡標示測驗上的錯誤次數會越低、完成時間越短。

第三，Wisconsin card sorting test 得以反映個體在規則固著上的認知失敗，常用於臨床神經心理學測驗中。參與者須不斷捨棄舊規則並抽換新規則以持續完成作業，抽換歷程歷程涉及更新能力（Ecker et al., 2010）。因此本研究中採用此類測驗作為效標，我們預期更新能力越佳者，在此作業上的表現越好。

最後，更新能力與學習表現存在緊密的關聯性（Gijsselaers et al., 2017; Lechuga et al., 2016），更新能力對閱讀理解（Iglesias-Sarmiento et al., 2015; St Clair-Thompson & Gathercole, 2006; Zaccoletti et al., 2020）、寫作能力（Gao et al., 2023）、數學表現（Lechuga et al., 2016; Lee & Bull, 2016）都有正向影響，甚或針對學習障礙學生的研究中亦發現更新能力的訓練有助於提升其數學表現（Zhang et al., 2018）。因此本研究分別採用語文及數理相關成就表現作為效標以評估更新測驗的有效性。

#### （四）更新能力的發展軌跡與常模建置

更新能力隨著年齡發展而逐漸進步，在幼兒階段（4 至 7 歲），隨著年齡增加，更新能力越來越好（Cheng & Kibbe, 2022）；接續幼兒階段的研究，學齡階段的研究發現更新能力從 7 歲開始到 15 歲逐漸提升（Lechuga et al., 2006），並在成年早期達到穩定（Carriedo et al., 2016; Linares et al., 2016），此與前額葉皮質的突觸修剪發展軌跡相仿（Huttenlocher, 1990）。由於用以評估更新能力的作業相當多元，因此其發展軌跡的評估尚須視作業需求而定：當作業涉及高度執行控制時，發展歷程將會持續到 16 歲並在 18 至 20 歲左右穩定下來；當任務需要低度執行控制時，發展在 11 至 12 歲時即已穩定（Conklin et al., 2007; Luciana et al., 2005）。例如在 Carriedo 等人（2016）的研究中使用需要高度執行控制的作業評估更新能力，研究中包含七歲到成年前期（20 至 35 歲）不同年齡階段的參與者，結果發現更新能力隨年齡增加而不斷變好。這的發展趨勢可能源自於更新能力仰賴工作記憶的整體發展，兒童期時工作記憶的核心區域（如前額葉皮層）主要負責基本的記憶更新和維持功能，隨著年齡增長，這些核心區域變得更高效、更專門化，能夠更有效地處理信息，從而提高工作記憶的精確性和速度，進而展現較佳的更新能力（Pelegrina et al., 2020）。

更新能力到成年前期逐漸發展穩定後，高齡階段是否會逐漸減退？此部分的文獻仍有不一致的證據，部分研究認為進入高齡階段後，更新能力會隨之減退（Basak & O'Connell, 2016; Basak & Zelinski, 2013; Braver et al., 2005; De Beni & Palladino, 2004），並對執行控制能力產生負面的影響（Podell et al., 2012）。神經生理的證據亦顯示高齡者在執行更新測驗時，左 precentral gyrus 與右 cerebellum 的活化減少（Qin & Basak, 2020），且其與更新測驗的表現下降相關。亦有部分研究顯示更新能力並不隨年齡增長而退化（Lange & Verhaeghen, 2009）。隨年齡而出現衰退效果可能因情境而異，Van der Linden 等人（1994）則發現在記憶負荷量低的時候，高齡者和年輕人在更新測驗上的表現相當，但記憶負荷提升時，高齡者的表現會較年輕人差。

上述的實徵研究證據都展現了更新能力隨著年齡逐漸發展已臻成熟的軌跡，然而，當我們回顧台灣相關的文獻時，卻發現因著評估工具的不便利與發散，而難以在研究間彼此對照，亦難取得跨年齡階段的資料來觀察其發展趨勢。因此除了工具的發展外，我們也同時建立了發展階段中的本土化常模，透過完整學齡階段的常模資料，未來的研究者及教育或臨床實務工作者可以將所獲得的資料與常模資料進行對照，這將能夠更清晰地評估個體在相對應群體中的表現以便正確詮釋數據。

## 方法

### （一）參與者

本研究依分層叢集抽樣，招募臺灣北、中、南、東四個地區之在學學生為研究對象，依地區人口比例 4 : 3 : 3 : 1 進行抽樣，其教育程度分組參考 D-KEFS（Delis et al., 2001），將參與者之教育程度分為：國小中年級、國小高年級、國中、高中、大學（含）以上。根據 Oosterhuis 等人（2016）之建議，發展測驗階段時每個子群體應招募至少 100 ~ 200 人，因此本研究於每個教育程度階段預計招募 200 位參與者，然而實際招募情況中，國小至高中階段皆以班級為單位進行招募，因此實際招募人數與預計招募人數略有不同。本研究據此建立更新測驗之本土化常模。參與者排除矯正後視

覺障礙（含色盲／色弱）、聽力障礙、上肢／上半身運動障礙、精神或神經疾病、一週曾有兩天以上飲用三份以上之酒精飲品、曾無故失去意識 5 分鐘以上、曾因腦傷住院、曾有認知或記憶障礙病史、曾確診為學習障礙者。共蒐集 1,222 位自願者參與此研究，排除資料填答不完整及無法正確完成測驗者，有效樣本共計 1,164 人，年齡區間為：11 ~ 33 平均年齡為 17.93 歲，標準差為 3.94；其中女性占 51%。參與者分布如表 1。

**表 1**  
參與者之人口學變項（地區、性別、教育年組）分布情形

	北		中		南		東		總計
	男	女	男	女	男	女	男	女	
國小	58	55	55	53	42	38	22	19	342
國中	62	61	25	26	60	63	20	21	338
高中	55	70	40	25	32	39	17	16	294
大學（含）以上	42	31	20	20	17	37	9	14	190
總計	217	217	140	124	151	177	68	70	1164

信、效度評估之參與者資料如下：50 位參與者在第一次受測後約六個月再次進行更新測驗，用以評估再測信度。其中女性比例為 44%，平均年齡為 16.10 歲，標準差為 2.94。效標關聯效度部分，我們選用以下四種作業表現作為效標：處理速度作業表現、軌跡標示測驗表現、分類測驗表現、學習成效。本研究邀請 641 位參與者同時參與處理速度作業（女性比例為 49%，平均年齡為 19.04 歲，標準差為 4.18），並計算簡單反應時間及選擇反應時間作為處理速度的指標，該指標越大表示處理速度越慢。729 位參與者同時參與軌跡標示測驗（女性比例為 50%，平均年齡為 18.87 歲，標準差為 4.47），並計算表現錯誤次數與完成時間作為指標。767 位參與者同時參與分類測驗（女性比例為 50%，平均年齡為 18.74 歲，標準差為 4.41），並計算表現正確率作為指標。在學習成效部分，本研究蒐集了 118 位 7 至 12 年級學生之學科學習表現，其中女性比例為 43%，平均年齡為 14.25 歲，標準差為 1.49。

## （二）研究工具

### 1. 更新測驗

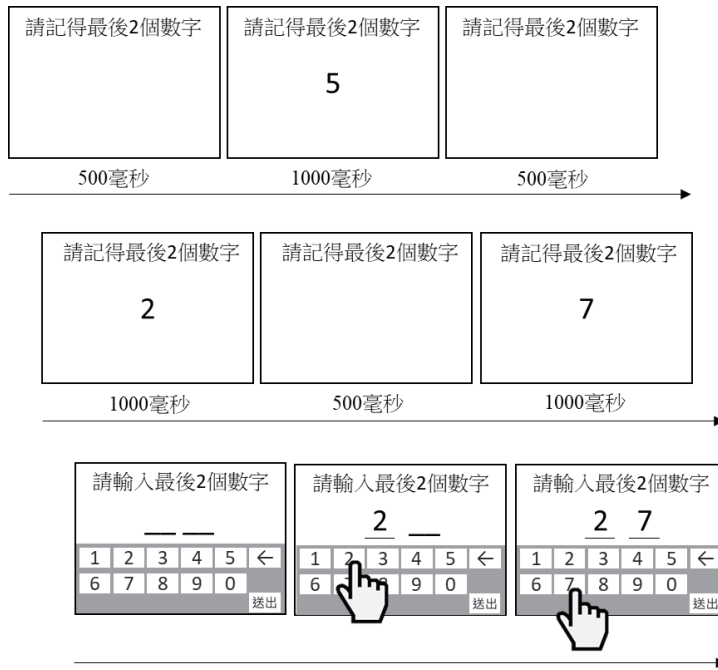
（1）作業設計。參與者須根據難度指示記憶相關數量的數字，並在每個嘗試的最末輸入答案。在難度 1 時，參與者須記憶所出現數字序列中的最末一個數字，例如：數字序列為：5 → 2 → 7 → 3 → 6 → 4，則參與者須輸入 4 為正確答案；在難度 2 時，參與者須記憶所出現數字序列中的最末兩個數字，因此在上述的數字序列中，參與者須輸入 6、4 為正確答案；在難度 3 時，正確答案則為 3、6、4，依此類推。難度等級每增加 1，所需記憶的數字序列則增加 1。本更新測驗由練習階段與測驗階段所構成。練習階段有 6 個嘗試，前 4 個嘗試難度為 1，後 2 個嘗試難度為 2。練習階段中的嘗試若答對則自動進行下一個嘗試；答錯將會重覆呈現該嘗試，直到答對為止。6 個練習嘗試全數回答正確後即開始正式階段。正式階段的難度由 1 開始，最高至 8 為止，每個難度中以兩個嘗試為一組，兩個嘗試均答對則難度 +1；兩個嘗試中答對一個則維持在同一難度，再進行一組（兩個嘗試）；兩個嘗試均答錯則難度 -1。當第四次進入某特定難度後則測驗中止。例如某參與者的作答軌跡為：難度 1，正確、正確（兩個嘗試均正確，難度 +1）→ 難度 2，正確、正確（兩個嘗試均正確，難度 +1）→ 難度 3，正確、正確（兩個嘗試均正確，難度 +1）→ 難度 4，正確、錯誤（一個嘗試正確，難度維持不變）→ 難度 4，正確、正確（兩個嘗試均正確，難度 +1）→ 難度 5，錯誤、正確（一個嘗試正確，難度維持不變）→ 難度 5，錯誤、錯誤（兩個嘗試錯誤，難度 -1）→ 難度 4，正確、錯誤（一個嘗試正確，難度維持不變）→ 中止測驗（第四次進入特定難度則中止測驗）。

（2）版面設計。更新測驗數字呈現畫面如圖 1 的第二個畫面示意圖。螢幕底色為白色，文字與數字均以黑色呈現。視角以眼睛至螢幕距離 65 公分來計算，每個中文字體大小為 0.8° × 0.8°，螢幕中

央呈現之數字大小為  $1.7^\circ \times 1.7^\circ$ 。更新測驗作答畫面呈現如圖 1 的最後一個畫面示意圖。螢幕上半部為答案輸入區塊，螢幕底色為白色；螢幕下半部為按鍵區塊，螢幕底色為淺灰色，按鍵底色為白色，文字與數字均以黑色呈現。答案輸入區塊的中文字體與數字大小與位置均與數字呈現畫面相同，作答區底線 2.3°寬。數字按鍵大小： $3^\circ \times 2.5^\circ$ ，每個按鍵間隔  $0.5^\circ$ 。

(3) 嘗試流程與指導語說明。作業一開始螢幕上方將會出現「請記得最後 n 個數字」的字樣，n 會視當時需記憶的數量而定。例如，在難度 2 的情況中，會出現「請記得最後 2 個數字」。接著螢幕中央會依序出現數字，每個數字呈現 1,000 毫秒，並間隔 500 毫秒後出現下一個。出現的數量將隨機由 10 ~ 25 之間抽選，且必定大於所需記憶數量。當所有數字出現完畢後，螢幕上方會出現「請輸入最後 n 個數字」，且螢幕中央會出現兩個底線，螢幕下方會出現 0 ~ 9 的數字按鍵，參與者須以觸碰數字按鍵的方式依序輸入最末兩個數字，輸入的答案將會即時呈现在螢幕中央。若輸入錯誤可按「←」鍵清除後重新輸入，輸入完成後按「送出」鍵送出答案，即完成一個嘗試，螢幕上會自動出現下一個嘗試。

圖 1  
更新測驗嘗試範例



(4) 指標與計分。更新測驗一共有兩個指標：最高廣度指標、更新分數。最高廣度指標指參與者能夠連續正確完成兩個嘗試的最高難度，最低分為 0，最高分為 8。除了傳統的最高廣度指標外，我們額外計算更新分數指標，更新分數可以反應在同一最高廣度下個體的穩定度，最低分為 0，最高分為 36，在同樣最高廣度的情況下，該數值越高表示其更新能力的穩定度越理想，其計算公式如下：

$$\text{更新分數} = \sum_{i=1}^n \frac{\text{難度 } i \text{ 的正確嘗試數}}{\text{難度 } i \text{ 的總嘗試數}} \times \text{難度值 } i$$

其中難度值  $i$  表示該嘗試需記憶的項目數量， $n$  為最高廣度。例如，某參與者的作答軌跡為：難度值 1，正確、正確（兩個嘗試均正確，難度 +1）→ 難度值 2，正確、正確（兩個嘗試均正確，

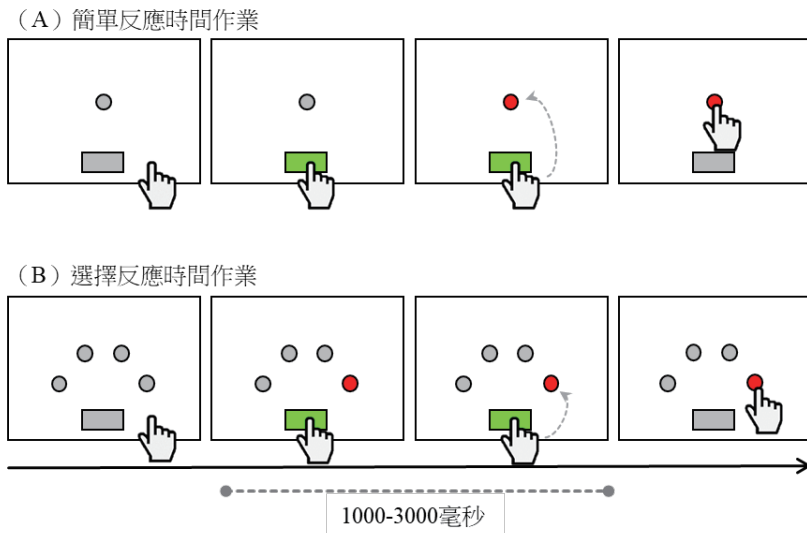
難度 +1) → 難度值 3, 正確、正確 (兩個嘗試均正確, 難度 +1) → 難度值 4, 正確、錯誤 (一個嘗試正確, 難度維持不變) → 難度值 4, 正確、正確 (兩個嘗試均正確, 難度 +1) → 難度值 5, 錯誤、正確 (一個嘗試正確, 難度維持不變) → 難度值 5, 錯誤、錯誤 (兩個嘗試錯誤, 難度 -1) → 難度值 4, 正確、錯誤 (一個嘗試正確, 難度維持不變) → 中止測驗 (第四次進入特定難度則中止測驗)。則最高廣度為 4, 更新分數為:

$$\text{更新分數} = \left(\frac{2}{2} \times 1\right) + \left(\frac{2}{2} \times 2\right) + \left(\frac{2}{2} \times 3\right) + \left(\frac{4}{6} \times 4\right) + \left(\frac{1}{4} \times 5\right) = 8.12$$

## 2. 效度考驗工具

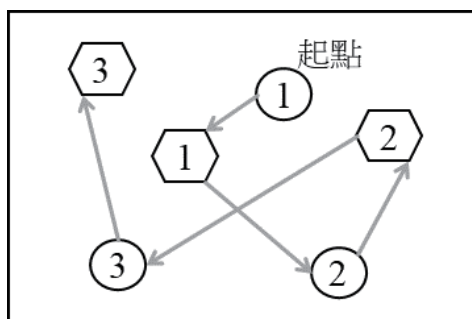
(1) 處理速度測驗。本研究採用簡單反應時間作業與選擇反應時間測驗來評估參與者的處理速度, 此二測驗最早由 Luce (1991) 所設計, 並由郭郡羽等人 (2004 年) 改編為中文版。在簡單反應時間測驗中, 參與者須將手指放置在準備位置上, 並在目標物變色時快速點擊目標物位置; 在選擇反應時間測驗中, 參與者須將手指放置在準備位置上, 並在四個目標物中其中一個變色時快速點擊變色目標物的位置。圖 2 為處理速度測驗之流程示意圖, 在該範例中參與者須先把手放置在下方方形的準備區, 等待上方圓形目標物變色後, 盡量快速且正確地點及變色的目標物。其中從手正確放置在準備區到目標物變色的間隔時間將由 1,000、1,250、1,500、1,750、2,000、2,250、2,500、2,750、3,000 間隨機抽選。主要產生的指標分別為簡單反應時間與選擇反應時間, 計算方式為: 刪去錯誤反應的嘗試、刪去反應時間在 80 毫秒以下、反應時間在  $\pm 2.5$  個標準差以外的嘗試 (Van Selst & Jolicoeur, 1994), 計算餘下之嘗試之平均數。該數值越小表示處理速度越快。

圖 2  
處理速度測驗程序示意圖



(2) 軌跡標示測驗。本研究選用黃詩媛等人 (2004) 所編製之「軌跡標示測驗平板版」作為工具, 該版本之測驗再測信度良好 ( $r = .40 \sim .55$ ), 且與彩色路徑描繪測驗、西蒙作業、塔測驗、反應時間測驗、中文版執行功能都有良好的效標關聯效度 ( $r = .16 \sim .38$ )。參與者須在形狀與數字間進行連線, 每一個選擇項目需轉換不同的形狀及更新數字, 例如: 圓形 1 → 六邊形 1 → 圓形 2 → 六邊形 2 → 圓形 3 → 六邊形 3……, 示意圖請參考圖 3。本研究中此測驗之使用正確率及完成時間作為指標。

圖 3  
軌跡標示測驗示意圖



(3) 分類測驗。本研究選用已中文化的分類測驗（原始版本為 Delis et al., 2001）進行施測。在此測驗中參與者將看到六張卡片，每張卡片上都有不同顏色、形狀、字詞所組成的群組刺激，參與者需將卡片依規則分為兩類，接著打亂卡片，再次進行分類，接續進行。例如：先依顏色進行分類，將黃色與藍色卡片分為兩堆；接著更換規則，改以是否有生命進行分類，將有生命與無生命的卡片分為兩堆；以此類推持續進行類別更新與分類。已用過的分類方式將不可再次被使用。本研究使用分類測驗之準確率為指標，其計算方式為：

$$\text{分類測驗準確率} = \frac{\text{總正確分類}}{\text{總嘗試分類}} \times 100\%$$

其中，總嘗試分類表示參與者所有進行的分類次數，總正確嘗試表示參與者正確進行分類且不重覆之分類次數。

(4) 學習成效。在學習成效部分，本研究蒐集了 118 位 7 至 12 年級之參與者各學科學習表現，並依據文、理科分別平均，接著在班級內計算標準化分數，將此標準化分數作為學習成效指標。

### (三) 資料整理與分析

為回應上述研究目的，本研究之資料分析策略與順序如下：針對資料進行前處理後，分別計算最高廣度指標與更新分數指標之平均數與標準差以反映全體樣本之表現。接著繪製上述指標隨著年齡的變化趨勢，並檢驗年齡效果。在信度分析中，將進行再測信度的檢驗，確認第一次施測與第二次施測在不同指標上之相關。在效度考驗部分，本研究共分為三個部分：第一、內部相關性，我們將檢驗上述作業指標間的相關情形；第二、效標關聯效度，我們將檢驗處理速度測驗表現、軌跡標示測驗表現、分類測驗表現、學習成效與更新能力之相關。所有分析均使用 JASP Version 0.19 (JASP Team, 2023) 進行分析。

## 結果

### (一) 更新測驗結果

#### 1. 最高廣度指標與更新分數指標

整體而言本研究發現之最高廣度指標之平均值為 4.70，標準差為 1.43。根據過去文獻顯示，因著更新測驗的複雜度較一般廣度作業難，因此最高廣度會低於簡單廣度作業 (Mathy et al., 2018)。無論總呈現刺激數量及要求回應之刺激數量，更新測驗之最高廣度均穩定為 4 (Jahanshahi et al.,



2009)，操弄不同刺激呈現時間與作答前提示發現，此二因素會影響整體表現之正確率，但對最高廣度亦沒有影響，平均最高廣度約為 4 (Broadway & Engle, 2010)。本研究發現複製過往研究所觀察到之更新測驗最高廣度為 4 左右之發現。然更新測驗之表現隨年齡而有所變化，因此在表 2 中同時呈現不同學齡階段之結果，在下個段落中我們將更細緻地區分年齡的效果，並進一步討論其發展趨勢。更新分數指標之平均值為 9.92，標準差為 6.36。其與最高廣度指標有一定程度的共變，因此以下將進一步針對同一最高廣度下之更新分數進行更細緻的討論。

**表 2**  
跨學齡階段之更新測驗平均表現

	1 至 6 年級	7 至 9 年級	10 至 12 年級	大學以上	全體參與者
最高廣度指標	3.79 (1.08)	4.67 (1.27)	5.12 (1.30)	5.75 (1.41)	4.70 (1.43)
更新分數指標	6.02 (3.96)	9.68 (5.74)	11.67 (6.05)	14.64 (7.12)	9.92 (6.36)

註：括號內為標準差。

## 2. 更新能力的發展軌跡

根據學齡階段將參與者分為 1 至 6 年級、7 至 9 年級、10 至 12 年級及大學以上四個年齡層，並分別對最高廣度指標與更新分數指標進行單因子變異數分析，結果發現：最高廣度指標有年齡差異， $F(3, 1,160) = 115.40, p < .001, \eta_p^2 = 0.23$ ，Tukey 事後比較發現各階段間之最高廣度指標均有差異， $ps < .001$ ，其表現依序為：大學以上階段 > 10 至 12 年級 > 7 至 9 年級 > 1 至 6 年級。此結果顯示最高廣度指標反應了隨年齡增加而提升的更新能力。

更新分數指標亦有年齡差異， $F(3, 1,160) = 108.90, p < .001, \eta_p^2 = 0.22$ ，Tukey 事後比較發現各階段之更新分數指標間均有差異， $ps < .001$ ，其表現依序為：大學以上階段 > 10 至 12 年級 > 7 至 9 年級 > 1 至 6 年級。更新分數指標得以反映在相同最高廣度下時更新能力的穩定度，因此表 3 中呈現在不同年齡區段中不同最高廣度下的更新分數。因某些情境中樣本數過少，例如並未有任何大學以上階段者之最高廣度為 1、1 至 6 年級階段之最高廣度僅有一人可達 7 或 8，因此我們將最高廣度分為低、中、高三組合併資料進行分析：低廣度組為最高廣度 1 至 3、中廣度組為最高廣度 4 至 5、高廣度組為最高廣度 6 至 8。結果發現在低廣度組中，更新分數在年齡階段中並沒有顯著差異， $F(3, 260) = 1.31, p = .270, \eta_p^2 = 0.02$ 。在中廣度組中，更新分數有年齡差異， $F(3, 568) = 12.98, p < .001, \eta_p^2 = 0.06$ ，事後比較顯示：1 至 6 年級之更新分數低於其他各組， $ps < .001$ ；7 至 9 年級之更新分數低於 10 至 12 年級及大學以上組， $ps < .05$ ；10-12 年級及大學以上組則沒有差異， $p = .994$ 。在高廣度組中，更新分數有年齡差異， $F(3, 324) = 2.66, p = .048, \eta_p^2 = 0.02$ ，事後比較顯示：1 至 6 年級之更新分數低於大學以上組， $p = .049$ ；其他各組間則沒有差異， $ps > .234$ 。此結果進一步展現了即便在最高廣度相同的情況下，我們仍能進一步透過更新分數觀察到更新能力穩定度的個別差異，而此穩定度隨著年齡增加亦有提升的趨勢，尤其在中最高廣度組的情境中尤為明顯。

**表 3**  
不同最高廣度與年齡區間下之平均更新分數

	1 至 6 年級	7 至 9 年級	10 至 12 年級	大學以上
最高廣度 = 1	1.09 (0.12) n = 2	1.67 (-) n = 1	1.92 (-) n = 1	- n = 0
最高廣度 = 2	2.74 (0.51) n = 28	2.57 (0.47) n = 9	2.92 (0.20) n = 6	- n = 0
最高廣度 = 3	3.14 (1.17) n = 121	3.44 (1.24) n = 51	3.10 (1.14) n = 28	3.47 (0.98) n = 17
低廣度組平均	3.04 (1.11) n = 151	3.28 (1.21) n = 61	3.04 (1.04) n = 35	3.47 (0.98) n = 17
最高廣度 = 4	5.88 (1.22) n = 103	6.02 (1.51) n = 92	6.01 (1.33) n = 46	5.98 (1.34) n = 19

(續下頁)

表 3  
不同最高廣度與年齡區間下之平均更新分數) (續)

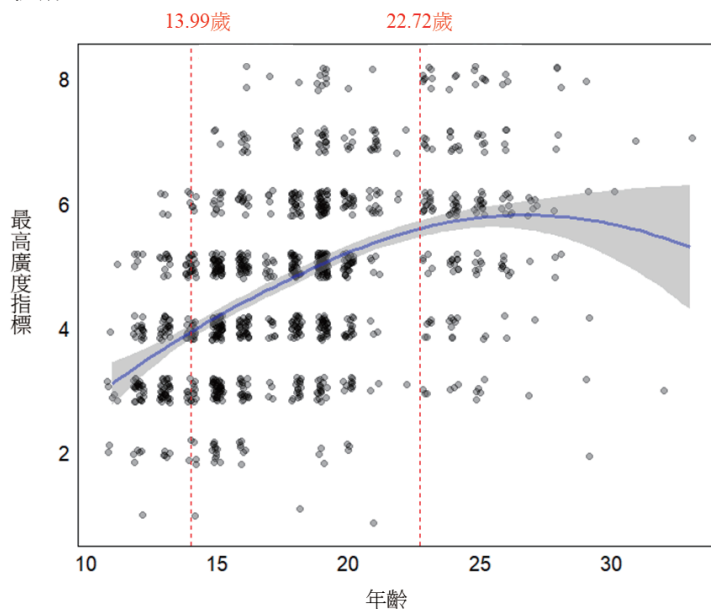
最高廣度 = 5	9.96 (1.67) n = 68	10.46 (1.68) n = 106	10.73 (1.84) n = 102	10.86 (1.92) n = 36
中廣度組平均	7.50 (2.45) n = 171	8.39 (2.74) n = 198	9.26 (2.77) n = 148	9.18 (2.91) n = 55
最高廣度 = 6	14.66 (2.92) n = 18	15.25 (1.70) n = 50	15.21 (2.06) n = 73	15.21 (2.21) n = 65
最高廣度 = 7	22.67 (-) n = 1	21.27 (2.62) n = 24	20.37 (2.50) n = 29	19.90 (3.31) n = 29
最高廣度 = 8	31.67 (-) n = 1	27.63 (2.02) n = 5	28.06 (3.10) n = 9	27.17 (2.33) n = 24
高廣度組平均	15.91 (4.96) n = 20	17.86 (4.26) n = 79	17.60 (4.45) n = 111	18.79 (5.31) n = 118

註：括號內為標準差。

為了檢驗年齡與更新能力之間的非線性關係，本研究進一步採用了廣義加性模型 (generalized additive model, GAM) (Wood, 2017)。使用了高斯分布 (Gaussian distribution) 作為誤差結構，並通過平滑函數對年齡進行處理，以捕捉可能存在的非線性趨勢。模型結果顯示，年齡對最高跨度的影響具有顯著的非線性關係 ( $p < .001$ )，模型解釋了 24% 的變異。此結果表示年齡是影響最高跨度的重要因素，且其影響隨年齡的變化而變化。此外，模型中的平滑項自由度為  $2.85 > 1$ ，進一步支持了年齡與最高跨度之間的非線性關係，如圖 4。

在年齡與更新能力存在非線性關係的前提下，為了深入探討更新能力的發展軌跡，本研究採用了分段迴歸分析 (segmented regression analysis) (Al-Azzeh et al., 2022) 以檢驗不同年齡段中更新能力的結構性變化。分段迴歸模型識別出兩個顯著的斷點 ( $R^2 = 0.24$ )，分別位於年齡約 13.99 歲和 22.72 歲處。具體而言，在 13.99 歲之前，更新能力隨年齡增加而增加，最高廣度指標以每年 0.30 單位的速度增加。然而，在 13.99 歲和 22.72 歲之間，更新能力隨年齡提升的速度減緩，斜率下降了 0.10 單位。在 22.72 歲之後，斜率顯著下降了 0.22 單位，顯示出更新能力至此後相對穩定。

圖 4  
更新能力的發展軌跡



## (二) 再測信度

50 位參與者間隔約六個月後再次進行更新測驗，結果發現第一次最高廣度指標與第二次最高廣度指標之相關為  $.50, p < .001$ ；第一次更新分數指標與第二次更新分數指標的相關為  $.54, p < .001$ 。兩指標均具有良好的再測信度。

## (三) 效度檢驗

### 1. 內部相關性

以最高廣度指標與更新分數指標的表現進行相關分析。結果發現兩指標間有顯著正相關 ( $r(1,164) = .94, p < .001$ )。此結果顯示更新測驗具有良好的內部相關性。

### 2. 效標關聯效度

(1) 處理速度測驗。以 641 位參與者之處理速度為效標發現：無論最高廣度指標或更新分數指標均與簡單反應時間及選擇反應時間達顯著相關，其相關落在  $-.22$  至  $-.29$  之間，更新能力越佳者，處理速度越快。相關數據整理如表 4。

(2) 軌跡標示測驗。以 729 位參與者之軌跡標示測驗表現為效標發現：最高廣度指標、更新分數指標均與軌跡標示測驗之錯誤次數與完成時間達顯著相關，最高廣度指標與更新分數指標與軌跡標示測驗錯誤次數之相關係數分別為  $r(727) = -.25, p < .001$ ； $r(727) = -.24, p < .001$ ；最高廣度指標與更新分數指標與軌跡標示測驗完成時間之相關係數分別為  $r(727) = -.31, p < .001$ ； $r(727) = -.31, p < .001$ 。更新能力越佳者，在軌跡標示測驗上的正確率越高，且完成時間越快。

(3) 分類測驗。以 767 位參與者之分類測驗準確率為效標發現：最高廣度指標與更新分數指標均與分類測驗正確率達顯著相關，最高廣度指標、更新分數指標與分類測驗正確率之相關係數分別為  $r(725) = .22, p < .001$ ； $r(725) = .22, p < .001$ 。更新能力越佳者，在分類測驗上的正確率越高。

(4) 學習成效。本研究收集 118 位參與者之文科及理科之學習成效並在年級區段內進行標準化，結果發現無論是最高廣度指標或更新分數指標均發現與學習成效顯著正相關，其相關落在  $.20$  至  $.26$  之間，更新能力越好，學習成效亦越好。而上述證據顯示更新測驗之效標關聯效度良好。

表 4  
更新測驗之效標關聯效度

效標	平均值 (標準差)	最高廣度指標	更新分數指標
簡單反應時間	1.06 (0.37)	-.23***	-.25***
選擇反應時間	1.13 (0.29)	-.29***	-.28***
軌跡標示錯誤次數	2.12 (5.22)	-.25***	-.24***
軌跡標示完成時間	65.64 (30.89)	-.31***	-.31***
分類測驗準確率	0.54 (0.27)	.22***	.22***
文科學習成效	-	.24**	.20*
理科學習成效	-	.26**	.25**

註：時間均以秒為單位。

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

## 討論

更新能力作為重要的執行功能之一，且與許多高階認知功能有高度關聯性，因此如何便利且準確地評估更新能力成為一個重要的研究議題。過去用以評估更新能力的作業主要分為電腦化的評估工具與紙本版的評估工具。前者多於實驗室內進行，其需要程式語言的能力方能建置作業情境，對資料處理及計分的要求也相當嚴謹，因此除研究場域外流通性偏低。紙筆版的評估工具多用於臨床

場域中以偵測腦損傷患者的認知功能表現，因主要目的為鑑別診斷，因此其難度偏低，在一般健康個體上較難精準展現更新能力之個別差異。因此本研究的首要目的在於發展一套便於施測與計分的更新能力評估作業，其可在平板電腦上施行，在施測後亦可產生對應於常模樣本的指標分數。其難度設計亦可依據參與者的能力而往上提升，以便精準地測量出更新能力的個別差異。在此目標下，本研究所設計之更新測驗在信度表現上有良好之再測信度。

在效度表現中，本研究發現更新測驗與處理速度作業表現、軌跡標示測驗表現、分類測驗表現與學習成效間都存在顯著關聯性。首先，我們複製了更新能力越佳者亦有較快的反應速度的發現 (Albinet et al., 2012; Chen & Li, 2007; Schubert et al., 2018)，此發現同時呼應處理速度理論 (Salthouse, 1996) 中的預期。軌跡標示測驗及分類測驗都是臨床情境中常選擇的測驗，前者涉及工作記憶中訊息的維持、轉換與抑制，與更新能力間存在關聯性 (Fellows et al., 2017)；後者常用的版本為 Wisconsin card sorting test，其與前額葉功能緊密相關，在認知歷程上涉及更新能力 (Doiseau & Isingrini, 2005)。本研究所設計之更新測驗與此二作業表現均有顯著相關。最後在學習成效的部分，本研究發現更新能力較佳者伴隨有更好的文科及理科學習表現，這個發現複製了更新能力與語文理解、數理能力間存在正相關的發現 (Lee et al., 2012; Palladino et al., 2001; Toll et al., 2011; Zaccoletti et al., 2020; Zheng et al., 2011)。總結而言，本研究中所設計之更新測驗具良好的效標關聯效度。

本研究第二個研究目的是要建置更新能力的本土化常模並呈現其發展軌跡。我們發現隨著年齡變化，更新能力逐漸上升。年齡與更新能力之間的關係並非線性，而是在不同的年齡段存在動態變化。特別是在青少年中期 (約 14 歲) 和成年早期 (約 23 歲) 這兩個關鍵階段，更新能力的發展軌跡有著不同的趨勢，在 14 歲以前，更新能力發展的速度較快，14 ~ 23 歲之間漸趨平緩，23 歲後則相對穩定。這個發現與 Carriedo 等人 (2016) 所觀察到的變化趨勢相仿。Luciana 等人 (2005)、Luna 等人 (2004)、Conklin 等人 (2007) 的研究中都發現工作記憶相關認知功能在青少年階段後會持續發展直到成年階段漸趨穩定。雖然根據資源有限模型 (limited resources models)，短期記憶系統的容量上限是穩定地，並不會隨著年紀增長而有所改變，抑或有實徵研究發現在青少年階段後短期記憶系統的容量就已進入高原期 (Lechuga et al., 2006)，但訊息處理的效率卻會隨著年齡增加而持續增長 (Bjorklund & Harnishfeger, 1990; Harnishfeger, 1995)，尤其當我們更新訊息時同時需要維持新訊息及抑制舊訊息的干擾時，這個效率的提升就顯得尤為重要 (Kessler & Meiran, 2008; Kessler & Oberauer, 2014)。同時回應 Cowan (2017) 的文章中針對短期記憶與工作記憶的討論，前者被定義為被動地在短時間內存放訊息的容量，而後者則須針對記憶表徵進行主動操弄。因此即便短期記憶容量隨年齡變化相對穩定，但其進行主動操弄的能力 (例如：更新能力) 隨年齡增加而逐漸進步，進而展現較佳的工作記憶作業表現，此論述與本研究觀察到更新能力發展直至成年階段才進入高原期與此論點彼此呼應。

本研究之貢獻在於設計出一個便於施測與計分的更新能力評估系統，在使用上參與者僅需依照指導語依序完成練習與正式測驗階段，在完成作業後，系統將會自動化地進行資料前處理並產生最高廣度指標與更新分數指標來反映參與者之更新能力。此外，使用者可將個體資料對應到本研究所建置之年齡常模中進行比較，報表分析系統會同時呈現指標對應之 PR 值，此報表數值將更便於解釋，因此能夠最大程度地適用在教育實務情境中。

本研究提出以下研究建議及未來研究方向：首先，本研究所設計之更新測驗適用於平板電腦作業情境中，雖可藉由 APP 模擬器在電腦或筆記型電腦上執行，但仍建議改用平板電腦作為載具進行施測最為直觀。第二，本更新測驗設計有練習階段以確認參與者是否能夠正確理解作業要求，需正確完成練習階段測試後方能開始正式施測，因此遇到無法通過練習階段的參與者時，主試者可從旁協助說明作業要求以利參與者能夠正確執行作業。第三，本研究以健康樣本為主，其年齡區段從學齡階段到成年期，因此在此階段中之樣本有完整的常模資料以供對應，然過去研究中已然發現幼兒階段之工作記憶發展的重要性 (周姍姍、王馨敏, 2021)，但目前本常模中尚缺乏幼兒及高齡階段之常模資料，此部分須待後續研究進行增補，以提升常模的普及性。第四，目前所選用之樣本均為健康樣本，並未收集有認知障礙者之資料，臨床或實務工作者可將個案之資料與常模對照以進行篩選，但尚未有足夠的特殊樣本資料以供分析，此部分亦期待後續研究能透過本研究所發展更新測驗進一步擴展對特殊個案之了解。

## 參考文獻

- 周姍姍、王馨敏（2021）：〈幼兒視覺空間工作記憶與翻牌遊戲之關聯性研究〉。《教育科學研究期刊》，66（3），191–212。[Chou, S.-S., & Wang, S.-M. (2021). Relationship between visuospatial working memory and concentration game performance in young children. *Journal of Research in Education Sciences*, 66(3), 191–212.]  
[https://doi.org/10.6209/JORIES.202109\\_66\(3\).0006](https://doi.org/10.6209/JORIES.202109_66(3).0006)
- 郭郡羽、陳珮臻、陳志洪、陳學志（2024）：〈平板處理速度作業發展暨本土常模建置：8–27 歲的發展軌跡〉。《測驗學刊》，71（4），479–510。<https://www.airitilibrary.com/Article/Detail?DocID=16094905-N202412250005-00003>[Kuo, C.-Y., Chen, P.-Z., Chen, Z.-H., & Chen, H.-C. (2024). The measurement and developmental trajectories of processing speed from 7–22 years old. *Psychological Testing*, 71(4), 479–510.]
- 黃詩媛、黃博聖、陳志洪、陳珮臻、陳學志（2024）：〈「軌跡標示測驗平板版」的信、效度研究暨本土常模建置〉。《教育心理學報》，56（2），231–259。  
[https://doi.org/10.6251/BEP.202412\\_56\(2\).0001](https://doi.org/10.6251/BEP.202412_56(2).0001) [Huang, S.-Y., Huang, P.-S., Chen, Z.-H., Chen, P.-Z., & Chen, H.-C. (2024). Reliability and validity of tablet-based trail making test and establishment of normative data. *Bulletin of Educational Psychology*. 56(2), 231–259.]
- Al-Azzeh, J., Mesleh, A., Zaliskyi, M., Odarchenko, R., & Kuzmin, V. (2022). A method of accuracy increment using segmented regression. *Algorithms*, 15(10), Article 378.  
<https://doi.org/10.3390/a15100378>
- Albinet, C. T., Boucard, G., Bouquet, C. A., & Audiffren, M. (2012). Processing speed and executive functions in cognitive aging: How to disentangle their mutual relationship? *Brain and Cognition*, 79(1), 1–11. <https://doi.org/10.1016/j.bandc.2012.02.001>
- Barrouillet, P. (1996). Transitive inferences from set-inclusion relations and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1408–1422.  
<https://doi.org/10.1037/0278-7393.22.6.1408>
- Barrouillet, P., & Lecas, J. (1999). Mental models in conditional reasoning and working memory. *Thinking & Reasoning*, 5(4), 289–302. <https://doi.org/10.1080/135467899393940>
- Basak, C., & O’Connell, M. A. (2016). To switch or not to switch: Role of cognitive control in working memory training in older adults. *Frontiers in Psychology*, 7, Article 230.  
<https://doi.org/10.3389/fpsyg.2016.00230>
- Basak, C., & Zelinski, E. M. (2013). A hierarchical model of working memory and its change in healthy older adults. In T. Packiam Alloway & R. G. Alloway (Eds.), *Working memory: The connected intelligence* (pp. 83–106). Psychology Press.
- Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review*, 10(1), 48–71.  
[http://doi.org/10.1016/0273-2297\(90\)90004-N](http://doi.org/10.1016/0273-2297(90)90004-N)

- Braver, T., Satpute, A., Rush, B., Racine, C., & Barch, D. (2005). Context processing and context maintenance in healthy aging and early stage dementia of the Alzheimer's type. *Psychology and Aging, 20*(1), 33–46. <https://doi.org/10.1037/0882-7974.20.1.33>
- Broadway, J., & Engle, R. (2010). Validating running memory span: Measurement of working memory capacity and links with fluid intelligence. *Behavior Research Methods, 42*(2), 563–570. <https://doi.org/10.3758/BRM.42.2.563>
- Carriedo, N., Corral, A., Montoro, P., Herrero, L., & Rucián, M. (2016). Development of the updating executive function: From 7-year-olds to young adults. *Developmental Psychology, 52*(4), 666–678. <https://doi.org/10.1037/dev0000091>
- Chen, T., & Li, D. (2007). The roles of working memory updating and processing speed in mediating age-related differences in fluid intelligence. *Aging, Neuropsychology, and Cognition, 14*(6), 631–646. <https://doi.org/10.1080/13825580600987660>
- Cheng, C., & Kibbe, M. M. (2022). Development of updating in working memory in 4–7-year-old children. *Developmental Psychology, 58*(5), 902–912. <https://doi.org/10.1037/dev0001337>
- Conklin, H. M., Luciana, M., Hooper, C. J., & Yarger, R. S. (2007). Working memory performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology, 31*(1), 103–128. [http://doi.org/10.1207/s15326942dn3101\\_6](http://doi.org/10.1207/s15326942dn3101_6)
- Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review, 24*(4), 1158–1170. <https://doi.org/10.3758/s13423-016-1191-6>
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450–466. [https://doi.org/10.1016/S0022-5371\(80\)90312-6](https://doi.org/10.1016/S0022-5371(80)90312-6)
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422–433. <https://doi.org/10.3758/BF03214546>
- De Beni, R., & Palladino, P. (2004). Decline in working memory updating through ageing: Intrusion error analyses. *Memory, 12*(1), 75–89. <https://doi.org/10.1080/09658210244000568>
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t15082-000>
- Doiseau, F., & Isingrini, M. (2005). Updating information in verbal working memory and executive functioning. *Psychological Reports, 96*(1), 67–76. <https://doi.org/10.2466/pr0.96.1.67-76>
- Ecker, U. K., Lewandowsky, S., Oberauer, K., & Chee, A. E. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 170–189. <https://doi.org/10.1037/a0017891>
- Ecker, U. K., Oberauer, K., & Lewandowsky, S. (2014). Working memory updating involves item-specific removal. *Journal of Memory and Language, 74*, 1–15. <https://doi.org/10.1016/j.jml.2014.03.006>
- Fellman, D., Soveri, A., Waris, O., & Laine, M. (2017). Training of verbal working memory at sentence

- level fails to show transfer. *Frontiers in Communication*, 2, Article 14.  
<https://doi.org/10.3389/fcomm.2017.00014>
- Fellows, R., Dahmen, J., Cook, D., & Schmitter-Edgecombe, M. (2017). Multicomponent analysis of a digital Trail Making Test. *The Clinical Neuropsychologist*, 31(1), 154–167.  
<https://doi.org/10.1080/13854046.2016.1238510>
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, 17(2), 172–179.  
<https://doi.org/10.1111/j.1467-9280.2006.01681.x>
- Gao, J., Li, G., Yang, Z., Li, F., Wang, T., & Wen, S. (2023). The effect of working memory updating training on the Chinese writing ability of primary school students. *Frontiers in Psychology*, 14, Article 1163132. <https://doi.org/10.3389/fpsyg.2023.1163132>
- Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, 18(1), 1–16. <https://doi.org/10.1002/acp.934>
- Gijsselaers, H. J., Meijs, C., Neroni, J., Kirschner, P. A., & de Groot, R. H. (2017). Updating and not shifting predicts learning performance in young and middle aged adults. *Mind, Brain, and Education*, 11(4), 190–200. <https://doi.org/10.1111/mbe.12147>
- Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404–411. <https://doi.org/10.1037/h0059831>
- Harnishfeger, K. K. (1995). The development of cognitive inhibition. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175–204). Academic Press.  
<http://doi.org/10.1016/B978-012208930-5/50007-6>
- Huttenlocher, P. R. (1990). Morphometric study of human cerebral cortex development. *Neuropsychologia*, 28(6), 517–527. [http://doi.org/10.1016/0028-3932\(90\)90031-I](http://doi.org/10.1016/0028-3932(90)90031-I)
- Iglesias-Sarmiento, V., Carriedo-López, N., & Rodríguez-Rodríguez, J. (2015). Updating executive function and performance in reading comprehension and problem solving. *Anales de Psicología*, 31(1), 298–309. <https://doi.org/10.6018/analesps.31.1.158111>
- Jahanshahi, M., Saleem, T., Ho, A., Fuller, R., & Dirnberger, G. (2009). A preliminary investigation of the running digit span as a test of working memory. *Behavioural Neurology*, 20(1–2), 17–25.  
<https://doi.org/10.3233/BEN-2008-0212>
- JASP Team. (2023). *JASP (Version 0.19)* [Computer software]. JASP. <https://jasp-stats.org/>
- Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, 36(2), 138–202. <https://doi.org/10.1006/cogp.1998.0684>
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122–149.  
<https://doi.org/10.1037/0033-295X.99.1.122>

- Kessler, Y., & Meiran, N. (2008). Two dissociable updating processes in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1339–1348.  
<http://doi.org/10.1037/a0013078>
- Kessler, Y., & Oberauer, K. (2014). Working memory updating latency reflects the cost of switching between maintenance and updating modes of operation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 738–754. <http://doi.org/10.1037/a0035545>
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358. <https://doi.org/10.1037/h0043688>
- Lange, E. B., & Verhaeghen, P. (2009). No age differences in complex memory search: Older adults search as efficiently as younger adults. *Psychology and Aging*, *24*(1), 105–115.  
<https://doi.org/10.1037/a0013751>
- Lechuga, M. T., Moreno, V., Pelegrina, S., Gómez-Ariza, C. J., & Bajo, M. T. (2006). Age differences in memory control: Evidence from updating and retrieval-practice tasks. *Acta Psychologica*, *123*(3), 279–298. <http://doi.org/10.1016/j.actpsy.2006.01.006>
- Lechuga, M. T., Pelegrina, S., Pelaez, J. L., Martín-Puga, M. E., & Justicia, M. J. (2016). Working memory updating as a predictor of academic attainment. *Educational Psychology*, *36*(4), 675–690.  
<https://doi.org/10.1080/01443410.2014.950193>
- Lee, K., & Bull, R. (2016). Developmental changes in working memory, updating, and math achievement. *Journal of Educational Psychology*, *108*(6), 869–882. <https://doi.org/10.1037/edu0000090>
- Lee, K., Ng, S. F., Pe, M. L., Ang, S. Y., Hasshim, M. N. A. M., & Bull, R. (2012). The cognitive underpinnings of emerging mathematical skills: Executive functioning, patterns, numeracy, and arithmetic. *British Journal of Educational Psychology*, *82*(1), 82–99.  
<https://doi.org/10.1111/j.2044-8279.2010.02016.x>
- Linares, R., Bajo, M. T., & Pelegrina, S. (2016). Age-related differences in working memory updating components. *Journal of Experimental Child Psychology*, *147*, 39–52.  
<https://doi.org/10.1016/j.jecp.2016.02.009>
- Luce, R. D. (1991). *Response times: Their role in inferring elementary mental organization*. Oxford University Press.
- Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, *76*(3), 697–712.  
<http://doi.org/10.1111/j.1467-8624.2005.00872.x>
- Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*(5), 1357–1372.  
<http://doi.org/10.1111/j.1467-8624.2004.00745.x>
- Mathy, F., Chekaf, M., & Cowan, N. (2018). Simple and complex working memory tasks allow similar benefits of information compression. *Journal of Cognition*, *1*(1), Article 31.  
<https://doi.org/10.5334/joc.31>



- 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>
- Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*(2), 111–121. <http://doi.org/10.1111/j.2044-8295.1990.tb02349.x>
- Nagahama, Y., Okina, T., Suzuki, N., Nabatame, H., & Matsuda, M. (2005). The cerebral correlates of different types of perseveration in the Wisconsin Card Sorting Test. *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(2), 169–175. <https://doi.org/10.1136/jnnp.2004.039818>
- Oberauer, K. (2001). Removing irrelevant information from working memory: A cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(4), 948–957. <https://doi.org/10.1037/0278-7393.27.4.948>
- Oosterhuis, H. E., van der Ark, L. A., & Sijtsma, K. (2016). Sample size requirements for traditional and regression-based norms. *Assessment*, *23*(2), 191–202. <https://doi.org/10.1177/1073191115580638>
- Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition*, *29*(2), 344–354. <https://doi.org/10.3758/BF03194929>
- Pelegrina, S., Molina, R., Rodríguez-Martínez, E., Linares, R., & Gómez, C. (2020). Age-related changes in selection, recognition, updating and maintenance information in WM. An ERP study in children and adolescents. *Biological Psychology*, *157*, Article 107977.  
<https://doi.org/10.1016/j.biopsycho.2020.107977>
- Podell, J., Sambataro, F., Murty, V., Emery, M., Tong, Y., Das, S., Goldberg, T., Weinberger, D., & Mattay, V. (2012). Neurophysiological correlates of age-related changes in working memory updating. *NeuroImage*, *62*(3), 2151–2160. <https://doi.org/10.1016/j.neuroimage.2012.05.066>
- Qin, S., & Basak, C. (2020). Age-related differences in brain activation during working memory updating: An fMRI study. *Neuropsychologia*, *138*, Article 107335.  
<https://doi.org/10.1016/j.neuropsychologia.2020.107335>
- Reitan, R. M. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology*, *19*(5), 393–394. <https://doi.org/10.1037/h0044509>
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. <https://doi.org/10.1037/0033-295X.103.3.403>
- Santarnecchi, E., Momi, D., Mencarelli, L., Plessow, F., Saxena, S., Rossi, S., Rossi, A., Mathan, S., & Pascual-Leone, A. (2021). Overlapping and dissociable brain activations for fluid intelligence and executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(2), 327–346.  
<https://doi.org/10.3758/s13415-021-00870-4>
- Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1089–1096. <https://doi.org/10.1037/a0015730>

- Schubert, A., Nunez, M., Hagemann, D., & Vandekerckhove, J. (2018). Individual differences in cortical processing speed predict cognitive abilities: A model-based cognitive neuroscience account. *Computational Brain & Behavior*, 2(2), 64–84. <https://doi.org/10.1007/s42113-018-0021-5>
- 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>
- Swanson, H. L., & Berninger, V. (1995). The role of working memory in skilled and less skilled readers' comprehension. *Intelligence*, 21(1), 83–108. [https://doi.org/10.1016/0160-2896\(95\)90040-3](https://doi.org/10.1016/0160-2896(95)90040-3)
- Toll, S. W. M., Van der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2011). Executive functions as predictors of math learning disabilities. *Journal of Learning Disabilities*, 44(6), 521–532. <https://doi.org/10.1177/0022219410387302>
- Van der Linden, M., Brédart, S., & Beerten, A. (1994). Age related differences in updating working memory. *British Journal of Psychology*, 85(1), 145–152. <https://doi.org/10.1111/j.2044-8295.1994.tb02514.x>
- Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A*, 47(3), 631–650. <https://doi.org/10.1080/14640749408401131>
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 32(6), 1436–1451. <https://doi.org/10.1037/0096-1523.32.6.1436>
- Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. Psychological Corporation.
- Wood, S. N. (2017). *Generalized additive models (2nd ed.): An introduction with R*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781315370279>
- Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors*, 5(1), 7–17. <https://doi.org/10.1177/001872086300500102>
- Yntema, D. B., & Mueser, G. E. (1962). Keeping track of variables that have few or many states. *Journal of Experimental Psychology*, 63(4), 391–395. <https://doi.org/10.1037/h0045706>
- Zaccoletti, S., Altoé, G., & Mason, L. (2020). The interplay of reading-related emotions and updating in reading comprehension performance. *The British Journal of Educational Psychology*, 90(3), 663–682. <https://doi.org/10.1111/bjep.12324>
- Zhang, H., Chang, L., Chen, X., Ma, L., & Zhou, R. (2018). Working memory updating training improves mathematics performance in middle school students with learning difficulties. *Frontiers in Human Neuroscience*, 12, Article 154. <https://doi.org/10.3389/fnhum.2018.00154>
- Zheng, X., Swanson, H. L., & Marcoulides, G. A. (2011). Working memory components as predictors of children's mathematical word problem solving. *Journal of Experimental Child Psychology*, 110(4), 481–498. <https://doi.org/10.1016/j.jecp.2011.06.001>

收稿日期：2024年08月21日  
一稿修訂日期：2024年08月29日  
二稿修訂日期：2024年10月04日  
三稿修訂日期：2024年10月31日  
接受刊登日期：2024年11月21日

Bulletin of Educational Psychology, 2025, 56(3), 711–732  
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

## Development of the Updating Ability Task: A Trajectory from 7-Year-Olds to Young Adults

Chun-Yu Kuo<sup>1</sup>, Pei-Zhen Chen<sup>2</sup>, Zhi-Hong Chen<sup>3</sup>, and Hsueh-Chih  
Chen<sup>2, 4, 5, 6</sup>

Updating, a key executive function, involves continuously adjusting and refreshing information within working memory to adapt to changing environments. This cognitive function is crucial for ensuring that relevant information is maintained while outdated or irrelevant information is discarded. The present study aimed to develop a standardized assessment tool for accurately measuring updating ability and to explore the developmental trajectory of this cognitive function across different age groups.

Initially, a review of existing methods for measuring updating ability was conducted, revealing several limitations, particularly in distinguishing updating from other executive functions such as inhibition and shifting. Most existing tests, including the n-back task and complex working memory span tasks, tend to confound updating with other working memory components, making it difficult to isolate the specific contribution of updating ability. To address these limitations, this study developed a tablet-based updating test using the running memory task paradigm. This test calculates both maximum span and updating scores to comprehensively reflect an individual's updating capacity. The maximum span index represents the highest level of difficulty at which participants can accurately remember the last items presented in a sequence, while the updating score captures the consistency of performance at different levels of difficulty, reflecting stability in updating ability.

The validity of the newly developed updating test was assessed through various criterion-related validity analyses. Significant correlations were observed between the updating test scores and a range of cognitive measures, including processing speed, the Trail Making Test, classification tasks, and academic learning outcomes. These findings indicate that the updating test is a valid tool for assessing updating ability, and that updating ability is closely related to general cognitive processes such as speed, flexibility, and reasoning. Moreover, the test-retest reliability of the updating test was examined over a six-month interval, with results showing strong stability in both the maximum span and updating scores. This suggests that the updating test can reliably assess an individual's updating ability over time.

The developmental trajectory of updating ability was further examined in this study, which involved participants ranging from 7 years old to early adulthood. The results indicated that updating ability improves steadily from childhood through adolescence and reaches a plateau during adulthood. Specifically, the ability develops rapidly during childhood, with significant

<sup>1</sup>Department of Adult and Continuing Education, National Taiwan Normal University

<sup>2</sup>Department of Educational Psychology and Counseling, National Taiwan Normal University

<sup>3</sup>Department of Graduate Institute of Information and Computer Education, National Taiwan Normal University

<sup>4</sup>Institute for Research Excellence in Learning Sciences, National Taiwan Normal University

<sup>5</sup>Chinese Language and Technology Center, National Taiwan Normal University

<sup>6</sup>Social Emotional Education and 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](mailto:hcjyh@ntnu.edu.tw)

gains observed from ages 7 to 15, and continues to improve, albeit at a slower rate, until early adulthood (ages 20–25), after which it stabilizes. This developmental pattern is consistent with the synaptic pruning observed in the prefrontal cortex, which plays a key role in supporting working memory processes. The findings align with previous research suggesting that the maturation of the prefrontal cortex is associated with improvements in executive functions, including updating.

The updating test is based on the running memory task, which requires participants to remember and report the last few items presented in a sequence, such as numbers or letters. The test starts with a practice phase to ensure participants understand the task requirements and then consists of increasing levels of difficulty that adjust based on the participant's performance. In each trial, participants must remember the last  $n$  items presented, where  $n$  increases with each correct response. This adaptive approach ensures that the test is challenging yet appropriate for each individual's ability level. The two primary indices obtained from the test are the maximum span, which indicates the highest level at which participants can consistently perform well, and the updating score, which captures the stability and reliability of their performance across different difficulty levels.

The validity of the updating test was established through its correlations with various cognitive measures. Specifically, the study employed multiple validity assessments, including processing speed measures (simple and choice reaction times), the Trail Making Test, the Wisconsin Card Sorting Test, and academic performance in linguistic and mathematical subjects. The results demonstrated significant associations between updating test performance and these criterion tasks, suggesting that individuals with better updating ability tend to have faster processing speeds, greater cognitive flexibility, and higher academic achievement. Processing speed, as assessed by reaction time tasks, was found to be significantly correlated with both maximum span and updating scores, indicating that faster cognitive processing is linked to more efficient updating. Similarly, performance on the Trail Making Test, which measures cognitive flexibility and the ability to shift between different sets of stimuli, was significantly associated with updating scores, further supporting the role of updating in complex cognitive tasks.

The study utilized a stratified sampling approach across northern, central, southern, and eastern regions of Taiwan, recruiting participants from elementary school to university levels. A total of 1,164 participants provided valid data for establishing local norms for the updating test. The final sample was well-distributed across age groups, with participants ranging from 11 to 33 years of age. The local norms were developed to facilitate comparisons between individuals and the general population, allowing for a more nuanced interpretation of updating ability in different age groups. The study also developed an automated scoring module to enhance usability, allowing users to obtain scores that are directly aligned with these norms. This advancement makes the updating test suitable for both research purposes and practical applications in educational and clinical settings.

To further investigate the developmental trajectory of updating ability, participants were divided into four age groups: Elementary school (grades 1–6), middle school (grades 7–9), high school (grades 10–12), and college level and above. A one-way analysis of variance (ANOVA) revealed significant differences in updating performance across these age groups, indicating a steady improvement in updating ability throughout childhood and adolescence. The ANOVA results were further supported by post-hoc analyses, which showed that each successive age group outperformed the previous one in both maximum span and updating scores, highlighting the gradual maturation of updating ability during these critical developmental periods.

In addition to the ANOVA, the study employed segmented regression and generalized additive models (GAM) to explore potential nonlinear relationships between age and updating ability. The segmented regression analysis identified two significant breakpoints at approximately ages 14 and 23, indicating distinct phases in the developmental progression of updating ability. Before age 14, updating ability increased rapidly, suggesting significant cognitive development during early adolescence. Between ages 14 and 23, the rate of improvement slowed, reflecting the gradual maturation of the prefrontal cortex and the stabilization of executive functions. After age 23, updating ability reached a plateau, consistent with the idea that the cognitive systems supporting working memory and executive function become fully developed by early adulthood. The GAM analysis further confirmed the nonlinear trajectory, showing a significant age effect that explained 24% of the variance in updating performance.

The test-retest reliability of the updating test was supported by a follow-up assessment conducted six months after the initial test. Fifty participants completed the updating test twice, and the results showed a significant correlation between the two testing occasions for both the maximum span and updating scores. These findings indicate that the updating test provides consistent and reliable measures of updating ability over time, making it a valuable tool for longitudinal studies of cognitive development.

In summary, the present study contributes to the field by developing a reliable and valid tool for evaluating updating ability and by providing a detailed understanding of the developmental trajectory of this critical executive function. The tablet-based updating test addresses previous limitations by isolating updating from other working memory components, offering practical advantages for use in both research and applied contexts. The establishment of local norms further enhances the applicability of the test, allowing practitioners to benchmark individual performance against the general population. This tool holds promise for use in educational and clinical settings, providing valuable insights into cognitive functioning and aiding in the identification of individuals who may benefit from targeted interventions to enhance executive function skills.

The study's findings also underscore the importance of accurately assessing updating ability, given its strong association with various higher-order cognitive abilities, including fluid intelligence, reading comprehension, and mathematical problem-solving. By providing a comprehensive assessment of updating ability and its developmental trajectory, this research offers a foundation for future studies aimed at understanding the mechanisms underlying cognitive development and the factors that contribute to individual differences in executive function performance. The updating test developed in this study can serve as a useful tool for educators, clinicians, and researchers interested in exploring the role of updating in cognitive development, academic achievement, and everyday functioning.

*Keywords:* tablet computer, updating, executive function, developmental trajectory