

# 以電腦化複雜廣度作業測量兒童工作記憶

陳湘淳<sup>1</sup>、唐藝<sup>2</sup>

本研究從記憶理論和認知發展的角度，探討如何以電腦化廣度作業有效評估兒童的工作記憶發展。據此，本研究檢驗改編自 Camos 與 Barrouillet（2011）的動物顏色廣度作業的信度與效度，同時，藉由此作業了解 4 至 7 歲兒童工作記憶的發展情形。臺灣北部典型發展兒童共 333 名（男生 167 名）參與本研究，依年齡分為四歲、五歲、六歲和七歲四組。在此作業中，兒童需在電腦控制的時間限制下，交替完成視覺圖像呈現的儲存作業（動物記憶）與處理作業（顏色叫名），最後依序回憶記憶項目。信度分析顯示此作業有理想的內在信度與折半信度。在效度方面，幅合效度分析顯示此作業與數字順背廣度作業、數字逆背廣度作業、魏氏幼兒智力測驗工作記憶分測驗和托尼非語文智力測驗有顯著正相關；試題分析以及 Rasch 模式分析指出，此作業的難度隨廣度增加而增加，試題鑑別度適中且與 Rasch 模式有良好的適配度。此外，單因子變異數分析發現工作記憶廣度隨年齡增長而有顯著增加，顯示此作業能夠反映成熟造成的發展差異。綜合上述結果本研究建議，動物顏色廣度作業可有效評估幼小兒童的工作記憶發展，而電腦化複雜廣度作業準確控制時間與測驗程序，可增進兒童工作記憶測量的信度與效度。

**關鍵詞：**工作記憶容量測量、兒童工作記憶、電腦化記憶測驗、複雜廣度作業

---

<sup>1</sup> 國立清華大學幼兒教育學系

<sup>2</sup> 國立清華大學教育心理與諮商學系

**通訊作者：**唐藝，國立清華大學教育心理與諮商學系，[tangyi@gapp.nthu.edu.tw](mailto:tangyi@gapp.nthu.edu.tw)。  
本研究經費由第一作者之國科會專題研究計畫補助（計畫編號：MOST106-2410-H007-001、MOST108-2410-H007-048）。本研究感謝鄭雨君、翁怡婷、林珮雯、馮翠文提供部分碩士論文收集之幼兒資料。

工作記憶 (working memory) 是一個容量有限的認知系統，用來暫時儲存 (storage) 與處理 (processing) 和手邊正在進行的活動有關的訊息 (Baddeley, 2000)。工作記憶不僅是兒童與外在互動和學習新知識的重要認知能力，也與兒童的日常生活息息相關。例如，記住什麼時間應該要做什麼事 (前瞻性記憶 prospective memory)，便需要工作記憶的支援 (Cheie et al., 2017)。

越來越多研究人員和教育人員意識到，許多我們想要理解、測量和操控的行為和認知能力，都與工作記憶的運作有關 (周姍姍、王馨敏, 2021; Cowan, 2014; Slattery et al., 2021)。兒童工作記憶容量的個別差異，可以有效地預測他們的學業表現 (簡馨瑩, 2020; Jaroslawska et al., 2016)。智能較高的兒童有較高的工作記憶容量 (Kornmann et al., 2015; Swanson, 2006)。相反地，許多神經性發展障礙，像是閱讀障礙 (Jeffries & Everatt, 2004)、特定語言障礙 (Montgomery et al., 2010) 和注意力缺陷過動障礙 (Alloway, 2011; Holmes et al., 2014) 的兒童有工作記憶缺失 (deficiency)。英國一項大規模研究甚至發現，約有 10% 至 15% 的學齡兒童顯現工作記憶低弱 (Alloway et al., 2009)。正因為如此，無論在學術上或實務上都有更多的需求，想要了解如何評估兒童的工作記憶容量。

## 文獻探討

### (一) 工作記憶發展理論

研究發現，兒童在同時處理另一項認知作業時，可以保留的記憶項目在 4 至 16 歲之間逐漸增加，並和短期記憶有不同的發展軌跡 (Gathercole, 1999)。工作記憶比短期記憶經歷更長的發展期，與大腦額葉的成熟有關 (Cowan, 1997)。許多心理學與教育學研究致力於探討兒童工作記憶的發展，想要釐清哪些因素促進工作記憶容量的增加，以及這可能對不同年齡兒童的認知發展與生活適應產生什麼影響。目前有三種理論取向說明工作記憶的運作機制，以及如何隨年齡而發展。

第一種理論取向，探討工作記憶如何利用有限的認知資源來保留及處理訊息。Baddeley (2000) 的多成分模型 (multi-component model)，包含一個有限容量的中央執行器 (central executive)，及三個輔助系統：語音迴圈 (phonological loop)、視覺空間畫版 (visual-spatial sketchpad) 及事件緩衝器 (episodic buffer)。中央執行器是整個模式的監督系統，負責協調子系統、維持及轉換注意力、活化長期記憶中的表徵及高階的認知運作 (如：理解、推理、計畫和問題解決)。語音迴圈和視覺空間畫版兩個子系統，負責處理及暫存語言或視覺空間之特定領域的訊息，而事件緩衝器則負責整合來自於不同子系統間，以及子系統與長期記憶間的訊息 (Baddeley, 2000)。這個理論模式已廣泛應用於兒童研究 (Gathercole, 1998; Henry & Millar, 1993)。實證結果顯示，兒童大約在 4 歲左右已發展語文和視覺空間兩個特定領域 (domain-specific) 的子系統，和一個一般領域 (domain-general) 的中央執行處理 (Alloway et al., 2006)。

第二種理論取向，強調注意力的控制能力與記憶表徵的活化程度。其中，Cowan (1988) 提出注意力鑲嵌模式 (embedded-processes model) 認為，工作記憶是一種認知處理，能將訊息保留在特別容易接觸的狀態，以完成任何需要心智運作的工作。長期記憶、目前被活化的長期記憶知識，以及被活化的長期記憶知識中注意力的焦點，這三個成分形成一個鑲嵌的處理模式。而工作記憶的運作，就是中央執行器對注意力焦點的控制 (Cowan, 1999)。根據注意力鑲嵌模式，工作記憶的發展包含注意力容量的增加和注意力控制性的增加 (Cowan et al., 2010)。而 Engle 等人 (1999) 提出的容量個別差異模式 (individual differences model) 則是主張，工作記憶是短期儲存能力加上面對干擾時維持注意的控制能力。工作記憶容量是高階認知系統運作時，能暫時將訊息保留在活化狀態的數量，這樣的容量限制是個人穩定的特質，不受個人知識結構改變的影響 (Engle et al., 1992)。

第三種理論取向，強調有限的工作記憶資源如何彈性的分配至處理和儲存的運作，並認為兒童工作記憶發展受限於認知處理的效能不佳。其中，資源分享模式 (resource sharing model) (Case et al., 1982) 從認知資源有限的角度，主張處理會對儲存造成干擾，而記憶衰退模式 (memory decay model) (Towse & Hitch, 1995) 主張儲存的記憶訊息會隨著處理的時間長短而衰退。近幾年較被接受的是時間為基的資源分享模式 (time-based resource sharing model, 簡稱 TBRS 模式) (Barrouillet et al., 2007)，綜合考慮認知資源與時間的因素，對記憶訊息被再恢復 (refreshing) 的影響，並認

為工作記憶發展的改變，是來自記憶痕跡衰退的速度變慢、再恢復記憶痕跡的速度變快，以及抵抗干擾能力增加。

## （二）複雜廣度作業

雖然不同工作記憶理論對工作記憶發展的解釋不同，但大部分理論認為，工作記憶包含儲存和處理兩種認知運作（Baddeley, 2000; Engle & Kane, 2004; Gavens & Barrouillet, 2004）。因此，測量工作記憶的廣度作業通常要求參與者暫時記住記憶項目，並同時完成另一項處理作業（或對記憶項目執行某種形式的處理）。這類作業與只要求參與者將記憶項目保留一段時間後依序回憶的簡單廣度作業（simple span tasks）不同，特別稱為複雜廣度作業（complex span tasks）（Redick et al., 2012）。簡單廣度作業，例如，非字廣度作業（children's test of non-word repetition）（Gathercole et al., 1994），有助我們了解兒童工作記憶系統中語音迴圈的保留方式及限制。複雜廣度作業，例如，運作廣度作業（operation span）（Turner & Engle, 1989），因涉及儲存和處理交替進行的動態歷程，反映的是工作記憶系統內中央執行器的容量限制。使用複雜廣度作業所測得的工作記憶容量，能穩定的預測許多高階認知能力，如：流體智力（Unsworth et al., 2014）、閱讀理解與閱讀速度（Könen & Karbach, 2020）、推理能力（Demetriou et al., 2022），也能夠預測學齡階段的學業成績（Peng & Kievit, 2020）。

## （三）兒童工作記憶複雜廣度作業

以複雜廣度作業評估兒童的工作記憶，需要簡單、清楚、容易理解的指導語，更重要的是，作業的刺激、內容、呈現形式與反應方式，也要符合兒童的能力。對於語文閱讀能力尚未成熟的幼小兒童而言，視覺圖像是經常使用的作業刺激與呈現方式。工作記憶發展研究，常以兒童熟悉的記憶項目作為儲存作業，簡單的判斷任務作為處理作業，要求兒童以口語念出或手指點選的方式來回憶。例如，前文提到第三種理論取向中的 TBRs 模式學者，Barrouillet 等人（2009）以及 Camos 與 Barrouillet（2011）以動物名稱記憶（儲存作業）和顏色判斷（處理作業）交替呈現的複雜廣度作業，探討 6 歲和 7 歲兒童的工作記憶發展機制；陳湘淳（2012）以故事情境引導 5 歲和 6 歲幼兒進行數字記憶（儲存作業）和顏色判斷（處理作業）交替呈現的複雜廣度作業，探討記憶策略對工作記憶容量發展的影響。

由於處理作業的目的是在干擾或中斷兒童對記憶項目的保留（Cowan et al., 2005），處理作業的複雜度也是影響工作記憶容量評估準確性的因素之一。TBRs 模式學者 Barrouillet 等人（2004）發現即使是簡單的處理活動，只要有時間限制，已能佔住足夠的注意力來干擾記憶項目的保留，使成人的記憶廣度下降至 3 左右。其他的研究也發現，和呈現時間有限（電腦調控，computer-paced）的處理作業相比，使用不限呈現時間（自己調控，self-paced；實驗者調控，experimenter-paced）的處理作業，所評估得到的工作記憶容量與高階認知能力的相關程度較低（Bailey, 2012; St Clair-Thompson, 2007）。而且，使用簡單但有時間限制的處理作業，較能降低作業複雜性所造成的個別差異，以及參與者策略性的延長處理作業時間，對工作記憶容量評估的影響（Lépine et al., 2005）。

另外，在測驗程序上，一方面要考量工作記憶容量非常小的兒童，避免使能力較弱的兒童在施測過程中感到無能為力，另一方面也要考量工作記憶容量高的兒童，使能力較好的兒童能夠有最佳表現。這種情況下，通常會使用逐漸提升作業難度（要記住的項目逐漸增多），並在兒童未通過某個難度時停止測試的測驗程序（Gonthier et al., 2018; Pickering, 2006）。目前，大部分用於測量兒童工作記憶的測驗都使用這樣的程序，例如，國內的魏氏兒童智力測驗（第四版中文版）的數字廣度分測驗（digit span subtests），以及幾個國外的工作記憶評量系統，例如：AWMA（Alloway, 2007）和 WMTB-C（Gathercole & Pickering, 2001），還有許多研究者自編的工作記憶作業（例如：Barrouillet et al., 2009; Bayliss et al., 2005）。

然而，上述以未通過特定級別即停止測試為程序的記憶廣度測驗，經常使用全有全無評分（all-or-nothing scoring），例如，所有的記憶項目需按照正確順序回憶才算對（Conway et al., 2005）。

如此一來，只要兒童有一個記憶項目沒有回憶出來或順序錯誤（可能是暫時恍神），該題回憶就會被視為失敗，也可能因此導致測驗終止，影響兒童工作記憶容量的評估結果。而另一種測驗程序，是每個兒童皆完成不同難度的所有測驗試題。在成人的工作記憶測驗中經常使用這種程序，通常是將不同難度的試題打散後隨機呈現（Conway et al., 2005; Unsworth et al., 2009），但這樣容易讓年齡較小或是低工作記憶容量的兒童，在遇到難度高的試題時連續產生挫折，而降低完成動機（Gonthier et al., 2018）。

有鑑於此，本研究認為，評估兒童工作記憶容量的複雜廣度作業，若使用一次增加一個廣度且未通過特定級別即停止的測驗程序，可藉由：（1）增加每一個廣度的嘗試次數；（2）不只以最大廣度來評量記憶容量；以及（3）考量兒童在同一廣度底下所有試題的正確率的方法，來減少因全有全無計分可能產生低估兒童工作記憶容量的情形。例如，TBRs 模式學者，Barrouillet 等人一系列的研究（Barrouillet et al., 2004; Barrouillet et al., 2009; Camos & Barrouillet, 2011; Lépine et al., 2005），使用自編的工作記憶廣度作業，每個廣度底下有四題（或三題），計分時以每個廣度 1 分，同一廣度底下每個嘗試題為  $1/4$  分（或  $1/3$ ，分母視每個廣度之嘗試題的數量而定），來計算兒童的部分廣度（partial span）。部分廣度與一般智力測驗中所採用的最大廣度不同，最大廣度是以同一廣度底下的任一嘗試題答對即算成功，來計算兒童最大可正確記憶的項目數。由於兒童的最大廣度大多落在 2 至 3 之間，只以最大廣度計分會影響測量的敏感度（Oberauer & Süß, 2000）。此外，若以答對 1 題嘗試題得 1 分的「總分」來計分（例如：簡馨瑩等人，2014），雖然能顯示得分越高表示工作記憶能力越好，但無法反映兒童工作記憶的廣度大小。因此，上述「部分廣度」的計分方法，能保有記憶廣度的計分概念，同時，也較能敏感的比較出兒童工作記憶容量的個別差異（Conway et al., 2005）。

綜上所述，目前雖然有魏氏幼兒智力量表第四版中文版（陳心怡、陳榮華，2013），內含適合測量 2 歲至 7 歲幼小兒童的工作記憶分測驗。但由於編列在整套智力量表中，無論是測驗中的操作工具或是計分紙，皆無法分開購置，對於只想了解工作記憶發展差異的研究者而言，仍缺乏專門用來測量工作記憶的測驗。而簡馨瑩等人（2014）編製的幼兒工作記憶測驗，是另一個適合評估 5 歲至 7 歲幼小兒童語文工作記憶的標準化測驗，但由於處理作業未限制反應時間，可能影響工作記憶容量的評估。重要的是，上述兩項測驗皆以總分的高低來評估兒童工作記憶的個別差異，無法提供兒童在處理和儲存的動態歷程下「記憶廣度」的發展訊息。

本研究旨在介紹一項適合測量 4 至 7 歲兒童工作記憶容量的電腦化複雜廣度作業，並綜合多個研究計畫與論文中收集的資料進行此作業的信度與效度分析，最後，藉由以此作業了解臺灣 4 至 7 歲兒童工作記憶的發展情形。

## 方法

### （一）研究對象

參與本研究的兒童來自新竹市、新竹縣和新北市地區，在 2018 年至 2023 年間年齡介於 4 歲至 7 歲且母語為國語的典型發展兒童共 333 名（167 名男性；平均月齡為 70.67 個月， $SD = 10.06$  個月），分為四個年齡段：4 歲組（59 個月以下；平均月齡為 55.06 個月， $SD = 3.83$  個月）、5 歲組（60 個月至 71 個月；平均月齡為 66.37 個月， $SD = 3.47$  個月）、6 歲組（72 個月至 83 個月；平均月齡為 75.82 個月， $SD = 3.03$  個月）及 7 歲組（84 至 99 個月；平均月齡為 89.85 個月， $SD = 3.95$  個月）。各組人數分別為：4 歲組 49 位（22 位男性）、5 歲組 123 位（67 位男性）、6 歲組 128 位（60 位男性）、7 歲組 33 位（18 位男性）。

本研究兒童在家長填寫通過研究倫理審查的知情同意書後，在幼兒園安靜不受干擾的教室進行個別測驗，或由家長陪同至實驗室參與研究。每位兒童在完成實驗後皆可獲得一份 100 元的等值紀念禮物。

## (二) 研究工具

### 1. 動物顏色廣度作業 (animal-color span task)

本研究所採用的動物顏色廣度作業，改編自前文所提及的第三種理論取向中，TBRs 模式的學者 Camos 與 Barrouillet (2011)，在探討兒童工作記憶發展機制之系列研究中所採用的複雜廣度作業。此作業的結構符合本研究前文所述之兒童工作記憶複雜廣度作業的設計原則，且以電腦控制儲存和處理作業的呈現時間，故本研究以此作業作為電腦化複雜廣度作業的範例。本研究以心理學實驗軟體 e-prime 3.0 撰寫測驗程式，由研究人員以 15 吋筆記電腦進行個別施測，並以電腦按鍵和計分紙紀錄兒童表現。

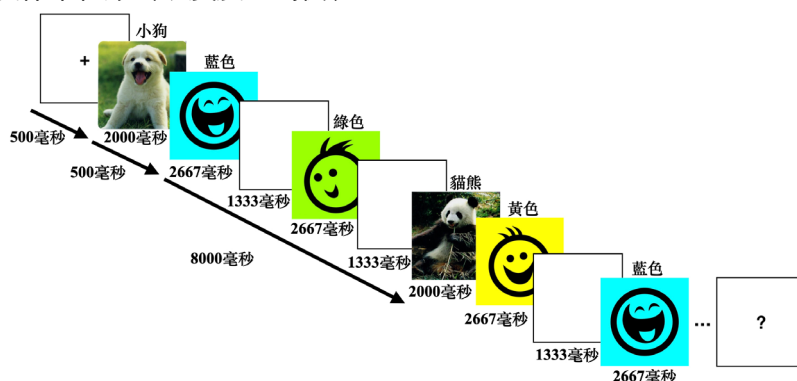
(1) 測驗材料。處理作業材料為綠色、黃色和藍色的笑臉，儲存作業材料為彩色的動物圖片，選自 Baby's 100 張全方位圖卡：動物王國 (幼福文化出版社，2011) 圖卡。以兒童需回憶的動物數為記憶廣度，由廣度 1 開始，每一廣度有 4 題，最高至廣度 5，共 20 題。每個處理作業包含兩個不同顏色笑臉，平均每個顏色的笑臉各出現 40 次。每個記憶作業使用不同的動物圖片，共使用 60 個不重複的動物圖片作為記憶項目，以避免產生試題間的干擾 (between-lists interferences) (Barrouillet et al., 2011)。各廣度的測驗題目以固定順序的方式呈現。

(2) 測驗程序：測驗前練習。本研究在正式測驗前，先個別讓兒童熟悉動物顏色廣度作業中使用的材料和電腦呈現速度。首先，以電腦呈現「顏色命名作業」，請兒童在時間 (2,667 毫秒) 內，說出螢幕上笑臉的顏色 (綠色、黃色或藍色)，目的在使兒童熟悉顏色與速度，並篩除辨色力有困難的兒童。接著，使用「動物王國圖卡」(60 個常見動物)，請兒童進行兩次圖卡命名。第一次請兒童跟著研究人員說出每一張圖卡上動物的名稱，第二次請兒童自己說出圖卡上動物的名稱，確認兒童認識且能正確說出所有的動物名稱後，練習結束。

(3) 測驗程序：正式測驗。兒童在測驗前練習階段結束後，休息 5 分鐘，接著進行正式測驗。兒童在研究人員說明指導語之後，先進行兩題練習題，確實了解實驗進行的方式之後，開始測驗。每題開始時螢幕會出現「+」，由研究人員按下空白鍵後間隔 500 毫秒螢幕上出現動物圖片 (呈現 2000ms)，兒童需立即說出動物的名稱 (例如：獅子)，並且記住，間隔 500 毫秒後，螢幕出現笑臉圖片，兒童需依序說出兩個笑臉的顏色 (例如：藍色、黃色)。每個笑臉呈現 2,667 毫秒<sup>1</sup>，空白 1,333 毫秒<sup>2</sup>，共 8,000 毫秒。當螢幕出現「？」時，兒童需立即依序回憶出動物名稱。由廣度 1 開始，每一廣度只要答對 1 題，即可繼續進行下一個廣度的測驗，直到同一廣度的 4 題全錯，則停止測驗。實驗程序如圖 1。施測時間約 15 分鐘至 20 分鐘。

(4) 計分方式。計分時需笑臉顏色判斷正確且動物名稱回憶順序正確才能計分，若其中一個顏色唸錯則該題無論動物名稱回憶順序正確與否皆不予以計分。本研究除了計算兒童的最大廣度，另以每個廣度下每答對 1 題得 1 / 4 分，來計算兒童的部分廣度。最大廣度和部分廣度皆是最小 0 分，最大 5 分。

圖 1  
動物顏色廣度作業程序 (以廣度 2 為例)



## 2. 數字廣度作業 (digit span tasks)

數字廣度作業是常用來測量語文短期記憶與語文工作記憶的測驗，先以視覺或聽覺方式呈現一系列的數字串後，再請兒童將看到或聽到的數字串，按照順序或以相反順序進行口頭回憶，前者為數字順背 (digit forward, 簡稱 DF) 廣度作業，後者為數字逆背 (digit backward, 簡稱 DB) 廣度作業。DF 作業僅涉及訊息暫存，測量的能力為短期記憶；DB 作業除了訊息暫存，另需要對訊息進行處理（將數字重新排列），測量的能力為工作記憶。

本研究的數字廣度作業，改編自陳湘淳與李玉琇（2005）的簡單廣度作業，以心理學實驗軟體 e-prime 3.0 撰寫，呈現由數字 1 ~ 9 隨機組合的記憶項目，由研究人員以 15 吋筆記電腦進行個別施測，以電腦按鍵和計分紙紀錄兒童表現。在指導語和兩題練習題之後，開始正式測驗。由電腦以每秒 1 個數字的速度播放語音數字，當電腦螢幕上出現問號時，兒童需立刻將剛剛聽到的數字按照順序（或以相反的順序）唸出來。從 2 個數字開始，最高進行到 7 個數字，每一廣度各有 3 題，每一廣度只要答對 1 題，即可繼續進行下一個廣度的測驗，直到同一廣度的 3 題全錯，則停止測驗。各廣度的測驗題目以固定順序的方式呈現。所有兒童先進行 DF 作業再進行 DB 作業。回憶時，數字和順序皆正確才予以計分。本研究計算最大廣度和部分廣度，最小 0 分，最大 7 分。

## 3. 魏氏幼兒智力量表第四版 (Wechsler preschool and primary scale of intelligence, 簡稱 WPPSI-IV)

WPPSI-IV 由陳心怡與陳榮華（2013）修訂編製，包含 2 歲 6 個月至 3 歲 11 個月、4 歲至 7 歲 11 個月兩個不同年段的測驗，具良好的折半信度、重測信度、建構效度、臨床區辨效度及效標關聯效度，與魏氏幼兒智力量表修訂版 (WPPSI-R) 和魏氏兒童智力量表第四版 (WISC-IV) 皆有高度相關。

WPPSI-IV 有五種分測驗：語言理解 (VCI)、視覺空間 (VSI)、流體智力 (FRI)、工作記憶 (WMI) 及處理速度 (PSI)。本研究採用工作記憶分測驗，包含圖畫記憶與動物園兩項測驗。本研究以個別施測方式，依指導手冊進行測驗與計分，並以兩項測驗原始分數的總分，作為工作記憶分測驗的得分。

## 4. 托尼非語言智力測驗第四版中文版 (test of nonverbal intelligence-fourth edition, 簡稱 TONI-4)

TONI-4 由林幸台等人（2016）修訂編製，分為普及版（共 60 題，適合 7 歲 6 個月至 15 歲 11 個月）與幼兒版（為普及版前 48 題，適合 4 歲至 7 歲 11 個月），並有甲乙兩種複本，具有良好信度與效度。本研究 4 至 6 歲兒童使用幼兒版，7 歲以上兒童使用普及版，避免產生天花板效應。本研究以個別施測方式，依指導手冊進行測驗與計分，並以原始分數作為非語言智力測驗的得分。

### （三）資料分析

本研究以內部一致信度與折半信度檢驗動物顏色廣度作業的信度。

效度分析部分包含：（1）幅合效度：以皮爾森相關分析檢驗動物顏色廣度作業與 DF 作業、DB 作業、WPPSI-IV-WM 及 TONI-4 分數的相關情形；（2）區辨效度：以單因子變異數分析比較不同年齡兒童的動物顏色廣度作業表現，以檢驗測驗是否能反映成熟造成的發展差異；（3）試題分析：分析每一題的通過率與鑑別指數，並以試題反應理論 (item response theory, 簡稱 IRT) Rasch 模式分析，計算題目的難度值、測量標準誤與適配度指標，以檢驗測驗的難度及鑑別度。

## 結果

### （一）測驗基本分析

本研究各變項之描述性統計分析數據如表 1 與表 2 所示，所有變項的分佈接近常態分配（偏態 < 2，峰度 < 4）（Kline, 1998）。

表 1  
各項測驗描述性統計分析

		平均	標準差	偏態		峰態	
				統計量	標準誤	統計量	標準誤
月齡		70.66	10.06	0.15	0.13	0.06	0.27
動物顏色廣度作業	最大廣度	2.49	1.00	0.29	0.13	-0.44	0.27
	部分廣度	1.56	0.75	0.57	0.13	0.26	0.27
	正確比率	0.79	0.25	-1.14	0.13	0.12	0.27
	動物顏色	0.83	0.25	-1.28	0.13	0.34	0.27
DF	最大廣度	5.27	1.27	-0.22	0.15	-0.28	0.29
	部分廣度	3.76	1.53	-0.15	0.15	-0.83	0.29
DB	最大廣度	3.00	1.29	0.32	0.15	1.60	0.29
	部分廣度	1.89	1.14	0.64	0.15	0.40	0.29
WPPSI-IV-WM		34.06	6.26	0.34	0.23	-0.12	0.46
TONI-4		24.58	7.29	0.05	0.16	-0.26	0.31

註：DF = 數字順背廣度作業；DB = 數字逆背廣度作業；WPPSI-WM = 魏氏幼兒智力量表第四版 - 工作記憶分測驗；TONI-4 = 托尼非語言智力測驗第四版。

表 2  
各年齡段兒童之各項測驗平均數（標準差）

		4 歲	5 歲	6 歲	7 歲
動物顏色廣度作業	最大廣度	1.82 (0.81)	2.35 (0.90)	2.59 (0.90)	3.64 (0.93)
	部分廣度	1.04 (0.57)	1.38 (0.58)	1.70 (0.72)	2.50 (0.68)
	正確比率	0.92 (0.02)	0.77 (0.24)	0.75 (0.28)	0.93 (0.13)
	儲存作業處理作業	0.96 (0.02)	0.82 (0.24)	0.79 (0.28)	0.95 (0.12)
DF	最大廣度	4.11 (1.26)	5.11 (1.15)	5.64 (1.12)	6.67 (0.65)
	部分廣度	1.91 (1.14)	3.61 (1.32)	4.25 (1.29)	5.89 (0.50)
DB	最大廣度	1.43 (1.04)	2.86 (0.86)	3.47 (1.21)	4.42 (1.31)
	部分廣度	0.51 (0.54)	1.70 (0.75)	2.25 (1.06)	3.72 (1.25)
WPPSI-IV-WM		28.20 (3.27)	32.15 (5.59)	35.74 (6.21)	38.75 (4.71)
TONI-4		18.37 (5.42)	23.19 (7.38)	25.17 (7.38)	30.42 (5.85)

註：DF = 數字順背廣度作業；DB = 數字逆背廣度作業；WPPSI-WM = 魏氏幼兒智力量表第四版—工作記憶分測驗；TONI-4 = 托尼非語言智力測驗第四版。

## （二）信度分析

動物顏色廣度作業具有理想的信度，內部一致性分析之 Cronbach's  $\alpha$  係數為 .76，將題目隨機折半之折半信度為 .77。四歲組 Cronbach's  $\alpha$  係數 .71、折半信度 .83；五歲組 Cronbach's  $\alpha$  係數 .63、折半信度 .66；六歲組 Cronbach's  $\alpha$  係數 .75、折半信度 .80；七歲組 Cronbach's  $\alpha$  係數 .64、折半信度 .81。

## （三）效度分析

### 1. 幅合效度

皮爾森積差相關分析發現，動物顏色廣度作業與 DF 作業、DB 作業、WPPSI-IV-WM 和 TONI-4，皆有顯著正相關，相關值介於 .33 到 .45，屬於中度相關（如表 3）。此結果顯示動物顏色廣度作業具有良好的幅合效度，兒童在動物顏色廣度作業的表現，能反映他們在短期記憶（DF 作業）和工作記憶（DB 作業、WPPSI-WM），以及非語文智力（TONI-4）的個別差異。

表 3  
動物顏色廣度作業分數與其他測驗分數之相關分析

		DF		DB		WPPSI-IV- WM	TONI-4
		最大 廣度	部分 廣度	最大 廣度	部分 廣度		
所有幼兒 (N)		272	272	114	241		
動物顏色	最大廣度	.38***	.33***	.45***	.36***	.45***	.35***
廣度作業	部分廣度	.44***	.37***	.52***	.41***	.53***	.43***
4—5 歲幼兒 (N)		141	141	46	101		
動物顏色	最大廣度	.24**	.21**	.34***	.37***	.25	.15
廣度作業	部分廣度	.25**	.15*	.25**	.20**	.32**	.17
6—7 歲幼兒 (N)		131	131	65	141		
動物顏色	最大廣度	.43***	.33***	.45***	.34***	.48***	.33***
廣度作業	部分廣度	.50***	.41***	.51***	.39***	.50***	.40***

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

## 2. 建構效度

本研究以年齡區辨性與試題分析，檢驗動物顏色廣度作業的建構效度。

(1) 年齡差異分析。本研究以單因子變異數分析，檢驗兒童在動物顏色廣度作業的最大廣度與部分廣度是否因不同年齡有顯著差異，結果發現年齡效果顯著（最大廣度： $F(3, 329) = 29.29, p < .001$ ；部分廣度： $F(3, 329) = 37.69, p < .001$ ），進一步以 Scheffe 法進行事後比較（如表 4），發現 4 歲至 7 歲兒童在動物顏色廣度作業的表現隨著年齡而顯著增加，顯示動物顏色廣度作業能夠反映工作記憶發展的成熟差異，具有良好的建構效度。此外，5 歲與 6 歲兒童的最大廣度沒有顯著差異，但部分廣度有顯著差異，顯示部分廣度的計分方式較敏感於反映兒童工作記憶的發展差異。

表 4  
各年齡組之動物顏色廣度作業最大廣度與部分廣度差異檢定（ $t$  值）

		最大廣度			部分廣度		
		5 歲	6 歲	7 歲	5 歲	6 歲	7 歲
4 歲		-0.53**	-0.78*	-1.82***	-0.34**	-0.66***	-1.44***
5 歲		—	-0.24	-1.29***	—	-0.32***	-1.10***
6 歲			—	-1.04***		—	-0.78***

結果整理：

最大廣度：7 歲組 > 6 歲組 = 5 歲組 > 4 歲組

部分廣度：7 歲組 > 6 歲組 > 5 歲組 > 4 歲組

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

## 3. 試題分析

本研究根據古典測驗理論，計算動物顏色廣度作業逐題的通過率與鑑別指數，通過率以該題目答對人數的比例計算；鑑別指數以測驗總分前百分之 25 之兒童的通過率減去後百分之 25 兒童的通過率計算。各廣度之分析結果如表 5 所示，其中，廣度 2 至廣度 3 的鑑別指數很高，顯示這兩個廣度的題目鑑別力很好，雖然其他廣度的鑑別力偏低，但可視為受題目難度極端之影響所致。由此結果可以得知，本研究大部分兒童能夠通過廣度 1，能力介於廣度 2 至廣度 3 之間。

在 IRT 的分析，本研究使用 jamovi 2.3 軟體內建的 snowIRT（Martinková & Drabinová, 2018; Robitzsch et al., 2020; Seol, 2023; The jamovi project, 2023）程式計算逐題的題目難度、測量標準誤與適配度指標，其中以試題的 Infit 與 Outfit 均方誤（mean-squares error, MnSq）作為 Rasch 模式適配度的指標。結果如表 5 所示，廣度 1 至廣度 3 的 MnSq 無論是 Infit 或 Outfit，皆介於理想值 0.5 ~ 1.5（Linacre, 2002）之間，顯示具有良好的適配度，而廣度 4、廣度 5 的 Outfit 的數值偏低（除了廣度



5 的最後一題)，顯示作答誤差的隨機性低，可能因為廣度 4、廣度 5 的題目對兒童而言難度過高，大部分的兒童無法答對，導致題目訊息太過雷同。但反觀這些題目的 Infit 都在可接受的範圍中，表示這些題目仍符合 Rasch 模型。

本研究進一步將兒童的能力估計值與題目難度位置繪製成圖形（item-person map，如圖 2），進行試題與兒童的比對。圖 2 最左邊為測量尺度，單位為 logit；圖 2 左邊部分為兒童的能力位置分佈圖，越上方表示工作記憶越好；圖 2 右邊部分為題目的難度位置，越上方表示題目難度越難。從圖 2 可以看出，兒童能力大致符合常態分配，能力值介於 -5 ~ 2；題目難度散布於 -4.5 ~ 5，而廣度 4 與廣度 5 的題目難度高，很少兒童能答對。

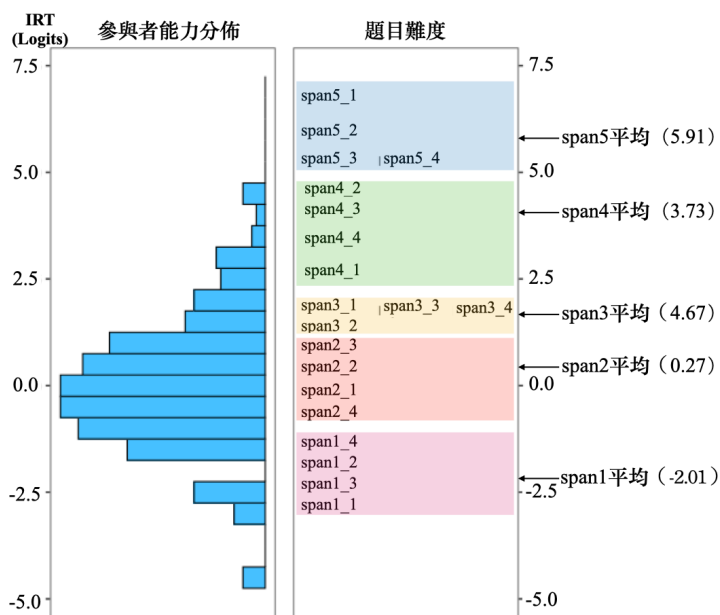
最後，IRT 難度分析結果發現，隨著廣度增加平均難度也漸增（如表 5），顯示動物顏色廣度作業具有良好的建構效度。

此外，以年齡來分析的結果發現，4 歲兒童在廣度一、廣度二、廣度三、廣度四、廣度五的平均通過率為 72.96%、23.47%、15.03%、7.14%、0%；5 歲兒童在廣度一、廣度二、廣度三、廣度四、廣度五的平均通過率為 79.05%、40.23%、15.03%、0%、0%；6 歲兒童在廣度一、廣度二、廣度三、廣度四、廣度五的平均通過率為 85.55%、50.59%、27.93%、5.66%、0%；7 歲兒童在廣度一、廣度二、廣度三、廣度四、廣度五的平均通過率為 92.43%、75.00%、50.00%、25.00%、5.30%，表示 4、5、6、7 歲兒童的最大廣度分別約為廣度一、廣度二、廣度二、廣度三，符合發展趨勢，顯示動物顏色廣度作業具有良好的建構效度。

**表 5**  
動物顏色廣度作業之各題試題分析

廣度	試題	古典 通過率	鑑別 指數	MnSq		IRT 難度	測量標準誤
				outfit	infit		
一	1	.86	.22	1.66	1.12	-2.40	0.18
	2	.81	.36	1.10	1.02	-1.93	0.16
	3	.84	.29	1.38	1.04	-2.14	0.16
	4	.76	.43	1.40	1.05	-1.55	0.15
	四題平均	.82	.33	—	—	-2.01	—
二	1	.56	.62	1.02	1.05	-0.34	0.13
	2	.37	.50	1.19	1.12	0.68	0.13
	3	.30	.56	1.03	1.05	1.11	0.14
	4	.57	.63	1.10	1.02	-0.37	0.13
	四題平均	.45	.58	—	—	0.27	—
三	1	.20	.44	0.90	1.00	1.81	0.16
	2	.26	.59	0.80	0.91	1.37	0.14
	3	.22	.48	0.82	0.94	1.69	0.15
	4	.20	.46	0.82	0.95	1.81	0.16
	四題平均	.22	.49	—	—	1.67	—
四	1	.11	.32	0.50	0.82	2.76	0.20
	2	.03	.05	0.52	0.94	4.48	0.36
	3	.04	.11	0.40	0.92	4.15	0.31
	4	.06	.15	0.60	0.88	3.53	0.25
	四題平均	.06	.16	—	—	3.73	—
五	1	<.01	.01	0.32	1.01	6.81	1.01
	2	.01	.01	0.41	0.99	6.10	0.72
	3	.01	.04	0.17	0.89	5.37	0.52
	4	.01	.04	0.28	0.97	5.37	0.52
	四題平均	.02	.03	—	—	5.91	—

圖 2  
動物顏色廣度作業之試題與兒童分佈關係圖



## 討論

雖然臨床上經常使用數字逆背廣度作業測量兒童的工作記憶容量，但由於數字逆背廣度作業屬於簡單廣度作業，在理論上是否測量到工作記憶系統的動態運作機制仍有爭議（Conway et al., 2005; Hutton & Towse, 2001）。本研究從工作記憶理論與認知發展的角度，探討如何以複雜廣度作業評估兒童的工作記憶，並檢驗改編自 Camos 與 Barrouillet (2011) 的動物顏色廣度作業的信、效度，最後，以此電腦化複雜廣度作業評估臺灣 4 ~ 7 歲兒童的工作記憶發展。以下針對本研究的發現，做進一步的討論。

### （一）動物顏色廣度作業適合作為兒童工作記憶測量工具

本研究使用的動物顏色廣度作業，相較於數字逆背廣度作業，更符合工作記憶理論概念（Baddeley, 2003; Barrouillet et al., 2007; Cowan, 1999; Engle et al., 1999），測量的是同時進行儲存和處理之動態歷程下中央執行器的容量。此外，動物顏色廣度作業也是一種電腦版的複雜廣度作業，以電腦控制處理和儲存作業呈現的時間和順序。最初，Barrouillet 等人（2009）和 Camos 與 Barrouillet (2011) 在其研究中設計使用的動物顏色廣度作業（一種複雜廣度作業），操弄三種不同的處理作業情境，目的在比較 5 歲至 7 歲兒童在不同的處理作業認知負荷下的記憶表現差異。本研究改編使用其中一種處理作業情境，即兩個顏色判斷且總呈現時間較長（2 colors long）的處理作業情境，而未使用其他兩個處理作業情境（1 color short; 2 colors short）的原因，主要是考量參與兒童的認知能力，使作業對 4 歲兒童不致太難且對 7 歲兒童不致太簡單（另外兩種情境皆採用步調較快的呈現形式，即 1,333 毫秒呈現有顏色笑臉、間隔 667 毫秒）。由本研究的 IRT 難度分析結果可以證實，本研究改編的作業版本難度適中，參與本研究大多數的兒童可以完成廣度 1 至廣度 3 的嘗試題。而由廣度 4 和廣度 5 對 4 ~ 7 歲兒童而言難度較高的結果也反映，此作業適合測量的兒童年齡層應該更廣，未來研究可以 8 ~ 12 歲兒童為對象，分析較大年齡兒童在此作業下的工作記憶表現。

進一步就作業刺激安排來看，本研究改編之動物顏色廣度作業的處理作業（2,667 毫秒內正確說出笑臉的顏色），對 4 ~ 7 歲兒童而言有趣、簡單但有時間限制，儲存作業（說出並記住動物圖

片的名稱)亦是以4~7歲兒童熟悉的視覺圖片作為記憶項目。這些作業特性,除了能增進較小兒童對作業投入的動機之外,如文獻探討曾提到的,可以降低作業複雜性以及策略使用對工作記憶評估的影響(Lépine et al., 2005),亦能達到使用處理作業來佔住注意力,以干擾或中斷兒童對記憶項目的保留的效果(Barrouill et al., 2004)。

接著,在測驗程序方面,由於本研究參與者主要為學齡前兒童,動物顏色廣度作業和大部分測量兒童工作記憶的作業一樣,使用逐漸提升作業難度的方式,在兒童無法通過某一廣度的所有嘗試題時停止測驗(Alloway, 2007; Gonthier et al., 2018; Pickering, 2006)。如此一來,不僅有利於了解兒童的最佳能力,也可以節省施測的時間(同一廣度嘗試題全錯後,不用再進行後面更高廣度的嘗試題)。更重要的是,能避免因隨機呈現廣度順序,可能使年齡較小兒童因連續錯誤產生挫折而降低完成動機的現象(Gonthier et al., 2018),進而提升作業的效度。

最後,在計分程序部分,本研究採用全有全無的判斷方式,所有的記憶項目需按照正確順序回憶才算對(Conway et al., 2005)。考量年齡較小的兒童,在回憶時可能出現項目順序正確但漏失部分項目(omission)的比例高的情形(Camos & Barrouillet, 2011)。動物顏色廣度作業的每一個廣度的嘗試題增加至四題(成人的複雜廣度作業通常一個廣度只有兩個嘗試題),且在計分時,除了計算最大廣度(無論同一廣度下答對幾個嘗試題,都計為廣度1分,每一個廣度不是0分就是1分)之外,另計算兒童的部分廣度(同一廣度下每一個嘗試題答對計為1/4分,依照答對的嘗試題計算該廣度得分,每一個廣度可能有0分、1/4分、2/4分、3/4分或4/4分)。如此不僅可減少因全有全無計分而低估兒童工作記憶容量的可能性,更能增加作業對幼小兒童工作記憶個別差異測量的敏感度(Oberauer & Süß, 2000)。此外,部分廣度的計分方式也有助於了解兒童在不同廣度下的正確回憶比率。

綜合上述討論,本研究認為動物顏色廣度作業在作業的刺激、內容、呈現形式與反應方式,符合4~7歲兒童的發展能力;以電腦化的操作,使用簡單有趣但有時間限制的處理作業,能佔住注意力並對儲存作業產生干擾,符合工作記憶理論(Baddeley, 2003; Barrouillet et al., 2007; Cowan, 1999; Engle et al., 1999);測驗程序和計分方式皆考量幼小兒童在執行複雜廣度作業時,可能受到因作業特性而產生外在與內在的負面影響,有助提升作業的信效度。故本研究認為,動物顏色廣度作業,作為一項電腦化的複雜廣度作業,適合用來測量4~7歲兒童的工作記憶發展。

## (二) 動物顏色廣度作業具有良好的心理計量特性

本研究信效度分析結果顯示,動物顏色廣度作業具有良好的心理計量特性。在信度方面,有良好的內在一致性與折半信度。在同時效度方面,兒童在動物顏色廣度作業的表現,與他們在短期記憶(DF作業)和工作記憶(DB作業、WPPSI-WM),以及非語文智力(TONI-4)的分數有顯著相關。試題分析結果也證實,動物顏色廣度作業的廣度1至廣度3的試題鑑別度高,能夠區別出不同能力的兒童,且廣度越大難度越高,符合Rasch模型。

除此之外,4至7歲兒童在此作業上的表現符合常態分配,且隨著年齡增加,相鄰年齡的兒童在此作業上的部分廣度皆有顯著的增加。雖然,以最大廣度來看,5歲與6歲兒童之間沒有顯著差異。此結果正呼應本文文獻探討所提到過去研究的發現,最大廣度計分可能影響兒童工作記憶測量的敏感性(Oberauer & Süß, 2000),部分廣度的計分較能反映兒童工作記憶容量的個別差異(Conway et al., 2005)。

## (三) 臺灣4歲至7歲兒童工作記憶發展

由333位來自臺灣北部的4歲至7歲兒童,接受本研究動物顏色廣度作業測量的結果顯示,兒童工作記憶的最大廣度在4歲時平均為1.82、5歲時平均為2.35、6歲時平均為2.59、7歲時平均為3.64,4~7歲的平均最大廣度為2.49。若以部分廣度來看,4~7歲兒童的工作記憶平均部分廣度為1.56。值得注意的是,其中5歲組和7歲組的部分廣度分別為1.38和2.50,與Barrouillet等人(2009)實驗三中5歲組和7歲組的部分廣度為1.10和2.40的結果非常接近。由此顯示,使用動物顏色廣度作業所測得的兒童工作記憶發展,具有跨文化的一致性,臺灣兒童與法國兒童的工作記憶廣度發展情形相近。

許多工作記憶理論模式，從不同角度探討造成兒童在複雜廣度作業上所顯現的發展改變或個別差異的因素，例如，前文曾提及的第二種理論取向中 Cowan（1999）的注意力鑲嵌模式、Engle 等人（1999）的容量個別差異模式，以及第三種理論取向的 Case 等人（1982）的資源共享模式、Towse 與 Hitch（1995）的時間衰退模式，以及 Barrouillet 等人（2009）的 TBRs 模式。這些理論使用複雜廣度作業要求兒童同時處理和儲存，並著眼於探討兒童在複雜廣度作業的儲存表現如何受到處理作業的干擾。綜合這些理論取向的看法，兒童工作記憶的廣度隨著大腦額葉皮質的成熟，使兒童注意力容量增加，及認知資源在處理和儲存作業之間更有效分配而增加。TBRs 模式的 Camos 與 Barrouillet（2011）進一步發現，幼小兒童和較大兒童的工作記憶運作機制不同，7 歲以上的兒童才會在工作記憶的處理作業期間，以轉換注意力的方式主動地恢復正在衰退的記憶痕跡，而 5 歲及 6 歲的幼兒只是被動地保留記憶項目。此外，他們也發現 5 歲和 6 歲的兒童的工作記憶廣度沒有顯著差異，但 5 歲和 7 歲兒童的工作記憶廣度則有顯著差異，亦即 5 至 7 歲兒童的工作記憶發展呈現質的改變。

本研究使用與 TBRs 模式相同的工作記憶複雜廣度作業，並同樣以部分廣度來分析兒童的工作記憶發展，發現 4 歲至 7 歲兒童的工作記憶部分廣度隨年齡增長逐年有顯著的進步，除了顯示工作記憶的中央執行器容量可能從 4 歲開始逐漸發展（與過去認知神經科學的研究發現一致）之外，也顯示兒童的工作記憶發展在 4 歲至 7 歲呈現量的改變，而後者與 TBRs 的發現不同，可能是文化差異的因素所致。

## 結論與建議

本研究改編之動物顏色廣度作業，是一項電腦化的複雜廣度作業，適合測量 4 歲至 7 歲的兒童，有良好的信度與效度，且能反映成熟造成的工作記憶發展差異。除此之外，本研究改編使用的電腦化複雜廣度作業有許多優點與貢獻：（1）在理論上，符合工作記憶理論模式（包含儲存作業和處理作業），能控制處理作業的認知負荷（Camos & Barrouillet, 2011），並測量到兒童在記憶的保留階段，抵抗干擾並維持注意的控制能力（Jarrold et al., 2011）；（2）在研究方法上：測量所使用的儲存作業和處理作業的刺激大小、呈現時間、間隔時間，以及測驗的終止，都由電腦控制，有助標準化測驗的程序、減少人為主觀導致的誤差；（3）在實務應用上：使不同地點的不同研究者之間能採用一致的施測程序（Bailey, 2012），有利於跨文化及跨研究間的比較；（4）在未來研究上：電腦化的複雜廣度作業因具有上述特色，可以進一步發展成不需要實驗者在旁協助的線上測驗，或是團體測驗。

然而，後續兒童工作記憶發展與測量研究，使用本研究所改編的電腦版複雜廣度作業，仍有幾點需要注意：（1）本研究採用的測驗程序，每個兒童依其能力完成的題目數量不同，可能產生不同程度的前向干擾（Gonthier et al., 2018）；（2）本研究僅改編使用 Barrouillet 等人（2009）和 Camos 與 Barrouillet（2011）設計的其中一種版本的動物顏色廣度作業，另外兩種不同處理作業情境的動物顏色廣度作業，是否也適合用來測量 4～7 歲兒童的工作記憶容量，本研究並未進行分析；（3）承上，前述兩個研究最初設計的複雜廣度作業，主要用來探討 5 至 7 歲期間兒童工作記憶發展的關鍵機制，本研究將其中一個版本的作業之適用年齡向下延伸至 4 歲，但此作業是否適合年齡更小或是更大的兒童，仍需要後續研究進一步探討。不過就目前的資料看來，4 歲兒童的平均最大廣度僅有 1.82，4 歲以下的兒童進行此作業可能會出現地板效應，而由 7 歲兒童的平均最大廣度為 3.7 看來，此作業可能有機會評量 8 歲以上兒童的工作記憶廣度；（4）本研究所介紹之電腦版複雜廣度作業（動物顏色廣度作業），以心理學實驗付費軟體 e-prime 3 撰寫，用以控制儲存作業和處理作業的刺激呈現時間。未來有興趣進行兒童工作記憶發展測量之研究，欲使用此作業作為研究工具，在無法取得 e-prime 3 實驗軟體的情況下，仍可使用其他免費的實驗軟體，或是直接使用微軟的簡報軟體 PowerPoint 來呈現刺激的畫面、設定刺激的呈現時間，並配合計分紙，來完成此作業的施測與計分。

## 註釋

<sup>1</sup> 此呈現時間的毫秒數專為 60Hz 的電腦顯示頻率所設計（2,667 毫秒為 160 個週期）。

<sup>2</sup> 此呈現時間的毫秒數專為 60Hz 的電腦顯示頻率所設計（1,333 毫秒為 80 個週期）。

## 參考文獻

- 幼福文化出版社（2011）：《Baby's 100 張全方位圖卡：動物王國》。[You Fu Culture Publishing Company. (2011). *Baby's 100 Zhang quanfangwei tuka: Dongwu wangguo.*]
- 林幸台、吳武典、胡心慈、郭靜姿、蔡崇建、王振德（2016）：《托尼非語文智力測驗（第四版中文版）》。心理出版社。[Lin, H.-T., Wu, W.-D., Hu, H.-T., Kuo, C.-T., Tsai, T.-C., & Wang, T.-T. (2016). *Test of Nonverbal Intelligence* (4th ed.). Psychological Publishing.]
- 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)
- 陳心怡、陳榮華（2013）：《魏氏幼兒智力量表第四版（WPPSI-IV）中文版》。中國行為科學社。[Chen, H.-Y., & Chen, Y.-H. (2013). *Wechsler Preschool and Primary Scale of Intelligence* (4th ed.). Chinese Behavioral Science Corporation.]
- 陳湘淳（2012）：《兒童工作記憶發展縱貫研究：複誦策略的效果（未出版博士論文）》，國立中正大學。[Chen, H.-C. (2012). *A longitudinal study of working memory development in children: The effect of repetition strategies* (Unpublished doctoral dissertation). National Chung Cheng University.]
- 陳湘淳、李玉琇（2005）：〈記憶策略訓練對工作記憶容量的影響〉。《教育心理學報》，37，41–59。[Chen, H.-C., & Lee, Y.-S. (2005). The effect of mnemonic training upon the working memory capacity. *Bulletin of Educational Psychology*, 37, 41–59.] <https://doi.org/10.6251/BEP.20050813>
- 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。[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)
- 簡馨瑩、趙子揚、王繼伶（2014）：〈「幼兒工作記憶測驗」之編製〉。《測驗學刊》，61，159–181。[Chien, H.-Y., Chao, T.-Y., & Wang, J.-L. (2014). The development of Chinese Working Memory Test for Young Children. *Psychology Testing*, 61, 159–181.]
- Alloway, T. P. (2007). Working memory, reading, and mathematical skills in children with developmental coordination disorder. *Journal of Experimental Child Psychology*, 96(1), 20–36.

- <https://doi.org/10.1016/j.jecp.2006.07.002>
- Alloway, T. P. (2011). A comparison of working memory profiles in children with ADHD and DCD. *Child Neuropsychology*, 17(5), 483–494. <https://doi.org/10.1080/09297049.2011.553590>
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, 80(2), 606–621. <https://doi.org/10.1111/j.1467-8624.2009.01282.x>
- Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short term and working memory in children: Are they separable? *Child Development*, 77(6), 1698–1716. <https://doi.org/10.1111/j.1467-8624.2006.00968.x>
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829–839. <https://doi.org/10.1038/nrn1201>
- Bailey, H. (2012). Computer-paced versus experimenter-paced working memory span tasks: Are they equally reliable and valid? *Learning and Individual Differences*, 22(6), 875–881. <https://doi.org/10.1016/j.lindif.2012.06.004>
- Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, 133(1), 83–100. <https://doi.org/10.1037/0096-3445.133.1.83>
- Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 570–585. <https://doi.org/10.1037/0278-7393.33.3.570>
- Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental Psychology*, 45(2), 477–490. <https://doi.org/10.1037/a0014615>
- Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review*, 118(2), 175–192. <https://doi.org/10.1037/a0022324>
- Bayliss, D. M., Jarrold, C., Baddeley, A. D., Gunn, D. M., & Leigh, E. (2005). Mapping the developmental constraints on working memory span performance. *Developmental Psychology*, 41(4), 579–597. <https://doi.org/10.1037/0012-1649.41.4.579>
- Camos, V., & Barrouillet, P. (2011). Developmental change in working memory strategies: From passive maintenance to active refreshing. *Developmental Psychology*, 47(3), 898–904. <https://doi.org/10.1037/a0023193>
- Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, 33(3), 386–404. [https://doi.org/10.1016/0022-0965\(82\)90054-6](https://doi.org/10.1016/0022-0965(82)90054-6)
- Cheie, L., MacLeod, C., Miclea, M., & Visu-Petra, L. (2017). When children forget to remember: Effects of

- reduced working memory availability on prospective memory performance. *Memory & Cognition*, 45(4), 651–663. <https://doi.org/10.3758/s13421-016-0682-z>
- Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769–786. <https://doi.org/10.3758/BF03196772>
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104(2), 163–191. <https://doi.org/10.1037/0033-2909.104.2.163>
- Cowan, N. (1997). The development of working memory. In N. Cowan (Ed.), *The development of memory in childhood* (pp. 163–199). Psychology Press/Erlbaum (UK) Taylor & Francis.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. <https://doi.org/10.1017/CBO9781139174909.006>
- Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, 26(2), 197–223. <https://doi.org/10.1007/s10648-013-9246-y>
- Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51(1), 42–100. <https://doi.org/10.1016/j.cogpsych.2004.12.001>
- Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven year olds allocate attention like adults unless working memory is overloaded. *Developmental Science*, 13(1), 120–133. <https://doi.org/10.1111/j.1467-7687.2009.00864.x>
- Demetriou, A., Mougi, A., Spanoudis, G., & Makris, N. (2022). Changing developmental priorities between executive functions, working memory, and reasoning in the formation of g from 6 to 12 years. *Intelligence*, 90, Article 101602. <https://doi.org/10.1016/j.intell.2021.101602>
- Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), 972–992. <https://doi.org/10.1037/0278-7393.18.5.972>
- Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge University Press. <https://doi.org/10.1017/CBO9781139174909.007>
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 44, pp. 145–199). Elsevier Science. [https://doi.org/10.1016/S0079-7421\(03\)44005-X](https://doi.org/10.1016/S0079-7421(03)44005-X)
- Gathercole, S. E. (1998). The development of memory. *The Journal of Child Psychology and Psychiatry*, 39(1), 3–27. <https://doi.org/10.1111/1469-7610.00301>

- Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*, 3(11), 410–419. [https://doi.org/10.1016/s1364-6613\(99\)01388-1](https://doi.org/10.1016/s1364-6613(99)01388-1)
- Gathercole, S. E., & Pickering, S. (2001). Research section: Working memory deficits in children with special educational needs. *British Journal of Special Education*, 28(2), 89–97. <https://doi.org/10.1111/1467-8527.00225>
- Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, 2(2), 103–127. <https://doi.org/10.1080/09658219408258940>
- Gavens, N., & Barrouillet, P. (2004). Delays of retention, processing efficiency, and attentional resources in working memory span development. *Journal of Memory and Language*, 51(4), 644–657. <https://doi.org/10.1016/j.jml.2004.06.009>
- Gonthier, C., Aubry, A., & Bourdin, B. (2018). Measuring working memory capacity in children using adaptive tasks: Example validation of an adaptive complex span. *Behavior Research Methods*, 50(3), 910–921. <https://doi.org/10.3758/s13428-017-0916-4>
- Henry, L. A., & Millar, S. (1993). Why does memory span improve with age? A review of the evidence for two current hypotheses. *European Journal of Cognitive Psychology*, 5(3), 241–287. <https://doi.org/10.1080/09541449308520119>
- Holmes, J., Hilton, K. A., Place, M., Alloway, T. P., Elliott, J. G., & Gathercole, S. E. (2014). Children with low working memory and children with ADHD: Same or different? *Frontiers in Human Neuroscience*, 8, Article 976. <https://doi.org/10.3389/fnhum.2014.00976>
- Hutton, U. M., & Towse, J. N. (2001). Short-term memory and working memory as indices of children's cognitive skills. *Memory*, 9(4-6), 383–394. <https://doi.org/10.1080/09658210042000058>
- Jaroslawska, A. J., Gathercole, S. E., Logie, M. R., & Holmes, J. (2016). Following instructions in a virtual school: Does working memory play a role? *Memory & Cognition*, 44(4), 580–589. <https://doi.org/10.3758/s13421-015-0579-2>
- Jarrold, C., Tam, H., Baddeley, A. D., & Harvey, C. E. (2011). How does processing affect storage in working memory tasks? Evidence for both domain-general and domain-specific effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(3), 688–705. <https://doi.org/10.1037/a0022527>
- Jeffries, S., & Everatt, J. (2004). Working memory: Its role in dyslexia and other specific learning difficulties. *Dyslexia*, 10(3), 196–214. <https://doi.org/10.1002/dys.278>
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. The Guilford Press.
- Könen, T., & Karbach, J. (2020). Self-reported cognitive failures in everyday life: A closer look at their relation to personality and cognitive performance. *Assessment*, 27(5), 982–995. <https://doi.org/10.1177/1073191118786800>
- Kornmann, J., Zettler, I., Kammerer, Y., Gerjets, P., & Trautwein, U. (2015). What characterizes children nominated as gifted by teachers? A closer consideration of working memory and intelligence. *High*



- Ability Studies*, 26(1), 75–92. <https://doi.org/10.1080/13598139.2015.1033513>
- Lépine, R., Barrouillet, P., & Camos, V. (2005). What makes working memory spans so predictive of high-level cognition? *Psychonomic Bulletin & Review*, 12(1), 165–170.  
<https://doi.org/10.3758/BF03196363>
- Linacre, J. (2002) What do infit, outfit, mean-square, and standardization mean? *Archives of Rasch Measurement*, 16(2), 871–882.
- Martinková, P., & Drabinová, A. (2018). Shiny item analysis for teaching psychometrics and to enforce routine analysis of educational tests. *The R Journal*, 10(2), 503–515.
- Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology*, 19(1), 78–94.  
[https://doi.org/10.1044/1058-0360\(2009/09-0028\)](https://doi.org/10.1044/1058-0360(2009/09-0028))
- Oberauer, K., & Süß, H. M. (2000). Working memory and interference: A comment on Jenkins, Myerson, Hale, and Fry (1999). *Psychonomic Bulletin & Review*, 7(4), 727–733.  
<https://doi.org/10.3758/BF03213013>
- Peng, P., & Kievit, R. A. (2020). The development of academic achievement and cognitive abilities: A bidirectional perspective. *Child Development Perspectives*, 14(1), 15–20.  
<https://doi.org/10.1111/cdep.12352>
- Pickering, S. J. (2006). Working memory in dyslexia. In T. P. Alloway & S. E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 7–40). Psychology Press.
- Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, 28(3), 164–171. <https://doi.org/10.1027/1015-5759/a000123>
- Robitzsch, A., Kiefer, T., & Wu, M. (2020). *TAM: Test analysis modules* (Version 4.1.4) [Computer software]. <https://CRAN.R-project.org/package=TAM>
- Seol, H. (2023). *snowIRT: Item Response Theory for jamovi* (Version 4.8.8) [jamovi module]. <https://github.com/hyunsooseol/snowIRT>
- Slattery, E. J., Ryan, P., Fortune, D. G., & McAvinue, L. P. (2021). Contributions of working memory and sustained attention to children's reading achievement: A commonality analysis approach. *Cognitive Development*, 58, Article 101028. <https://doi.org/10.1016/j.cogdev.2021.101028>
- St Clair-Thompson, H. L. (2007). The influence of strategies on relationships between working memory and cognitive skills. *Memory*, 15(4), 353–365. <https://doi.org/10.1080/09658210701261845>
- Swanson, H. L. (2006). Cross-sectional and incremental changes in working memory and mathematical problem solving. *Journal of Educational Psychology*, 98(2), 265–281.  
<https://doi.org/10.1037/0022-0663.98.2.265>
- The jamovi project. (2023). *jamovi* (Version 2.3) [Computer Software]. <https://www.jamovi.org>
- Towse, J. N., & Hitch, G. J. (1995). Is there a relationship between task demand and storage space in tests

- of working memory capacity? *The Quarterly Journal of Experimental Psychology*, 48(1), 108–124.  
<https://doi.org/10.1080/14640749508401379>
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28(2), 127–154. [https://doi.org/10.1016/0749-596X\(89\)90040-5](https://doi.org/10.1016/0749-596X(89)90040-5)
- Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, 71, 1–26.  
<https://doi.org/10.1016/j.cogpsych.2014.01.003>
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, 17(6), 635–654. <https://doi.org/10.1080/09658210902998047>

收稿日期：2023 年 08 月 30 日  
一稿修訂日期：2023 年 09 月 19 日  
二稿修訂日期：2023 年 09 月 25 日  
三稿修訂日期：2024 年 02 月 16 日  
四稿修訂日期：2024 年 03 月 12 日  
五稿修訂日期：2024 年 06 月 05 日  
接受刊登日期：2024 年 06 月 06 日

Bulletin of Educational Psychology, 2024, 56(2), 261–280  
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

# Computerized Complex Span Measurement of Children's Working Memory

Hsiang-Chun Chen<sup>1</sup> and Yi Tang<sup>2</sup>

Psychology researchers and educators have determined that many behavioral and cognitive abilities that we want to understand, measure, and manipulate are related to the function of working memory. Working memory, a cognitive system with limited capacity, temporarily stores and processes information relevant to ongoing activity (Baddeley, 2000). Unlike short-term memory, working memory has a longer developmental period and is affected by the maturation of the brain's frontal lobes (Cowan, 1997). This cognitive ability is essential not only for children's interactions with the external world and acquisition of new knowledge but also for their daily functioning.

Although different theories provide various explanations for the development of working memory, most theories hold that working memory involves two cognitive operations: storage and processing (Baddeley, 2000; Engle & Kane, 2004; Gavens & Barrouillet, 2004). Therefore, tasks that measure a person's working memory span typically require the person to temporarily remember a memory item while performing another processing task or to engage in some form of processing related to the memory item. They are specifically called dual-task or complex span tasks (Redick et al., 2012). Such tasks differ from simple span tasks, which require the testee to retain memory items for a period of time and then recall them sequentially.

Assessing children's working memory by using complex span tasks requires clear, simple, and easy-to-understand instructions. The tasks must be designed with stimuli, content, presentation formats, and response methods that are consistent with children's cognitive abilities. In addition, test procedures should account for children with varying levels of working memory capacity. For those with low working memory capacity, the tasks should avoid inducing a sense of frustration or helplessness. By contrast, for children with high working memory capacity, the test should provide opportunities for them to fully demonstrate their abilities. A common approach involves gradually increasing the difficulty of the tasks, requiring the testee to memorize more items, and terminating the task when a specific level of difficulty cannot be completed (Gonthier et al., 2018; Pickering, 2006).

Memory span tests use the all-or-none scoring method (Conway et al., 2005). In this approach, if a child fails to recall a memory item or recalls it in the incorrect order, that item is marked as a failure. Such failures can lead to task termination, potentially affecting the accuracy of assessments of children's working memory abilities. Alternatively, one procedure involves presenting all test questions of various difficulty levels in random order to each child (Conway et al., 2005; Unsworth et al., 2009). However, this method may frustrate younger children or those with lower working memory capacity when they encounter difficult questions, reducing their motivation to complete the task (Gonthier et al., 2018).

This study introduced a computerized complex span task designed to measure the working memory abilities of children aged 4 to 7 years and analyzed its reliability and validity by integrating data from multiple research projects and a master's

<sup>1</sup> Department of early childhood education, National Tsing-Hua University

<sup>2</sup> Department of Educational Psychology and Counseling, National Tsing-Hua University

**Corresponding author:**

Yi Tang, Department of Educational Psychology and Counseling, National Tsing-Hua University. Email: tangyi@gapp.nthu.edu.tw

thesis. In addition, this study explored the development of working memory in Taiwanese children aged 4 to 7 years. A total of 323 typically developing children (164 boys) from North Taiwan participated in the study, and they were divided into age groups of 4, 5, 6, and 7 years. The study employed the animal-color span task, adapted from Camos and Barrouillet (2011), in which the children alternated between a storage task (animal memory) and a processing task (color naming) within a limited time. At the end of the task, the children were required to recall the memory items in the correct sequence. Scoring for the task was strict. Points were awarded only if both the color naming and animal recall sequences were correct. In addition to the animal-color span task, all participants completed four additional memory and cognitive assessments, namely the Digit Forward Span Task (DF), the Digit Backward Span Task (DB), the Working Memory Subtest of the WPPSI-IV, and the TONI-4 Nonverbal Intelligence Test.

The distribution of the children's performance on this task followed a normal curve, suggesting that the sampling was effective. Reliability analysis demonstrated strong internal consistency for the task, with a Cronbach's  $\alpha$  of 0.76 and good split-half reliability ( $r = .77$ ). Age-based analyses revealed consistent reliability across the age groups. Convergent validity analysis indicated significant positive correlations of this task with the DF task, the DB task, the working memory subtest of the WPPSI-IV, and the TONI-4 Nonverbal Intelligence Test. These findings indicate that the task effectively measures individual differences in short-term memory, working memory, and nonverbal intelligence.

Age discrimination and item response theory (IRT) were employed to evaluate the construct validity of the animal-color span task. On the basis of IRT, the task's difficulty increased progressively with the span levels, aligning with the Rosch model and demonstrating strong construct validity. Moreover, the conversion of item pass probabilities into item difficulty and a participant ability distribution yielded results that approximated a normal distribution. According to classical test theory, discrimination indices for spans 2 and 3 were high, indicating strong item discrimination at these levels. Most children successfully completed span 1 but faced limitations at spans 2 and 3. They considered span 4 to be challenging. Age-based analysis further revealed that children aged 4, 5, 6, and 7 years achieved the approximate maximum spans 1, 2, 2, and 3, respectively, which aligns with expected developmental trends. One-way ANOVA showed a significant increase in working memory capacity from ages 4 to 7 years, indicating that the task effectively captures the maturation of working memory.

These results demonstrate that the animal-color span task developed in this study, a computerized complex span task, is suitable for assessing working memory in children aged 4 to 7 years. The task demonstrates good reliability and validity and reflects developmental differences in working memory. In addition, the computerized complex span task offers several other advantages. (1) Theoretically, it aligns with the working memory model by incorporating both storage and processing tasks, controlling the cognitive load of processing tasks (Camos & Barrouillet, 2011), and measuring children's ability to resist interference and maintain attentional control during the memory retention stage (Jarrold et al., 2011). (2) Methodologically, the stimulus size, presentation time, interval time, and task termination for the storage and processing tasks are all computer-controlled, which ensures standardized test procedures and reduces errors caused by human subjectivity. (3) In terms of practical application, the standardized format enables researchers in different locations to implement consistent testing procedures (Bailey, 2012), facilitating cross-cultural and cross-study comparisons. (4) In future applications, the task's standardized features can be further developed into online or group testing formats that do not require experimenter assistance.

This study adapted only one version of the animal-color span task designed by Barrouillet et al. (2009) and Camos and Barrouillet (2011). The other two versions of the task, which involve different processing conditions, were not included. In addition, the complex span tasks developed in the aforementioned studies were originally intended to investigate the mechanisms of working memory development in children aged 5 to 7 years. This study extended the applicability of one version of the task downward to include 4-year-old children. However, additional research is needed to determine whether this task is suitable for children younger than 4 years or those older than 7 years.

**Keywords:** working memory capacity measurement, children's working memory, computerized memory tests, complex-span tasks