

# 台灣WAIS-III中文版算術、記憶廣度測驗及其組合估算 工作記憶指數在臨床上之適用性：回溯性研究

王瑋瀚<sup>1</sup> 花茂琴<sup>1,2</sup> 楊啓正<sup>1,3</sup> 朱怡娟<sup>4</sup> 鄭婷文<sup>2</sup> 葉炳強<sup>2</sup>  
邱銘章<sup>1,2</sup> 陳達夫<sup>2</sup> 黃勝堅<sup>3</sup> 陳獻宗<sup>4</sup> 徐文俊<sup>4</sup>

<sup>1</sup> 台灣大學心理學系暨研究所

<sup>2</sup> 台灣大學附設醫院神經部

<sup>3</sup> 台灣大學附設醫院外科部神經外科

<sup>4</sup> 林口長庚醫院神經內科

論文編號：07B10；初稿收件：2007年8月7日；第一次修正：2008年1月11日；第二次修正：2008年1月31日；

正式接受：2008年2月5日

通訊作者：花茂琴 106 台北市羅斯福路四段1號國立台灣大學心理學系暨研究所 (E-mail: huams@ntu.edu.tw)

受測驗題材適用性之限制，國內臨床在取得WAIS-III中文版之工作記憶指數時，仍只慣用算術與記憶廣度兩項分測驗組合進行估算，其有效性有待檢驗。此外，由於算術分測驗相當受語文理解能力干擾，而記憶廣度分測驗之順、逆序背誦廣度值在檢測腦部是否病變的有效性也仍有爭議，因此本研究目的在檢驗台灣WAIS-III中文版工作記憶指數、單一算術與記憶廣度分測驗量尺分數、及記憶廣度分測驗之各項廣度值能否有效反映工作記憶。

本研究以回溯研究方法，將179名各類中樞神經病變患者之腦造影結果當為參照效標，WAIS-III中文版之各項工作記憶相關測驗分數當為檢驗標的，進行ROC (Receiver Operating Characteristic) 分析，結果顯示僅由算術與記憶廣度分測驗組合所推估之工作記憶指數的特異性、陽性預測值、及陰性預測值皆在94%以上，但其敏感度僅有63%。單一算術與單一記憶廣度分測驗之敏感度皆為50%，但在陽性預測值上，算術分測驗為67%，記憶廣度分測驗則為86%。記憶廣度分測驗之順、逆序背誦廣度值的敏感度皆在58%以下，陰性預測值則皆在92%以上。

綜合上述，顯見WAIS-III中文版中之工作記憶指數、單一算術及單一記憶廣度分測驗、或記憶廣度分測驗之各項廣度值在區分患者有無工作記憶功能異常上皆擁有良好之區辨力，但對偵測工作記憶功能異常患者的敏感度普遍不佳，建議臨床工作者應儘量取得算術、記憶廣度、與數-字序列等三項分測驗組合所估算之工作記憶指數；若在有限條件下，臨床工作者僅能以算術與記憶廣度兩項分測驗之組合所估算之工作記憶指數來評估受試者工作記憶，應留意假陰性比例上升的問題。有鑑於本研究受限於回溯性研究方法而延伸出的問題，我們建議採取前瞻性後續研究，重新探討這個議題有其必要性。

關鍵詞：魏氏成人智力量表、工作記憶、算術測驗、記憶廣度測驗



人類記憶究竟是單一運作系統抑或多系統分工運作？一直是1960年代認知心理學者爭議的焦點。1968年，Atkinson 和Shiffrin以訊息處理觀點指出短期記憶仿如一個忙碌的工作者，負責將感官收錄之訊息暫存、編碼後，再送往長期記憶儲存（見Matlin, 1994）；一旦受損，預期個體在新事物學習及舊訊息回憶等功能運作上將全面受阻。然而，來自臨床短期記憶受損患者的研究發現卻不然，意謂短期記憶可能由一個以上之次系統分工執行（Baddeley, 1992）；1974年，Baddeley與Hitch（見Baddeley, 1992）以雙重作業技術（dual-task technique）驗證了上述推論，並嘗試以多系統運作之工作記憶模式取代傳統單一結構之短期記憶概念，Baddeley（1992, p.556）並指出：「工作記憶代表個體在執行複雜認知作業時（如語文理解、學習與推理等），將訊息暫留，同時予以分析處理的能力。它分別由語音迴路（phonological loop，負責聲韻訊息的覆誦與暫存）、視空間模版（visuospatial sketchpad，負責視覺影像之暫存與分析）及中央執行區（central executive，職司訊息處理歷程的監控與整合）等三個次系統組成」。自此，工作記憶及其相對應之腦部區域，遂為後續學者欲進一步探討之課題。

1970年代末期，核磁共振攝影（Magnetic Resonance Imaging, MRI）的問世，到近年功能性核磁共振攝影（Functional Magnetic Resonance Imaging, fMRI）技術的突破性發展，雖在認知功能與腦部區域間提供良好的時序及空間定位，但多數學者仍認為來自實驗室的腦造影研究結果僅能呈現相關推論，唯有結合神經病變患者研究，方能進一步檢驗其間的因果關係（Rorden & Karnath, 2004; Müller & Knight, 2006）。據此，Müller 與 Knight（2006）在回顧一系列與工作記憶相關之腦造影與神經病變研究後指出：大腦頂葉下方（含Brodmann's Area, 以下簡稱BA, 39, 40）及腹外側前額葉皮質區（含BA 44, 45, 47）與Baddeley工作記憶模式下的語音迴路運作有關，枕葉部分區域（含BA 19）及頂葉部分區域（含部分BA 7, 39）和視空間模版之運作有關，而背外側前額葉皮質區（含BA 46, 9, 及部分BA 8, 10）則與中央執行系統之運作有關。此外，近年有越來越多研究發現上小腦區域（superior cerebellum）中的simplex lobule、crus I與前額葉皮質區（包括背外及腹外側區域）之間、下小腦區域（inferior cerebellum）與頂葉之間存在許多神經連結路徑（Chen & Desmond, 2005; Ziemus et al., 2007）。透過上列路徑的連結，小腦與工作記憶相關大腦皮質區之間形成密切聯繫的神經網絡。因此，該網絡中任一部位的受損，都將造成工作記憶缺損

（Gottwald, Wilde, Mihajlovic, & Mehdorn, 2004; Chen & Desmond, 2005; Ziemus et al., 2007）。

工作記憶在神經結構定位上的研究發現，不僅進一步提供Baddeley工作記憶模式之相關腦部基礎，其對於神經病變患者記憶功能的評估亦有重要參考價值。目前，國內臨床最廣為使用之工作記憶評估工具仍屬魏氏成人智力量表第三版（Wechsler Adult Intelligence Scale- Third Edition，以下簡稱WAIS-III）中文版裡，構成工作記憶指數之算術（Arithmetic）、記憶廣度（Digit Span）、及數一字序列（Number-Letter Sequence）等三項分測驗，其中，新增列之數一字序列分測驗乃參照Gold、Carpenter、Randolph、Goldberg及Weinberger（1997）在其研究中編製“Letter-Number (LN) Span”作業，以口語方式，隨機、交替呈現數字、英文字母，並要求受試者按數字在前、英文字母在後之順序，從小到大，依序以口語方式作答，Gold等人（1997）認為該作業內容比記憶廣度分測驗複雜，因此更能充分反映個體在訊息暫存與操弄上的能力。考量測驗材料對國人的適用性，數一字序列分測驗在編譯之初，是以十二生肖取代原美國版本中的英文字母，然而陳榮華與陳心怡（2002）在常模建製過程中發現：仍有11%的受試者無法自動化背誦十二生肖順序，因此該測驗被列為選擇性替代測驗，換句話說，WAIS-III中文版裡的工作記憶指數，在缺少數一字序列測驗成分下，其臨床適用性究竟為何，尚待檢驗。

此外，相較前一版魏氏成人智力量表（WAIS-R），WAIS-III除了保留原有語文、作業及全量表等三項智商分數，更進一步以因素分析方式，從所有分測驗分數中抽取出語文理解、知覺組織、工作記憶、及處理速度等四項指數，該結果雖在加拿大版（Saklofske, Hildebrand, & Gorsuch, 2000）以及台灣版（陳榮華、陳心怡，2002）之WAIS-III修訂研究中獲得支持，然而在一項以法國人為受試樣本之修訂研究中卻發現：同樣的四項因素結構雖可被抽取出來，但其中算術分測驗卻無法清楚地歸類在任何一項因素指標之下（Grégoire, 2004）。Burton、Ryan、Axelrod、Schellenberger及Richards（2003）以正常成年人及各類臨床個案（包括精神及中樞神經疾病患者）進行WAIS-III之因素及效度分析研究，發現語文記憶、語文推理、視覺推理、工作記憶、處理速度、及空間建構等六項因素指標更能適切地歸納各分測驗分數，唯算術分測驗在語文推理及工作記憶兩項指標上，具有相當之因素負荷量。該結果與Wechsler（1991）曾於WAIS-R常模建置研究中提及：「算術分測驗的因素負荷量分割在語文量表及工作記憶（免於分心）指數

上」之論述相一致（見陳榮華、陳心怡，2002）。

由於WAIS-III算術分測驗是透過語句陳述，請受試者在聽完題目後以心算方式作答，其內容編製相當於小學至初中一年級程度（Kaufman & Lichtenberger, 2006），因此除了注意力專注、工作記憶功能，和教育程度息息相關的四則運算及語文能力（如一般常識、語意理解等）亦會影響受試者在該測驗上的表現（Kaufman & Lichtenberger, 1999）。回顧美國版WAIS-III之內部相關資料，亦可發現算術分測驗和語文理解指數間的相關（ $r = .66$ ）反而略高於算術分測驗和工作記憶指數間的相關（ $r = .60$ ），而這樣的差異在國內中文版WAIS-III中更為明顯（算術分測驗和語文理解指數間的相關值為 .71，和工作記憶指數間的相關值為 .62）。反觀同為工作記憶指數下的記憶廣度與數-字序列分測驗，皆無此情況（陳榮華、陳心怡，2002）。此外在試探性因素分析資料中，美國版WAIS-III之工作記憶指數下的三項分測驗，記憶廣度分測驗擁有最高因素負荷量：0.71，算術分測驗最低：0.51；國內WAIS-III中文版亦然，記憶廣度分測驗在工作記憶指數中擁有最高因素負荷量：0.72，算術分測驗則擁有最低因素負荷量：0.37（陳榮華、陳心怡，2002）。據此，我們預期記憶廣度分測驗在工作記憶評估上，受語文或學習成就等因素干擾較少，因此相對具有較佳之檢測力。

記憶廣度分測驗在心智銜鑑上的應用雖有數十年歷史，然該測驗之順序數字背誦與逆序數字背誦是否反映相同之心理運作機制？目前研究仍有爭議（Wilde, Strauss, & Tulsy, 2004）。Banken 於1985年、Kaplan於1991年相繼指出相較順序數字背誦，逆序數字背誦需要更多工作記憶的支援，受老化或神經病變等因素的影響相對較敏感，因此在解釋上，順序及逆序背誦應被視為兩項不同功能指標（見Lezak, Howieson, Loring, Hannay, & Fischer, 2004），Kaplan甚至進一步指出將最長順序數字背誦廣度值（Forward Span Score, 以下簡稱FSS）減去最長逆序數字背誦廣度值（Backward Span Score, 以下簡稱BSS），所得到的差值（以下簡稱FSS - BSS）最能夠反映老化對注意力及工作記憶造成的影響，Lezak等人（2004）則進一步指出腦部損傷患者常見其BSS值較FSS值少3個數字位數（含）以上（意即FSS - BSS  $\geq 3$ ）；然而Wilde等人（2004）在分析原版WAIS-III之常模資料後發現：正常老年群體與神經疾患群體在順序及逆序數字背誦上的得分，呈現相同斜率之下降走勢，而類似結果亦在Myerson、Emery、White及Hale（2003）的研究中發現，因此，上述學者認為順序及逆序數字背誦能力對老化或神經病變等因素帶來的影響同樣敏感，

顯見該議題尚待更多研究釐清。

基於上述，本研究目的在藉各類神經病變患者之腦造影資料，檢驗台灣WAIS-III中文版中：（1）僅由算術與記憶廣度兩項分測驗組合所估算之工作記憶指數能否有效反映工作記憶；（2）單一算術分測驗量尺分數（the scaled score of the Arithmetic subtest, ARIs）、記憶廣度分測驗量尺分數（the scaled score of the Digit Span subtest, DSs）能否有效代表工作記憶指數反映工作記憶；以及（3）記憶廣度分測驗之最長順序背誦廣度值、最長逆序數字背誦廣度值、最長順序與逆序數字背誦廣度差值、及該三項數值之各種組合在工作記憶評估上的檢測力。

## 研究方法

### （一）受試者

本研究以回溯方式（retrospective）取得2005年7月～2007年2月期間，來自台灣大學附設醫院神經內、外科及林口長庚醫院神經內科，具腦造影或相關神經學檢查資料之各類神經病變患者共179名，平均年齡約49歲，平均受教育年數約12年（詳見表1），其中包括失憶型輕度認知功能損傷（amnesic Mild Cognitive Impairment, aMCI）患者12名、早期阿茲海默型失智症（Early stage Alzheimer's Disease, Early stage AD）患者13名、阿茲海默型失智症（Alzheimer's Disease, AD）患者9名、巴金森氏症（Parkinson's Disease, PD）患者11名、帕金森氏症候群（Parkinsonism）患者9名、額-顳葉型失智症（Frontotemporal Dementia, FTD）患者3名、亨汀頓氏症（Huntington's Disease, HD）患者3名、血管型失智症（Vascular Dementia, VaD）患者15名、正常腦壓水腦症（Normal Pressure Hydrocephalus, NPH）患者1名、腦部創傷（Traumatic Brain Injury, TBI）患者28名、癲癇（Epilepsy）患者21名、腦瘤（Brain Tumor）患者5名、腦神經中毒（Neurointoxication）患者1名、威爾森氏症（Wilson's Disease）患者1名、神經性梅毒（Neurosyphilis）患者1名、腦血管病變（Cerebral Vascular Disease）患者8名、缺氧性腦病變（Hypoxic Encephalopathy）患者3名、腦炎（Encephalitis）患者8名、神經髓鞘病變（Demyelinating Disease）患者2名、大腦白質病變（Leukoencephalopathy）患者2名、小腦脊髓運動失調症候群（Spinocerebellar Ataxia, SCA）患者3名、肌萎縮性脊髓側索硬化症（Amyotrophic Lateral Sclerosis, ALS）患者1名、貝克型肌肉萎縮症（Becker Muscular Dystrophy）患者1名、酒精中毒（Alcoholic）患

表1 研究樣本之人口學資料

	平均值	標準差
性別（男性：122人；女性：57人）		
年 齡（範圍：18歲~85歲）	49.68歲	20.32歲
教育年數（範圍：6年~18年）	12.56年	2.73年

者2名、進行性上眼神經核麻痺症（Progressive Supranuclear Palsy, PSP）患者1名、多發性硬化症（Multiple Sclerosis, MS）患者1名、粒腺體病變症候群（Mitochondrial Encephalomyopathy Lactic Acidosis, and Strokelike, MELAS）患者1名、及其它複合型神經疾病患者13名。每位患者皆由神經內、外科專科醫師整合多項評估報告後（如神經心理衡鑑、大腦造影、相關神經學檢查等），確立診斷。

## （二）研究工具

本研究主要研究工具為陳榮華與陳心怡編譯之中文版WAIS-III。WAIS-III中文版共含十四項分測驗，每項分測驗的量尺平均數皆為10，標準差為3。為求能更精細且結構化地呈現受試者認知功能，該版測驗以因素分析方式，從所有分測驗分數中抽取四項因素結構，分別代表4種認知功能，包括：（1）語文理解指數—由詞彙（Vocabulary）、類同（Similarities）和常識（Information）三項分測驗之量尺分數合計而成；（2）知覺組織指數（Perceptual Organization Index, POI）—由圖畫補充（Picture Completion）、圖形設計（Block Design）與矩陣推理（Matrix Reasoning）三項分測驗之量尺分數合計而成；（3）工作記憶指數（Working Memory Index, WMI）—由算術（Arithmetic）、記憶廣度（Digit Span）及數字序列（Letter-Number Sequencing）三項分測驗之量尺分數合計而成；及（4）處理速度指數（Processing Speed Index, PSI）—由數字-符號替代（Digit Symbol-Coding）和符號尋找（Symbol Search）兩項分測驗之量尺分數合計而成，每項指數之量尺平均數皆為100，標準差為15（陳榮華、陳心怡，2002）。誠如文獻回顧所述，由於WAIS-III中文版之數-字序列被列為替代性分測驗，每位神經病變患者在受測當時皆未施予該分測驗，因此依據研究目的，本研究僅擷取其中語文理解指數、工作記憶指數、算術、與記憶廣度分測驗分數進行分析。

此外，本研究亦將針對患者之腦造影或相關神經學檢查結果進行分析，包括核磁共振攝影

（Magnetic Resonance Imaging, MRI）、電腦斷層掃描（Computer Tomography, CT）、正子斷層掃描（Positron Emission Tomography, PET）、單光子放射電腦斷層掃描（Single-Photon Emission Computed Tomography, SPECT）等資料，所有造影資料皆由影像醫學部主治醫師予以判讀。

## （三）進行步驟

由修習臨床神經心理實習課程之臨床心理學研究生，在具有證照之臨床心理師及任課教授督導下，針對來院就診轉介之個案進行神經心理衡鑑（包括WAIS-III中文版、語言、空間建構及執行功能等測驗），在與臨床督導及神經內、外科專科醫師充分討論並確立診斷後，隨即將完成之衡鑑報告歸檔存置。研究者再根據研究目的，以回溯方式擷取台灣版WAIS-III測驗資料進行分析。

## （四）統計分析

本研究係利用SPSS視窗第10版統計套裝軟體及Microsoft Excel軟體進行各項統計分析，並以Receiver Operating Characteristic（簡稱ROC）法進行各項分析。ROC分析法早在1950年代即被醫學界廣泛應用在疾病診斷之適切性考驗（diagnostic test），方法如下（請參見表2）：首先以二元分類方式（binary classification）將某疾病之臨床診斷結果（是為參照效標）與神經心理功能檢查結果（是為檢驗標的）定義為陽性（positive，代表罹病或異常）或陰性（negative，代表未罹病或正常），形成如表2之四種預測組合，分別是：（1）真陽性（True Positive, TP）：代表神經心理功能檢查結果與臨床疾病診斷結果一致為異常（罹病）的人數；（2）真陰性（True Negative, TN）：代表神經心理功能檢查結果與臨床疾病診斷結果一致為正常的人數；（3）假陽性（False Positive, FP）：代表神經心理功能檢查結果異常，但臨床疾病診斷結果卻為正常之人數；及（4）假陰性（False Negative, FN）：代表神經心理功能檢查結果

表2 ROC分析法在診斷測試 (diagnostic test) 上的應用

		臨床疾病診斷結果 (參照效標)		
		陽性 (positive)	陰性 (negative)	
功能檢查結果 (檢驗標的)	陽性 (positive)	① 真陽性人數	② 假陽性人數	陽性預測值 <sup>c</sup> (PPV)
	陰性 (negative)	③ 假陰性人數	④ 真陰性人數	陰性預測值 <sup>d</sup> (NPV)
		敏感度 <sup>a</sup> (Sensitivity)	特異性 <sup>b</sup> (Specificity)	

註：<sup>a</sup>敏感度 = 真陽性人數 / (真陽性人數 + 假陰性人數)

<sup>b</sup>特異性 = 真陰性人數 / (假陽性人數 + 真陰性人數)

<sup>c</sup>陽性預測值 = 真陽性人數 / (真陽性人數 + 假陽性人數)

<sup>d</sup>陰性預測值 = 真陰性人數 / (假陰性人數 + 真陰性人數)

正常，但臨床疾病診斷結果卻為罹病之人數 (Pepe, 2003)。其次根據上列累積人數，分別求算 (求算方式請參見表2) 下列指標 (Altman & Bland, 1994a, 1994b; Swets, 1988)：(1) 敏感度 (Sensitivity)：代表在所有臨床疾病診斷為罹病的患者中 (亦即分母)，神經心理功能檢查結果亦呈現異常 (亦即分子) 的比例；(2) 特異性 (Specificity)：在所有臨床疾病診斷為未罹病的個案中 (亦即分母)，神經心理功能檢查結果亦呈現正常 (亦即分子) 的比例；(3) 陽性預測值 (Positive Predictive Value, PPV)：代表在所有神經心理功能檢查結果異常的患者中 (亦即分母)，臨床疾病診斷亦確實為罹病 (亦即分子) 的比例；(4) 陰性預測值 (Negative Predictive Value, NPV)：代表在所有神經心理功能檢查結果正常的個案中 (亦即分母)，臨床疾病診斷亦確實為未罹病 (亦即分子) 的比例；最後再依據下列分類標準 (Swets, 1988) 逐一判辨各項指標之優劣，指標值落在：1.0 ~ 0.9：為極佳 (excellent)；0.9 ~ 0.8：為良好 (good)；0.8 ~ 0.7：為尚可 (fair)；0.7 ~ 0.6：為不佳 (poor)；0.6 ~ 0.5：為極差 (fail)。經由上列四項指標值之判讀，遂可檢視測驗工具在某疾病篩檢上的有效性。

受回溯研究方法的限制，研究者已無法取得現有資料以外之工作記憶效標，依據本研究目的，上述「臨床疾病診斷結果 (參照效標)」將由「工作記憶相關腦部區域之異常與否」代表：若腦損傷部位落於

大腦左或右前額葉背外側、左或右前額葉腹外側、左或右頂葉下方區域、或上小腦、下小腦等任一處者 (Müller & Knight, 2006)，將被定義為「工作記憶相關腦部區域受損」，反之，腦損傷部位落於上述各區域以外，將定義為「工作記憶相關腦部區域未受損」。

「神經心理功能檢查結果 (檢驗標的)」的部分，為避免工作記憶功能檢查結果可能因語文理解因素干擾受到低估 (由於算術分測驗緣故)，首先需將語文理解指數和工作記憶指數同時落於常模平均數兩個標準差 (含) 以下者剔除。研究目的 (1) 中，WAIS-III 中文版中之工作記憶指數即為檢驗標的，依據 WAIS-III 分數分類系統 (陳榮華、陳心怡, 2002) 及 Lezak 等人 (2004) 建議之臨床常用切點分數 (cut-point)，工作記憶指數若落於常模平均數兩個標準差 (含) 以下者 (即 WMI ≤ 70)，將被定義為功能檢查結果異常，反之，定義為功能檢查結果正常。研究目的 (2) 中，單一算術、記憶廣度分測驗之量尺分數即為檢驗標的，若量尺分數落於常模平均數兩個標準差 (含) 以下者 (即 Scaled Score ≤ 4)，將被定義為工作記憶功能異常，反之，定義為工作記憶功能正常。研究目的 (3) 中，檢驗標的則包括記憶廣度分測驗之最長順序 (FSS) 與逆序 (BSS) 數字背誦廣度值、最長順序與逆序數字背誦廣度差值 (FSS - BSS)、及上列各值的組合，其中，最長順序數字背誦廣度值、或最長逆序數字背誦廣度值若落在常模平均數兩個標

準差（含）以下者（各年齡層之平均數與標準差請參閱WAIS-III中文版指導手冊p.383，表乙6），代表功能檢查結果異常，反之，代表功能檢查結果正常；至於最長順序與逆序數字背誦廣度差值則根據Lezak等人（2004）的報告，將FSS - BSS  $\geq 3$ 者定義為功能檢查結果異常，反之則為功能檢查結果正常。其餘各組合包括：記憶廣度分測驗量尺分數與最長順序數字背誦廣度值組合（DSs & FSS）、記憶廣度分測驗量尺分數與最長逆序數字背誦廣度值組合（DSs & BSS）、記憶廣度分測驗量尺分數與最長順序、逆序數字背誦廣度差值組合（DSs & FSS-BSS）、最長順序數字背誦廣度值與最長逆序數字背誦廣度值組合（FSS & BSS）、最長順序數字背誦廣度值與最長順序、逆序數字背誦廣度差值組合（FSS & FSS-BSS）、最長逆序數字背誦廣度值與最長順序、逆序數字背誦廣度差值（BSS & FSS-BSS）、及最長順序數字背誦廣度值與最長逆序數字背誦廣度值與最長順序、逆序數字背誦廣度差值組合（FSS & BSS & FSS-BSS）中，每一項目數值皆須同時落於切點值以下，該組合方能定義為功能檢查結果異常，反之，則定義為功能檢查結果正常。

## 結果

表3顯示179名神經病變患者之WAIS-III中文版測驗量尺分數平均值，整體而言，其各智商指數及測驗量表平均術皆落於常模平均數以下0~1個標準差之間。其中，可發現最長順序與逆序數字廣度差異平均值（即FSS - BSS）為2.66，未達Lezak報告之腦損傷患者常見差值：3位數。

表4顯示179名神經病變患者中，總計有26人之工作記憶指數落於常模平均數兩個標準差（含）以下（即WMI  $\leq 70$ ，代表功能檢查結果異常），其中有10人之語文理解指數和工作記憶指數同時落於常模平均數兩個標準差（含）以下，為避免工作記憶功能檢查結果可能因語文理解因素干擾受到低估，在以下分析中，該10名個案資料將予以剔除。

表5顯示台灣WAIS-III中文版中，僅由算術與記憶廣度分測驗組合所推估之工作記憶指數，雖具極佳之特異性（specificity = 99%）、陽性預測值（PPV = 94%）、及陰性預測值（NPV = 94%），但其敏感度（sensitivity = 63%）卻不佳。

表6顯示單一算術分測驗量尺分數在工作記憶評估上擁有極佳之特異性（specificity = 96%）與陰性預測值（NPV = 92%），但其敏感度（sensitivity = 50%）與陽性預測值（PPV = 67%）皆不佳。

表7顯示單一記憶廣度分測驗量尺分數在工作記憶評估上擁有極佳之特異性（specificity = 99%）與陰性預測值（NPV = 92%），以及良好之陽性預測值（PPV = 86%），然而其敏感度（sensitivity = 50%）卻不佳。此外，綜合表6及表7結果，可見記憶廣度分測驗之陽性預測值（PPV = 86%）明顯優於算術分測驗（PPV = 67%）。

記憶廣度分測驗之各廣度值（組合）的ROC分析結果如表8。整體而言，各廣度值（組合）的敏感度均不佳（sensitivity皆落於58%以下），而除了最長順序與逆序數字背誦廣度差值（FSS-BSS），其餘各廣度值（組合）皆具極佳之特異性（其值皆落在97%以上）。此外，各廣度值（組合）皆擁有良好之陰性預測值（其值介於86%~89%），至於在陽性預測值上，各廣度值（組合）分佈差異甚大，範圍從15%~83%。

## 討論

本研究目的主要在藉各類神經病變患者之腦造影或相關神經學檢查資料，檢驗台灣WAIS-III中文版中：（1）僅由算術與記憶廣度兩項分測驗組合所估算之工作記憶指數能否有效反映工作記憶；（2）單一算術、記憶廣度分測驗量尺分數能否有效代表工作記憶指數反映工作記憶；以及（3）記憶廣度分測驗之最長順序背誦廣度值、最長逆序數字背誦廣度值、最長順序與逆序數字背誦廣度差值、及該三項數值之各種組合在工作記憶評估上的檢測力。研究結果發現僅由算術與記憶廣度分測驗組合所估算之工作記憶指數，雖對區分患者有無工作記憶功能異常具極佳之區辨力（特異值為94%），但對偵測工作記憶功能異常患者的敏感度卻不佳。此外，單一記憶廣度與單一算術分測驗之敏感度雖不佳，但兩者皆具極佳之特異性及陰性預測值，同時記憶廣度分測驗更較算術分測驗擁有良好之陽性預測值。最後，記憶廣度分測驗之各廣度值的ROC分析結果發現：各廣度值（組合）雖具有極佳之陰性預測值，但對偵測工作記憶功能異常患者的敏感度卻不佳，而除了最長順序與逆序數字背誦廣度差值（FSS-BSS），其餘各廣度值（組合）皆具極佳之特異性。以下將分別針對上列結果進行討論：

### （一）WAIS-III中文版工作記憶指數之臨床適用性

回顧Baddeley 與 Hitch 1974年提出之工作記憶概念，乃是個體透過心像或聲韻方式將訊息暫留，並予以分析處理的能力（見Baddeley, 1992），依據該定

表3 179名患者之WAIS-III中文版測驗結果

WAIS-III中文版各分測驗項目	平均量尺分數	標準差
全量表智商	90.93	18.14
語文量表智商	93.45	18.31
作業量表智商	88.51	18.99
詞彙分測驗	9.39	3.86
類同分測驗	8.55	3.36
常識分測驗	9.19	3.29
語文理解指數	94.75	17.95
圖畫補充分測驗	8.29	3.19
圖形設計分測驗	8.17	3.51
矩陣推理分測驗	8.66	3.52
知覺組織指數	90.69	17.61
算術分測驗	8.44	3.25
記憶廣度分測驗	8.97	3.43
最長順序數字背誦廣度	7.16	1.45
最長逆序數字背誦廣度	4.50	1.50
最長順序與逆序數字廣度差值	2.66	1.37
工作記憶指數	92.08	17.40
數字—符號替代分測驗	7.14	3.86
符號尋找分測驗	7.48	3.65
處理速度指數	85.86	18.91
理解分測驗	8.92	3.81
連環圖系分測驗	8.67	3.61

表4 工作記憶指數落於常模平均數兩個標準差（含）以下者之腦造影檢查結果

腦造影檢查結果	人數
僅工作記憶指數落於常模平均數兩個標準差（含）以下	
腦受損區域涵蓋前額葉背外側或腹外側區域、頂葉下方、上小腦區域或下小腦區域	15
腦受損區域未涵蓋前額葉背外側或腹外側區域、頂葉下方、上小腦區域或下小腦區域等處	1
工作記憶和語文理解指數同時落於常模平均數兩個標準差（含）以下	
腦受損區域涵蓋前額葉背外側或腹外側區域、頂葉下方、上小腦區域或下小腦區域	8
腦受損區域未涵蓋前額葉背外側或腹外側區域、頂葉下方、上小腦區域或下小腦區域等處	2
總計	26

表5 工作記憶指數之ROC分析結果

	涉及工作記憶相關腦部區域		
	受損	未受損	
工作記憶指數落於常模平均數兩個標準差（含）以下者（WMI $\leq$ 70）	真陽性人數 = 15	假陽性人數 = 1	陽性預測值 = 94%
工作記憶指數落於常模平均數兩個標準差以上者（WMI $>$ 70）	假陰性人數 = 9	真陰性人數 = 144	陰性預測值 = 94%
	敏感度 = 63%	特異性 = 99%	

表6 單一算術分測驗量尺分數之ROC分析結果

	涉及工作記憶相關腦部區域		
	受損	未受損	
算術分測驗量尺分數落於常模平均數兩個標準差（含）以下者（ARIs $\leq$ 4）	真陽性人數 = 12	假陽性人數 = 6	陽性預測值 = 67%
算術分測驗量尺分數落於常模平均數兩個標準差以上者（ARIs $>$ 4）	假陰性人數 = 12	真陰性人數 = 139	陰性預測值 = 92%
	敏感度 = 50%	特異性 = 96%	

表7 單一記憶廣度分測驗量尺分數之ROC分析

	涉及工作記憶相關腦部區域		
	受損	未受損	
記憶廣度分測驗量尺分數落於常模平均數兩個標準差（含）以下者（DSs $\leq$ 4）	真陽性人數 = 12	假陽性人數 = 2	陽性預測值 = 86%
記憶廣度分測驗量尺分數落於常模平均數兩個標準差以上者（DSs $>$ 4）	假陰性人數 = 12	真陰性人數 = 143	陰性預測值 = 92%
	敏感度 = 50%	特異性 = 99%	



表8 記憶廣度分測驗各廣度值（組合）之ROC分析結果

各分數項目（組合）	敏感度	特異性	陽性預測值	陰性預測值
最長順序數字背誦廣度值（FSS）	29%	97%	58%	89%
最長逆序數字背誦廣度值（BSS）	21%	98%	63%	88%
最長順序、逆序數字背誦廣度差值 （FSS-BSS）	58%	43%	15%	86%
記憶廣度分測驗量尺分數與最長順序 數字背誦廣度值組合（DSs & FSS）	29%	99%	78%	89%
記憶廣度分測驗量尺分數與最長逆序 數字背誦廣度值組合（DSs & BSS）	17%	99%	80%	88%
記憶廣度分測驗量尺分數與最長順序、逆序 數字背誦廣度差值組合（DSs & FSS-BSS）	21%	99%	83%	88%
最長順序數字背誦廣度值與最長逆序 數字背誦廣度值組合（FSS & BSS）	13%	99%	60%	87%
最長順序數字背誦廣度值與最長順序、逆序 數字背誦廣度差值組合（FSS & FSS-BSS）	4%	99%	33%	86%
最長逆序數字背誦廣度值與最長順序、逆序 數字背誦廣度差值（BSS & FSS-BSS）	13%	99%	60%	87%
最長順序數字背誦廣度值與最長逆序數字背 誦廣度值與最長順序、逆序數字背誦廣 度差值組合（FSS & BSS & FSS-BSS）	4%	99%	33%	86%

義，國內WAIS-III中文版裡的算術與記憶廣度分測驗皆以口語方式施測，並要求受試者以心算及口語方式作答，依據該兩項分測驗之作答結果，可分別取得其量尺分數，予以合計後，即可估算受試者之工作記憶指數。然而本研究結果發現僅由算術與記憶廣度分測驗組合所估算之工作記憶指數，對工作記憶功能異常者的偵測敏感度並不佳（sensitivity = 63%），其可能原因有：（1）依據Baddeley 與 Hitch 1974年之工作記憶架構（見Baddeley, 1992），部分學者認為記憶廣度分測驗所涉及之視空間（心像）暫存能力較少，例如Gold等人（1997）在其研究中提出“LN Span task”，認為該作業內容比記憶廣度分測驗更能充分反映個體在訊息暫存與操弄上的能力；而Crowe（2000）分析102位大學生在11項神經心理測驗上的表現後指出：WAIS-III之算術、記憶廣度與數-字序列分測驗皆能測得個體對聽覺訊息的工作記憶能力，只是相較於算術與記憶廣度分測驗，數-字序列分測驗尚與個體視空間工作記憶（visual spatial working memory）能力具密切關係。因此，僅由算術與記憶廣度分測驗組合所

估算之工作記憶指數，由於涉及較少之視空間工作記憶成分，致其可能高估（over-estimation）患者之工作記憶功能，進而使假陰性人數上升。（2）回顧過去文獻，Baddeley 與 Hitch 1974年提出之工作記憶模式中（見Baddeley, 1992），各次系統究竟是由單側半腦或雙側半腦主負責，尚無明確定論（Müller & Knight, 2006），因此本研究在界定工作記憶相關腦部區域是否受損（參照效標）時，只要患者在工作記憶相關之任一腦區受損（不論左或右側半腦），遂將其界定為工作記憶相關腦部區域受損。然而就算術分測驗為例，Fehr、Code 及 Herrmann（2007）以功能性核磁造影技術（fMRI）發現受試者在進行心算作業時，除了右側precuneus區域，其雙側額葉內側、與雙側背外側前額葉皮質區皆會被活化，因此，若患者僅是單側背外側前額葉區域受損，由於其他未受損腦區可能產生之代償作用（compensation），其心算能力未必就此完全喪失。換句話說，藉由算術測驗來評估患者工作記憶功能時，患者之雙側額葉內側、或雙側背外側前額葉區域皆須受損，才可能導致心算能力喪失，並

進而反映出工作記憶功能之異常。因此，在參照效標的界定上，本研究將單側或雙側工作記憶相關腦區受損患者皆定義為工作記憶相關腦部區域受損，實難以完全確定其所反映的全是工作記憶功能異常者（意即有over-inclusion之可能）。綜合上述，由於檢驗標的（工作記憶指數）之“over-estimation”、以及參照效標之“over-inclusion”兩項可能性，將使真陽性人數（hit rate）下降，進而拉低敏感值。（3）依據表5，研究者進一步檢視其中九名工作記憶指數70分以上，但工作記憶相關腦區卻受損（即假陰性）之患者，發現多數是非神經退化性疾病患者（五名是腦部創傷患者，一名為腦血管病變患者），意謂該群患者的認知功能可能因為各種臨床處遇（例如藥物治療、手術治療等）、或腦部功能代償機制，隨時間日漸改善。因此，接續比對該六位患者之腦造影與神經心理衡鑑日期後，確實發現其中五名患者之神經心理衡鑑，是在取得腦造影資料半年後（以上）才進行，換句話說，該群患者在初接受腦造影時，雖被發現其工作記憶相關腦區受損，但隨各項醫療處置之介入，其整體認知功能可能逐漸獲得改善，因而在半年後的工作記憶功能檢查中，表現正常。（4）最後，在判斷工作記憶指數異常與否時，所選取之切點分數亦可能攸關敏感度等數值的高低，例如，若將切點分數提高，則落在該切點分數以下（亦即工作記憶功能檢查結果異常）的人數將會增加，而真陽性人數亦可能因此隨之增加，進而使敏感度獲得提升（請參見表2）。因此，是否有更佳之切點分數？研究者曾嘗試將該切點分數分別界定在平均數以下1.65個標準差（約含括常態分配5個百分位數以下範圍）及1.5個標準差（約含括常態分配6個百分位數以下範圍），並依據本研究對檢驗標的之界定規則，將語文理解指數亦同時落在該切點分數以下的個案剔除，ROC分析結果發現：相較於本研究選取之切點分數（平均數以下2個標準差），當切點分數界定在平均數以下1.65及1.5個標準差時，敏感度並未如預期有所提升（反而略微下降）；同時由於假陽性（false positive）人數的增加，使陽性預測值大幅下降。因此，研究者最後仍採WAIS-III分數分類系統及Lezak等人（2004）之建議，將切點分數界定在平均數以下2個標準差。

綜合上述，我們提醒臨床工作者在使用僅由算術與記憶廣度兩項分測驗估算之工作記憶指數評估受試者工作記憶時，需留意假陰性比例（false negative rate）可能上升（意即增加犯第二類型錯誤的可能性），而為能提升工作記憶指數在工作記憶評估上之適切性及完整性，我們建議臨床工作者應盡量取得算術、記憶廣度、與數-字序列等三項分測驗估算而得之工作記憶指數。

此外，隨著教育水準提高，年輕族群對英文字母的熟悉度亦隨之提升，若面對不熟悉十二生肖順序之年輕族群，我們建議臨床工作者亦可將原文版WAIS-III將數-字序列測驗、算術與記憶廣度分測驗合併在一起，進行常模的建置，以利臨床衡鑑之用。

## （二）單一算術分測驗與單一記憶廣度分測驗之臨床適用性比較

回顧組成工作記憶指數之三項分測驗中，算術分測驗題目之編製是以小學數學科目內容為基礎，透過語句陳述方式施測，因此相當受到個案四則運算及語文成就（如一般常識、語意理解等）等混淆因素的影響。根據原版及中文版WAIS-III之常模資料，即可發現算術分測驗和語文理解指數間的內部相關，皆高於算術分測驗和工作記憶指數間的相關，同時在工作記憶指數上之因素負荷量，算術分測驗皆是最低（陳榮華、陳心怡，2002）；此外，Grégoire（2004）、Ryan與Paolo（2001）、Burton等人（2003）的研究則皆發現算術分測驗無法明確歸類在工作記憶指數下。因此，我們預期記憶廣度分測驗在工作記憶評估上，由於受語文或學習成就等混淆影響因素較少，應相對具有較佳之檢測力。研究結果發現雖然單一記憶廣度與單一算術分測驗之敏感度雖不佳（皆為50%），且兩者皆具極佳之特異性及陰性預測值（皆在92%以上），但是記憶廣度分測驗（ $PPV = 86\%$ ）卻較算術分測驗（ $PPV = 67\%$ ）擁有良好之陽性預測值。

此外，根據陳心怡、花茂琴及朱建軍（2007）的研究顯示，就一般正常成年人而言，算術分測驗所需之平均施測時間約六分半鐘，記憶廣度分測驗則約需五分鐘，在臨床實務工作上，常需面對認知功能缺損、高齡、或教育貧乏之族群，上述所需施測時間往往會被拉長，此外對於教育程度偏低者而言，算術分測驗更可能受四則運算或語文理解能力等因素干擾，影響其表現。因此，臨床上，在數-字序列分測驗無法適用（即受試者無法自動化背誦十二生肖順序）、及評估時間受限情況下，我們建議臨床工作者優先採用陽性預測值（PPV）相對較佳、施測所需時間相對較短之記憶廣度分測驗，唯仍需留意偵測工作記憶功能異常時，假陰性比例的上升。

## （三）記憶廣度分測驗之各廣度值（組合）的適用性

雖然Kaplan、Banken等學者（見Lezak, 2004）指出逆序數字背誦由於需要更多工作記憶的支援，對



老化或神經病變等因素的影響相對較敏感，但近年 Myerson 等人 (2003)、Wilde 等人 (2004) 的研究中卻有不一致發現。亦即 FSS、BSS、及 FSS - BSS 等三項數值，何者對老化或神經病變等因素的影響較敏感，目前仍有爭議 (Wilde, et al., 2004)。本研究針對各類腦部病變患者，利用 ROC 分析方法，檢驗記憶廣度分測驗之各廣度值 (及其各種組合) 的適用性，結果發現所有廣度值對腦部病變患者工作記憶功能異常的偵測力皆不佳 (敏感度皆在 58% 以下)，顯見該三項數值之適用性尚待更多研究釐清。此外，本研究結果亦發現 FSS - BSS 值在偵測工作記憶功能異常之敏感度 (58%) 及特异性 (43%) 上皆偏低。歸咎其可能原因是研究者以  $FSS - BSS \geq 3$  來定義工作記憶功能異常時，並未將 FSS 及 BSS 需落在正常值範圍之設定當為前提條件，以致於未包括實際上工作記憶功能已呈現異常的腦損傷患者，例如，患者之 FSS 值為 4，BSS 值為 3，儘管兩者差值小於 3 位數，然而其 FSS 值與 BSS 值卻分別低於 Lezak 等人 (2004) 所報告之正常值範圍 (一般成人 FSS 值之正常範圍約  $6 \pm 1$  個數字位數，BSS 值則較 FSS 值少 1 位數)，因此，有待後續進一步探討。

綜合上述，研究結果發現僅由算術與記憶廣度分測驗組合所估算之工作記憶指數，在工作記憶功能異常的偵測上雖具有極佳之特異值 (94%)，但其敏感度卻不佳 (63%)。其中，單一記憶廣度與單一算術分測驗之敏感度雖不佳，但兩者皆具極佳之特异性及陰性預測值，同時記憶廣度分測驗更較算術分測驗擁有良好之陽性預測值。建議臨床工作者應盡量取得算術、記憶廣度、與數-字序列等三項分測驗估算而得之工作記憶指數，或使用臨床上已具良好信、效度之工作記憶測驗，然而若在測驗題材不適用及評估時間有限等因素限制下，臨床工作者僅能以算術與記憶廣度兩項分測驗之組合所估算之工作記憶指數來評估受試者工作記憶時，應留意假陰性比例上升的可能性。本研究仍有下列限制：(1) 在 16 名工作記憶功能檢查結果異常患者中，仍有 1 人之工作記憶相關腦部區域並未受損 (約佔 6%，即為假陽性比例)，進一步檢視該筆造影資料，發現該名患者之造影結果僅呈現輕微腦部萎縮 (mild brain atrophy)，沒有明確損傷定位，但卻被認定為工作記憶相關腦部區域未受損。因此針對該類患者，除了結構性造影檢查，有必要輔以功能性造影檢查 (例如 PET 或 SPECT)，以改善腦損傷區域之定位。(2) 本研究係以回溯研究方法進行資料蒐集與分析，已無法取得現有工作記憶指數以外之測驗結果當為參照效標，因此，在參照效標的界定上，是以腦造影資料為主，但由於少部分患者之電子影像檔已

不存在醫院資料庫，因此，僅能就影像醫學部單一醫師的判讀報告進行分析，其精確性與客觀性皆受到限制。有鑑於此，在參照效標的界定上，我們建議後續研究者可透過前瞻性研究 (prospective study) 方法予以改善，包括請兩位 (以上) 之影像醫學部醫師協助造影資料的判讀，以提升判讀結果之客觀性與準確性，此外，研究者在進行 WAIS-III 工作記憶相關測驗的同時，亦應給予其它具良好信、效度之工作記憶測驗 (例如 N-Back task)，輔以腦造影資料，來當為參照效標。

## 參考文獻

- 陳榮華、陳心怡 (2002)。《魏氏成人智力量表第三版 (中文版) 指導手冊》。台北：中國行為科學社。
- 陳心怡、花茂琴、朱建軍 (2007)。〈台灣 WAIS-III 二分測驗組合之簡式版本建立研究〉。《測驗年刊》，54，305-330。
- Altman, D. G., & Bland J. M. (1994a). Diagnostic tests 1: Sensitivity and specificity. *British Medical Journal*, 308, 1552.
- Altman, D. G., & Bland J. M. (1994b). Diagnostic tests 2: Predictive values. *British Medical Journal*, 309, 102.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556-559.
- Burton, D. B., Ryan, J. J., Axelrod, B. N., Schellenberger, T., & Richards, H. M. (2003). A confirmatory factor analysis of the WMS-III in a clinical sample with crossvalidation in the standardization sample. *Archives of Clinical Neuropsychology*, 18, 629-641.
- Chen, S. H., & Desmond, J. E. (2005). Cerebrocerebellar networks during articulatory rehearsal and verbal working memory tasks. *Neuroimage*, 24, 332-338.
- Crowe, S. F. (2000). Does the letter number sequencing task measure anything more than digit span? *Assessment*, 7, 113-117.
- Fehr, T., Code, C., & Herrmann, M. (2007). Common brain regions underlying different arithmetic operations as revealed by conjunct fMRI-BOLD activation. *Brain Research*, 1172, 93-102.
- Gold, J. M., Carpenter, C., Randolph, C., Goldberg, T. E., & Weinberger, D. R. (1997). Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Archives of general*

- psychiatry*, 54, 159-165.
- Gottwald, B., Wilde, B., Mihajlovic, Z., & Mehdorm, H. M. (2004). Evidence for distinct cognitive deficits after focal cerebellar lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, 75, 1524-1531.
- Grégoire, J. (2004). Factor structure of the French version of the Wechsler Adult Intelligence Scale-III. *Educational and Psychological Measurement*, 64, 463-474.
- Kaufman, A. S., & Lichtenberger, E. O. (1999). *The Essentials of WAIS-III Assessment*, New York: John Wiley & Sons.
- Kaufman, A. S., & Lichtenberger, E. O. (2006). *Assessing adolescent and adult intelligence* (3rd ed.), New York: John Wiley & Sons.
- Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2004). *Neuropsychological Assessment* (4th ed.). New York: Oxford University Press.
- Matlin, M. W. (1994). *Cognition* (3rd ed.). Harcourt Brace.
- Müller, N. G., & Knight, R. T. (2006). The functional neuroanatomy of working memory: Contributions of human brain lesion studies. *Neuroscience*, 139, 51-58.
- Myerson, J., Emery, L., White, D. A., & Hale, S. (2003). Effects of age, domain, and processing demands on memory span: Evidence for differential decline. *Aging Neuropsychology and Cognition*, 10, 20-27.
- Pepe, M. S. (2003). *The statistical evaluation of medical tests for classification and prediction*. New York: Oxford.
- Rorden, C., & Karnath, H. O. (2004). Using human brain lesions to infer function: A relic from a past era in the fMRI age? *Nature Reviews Neuroscience*, 5, 813-819.
- Ryan, J. J., & Paulo, A. M. (2001). Exploratory factor analysis of the WAIS-III in a mixed patient sample. *Archives of Clinical Neuropsychology*, 16, 151-156.
- Saklofske, D. H., Hildebrand, D. K., & Gorsuch, R. L. (2000). Replication of the factor structure of the Wechsler Adult Intelligence Scale – third edition with a Canadian sample. *Psychological Assessment*, 12, 436-439.
- Swets, J. A. (1988). Measuring the Accuracy of Diagnostic Systems. *Science*, 240, 1285-1293.
- Wilde, N. J., Strauss, E., & Tulskey, D. S. (2004). Memory Span on the Wechsler Scales. *Journal of Clinical and Experimental Neuropsychology*, 26, 539-549.
- Ziemus, B., Baumann, O., Luerding, R., Schlosser, R., Schuierer, G., Bogdahn, U., et al. (2007). Impaired working-memory after cerebellar infarcts paralleled by changes in BOLD signal of a cortico-cerebellar circuit. *Neuropsychologia*, 45, 2016-2024.

## The Clinical Applicability of the Arithmetic, Digit Span Subtests of the Taiwan WAIS-III, and their Composite in Reflecting the Working Memory Index: A Retrospective Study

Wei-Han Wang<sup>1</sup>, Mau-Sun Hua<sup>1, 2</sup>, Chi-Cheng Yang<sup>1, 3</sup>, Yi-Chuan Chu<sup>4</sup>,  
Ting-Wen Cheng<sup>2</sup>, Ping-Keung Yip<sup>2</sup>, Ming-Chang Chiu<sup>1, 2</sup>, Ta-Fu Chen<sup>2</sup>,  
Sheng-Jean Huang<sup>3</sup>, Hsien-Tsung Chen<sup>4</sup>, and Wen-Chun Hsu<sup>4</sup>

<sup>1</sup> Department of Psychology, National Taiwan University

<sup>2</sup> Department of Neurology, National Taiwan University Hospital

<sup>3</sup> Division of Neurosurgery, Department of Surgery, National Taiwan University Hospital

<sup>4</sup> Department of Neurology, Chang-Gung Memorial Hospital

There are many issues in regards to whether or not the composite of Arithmetic and Digit Span subtests used to determine the working memory index of the Taiwan WAIS-III is compatible with that of the English WAIS-III, and whether the individual Arithmetic and Digit Span subtests, and the span scores of the Digit Span subtest are adequate for measuring working memory. This study was thus to make an attempt to examine these issues.

In this retrospective study, receiver operating characteristic (ROC) analyses were employed, using brain imaging data of 179 patients with various CNS diseases as the gold standard, to identify the working-memory-related scores in the Taiwan WAIS-III. The results, as estimated by a composite of the Arithmetic and the Digit Span subtests, revealed all the specificity, negative predictive value, and positive predictive value of the Taiwan WAIS-III working memory index were each at or above 94%, yet the sensitivity was only 63%. The sensitivities of the Arithmetic and the Digit Span subtests were even poorer, measuring only 50%. However, the positive predictive value of the Digit Span subtest was 86% while the Arithmetic subtest was 67%. In addition, the forward and backward span scores of the Digit Span subtest had good negative predictive value (all of them were above 92%), but poor sensitivity (all of them were below 58%).

Based on the present findings, it appears that all the Taiwan WAIS-III working memory index, the single Arithmetic and the Digit Span subtests, and the span scores of the Digit Span subtest could adequately discriminate individuals without working memory impairment, but had poor sensitivity on detecting the patients who evidenced working memory impairment. Accordingly, it is suggested that clinicians should gauge the working memory index by a composite of the Arithmetic, the Digit Span, and the Letter-Number Sequencing subtests if possible. Otherwise, a caveat should be taken with care to interpret the working memory index, as estimated by a composite of the Arithmetic and the Digit Span subtests, because the chance of the increasing false negative rate becomes ineludible. Respecting the limitations of retrospective study, however, it is necessary to make a prospective study to re-approach this issue.

**Keywords:** *WAIS-III, working memory, arithmetic, digit span*

