Volume 36(2), 2018
ISSN 1013-7661
DOI：10.6594/JTOTA

中華民國一〇七年 第三十六卷第二期
十二月出刊

職能治療學會雜誌

Journal of Taiwan Occupational Therapy Association

臺灣職能治療學會出版
Publlshed by Taiwan Occupational Therapy Association

職能治療學會雜誌
民國 107 年 36 卷第 2 期頁 133-151
DOI:10.6594/JTOTA.201812_36(2).0001

原著

學齡前孩童認知及情感心智理論與執行功能之關係

吳庭嘉 [1] 李雅珍 [2] 傅奕寧 [1,3] 尤姿婷 [1] 陳官琳 [1,4,*]

摘要

背景與目的：心智理論 (theory of mind) 與執行功能 (executive function) 為學齡前孩童二項重要發展能力，於孩童發展過程可能互相影響。近年神經影像研究發現心智理論可區分為認知及情感二層面。然而，至今仍無研究探討認知及情感心智理論個別與執行功能之關係。因此，本研究目的為探究學齡前孩童認知及情感心智理論與執行功能之關係。**研究方法：**以學齡前兒童心智理論評估工具–簡版、卡片向度改變分類作業–電腦版與魏氏幼兒智力測驗第四版中文版之語言理解指數評估 124 位 3 歲至 6 歲 5 個月之學齡前孩童之心智理論能力、執行功能能力及語言理解能力。**結果：**皮爾森相關係數 r (Pearson's correlation coefficient, r) 結果顯示學齡前孩童之認知及情感心智理論二分量表與執行功能測驗之單、雙向度正確率與反應時間之間，多個達顯著相關 ($|r|$=0.197~0.318, p<0.05)，且情感心智理論分量表與執行功能反應時間之相關性皆較認知心智理論分量表大。**討論與結論：**總體來說，本研究結果呈現認知及情感心智理論與執行功能間為顯著正向的關係，且情感心智理論與執行功能之相關性較認知心智理論大。本研究結果可增進臨床工作者對於心智理論與執行功能之兒童發展的知識基礎，亦有助於研究工作者進一步探究二者機制及介入方式。

關鍵字：學齡前兒童，認知心智理論，情感心智理論，執行功能

國立成功大學醫學院職能治療學系 [1]
亞洲大學職能治療學系 [2]
臺北市立聯合醫院兒童發展評估療育中心 [3]
國立成功大學醫學院附設醫院復健部 [4]

*通訊作者：陳官琳
701 台南市大學路一號
電話：886-6-2353535 分機 5906
電子信箱：klchen@mail.ncku.edu.tw

受文日期：民國 107 年 07 月 20 日
接受刊載：民國 107 年 10 月 30 日

前言

一、心智理論的定義及學齡前孩童心智理論發展的重要性

Premack 和 Woodruff 於 1978 年首次將「心智理論 (theory of mind) 」定義為理解自己及他人心理狀態的社會認知能力（例如情緒、觀念、感覺、慾望和意圖），並預測他人行為的推論系統。因此，具備「心智理論」能力可使個體站在他人的立場思考，理解他人想法 (Premack & Woodruff, 1978) ，以幫助個體有效地與他人社交互動。近年來，Shamay-Tsoory 等學者根據影像學的證據發表了一系列心智理論研究，將心智理論能力分為「認知心智理論 (cognitive theory of mind) 」和「情感心智理論 (affective theory of mind) 」 (Kalbe et al., 2010; Shamay-Tsoory & Aharon-Peretz, 2007; Shamay-Tsoory et al., 2007) 。「認知心智理論」關乎理解他人的認知狀態（例如：信念、想法、意圖與需求），並進而解釋與揣測他人行為反應 (Brothers & Ring, 1992; Coricelli, 2005; Duval et al., 2012; Kalbe et al., 2010) ；而「情感心智理論」則關乎理解他人的情感狀態（例如：情緒與感受），並進而辨識與推測他人情感狀況 (Brothers & Ring, 1992; Duval et al., 2012) 。由於情感心智理論可能牽涉乎理解他人的認知狀態，因此認知心智理論被認為是情感心智理論的基礎 (Dvash & Shamay-Tsoory, 2014) 。

心智理論的發展可視為孩童社會互動及建立社會關係的先備條件之一。在以群體生活為主的人類社會中，理解他人並建立有效的溝通蔚為重要，也因此心智理論之重要性不言而喻。正常發展孩童的心智理論發展有其順序與時間 (Callaghan et al., 2005) ，且學齡前是孩童心智理論發展的重要時期。正常發展孩童三歲開始可以分辨假裝與現實，已能根據他人的想法來預測對方的行動，並能依據他人的意圖是否得到滿足，來辨認他人的情緒起伏；四歲開始發展出初級錯誤信念能力 (first-order false belief) ，瞭解人們對同一件事物可能會有不同的看法；五歲開始瞭解兩人的觀點或信念可能是相異的；心智理論於六到七歲逐漸發展完備，具備次級錯誤信念 (second-order false belief) ，個體開始能考量他人如何看待其他人的想法 (Astington & Gopnik, 1991; Perner & Wimmer, 1985) 。根據以上發

展順序，過去研究者常以情緒辨識 (emotion distinction) 、個體有不同的慾望 (diverse desires) 、初級錯誤信念 (first-order false belief) 、次級錯誤信念 (second-order false belief) 等測驗作為瞭解學齡前孩童心智理論的指標。這些心智理論能力的發展，幫助孩童瞭解自身與他人之間的關係，並且察覺自己和他人可能擁有不一樣的信念，使孩童與他人的信念不同時，不會感到衝突。若缺乏心智理論能力，孩童將無法與他人有效地互動，進而影響交友能力與人際關係（Jenkins & Astington, 2000; 蔡淑玲，2001）。

二、執行功能的定義及學齡前孩童執行功能發展的重要性

「執行功能 (executive function) 」幫助個體執行有目的的行為，其不是單一的能力，它代表的是更高階一系列的能力，包含計畫 (planning) 、工作記憶 (working memory) 、認知彈性 (mental flexibility) 、反應抑制 (response inhibition) 、起始及監控行動 (initiation and monitoring of action) 等基本認知能力，這些能力的整合幫助個體應對需更高階處理程序之資訊、目標，以及適當的情緒反應 (Barkley & Lombroso, 2000; Pennington, 1997; Stuss, 1992) 。換句話說，執行功能是個體在面臨一個新的情況時，能夠迅速判斷周邊相關訊息，在腦中分析，並形成合理的各種選項，再進一步判斷最恰當的選擇，以最合適的方式進行，並利用外界的回饋來修正其反應的能力 (Lezak, Howieson, Loring, Hannay, & Fischer, 1995; Tranel, Anderson, & Benton, 1994) 。而其中認知轉換 (cognitive shifting) 被視為執行功能的核心成份，即根據不同的規則轉換分類方式，是兒童發展領域中最常被研究的認知彈性 (mental flexibility) 之一 (P. J. Anderson & Reidy, 2012; V. Anderson, Jacobs, & Anderson, 2010; Diamond, Barnett, Thomas, & Munro, 2007; Zelazo, 2006) 。認知轉換可幫助個體適應新情境的轉換，對習慣化的事物有改變與因應的彈性。

孩童於學齡前階段的發展即是為未來進入學校做預備，研究顯示執行功能之發展甚至比智力商數 (intelligence quotient, IQ) 或是閱讀、數學技巧更重要 (Blair & Razza, 2007; McClelland, Morrison, & Holmes, 2000) 。就認知轉換而言，4 歲前的幼兒傾向沿用舊的規則，無法彈性轉換規則 (Diamond, Carlson, & Beck, 2005;

Frye, Zelazo, & Palfai, 1995; Zelazo, Frye, & Rapus, 1996) 。這種重複出現不適用於當下的行為，稱作固著行為 (perseverative behavior) ，需約 4 至 5 歲才能比較彈性地轉換規則 (Zelazo et al., 1996; Zelazo et al., 2003) 。若孩童有執行功能障礙，則可能造成行為或概念轉換困難、抑制困難，以及無法辨識自己的行為錯誤等行為問題 (Lezak et al., 1995) ，造成孩童面對日常生活中的事件無法適當地反應 (Perner & Lang, 1999) 。生活中無時無刻都會接收到龐大的資訊，學齡前階段是學習如何適當反應、執行目標導向任務的重要時期 (Diamond, 2002; Zelazo, Frye, & Rapus, 1996) ，此階段執行功能的發展為未來之重要基礎，幫助孩童於日漸複雜的社會環境中發展出良好的應對策略 (Stuss, 1992; Travis, 1998) 。

三、心智理論及執行功能之關係

回顧相關文獻，可以發現心智理論及執行功能皆於學齡前階段快速成熟發展 (Diamond, 2002; Wimmer & Perner, 1983) ，且已有許多研究對學齡前孩童心智理論及執行功能之關係提出看法。Frye 等學者於 1995 年針對學齡前孩童認知心智理論能力與執行功能研究中發現，學齡前孩童需要依靠執行功能，根據不同情況進行推理及判斷，才能展現出心智理論能力。Hughes (1998) 的研究結果發現，學齡前孩童心智理論中說謊的能力和執行功能中抑制的能力之良好表現有極高的關係。Hughes 等學者 (1998) 的另一個學齡前孩童研究結果亦顯示，心智理論中的初級與次級錯誤信念是執行功能中制定策略、心理彈性的基礎 (Frye, Zelazo, & Palfai, 1995; Hughes, 1998a, 1998b) 。Sabbagh 等學者 (2006) 研究美國及中國之學齡前孩童，發現兩個文化之下的孩童其心智理論及執行功能之間皆有顯著之關係，研究結果顯示此相關性既強健且跨文化 (Sabbagh, Xu, Carlson, Moses, & Lee, 2006) 。此外，近年來有學者對學齡前兒童進行執行功能訓練，發現除了可提升執行功能表現，同時也提升了心智理論能力的表現 (Fisher & Happé, 2005; Kloo & Perner, 2003) 。綜合上述的研究結果，學齡前孩童的心智理論與執行功能有密切的關係。此外，根據上述研究主要可知認知心智理論應與執行功能相關，以及認知心智理

論被認為是情感心智理論的基礎 (Dvash & Shamay-Tsoory, 2014) ，因此認知心智理論與情感心智理論可能分別與執行功能相關。

然而，綜合上述學齡前孩童心智理論及執行功能的研究結果亦可發現，過去研究僅探討認知心智理論與執行功能之關係，未將心智理論分為認知及情感心智理論，亦無同時探討二種心智理論與執行功能之關係。未分開探討認知心智理論或情感心智理論與執行功能的相關性，混雜二種心智理論的情況下無法比較二種心智理論與執行功能之關係，造成臨床或研究人員及教育工作者無法全面掌握學齡前孩童心智理論及執行功能之確切關係，且缺乏心智理論與執行功能之兒童發展的知識基礎作為依據。

四、研究目的

為補足上述所提研究之不足，本研究目的為同時探討學齡前孩童認知及情感心智理論與執行功能關係。本研究假設為學齡前孩童的心智理論與執行功能具相關性，且認知心智理論以及情感心智理論亦分別與執行功能相關。本研究結果可增進臨床與教育工作者對於心智理論與執行功能之兒童發展的知識基礎，介入時亦可考量二者之關係。此外，本研究結果亦有助於研究工作者更進一步研究心智理論與執行功能相關性及介入方式之基礎。

方法

一、參與者

研究對象為小班至大班學齡前孩童，收案地點在臺南之幼兒園。孩童納入條件包含： (1) 生理年齡介於 3 歲至 6 歲 5 個月； (2) 無其他神經科、精神科或發展上的問題； (3) 未接受過任何早期療育或相關藥物治療； (4) 孩童家長同意，且孩童願意配合參與本研究。

二、研究工具

本研究使用三項工具，包含學齡前兒童心智理論評估工具-簡版、卡片向度改變分類作業-電腦版與魏氏幼兒智力測驗第四版中文版之語言理解指數，前二者為主要研究工具，後者之語言理解指數用以確認孩童之語言理解能力。

（一）學齡前兒童心智理論評估工具-簡版 (Brief Preschool Theory of Mind Assessment, BP-ToMA) ：

本研究使用 BP-ToMA 評估孩童的心智理論能力（傅奕寧，2017）。BP-ToMA 評估方式為電腦施測，題目涵蓋 3 至 6 歲 5 月之心智理論發展，共 8 題，並分為認知心智理論分測驗與情感心智理論分測驗。認知心智理論分測驗共 3 題，包含個體有不同的慾望測驗 (diverse desires) 、非預期的移位 (unexpected location)（初級錯誤信念）測驗、次級錯誤信念測驗 (second-order false belief) 各 1 題；情感心智理論分測驗共 5 題，包含 2 題情緒辨識 (emotional distinction) 測驗、1 題個體有不同慾望測驗、1 題非預期的移位（初級錯誤信念）測驗與 1 題次級錯誤信念測驗。題目以連環圖畫顯示，問題涵蓋 4 種題目類型： (1) 記憶題 (memory question) ：確認孩童記得題目的故事； (2) 前置題 (pre-question) ：確認孩童的喜好，以接續辨識題； (3) 辨識題 (test question) ：測驗孩童是否知道題目主角的心理狀態、行為、情緒； (4) 確認題 (justification question) ：確認孩童是知曉答案所以答對辨識題，而非猜對。情緒辨識測驗題包含了 1 題辨識題；個體有不同慾望測驗包含了記憶題、前置題、辨識題各 1 題；非預期的移位測驗題包含了記憶題、辨識題、確認題各 1 題；次級錯誤信念測驗題包含了，記憶題、辨識題、確認題各 1 題。表 1 為 BP-ToMA 題目結構的整理。整題測驗完整答對可得 1 分，答錯 1 小題(含)以上即為 0 分。BP-ToMA 之信效度皆已建立，兼具發展性及多元性架構，且適用於臺灣的學齡前孩童。除了經過專家檢驗之專家效度，建構效度之向度效度以羅序測量模式 (Rasch model) 檢驗，結果顯示 BP-ToMA 可分為二個向度 (p<.05) ，且皆顯示中度的內部一致性（個案分離係數[person separation reliability]=0.58 及 0.58）。題目難易度遵循發展順序，且符合認知心智理論為情緒心智理論之先備條

件的假說。題目在不同的性別上皆無顯著的差別試題功能 (p>.05) 。關於建構效度之收斂效度，魏氏幼兒智力量表第四版之語言理解分數 (Verbal Comprehension Index of the Wechsler Preschool and Primary Scale of Intelligence-IV) 、文蘭適應行為量表 (Vineland Adaptive Behavior Scale) 之溝通及社會化領域原始分數顯著 (p<.05) 預測 BP-ToMA 分數。此外，以全彩且富有互動性之連環漫畫呈現的 BP-ToMA 備受孩童喜愛，內容可被學齡前孩童容易理解，且作答時沒有障礙，有良好的適用性（傅奕寧, 2017）。

表 1
學齡前兒童心智理論評估工具-簡版 (Brief Preschool Theory of Mind Assessment, BP-ToMA) 題目結構

學齡前兒童心智理論評估工具-簡版	分測驗	
	情感心智理論	認知心智理論
情緒辨識	2	-
個體有不同慾望	1	1
非預期的移位	1	1
次級錯誤信念	1	1
總題數	5	3

註：數字代表題數

（二）卡片向度改變分類作業-電腦版 (Dimensional Card Change Sort task, DCCS) ：

本研究使用電腦版 DCCS 評估孩童的執行功能 (Zelazo, 2006) ，電腦版 DCCS 被視為威斯康辛測驗 (Wisconsin Card Sorting Testing) 的兒童版，以電腦施測，適用於 3 歲以上的孩童。電腦版 DCCS 測驗包含 2 張目標卡片和 14 張測驗卡片，分為形狀和顏色兩向度，請孩童依向度將測驗卡片與目標卡片配對，包含有回饋的練習題 15 題（正確與不正確）與 7 個段落的正式測驗題，共 80 題。正式測驗題除了第 4 段落 20 題，顏色、形狀向度混合出現，為雙向度的段落（包含二規則），其他段落皆為各 10 題，固定分顏色或形狀的單向度段落（僅單一規則）。其中，第 1、3 和 6 段落是以顏色分類，第 2、5 和 7 段落是以形狀分類。評估結果會包含正確率和反應時間，正確率是指孩童分類正確的比率，反應時間是指從刺激開

始到按鍵的時間，以毫秒為單位測量，但僅計算正確答案的反應時間，並且排除超過平均值 2.5 個標準差或是低於 200 毫秒的題目的反應時間。正確率愈高執行功能愈好，反應時間則愈短執行功能愈好 (Zelazo, 2006; Zelazo et al., 1996) 。本研究採取與 Diamond 和 Kirkham (2005) 的標準一致，反應時間分析中僅包括正確答案的試驗，亦排除反應時間超過平均值超過 2.5 標準偏差和反應時間低於 200 毫秒的試驗。此研究工具易於施測且廣泛使用於孩童，是探討孩童執行功能之認知轉換最常使用的工具之一 (Zelazo, 2006) 。

（三）魏氏幼兒智力測驗第四版中文版 (Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition, WPPSI-IV) 之語言理解指數 (Verbal Comprehension Index, VCI) ：

本研究使用 WPPSI-IV 中的 VCI 評估孩童的語文理解能力，以確認孩童有足夠的語文理解能力作答 BP-ToMA 與電腦版 DCCS，確保評估測驗結果正確，未受到語言理解能力影響而低估。本研究僅施測 WPPSI-IV 的 VCI 相關項目，2 歲 6 個月到 3 歲 11 個月之孩童施測聽詞指圖及常識二項目，4 歲到 7 歲 11 個月之孩童施測常識及類同二項目。WPPSI-IV 施測方式是請受試者以口頭回答或用手指指出答案來回答。評估 WPPSI-IV 的 VCI 項目約需時二十分鐘，依孩童實際作答狀況而不同。WPPSI-IV 已建立完整之信效度驗證，其再測信度為 0.72 至 0.89；內部一致性為 0.66；折半信度為 0.86 至 0.96；與魏氏孩童智力測驗第四版中文版之相關為 0.55 至 0.84（陳心怡、陳榮華，2013）。

三、收案程序

孩童將在安靜的環境下進行施測，施測程序如下：首先，研究者向孩童及家長說明本研究的目的與程序，在說明並取得同意後，請家長填寫同意書及基本資料表。接著，孩童將接受 BP-ToMA、WPPSI-IV、電腦版 DCCS，三項測驗的評估。此三項測驗以隨機順序施測於孩童。完成所有評估所需時間約 35 分鐘。本研究已送審教學醫院的人體試驗委員會，並於審查通過後執行計畫。

四、資料分析

先以描述性統計分析受試者之基本資料，並以皮爾森相關係數 *r* (Pearson's correlation coefficient, *r*) 檢驗認知心智理論、情感心智理論與執行功能彼此間的相關性強度。其中認知心智理論與情感心智理論是採用 BP-ToMA 的原始分數，執行功能是採用 DCCS 單向度與雙向度任務的正確率與反應時間。相關係數介於－1~＋1 之間，相關係數判斷標準為 0.70~0.99 為高度相關 (highly correlated) ，0.40~0.69 為中度相關 (moderately correlated) ，0.10~0.39 為低度相關 (mildly correlated) ，0.01~0.09 為接近無相關 (weakly correlated) ，0 為無相關 (Team, 2014) 。本研究使用 SPSS17 之軟體進行分析。

結果

本研究共自台南的五家幼兒園募集 128 位學齡前孩童，其中 4 位孩童因施測過程無法配合因而未能完成 DCCS 評估，已將其排除，其餘 124 位孩童（70 位男童與 54 位女童），平均年齡為 59.30 個月（標準差＝9.36），3 歲 0 月~3 歲 5 月之人數較少 (0.8%) ，孩童之人口學資料與各研究變項之描述型結果請見表 2。孩童之語言理解指數 (Verbal Comprehension Index, VCI) 項目百分等級平均為 73.1，且全數孩童之 VCI 整體評語皆為中下以上，並無能力邊緣或落後之孩童。進一步比較認知、情感心智理論發展階段之分布（圖 1），可發現孩童的認知心智理論發展階段分布較情感心智理論趨向進階，亦認知心智理論的發展較情感心智理論快。此外，DCCS 單向度之正確率平均值 (0.873) 高於雙向度正確率平均值 (0.802) ，DCCS 單向度反應時間平均值（1908.721 秒）較雙向度反應時間平均值（2341.156 秒）短。

心智理論與執行功能之相關性請見表 3，BP-ToMA 認知分量表與 DCCS 單向度正確率 (r=0.203, p=0.024) 、DCCS 單向度反應時間 (r=-0.268, p=0.003) 及 DCCS 雙向度反應時間 (r=-0.197, p=0.028) 皆達顯著相關 (p<0.05) ；BP-ToMA 情感分量表與 DCCS 單向度反應時間 (r=-0.318, p=0.000) 及 DCCS 雙向度反應時間 (r=-0.264, p=0.003) 皆達顯著相關 (p<0.01) ；BP-ToMA 認知分量表 (r=0.157,

p=0.083) 、情感分量表 (*r*=0.050, *p*=0.579) 與 DCCS 雙向度正確率之相關性未達顯著相關；比較數據可發現，BP-ToMA 認知、情感分量表與 DCCS 單向度（分別比較正確率、反應時間）的相關性皆大於雙向度，而 BP-ToMA 認知及情感分量表與 DCCS 單、雙向度反應時間的相關性皆大於 DCCS 單、雙向度正確率，且 BP-ToMA 情感分量表與 DCCS 單、雙反應時間之相關性皆大於認知分量表。

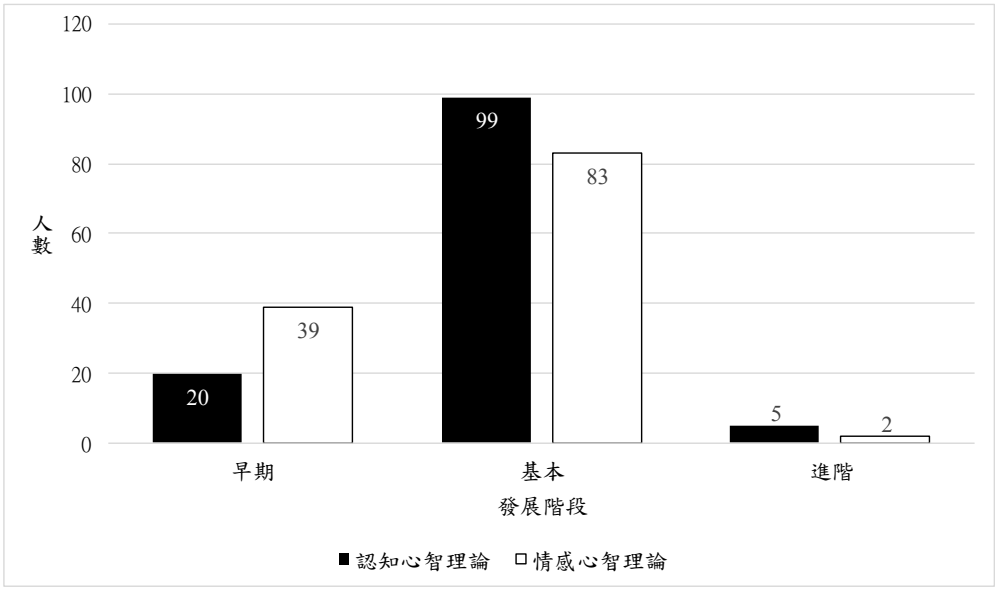

圖 1
認知、情感心智理論發展階段分布圖

討論

本研究旨在探討學齡前孩童認知及情感心智理論與執行功能之關係，以彌補過去研究僅探討認知心智理論與執行功能之關係，而未探討情感心智理論與執行功能之關係。根據本研究的結果，驗證了認知及情感心智理論與執行功能間為正相關，本研究不但再次證實認知心智理論與執行功能之間具相關性 (Frye et al., 1995; Hughes, 1998a, 1998b; Sabbagh et al., 2006) ，顯示孩童理解他人的認知狀態並進而解釋與揣測他人行為反應的能力，與根據不同的規則轉換分類方式之認知轉換相關。本研究結果亦提供了情感心智理論與執行功能的相關性，顯示孩童瞭

解他人情感狀態（例如：情緒與感受）並進而辨識與推測他人情感狀況的能力，亦與根據不同的規則轉換分類方式進行認知轉換相關。因此，本研究結果支持理解他人認知和情感狀態的心智理論與孩童順利執行認知轉換的執行功能之相關性。

表 2

孩童之人口學資料與各研究變項結果 (*N*=124)

項目	結果
孩童年齡（月）(平均值，標準差)	59.30/9.36
年齡族群（人數，%)	
3 歲 0 月~3 歲 5 月	1/0.8
3 歲 6 月~3 歲 11 月	14/11.3
4 歲 0 月~4 歲 5 月	20/16.1
4 歲 6 月~4 歲 11 月	31/25.0
5 歲 0 月~5 歲 5 月	25/20.2
5 歲 6 月~5 歲 11 月	20/16.1
6 歲 0 月~6 歲 5 月	13/10.5
孩童性別（男/女，%)	70/54 (56.5/43.5)
語文理解能力：WPPSI-IV	
VCI 組合分數（平均值，標準差）	114.24/16.376
VCI 百分等級（平均值，標準差）	73.1/26.9
VCI 整體評語(落後/邊緣/中下/中等普通/中上/優秀/非常優秀，%)	0/0/9/37/26/25/27 (0.0/0.0/7.3/29.8/21.0/20.2/21.8)
心智理論：BP-ToMA（平均值，標準差）	
認知分量表	1.15/0.722
情感分量表	2.85/0.722
發展階段	
認知心智理論（早期/基本/進階，%）	20/99/5 (16.1/79.8/4.0)
情感心智理論（早期/基本/進階，%）	39/83/2 (31.5/66.9/1.6)
執行功能：DCCS（平均值，標準差）	
單向度正確率	0.873/0.098
雙向度正確率	0.802/0.152
單向度反應時間（毫秒）	1908.721/1083.006
雙向度反應時間（毫秒）	2341.156/1286.492

註：WPPSI-IV=魏氏幼兒智力測驗第四版中文版 (Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition) ；VCI=語言理解指數 (Verbal Comprehension Index) ；BP-ToMA=學齡前兒童心智理論評估工具-簡版 (Brief Preschool Theory of Mind Assessment) ；DCCS=卡片向度改變分類作業-電腦版 (Dimensional Card Change Sort task)

表 3
心智理論與執行功能之關係 (*N*=124)

	BP-ToMA		DCCS			
	認知分量表	情感分量表	單向度正確率	雙向度正確率	單向度反應時間	雙向度反應時間
BP-ToMA 認知分量表	1					
BP-ToMA 情感分量表	0.390*** 0.000	1				
DCCS 單向度正確率	0.203* 0.024	0.132 0.144	1			
DCCS 雙向度正確率	0.157 0.083	0.050 0.579	0.728*** 0.000	1		
DCCS 單向度反應時間	-0.268** 0.003	-0.318*** 0.000	-0.448*** 0.000	-0.213* 0.018	1	
DCCS 雙向度反應時間	-0.197* 0.028	-0.264** 0.003	-0.307** 0.001	-0.134 0.136	0.885*** 0.000	1

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

有關認知心智理論與執行功能之關係，本研究結果呼應 Frye 等學者的研究，即正常發展學齡前孩童執行功能於卡片分類簡版 (Short-Form Card Sort) 的表現與認知心智理論的三個測驗表現，即錯誤信念 (false belief) 、外表/真實 (appearance-reality) 、表徵轉換 (representational change) 皆呈顯著正相關 (Frye et al., 1995) 。學齡前孩童需要依靠執行功能，根據各測驗中不同情況進行推理及判斷，以瞭解他人的心智狀態，並進一步預測他人的行為，以展現出相關測驗內容的心智理論能力；亦呼應 Hughes 的研究結果：認知心智理論的錯誤信念測驗與執行功能之工作記憶 (working memory) 、抑制控制 (inhibitory control) 測驗皆呈顯著正相關 (Hughes, 1998a) ，心智理論中的初級與次級錯誤信念即為瞭解他人所想的與實際狀況不同，過程中所需要的認知轉換、工作記憶與抑制控制能力都是執行功能的基礎。綜合以上可以發現，認知心智理論與執行功能之相關性皆呈顯著正相關。

比較達顯著相關的執行功能反應時間與認知、情感心智理論相關性，可發現情感心智理論與執行功能單、雙反應時間之相關性皆大於認知心智理論，意即情

感心智理論與執行功能的相關性大於認知心智理論。此結果可能的主要原因為認知心智理論被認為是情感心智理論的基礎 (Dvash & Shamay-Tsoory, 2014) ，而有二推測。其一，孩童認知心智理論表現較情感心智理論好，且孩童情感心智理論較認知心智理論晚發展，其分數分布與執行功能工具的分數分布較為相近，因此相較於認知心智理論可能有較佳的相關性。其二，考量認知心智理論被認為是情感心智理論的基礎來探討研究結果的行為意義可得知，相較於理解他人的認知狀態進而解釋與揣測他人行為反應，孩童可能更需要透過認知轉換來完成理解他人情感狀態（例如：情緒與感受）並進而辨識與推測他人情感狀況的過程 (Liebermann, Giesbrecht, & Müller, 2007) 。

比較心智理論與執行功能反應時間及正確率的相關性，可得知認知及情感心智理論與執行功能之反應時間的相關性皆大於正確率，意即理解自己及他人心理狀態並預測他人行為與訊息處理速度之相關性較大，與規則的順利使用及規則轉換之相關性較小。這也可能是因為反應時間較正確率能敏感地反映出孩童執行功能表現，反應時間呈現連續性的數據，而正確率只呈現答對或答錯，因此敏感性較好的反應時間能與心智理論有較高的相關性。

認知及情感心智理論與執行功能之單向度（包含正確率、反應時間）的相關性皆大於雙向度。由此可推測，在理解並推測他人認知及情感狀態時，相較於在兩個規則之間轉換，孩童需要較專注於使用同一規則。認知心智理論與執行功能單、雙向度正確率的相關性這部分的結果與 Frye 等學者之研究結果相同 (Frye et al., 1995) 。綜合以上，相較於順利於兩個規則間轉換以及於兩個規則間轉換時的心理歷程、訊息處理速度，認知及情感心智理論與順利使用單一規則以及專注於單一規則上不受干擾所需的心理歷程和訊息處理速度兩者較為相關。

單獨比較認知及情感心智理論，發現學齡前正常發展孩童認知心智理論表現比情感心智理論好，因此推測情感心智理論難度較難，且情感心智理論較認知心智理論晚發展成熟。執行功能單向度表現比雙向度好，則推測是由於雙向度作業需要在兩個向度之間轉換，需要有抑制已習慣的分類向度之能力，相較於單向度作業需要較多的心智歷程 (Diamond & Gilbert, 1989; Mayr & Keele, 2000) 。

本研究招募之學齡前孩童語言理解能力表現大多落在正常範圍，確保孩童有能力進行心智理論及執行功能之評估。此外，分析孩童人口學資料之年齡族群，發現 3 歲 0 月~3 歲 5 月孩童人數較少。推測這是因為此年齡區間孩童本就佔幼兒園的少數，且一部分（4 位）又因無法完成 DCCS 而排除。此情形亦符合過去研究發現三到四歲的學齡前孩童，在規則改變後仍會繼續使用上一個分類方法，出現明顯的固著現象(Zelazo, 2006; Zelazo et al., 1996; Zelazo et al., 2003)，且於實際施測時亦有發現此情況。回顧施測過程發現，雖施測時間略長（大約 35 分鐘）但大部分孩童皆能保持專注。推測這是由於本研究使用之評估工具能引起孩童的興趣，如 BP-ToMA 是以說故事、問問題的形式呈現，與平時幼兒園老師說故事的方式相似，對孩童來說較熟悉，也很有動機想要回答問題；此外，DCCS 使用電腦施測，實際施測時觀察孩童反應，發現此形式對他們來說就像玩遊戲一樣。總體來說本研究使用的心智理論與執行功能評估工具能引起孩童興趣，且十分適用於學齡前孩童。

回顧施測過程發現，雖施測時間略長（大約 35 分鐘），但大部分孩童皆能保持專注。施測過程中若孩童要求停下，中途會休息。雖然使用之評估工具皆適用於學齡前孩童，且可引起興趣，但仍可觀察到孩童之注意力隨著時間下降，較後順序施測之評估項目可能受到注意力下降之影響造成表現較差，即使本研究三項測驗以隨機順序施測於孩童，仍不免有些影響。建議未來在施測時可增加休息時間或是分為二次施測，使孩童能以較一致的狀態進行測驗。此外，雖然本研究結果可提供臨床人員關於學齡前孩童認知及情感心智理論與執行功能之相關性更充足的理論基礎和架構，卻未能直接確認二者因果關係，建議未來研究者可進一步以縱貫性研究與介入研究探究二者之因果關係。

總體來說，本研究結果呈現認知及情感心智理論與執行功能為顯著正向的關係，且情感心智理論與執行功能之相關性較認知心智理論大。因此推論在學齡前階段，情感心智理論與執行功能發展的交互關係較密切。此資訊有助於臨床人員與教育工作者更清楚掌握學齡前兒童心智理論與執行功能發展的實證基礎，在分析兒童相關之發展能力能有所依據以提供個案適切的介入。此外，在心智理論分為認知心智理論及情感心智理論之初期，仍有許多理論基礎還未建構，本研究結果有助於研究工作者更進一步研究心智理論與執行功能相關性及介入方式之基礎。

致謝

感謝所有參與本研究的孩童、家長與幼兒園，以及國立成功大學醫學院暑期研究獎助計畫 (NCKUMCS2016059) 與科技部大專研究計畫 (106-2813-C-006-078-H) 提供經費，使研究得以完成。

參考文獻

陳心怡、陳榮華（2013）。**魏氏幼兒智力量表第四版(中文版)技術與解釋手冊**。台北市：中國行為科學社股份有限公司。

傅奕寧（2017）。**新發展之「學齡前兒童心智理論評估工具-簡版」：發展、信效度及適用性驗證**。國立成功大學職能治療學系碩士論文，未出版，台南市。

蔡淑玲 （2001）。 高功能自閉症兒童之心智理論教學。**國小特殊教育**。

Anderson, P. J, & Reidy, N. (2012). Assessing executive function in preschoolers. *Neuropsychology Review, 22*, 345-360.

Anderson, V., Jacobs, R., & Anderson, P. J. (2010). *Executive functions and the frontal lobes: A lifespan perspective*: Psychology Press.

Astington, J. W., & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology, 9*, 7-31.

Barkley, R. A, & Lombroso, P. J. (2000). Genetics of childhood disorders: XVII. ADHD, Part 1: The executive functions and ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*, 1064-1068.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*, 647-663.

Brothers, L., & Ring, B. (1992). A neuroethological framework for the representation of minds. *Journal of Cognitive Neuroscience, 4*, 107-118.

Callaghan, T., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., Tapanya, S., & Singh, S. (2005). Synchrony in the onset of mental-state reasoning evidence from five cultures. *Psychological Science, 16*, 378-384.

Coricelli, G. (2005). Two-levels of mental states attribution: from automaticity to voluntariness. *Neuropsychologia, 43*, 294-300.

Dvash, J., & Shamay-Tsoory, S. G. (2014). Theory of mind and empathy as multidimensional constructs: Neurological foundations. Topics in Language Disorders, 34, 282-295.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. *Principles of frontal lobe function*, 466-503.

Diamond, A., Carlson, S. M., & Beck, D. M. (2005). Preschool children's performance in task switching on the dimensional change card sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology*, 28, 689-729.

Diamond, A., Barnett, S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science (New York, NY), 318*, 1387.

Diamond, A., & Gilbert, J. (1989). Development as progressive inhibitory control of action: Retrieval of a contiguous object. *Cognitive Development, 4*, 223-249.

Duval, C., Bejanin, A., Piolino, P., Laisney, M., de La Sayette, V., Belliard, S., Eustache, F. & Desgranges, B. (2012). Theory of mind impairments in patients with semantic dementia. *Brain, 135*, 228-241.

Fisher, N., & Happé, F. (2005). A training study of theory of mind and executive function in children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 35*, 757.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*, 483-527.

Hughes, C. (1998a). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*, 233-253.

Hughes, C. (1998b). Finding your marbles: Does preschoolers' strategic behavior predict later understanding of mind? *Developmental Psychology, 34*, 1326.

Jenkins, J. M, & Astington, J. W. (2000). Theory of mind and social behavior: Causal models tested in a longitudinal study. *Merrill-Palmer Quarterly (1982-)*, 203-220.

Kalbe, E., Schlegel, M., Sack, A. T, Nowak, D. A, Dafotakis, M., Bangard, C., . . .

Kessler, J. (2010). Dissociating cognitive from affective theory of mind: a TMS study. *Cortex, 46*, 769-780.

Kloo, D., & Perner, J. (2003). Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. *Child Development, 74*, 1823-1839.

Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, J. H., & Fischer, J. S. (1995). Neuropsychological assessment. 3 Oxford University Press. *New York*.

Liebermann, D., Giesbrecht, G. F, & Müller, U. (2007). Cognitive and emotional aspects of self-regulation in preschoolers. *Cognitive Development, 22*, 511-529.

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: the role of backward inhibition. *Journal of Experimental Psychology: General, 129*, 4.

McClelland, M. M, Morrison, F. J, & Holmes, D. L. (2000). Children at risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly, 15*, 307-329.

Pennington, B. F. (1997). Dimensions of executive functions in normal and abnormal development. *Development of the Prefrontal Cortex: Evolution, Neurobiology, and Behavior, 265-281*.

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3*, 337-344.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of second-order beliefs by 5-to 10-year-old children. *Journal of Experimental Child Psychology, 39*, 437-471.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 515-526.

Sabbagh, M. A, Xu, F., Carlson, S. M, Moses, L. J, & Lee, K. (2006). The development of executive functioning and theory of mind a comparison of Chinese and US preschoolers. *Psychological Science, 17*, 74-81.

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia, 45*, 3054-3067.

Shamay-Tsoory, S. G, Shur, S., Barcai-Goodman, L., Medlovich, S, Harari, H., &

Levkovitz, Y. (2007). Dissociation of cognitive from affective components of theory of mind in schizophrenia. *Psychiatry Research, 149*, 11-23.

Stuss, D. T. (1992). Biological and psychological development of executive functions. *Brain and Cognition, 20*, 8-23.

Team, R. C.. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013: ISBN 3-900051-07-0.

Tranel, D., Anderson, S. W., & Benton, A. (1994). Development of the concept of "executive function" and its relationship to the frontal lobes. *Handbook of Neuropsychology, 9*, 125-148.

Travis, F. (1998). Cortical and cognitive development in 4th, 8th and 12th grade students: The contribution of speed of processing and executive functioning to cognitive development. *Biological Psychology, 48*, 37-56.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103-128.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*, 297.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37-63.

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., . . . Sutherland, A. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, i-151.

Journal of Taiwan Occupational Therapy Association
2018, 36(2),133-151
DOI:10.6594/JTOTA.201812_36(2).0001

Original article

The Relationships Between Both Cognitive and Affective Theory of Mind and Executive Function in Preschool Children

Ting-Jia Wu[a], Ya-Chen Lee[b], I-Ning Fu[a,c],
Tzu-Ting Yu[a], Kuan-Lin Chen[a,d,*]

Abstract

Background: Theory of mind (ToM) and executive function (EF) are important to preschool children and may interact throughout preschool development. Recent neuroimaging research has shown ToM to have cognitive and affective aspects. However, so far, no researchers have separately investigated the relationships between these two aspects of ToM and EF. Hence, the purpose of this research was to explore the relationships between both cognitive and affective ToM and EF in preschool children. **Methods:** A total of 124 children aged between 3 years old and 6 years 5 months were assessed with the Brief Preschool Theory of Mind Assessment (BP-ToMA) , Dimensional Change Card Sort task (DCCS) , and Verbal Comprehension Index of the Wechsler Preschool and Primary Scale of Intelligence–Fourth Edition respectively for their ToM, EF, and verbal comprehension abilities. **Results:** Most of the correlations between the cognitive and affective ToM subscale scores of the BP-ToMA and the reaction times of the one- and two-dimension tests of the DCCS were significant ($|r|$ = 0.197–0.318, $p < 0.05$) . The correlations between affective ToM scores of the BP-ToMA and reaction times of the DCCS were stronger than those between the cognitive ToM scores of the BP-ToMA and the reaction times of the DCCS. **Conclusion:** Both affective and cognitive ToM are significantly correlated with EF, and the relationship between affective ToM and EF is the stronger of the two. These findings clarify the relationships between ToM and EF for better practice and as a basis for further exploration of their mechanisms.

Keywords: *Affective Theory of Mind, Cognitive Theory of Mind, Executive Function Preschool Children*

[a] Department of Occupational Therapy, College of Medicine, National Cheng Kung University
[b] Department of Occupational Therapy, College of Medical and Health Science, Asia University
[c] Child Developmental Assessment & Intervention Center, Taipei City Hospital
[d] Department of Physical Medicine and Rehabilitation, National Cheng Kung University Hospital

*Correspondence: Kuan-Lin Chen
1 University Rd., Tainan 701, Taiwan
TEL: 886-6-2353535 ext 5906
E-mail: klchen@mail.ncku.edu.tw

Received: 20 July 2018
Accepted: 30 October 2018

Journal of Taiwan Occupational Therapy Association
2018, 36(2),152-176
DOI:10.6594/JTOTA.201812_36(2).0002

Original article

The Role of Occupational Therapists in Special Educational Professional Teams for Improving the Adaptive Functioning of Elementary School Children With Autism Spectrum Disorder in Tainan City

Wei-Dai Sun[a], Ling-Yi Lin[a,*]

Abstract

Special educational professional teams have developed for more than twenty years in Taiwan, and increasingly more related research outcomes have been focused on the satisfaction and interactions between members. However, few studies have investigated the effects of occupational therapists (OTs) in professional teams. Therefore, this study investigated whether the adaptive functioning of elementary school students with autism spectrum disorder (ASD) improved after receiving services from OTs in professional teams. The study recruited 19 elementary school students with ASD (16 boys, 3 girls; mean age: 8.79 years) to receive services by OTs in professional teams over two semesters during one academic year. The Vineland Adaptive Behavior Scale-Chinese (VABS-C) and the Goal Attainment Scale (GAS) were used. The results of the VABS-C indicated significant improvements in the students' composite scores (Z = -2.069, p = .039) and daily living skills (Z = -2.536, p = .011). The GAS results showed significant improvements from 37.60 (SD = 1.47) at baseline to 52.55 (SD = 6.98; Z = -3.417, p = .001) after receiving services from the OTs. The difference in the socialization score significantly correlated with the amount of time that the OTs spent in school (r = .489, p = .046). The differences in the raw scores of daily living skills, socialization, and motor correlated significantly with the amount of time that the OT spent at the school. The present study provides evidence to support the current policy of the special educational professional team and highlights the importance of school-based OTs.

Keywords: *Adaptive Functioning, Autism, Occupational Therapist, Special Educational Professional Team*

[a] Department of Occupational Therapy, College of Medicine, National Cheng Kung University

*Correspondence: Ling-Yi Lin
1 University Road, Tainan 701, Taiwan
TEL: 06-235-3535 ext. 5917
E-mail: lingyi@mail.ncku.edu.tw

Received: 31 August 2018
Accepted: 07 November 2018

1. Introduction

Autism spectrum disorder (ASD) is a permanent impairment that can affect a person's independence throughout their entire life. An ASD diagnosis is determined according to a person's social communication dysfunction and repetitive or restrained behaviors and interests (American Psychiatric Association, 2013). When children attend elementary school, they are expected to possess a certain level of independence and control over their lives. For example, children can participate in the academic, social, extracurricular, and independent living activities needed for student success in school. However, elementary school children with ASD typically have underdeveloped or inappropriate adaptive functions, affecting their ability to perform activities of daily living independently even if they are high-functioning ASD. The adaptive functions (communication, social interaction, daily living, and motor skills) of people with ASD are considerably impaired regardless of the symptom severity or intelligence level (Kanne et al., 2011; Klin et al., 2007; Perry, Flanagan, Geier, & Freeman, 2009).

Adaptive functioning, which is crucial to independent functioning, is divided into four general abilities: communication, socialization, daily living, and motor skills (Sparrow, Balla, & Cichetti, 1985; Sparrow, Cicchetti, & Balla, 2005). In the communication and socialization domains, children with ASD have major deficits (American Psychiatric Association, 2013). Difficulties with comprehension and verbal expression affect not only children with low-functioning ASD, but also children with high-functioning ASD, particularly

those experiencing problems in developing adequate social communication ability (Kanne et al., 2011). In the daily living domain, low-functioning children with ASD cannot execute age-appropriate behaviors because of their obvious sensory processing problems and self-stimulus behaviors (Patel, Preedy, & Martin, 2014). Children with high-functioning ASD do not perform as well as children of average intelligence, and their Vineland Adaptive Behavior Scales-II (VABS-II) scores are significantly lower than those of typically developing children (Farley et al., 2009; Hume, Loftin, & Lantz, 2009). In the motor skill domain, the motor coordination abilities of children with ASD, such as balance and ball skills, are poorer than those of typically developing children (Green et al., 2009); moreover, incoordination in handwriting is common in children with ASD (Ashburner, Ziviani, & Rodger, 2008; Cartmill, Rodger, & Ziviani, 2009). As the impairments in adaptive functioning in children with ASD affect their level of independence, the appropriate intervention by professionals aims to improve these children's adaptive functioning performance.

Special educational professional teams (hereinafter referred to as professional teams), as a type of school-based intervention, are organized to provide a holistic perspective for improving the situation of students with special needs. When children with ASD reach school age, the school environment has a critical role in their lives because they spend considerable time accessing compulsory education. The goal of school-based occupational therapy services in professional teams is enabling participation in school-related occupations (Bonnard & Anaby, 2016). Furthermore, according to

Kasari and Smith (2013), school-based interventions are more effective than clinic-based interventions. Children with ASD can easily learn the skills required for general school contexts through a school-based intervention. In Taiwan, the concept of professional teams originated from the Education for All Handicapped Children Act (EAHCA, 1975) and the Individual with Disabilities Education Act (IDEA, 1982) in the United States, which were merged into the 1997 Special Education Act in Taiwan (Special Education Act, 1997). Professional teams were introduced by local educational agencies and provided team-based collaborative services. In 1998, Occupational Therapy was identified as a team-based collaborative service based on the Special Education Act and regulations (Enforcement Rules of the Special Education Act, 1998). Professional teams have five responsibilities: (a) assess a student's present status regarding abilities, family functioning, and needs; (b) provide related services and support strategies for students with special education needs; (c) determine school-year and per-semester educational goals, assessment means, dates, and educational objective criteria each semester; (d) provide functional behavioral intervention and administrative support for students with emotional and behavioral problems; and (e) provide transitional counseling and services for students with special needs (Special Education Act, 2014; Enforcement Rules of the Special Education Act, 2013). While the professional teams typically comprise multiple professionals (physical therapists, occupational therapists [OTs], speech therapists, psychological therapists, teachers, special education teachers, and parents), their composition differs to suit students' individual needs. OTs focus on children's

occupational performance such as education, the activities of daily living, social participation, play, leisure, and vocational activities.

The concept of professional teams has been developed for more than 20 years in Taiwan, and increasingly more related research outcomes have focused on the internal effects, which include how members interact with each other (Lai & Yen, 2006; Wuang & Wang, 2002). However, only a few studies have researched the relevant external effects, such as the quality of teamwork (Wuang & Wang, 2002; Yang & Cherng, 2013). Some researchers have investigated the subjective satisfaction of parents and teachers for evaluating the quality of team services (Lai & Yen, 2006; Yang & Cherng, 2013; Yen, 2005); however, clinicians and parents lack objective terms for understanding the extent to which the teams improve the adaptive performance of children with ASD.

This study investigates whether the adaptive functioning of elementary school children with ASD improves after receiving services by OTs in professional teams. The results could clarify the role of occupational therapy services in professional teams. Thus, this study addressed the following questions: (1) Does a significant difference exist between the pre-test and post-test adaptive functioning of children with ASD? (2) Was the amount of time that the OTs spent in school over two semesters related to the change in students' scores for adaptive functioning? Subsequently, the research hypotheses of this study were (1) the adaptive functioning of elementary students with ASD was significantly different between the pre-test and the

post-test, and (2) the amount of time that the OTs spent in school over two semesters correlated with improvements in students' adaptive functioning.

2. Methods

2.1 Participants

This study used a repeated measure design and purposive sampling. Nineteen participants (16 boys, 3 girls) were recruited from regular elementary schools in Tainan City. The criteria for inclusion were (a) possessing formal disability cards for ASD, (b) studying in regular elementary schools in Tainan, and (c) receiving occupational therapy services in professional teams. The mean age of the participants was 8.79 years old, ranging from 6.46 to 10.87 years old. The school days for two semesters were 179 days. Students received an average of 3.75 hours of occupational therapy services over two semesters. The students were pulled out of the class to receive services without the presence of teachers or other professional members. According to the participants' disability certifications, the severities were divided into three levels: mild (8 students), moderate (10 students), and severe (1 student). The students attended three different types of classes: regular class (1 student), inclusive class (8 students are extracted for individual lessons), and special education class (10 students). Seven participants had no extracurricular treatment. Twelve participants also received extracurricular treatments such as occupational therapy (10

students), physical therapy (5 students), speech therapy (8 students), and other treatments (1 student received psychological treatment and 1 received acupuncture) (Table 1).

To understand the students with ASD comprehensively, we recruited their OTs. The inclusion criteria for the OTs were (a) members of professional teams, (b) certificated by the Tainan Bureau of Education, and (c) with at least two years of related pediatric experience. The present study recruited 10 part-time OTs, with a mean pediatric experience of 9.6 years.

2.2 Measurements

2.2.1 Basic Information Questionnaire

This questionnaire was designed to investigate the demographic data, such as the birth date, gender, autism severity, types of classes, and socioeconomic status of the family, and the ongoing treatments in clinics. The main classroom teacher was asked to provide information about the related services in the professional teams.

2.2.2 Vineland Adaptive Behavior Scales

The Vineland Adaptive Behavior Scales (VABS; Sparrow, Cicchetti, & Balla, 1984) is commonly used to evaluate the adaptive functioning of children with ASD and is suitable for use with children aged 3 to 12 years. There are two different editions for teachers and parents to assess people with ASD. This study used the Classroom Edition-Chinese version (Wu, Chang, Lu & Chiu, 2004) for the main classroom teachers who were familiar with the

child with ASD to assess their performance at school. The following four domains are key parts of the VABS: communication (receptive, expressive, and written), activity of daily living (personal, domestic, and community), socialization (interpersonal relationships, play and leisure, and coping skills), and motor skills (gross and fine motor). The scores according to the teachers' observation are divided into three levels (i.e., 0 = never/ seldom, 1 = sometimes, and 2 = very often). The total raw scores can be analyzed by conversion to standard scores, percentile rank, and developmental age. The results showed the Chinese version of VABS (VABS-C) having a good split-half reliability (0.91–0.99), good test-retest reliability (0.62–0.95), good interrater reliability (0.74–0.89), and good construct validity (Wu, Chang, Lu & Chiu, 2004).

2.2.3 Goal Attainment Scale

The Goal Attainment Scale (GAS) was originally developed as a measurement tool for program evaluation purposes in 1968 (Kiresuk & Sherman, 1968). Researchers used the GAS to capture the extent to which the individual goals for treatment were achieved. The GAS is rated on five outcome score levels (ranging from -2 to +2) with the degree of attainment captured for each goal area (i.e., 0 indicates the achievement of the expected level, ±1 indicates the achievement of somewhat less than/more than the expected level, ±2 indicates the achievement of much less/more than the expected level). In the practical guide for the GAS, raters and clients could rate the varied importance and difficulty for each goal to weigh the scores.

However, as the OTs in this study were new GAS users, they were asked to rate the importance and difficulty as one and put no weight on the goals. The baseline score of each goal was -1, and the OTs rated the students' changes as outcome scores. The following is the formula for the GAS:

$$\text{Overall Goal Attainment Score} = 50 + \frac{10\sum(W_i X_i)}{\sqrt{(0.7\sum W_i^2 + 0.3(\sum W_i)^2)}}$$

where W_i is the score weighting, which is equal to 1 in this study; and X_i is the score from the OTs, ranging from -2 to +2. Researchers have claimed that, as the reliability of the GAS varies depending on the raters, it should be assessed on a case-by-case basis because the goals are individual and the reliability could change (Ruble, McGrew, & Toland, 2012; Schlosser, 2004).

2.3 Experimental Procedure

After receiving approval from the Institutional Review Board (IRB) in National Cheng Kung University Hospital, the elementary schools in Tainan City were contacted to recruit the students with ASD who received services from OTs in professional teams. The investigator distributed flyers to the parents of the students receiving services. Following agreement by parents and students to participate in this study, the OT in the professional team was invited to participate, and the teachers of the students with ASD were asked to complete the VABS-C. In the beginning of the semester, the OTs in the professional teams set appropriate goals for the students with ASD as the baseline of the GAS. Over two semesters, the OT offered these students suitable interventions or adjustments and provided their teachers and

parents with supportive suggestions to improve these students' adaptive functioning. After the students received services from the OTs, their scores of the VABS-C and GAS were collected again to determine any changes following the intervention. All students were assessed in September for the baseline and in June for the end of services.

2.4 Data analysis

The collected data were analyzed in SPSS 17.0 for the descriptive analysis and hypothesis testing. The demographic data were presented using descriptive analysis in means, percentages, and standard deviations. The pre-test scores of the VABS-Chinese version and GAS data were compared with the post-test scores using the Wilcoxon signed rank test analysis. The significant level was set at $p < .05$. In order to determine whether a confounding factor (i.e., the extracurricular treatments) might affect the students' improvements, the Mann-Whitney U analysis was conducted to see whether any significant differences exist between two groups—with versus without clinical treatments. The Pearson and Spearman's correlation analyses were computed among the data from the demographic data and the scores of VABS. The cut-off points for Pearson correlation coefficients were 0.25, 0.5, and 0.75 to present as mild, moderate, and excellent levels of correlation, respectively (Portney & Watkins, 2007)

Table 1

Overview of Demographic Data

	Mean	SD	Range
Chronological age (years)	8.79	1.23	6-11
The amount of time spent in school for two semesters (hours)			
OTs	3.75	2.73	1-10
PTs	3.33	3.01	0-10
STs	2.23	2.93	0-10
	n	**%**	
Boys	16	84.2	
Girls	3	15.8	
Level of disability			
Mild	8	42	
Moderate	10	53	
Severe	1	5	
Types of class			
Regular class	1	5	
Inclusive class	8	42	
Special education class	10	53	
Professional teams			
OT	5	26	
OT and PT	4	21	
OT and ST	3	16	
OT, PT, and ST	7	37	
Extracurricular treatments			
OT	10	53	
PT	5	26	
ST	8	42	
Others			
Psychological therapy	1	5	
Acupuncture	1	5	

NOTE. OTs = Occupational therapists; PTs = Physical therapists; STs = Speech therapists; OT = Occupational therapy; PT = Physical therapy; ST = Speech therapy.

3. Results

Table 2 shows the pre-test results of the descriptive analysis of VABS-C. The students showed relative weakness among all domains of the VABS-C. In the domain of communication, 72.2% of the students were at borderline (-1 SD~-2 SD) or disability (-2SD below). In the domains of daily living skills and socialization, only 5.6% of the students had age-appropriate skills at school. The domain of motor skills showed similar results as communication: 77.8% of the students were at borderline or disability performance. The students' mean developmental age of adaptive functions was 4.32 years old (ranging from 0.5 to 8.5 years old), which was lower than their chronological age.

After the students received services from the OTs in professional teams, they showed some subtle improvements during their school days. A comparison of the raw pre-test and post-test scores in the VABS-C showed that the scores of daily living skills ($Z = -2.536$, $p = .011$) and the composite scores ($Z = -2.069$, $p = .039$) reached a significant level. Overall, the mean age equivalent increased significantly from 4.25 ($SD = 2.1$) to 4.60 years old ($SD = 2.28$; $Z = -1.970$, $p = .049$). However, no significant differences were found in the standard scores of the VABS-C at the pre-test and post-test.

The students' difficulties of adaptive functioning existed in their daily lives at school, and their OTs set individual goals for them through the GAS. According to the OTs' documents, most of the students' goals were related with adaptive functioning. Depending on the students' individual needs, their goals could be set by the OTs as one to five for each student. Socialization

was considered a major problem because it formed 50% of the goals listed. The other problems belonging to adaptive functioning were motor (33.33%), daily living skills (12.49%), and communication (4.17%). In the GAS, the OTs rated the students' changes from -2 to +2. The baseline and outcome scores for each student were adjusted according to the formula for the GAS. A comparison of the differences before and after receiving the OTs' services using the Wilcoxon rank-signed test showed a significant increase in the post-test scores from the pre-test ($Z = -3.417$, $p = .001$).

Because outcomes from the extracurricular treatments might have been confounding factors that affected the students' improvements, we separated the data into two groups—with versus without clinical treatments—and compared their standard scores on the VABS-C at pre-test and post-test. The results of the Mann-Whitney U analysis showed no significant differences in the demographic data between the two groups. Furthermore, a comparison of the differences between the standard scores of the two groups showed a significant difference in the communication domain ($Z = -2.175$, $p = .030$). While the communication domain might be affected by extracurricular treatments, other domains of improvement might not.

To investigate the relationship between other factors and the change in scores for adaptive functioning, we included factors, such as the amount of time that the OTs spent at the school over two semesters and the receiving of extracurricular treatments, in the correlation analysis. A difference in the socialization standard score correlated significantly with the amount of time that the OTs spent at the school over two semesters ($r = .489$, $p = .046$), and

the results showed a moderate positive correlation. Additionally, the differences in the raw scores of daily living skills (r = .426, p = .049), socialization (r = .504, p = .028), and motor (r = .525, p = .021) correlated significantly with the amount of time that the OT spent at the school over two semesters. A difference in the communication standard score correlated significantly with receiving extracurricular treatments, with a moderate positive correlation. No correlations were found between the remaining variables and differences in the VABS-C scores (Table 3).

Table 2

Comparing Pre-test and Post-test of VABS-C and GAS by Wilcoxon Signed-rank Test

Subtests	Pre-test	Post-test	Z value	*p* value
VABS-C				
Communication	60.67 ± 34.34	59.94 ± 35.03	-.129	.897
Activity of Daily Living	86.83 ± 36.88	95.29 ± 44.98	-2.536	.011*
Socialization	34.56 ± 20.47	36.53 ± 22.42	-1.336	.182
Motor	40.06 ± 12.89	41.59 ± 11.19	-1.114	.265
Composited	222.11 ± 98.89	233.35 ± 107.17	-2.069	.039*
Age equivalent (years)	4.32 ± 2.12	4.60 ± 2.35	-1.970	.049*
GAS	37.60 ± 1.47	52.55 ± 6.98	-3.417	.001*

NOTE. VABS-C = Vineland Adaptive Behavior Scales-Chinese Version; GAS = Goal attainment scaling

*$p < .05$.

Table 3
Correlations of the Pre-test and Post-test Change Scores of VABS-C, the Amount of Time that OTs, PTs, and STs Spent in School, and Having Extracurricular Treatments

Change Scores	The amount of time that the OTs spent in school for two semesters[a]	The amount of time that the PTs spent in school for two semesters[a]	The amount of time that the STs spent in school for two semesters[a]	Having extracurricular treatments[b]
VABS-C Standard Scores				
Communication	0.167	0.469	0.557*	0.544*
Activity of Daily Living	0.427	0.097	0.171	0.238
Socialization	0.489*	0.206	0.481	0.331
Motor	0.422	0.402	0.374	0.252
Composited	0.349	0.487	0.436	0.465
VABS-C Raw Scores				
Communication	0.033	0.325	0.501*	0.318
Activity of Daily Living	0.482*	0.151	0.139	0.012
Socialization	0.464*	0.212	0.467*	0.275
Motor	0.459*	0.078	0.263	0.142
Composited	0.412	0.246	0.335	0.263

NOTE. VABS-C = Vineland Adaptive Behavior Scales-Chinese Version; OTs = Occupational therapists; PTs = Physical therapists; STs = Speech therapists.

[a]Pearson correlation coefficients. [b]Spearman's correlation coefficients. *$p < .05$.

4. Discussion

The present study examined the differences between students' pre-test and post-test adaptive functioning scores after receiving services from OTs in professional teams and investigated the relationships among the changes in the scores of adaptive functioning, the amount of time that the OTs spent in school over two semesters, and the receiving of extracurricular treatments. The results indicated that students' adaptive functioning made significant improvements, evident in the students' raw scores for the VABS-C and GAS. Moreover, the relationships between the changes of adaptive functioning and the amount of time that the OT spent in school over two semesters and the receiving of extracurricular treatments were found to be significant. The present study fills important gaps in the previous research to provide a better understanding of the role of OTs in professional teams and the changes in the adaptive functioning of students with ASD.

Consistent with the previous research, this study showed that the ASD students' adaptive functioning had mild or moderate impairments concerning not only their language and socialization skills but also their daily living and motor skills (Wang, Hsu, & Huang, 2009; Matthews et al., 2015). We used VABS-C as a standard measurement to investigate the difference between the pre-test and post-test adaptive functioning scores. The raw scores of the composite and domain of daily living skills showed significant differences. In the GAS, the students showed significant improvements after receiving the OTs'

services. The results indicated that the students' adaptive functioning improved distinctly but in a slow and subtle way. Consistent with the previous study of professional teams in Kaohsiung (Wuang & Wang, 2002), this study showed that the main problems for OTs in professional teams to overcome include (a) poor motor coordination, (b) inattention, and (c) poor daily living skills. The results from this study support the role of OTs in professional teams for improving students' daily living skills.

When comparing the contents of the GAS with those of the VABS-C, we found that the improvements of socialization were not evident in the changes of the VABS-C scores, despite finding a significant relationship between the change in scores for socialization and the amount of time that the OTs spent at the school over two semesters. The reasons for this discrepancy might be that the VABS-C is not sensitive enough to detect the micro changes; the standard scores and norms based on typical children's speedy development might not distinguish between the micro changes of the ASD students. Additionally, the low amount of time that the OTs spent at the school was not enough to make obvious improvements among the students. For school-based services, therapists tend to design problem-oriented programs to deal with the predicaments encountered by the teachers directly (Bazyk & Case-Smith, 2010). However, limited services and inefficient treatments could lead OTs in professional teams to abandon socialization training for students with ASD because it usually requires intensive treatments. The amount of time that the OTs spent at the school over two semesters was relatively low. In two previous studies about the effects and collaboration of professional teams, OTs and

physical therapists in the teams provided school-based services for children with special needs once to twice a week at school and spent extra time discussing and having meetings (Lu, 2006; Wang, Hsu, & Huang, 2009). The clients that they serviced improved at schools. Additionally, they indicated that the more efficient professional teams need sufficient time to supply services to students and effect change.

Having extracurricular treatments was a major confounding factor in the present study. The changes in scores for communication correlated with receiving extracurricular treatments. Communication is a key deficit of individuals with ASD and one of the most common concerns of parents and teachers (Azad & Mandell, 2016), and it may be treated reasonably as a primary goal in multidisciplinary clinics.

Do OTs in professional teams in Tainan play their roles efficiently? Education is an essential occupation for students, and school is the key environment. Students with ASD could have disabilities related to academic (e.g., math, reading, writing), nonacademic (e.g., recess, lunch, self-help skills), extracurricular (e.g., sports, clubs), and prevocational and vocational activities. The professional competence of OTs is qualified in the school-based situation. Occupational therapy services could promote self-help skills, positioning, sensorimotor processing, fine motor performance, psychosocial functioning, and life skills training (Bazyk & Case-Smith, 2010). Adaptive functioning is a key ability for independent living that needs to be practiced in varied environments. Promoting adaptation is the central theme for OTs. Ideally, OTs could play a more valuable role in school-based services by broadening the

participation and achievements of students with special needs. However, most of the OTs in Tainan face many challenges related to limited time for services and inconvenient communication. The OTs in this study recommended that increasing the frequency of services, providing direct communication, or recruiting full-time OTs might help students with ASD more efficiently.

Students' changes should be assessed regularly through interventions to make appropriate adjustments to the programs and services. Even though students do not sometimes receive direct treatments, their therapists, teachers, and parents should identify their improvements or understand the reason of stagnation to render the service more efficient. For now, no unified measurements are used in professional teams. In this study, we attempted to promote the GAS to OTs to evaluate students' changes, which demonstrated concrete and measurable results. While using the GAS to specify every single goal, OTs could refine the goals rather than describe a narrative-style note. The GAS can provide evidence for therapists' recommendations to initiate, increase, decrease, or discontinue therapy services (Bazyk & Case-Smith, 2010). The GAS could also act as a medium to communicate with parents and teachers to show the effects of professional teams

The present study has three limitations. First, the findings could not be generalized to all students with ASD in elementary schools because the present study included only the students who received occupational therapy services in school. As some of the participants received multiple-profession services including occupational therapy, speech therapy, and/or physical therapy in school, some improvements may have derived from the team-based

collaborative services. The effects of a multiple-profession group in school services should be examined in the future as well. Second, the design of this study did not have a control group with which to compare the effects of OTs in professional teams; thus, some latent factors might have affected the results. In addition, we could not eliminate the effects for natural maturity. Third, as the GAS scores were rated only by the OTs at the schools, the students' behavioral changes may not be consistent in other contexts. Therefore, future studies should recruit students with ASD receiving and not receiving occupational therapy services in schools, include control groups, and recruit parents and teachers to rate the GAS scores to ensure valid improvements.

5. Conclusion

This research study investigated the effects of OTs in professional teams on students with ASD and examined the relationships between other factors that may influence the effects. The present study used a standard measurement to investigate the changes in adaptive functioning among these students. Meanwhile, this study provided a quantitative measurement, the GAS, to monitor the improvements of students with ASD at school. The results indicated significant improvements in the students' adaptive functioning based on the changes of the raw scores for the VABS-C and GAS. The present study fills a gap in the previous research to provide a better understanding of the role of OTs in professional teams.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Azad, G., & Mandell, D. S. (2016). Concerns of parents and teachers of children with autism in elementary school. *Autism, 20*, 435-441.

Ashburner, J., Ziviani, J., & Rodger, S. (2008). Sensory processing and classroom emotional, behavioral and educational outcomes in children with autism spectrum disorder. *American Journal of Occupational Therapy, 62*, 556- 565.

Bazyk, S. & Case-Smith, J. (2010). School-based occupational therapy. In J. Case-Smith and J. O'Brien (Eds.), *Occupational therapy for children* (6th ed., pp. 713-743). St. Louis, MO: Elsevier Mosby.

Bonnard, M., & Anaby, D. (2016). Enabling participation of students through school-based occupational therapy services: Towards a broader scope of practice. *British Journal of Occupational Therapy, 79*, 188-192.

Cartmill, L., Rodger, S., & Ziviani, J. (2009). Handwriting of eight-year-old children with autism spectrum disorder: An exploration. *Journal of Occupational Therapy, Schools, & Early Intervention, 2*, 103-118.

Farley, M., A., McMahon, W. M., Fombonne, E., Jenson, W. R., Miller, J., & Gardner, M. (2009). Twenty-year outcome for individuals with autism and average or near-average cognitive abilities. *Autism Research, 2*, 109-118.

Green, D., Charman, T., Pickles, A., Chandler, S., Loucas, T., Simonoff, E., & Baird, G. (2009). Impairment in movement skills of children with autistic spectrum disorders. *Developmental Medicine & Child Neurology, 51*, 311-316.

Hume, K., Loftin, R., & Lantz, J. (2009). Increasing independence in autism spectrum disorder: A review of three focused interventions. *Journal of Autism and*

Developmental Disorders, 39, 1329-1338.

Kanne, S. M., Gerber, A. J., Quirmbach, L. M., Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2011). The role of adaptive behavior in autism spectrum disorders: Implications for functional outcome. *Journal of Autism and Developmental Disorders, 41*, 1007-1018. doi: 10.1007/s10803-010-1126-4

Kasari, C., & Smith, T. (2013). Interventions in schools for children with ASD: Methods and recommendations. *Autism, 17*, 254-267.

Kiresuk, T. J., & Sherman, R. E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community Mental Health Journal, 4*, 443-453.

Klin, A., Saulnier, C. A., Sparrow, S. S., Cicchetti, D. V., Volkmar, F. R., & Lord, C. (2007). Social and communication abilities and disabilities in higher functioning individuals with autism spectrum disorders: The Vineland and the ADOS. *Journal of Autism and Developmental Disorders, 37*, 748-759. doi: 10.1007/s10803-006-0229-4

Lai, M. L., & Yen, C. H. (2006). Analysis on the satisfaction of elementary school special education teachers toward the professional team operation. *Journal of Disability Research, 4*, 18-31.

Lu, Y. T. (2006). The effectiveness of implementing transdisciplinary teamwork approach in a special school: A case study of a student with severe and multiple disability. *Journal of Disability Research, 4*, 83-96.

Matthews, N. L., Pollard, E., Ober-Reynolds, S., Kirwan, J., Malligo, A., & Smith, C. J. (2015). Revisiting cognitive and adaptive functioning in children and adolescents with autism spectrum disorder. *Journal of Autism & Developmental Disorders, 45*, 138-156. doi: 10.1007/s10803-014-2200-0

Patel, V. B., Preedy, V. R., & Martin, C. R. (2014). *Comprehensive guide to autism*. New York: Springer. doi: 10.1007/978-1-4614-4788-7

Perry, A., Flanagan, H. E., Geier, J. D., & Freeman, N. L. (2009). Brief report: The Vineland Adaptive Behavior Scales in young children with autism spectrum disorders at different cognitive levels. *Journal of Autism & Developmental Disorders, 39*, 1066-1078. doi: 10.1007/s10803-009-0704-9

Portney, L. G., & Watkins, M. P. (2007). *Foundations of clinical research applications to practice, Third Edition:* Prentice Hall.

Ruble, L., McGrew, J. H., & Toland, M. D. (2012). Goal attainment scaling as an outcome measure in randomized controlled trials of psychosocial interventions in autism. *Journal of Autism and Developmental Disorders, 42*, 1974-1983.

Schlosser, R. W. (2004). Goal attainment scaling as a clinical measurement technique in communication disorders: a critical review. *Journal of Communication Disorders, 37*, 217-239.

Sparrow, S. S., Balla, D. A., & Cichetti, D. V. (1985). *Vineland Adaptive Behavior Scales: Survey Form.* Circle Pines, MN: American Guidance Service.

Sparrow, S. S., Cicchetti, D., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales-2nd edition manual*. Circle Pines, MN: AGS Publishing.

Wang, T. N., Hsu, Y. W., & Huang, T. H. (2009). Sensory motor and functional performances of children with special needs referred to occupational therapy by school teachers in Tainan City. *Journal of Taiwan Occupational Therapy Research and Practice, 5*, 15-26.

Wu, W. T., Chang, C. F., Lu, T. H., & Chiu, S. C. (2004). *Vineland adaptive behavior scales: Chinese version manual.* (Sparrow, S. S., Balla, D. A., & Cicchetti, D. V.) Taipei, Taiwan: Psychological (Original work published 1984)

Wuang, Y. P., & Wang, C. C. (2002). Occupational therapy in collaborative teams for children with special needs in Kaohsiung City. *Journal of Occupational Therapy Association, R. O. C., 20*, 75-94.

Yang, G. W., & Cherng, R. J. (2013). The efficiency of the professional team model in Taiwan special education system: A systematic review. *Formosan Journal of Physical Therapy, 38*, 286-298.

Yen, C. H. (2005). An exploratory study of the evaluations on special education disciplinary teamwork. *Journal of Disability Research, 3*, 216-228.

職能治療學會雜誌
民國 107 年 36 卷第 2 期頁 152-176
DOI:10.6594/JTOTA.201812_36(2).0002

原著

探討特殊教育專業團隊中職能治療師改善台南市國小自閉症學童適應性行為的角色

OCCUPATIONAL THERAPY

孫維黛 [1] 林玲伊 [1,*]

摘要

特殊教育專業團隊模式在台灣推行至今已近二十年，過去文獻探討此介入模式時，多數著重於內部成員的合作現況、相互滿意度之調查，卻少深入探討專業團隊或職能治療師的介入是否有助於提升特殊需求學童的各項能力。故此研究旨在探討特殊教育專業團隊中職能治療師是否能有效提升自閉症學童在校的適應性行為能力。此研究招募台南市十九位在一般國小就學自閉症學童（16 位男童，3 位女童，平均年齡 8.79 歲）。分別於職能治療師介入前以及介入二學期後，以文蘭適應行為量表中文編譯版和目標達成量表評估自閉症學童在校適應行為的變化。研究結果顯示，自閉症學童在接受學校職能治療師服務前後，其適應性行為的合成分數 ($Z = -2.069, p = .039$) 與日常生活能力 ($Z = -2.536, p = .011$) 的變化有顯著性差異。學童的進步亦顯著呈現於目標達成量表中從介入前的 37.60 分 ($SD = 1.47$) 到介入後的 52.55 分 ($SD = 6.98; Z = -3.417, p = .001$)。職能治療師到校服務的時間顯著與社會互動有正相關 ($r = .489, p = .046$)。日常生活能力、社會互動和動作的原始得分差異與職能治療師到校服務的時間顯著相關。職能治療師在專業團隊服務中對於學童在校的適應性行為扮演重要的角色，此研究利用量化評估工具可為特殊教育專業團隊中的職能治療服務提供實證基礎。

關鍵字：自閉症，適應性行為，職能治療師，特殊教育專業團隊

國立成功大學醫學院職能治療學系 [1]

*通訊作者：林玲伊
701 台南市東區大學路一號
電話：06-2353535 分機 5917
電子信箱：lingyi@mail.ncku.edu.tw

受文日期：民國 107 年 08 月 31 日
接受刊載：民國 107 年 11 月 07 日

職能治療學會雜誌
民國 107 年 36 卷第 2 期頁 177-202
DOI:10.6594/JTOTA.201812_36(2).0003

文獻評論

洛文斯坦職能治療認知評估第二版應用於臺灣不同障礙族群之心理計量特性回顧

李沛綺[1] 蔡沛潔[2] 黃怡靜[1] 李士捷[1] 董雅綺[1] 謝清麟[1,3,4] 黃小玲[1,3,*]

摘要

認知功能是維持日常生活功能及工作的關鍵能力，但腦部疾病多伴隨認知功能的缺損。洛文斯坦職能治療認知評估第二版為目前臨床常用之認知功能評估工具組，然而國內在各障礙族群之心理計量特性驗證文獻相當有限。本研究回顧並彙整國內應用 LOTCA-II 於不同個案群之心理計量特性文獻。

作者合併檢索中、英文之電子期刊資料庫中驗證於臺灣各障礙族群之LOTCA-II 之心理計量特性文獻。結果顯示 LOTCA-II 總分應用於臺灣中風及創傷性腦傷個案整體信度大致良好，但僅部份分量表之內在一致性達良好。其與魏式智力測驗之收斂效度良好，且可區辨個案和健康成人間之認知功能的差異。於思覺失調症個案信度大致良好，其與魏式智力測驗及日常生活功能測驗之收斂效度良好，但無法區辨個案之就業情況。於智能障礙個案中內在一致性大致良好，其與圖形智力測驗之收斂效度良好，但無法區辨個案認知功能受損的嚴重程度。另外，雖有驗證因素效度的研究，但因採用混合樣本，結果難以概化。

LOTCA-II 於臺灣中風、創傷性腦傷、思覺失調症及智能障礙等族群皆尚未驗證同時效度、預測效度及反應性，且部份已驗證之信、效度結果不佳，亦缺乏驗證各障礙族群之因素效度。未來研究需更深入驗證 LOTCA-II 之心理計量特性，以提供更完整的實證參考數據。

關鍵字：洛文斯坦職能治療認知評估第二版，中風，思覺失調，認知功能，心理計量特性

臺灣大學醫學院職能治療學系[1]
新光醫療財團法人新光吳火獅紀念醫院復健科職能治療[2]
臺大醫院復健部職能治療[3]
亞洲大學醫學暨健康學院職能治療學系[4]

*通訊作者：黃小玲
台北市中正區徐州路 17 號 4 樓
電話：02-33668179
電子信箱：cathy@ntu.edu.tw

受文日期：民國 107 年 02 月 22 日
接受刊載：民國 107 年 07 月 14 日

前言

認知功能是執行日常活動及維持工作的重要基本能力。無論是中風、創傷性腦傷、思覺失調症或智能障礙等個案，多伴隨不等程度的認知功能障礙；輕微的可能僅僅注意力渙散、精神無法集中；嚴重的會影響理解力及記憶力，甚至生活無法自理 (Lezak, 2004) 。全面客觀的評估及確認個案認知功能障礙之面向與程度，有助臨床及研究人員釐清個案的問題，提出適當的治療計畫，改善認知功能，進而提升日常生活功能及品質。因此，臨床人員應使用具良好心理計量特性及臨床適用性之認知評估工具，以提升臨床效能。

心理計量特性（包括信度、效度及反應性）是臨床及研究人員挑選評估工具之參考重點。信度為評估工具重複評估之一致程度及穩定性，評估工具若有良好的穩定性，則代表此工具所收集的數據可被相信 (Hobart, Lamping, & Thompson, 1996; Kline, 1998) 。良好的信度是評估工具發展之必要條件，若評估工具無良好之信度，就算有良好之效度亦為空談。信度包含內在一致性、再測信度及施測者間信度。內在一致性代表一評估工具中所有項目之同質性。項目同質性越高，評估結果越趨穩定 (Hobart et al., 1996; Portney & Watkins, 2009) 。再測信度是評估工具追蹤評估及驗證療效的先決條件，可確保評估工具所反應的改變量是否主要來自於受測者本身的改變。施測者間信度可確認評估工具評估結果的客觀性。效度為評估工具是否能正確評測到欲評量之特質，亦可藉由其數據瞭解個案目前之能力，並且推論個案預後，這也是評估工具發展之必要條件 (Aaronson et al., 2002; Hobart et al., 1996; Kline, 1998; Portney & Watkins, 2009) ，常用之效度包含同時效度、收斂效度、預測效度、已知族群效度及因素效度。同時效度為確認評估工具所評估之概念，是否與同一時間施測相同概念之黃金標準 (gold standard) 一致 (Bowling, 2001; Hobart et al., 1996) 。若欲評估之概念無一黃金標準，則可用收斂效度驗證評估工具與欲評估之特質相關聯的程度 (Van Deusen & Brunt, 1997; Wade, 1992) 。預測效度為評估工具評估之結果預測未來的健康狀態之程度 (Kline, 1998) 。已知族群效度為評估工具可區辨不同族群間之能力程度差異，例如：個案群與健康群、能力低與能力高之個案群 (McCarthy et al., 2002; Wade, 1992)。因素效度包含探索性因

素分析 (exploratory factor analysis) 及驗證性因素分析 (confirmatory factor analysis)，目的為發掘或確認特定量表所包含的因素結構，以協助驗證量表所評量的概念是否吻合理論模型。探索性因素分析適用於尚無可能的理論模型，需以實證資料探索項目背後可能的潛在結構 (Costello & Osborne, 2005)；而驗證性因素分析則適用於已有理論模型，目的為驗證該模型能否解釋實證資料 (Schreiber, Nora, Stage, Barlow, & King, 2006)。反應性為評估工具可偵測個體或群體欲評估之目標特質變化的能力，當療效評估工具有適當的反應性，才能靈敏的偵測出治療效果，因此反應性也是評估工具良莠的重要指標 (Barak & Duncan, 2006; Beaton et al., 2001; Hutchinson, 1992)。

洛文斯坦職能治療認知評估第二版 (Loewenstein Occupational Therapy Cognitive Assessment, LOTCA-II) (Itzkovich, Elazar, Averbuch, & Katz, 2000) 是 Itzkovich 等學者改良自第一版 LOTCA (Itzkovich, Elazar, Averbuch, & Katz, 1990)，其修改內容主要有二：(1) 調整計分方式；(2) 增加評估面向。LOTCA-II 主要依據神經心理學、發展理論與臨床經驗等為基礎所編製而成，可用於評估個案認知功能損傷之程度，亦可用於檢視認知功能介入之成效(Annes, Katz, & Cermak, 1996)。LOTCA-II 主要評估的內容為每日執行活動所需的基本認知能力，包括：定向感 (orientation)、視知覺 (visual perception)、空間知覺 (spatial perception)、動作計畫 (motor praxis)、視覺動作組織 (visuomotor organization) 以及思考操作 (thinking operations) 等 6 個分量表，共 26 個子題；施測所需時間平均約 45 分（30-90 分鐘）。LOTCA-II 之計分，將原本「定向感」分量表的測驗子題改為 1-8 分，其他測驗子題維持 1-4 分；「思考操作」分量表中有 3 個子題為 1-5 分，總分為 26-115 分，分數愈低代表認知功能愈差 (Itzkovich et al., 2000)。目前已應用於中風、失智症、創傷性腦傷、智能障礙、精神疾病、成癮等個案 (Annes et al., 1996; Jang, Chern, & Lin, 2009; Josman & Katz, 2006; Rojo-Mota et al., 2017)。臺灣目前有兩組學者將 LOTCA-II 的指導語翻譯為中文，施測用的題本及評分方式相同。此兩版本僅部份名詞翻譯不同，但詞義相同。

目前雖有諸多國家研究 LOTCA-II 之心理計量特性（Avila et al., 2015; Jang et al., 2009; Josman, Abdallah, & Engel-Yeger, 2010; Josman, Abdallah, & Engel-Yeger,

2011; Lund, Oestergaard, & Maribo, 2013, 2014; Natar, Nagappan, Ainuddin, Masuri, & Thanapalan, 2015; Pyun et al., 2009; Rojo-Mota et al., 2017; Su, Chen, Tsai, Tsai, & Su, 2007; Su, Lin, Chen-Sea, & Yang, 2007; 蔡佳殷, 2005; 蔡佳殷, 陳惠媚, & 蘇純瑩, 2007; 蘇純瑩, 陳惠媚, 張雅雯, 蔡佳殷, & 林月仙, 2007）。然而，因為心理計量特性易受樣本或族群差異之影響 (Jongbloed, 1986) ，故國外的心理計量特性證據難以直接類推至本國國人，且不同診斷之個案群雖使用同一評估工具，但無法互相類推評估工具之心理計量特性。在臺灣，LOTCA-II 是臨床實務上常施測的認知評估工具組，陸續應用在多個障礙族群，為全面瞭解截至目前為止有關 LOTCA-II 應用於國人之心理計量特性相關文獻，提供實證結果給臨床或研究使用者正確判讀 LOTCA-II 的評估結果。因此本研究回顧並彙整國內驗證 LOTCA-II 之心理計量特性文獻，裨以提供完整的實證基礎供後續臨床及研究人員參考。

研究方法

本研究主要目的為檢索並彙整（2000 年 1 月至 2017 年 10 月）LOTCA-II 應用於國內不同障礙族群之心理計量特性，包括以下三步驟：「搜尋資料庫」、「篩選文獻」與「彙整/評析心理計量特性」，詳述如下：

一、搜尋資料庫

作者分別於 PubMed、MEDLINE、PsycINFO、CINAHL 與華藝線上圖書館（包含期刊論文及碩博士論文）等五個資料庫檢索 2000 年 1 月至 2017 年 10 月間 LOTCA-II 之心理計量特性研究。檢索策略為：在各資料庫中分別檢索 LOTCA-II 及心理計量特性（如：信度、效度、反應性及天花板/地板效應）相關之醫學標題詞彙及關鍵字，再取檢索結果之交集，以獲得同時涵蓋二者之文獻。詳細之 MEDLINE 及 PubMed 文獻檢索策略列於表 1、表 2。

表 1

MEDLINE 檢索策略

編號	檢索策略
S1	TX ((Loewenstein N1 Occupational N1 Therapy N1 Cognitive N1 Assessment) or LOTCA$)
S2	MH "psychometrics"
S3	MH " reproducibility of results "
S4	MH " exp sensitivity" and MH "specificity"
S5	reliabilit*
S6	validit*
S7	responsiveness
S8	sensitivit*
S9	specificit*
S10	(psychometric N1 properti*)
S11	(psychometr* or clinimetr* or clinometr*)
S12	(reliab* or reproducib$* or test-retest or interater or inter-rater or intrarater of intra-rater)
S13	(valid* or responsive*)
S14	S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13
S15	S1 AND S14
S16	limit S15 to English language
S17	limit S16 to evaluation studies
S18	limit S16 to validation studies
S19	S17 OR S18

S1: LOTCA 相關詞彙，S2–S13: 心理計量特性相關詞彙，S16: 限制英文撰寫之文獻，S17–S19: 限制文獻類型

表 2

PubMed 檢索策略

編號	檢索策略
#1	psychometrics[MeSH] OR reproducibility of results[MeSH] OR Sensitivity and Specificity[MeSH] OR realibilit*[TW] OR validit*[TW] OR responsiveness[TW] OR specifict*[TW] OR sensitivit*[TW]
#2	" Loewenstein Occupational Therapy Cognitive Assessment "[TW] OR LOTCA *[TW]
#3	1 AND 2

#1: 心理計量特性相關詞彙，#2: LOTCA 相關詞彙

二、篩選文獻

同時符合下列三個篩選條件的文獻，則予以收錄：1.明確指出研究目的為驗證 LOTCA-II 心理計量特性之文獻。2.主要樣本為臺灣人。3.英文（或中文）期刊論文（或碩博士論文）。但若該論文為文獻回顧、評論、信函、演說或社論則予以排除。

三、心理計量特性之種類及評析標準

作者彙整並評析所得論文之心理計量特性。評析項目及標準如下：

（一）信度

包括內在一致性、再測信度及施測者間信度。內在一致性以克隆巴赫係數 (Cronbach's α) 為指標，$\alpha > .70$ 為良好，可用於群體分數之比較；>.90 為優，可用於個案間分數的比較 (Aaronson et al., 2002) 。再測信度與施測者間信度以組內相關係數 (intra-class correlation coefficient, ICC) 為指標，ICC ≥ .75 為良好；介於.40~ .74 為中等；< .40 為差 (Shrout & Fleiss, 1979) 。再測信度亦可以皮爾森積差相關係數 (Pearson's r) 或斯皮爾曼相關係數 (Spearman's ρ) 為指標， $r/\rho \geq .75$ 為良好；介於.40~ .74 為中等；< .40 為差 (Van Deusen & Brunt, 1997) 。

（二）效度

包含同時效度、收斂效度、預測效度、已知族群效度及因素效度。同時效度以 Pearson r/Spearman ρ 為指標，r/ρ >.75 為良好；介於.40~ .74 為中等；< .40 為差 (Van Deusen & Brunt, 1997) 。收斂效度同樣以 Pearson r/Spearman ρ 為指標，r/ρ > .50 為良好，或依據各次量表與效標工具理論上之相關程度判斷之 (Salter et al., 2005) 。若收錄論文驗證同時效度所使用的效標工具並非黃金標準評估工具，在本研究則併入收斂效度的驗證報告。已知族群效度以 t 檢定值或 F 檢定值是否達顯著為指標，若有顯著差異則代表評估工具可區辨不同族群或嚴重程度的個案相同特質 (McCarthy et al., 2002) 。

驗證性因素效度為建構效度之一，是藉由因素分析以檢驗某一評估工具與其理論模型是否相符，以驗證工具是否測量到其欲測量之特性。驗證性因素分析 (confirmatory factor analysis, CFA) 之模型適配度指標為卡方檢定值 (chi-square index) 、適配度指標 (goodness-of-fit index, GFI) 、調整後適配度指標 (adjusted goodness-of-fit index AGFI) 、規範適配指標 (normed fit index, NFI) 、非規範適配指標 (Tucker-Lewis index, TLI) 、增值適配指標 (Bollen's incremental fit index, IFI) 、均方根近似誤 (root mean square error of approximation, RMSEA) 、標準差均方根殘差 (standardized root mean square residual, SRMR) 及比較適配指標 (comparative fit index) (Bentler, 1990; Bollen, 1990) 。適配度指標 (GFI, AGFI, CFI, TLI, IFI, NFI) 高於.95 為極佳的模型適配度，介於.85~.90 為模型適配度尚可；RMSEA < .05 則為模型有良好的適配程度，RMSEA > .1 則適配度較差；SRMR < .08 則模型適配度較佳 (Bentler, 1990; Bollen, 1990; Hu & Bentler, 1999) 。若模型適配度佳，則代表該模型符合可解釋評估結果之分數，故支持使用者以分量表之分數反應其潛在的影響因素。

（三）反應性

反應性之判定指標為效應值 (effect size, ES) ；ES= .2 代表低度效應；ES= .5 代表中度效應；ES= .7 代表高度效應 (Salter et al., 2005) 。

（四）天花板/地板效應

以描述性統計呈現各量表最高/最低分之人數比例。若超過 20%的受試者在該分量表/總分獲得最高或最低分，則代表此評估工具有顯著的天花板或地板效應 (Holmes, Bix, & Shea, 1996) 。

結果

作者於 PubMed、MEDLINE、PsycINFO、CINAHL 與華藝線上圖書館等五個資料庫檢索 2000 年 1 月至 2017 年 10 月間驗證 LOTCA-II 之心理計量特性研究，

檢索之後共發現 18 篇論文。我們再依據文章標題、摘要或全文逐一檢視這 18 篇論文，刪除非應用於臺灣人之論文 12 篇，最後收錄 6 篇論文（5 篇期刊論文及 1 篇碩士論文）（圖 1）。這 6 篇論文分別應用於健康成人（蔡佳殷，2005；蔡佳殷等人，2007；蘇純瑩等人，2007）、慢性中風個案（Su, Lin, et al., 2007;蔡佳殷，2005；蔡佳殷等人，2007；蘇純瑩等人，2007）、創傷性腦傷個案 （Su, Lin, et al., 2007; 蔡佳殷等人， 2005;蘇純瑩等人，2007）、思覺失調症個案 (Su, Chen, et al., 2007; Su, Lin, et al., 2007) 、及智能障礙者 (Jang et al., 2009) 。我們整理 LOTCA-II 應用於各個障礙族群之心理計量特性如下（表 3）：

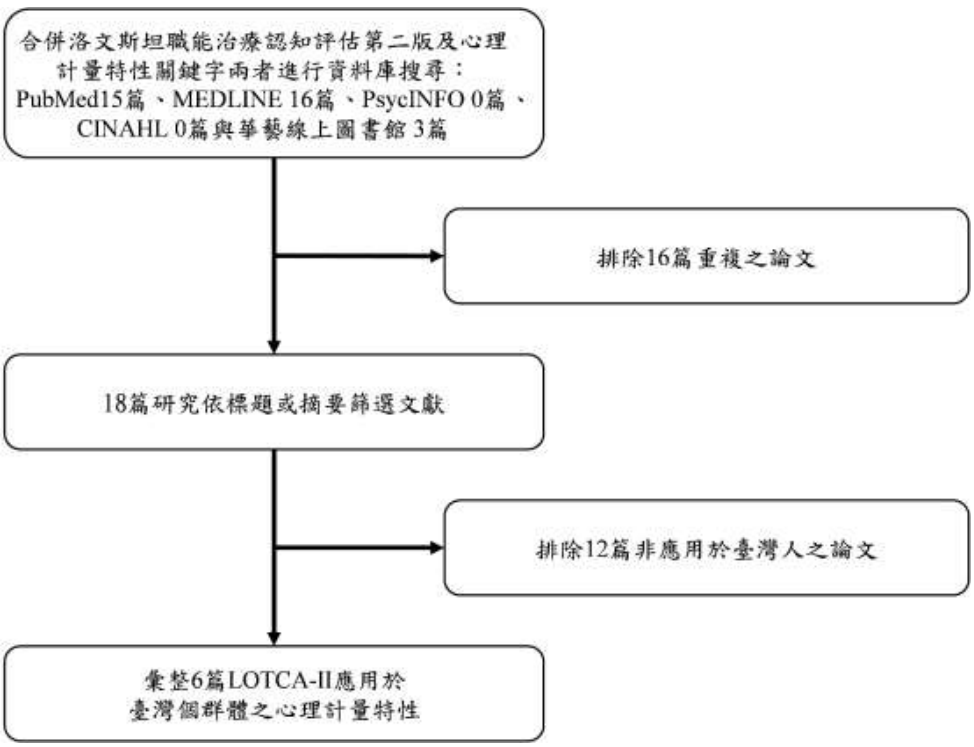

圖 1
文獻回顧之檢索流程圖

一、慢性中風個案

LOTCA-II 應用於臺灣慢性中風個案（指中風超過三個月以上者），信度的驗證包括：內在一致性、再測信度及施測者間信度（蔡佳殷，2005）。效度的驗證包括收斂效度、已知族群效度及因素效度（蔡佳殷，2005）。此個案群亦有驗證天花板/地板效應。

信度方面，LOTCA-II 測驗之內在一致性為優 (α= .91) ，但各分量表之內在一致性參差不齊 (α= .43~ .84) ，其中「視覺動作組合」(α= .84) 及「思考操作」(α= .84) 兩個分量表較為良好。LOTCA-II 的再測信度良好 (r= .95) ，僅「動作計畫」分量表 (r= .69) 與「定向感」分量表 (r= .77) 為中度，其餘分量表都良好 (r= .86~1.00) 。LOTCA-II 的施測者間信度良好 (r= .99) ，各分量表亦皆為良好 (r= .91~1.00) （蔡佳殷，2005）。

收斂效度方面，LOTCA-II 測驗與魏氏成人智力測驗第三版量表 (Wechsler Adult Intelligence Scale-III , WAIS-III) 之相關性達良好等級 (r= .70) ，代表收斂效度良好，分量表中僅「空間知覺」、「視覺動作組合」及「思考操作」達良好等級 (r= .55~.71) ；其與 WAIS-III 語言智商之相關性達良好等級 (r= .64) ，分量表中「視知覺」、「視覺動作組合」及「思考操作」達良好等級 (r= .51~.67) ；與 WAIS-III 操作智商量表之相關性達良好等級 (r= .64) ，分量表中則有「空間知覺」、「視覺動作組合」及「思考操作」達良好等級 (r= .50~.63) （蔡佳殷，2005）。LOTCA-II 測驗與簡短智能測驗 (Mini-Mental State Examination; MMSE) 總分之相關性達良好等級 (r= .69) ，代表收斂效度良好，但分量表中僅「視知覺」、「視覺動作組合」及「思考操作」達良好等級 (r= .56~.66) （蔡佳殷，2005）。

在已知族群效度的部分，蔡佳殷發現健康成人在 LOTCA-II 的總分及分量表分數都高於中風個案，但僅在「思考操作」分量表有顯著差異。中風組 (n=31) 與正常組 (n=31) 整體區別正確率為 66.1%，中風組正確預測率為 58.1%，正常組的正確預測率為 67.2%（蔡佳殷，2005）。

天花板/地板效應方面：蘇純瑩的研究結果顯示「定向感」、「視知覺」及「空間知覺」分量表有顯著的天花板效應；於「動作計畫」分量表之 2 項子題、「視覺

動作組合」6 項子題及「思考操作」4 項子題都有顯著的天花板效應。地板效應的部分，「思考操作」的「邏輯問題」子題有顯著的地板效應（蔡佳殷等人，2007）。

二、創傷性腦傷個案

LOTCA-II 應用於臺灣創傷性腦傷個案的研究，信度的驗證包括內在一致性、再測信度、施測者間信度（蔡佳殷，2005）。效度的驗證包括收斂效度、已知族群效度及因素效度（蔡佳殷，2005）。

LOTCA-II 測驗之內在一致性為優 (α= .94) ，其中「空間知覺」分量表的內在一致性較差 (α= .56)，其餘分量表的內在一致性有中等至良好 (α= .71~ .94) 。LOTCA-II 整體測驗與各分量表的再測信度皆良好 (相關係數分別為 r= .98; r= .89~.98) ，施測者間信度亦良好 (相關係數分別為 r= .99; r= .99~1.00) （蔡佳殷，2005）。

收斂效度方面，LOTCA-II 測驗與魏氏成人智力測驗第三版全量表智商之相關性達良好等級 (r=.61) ，分量表中「視知覺」、「空間知覺」、「視覺動作組合」及「思考操作」達良好等級 (r= .50~.61) ；其語文智商量表之相關性達良好等級 (r=.54) ，代表收斂效度良好，但分量表中僅「空間知覺」達良好等級 (r=.63) ；其與操作智商量表之相關性達良好等級 (r=.56) ，但分量表中僅「視覺動作組合」及「思考操作」達良好等級 (r= .56~.58) （蔡佳殷，2005）。LOTCA-II 測驗與 MMSE 總分之相關性達良好等級 (r=.81) ，代表收斂效度良好，分量表中「定向感」、「視覺動作組合」及「思考操作」達良好等級 (r=.65~.68) （蔡佳殷，2005）。

在已知族群效度的部分，蔡佳殷發現健康成人在 LOTCA-II 的總分及分量表分數都高於創傷性腦傷個案，而且在所有分量表都有顯著差異（蔡佳殷，2005）。

三、思覺失調症個案

LOTCA-II 已驗證於臺灣之思覺失調症個案之內在一致性及間隔 7 天之再測信度、收斂效度、已知族群效度及天花板/地板效應。

信度部分，LOTCA-II 之「定向感」、「視覺動作組合」及「思考操作」分量表之內在一致性介於中等至良好 (α= .77~ .91) ；而「視知覺」、「空間知覺」及「動作計畫」分量表之內在一致性差 (α= .20~ .45) 。再測信度方面，「視知覺」、「視覺動作組合」及「思考操作」分量表之再測信度介於中等至良好 (ICC= .77~ .89) ；而「定向感」、「空間知覺」及「動作計畫」分量表之再測信度較差 (ICC= .49~ .67) (Su, Chen, et al., 2007) 。

收斂效度方面，LOTCA-II 總分與魏式智力測驗相關性達良好等級 (r= .53~ .69) ，代表收斂效度良好；分量表僅「視覺動作組合」及「思考操作」有良好收斂效度 (r= .55~ .64) 。LOTCA-II 總分與評估執行功能的威斯康辛卡片分類測驗 (Wisconsin Card Sorting Test, WSCT) 相關性達良好等級 (r= .56) ，有良好收斂效度；分量表中僅「思考操作」有良好收斂效度 (r= .55) 。LOTCA-II 總分與評估認知功能的艾倫認知階層測驗 (Allen Cognitive Level Screen, ACLS) 相關性達良好等級 (r= .55) ，有良好效度；分量表中僅「視覺動作組合」及「思考操作」有良好效度 (r= .51~ .58) ；LOTCA-II 總分與褚氏日常生活功能表相關性達良好等級 (r= .55) ，有良好效度；分量表中僅「思考操作」有良好效度 (r= .57) (Su, Chen, et al., 2007) 。

已知族群效度方面，蘇純瑩以就業情況將思覺失調症個案分為三組，發現三組個案之 LOTCA-II 各分量表之間無顯著差異 (p= .13) (Su, Chen, et al., 2007) 。

天花板/地板效應方面，思覺失調症個案在各分量表子題中，「定向感」2 項子題、「視知覺」的 3 項子題、「空間知覺」的 2 項子題、「動作控制」的 2 項子題及「視覺動作組合」的 1 項子題都有顯著的天花板效應，而在「思考操作」的後 2 項子題有顯著的地板效應 (Su, Chen, et al., 2007) 。

四、智能障礙個案

LOTCA-II 已驗證於智能障礙個案之內在一致性，亦有驗證收斂效度、已知族群效度及天花板/地板效應。

信度部分，LOTCA-II 之內在一致性除了「動作計畫」分量表為低度 (α= .48)，其他五項分量表皆有中度至良好之內在一致性 (α= .74~ .86)。

收斂效度方面，圖形 IQ 測驗 (Pictorial IQ test) 與「視覺動作組合」及「思考操作」分量表相關性達良好等級 (ρ= .61~ .63)，代表有良好收斂效度 (Jang et al., 2009)。

在已知族群效度的部分，張彧學者等人將智能障礙個案依照障礙程度分成三組，並比較此三組與健康成人的 LOTCA-II 分數，健康組的得分為四組中最高，嚴重智能障礙組為得分最低，健康組與智能障礙組之間有顯著的差異，但三組不同障礙級別之間的差異相較於健康組與智能障礙組之間的差異較不顯著 (Jang et al., 2009)。

天花板/地板效應方面，智能障礙族群在「定向感」、「視知覺」、「空間知覺」及「動作計畫」有天花板效應，「視覺動作組合」的 5 項子題及「思考操作」的 6 項子題有顯著的地板效應 (Jang et al., 2009)。

五、混合樣本（慢性中風、創傷性腦傷及思覺失調症個案）

蘇純瑩等學者利用 CFA 驗證 LOTCA-II 應用於混合樣本（中風、創傷性腦傷及思覺失調症）之因素效度 (Su, Lin, et al., 2007)，包括 2 種因素模型： (1) 單因素模型 (one-factor model)：認為 LOTCA-II 的 6 個次量表都評量到單一認知面向；(2) 雙因素模型 (oblique two-factor model)：認為 LOTCA-II 共評估到二種認知面向，即「知覺記憶或應用功能（低階知覺認知技巧）」與「組織及概化功能（高階知覺認知技巧）」 (Su et al., 2000; Su, Lin, et al., 2007)。結果顯示：雙因素模型提供良好的適配度；相對地，單因素模型並未提供好的適配度，因為所有適配指標都遠低於標準值。

接著蘇學者進一步使用階層式雙因素模型 (hierarchical two-factor model) 分別檢驗在中風及創傷性腦部傷害個案及思覺失調個案的適配度。結果顯示：模型適配度都極佳，適配指標在神經個案 (GFI= .982, AGFI= .933, TLI=1.01, CFI=1, NFI= .986, RMSEA= .000, IFI=1.01, SRMA= .02) 及思覺失調症個案 (GFI= .980,

AGFI= .924, TLI=1.01, CFI=1, NFI= .979, RMSEA= .000, IFI=1.01, SRMA= .03) 都很好 (Su, Lin, et al., 2007) 。

另外，就我們所知，目前並無單獨驗證 LOTCA-II 應用於中風、創傷性腦傷、思覺失調症或智能障礙個案之因素效度研究。

表 3

LOTCA-II 心理計量特性信、效度彙整

心理計量特性/群體	中風	創傷性腦傷	思覺失調	智能障礙
相關文獻	蔡佳殷， 2005	蔡佳殷，2005	Su, Chen, et al., 2007	Jang et al., 2009
內在一致性 Cronbach's α	(*n*=31) 總分為良好 (α= .91)	(*n*=33) 總分為良好 (α= .94)	總分為良好(α= .90) ，定向感、視覺動作組合及思考操作之分量表為中等至良好 (α= .77~ .91) ，視知覺、空間知覺及動作計畫之分量表為低等 (α= .20~ .45)	定向感、視知覺、空間知覺、視覺動作組合、及思考操作有中等至良好 (*α*= .74~ .86) ，動作計畫分量表為低等 (*α*= .48)
施測者間信度 Pearson *r*	(*n*=20) 良好 (*r* = .99)	(*n*=20) 良好 (*r* = 1.00)	-	-
再測信度 ICC/ Pearson *r*	(*n*=22) 總分為良好 (*r* = .95)	(*n*=20) 總分為良好 (*r* = .98)	總分為良好 (ICC= .95) ，視知覺、視覺動作組合及思考操作分量表為中等至良好 (ICC= .77~ .89) ，定向感、空間知覺及動作計畫為低至中等 (ICC= .49~ .67)	-
地板效應 N (%)	-	-	思考操作有顯著的地板效應	部分視覺動作組合及思考操作子題有顯著的地板效應
天花板效應 N (%)	定向感、視知覺及空間知覺有顯著的天花板效應	-	定向感、視知覺、空間知覺及動作計畫有顯著的天花板效應	定向感、視知覺、空間知覺及動作計畫有天花板效應
同時效度 Pearson *r*/Spearman *ρ*	-	-	-	-

（表 3 續）

表 3 (續)

LOTCA-II 心理計量特性信、效度彙整

心理計量特性/群體	中風	創傷性腦傷	思覺失調	智能障礙
收斂效度 Pearson *r*/Spearman ρ	(*n*=20) 與 WAIS-III 總量表效度良好 (*r*= .70) 與 MMSE 總分效度良好 (*r*= .69) 與 BVRT 總分效度良好 (*r*= .69)	(*n*=20) 與 WAIS-III 總量表效度良好 (*r*= .61) 與 MMSE 總分效度良好 (*r*= .81) 與 BVRT 總分效度良好 (*r*= .65)	WAIS-III：與總分效度良好 (*r*= .53~ .69)，視覺動作組合及思考操作效度良好 (*r*= .55~ .64) WCST：與總分效度良好 (*r*= .56) ，思考操作效度良好 (*r*= .55) ACLS：與總分效度良好 (*r*= .55)，視覺動作組合及思考操作效度良好 (*r*= .51~ .58) DLFS：與總分效度良好 (*r*= .55) ，思考操作效度良好 (*r*= .57)	Pictorial IQ test：視覺動作組合及思考操作效度良好 (ρ= .61~ .63) 分量表間與思考操作有良好之收斂效度 (ρ= .60~ .73)
已知族群效度 Cohen's d/MANOVA	(n=31) 正常組之比較區別正確率為 66.1%，中風組正確預測率為 58.1%，正常組的正確預測率為 67.2%	(n=33) 與正常組之比較，區別正確率為 78.1%，頭部外傷組正確預測率為 63.6%，正常組的正確預測率為 93.5%	依就業度分為三組。整體而言分析結果顯示三組間無顯著差異 (Wilk's Lambda=.74, F(12,112)=1.52, *p*=.13)	智能障礙個案依障礙程度分成 4 組，整體而言以 K-W 檢定結果顯示健康組與智能障礙組之間有顯著的差異。
因素效度 EFA/CFA	-	-	以 CFA 檢驗結果顯示雙因素模型之模型適配度較好	-

討論

本研究為第一篇彙整 LOTCA-II 驗證於臺灣不同障礙族群之心理計量特性的論文。LOTCA-II 是臺灣臨床使用率極高的認知評估工具，經檢索分析，LOTCA-II 曾驗證於臺灣的四種障礙族群（包括中風、創傷性腦傷、思覺失調症及智能障礙者）。結果發現：雖有心理計量特性之證據呈現，但部分驗證結果不佳，部分尚未驗證，造成臨床應用及學術研究缺乏完整實證之支持。

目前四種障礙族群皆缺乏同時效度之驗證。因認知功能包含太多面向，至今尚無公認之黃金標準，故 LOTCA-II 目前僅能驗證收斂效度。

一、慢性中風個案

LOTCA-II 應用於中風個案的心理計量驗證研究不多，仍須後續研究深入探究。在臺灣，醫院復健部治療對象以中風個案為大宗，中風個案又多伴隨不等程度的認知功能障礙問題，因此 LOTCA-II 的使用非常頻繁，但目前僅有四篇研究探究 LOTCA-II 應用於本國中風個案之心理計量特性，且缺乏研究驗證同時效度、預測效度及因素效度，亦缺乏研究驗證其反應性。

在信度方面，LOTCA-II 於各分量表之內在一致性及再測信度分布在中等至良好等級。造成上述信度不均的可能原因有二： (1) 認知功能包含了許多面向，可能造成題目間相關性較低，進而導致內在一致性差； (2) 樣本數少，可能錯估樣本之變異程度，進而低估信度。例如，於我們所回顧之文獻中，再測信度及施測者間信度之驗證僅有 20 位受測者。未來研究應增加樣本數，以擴大驗證 LOTCA-II 應用於慢性中風個案之穩定性。收斂效度部分，目前僅驗證 LOTCA-II 與 WAIS-III 之相關性。收斂效度之驗證結果顯示為良好等級，表示 LOTCA-II 可評估到個案之智商能力。然而，目前尚缺乏研究驗證 LOTCA-II 與其他與認知功能相關構念評估工具之關聯性，如日常生活功能 (ADL) 、工具性日常生活功能 (IADL) 、執行功能等，未來研究可使用其他認知相關評估工具驗證 LOTCA-II 之收斂效度，以提供更完整的效度實證。在已知族群效度的部分，中風個案之 LOTCA-II 的總分及

分量表分數中，僅在「思考操作」分量表與健康成人達顯著差異（蔡佳殷，2005），且正確區辨率僅達 66.1%，代表只有三分之二的機率區辨成功。LOTCA-II 為目前臺灣最常用之認知評估工具，但是我們所使用的認知評估工具正確區辨率並不理想，這可能造成臨床人員實務上的限制，無法有效判斷個案是否有認知受損，並且給予有效的治療。另外，LOTCA-II 應用於慢性中風患者之族群，在四種分量表中有顯著的天花板效應，代表此一測驗對於慢性中風患者過於簡單，無法區辨個案認知損傷之程度。因素效度方面，混合樣本中包含中風個案的驗證性因素分析 (Su, Lin, et al., 2007) ，但因是混合樣本，不能代表單獨應用在中風族群的結果。因 LOTCA-II 有六個分量表，且完成測驗需費時 30-90 分鐘，故不適合用來做為篩檢認知能力的評估工具，後續可以用驗證性因素分析再進一步探究 LOTCA-II 的因素結構。簡言之，LOTCA-II 應用於臺灣中風個案尚有同時效度、預測效度及反應性等心理計量特性尚未驗證，且收斂效度亦缺乏驗證 LOTCA-II 與日常生活功能、執行功能等評估工具之相關性，未來研究可更有系統地深究並驗證缺少或不全的特性，以提供臺灣在地的臨床及學術應用。

二、創傷性腦傷個案

LOTCA-II 應用於創傷性腦傷個案亦不充分。目前臺灣僅有三篇文獻探究其應用於創傷性腦傷個案之心理計量特性。雖已驗證此群體之信度，但效度仍缺乏驗證同時效度及因素效度，且也無研究驗證此族群之天花板/地板效應及反應性。

信度方面，創傷性腦傷個案各分量表之內在一致性分布於中等至良好等級，再測信度良好（蔡佳殷，2005）。內在一致性不佳之原因可能為 LOTCA-II 量表本身涵蓋太多認知面向，導致題目間相關性較低。收斂效度方面，目前也僅與認知功能之量表作為效標驗證收斂效度，且結果僅有部分分量表達良好等級，收斂效度不佳之分量表可能受其他構念（如：日常生活功能、執行功能）影響，但目前尚未有研究驗證 LOTCA-II 與這些構念之收斂效度。在已知族群效度的部分，創傷性腦傷個案在所有分量表都有顯著差異（蔡佳殷，2005），代表 LOTCA-II 可區辨創傷性腦傷之個案。因素效度方面，僅有包含創傷性腦傷個案的混和樣本驗證性因素

分析研究 (Su, Lin, et al., 2007) ，但因是混合樣本，不能代表單獨應用在創傷性腦傷族群的結果。所以，LOTCA-II 應用於臺灣創傷性腦傷個案之研究仍缺乏同時效度、預測效度及反應性之驗證，且未來值得再進一步驗證各個腦部創傷程度之個案群之心理計量特性。

三、思覺失調症個案

目前僅有兩篇文獻探究 LOTCA-II 應用於思覺失調症個案之心理計量特性，但驗證結果顯示其信、效度皆不佳，需要重新檢視以確認 LOTCA-II 對思覺失調症個案適用性，即此一評估工具是否適用於思覺失調症個案及其評估效能。至今仍缺乏研究驗證此群體之同時效度、預測效度及反應性。

信度方面，同中風及創傷性腦傷個案，各分量表之內在一致性僅部分達良好等級，可能此一評估工具所涵蓋之認知功能之面相過於廣泛，缺乏一致性。LOTCA-II 中「定向感」、「空間知覺」及「動作計畫」分量表於思覺失調症個案之再測信度不佳，表示此三種分量表應用於思覺失調症個案，不同次評估所收集之資料缺乏一致性，這可能影響臨床或研究人員解讀評估數據，也可能影響其他心理計量特性之結果。

收斂效度的部分，LOTCA-II 應用於臺灣思覺失調症個案之收斂效度使用魏式智力測驗 (WAIS-III) 、執行功能 (WCST) 、認知功能 (ACLS) 及日常生活功能 (DLFS) 等做為效標之相關性達良好程度，代表 LOTCA-II 用於思覺失調症族群，與認知功能及日常生活功能有良好的收斂效度。

已知族群效度的部分， LOTCA-II 於思覺失調症族群間之就業情況之已知族群效度有限。在國內，有學者探討 LOTCA-II 是否能區辨思覺失調個案目前是否就業，其結果顯示 LOTCA-II 無法區辨思覺失調症個案目前是否正在就業。國外學者亦針對思覺失調症個案、中風個案及健康成人做群組間的比較，結果顯示思覺失調症之個案與中風個案之 LOTCA-II 總分較有顯著的差異。此結果顯示 LOTCA-II 僅能區辨不同族群之個案，無法區辨族群內個案能力程度之差異 (Josman & Katz,

2006)。目前 LOTCA-II 驗證於國內思覺失調症個案對於不同族群之已知族群效度仍有待確認，日後之研究可做更進一步之探討。

LOTCA-II 題目對於思覺失調症族群有顯著的天花板效應。此顯示若評估能力較好的思覺失調個案，可能無法區隔他是否有認知功能相關的障礙，也可能無法呈現認知功能介入之後的進步。若要進一步了解功能較佳的個案其認知功能的差異，則要考慮選用其他評估工具，才能做更精確的評估。

總括而言，LOTCA-II 於思覺失調症族群之內在一致性、再測信度及已知族群效度在臺灣雖皆有被驗證，但結果僅部份分量表達良好，針對信度不佳之分量表，評估人員應謹慎解釋評估結果。LOTCA-II 於思覺失調症個案族群中，雖與認知功能、執行功能及日常生活功能達良好之收斂效度，但因部份分量表再測信度不佳，應針對信度不佳之分量表進行修改，或使用其他適合用於思覺失調症個案之認知評估工具。目前也缺乏同時效度、預測效度及反應性之驗證，未來研究可進一步驗證上述之心理計量特性，以探討 LOTCA-II 於思覺失調症族群之臨床評估之適用性。

四、智能障礙者

國內目前僅有一篇文獻驗證 LOTCA-II 應用於智能障礙個案群體的內在一致性、收斂效度及已知族群效度，但結果亦不甚理想。目前臺灣也缺乏文獻驗證此群體之再測信度、施測者間信度、內容效度、同時效度、因素效度及反應性。

信度方面，目前臺灣之研究僅驗證智能障礙族群之內在一致性，缺乏研究驗證 LOTCA-II 應用於智能障礙族群之再測信度。缺乏再測信度之評估工具，可能無法提供穩定的評估結果，進而驗證其他心理計量特性。

收斂效度方面， 張彧等人發現 LOTCA-II 於智能障礙等級及圖形 IQ 測驗（以圖形來評估個體的智商的測驗，多用於超過九歲的個體）做為效標驗證其相關性，結果顯示良好。目前亦無其他驗證於智能障礙族群之國外研究可比較，未來應用於臺灣智能障礙個案之研究可針對認知功能、執行功能及學習功能等構念，確認 LOTCA-II 於智能障礙者之收斂效度。已知族群效度部分，LOTCA-II 只能區辨健

康成人跟智能障礙者之認知功能差異，但在區辨智能障礙族群間認知程度差異則不顯著，代表 LOTCA-II 可能不適合用來區辨不同智能障礙程度，僅能篩檢個案是否有認知功能障礙。若需確認個案之智能障礙程度，需要使用其他評估工具進行確認。

在智能障礙族群中，大多數的分量表都呈現天花板效應。即 LOTCA-II 對於功能較佳的智能障礙個案，在評估不同的認知領域可能無法有效評估到個案能力的變化。若要進一步了解個案認知功能的差異，可能需要選擇其他的評估工具。

總體而言，LOTCA-II 於智能障礙族群應先驗證再測信度，進而驗證同時效度、預測效度及反應性，以驗證 LOTCA-II 應用於智能障礙個案之適用性。

五、混合樣本（中風、創傷性腦傷及思覺失調症個案）

CFA 驗證結果發現：雙因素模型相較於單因素模型提供較好的適配度；進一步使用階層式雙因素模型 (hierarchical two-factor model) 分別應用於中風及腦外傷混和樣本及思覺失調個案的適配度，結果顯示模型適配度也都極佳 (Su, Lin, et al., 2007) 。然而雙因素模型的驗證，對於具有 6 個分量表的 LOTCA-II，不論是在廣度或內容上，顯然是不盡合理的，無法涵蓋認知的所有面向，需要更多研究進一步探討。

LOTCA-II 為 6 個分量表的認知評估工具組，若要用 CFA 驗證 LOTCA-II 屬於幾個因素的模型較為適配，至少應該納入 6 個因素的模型進行比較。蘇等學者以 CFA 驗證 LOTCA-II，直接比較單因素及雙因素模型的適配度，而未加入六因素模型進行比較，著實可惜。所得結果雖顯示雙因素模型適配度較佳，但其結果之應用及概化有限，後續研究可再驗證何種模型是最適配的。

採用混合樣本的研究，容易因為樣本間的特質不同，使得結果的解讀不一及難以概化。蘇的研究採用的樣本包括中風、創傷性腦傷及思覺失調症個案，或許是為了增加樣本數，才匯集三種不同診斷的個案在同一研究中，但因此三種障礙族群的特質均不同，造成的認知損傷也不同，所得結果恐難概化，未來的研究宜分別驗證不同特質的障礙族群，以利後續臨床及研究之詮釋及應用。

本研究之限制有二：一、因本文僅回顧已刊登之 LOTCA-II 驗證於臺灣人之心理計量特性之文獻，可能遺漏發表於學術研討會而未刊登於期刊之研究結果。二、因本文未納入非以驗證心理計量特性為研究目的之文獻，然而有些其它類型研究可能包含心理計量驗證之論文並未納入本研究中。

結論

雖然 LOTCA-II 是臺灣臨床上常用的認知評估工具，且已廣泛應用於中風個案、創傷性腦傷、思覺失調症及智能障礙等族群。然而，LOTCA-II 應用於上述族群之心理計量特性不完整（尚缺同時效度、預測效度及反應性），且缺乏應用於部分族群（例如：失智症及成癮患者）、不同病程（例如：急性中風個案）之心理計量特性驗證，嚴重限制使用者對評估分數之解釋。未來研究宜更深入完整地驗證 LOTCA-II 應用於各障礙族群之心理計量特性，並修改信、效度不佳之分量表，以提供更完整的實證參考數據，目前臨床及研究人員對於 LOTCA-II 的評估結果宜審慎解讀及應用。

參考文獻

蔡佳殷（民 104）。洛文斯坦認知評估第二版國內信效度之檢驗：中風與頭部外傷病患。**高雄醫學大學行為科學研究所學位論文**，1–127。

蔡佳殷、陳惠媚、蘇純瑩（民 106）。“洛文斯坦職能治療認知評估”第二版於輕度血管性認知缺損的中風病人之適用性：初探研究. **職能治療學會雜誌**，**25**，15-31。

蘇純瑩、陳惠媚、張雅雯、蔡佳殷、林月仙（民 106）“洛文斯坦職能治療認知評估”第二版的南台灣初步常模之建立.**職能治療學會雜誌，25**，60-74。

Aaronson, N., Alonso, J., Burnam, A., Lohr, K. N., Patrick, D. L., Perrin, E., & Stein, R. E. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research, 11*, 193-205.

Annes, G., Katz, N., & Cermak, S. A. (1996). Comparison of younger and older healthy American adults on the Loewenstein Occupational Therapy Cognitive Assessment.

Occupational Therapy International, 3, 157-173.

Avila, A., Duran, M., Peralbo, M., Torres, G., Saavedra, M., & Viana, I. M. (2015). Effectiveness of an occupational therapy home programme in Spain for people affected by stroke. *Occupational Therapy International, 22*, 1-9.

Barak, S., & Duncan, P. W. (2006). Issues in selecting outcome measures to assess functional recovery after stroke. *NeuroRx, 3*, 505-524.

Beaton, D. E., Bombardier, C., Katz, J. N., Wright, J. G., Wells, G., Boers, M., . . . Shea, B. (2001). Looking for important change/differences in studies of responsiveness. OMERACT MCID Working Group. Outcome Measures in Rheumatology. Minimal Clinically Important Difference. *Journal of Rheumatology, 28*, 400-405.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Bollen, K. A. (1990). Overall fit in covariance structure models: Two types of sample size effects. *Psychological Bulletin, 107*, 256-259.

Bowling, A. (2001). *Measuring disease: A review of disease-specific quality of life measurement scales*. Philadelphia: Open University Press.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research & evaluation, 10*, 1-9.

Hobart, J. C., Lamping, D. L., & Thompson, A. J. (1996). Evaluating neurological outcome measures: The bare essentials. *Journal of Neurology, Neurosurgery and Psychiatry, 60*, 127-130.

Holmes, W., Bix, B., & Shea, J. (1996). SF-20 score and item distributions in a human immunodeficiency virus-seropositive sample. *Medical Care, 34*, 562-569.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling-a Multidisciplinary Journal, 6*, 1-55.

Hutchinson, A. (1992). Measures of need and outcome for primary health care. *The British Journal of General Practice, 42*, 262-262.

Itzkovich, M., Elazar, B., Averbuch, S., & Katz, N. (1990). *Loewenstein Occupational*

Therapy Cognitive Assessment (LOTCA) battery (1st ed.). Pequannock, NJ: Maddak Inc.

Itzkovich, M., Elazar, B., Averbuch, S., & Katz, N. (2000). *Loewenstein Occupational Therapy Cognitive Assessment-Second Edition Manual.* Pequannock, NJ: Maddak Inc.

Jang, Y., Chern, J. s., & Lin, K. c. (2009). Validity of the Loewenstein Occupational Therapy Cognitive Assessment in people with intellectual disabilities. *American Journal of Occupational Therapy, 63*, 414-422.

Jongbloed, L. (1986). Prediction of function after stroke: A critical review. *Stroke, 17*, 765-776.

Josman, N., Abdallah, T. M., & Engel-Yeger, B. (2010). Cultural factors affecting the differential performance of Israeli and Palestinian children on the Lewenstein Occupational Therapy Cognitive Assessment. *Research in Developmental Disabilities, 31*, 656-663.

Josman, N., Abdallah, T. M., & Engel-Yeger, B. (2011). Using the LOTCA to measure cultural and sociodemographic effects on cognitive cultural and sociodemographic effects on cognitive skills in two groups of children. *American Journal of Occupational Therapy, 65*, e29-37.

Josman, N., & Katz, N. (2006). Relationships of categorization on tests and daily tasks in patients with schizophrenia, post-stroke patients and healthy controls. *Psychiatry Research, 141*, 15-28.

Kline, P. (1998). *The new psychometrics: Science, psychology, and measurement*. Abingdon, Oxon: Routledge.

Lezak, M. D. (2004). *Neuropsychological assessment*. New York: Oxford University Press.

Lund, K., Oestergaard, L. G., & Maribo, T. (2013). Danish translation and adaptation of Loewenstein Occupational Therapy Cognitive Assessment 2nd Edition (LOTCA-II). *Scandinavian Journal of Occupational Therapy, 20*, 302-305.

Lund, K., Oestergaard, L. G., & Maribo, T. (2014). Reliability and internal consistency of the Danish version of Loewenstein Occupational Therapy Cognitive Assessment 2nd Edition (LOTCA-II/D). *Scandinavian Journal of Occupational Therapy, 21*,

473-478.

McCarthy, M. L., Silberstein, C. E., Atkins, E. A., Harryman, S. E., Sponseller, P. D., & Hadley-Miller, N. A. (2002). Comparing reliability and validity of pediatric instruments for measuring health and well-being of children with spastic cerebral palsy. *Developmental Medicine and Child Neurology, 44*, 468-476.

Natar, A. K. M., Nagappan, R., Ainuddin, H. A., Masuri, G., & Thanapalan, C. K. K. (2015). Psychometric properties of the Malay version of the Loewenstein Occupational Therapy Cognitive Assessment for Geriatrics (M-LOTCA-G) among the Malaysian elderly population. *Educational Gerontology, 41*, 27-40.

Portney, L. G., & Watkins, M. P. (2009). *Foundation of clinical research: Application to practice*. Upper Saddle River, NJ: Pearson Prentice Hall.

Pyun, S. B., Yang, H., Lee, S., Yook, J., Kwon, J., & Byun, E. M. (2009). A home programme for patients with cognitive dysfunction: A pilot study. *Brain Injury, 23*, 686-692.

Rojo-Mota, G., Pedrero-Perez, E. J., Ruiz-Sanchez de Leon, J. M., Leon-Frade, I., Aldea-Poyo, P., Alonso-Rodriguez, M., . . . Morales-Alonso, S. (2017). Loewenstein Occupational Therapy Cognitive Assessment to evaluate people with addictions. *Occupational Therapy International, 2017*, 2750328.

Salter, K., Jutai, J. W., Teasell, R., Foley, N. C., Bitensky, J., & Bayley, M. (2005). Issues for selection of outcome measures in stroke rehabilitation: ICF activity. *Disability and Rehabilitation, 27*, 315-340.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*, 323-338.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

Su, C. Y., Chang, J. J., Chen, H. M., Su, C. J., Chien, T. H., & Huang, M. H. (2000). Perceptual differences between stroke patients with cerebral infarction and intracerebral hemorrhage. *Archives of Physical Medicine and Rehabilitation, 81*, 706-714.

Su, C. Y., Chen, W. L., Tsai, P. C., Tsai, C. Y., & Su, W. L. (2007). Psychometric

properties of the Loewenstein Occupational Therapy Cognitive Assessment-Second Edition in Taiwanese persons with schizophrenia. *American Journal of Occupational Therapy, 61*, 108-118.

Su, C. Y., Lin, Y. H., Chen-Sea, M. J., & Yang, M. J. (2007). A confirmatory factor analysis of the Chinese version of the Loewenstein Occupational Therapy Cognitive Assessment--Second Edition as a Taiwanese mixed clinical sample. *OTJR: Occupation, Participation and Health, 27*, 71-80.

Van Deusen, J., & Brunt, D. (1997). *Assessment in occupational therapy and physical therapy*. Philadelphia, PA: WB Saunders Company.

Wade, D. T. (1992). Measurement in neurological rehabilitation. New York: Oxford University Press.

Journal of Taiwan Occupational Therapy Association
2018, 36(2),177-202
DOI:10.6594/JTOTA.201812_36(2).0003

Review article

A Review of Psychometric Properties of the Loewenstein Occupational Therapy Cognitive Assessment–Second Edition Used in Patients With Different Disabilities in Taiwan

Pei-Chi Li [a], Pei-Chieh Tsai[b], Yi-Jing Huang [a], Shih-Chieh Lee[a], Ya-Chi Tung[a], Ching-Lin Hsieh[a,c,d], Sheau-Ling Huang[a,c,*]

Abstract

Cognitive ability plays a key role in maintaining daily function. Cognitive deficits are common in patients with brain disease. Loewenstein Occupational Therapy Cognitive Assessment–Second Edition (LOTCA-II) is the most widely used cognitive ability battery in clinical practice. However, only few studies in Taiwan have evaluated the psychometric properties of the LOTCA-II. The aim of this study was to review and appraise the psychometric properties of the LOTCA-II in different disability groups in Taiwan. After searching electronic English and Chinese databases, results showed that the LOTCA-II's total score has good reliability in stroke and traumatic brain injury (TBI) , but has poor internal consistency in each subtest. In patients with stroke and TBI, LOTCA-II was highly correlated with cognitive tests, which indicates good convergent validity, and can discriminate between patients and healthy adults. In patients with schizophrenia, LOTCA-II has good reliability and good convergent validity, and can discriminate between patients and healthy adults. In patients with intellectual disabilities, LOTCA-II has good internal consistency and good convergent validity, and can discriminate between patients and healthy adults. A few studies investigated factor validity in patients with stroke and TBI collectively, but no studies investigated factor validity in individual disability groups. Our study found that the psychometric properties of the LOTCA-II were limited in patients with stroke, TBI, schizophrenia and intellectual disabilities in Taiwan. Future studies are needed to explore the applicability and psychometric properties of the test in all disability groups. For now, clinicians should be careful when explaining the scores of the LOTCA-II.

Keywords: *Cognition, Loewenstein Occupational Therapy Cognitive Assessment-Second Edition, Psychometric Properties, Schizophrenia, Stroke*

[a] School of Occupational Therapy, College of Medicine, National Taiwan University
[b] Occupational Therapy, Rehabilitation Department, Shin Kong Wu-Ho-Su Memorial Hospital
[c] Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital
[d] Department of Occupational Therapy, College of Medical and Health Science, Asia University

*Correspondence: Sheau-Ling Huang
17 Xuzhou Rd., Zhongzheng Dist.,
Taipei City 100, Taiwan
TEL: 06-33668179
E-mail: cathy@ntu.edu.tw

Received: 22 February 2018
Accepted: 14 July 2018

Journal of Taiwan Occupational Therapy Association
2018, 36(2),203-227
DOI:10.6594/JTOTA.201812_36(2).0004

Original article

Developing the Pictograms of the Contextual Memory Test for Cultural Suitability in Taiwan

Jing-Ru Sun[a], Huang-Min Lioa[a], Ling-Wei Chang[b],
Ko-Hsin Chen[c], Chien-Hsiou Liu[c,*]

Abstract

Background: The Contextual Memory Test (CMT) measures aspects of memory and metamemory of people with cognitive impairments. It is important to improve cultural suitability of the CMT and develop pictograms for users in Taiwan. The purposes of this study were (1) to develop and refine appropriate pictograms and objects to address the issue of cultural suitability for people in Taiwan, and (2) to understand the recognition performance of the new versions of the pictograms for people with cognitive impairments.

Methods: Three tests were designed, including a pictogram comprehension test for healthy adults, a pictogram recognition test for healthy adults, and a pictogram recognition test for people with cognitive impairments. The analysis of variance (ANOVA) and Fisher's least significant difference test (LSD test) were used.

Results: For healthy adults, the results indicated that 8 pictograms (razor, towel, coffee cup, and comb in the Morning version and flowers, fork, coffee cup, and plate in the Restaurant version) should be retained; the objects represented by two pictograms (deodorant and cash register) should be changed; and another 30 pictograms should be replaced with the modified versions. For example, the deodorant pictogram was changed to a eyeglasses pictogram, and the cash register pictogram was changed to the chopsticks pictogram. The correct recognition rates for all pictograms were higher than 90%. For participants with cognitive impairments, the results indicated that the recognition rates for all of the pictograms (lowest recognition rate: 40%) were higher in the new version than in the old version (lowest recognition rate: 0%), and mild improvement was found for the items of restaurant server, sugar, and menu (from 0–20% to 40%).

Discussion: The increased recognition rates suggest that cultural suitability of the CMT has improved for users in Taiwan.

Keywords: *Contextual Memory, Cultural Suitability, Pictogram Comprehension, Pictogram Recognition*

[a] Department of Rehabilitation, Fu Jen Catholic University Hospital
[b] Department of Rehabilitation, Cardinal Tien Hospital
[c] Department of Occupational Therapy, College of Medicine, Fu Jen Catholic University

*Correspondence: Liu Chien-Hsiou
510 Zhongzheng Rd., Xinzhuang Dist, New Taipei City 24205, Taiwan
TEL: 02-29053926
E-mail: 072223@mail.fju.edu.tw

Received: 23 April 2018
Accepted: 24 August 2018

1. Introduction

The Contextual Memory Test (CMT) measures aspects of the memory and metamemory of people with cognitive impairments. It has been used with a wide variety of diagnoses, such as head trauma, cerebral vascular disorders, and dementia. The strengths of the CMT include ease of administration, ease of transport, a short administration time of only five to ten minutes, and the ability to screen for memory dysfunction. The limitation of the CMT is that it is not a diagnostic memory assessment tool. In the United States, the reliability and validity of the CMT have been studied. Moderate validity was found on recall and on metamemory for people with brain injuries, and normative data have also been collected (Gil & Josman, 2001; Toglia, 1992).

The testing materials of the CMT contain 40 line drawings of objects. All of the objects to be remembered are familiar objects from daily life in the United States; meaningful everyday items are thought to relate more to memory function than do the items used in traditional memory tests (Toglia, 1992). Because the ”meaningful items” of the CMT were originally developed and standardized in the United States, the comprehension of the pictograms of people from other cultures should be examined. Previous research has indicated that the performance on the CMT of normal subjects in Taiwan tends to be lower than that of Americans (Tsai, 2006; Tsai & Chen, 2006). Another study also indicated that the memory scores of normal Israeli subjects on the CMT were slightly lower than those of American subjects (Naomi Josman, 2000). The poor comprehension of the CMT items lead to

lower ratings of the memory performance for normal subjects in Taiwan than for American subjects, so the issue of cultural suitability should be considered (Tsai & Chen, 2006). This research attempted to develop pictures and objects easily recognizable by natives of Taiwan to improve the cultural suitability of the CMT for users in Taiwan.

Previous research has found that age, education level, and cognitive ability may influence performance in determining the meanings of pictograms (Beaufils et al., 2014; Scialfa et al., 2008; Toglia, 1992). In other words, the comprehension of the pictograms depends on several factors, such as the concept, depiction quality, prior training, and target group knowledge. As compared with young adults, older adults tend to have poorer performance in determining the meanings of pictograms (Beaufils et al., 2014). Older adults with cognitive impairment also have poorer pictogram comprehension performance than that of healthy older adults (Scialfa et al., 2008). This research attempted to consider the subjects' age, education level, and cognitive ability to develop appropriate pictures and objects to improve the cultural suitability of the CMT for users in Taiwan.

Pictogram designs are culturally linked and are important in the interpretation of the pictograms (Naomi Josman, 2000). The issue of cultural suitability should be considered for the CMT; for example, the item deodorant is rarely used in daily life in Taiwan (Tsai, 2006; Tsai & Chen, 2006), and the format of a receipt is obviously different in Taiwan. Replacing some items and pictures would improve the cultural suitability.

Pictograms are sometimes called pictures, drawings, or graphics. In this research, the term pictogram will be used as our primary term. Easily recognizable pictograms should be provided for memory evaluation because the encoding and retrieval conditions in recall performance are more consistent with the theoretical perspectives of memory (Tulving, 1987). For presenting correct information, the American National Standards Institute (ANSI) has proposed comprehension testing methods and criteria. An item is considered "acceptable" if it is comprehended by 85% of a sample of 50 participants, with no more than 5% critical confusion errors (American National Standards Institute, 2011). Based on the "acceptable" consideration of no more than 5% critical confusion errors, our research defined that the pictograms should be changed if the recognition rate was lower than 95%. After comprehension testing, the end result was improved recognition and comprehension of pictogram-related information, leading to significant cognition processes.

The purpose of this research was to develop and refine appropriate pictures and objects which address the issue of cultural suitability for a pictogram comprehension test and recognition test for people living in Taiwan, based on the 40 line drawings of objects (pictograms) that are contained in the testing materials of the CMT. The second purpose of this research was to understand the recognition performance of the new versions of the pictograms for people with cognitive impairment in Taiwan.

2. Methods

Three experiments were designed to achieve our research purposes. First, a pictogram comprehension test was designed for healthy adults who are natives of Taiwan. Second, a recognition test was designed for healthy adults who are natives of Taiwan (the participants were different from those who participated in the comprehension test). However, in an examination of the possible cultural suitability of a pictogram comprehension test, the effects of cognitive impairment on pictogram identification must also be considered. Therefore, the third test, the pictogram recognition test, was designed for people with cognitive impairment. This research was approved by the local ethics committee of the Institutional Review Board.

2.1 Experiment 1: Pictogram comprehension test

2.1.1 Participants

All of the participants in this experiment were recruited via the internet, such as through Facebook, Line, and e-mail. The recruitment criteria were as follows: native of Taiwan, absence of cognitive impairment, normal or corrected-to-normal vision, and willingness to participate. People who were born outside Taiwan and had studied or lived outside Taiwan for over one year were excluded. The recruitment period was two months. Participants who did not complete the test were excluded from subsequent analyses.

2.1.2 Materials

Five sets of pictograms comprising the 40 line drawings of objects contained in the CMT were used. Equal numbers of studied objects were paired in two versions of the CMT: 20 line drawings of objects for the morning version, and 20 line drawings of objects for the restaurant version. The pictogram comprehension test was conducted in the form of a ranking questionnaire in this study. The ranking questionnaire was designed in Google Sheets. Regarding the pictograms in the ranking questionnaire, four designers of Taiwanese nationality were invited to draw 40 line drawings for the objects of the CMT. Therefore, five digital images of the pictures were shown for each of the 40 testing objects. The first was the original image of the CMT, and the following four images were drawn by designers of Taiwanese nationality. For example, Fig. 1 shows five line drawings of the object "bed". Forty line drawings for the objects of the CMT were designed in the ranking questionnaire and sequentially presented. Previous research has suggested that because deodorant is rarely used in daily life in Taiwan, the cultural suitability of the CMT (Tsai & Chen, 2006) needed to be considered; therefore, deodorant xwas replaced with eyeglasses after discussion and agreement by our research team.

2.1.3 Experimental procedure

The pictogram comprehension test was conducted to determine the most appropriate of the five pictograms of each object for participants living in Taiwan. First, the participants were asked to enter their demographic

information, including age, gender, and educational level, into a Google Sheets table before testing began. Then the five pictograms of each object were presented sequentially using a personal mobile phone. Participants were blind to the sequences of images. While viewing the five pictograms of each object, participants were asked to choose the most recognizable image and fill in a score of "5", and then to continue until all the images were scored from "5" to "1" to rank them from excellent to poor. In the pictogram recognition test, the 40 objects were presented sequentially. The participants could rest at any time if they felt tired. All participants completed the pictogram comprehension test on their personal mobile phones in a quiet setting.

2.1.4 Statistical analysis

All of the data were examined by normality test. Chi-square was used to evaluate if the gender, age, and education level were related in choosing the most recognizable pictograms. Analysis of variance (ANOVA) was used to determine the most recognizable pictogram of each object for participants from Taiwan, and Fisher's least significant difference test (LSD test) was used for post hoc analysis if any significant differences in pictogram recognition scores were found within each set of five pictograms. The significance level was set at $\alpha < 0.05$.

In addition, the results were summarized for the retention rate, replacement rate, and change rate of the pictograms. The retention rate was the number of retained pictograms divided by the total number of pictograms;

the replacement rate was the number of replaced pictograms divided by the total number of pictograms; and the change rate was the number of changed pictograms divided by the total number of pictograms.

Figure 1

Pictogram recognition test for the "Bed" item

2.2 Experiment 2: Pictogram recognition test for healthy adults

2.2.1 Participants

Fifty-four healthy adults (22 males and 32 females; mean age 30.40 (± 16.99) years) who were natives of Taiwan were recruited to participate in this

experiment. All of the participants were recruited via the internet, such as through Facebook, Line, and e-mail. The recruitment criteria were as follows: native of Taiwan, absence of cognitive impairment, normal or corrected-to-normal vision, and willingness to participate. People who were born outside Taiwan and had studied or lived outside Taiwan for over one year were excluded. The visual acuity of all participants was better than 0.8 to ensure that they could clearly see the pictograms and recognize them correctly.

2.2.2 Materials

The testing materials for the pictogram recognition test comprised the 40 pictograms of the modified CMT (called CMT-Taiwan, CMT-T) developed in Experiment 1. The 40 pictograms were printed on 2 sheets of A4 paper, each containing 20 pictograms. One was the Morning version, and the other was the Restaurant version (see Figure 2 and Figure 3).

2.2.3 Procedures and data analysis

First, the demographic data were collected and the corrected Snellen visual acuity at a three-meter viewing distance was measured while the participants wore corrective lenses. Then all the participants participated in the recognition test. The distribution of the two testing materials to the participants was randomized. The participant sat at a table, and an A4 sheet of paper with 20 pictograms was placed on the table in front of the participant. Then the participant was asked to name each pictogram from left to right, and from the top row to the bottom row. Each participant was encouraged to

name the pictogram or respond with his or her interpretation of its meaning but not given any verbal or nonverbal cues. The examiner recorded each response. Finally, the correct recognition score was calculated after all participants completed testing. Several responses were recorded for the same pictogram; for example, the "tap" picture elicited responses of faucet, tap, and water. Therefore, for reliability, three examiners determined which one was the correct name or meaning in a blind condition. The open-ended responses were then scored as dichotomous data (either correct or incorrect). Correct was defined as two or three examiners determining the name to be correct. All examiners were occupational therapists, were familiar with the testing procedures of the CMT, and understood the research purposes.

2.3 Experiment 3: Pictogram Recognition test for people with cognitive impairment

2.3.1 Participants

Five participants (3 males and 2 females, mean age 25.80 ± 5.53 (range from 20 to 34 years)) were recruited from local educational and nursing institutions. Three had been diagnosed with intellectual disabilities, one with Down Syndrome, and one with traumatic brain injury. They were thoroughly screened by interview to exclude individuals with expressive and recognizable communication limitations before completing the pictogram recognition test. The Standardized Mini Mental State Exam (SMMSE) was used to evaluate their cognitive functioning. The mean score of the SMMSE was 13.40 ± 3.77, ranging from 10 to 20. Before the start of the testing session, the corrected

Snellen visual acuity at a three-meter viewing distance was measured while the participants wore corrective lenses. The visual acuity of all the participants was better than 0.8.

2.3.2 Materials

Two sets of testing materials for the open-ended test were prepared. One set was the 40 pictograms of the CMT, and the other set was the 40 pictograms of the modified CMT (CMT-T) developed after Experiment 2. The 80 pictograms were printed on 4 sheets of A4 paper, each containing 20 pictograms.

2.3.3 Procedures and data analysis

First, the demographic data were collected, and then the cognitive level and corrected Snellen visual acuity were measured. Then all participants sat at a table to begin the pictogram recognition test. Each participant was given one of the four testing material sheets in randomized order. Each sheet was collected after completion and the next one was given until the four sheets had been completed. Each sheet of the testing materials containing 20 pictograms was placed on the table in front of the participant. The subsequent instructions, testing procedure, and scoring process were the same as in Experiment 2.

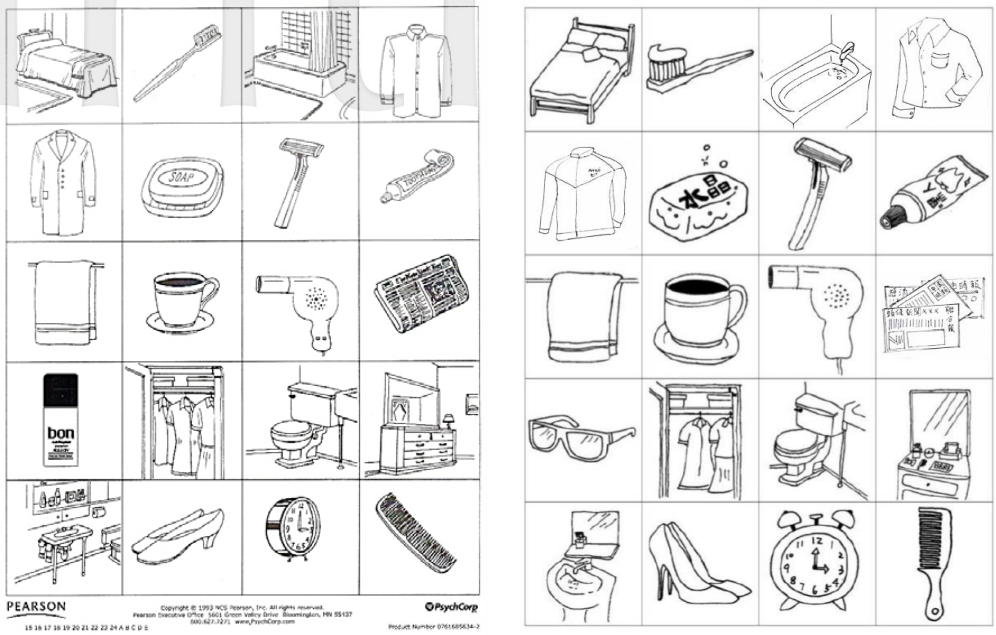

Figure 2 Testing materials for the Morning version: left, original; right, refined.

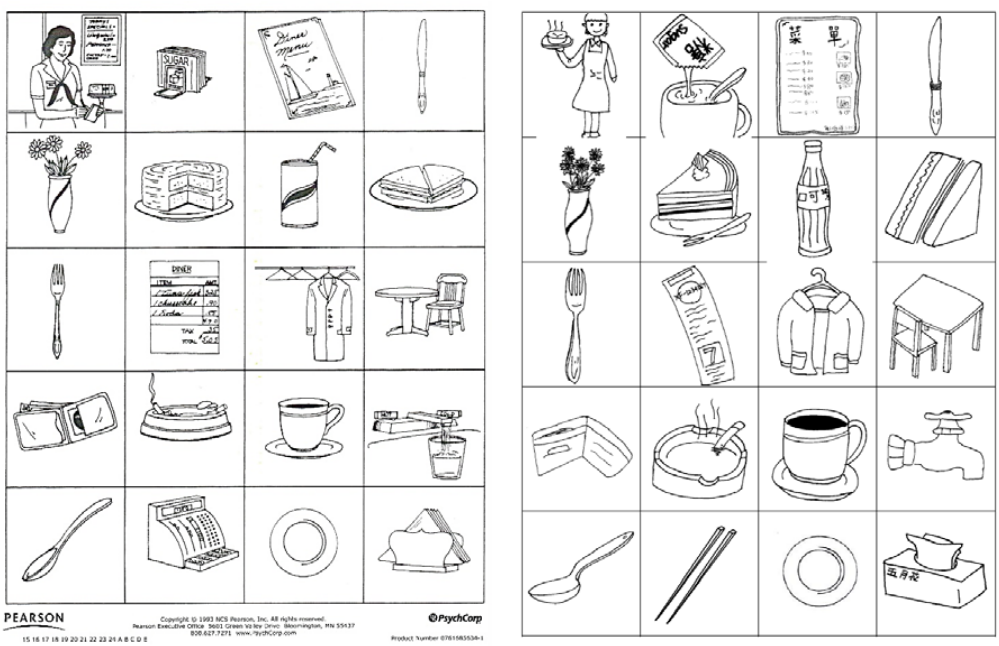

Figure 3

Testing materials for the Restaurant version: left, original; right, refined.

3. Results

3.1 Experiment 1. Pictogram comprehension test

For this experiment, 128 healthy adults native to Taiwan were recruited. All participants completed the test. Demographic data included information on gender, age, and years of education. In the sample, 60.75% (n=78) of the participants were female and 39.25% (n=50) were male. The percentages of participants that chose the most recognizable image did not differ by gender, X2 (1, n=128) = 4.762, p > 0.313. Divided by age, 26.17% (n=33) of the participants were 20–29 years old, 12.15% (n=16) were 30–39 years old, 27.10% (n=35) were 40–49 years old, and 34.58% (n=44) were 50–59 years old. The age groups were not significantly different, X2 (3, N = 128) = 6.026, p = 0.915. Regarding level of education, 17.76% (n=23) of the participants had more than 17 years of education, 72.90% (n=93) had 12–16 years of education, and 9.35% (n=12) had less than 12 years of education. The relation between levels of educational was significant, X2 (2, N = 128) = 28.815, p < 0.001. Participants with > 17 years of education rated image 1 lower than did those with 12–16 years of education or < 12 years of education.

Table 1 shows the mean pictogram comprehension score and 95% confidence interval for each object in the Morning version. The results of ANOVA for each object are shown in Table 3. All twenty objects were found to have significant differences within each set of five pictograms (p < 0.05), so the LSD post hoc test was used to compare the pictogram comprehension scores within each set of five pictograms. The results of the LSD post hoc test

indicated the most recognizable images ($p < 0.05$), which are shown in Table 3. The results indicated that four pictograms, those of a razor, a towel, a cup of coffee, and a comb, should be retained, and that the object in one pictogram, deodorant, should be changed. That object was changed to eyeglasses. Therefore, in the Morning version, the retention rate was 20% (4/20), the replacement rate was 75% (15/20), and the change rate was 5% (1/20).

Table 1

Mean pictogram comprehension score and 95% confidence interval for each object in the morning version

Pictogram / Object	1	2	3	4	5
	Mean(95%CI)	Mean(95%CI)	Mean(95%CI)	Mean(95%CI)	Mean(95%CI)
Bed	3.299(3.015-3.582)	2.299(2.102-2.496)	3.271(3.027-3.515)	**3.953(3.752-4.154)**	2.178(1.916-2.439)
Toothbrush	3.710(3.478-3.943)	3.514(3.306-3.722)	1.505(1.312-1.697)	**3.944(3.749-4.139)**	2.327(2.134-2.520)
Bathtub	3.710(3.493-3.927)	**3.822(3.614-4.031)**	2.495(2.278-2.713)	3.290(3.037-3.542)	1.682(1.463-1.902)
Shirt	2.710(2.440-2.980)	2.822(2.587-3.058)	2.766(2.547-2.986)	**3.925(3.702-4.148)**	2.776(2.452-3.100)
Overcoat	3.336(3.058-3.615)	2.879(2.641-3.116)	2.093(1.847-2.340)	**3.813(3.569-4.057)**	2.879(2.642-3.115)
Soap	3.654(3.058-3.615)	2.720(2.641-3.116)	2.467(1.847-2.340)	**4.336(3.569-4.057)**	1.822(2.642-3.115)
Razor	**4.346(4.177-4.515)**	3.000(2.799-3.201)	2.617(2.325-2.908)	3.084(2.890-3.278)	1.953(1.696-2.210)
Toothpaste	3.589(3.353-3.824)	2.813(2.630-2.996)	1.916(1.702-2.130)	**4.318(4.100-4.536)**	2.364(2.137-2.562)
Towel	**4.084(3.840-4.328)**	2.393(2.154-2.631)	2.804(2.509-3.099)	3.047(2.812-3.281)	2.673(2.450-2.896)
Coffee cup	**3.944(3.673-4.215)**	3.224(3.034-3.415)	2.103(1.883-2.323)	3.299(3.056-3.542)	2.430(2.159-2.701)
Hair dryer	**3.374(3.087-3.660)**	2.981(2.752-3.211)	2.710(2.434-2.986)	2.925(2.664-3.187)	3.009(2.722-3.297)
Newspaper	2.421(2.160-2.681)	3.121(2.894-3.349)	2.907(2.670-3.143)	**3.495(3.204-3.787)**	2.056(1.809-2.303)
Eyeglasses	**4.028(3.805-4.251)**	3.075(2.857-3.293)	2.308(2.050-2.567)	3.439(3.191-3.688)	2.150(1.917-2.382)
Closet	2.888(2.575-3.201)	**3.850(3.657-4.044)**	2.925(2.692-3.159)	2.570(2.353-2.787)	2.766(2.454-3.079)
Toilet	3.748(3.482-4.014)	**4.065(3.870-4.261)**	1.776(1.597-1.955)	3.140(2.950-3.331)	2.271(2.038-2.504)
Dresser	2.804(2.516-3.091)	**3.626(3.430-3.822)**	3.589(3.355-3.823)	2.458(2.234-2.682)	2.523(2.213-2.834)
Bathroom	2.439(2.148-2.730)	**4.290(4.115-4.464)**	2.477(2.264-2.689)	2.449(2.243-2.655)	3.346(3.084-3.608)
Shoes	3.047(2.765-3.328)	2.953(2.786-3.121	2.234(1.962-2.505)	**3.514(3.250-3.778)**	3.252(2.959-3.546)
Alarm clock	3.112(2.778-3.447)	3.280(3.057-3.503	2.121(1.910-2.333)	**3.785(3.562-4.008)**	2.701(2.464-2.938)
Comb	3.374(3.114-3.634)	3.421(3.208-3.633	1.991(1.760-2.221)	**3.692(3.450-3.933)**	2.523(2.257-2.790)

NOTE. Underline and bold indicate the highest value of the five scores.

Table 2 shows the mean pictogram comprehension score and 95% confidence interval for each object in the Restaurant version. The ANOVA results for each object are shown in Table 3. For this version as well, all twenty objects were found to have significant differences within each set of five pictograms ($p < 0.05$), so the LSD post hoc test was used to compare the pictogram comprehension scores within each set of five pictograms. The results of the LSD post hoc test indicated the most recognizable images ($p < 0.05$), which are also shown in Table 3. However, no significant difference was found for the image of a plate ($p=0.07$), so the original image was retained. The results indicated that four images, the images of flowers, a fork, a coffee cup, and a plate, should be retained. Therefore, in the Restaurant version, the retention rate was 25% (5/20) and the replacement rate was 75% (15/20).

In total, 9 pictograms were retained; the mean retention rate was 22.5% (9/40). Thirty pictograms were replaced; the mean replacement rate was 75% (30/40). Two pictograms were changed; the mean change rate was 2.5% (1/40).

3.2 Experiment 2: Pictogram recognition test for healthy adults

For the Morning version, the correct recognition score for all of the objects was 100% (54/54). For the Restaurant version, the correct recognition scores for wallet, plate, and cash register were 96.30% (52/54), 96.30% (52/54), and 90.74% (49/54), respectively. For the wallet pictogram, incorrect responses included cabinet or book; for plate, such responses included hula hoop or doughnut; for cash register, they were typewriter or calculator. Only one

pictogram, that of a cash register, should be rejected because the recognition rate was lower than 95%. Finally, the cash register was changed to chopsticks after discussion and agreement by all of our research team members.

Table 2

Mean pictogram recognition score and 95% confidence interval for each object in the Restaurant version

Pictogram / Object	1	2	3	4	5
	Mean(95%CI)	Mean(95%CI)	Mean(95%CI)	Mean(95%CI)	Mean(95%CI)
Restaurant server	3.271(2.964-3.578)	**3.869(3.677-4.061)**	2.542(2.311-2.773)	2.664(2.437-2.890)	2.654(2.357-2.951)
Sugar	2.963(2.665-3.260)	**4.056(3.815-4.297)**	2.551(2.344-2.759)	2.477(2.269-2.685)	2.953(2.673-3.234)
Menu	2.355(2.106-2.604)	2.822(2.595-3.050)	3.196(3.007-3.385)	**4.542(4.371-4.713)**	2.084(1.850-2.318)
Knife	2.626(2.386-2.867)	**3.879(3.660-4.097)**	2.131(1.895-2.366)	3.570(3.285-3.855)	2.794(2.564-3.025)
Flowers	**3.813(3.560-4.066)**	3.103(2.853-3.352)	2.383(2.119-2.647)	2.318(2.121-2.514)	3.383(3.110-3.656)
Cake	3.075(2.776-3.373)	**3.449(3.222-3.675)**	2.953(2.716-3.191)	3.364(3.130-3.599)	2.159(1.876-2.441)
Soda	2.850(2.617-3.084)	3.364(3.131-3.598)	2.271(2.058-2.484)	**4.037(3.768-4.307)**	2.477(2.216-2.738)
Sandwich	1.879(1.641-2.116)	**3.841(3.630-4.052)**	3.318(3.076-3.560)	3.654(3.423-3.885)	2.308(2.091-2.526)
Fork	**3.850(3.570-4.131)**	3.449(3.231-3.666)	1.636(1.415-1.856)	3.206(3.009-3.402)	2.860(2.633-3.087)
Receipt	2.187(1.898-2.475)	2.991(2.756-3.225)	2.589(2.364-2.814)	3.355(3.163-3.547)	**3.879(3.594-4.163)**
Coat	3.215(2.927-3.503)	3.290(3.036-3.545)	2.570(2.324-2.816)	2.533(2.274-2.791)	**3.393(3.127-3.658)**
Table and chair	3.439(3.219-3.660)	**4.421(4.240-4.601)**	1.776(1.597-1.955)	2.664(2.443-2884)	2.701(2.448-2.954)
Wallet	3.196(2.935-3.457)	3.308(3.062-3.555)	**3.318(3.064-3.571)**	3.047(2.774-3.320)	2.131(1.837-2.389)
Ashtray	2.738(2.491-3.457)	**3.794(3.578-4.011)**	1.692(1.453-1.931)	3.467(3.268-3.667)	3.308(3.054-3.563)
Coffee cup	**3.944(3.673-4.215)**	3.224(3.034-3.415)	2.103(1.883-2.323)	3.299(3.056-3.542)	2.430(2.159-2.701)
Tap	2.150(1.851-2.448)	3.449(3.223-3.674)	2.439(2.195-2.684)	3.449(3.236-3.661)	**3.514(3.264-3.764)**
Spoon	3.280(3.003-3.557)	2.869(2.660-3.078)	1.822(1.576-2.069)	**3.626(3.428-3.824)**	3.402(3.128-3.675)
Cash register	2.430(2.179-2.681)	**3.916(2.677-4.154)**	3.551(3.317-3.786)	2.682(2.476-2.889)	2.421(2.133-2.708)
Plate	3.103(2.783-3.423)	2.794(2.542-3.047)	**3.327(3.124-3.531)**	2.953(2.726-3.180)	2.822(2.499-3.146)
Napkins	1.364(1.203-1.525)	3.523(3.309-3.738)	3.168(2.950-3.386)	**3.804(3.574-4.034)**	3.140(2.888-3.392)

NOTE. Underline and bold indicate the highest value of the five scores.

Table 3

The results of ANOVA for 40 pictograms of the CMT

Morning Version					Restaurantversion				
Object	*F*	*p*	partial eta	most recognizable image	Object	*F*	*p*	partial eta	most recognizable image
Bed	30.586	<0.001	0.289	4	Restaurant server	15.440	<.001	0.146	2
Toothbrush	81.561	<0.001	0.769	4	Sugar	20.101	<.001	0.19	2
Bathtub	51.177	<0.001	0.483	2	Menu	62.326	<.001	0.588	4
Shirt	12.786	<0.001	0.121	4	Knife	27.070	<.001	0.255	2
Overcoat	20.577	<0.001	0.194	4	Flowers	21.166	<.001	0.2	1
Soap	69.588	<0.001	0.656	4	Cake	12.427	<.001	0.117	2
Razor	46.753	<0.001	0.441	1	Soda	27.110	<.001	0.256	4
Toothpaste	62.181	<0.001	0.587	4	Sandwich	44.836	<.001	0.423	2
Towel	21.588	<0.001	0.204	1	Fork	42.216	<.001	0.398	1
Coffee cup	29.216	<0.001	0.276	1	Receipt	22.146	<.001	0.209	5
Hair dryer	2.492	0.043	0.024	1	Coat	7.824	<.001	0.074	5
Newspaper	16.355	<0.001	0.154	4	Table and chair	68.127	<.001	0.643	2
Eyeglasses	34.527	<0.001	0.326	1	Wallet	11.676	<.001	0.11	3
Closet	11.528	<0.001	0.109	2	Ashtray	39.704	<.001	0.375	2
Toilet	63.440	<0.001	0.598	2	Coffee cup	29.216	<.001	0.276	1
Dresser	15.815	<0.001	0.149	2	Tap	21.779	<.001	0.205	5
Bathroom	38.717	<0.001	0.151	2	Spoon	27.099	<.001	0.256	4
Shoes	10.734	<0.001	0.101	4	Cash register	24.939	<.001	0.235	2
Alarm clock	19.722	<0.001	0.186	4	Plate	2.092	0.081	0.02	3
Comb	27.181	<0.001	0.256	4	Napkins	60.739	<.001	0.573	4

3.3 Experiment 3: Pictogram recognition test for people with cognitive impairment

Table 4 shows the correct recognition scores for each pictogram. On the CMT-T, 24 items, such as toothbrush, tube, and shirt, had correct recognition scores of 100%, and 13 items, such as soap, razor, and flowers, had correct recognition scores of 60%–80%. Three items, restaurant server, sugar, and menu in the Restaurant version, had correct recognition scores of 40%. Even though these three scores were low, they were greatly improved from 0% on the CMT to 40% on the CMT-T.

4. Discussion

In this study, three experiments were conducted. In the first experiment, it was expected that the comparisons of five pictures drawn by different designers would produce pictograms easily recognizable to people who live in Taiwan. For this experiment, 128 healthy adults were recruited, and the results indicated that 8 pictograms should be retained, 30 pictograms should be replaced, and one pictogram should be changed. The deodorant pictogram was changed to one of eyeglasses. In the second experiment, it was expected that the pictograms would be correctly recognized in the pictogram recognition test. The results indicated that the pictogram representing a cash register should be changed, so it was changed to one of chopsticks. The third experiment was conducted to measure the recognition performance of participants with cognitive impairment. The results indicated improved recognition performance for all of the pictograms,

but only mild improvement for restaurant server, sugar, and menu in the Restaurant version.

Table 4

Mean pictogram recognition score and 95% confidence interval for each object in the Restaurant version

Morning version	Correct recognition score (valid numbers / total responses)			Restaurant version	Correct recognition score (valid numbers / total responses)		
CMT		CMT-T		CMT		CMT-T	
Bed	100 (5/5)	Bed	100 (5/5)	Restaurant server	20 (5/5)	Restaurant server	40 (5/5)
Toothbrush	100 (5/5)	Toothbrush	100 (5/5)	Sugar	0 (0/5)	Sugar	40 (2/5)
Tub	100 (5/5)	Tub	100 (5/5)	Menu	0 (0/5)	Menu	40 (2/5)
Shirt	100 (5/5)	Shirt	100 (5/5)	Knife	80 (4/5)	Knife	80 (4/5)
Overcoat	100 (5/5)	Overcoat	100 (5/5)	Flowers	80 (4/5)	Flowers	80 (4/5)
Soap	40 (2/5)	Soap	80 (4/5)	Cake	80 (5/5)	Cake	100 (5/5)
Razor	60 (3/5)	Razor	60 (5/5)	Soda	100 (5/5)	Soda	100 (5/5)
Toothpaste	100 (5/5)	Toothpaste	100 (5/5)	Sandwich	60 (3/5)	Sandwich	80 (4/5)
Towel	100 (5/5)	Towel	100 (5/5)	Fork	100 (5/5)	Fork	80 (4/5)
Coffee cup	100 (5/5)	Coffee cup	100 (5/5)	Receipt	20 (1/5)	Receipt	60 (3/5)
Hair dryer	100 (5/5)	Hair dryer	100 (5/5)	Coat	100 (5/5)	Coat	100 (5/5)
Newspaper	100 (5/5)	Newspaper	100 (5/5)	Table and chair	80 (5/5)	Table and chair	100 (5/5)
Deodorant	0 (0/5)	Eyeglasses	100 (5/5)	Wallet	60 (3/5)	Wallet	60 (3/5)
Closet	80 (4/5)	Closet	80 (4/5)	Ashtray	100 (5/5)	Ashtray	100 (5/5)
Toilet	100 (5/5)	Toilet	100 (5/5)	Coffee cup	100 (5/5)	Coffee cup	100 (5/5)
Dresser	0 (0/5)	Dresser	60 (3/5)	Tap	60 (3/5)	Tap	80 (4/5)
Bathroom	40 (2/5)	Bathroom	100 (5/5)	Spoon	100 (5/5)	Spoon	100 (5/5)
Shoes	100 (5/5)	Shoes	100 (5/5)	Cash register	0 (0/5)	Chopsticks	100 (5/5)
Alarm clock	100 (5/5)	Alarm clock	100 (5/5)	Plate	80 (4/5)	Plate	80 (4/5)
Comb	100 (5/5)	Comb	100 (5/5)	Napkins	20 (1/5)	Napkins	80 (4/5)

NOTE. Grey cells indicated the changed objects.

Previous researchers found that the scores of the CMT for healthy adults in Taiwan were lower than those of their counterparts in the United States (Tsai & Chen, 2006). Based on the results of experiment 1, our research found that the retention rate was only 20%. In Experiment 3, which measured the performance of cognitively impaired participants, the recognition rate improved from 0% to 100%. The unfamiliar objects on the original CMT were difficult to encode, register, and retrieve, and cultural unsuitability may have limited the memory performance or decreased the effectiveness of contextual cues (Tsai & Chen, 2006). The low retention rate indicated that the subjects were unfamiliar with the objects or had difficulty processing the pictograms, and it also indicated the cultural unsuitability of the set of pictograms. The high recognition rate indicated that the subjects could comprehend the objects. Our findings indicate that the issue of cultural suitability has been considered, and the pictograms can be correctly recognized and applied to memory testing in Taiwanese subjects.

The percentages of participants that chose the most recognizable images did differ by years of education, but not by gender and age. Although the mean age of subjects was higher in experiment 1 than in experiments 2 and 3, we suggest that performance on experiment 1 may not have been affected by that difference because age did not affect the selection of the most recognizable image. For the years of education, participants with > 17 years of education rated image 1 lower than did those with 12–16 years of education or < 12 years of education. Evaluations of memory performance with the CMT should consider the effects of years of education in the future.

In Experiment 3, the results indicated improved recognition performance for all of the CMT-T pictograms for the participants with cognitive impairment. Previous researchers have also reported that the pictogram comprehension ability of adults with cognitive impairment is poorer than that of healthy adults (Scialfa et al., 2008). Individuals with limited cognitive ability have been reported to have low comprehension rates and more difficulty with pictogram materials (Hoonhout, 2000; Peckham, 2014). As mentioned above, because the people with cognitive impairment could recognize and comprehend the new versions of the 40 line drawings for the CMT-T, they could encode, register, and retrieve the memory materials in the CMT-T, so memory performance could be tested with the CMT-T.

The results of the third experiment indicated that recognition performance improved for all of the pictograms, but that the improvement for restaurant server, sugar, and menu (40%) was less than that of the others (60–100%). It is possible that, because the participants had lived in an educational institution for the cognitively impaired for extended periods, they had little experience of eating out. Because they rarely encountered restaurant servers, sugar, or menus, they were unable recognize the objects. The results of the second experiment indicated that the correct recognition scores for wallet and plate were both 96.30% (52/54). For the wallet pictogram, incorrect responses included cabinet or book; for plate, such responses included hula hoop or doughnut. These two pictograms should be retained because the recognition rate was higher than 95%. Three suggestions are proposed to overcome these limitations: (1) replace the pictograms with those of other items, such as a rice

spoon, which often appear in dining rooms, (2) develop a new version, such as a Shower version, to replace the Restaurant version for people with cognitive impairment, and (3) add the object names given by participants with impaired cognition to the list of acceptable alternatives on the testing manual.

Previous research on the CMT for healthy adults living in Taiwan indicated that sixteen objects (e.g., deodorant, receipt, and tap) were significantly difficult to identify (Tsai & Chen, 2006). Tsai et al.'s research suggested that the sixteen pictograms should be replaced or changed, and our study developed in depth appropriate pictograms to address issue of cultural suitability for users in Taiwan. The CMT-T can be applied in clinical memory assessment for people who live in Taiwan, and it can be used to determine memory response before and after cognitive intervention.

Our research had several research limitations. First, all of the participants were recruited via the internet; this recruitment method may have introduced bias into experiment 1. Second, only five subjects were recruited in the third experiment. Third, three items (restaurant server, sugar, and menu) should be considered for replacement in the future. Finally, the experimental design did not consider whether participants could understand the Chinese characters in the pictograms, such as the words for sugar and menu.

5. Conclusion

In this research, 80% of the pictograms of the original CMT were replaced or changed to improve the cultural suitability of the pictograms for Taiwanese

people. This research developed the pictogram materials for the CMT-T. Since picture recognition can affect image memory performance, information encoding, retrieval, and memory will be enhanced by easily recognizable pictures. It is expected that the clinical implications of the CMT-T will be further examined in the future.

References

Beaufils, E., Hommet, C., Brault, F., Marqué, A., Eudo, C., Vierron, E., . . . Mondon, K. (2014). The effect of age and educational level on the cognitive processes used to comprehend the meaning of pictograms. A*ging clinical and experimental research*, *26*, 61-65.

Gil, N., & Josman, N. (2001). Memory and metamemory performance in Alzheimer's disease and healthy elderly: The Contextual Memory Test (CMT). *Aging-Clinical and Experimental Research, 13*, 309-315.

Hoonhout, H. C. (2000). Evaluating graphical messages for mentally retarded adults; do pictures get the intended information across? Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

American National Standards Institute. (2011). Standards Z535.1-5. Retrieved Sep. 30, 2017, from https://www.ansi.org/

Naomi Josman, T. B., Tal Jarus. (2000). Cross-cultural assessment of the Contextual Memory Test. O*ccupational Therapy International*, *7*, 246-258.

Peckham, G. (2014). A New and Powerful Risk Reduction Technique: Creating OSHA/ANSI Safety Sign Systems as an Outcome of Risk Assessment. Paper

presented at the ASSE Professional Development Conference and Exposition.

Scialfa, C., Spadafora, P., Klein, M., Lesnik, A., Dial, L., & Heinrich, A. (2008). Iconic sign comprehension in older adults: The role of cognitive impairment and text enhancement. *Canadian Journal on Aging/La Revue canadienne du vieillissement, 27*, 253-265.

Toglia, J. P. (1992). Cognitive rehabilitation: Models for intervention in occupational therapy. Boston, MA: Andover Medical.

Tsai, P. L. (2006). Intoduction of Memory Function and Contextual Memory Test: A Review of Mechanism and Empirical Evidence. *Journal of Taiwan Occupational Therapy Research and Practic, 2*, 65-73.

Tsai, P. L., & Chen, Y. T. (2006). Contextual Memory Test Performance among Normal Adults and Brain-injured Patients in Taiwan: A Pilot Study. *Journal of Taiwan Occupational Therapy Research and Practic, 2*, 90-98. doi: doi:10.6534/jtotrp.2006.2(2).90

Tulving, E. (1987). Memory experiments: a strategy for research. Neurobehavioral recovery from head injury., JGHME Harvey S. Levin, ed.(Oxford University Press, New York, NY, US), 341-351.

職能治療學會雜誌
民國 107 年 36 卷第 2 期頁 203-227
DOI:10.6594/JTOTA.201812_36(2).0004

原著

以台灣文化適應性為基礎發展情境記憶測驗圖像

孫淨如 [1] 廖皇閔 [1] 張伶瑋 [2] 陳可欣 [3] 劉倩秀 [3,*]

摘要

研究背景：情境記憶測驗 (Contextual Memory Test, CMT) 是職能治療經常用以量測認知障礙者記憶與後設記憶能力之評估工具。本研究欲改善情境記憶測驗中圖像之文化適應性，獲得符合台灣文化的測試圖像，進以發展適合台灣人使用的版本。研究目的為 (1) 發展或修改 CMT 圖像，以符合台灣之文化適應性； (2) 瞭解認知障礙者對於新版本 CMT 圖像的辨識表現。

方法：本研究包含三個試驗，試驗一圖像理解測驗和試驗二圖像辨識測驗以正常成人為對象，試驗三之圖像辨識測驗則以認知障礙者為對象。並以變異數分析法和費雪爾最小顯著差異法進行資料分析。

結果：針對正常人之研究所獲得之結果，早晨版的刮鬍刀、毛巾、咖啡杯、梳子和餐廳版本的花瓶、叉子、咖啡杯和盤子等 8 個圖像獲得保留，早晨版本中的體香膏物件改成眼鏡，餐廳版本中的收銀機物件改成筷子，其餘 30 個圖像被新繪製的圖像取代，圖片辨識率皆高於 90%。認知障礙者的表現，新版的圖像辨識能力（最低辨識率 40%）皆高於舊版（最低辨識率 0％），辨識率微幅改善者為餐廳版本中的服務生、糖包、和菜單三項物件（皆由 0–20%進步到 40%）。

討論：基於台灣受試者對於圖片辨識率之提升，本研究結果所獲得的 CMT 新版圖像可符合台灣之文化適應性。

關鍵字：情境記憶，文化適應性，圖像理解，圖像辨識

輔仁大學附設醫院復健科 [1]
耕莘醫院復健科 [2]
輔仁大學醫學院職能治療學系 [3]

*通訊作者：劉倩秀
24205 新北市新莊區中正路 510 號
電話：02-29053926 分機 7517
電子信箱：072223@mail.fju.edu.tw

受文日期：民國 107 年 04 月 23 日
接受刊載：民國 107 年 08 月 24 日

Journal of Taiwan Occupational Therapy Association
2018, 36(2),228-245
DOI: 10.6594/JTOTA.201812_36(2).0005

Original article

Fabrication of Transparent Face Orthosis by Using 3D Printing Technology: A Feasibility Study

Wen-Lung Liang[a], Wan-Meng Lee[a], Shu-Hung Huang[b,c], Jyh-Jong Chang[d], Yu-Sheng Yang[d,e,*]

Abstract

3D printing is a process of creating a three-dimensional object from a digital file. It allows for printing customized and complicated models with relatively low cost. This study investigated the feasibility of fabrication of transparent face orthosis (TFO) by using a non-contact scanning with 3D printing. Twenty able-bodied participants participated in this study. After fabrication of TFOs by 3D printed mold and conventional plaster mold, participants were assigned to wear these two TFOs. The sequential order of the wearing model was randomly assigned with one-week interval. After one-hour wearing, participants were asked to indicate the comfort level and the discomfort region on face. The results showed that no differences were noticed between these two molds. But participants with 3D printed mold felt much more comfortable during the casting process. 3D printing technology was a promising method to help therapist fabricate the TFO. This method was faster and less stressful for patients. However, one limitation of this study was the lack of comparisons of comfort rating from patients with facial burns. Future research is needed to examine the comfortable rating for burn patients.

Keywords: *Assistive Technology, Burns, 3D Printing*

[a] Department of Rehabilitation Medicine, Kaohsiung Medical University Hospital
[b] Division of Plastic Surgery, Department of Surgery, Kaohsiung Medical University Hospital
[c] Department of Surgery, Faculty of Medicine, Kaohsiung Medical University
[d] Department of Occupational Therapy, Kaohsiung Medical University
[e] Department of Medical Research, Kaohsiung Medical University Hospital

[*] Correspondence: Yu-Sheng Yang
100 Shih-Chuan 1st Road, Kaohsiung 807, Taiwan
TEL: 07-3121101 ext. 2657
FAX: 07-3215845
E-mail: yusheng@kmu.edu.tw

Received: 09 July 2018
Accepted: 31 Dec 2018

1. Introduction

Hypertrophic scars are an unavoidable complication of burn injuries. It can limit a survivor's functional ability as well as affect their body image. The use of external pressure applied by non-invasive means of preventing and controlling of hypertrophic scars have been widely accepted (Cheng et al., 1984; Linares, Larson, & Willis-Galstaun, 1993; Mann, Yeong, Moore, Colescott, & Engrav, 1997; Staley & Richard, 1997; Ward, 1991). Hypertrophic scars could become softer and flatter if adequate external pressure is applied for a correct period of time. Otherwise scar tissue may continue to grow until the elasticity of the surrounding tissue is exhausted (Rivers, Strate, & Solem, 1979). Existing research showed that pressure reduced minor problems such as itchiness, discomfort and blistering (Leung & Ng, 1980).

Pressure garments or transparent face masks are widely used conventional methods for burn scars. Pressure garments seem to work well over areas of the body. However, the complex and contoured surface of facial area is a challenge for pressure garments, even for skilled therapists or manufacturers. If the pressure garment does not provide uniform pressure to contoured areas of the body, it may increase the tendency of those areas to form hypertrophic scars. Furthermore, pressure garments fail to apply enough pressure in areas of high contour or during movement and tend to lose their elasticity over time (Ahn, Monafo, & Mustoe, 1989; Cheng, et al., 1984; Harries & Pegg, 1989).

A transparent face orthosis (TFO), sometimes called a transparent facial mask, is widely used for facial burns treatment. TFO is effective to reduce hypertrophic scarring and to maintain the facial contours (Esselman, Thombs, Magyar-Russell, & Fauerbach, 2006; Powell, Haylock, & Clarke, 1985; Rivers, et al., 1979; Shons, Rivers, & Solem, 1981). Straps and anchors that secured TFO provide compression pressure. TFO fabricated from an accurate pattern of the head eliminates many of the disadvantages of pressure garments. The vascular blanching of scar beneath the TFO allows the therapist to ensure a proper fit acutely. Furthermore, Groce, Meyers-Paal, Herndon, and McCauley, (1999) found that although no significant difference in pressure under a TFO compared with a custom pressure garment, patients with noticeable scarring around the face preferred the TFO because it was accepted as normal to expose facial features in public.

Two methods of fabricating the facial mask could be practiced by the occupational therapist in rehabilitation unit. First is the conventional plaster method that involves multiple complicated steps and direct contact with the scars. The fabricating process includes using dental alginate as a casting material, applying plaster bandages for reinforcement, filling liquid plaster for molding the mask, and using high-temperature thermoplastics to vacuum the mask mold (Locke, 1991). This conventional method is a complex and labor-intensive task and often causes discomfort to the patient. 64% of therapists reported that it took 6 to 10 hours to fabricate a TFO, and 68% stated that the most difficult step of making a TFO was “casting” or “modifying the mold”. The survey also revealed that 86% of therapists needed to recast

the patient due to improper fit (Parry, Doyle, Hurlin-Foley, Palmieri, & Greenhalgh, 2002).

The second technology for fabricating TFO is using a 3D scanner that scans the facial area. Several types of 3D scanner, (e.g. 3D digitizers, laser scanners, white light scanner and others) were used to provide the cloud data. The point of cloud data from the scanning process is used to generate the mask using computer-aided design (CAD) software, and the part is exported to a stereolithography (STL) file for rapid prototyping process. The scanning process increases the fitting between the material and the face surface.

Face masks built using this 3D scanner technology have been reported in two case studies, and shown to yield better outcomes with accuracy and fitting of the mask in comparison with using the conventional method. The first study built the facial contour of a burn patient by 3D surface scanning. The ZCorp machine with the powder based rapid prototyping system was used to build the mask mold (Pilley, Hitchens, Rose, Alexander, & Wimpenny, 2011). The second study on two children with facial burns was conducted by a portable 3D scanning and CAD to produce 3D-printed transparent facemasks (Wei, Li, Liu, Xie, & Yue, 2017). The results indicated that scar thickness was decreased and the patients were more satisfied with their facial appearance. The 3D-printed facemasks were well fitted on two patients. However, the authors did not report the model of 3D-printer so we are unable to identify the 3D printed method used in their study. But they concluded that 3D-printed transparent facemask was a feasible means for treating pediatric facial hypertrophic scars after burn.

As 3D printing technology improves, more 3D printed methods are used to implement the project. In health care, rapid prototyping techniques like stereolithography (SLA), fused deposition modeling (FDM), and selective laser sintering (SLS) are commonly used to manufacture medical devices (Rengier et al., 2010). The FDM printers use a thermoplastic filament, which is heated to its melting point and then extruded, layer by layer, to form the object. Due to its application in varied fields such as architecture, engineering, and electronics medicine, 3D printing technology is getting more available to the consumer market today, as an inexpensive go-to machine. Therefore, the purpose of this study was to investigate the feasibility of fabrication of TFO by use of non-contact structured light scanning with 3D printing technology. It was intended to provide a practical means of new fabrication technology for TFO.

2. Methods

2.1 Participants

Twenty able-bodied participants were recruited through posters in public spaces. Participants were included in the study if they met the following criteria: 20 years of age or older; did not have previous history of musculoskeletal or nerve injuries causing involuntary facial movements such as facial tics. If the participant appeared facial skin irritation after wearing mask, the participant was excluded. All participants were informed of the study's risks and benefits, and their identity would not be disclosed. Their

participation was voluntary. Afterwards, participants signed the informed consent in accordance with the procedures approved by the Institutional Review Board of the Kaohsiung Medical University Hospital prior to participation in the study.

2.2 Fabrication of TFO using traditional method

Participants were positioned in semi-reclined position on the hospital bed. One experienced occupational therapist performed the casting produces for all participants. The total casting procedure took approximately 25 minutes for each participant, but the TFO took approximately 72 hours to complete due to plaster mold drying time. In the first step of making the mask, the participant's face was covered with dental alginate impression materials (Figure 1) and reinforced with plaster bandages. Afterwards, the impression was removed from the patient's face. The inside of the impression duplicated the curves of the person's face. The face impression then was filled with liquid plaster (Figure 2). When the plaster dried, it formed a solid head. Next, the therapist cleaned the plaster head and sanded it smooth. A sheet of transparent thermoplastic was then heated in an oven until the plastic was soft and pliable. The warm plastic sheet was then draped and completely pressed over the plaster head with a vacuum former. After the plastic cooled down, the therapist cut openings in the mask for eyes, nose, mouth and ears, and attached elastic straps to the sides.

Figure 1

Using dental alginate impression materials to get the facial profile

Figure 2

A positive mold of the face was fabricated with the use of liquid plaster

2.3 Fabrication of TFO using 3D printing technology

After participants finished the casting of face impression, they used face wipes to clean their faces. Participants were positioned in the semi-reclined position again. The three-dimensional face image of participant was digitally scanned by a handheld scanner (iSense™ 3D-Scanner, 3D Systems, Inc, South Carolina, USA). The acquisition time was less than one minute to complete the image. Afterwards, the imager was operated by Blender software package (www.blender.org). Blender is a professional open-source and free 3D modeling software for creating 3D pipeline, modeling, animation, and rendering. The data point clouds generated from the handheld scanner were then cleaned of extraneous data by trimming tool and fitted with 3D surfaces via Blender. The finished 3D facial model was then exported to a STL file for use in 3D printing. A 3D printer (Mbot Black Stone plus, Armsource Inc, Kaohsiung, Taiwan) with the quality 0.2 mm layer resolution was used to make a face mold (Figure 3). The printing process took 24 to 32 hours depending on the size of face mold. Afterward, the following steps were the same as described above for the traditional method including heating transparent thermoplastic, pressing over the face mold with a vacuum former, and cutting openings in the mask (Figure 4).

2.4 Questionnaire

After completing fabrication of TFO by using the two methods, the participants were asked to wear the TFO which were made from either the conventional plaster mold or 3D printed mold for one hour. After one-hour wearing TFO, the participants filled out a subjective questionnaire. Ten

questions including a question regarding recall of comfort level during fabrication and a question regarding facial discomfort regions (Figure 5) were constructed to form this subjective comfort questionnaire. Perceptions of TFO fitting comfort were measured by a visual analog scale of 10cm in length (Daniel, 1995). The left end of the scale (0cm) represented extreme discomfort and the right end of the scale (10cm) represented extreme comfort. Participants were asked to place an "X" on the visual analog scale to show their level of comfort for each TFO (DiGiovine, Cooper, Boninger, Lawrence, VanSickle, & Rentschler, 2000).

The wearing session of two TFOs was one week apart to minimize problems with recall over a more extended time interval. The sequential order of the wearing model was assigned by a random number generator.

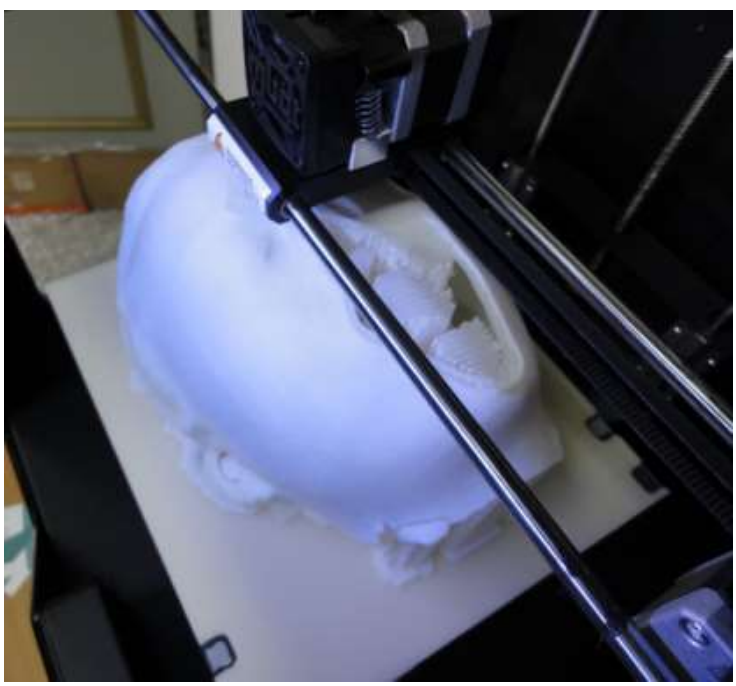

Figure 3
A positive mold of the face was fabricated with the use of 3D printing technology

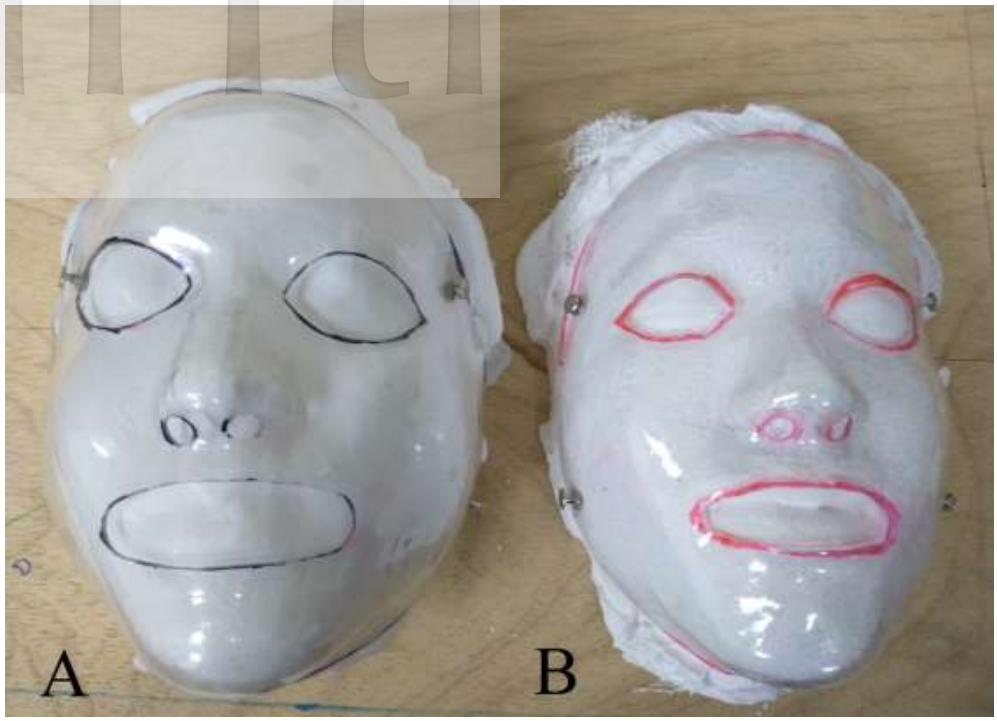

Figure 4

The final product of TFO by using 3D printed mold (A) and plaster mold (B)

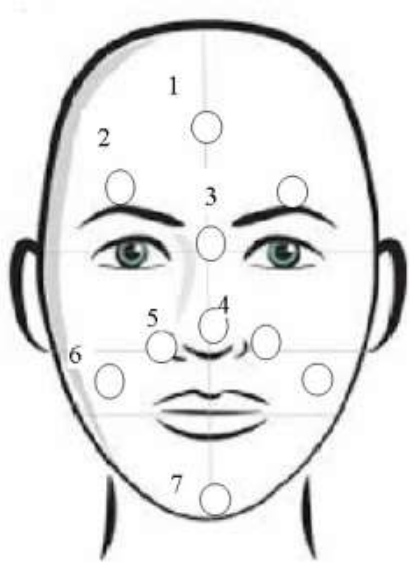

Figure 5

The regions of physical discomfort over facial areas

2.5 Statistical analyses

Descriptive statistics was used to describe the presence of physical discomfort in the regions of the face. A Wilcoxon Signed Rank Test was used to look for significant differences of comfort level between TFO made from the traditional plaster mold versus 3D printed mold. All statistical analyses were performed using IBM SPSS Statistics for Windows, Version 21.0 (IBM Corp, Armonk, NY, USA). The level of significance was set to 0.05.

3. Results

The average age, height and weight of participants were 23.13±2.17 years old, 164.20±5.34 cm, 53.00±5.98 kg, respectively. The gender ratio in participants was 1:1. There were statistically significant differences in reported comfort levels between plaster mold casting and 3D printed mold for casting TFO. Participants with 3D printed mold felt much more comfortable during the casting process (mean=8.5$\pm$0.7 versus 5.9$\pm$1.1, p<0.01). In term of comfortable fit with TFO, there are no significant differences (p=0.57). The comfort of the participants with 3D printed model ranged from 2.1 to 8.7 (SD=2.5, median=4.0), and ranged from 1.8 to 7.9 (SD=2.0, median=3.8) with traditional plaster mold. The presence of self-reported physical discomfort across seven facial regions is presented in Table 1. Most participants with 3D printed model reported experiencing discomfort in the root of nose (90%); 40% participants reported discomfort in supraorbital ridge; 20% participants reported in forehead region; and 5% participants reported in chin. On the other hand, 50% participants with conventional plaster mold reported

experiencing discomfort in the root of nose, 30% participants reported in forehead region, 20% participants reported in supraorbital ridge and chin region. No severe complaints such as stinging, itching, and irritation on facial skin were reported by any participant after wearing TFO for 1 hours.

Table 1

The presence of self-reported physical discomfort across facial regions (N=20)

	Forehead	Supraorbital ridge	Root of nose	Tip of nose	Wing of the nose	Zygoma	Chin
3D printed mold	4/20 (20%)	8/20 (40%)	18/20 (90%)	0/20 (0%)	0/20 (0%)	0/20 (0%)	1/20 (5%)
Traditional plaster mold	6/20 (30%)	4/20 (20%)	10/20 (50%)	0/20 (0%)	0/20 (0%)	0/20 (0%)	4/20 (20%)

4. Discussion

The use of 3D scanning systems has gained popularity in recent years. The handheld scanner rotated 360° around the participant's head and captured 3D models of the facial profile and contour. When using this scanning process, there was no direct contact with the participant in the process, which was different from the traditional method that was mostly painful to the participant (Rogers et al., 2003). The scanning process took just a few seconds, significantly reducing the amount of time and resources devoted to each step. Furthermore, a handheld 3D scanner was easy to use. The simple point-and-shoot action ensured that data can be collected easily by healthcare professionals. The noncontact nature of the 3D scanner

combined with the speed of shape acquisition eliminated most of the drawbacks of conventional fabrication.

According to the results, the root of nose was the most discomfort region reported by the participants with 3D printed model. Although the handheld scanner used in this study claimed resolutions as high as 0.5mm and an accuracy of up to 4mm, our 3D printed model resulted in poor accuracy in particular region of facial contour. Through structured light technology of 3D scanner, highlight bridge of nose (usually between eyebrow) might shade down the sides of the nose. Therefore, it might not be able to capture all details in the concavity of the nasal root. It was suggested that softening of a bony nose contour should be done carefully by the CAD software to create facial model. Besides, each year 3D scanning technology advances in terms of accuracy and speed. The technology improves significantly on the reliability and on perfect 3D modeling. Therefore, it can be expected that more handheld scanners will deliver a very high level of feature accuracy in the near future.

Our results show that the non-contact structured light scanning with 3D printing technology is feasible to fabricate TFO. No difference was noticed between wearing TFO fabricated by 3D printed mold and traditional plaster mold. The total cost of the handheld scanner with a 3D printing system was less than $50,000 NTD. The software used in this study, Blender, is a free 3D animation program. It integrates the steps required for TFO fabrication. It can import data from a scanned image and export 3D solid objects in the STL file format compatible with all major 3D printers. Therefore, it would be feasible for a burn care facility to have this low-cost system for mask fabrication. The

TFO could be designed onsite by a therapist, and fabrication of mask could be done locally at the same site. If a 3D printer is available in the facility, then a mask can be available in just few hours based on printing speed and quality settings.

The advantages of non-contact scanning with 3D printing technology for TFO fabrication include easy to use, rapid process, and requires minimal training. The major strength is that the proper face topology can be created quickly and accurately without discomfort to the patient. Generally speaking, it appears to be a less anxiety-provoking, less uncomfortable, and more efficient process than conventional methods. However, one limitation of this study was the lack of comparisons of comfort rating from patients with facial burns. Thus, the results from this study may not well reflect satisfaction of burn patients. Nevertheless, this study supports the feasibility of using 3D-printed TFO as a novel method for treating facial hypertrophic scars for burn patients. More clinical cases should be examined in the future to refine the methods and consolidate the knowledge.

5. Conclusion

3D scanners and printers are improving accuracy and speed each year. 3D printing technology has a great potential to improve the quality of life of people with disabilities by providing options for highly customized and affordable assistive devices for daily living activities. Over the years, major improvements have been made in 3D scanning and printing materials, which

made 3D facial model for TFO fabrication possible and reliable. Burn victims may be benefited greatly from the fabrication of face masks utilizing this technology with less uncomfortable process.

References

Ahn, S. T., Monafo, W. W., & Mustoe, T. A. (1989). Topical silicone gel: a new treatment for hypertrophic scars. *Surgery, 106*, 781-786.

Cheng, J. C., Evans, J. H., Leung, K. S., Clark, J. A., Choy, T. T., & Leung, P. C. (1984). Pressure therapy in the treatment of post-burn hypertrophic scar-a critical look into its usefulness and fallacies by pressure monitoring. *Burns, including thermal injury , 10*, 154-163.

Daniel, W. W. (1995). *Biostatistics: A foundation for analysis in the health sciences*. New York: Wiley.

DiGiovine, M.M., Cooper, R.A., Boninger, M.L., Lawrence, B.M., VanSickle, D.P., & Rentschler, A.J. (2000) User assessment of manual wheelchair ride comfort and ergonomics. *Archives of Physical Medicine and Rehabilitation, 81,* 490-494.

Esselman, P. C., Thombs, B. D., Magyar-Russell, G., & Fauerbach, J. A. (2006). Burn rehabilitation: state of the science. *American Journal of Physical Medicine and Rehabilitation, 85*, 383-413.

Gauglitz, G. G., & Jeschke, M. G. (2011). Pathophysiology of burn injury. In M. G. Jeschke, L.-P. Kamolz, F. Sjoberg & W. S.E. (Eds.), *Handbook of Burns Volume 1: Acute Burn Care* (pp. 131-148). New York: Springer.

Groce, A., Meyers-Paal, R., Herndon, D. N., & McCauley, R. L. (1999). Are your

thoughts of facial pressure transparent? *Journal of Burn Care and Research, 20*, 478-481.

Harries, C. A., & Pegg, S. P. (1989). Measuring pressure under burns pressure garments using the Oxford Pressure Monitor. *Burns, 15*, 187-189.

Leung, P. C., & Ng, M. (1980). Pressure treatment for hypertrophic scars resulting from burns. *Burns, 6*, 244-250.

Linares, H. A., Larson, D. L., & Willis-Galstaun, B. A. (1993). Historical notes on the use of pressure in the treatment of hypertrophic scars or keloids. *Burns, 19*, 17-21

Locke, S., Smith, S., Szeliski-Scott, B., & Lemaire, E. D. (1991). A Clear Polycarbonate Face Mask for the Treatment of Hypertrophic Scars. *JPO: Journal of Prosthetics and Orthotics, 3*, 182-190.

Mann, R., Yeong, E. K., Moore, M., Colescott, D., & Engrav, L. H. (1997). Do custom-fitted pressure garments provide adequate pressure? *Journal of Burn Care and Rehabilitation, 18*, 247-249.

Parry, I., Doyle, B., Hurlin-Foley, K., Palmieri, T. L., & Greenhalgh, D. G. (2002). Does the First Impression Count? A Survey of Current Practices and Satisfaction with the Use of Transparent Face Masks in the Treatment of Facial Scarring. *Journal of Burn Care and Rehabilitation, 23*, S55.

Pilley, M. J., Hitchens, C., Rose, G., Alexander, S., & Wimpenny, D. I. (2011). The use of non-contact structured light scanning in burns pressure splint construction. *Burns, 37*, 1168-1173.

Powell, B. W., Haylock, C., & Clarke, J. A. (1985). A semi-rigid transparent face mask in the treatment of postburn hypertrophic scars. *British Journal of Plastic Surgery, 38*, 561-566.

Rengier, F., Mehndiratta, A., von Tengg-Kobligk, H., Zechmann, C. M., Unterhinninghofen, R., Kauczor, H. U., & Giesel, F. L. (2010). 3D printing based on imaging data: review of medical applications. *International Journal for Computer Assisted Radiology and Surgery, 5,* 335-341.

Rivers, E. A., Strate, R. G., & Solem, L. D. (1979). The transparent face mask. *American Journal of Occupational Therapy, 33*, 108-113.

Rogers, B., Chapman, T., Rettele, J., Gatica, J., Darm, T., Beebe, M., Dilworth, D. & Walsh, N. (2003). Computerized manufacturing of transparent face masks for the treatment of facial scarring. *Journal of Burn Care and Rehabilitation, 24*, 91-96

Shons, A. R., Rivers, E. A., & Solem, L. D. (1981). A rigid transparent face mask for control of scar hypertrophy. *Annals of Plastic Surgery, 6*, 245-248.

Staley, M. J., & Richard, R. L. (1997). Use of pressure to treat hypertrophic burn scars. *Advances in Wound Care, 10*, 44-46.

Ward, R. S. (1991). Pressure therapy for the control of hypertrophic scar formation after burn injury. A history and review. *Journal of Burn Care and Rehabilitation, 12*, 257-262.

Wei Y, Li-Tsang CWP, Liu J, Xie L, Yue S. (2017). 3D-printed transparent facemasks in the treatment of facial hypertrophic scars of young children with burns. *Burns,* 43, e19-e26.

職能治療學會雜誌
民國 107 年 36 卷第 2 期頁 228-245
DOI: 10.6594/JTOTA.201812_36(2).0005

原著

運用 3D 列印技術製作全臉式燒傷壓力面罩之可行性研究

梁文隆 [1] 李萬盟 [1] 黃書鴻 [2,3] 張志仲 [4] 楊育昇 [4,5,*]

摘要

3D 列印是一種數位模型檔案為基礎，能直接製造三維實體的技術。它可用十分低廉的成本印出客製化及複雜的模型。據此，本研究的目的在於驗證透過非接觸體表的 3D 掃瞄與列印技術，來製作透明塑膠面罩的可行性。共計 20 位健康受測者參與此研究。在透明塑膠面罩分別透過 3D 列印模型與傳統石膏模型製作方法完成後，受測者以一星期為時間間隔，依隨機順序來分別穿戴這兩款透明塑膠面罩一小時。穿戴一小時後，對受測者進行問卷調查以瞭解穿戴的舒適性與造成臉部不舒服的部位。結果顯示這兩種製作方法對穿戴舒適性並無顯著的差異，然而受測者自覺 3D 列印製模的過程比較舒服。本研究結果顯示利用 3D 列印技術來協助治療師製作透明塑膠面罩是可行的方案；此方案對患者而言省時且較舒適。然而，本研究的限制是缺乏比較臉部燒傷患者穿戴後的感受，未來有待針對此族群長期穿戴的感受做進一步研究。

關鍵字：3D 列印，燒傷，輔具科技

高雄醫學大學附設醫院復健科 [1]
高雄醫學大學附設醫院整形外科 [2]
高雄醫學大學醫學系外科學科 [3]
高雄醫學大學職能治療學系暨研究所 [4]
高雄醫學大學附設醫院臨床醫學研究部 [5]

*通訊作者：楊育昇
高雄市十全一路 100 號
電話：07-3121101 分機 2657
傳真：07-3215845
電子信箱：yusheng@kmu.edu.tw

受文日期：民國 107 年 07 月 09 日
接受刊載：民國 107 年 12 月 31 日