

《測驗學刊》[DOI: 10.7108/PT.202503_72(1).0001]
第七十二輯第一期 2025 年 3 月 第 1~32 頁

「華語兒童圖畫詞彙理解與表達測驗」之編製與應用

劉惠美¹ 張鑑如² 陳柏熹³ 陳雅雯⁴
陳修元⁵ 林佳慧⁶

摘要

本研究之目的在編製 3~12 歲華語兒童圖畫詞彙理解與表達能力之標準化測驗工具，以了解當代華語學前和學齡兒童的詞彙能力。理解與表達兩個分測驗皆包括 A、B 兩式，理解測驗 A、B 兩式各為 52 題，表達測驗 A、B 兩式各為 35 題。測驗研發歷程包括選題、預試和正式施測三階段。正式施測人數共 1,708 人，含臺灣北、中、南、東四區 3 歲至國小六年級兒童，其中施以 A 式測驗兒童共 881 人，B 式測驗兒童共 827 人。研究結果顯示，圖畫詞彙理解測驗和圖畫詞彙表達測驗 A、B 兩式的內部一致性係數皆大於 .90，折半信度大於 .87，具良好的內部一致性。研究者以「修訂畢保德圖畫詞彙測驗」（PPVT-R）作為效標參照，兩個分測驗 A、B 兩式皆與其有高度正相關，顯示本測驗的效標關聯效度佳。此外，無論學前或學齡各年齡組，兩個分測驗的表現皆為高年齡組得分高於低年齡組，且有顯著的年齡組

1. 劉惠美，國立臺灣師範大學特殊教育學系、學習科學跨國頂尖研究中心、社會情緒發展研究中心教授，ORCID: 0000-0001-9586-7675
 2. 張鑑如，國立臺灣師範大學幼兒與家庭科學學系教授，ORCID: 0000-0001-7776-6678
 3. 陳柏熹，國立臺灣師範大學教育心理與輔導學系、學習科學跨國頂尖研究中心教授，ORCID: 0000-0002-5377-2077
 4. 陳雅雯，幼兒發展調查資料庫建置計畫研究助理
 5. 陳修元，臺北市立大學幼兒教育學系副教授，ORCID: 0000-0003-3554-3956
 6. 林佳慧，國立臺灣師範大學博士後研究員，ORCID: 0000-0002-1356-4214
- 收件日期：2023.06.02；完成修改：2024.01.24；正式接受：2024.01.24
通訊作者：張鑑如；Email：changch2@ntnu.edu.tw
地址：106308 臺北市大安區和平東路一段 162 號
國立臺灣師範大學幼兒與家庭科學學系

測驗學刊，第 72 輯第 1 期

間差異，顯示本測驗具有良好的區辨效度。試題反應理論分析結果顯示，兩個分測驗 A、B 兩式的題目適配度良好。本測驗為減少年幼兒童因施測時間和難度的壓力，根據兒童的能力值去對應題目難度，訂出適合各年齡組施測的題組，具有良好的信效度，並開發電腦化施測程序，以便於施測和計分。此工具應可廣泛應用於學前至學齡兒童的詞彙能力評估，以作為相關教學介入成效之評估參考。

關鍵詞：華語兒童、詞彙理解、詞彙表達、詞彙測驗、語言評估

Development and Application of Picture Vocabulary Test for Mandarin Chinese-Speaking Children: Receptive and Expressive Vocabulary Skills

Huei-Mei Liu¹ Chien-Ju Chang² Po-Hsi Chen³
Ya-Wen Chen⁴ Shiou-Yuan Chen⁵ Chia-Hui Lin⁶

Abstract

This study aimed to develop a standardized assessment tool to measure the receptive and expressive vocabulary skills of contemporary Mandarin-speaking children aged 3 to 12. The assessment comprises receptive and expressive vocabulary tests in

1. Huei-Mei Liu, Professor, Department of Special Education, Institute for Research Excellence in Learning Sciences, Social Emotional Education and Development Center, National Taiwan Normal University, ORCID: 0000-0001-9586-76
2. Chien-Ju Chang, Professor, Department of Child and Family Science, National Taiwan Normal University, ORCID: 0000-0001-7776-6678
3. Po-Hsi Chen, Professor, Department of Educational Psychological and Counseling, Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, ORCID: 0000-0002-5377-2077
4. Ya-Wen Chen, Research Assistant, Kids in Taiwan: National Longitudinal Study of Child Development and Care
5. Shiou-Yuan Chen, Associate Professor, Department of Early Childhood Education, University of Taipei, ORCID: 0000-0003-3554-3956
6. Chia-Hui Lin, Postdoctoral Researcher, National Taiwan Normal University, ORCID: 0000-0002-1356-4214

Received: 2023.06.02; Revised: 2024.01.24; Accepted: 2024.01.24

Corresponding Author: Chien-Ju Chang; Email: changch2@ntnu.edu.tw

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

Department of Child and Family Science, National Taiwan Normal University

two variants: A and B. Each version of the receptive vocabulary test contains 52 items, while each version of the expressive vocabulary test has 35 items. The test development process involved three phases, namely preliminary, pilot, and formal testing. In the formal study, 1,708 children from preschool through sixth grade, residing in the northern, central, southern, and eastern regions of Taiwan, participated. Among them, 881 children completed version A of the assessment, while 827 children completed version B. The results showed that both versions of the receptive and expressive vocabulary tests exhibited good internal consistency, with Cronbach's alpha being greater than .90 and split-half reliability being higher than .87. Both versions showed a strong positive correlation with the revised Peabody Picture Vocabulary Test, indicating good criterion-related validity. The scores of both tests were higher for older age groups, signifying robust discriminant validity. An analysis based on the item response theory showed that both test versions exhibited good item fit. To reduce testing complexity and time constraints for young children, the researchers developed age-appropriate question sets for each age group. Additionally, a computer-based version of the test was developed to facilitate scoring and administration. The findings showed that this tool can be widely used to assess the language skills of preschool and school-aged children, and can serve as a reference for evaluating educational interventions in the future.

Keywords: language assessment, Mandarin-speaking Chinese, vocabulary comprehension, vocabulary production, vocabulary test

壹、緒論

語言能力是有效溝通的基本能力，涵蓋了語音、語意、語法和語用等面向，各面向又可區分為理解和表達向度，面向間相互運作，以達到有效溝通的目的。其中，語意向度主要關注詞彙能力，反映出個體對指涉（referent）概念內涵和知識的了解（Golinkoff et al., 2019），是兒童早期語言發展的關鍵能力，與認知發展密切相關（Pérez-Pereira & Cruz, 2018; Rowe et al., 2012），更是預測閱讀能力和學業表現的重要指標（Bleses et al., 2016; NICHD Early Child Care Research Network, 2005; National Reading Panel, 2000; Ouellette & Beers, 2010; Protopapas et al., 2013; Sénéchal et al., 2006）。若能以具有良好信效度的詞彙評估工具，早期評估兒童的詞彙理解與表達能力，並持續追蹤跨年齡的發展，將有助於了解一般發展兒童語言能力的成長趨勢，也有利於語言障礙兒童的早期發現與鑑定，是值得投注心力的研究議題。

鑑於過去對華語兒童詞彙能力（含理解和表達）之實證評量工具並不多，加上需考量學前幼兒本身反應能力的特性與限制，能跨學前至學齡階段進行跨時追蹤與評量的工具尤其稀少。本研究之目的在於編製適用於評估3~12歲華語兒童詞彙理解與表達能力之標準化測驗工具，以了解臺灣華語學前和學齡兒童的詞彙能力，並作為未來兒童語言發展研究的本土化工具。

一、兒童詞彙能力的發展

詞彙學習是長期的發展與累積歷程。就向度而言，詞彙能力可以區分為理解性詞彙及表達性詞彙；就發展的觀點，理解性詞彙的發展通常早於表達性詞彙（Fenson et al., 2007）。國外研究指出，多數嬰兒在5~6個月大時就可以辨識或理解出現頻率較高的詞語，例如：自己的名字或生活中常用的物品；13個月大的幼兒大約能夠理解100個以上詞彙，16個月大時的理解詞彙量約為180個，6歲時的理解詞彙量已達20,000至24,000個之間（Fenson et al., 2007; Hoff, 2015; Owens, 2007）。在華語研究部分，劉惠美與曹峰銘（2010）利用「華語嬰幼兒溝通發展量表」〔Mandarin-Chinese Communicative Development Inventory (Taiwan) [MCDIT]，即華語版CDI〕大規模分析臺灣習華語嬰幼兒的家長觀察報告，發現9~10個月大的嬰兒約可理解50個詞彙，14~15個月大的嬰兒則可理解150個詞彙，認為臺灣嬰幼兒的詞彙發展

趨勢與其他語言背景嬰幼兒的發展趨勢大致相符。

在表達性詞彙方面，Hoff（2015）以及 MacWhinney（2015）指出，多數嬰幼兒會在 1 歲左右出現第一個有意義的詞彙。當幼兒剛開始說話時，詞彙產出的進展速度較慢，而且對特定詞彙的指涉對象和範圍通常比成人侷限得多，大多以表達即時即地的訊息為主。但隨著年齡和認知經驗增長，幼兒產出詞彙的意義會愈來愈去脈絡化，並與成人相像。有些幼兒在 2 歲前會有「詞彙爆發」（word spurt）的現象，意指詞彙累積的速率突飛猛進，但多數幼兒是以持續增長的速度在擴展詞彙（Hoff, 2015; MacWhinney, 2015），例如：說英語幼兒在 15 個月大時約有 15 個表達性詞彙，18 個月大時約有 50 個詞彙，到了 2 歲時，表達性詞彙已大幅增加，約有 200~300 個詞彙，3 歲時則約為 900~1,000 個詞彙，4 歲時接近 1,500 個詞彙，5 歲時大約已有 2,000 個詞彙。劉惠美與曹峰銘（2010）則發現，臺灣習華語嬰幼兒多數在 10~12 個月時發出第一個詞彙，在 16 個月大時的表達性詞彙約有 40 個，到了 20 個月大時已增加為 100 個詞彙，24 個月大時約有 200 個詞彙，26 個月大時約有 300 個詞彙，此發展趨勢也和其他語言背景的嬰幼兒相似。然而，該研究採用 696 個詞彙清單，請家長從中勾選孩子在特定月齡時曾說出過的詞彙，發現臺灣 36 個月幼兒在該量表測得的表達性詞彙僅約 500~600 個，少於文獻中習英語兒童的 900~1000 個詞彙，推測差異的主要原因是因為測量方法不同與工具特性所致之天花板效應。

除了探討詞彙量的發展趨勢外，兒童表達性詞彙的詞性類別也受到重視。當表達性詞彙達到 50 個之後，不同的詞性類別也會開始出現，例如：名詞、動詞或形容詞，以及只有語法功能而沒有語意內容的虛詞。過去文獻指出，英語為母語兒童的早期詞彙有明顯之「名詞偏好」（noun bias），意即指稱人物、食物、物品等的名詞比例最高，動詞比例則次之（劉惠美、陳昱君，2015; Choi & Gopnik, 1995; Ogura et al., 2006）。透過「麥克阿瑟貝茲溝通發展量表」（The MacArthur-Bates Communicative Development Inventories [MB-CDis]）所蒐集的家長觀察報告也發現，當習英語兒童的詞彙量在 100~200 個之間時，普通名詞大約占 55%，而動詞、形容詞和虛詞的比例則不到 20%（Fenson et al., 2007）。在發展速率上，動詞、形容詞和虛詞要等到詞彙量約達到 400 個之後，才會有明顯的提升。

劉惠美與陳昱君（2015）以「華語嬰幼兒溝通發展量表」（MCDIT）所蒐集的家長觀察報告則發現，在 1,897 名 16~36 個月大的臺灣幼兒中，當孩

子在 18~19 個月大，表達詞彙約達 50 個詞時，以指稱人物、動作和身體部位的名詞為最多，出現早期名詞偏好的跨語言普遍現象。然而，根據語言樣本的分析結果，華語兒童的「名詞偏好」不如說英語幼兒明顯，「名詞偏好」會受到語言和情境影響，例如：美國學者 Tardif 根據語言樣本分析結果，發現華語幼兒「名詞偏好」的現象沒有說英語的幼兒明顯（Tardif, 1996; Tardif et al., 2009; Tardif et al., 1999）。Hung 與 Chang（2023）也採用語言樣本分析，發現 26 及 36 個月大幼兒在親子共玩玩具的情境下，動詞的產出高於名詞，顯示語言樣本的蒐集情境式也可能會影響結果。

隨著年齡增長至 3 歲以上，雖然不容易實際測量出兒童能理解或表達的詞彙數量，但有敘事研究指出，3~6 歲兒童所使用的詞彙仍以實詞為主，其中動詞、名詞、副詞和代詞是最主要的詞彙類別，且又以名詞和動詞的數量增加最多（涂嘉新、陳姿蘭，2020）。也有跨語言研究以圖片指認和命名的方式測試，發現在大多數語言環境下，3~7 歲兒童對名詞的理解和表達正確率普遍高於動詞，且詞彙理解優於詞彙表達能力（Haman et al., 2017），但仍有一些例外的情形。

本研究所編製之華語兒童詞彙能力評估工具，將參考上述文獻提及實詞為兒童詞彙的主要詞性類別，選取以名詞、動詞和形容詞等實詞為主的詞彙，以期能適切評估華語兒童的詞彙能力發展。

二、華語兒童詞彙能力評估工具

如前所述，詞彙能力關乎溝通與閱讀能力的發展，也是評估兒童是否具有語言發展困難的早期重要指標。因此，建立一套具有信、效度的詞彙評量工具相當重要。目前，針對臺灣華語兒童詞彙發展的評估工具大致有三類：(1)家長的觀察報告，如上述「華語嬰幼兒溝通發展量表」（MCDI-T）（劉惠美、曹峰銘，2010）；(2)標準化的詞彙測驗，如「修訂畢保德圖畫詞彙測驗」（Peabody Picture Vocabulary Test-Revised [PPVT-R]；陸莉、劉鴻香修訂，1998）、「華語兒童理解與表達詞彙測驗」（Receptive and Expressive Vocabulary Test [REVT]；黃瑞珍等人，2010）；(3)語言樣本分析（language samples）（洪宜芳、張鑑如，2019；Bernstein Ratner & MacWhinney, 2016; Hung & Chang, 2023）。

「華語嬰幼兒溝通發展量表」（MCDI-T；劉惠美、曹峰銘，2010）主要是參考英語版 MB-CDIs（Fenson et al., 2007）的架構，編製適用於國內語

言環境的嬰幼兒語言和溝通發展評量工具，具有信效度和臨床應用性。該量表分為 8~16 個月嬰兒版和 16~36 個月幼兒版，在詞彙的理解和表達上，提供詞彙清單，由家長進行勾選。該工具是透過家長的觀察報告，因此可涵括較多生活情境，也可大樣本施測。然而，此量表的主要限制在於僅適用小於 36 個月大的嬰幼兒，且有研究指出這類以家長報告形式的評量結果較欠缺對未來語言發展的明確預測力（Dockrell, 2001）。此外，要求家長同時觀察孩子多面向的語言溝通能力，對有些家長而言要求過高，也會出現過於主觀的問題。

在標準化測驗工具上，國內研究比較常使用的 PPVT-R（陸莉、劉鴻香修訂，1998），是由幼兒指出一組圖片中符合所聽到詞彙的圖畫，以估計其理解性詞彙的發展。由於施測方法簡單，且涵括年齡範圍廣，適用於 3~12 歲，並具有常模可對照，因此常被用於評估兒童的口語能力或簡易智力的推估。但該測驗僅能評估兒童的理解性詞彙能力，且其題目內容是參照 1981 年的英文版 PPVT-R（Dunn & Dunn, 1981），其常模建置距今已逾四十年，所涵括的詞彙範圍或是圖畫的呈現均有更新的必要。而「華語兒童理解與表達詞彙測驗」（REVT；黃瑞珍等人，2010）包括命名、歸類、定義和推理四個分測驗，其主要目的在於了解幼兒的詞彙理解和表達能力，以鑑定詞彙發展遲緩幼兒，適合施測的年齡也僅限於 3~6 歲，對評估 7 歲以上一般發展兒童的詞彙能力則不適用。

語言樣本分析則是蒐集幼兒在自然或實驗情境的語料，在語意方面可根據語料分析詞彙的使用，例如：總詞彙數、相異詞彙數、詞彙多樣性等（洪宜芳、張鑑如，2019；Bernstein Ratner & MacWhinney, 2016；Hung & Chang, 2023；Spaulding et al., 2012）。語言樣本分析具有較佳的生態效度（Botting, 2002），但由於語料轉譯和分析的過程相當費時，因此不易做為大規模施測時的評量工具；且語言樣本分析也僅能針對表達性詞彙進行評估，無法評估理解性詞彙的能力。

基於以上詞彙評量工具在適用年齡、時效性和施測規模上均有其各自的限制，本研究之目的在於開發一套適用於 3~12 歲兒童的詞彙理解與表達能力評估工具——「華語兒童圖畫詞彙理解與表達測驗」，透過符合現今臺灣生活文化與情境之詞彙清單，以作答方式簡明且易施測的圖片指認和命名的方式，實際測量個別兒童的詞彙理解和表達能力，並建立具信度和效度的標準化程序和常模。期待此工具未來能運用於大規模樣本的施測，且提供同時

涵蓋理解性詞彙和表達性詞彙能力的評估，擴大其教育及臨床應用價值。

貳、研發歷程

「華語兒童圖畫詞彙理解與表達測驗」係參考英語版「畢保德圖畫詞彙測驗」（第四版）（Peabody Picture Vocabulary Test, 4th ed. [PPVT-4]; Dunn & Dunn, 2007）和「詞彙表達測驗」（Expressive Vocabulary Test, 2nd ed. [EVT-2]; Williams, 2007）的架構和題目，以及國內外相關論文、語料庫和資料庫，包括：中央研究院平衡語料庫、劉惠美與曹峰銘（2010）、余永吉（2006），以及張鑑如（Chang & Hung, 2016; Chang & Luo, 2020; Chang & McCabe, 2013）所蒐集的兒童語料、International Picture Naming Project at CRL-UCSD、Child Language Data Exchange System (CHILDES)、MRC Psycholinguistic Database、English Lexicon Project 等，並依據兒童詞彙發展順序、詞彙詞性、出現頻率，以及該詞彙概念可清晰以圖像呈現等之條件（如圖畫詞彙理解測驗的「郵差」、「推倒」、「傷心的」；圖畫詞彙表達測驗的「吹風機」、「游泳」、「半滿」），初步建立 3~6 歲各年齡的詞彙語單，並進行預試程序進行題目的刪修，以發展出正式版本。另外，適用於評估 6 歲以上至 12 歲兒童之詞彙選擇，除了參考上述國內外相關論文、英語版 PPVT-4 和 EVT-2、國內外華語兒童語料庫外，也參考 PPVT-R（陸莉、劉鴻香修訂，1998）中 6 歲以上使用之測驗題目、「國小學童常用字詞調查報告書」（教育部，2002），以及「八十七年口語統計詞頻表」（教育部，1998），以建立題目進行預試。以下說明預試和正式施測的歷程。

一、預試

（一）研究對象與程序

研究者自 2018 年 1 月起從全臺各縣市（包含臺北市、新北市、桃園市、宜蘭縣、臺南市、臺東縣）的學校和教育機構招募對象，請幼兒園、國小、安親班班級教師協助發放邀請函與同意書給家長，邀請 3 歲半~12 歲兒童參加。進行的測驗項目包括自編之詞彙理解和表達兩個分測驗，並以 PPVT-R（陸莉、劉鴻香修訂，1998）作為本自編測驗的效標。此預試版本之圖畫詞彙

測驗學刊，第 72 輯第 1 期

理解測驗的題數為 131 題，題目詞性有名詞 59 題、形容詞 26 題、動詞 44 題、數量詞 2 題；圖畫詞彙表達測驗的題數為 94 題，題目詞性有名詞 78 題、形容詞 1 題、動詞 10 題、數量詞 5 題。

圖畫詞彙理解測驗的施測方式採圖片 4 選 1 的施測方式，每一個題目會呈現張圖片，並詢問兒童「哪一個是○○」，例如：「哪一個是蝴蝶」。受測者需從 4 張圖片中選擇 1 張符合所聽到詞彙（如「蝴蝶」）的圖片。

圖畫詞彙表達測驗則是透過圖片來詢問受測者，例如：「這是什麼」、「他在做什麼」、「這是什麼形狀」、「他看起來怎麼樣」等，問題之問法會依據圖片所呈現詞彙類別或詞性（如名詞、動詞、形容詞等）來調整詢問，例如：圖片呈現的是「打哈欠」，則會以「他在做什麼」來詢問。

完成圖畫詞彙理解測驗的預試人數為 380 位 3 歲半～12 歲兒童（男 195 人、女 185 人），包括：3 歲半組 9 人、4～6 歲半每組約 19～22 人、7～12 歲每組約 41～43 人。完成圖畫詞彙表達測驗的預試人數則為 381 位 3 歲半～12 歲兒童（男 195 人、女 186 人），包括：3 歲半組 9 人、4～6 歲半每組約 19～22 人、7～12 歲每組約 41～43 人。其中 3 歲半～6 歲半以半歲為一組，7～12 歲以 1 歲為一組。

（二）信效度分析

1. 圖畫詞彙理解測驗

預試版本的內部一致性信度（Cronbach's alpha）為 .96、折半信度為 .93。效標關聯效度檢驗結果顯示，本測驗總分與 PPVT-R（陸莉、劉鴻香修訂，1998）原始分數有顯著相關（ $r = .91$ ， $p < .001$ ）。以 ConQuest 軟體進行試題反應理論（item response theory [IRT]）的 Rasch 模式檢測題目的難度估計值、均方誤（mean square error [MNSQ]），以了解題目的適配度，並參考 Wright 與 Linacre（1994）的試題評斷指標，infit MNSQ 值需介於 0.70～1.30 之間，表示題目適配度良好。在試題中，除了第一題所有兒童皆答對（難度值為 -77.545）無法估算之外，其餘題目難度值介於 -3.431～3.732 之間。全部題目的 infit MNSQ 值介於 0.78～3.63 之間，其中有 120 題是介於 0.70～1.30 之間，有 10 題超出 1.30。

為設計出更具經濟效益的施測題本，減少兒童因受測時間過久所造成的疲勞影響而進行刪題，並將全部試題分為 A、B 兩個平行題本，建有共同

題，後續經由合併分析得到相同量尺，以發展出平行測驗。

首先，從試題難度分布之受試兒童能力值（介於 $-0.337 \sim 6.279$ ）對應至題目難度，得知整體測驗題目偏簡單，因此依照題目難度估計值排序，從列於同一層難度值的題目中至少挑選一題，並考慮題目的詞性平均分配至 A、B 題本，使兩式題本的詞彙詞性分布平均，此歷程共刪除 37 題，保留了 94 題，其中有 14 題共同題目，以作為不同題本間的試題檢驗之用，A、B 兩式各分配 54 題。接著，參考前述之各項詞彙資料庫新增 22 題較難的詞彙，以增加本試題的難度，題數刪修為 A、B 兩式各 76 題進行施測。運用 36 題共同題合併兩式試題作答反應矩陣共 116 題（ $76+76-36$ ）進行同時估計，結果顯示，除了第一題所有兒童皆答對外，其餘題目難度介於 $-6.832 \sim 2.944$ 之間。適配度指標 infit MNSQ 值介於 $0.75 \sim 14.52$ 之間，其中 101 題介於 $0.70 \sim 1.30$ 之間，15 題超出 1.30。依照題目難度排序，考量題目難度（低於 3 歲組幼兒能力值平均值 -2.5 個標準差以下，共 12 題）、 infit MNSQ 值（前述難度過低的 12 題當中有 7 題的 infit MNSQ 值也超出 1.30），共刪除 12 題，保留 104 題，再從保留的題數中分屬為同一層能力值的題目，並且考量題目的詞性平均分配 A、B 兩式題目，最後詞彙理解題數 A、B 兩版各以 52 題作為正式施測版本。

2. 圖畫詞彙表達測驗

預試版本的內部一致性信度（Cronbach's alpha）為 .95、折半信度為 .93。效標關聯效度的檢驗發現，測驗總分與 PPVT-R（陸莉、劉鴻香修訂，1998）原始分數有顯著相關（ $r = .84, p < .001$ ）。IRT 的 Rasch 模式檢測結果顯示，94 題中除第一題所有兒童皆答對（難度值無法估算）、第二題僅 1 位兒童答錯（難度值 -55.169 ）外，其餘題目難度值介於 -4.112 至 5.928 之間。94 題的 infit MNSQ 值介於 0.72 至 1.51 之間，其中 86 題介於 0.70 至 1.30 之間，除第一題及第二題 2 題無法估算，另有 6 題超出 1.30。

採用與理解測驗相同的試題刪選標準，原有 94 題，共刪除 37 題，保留 57 題，其中 13 題為共同題目，A、B 兩式各分配 35 題，再新增 21 題較難的詞彙，圖畫詞彙表達題數刪修為 A、B 兩式各 56 題進行施測，接著運用 34 題共同題合併兩式試題作答反應矩陣合併共 78 題進行分析，結果顯示，題目難度介於 -5.667 至 6.254 之間， infit MNSQ 值介於 0.66 至 1.73 之間，其中 72 題介於 0.70 至 1.30 之間，1 題低於 0.70 ，5 題超出 1.30。再依照題目難度排序，

測驗學刊，第 72 輯第 1 期

考量題目難度（低於 3 歲組幼兒能力值平均值 -2.5 個標準差以下的題目）、*infit MNSQ* 值、題目的詞性、兒童回答狀況等因素。圖畫詞彙表達題目難度最低值該題的難度（-5.667）剛好與於 3 歲組幼兒能力值平均值的-2.5 個標準差數值（-5.559）只差 0.11，因此保留該題。兒童回答時出現多數不確定答案、圖片主題不明的題目（如爬行、燃燒、壓舌板、浮雕、單簧管、鑷子、渦湖、錐子）8 題，此 8 題其中的「燃燒」、「壓舌板」2 題 *infit MNSQ* 值也超出 1.30，總共刪除 8 題。再從列於同一層能力值的題目並考量題目的詞性分配 A、B 兩版題目，最後圖畫詞彙表達題數 A、B 兩版各以 35 題作為正式施測版本。

二、正式施測

（一）測驗題目

正式施測版本分為 A、B 兩式測驗題本，圖畫詞彙理解測驗兩式題數各為 52 題；圖畫詞彙表達測驗兩式題數各為 35 題。

（二）研究對象與程序

為使研究對象較具有代表性，取樣時參考臺灣社會變遷基本調查計畫，依據人口結構變項與人文區位的經濟變項，以「農林漁牧從業人口百分比」、「工業從業人口百分比」、「職業等級_專業及主管人員人口百分比」、「十五至六十四歲人口百分比」、「六十五歲及以上人口百分比」、「大學及以上教育人口百分比」、「人口密度」，以及「5 年人口成長數」等八個變項，將全臺鄉鎮分為七層之各層別鄉鎮列表（傅仰止等人，2015），參照其鄉鎮列表分層招募正式施測樣本，以期能選取涵蓋北、中、南、東四區域之不同社經條件層別的樣本資料。此外，並考量兒童性別比例、兒童是否為特殊學生等條件。

實際完成施測人數共 1,708 人（北、中、南、東四區人數分別為 805、402、395、106 人，比例為 47.13%、23.54%、23.13%及 6.21%；各層別人數分布為第一層 148 人、第二層 599 人、第三層 316 人、第四層 224 人、第五層 310 人、第六層 109 人、第七層 2 人，並盡量讓每個年齡層性別比例為男女生大約各半。3~6 歲半每組人數約 113~148 人、國小每組人數約

103~131 人。以 A、B 兩式測驗分開計算，A 式人數共 881 人，其中學前幼兒共 476 人（女 243 人）、學齡兒童共 405 人（女 215 人）；B 式人數共 827 人，其中學前幼兒共 544 人（女 263 人）、學齡兒童共 283 人（女 137 人）。為後續建立測驗的常模，在受試者取樣上，於發送邀請函時請教師先幫忙排除特殊兒童。在家長教育程度上，父母親以大學／二技最多，國小或以下者為最少（如表 1 所示）。正式施測方式與預試相同，亦以 PPVT-R（陸莉、劉鴻香修訂，1998）作為效標，於 2018 年 11 月起至 2019 年 7 月完成收案。

表 1 正式施測兒童之父母親教育程度次數分配

教育程度	全部 (N = 1708)		A 式 (N = 881)		B 式 (N = 827)	
	父親	母親	父親	母親	父親	母親
國小或以下	6	13	3	8	3	5
國（初）中	97	52	56	36	41	16
高中（職）	410	366	220	205	190	161
專科	259	289	150	146	109	143
大學／二技	546	671	276	349	270	322
碩士或以上	286	214	140	100	146	114
未填	104	103	36	37	68	66

參、正式施測結果

一、信度分析

研究者以 Cronbach's alpha 係數表示每一測驗題目間的內部一致性情形，圖畫詞彙理解測驗 A 式為 .94、B 式為 .92，圖畫詞彙表達測驗 A 式為 .93、B 式為 .92，顯示本研究編製之圖畫詞彙理解測驗和圖畫詞彙表達測驗的題目均有良好的內部一致性。由已知之信度分別計算出測量標準誤，圖畫詞彙理解測驗 A 式為 2.77、B 式為 2.83，圖畫詞彙表達測驗 A 式為 2.09、B 式為 2.01。

檢驗單數與雙數題號間的折半信度，圖畫詞彙理解測驗 A 式為 .88、B 式為 .87，圖畫詞彙表達測驗 A 式為 .89、B 式為 .87，顯示良好的內部一致性。

二、效度分析

（一）同時效標關聯效度

研究者以 PPVT-R（陸莉、劉鴻香修訂，1998）為效標參照，進行同時效度分析，結果顯示圖畫詞彙理解測驗 A、B 式總分與 PPVT-R 原始分數之間呈高度正相關（ $r = .86$ 和 $.74$ ）；圖畫詞彙表達測驗 A、B 式總分與 PPVT-R 原始分數之相關分別為 $.87$ 和 $.53$ ，呈中度到高度的正相關。整體而言，此項自編之圖畫詞彙理解測驗和圖畫詞彙表達測驗的 A、B 兩式均與國內研究常用之 PPVT-R 有合理的效標關聯效度。

（二）區辨效度

研究者針對兒童年齡變項，以單因子變異數分析及 T 檢定檢驗各個版本試題是否有顯著的年齡效果。

1. 圖畫詞彙理解測驗

學齡前各年齡（3~6 歲半）兒童在圖畫詞彙理解測驗 A 式與 B 式總分平均數、標準差及組間差異，如表 2 所示。以年齡為受試者間變項，每半歲為一區間，結果顯示圖畫詞彙理解測驗得分因兒童年齡不同而有所差異（A 式 $F(7, 468) = 55.50, p < .001$ ；B 式 $F(7, 536) = 58.19, p < .001$ ），且隨著年齡增長，得分也隨之增加，再以 Scheffe 法進行事後比較，也證實此一年齡發展趨勢（ $ps < .05$ ）。

學齡階段之不同年級兒童的圖畫詞彙理解測驗 A 式與 B 式總分平均數、標準差及組間差異，如表 3 所示。以年級為受試者間變項，每一年級為一區間，結果顯示圖畫詞彙理解測驗得分因年級不同而有所差異（A 式 $F(5, 399) = 100.15, p < .001$ ；B 式 $F(5, 277) = 69.26, p < .001$ ），且隨著年齡增長，得分也隨之增加。以 Scheffe 法進行事後比較，也大致支持此一年齡發展趨勢（ $ps < .05$ ）。學齡階段 A、B 兩式人數雖有差異，但在圖畫詞彙理解測驗各分數範圍的分布情形大致相同，顯示兩組樣本在研究者所關心的詞彙理解變項上並無明顯差異。

表 2 學齡前各年齡組兒童在圖畫詞彙理解測驗 A、B 兩式的總分平均數、標準差及組間差異

年齡	A 式測驗				B 式測驗			
	N	平均數	標準差	組間差異比較	N	平均數	標準差	組間差異比較
3.0	68	18.03	6.54	4~6.5 > 3	45	18.47	5.18	4~6.5 > 3
3.5	59	19.46	6.33	4.5~6.5 > 3.5	58	21.71	5.82	4.5~6.5 > 3.5
4.0	46	22.70	6.71	5~6.5 > 4; 4 > 3	74	24.93	5.84	5~6.5 > 4; 4 > 3
4.5	52	24.33	5.87	5.5~6.5 > 4.5; 4.5 > 3~3.5	95	27.08	5.34	5.5~6.5 > 4.5; 4.5 > 3~3.5
5.0	53	28.28	6.28	6.5 > 5; 5 > 3~4	72	30.04	5.36	6 > 5; 5 > 3~4
5.5	58	31.17	6.06	5.5 > 3~4.5	76	31.82	5.69	5.5 > 3~4.5
6.0	65	31.20	6.19	6 > 3~4.5	83	33.65	5.63	6 > 3~5
6.5	75	33.71	6.45	6.5 > 3~5	41	32.56	3.68	6.5 > 3~4.5
總計	476	26.36	8.47		544	27.97	7.18	

表 3 學齡階段不同年級兒童在圖畫詞彙理解測驗 A、B 兩式的總分平均數、標準差及組間差異比較

年級	A 式測驗				B 式測驗			
	N	平均數	標準差	組間差異比較	N	平均數	標準差	組間差異比較
一	86	35.65	5.76	2~6 > 1	45	36.02	5.45	2~6 > 1
二	60	39.22	5.25	3~6 > 2; 2 > 1	62	40.79	4.17	3~6 > 2; 2 > 1
三	48	44.25	4.71	5 > 6 > 3; 3 > 1、2	55	44.85	3.79	6 > 3; 3 > 1、2
四	72	45.49	3.71	6 > 4; 4 > 1、2	44	46.02	2.99	6 > 4; 4 > 1、2
五	72	47.00	3.12	5 > 1~3	40	47.08	2.79	5 > 1、2
六	67	48.90	2.49	6 > 1~4	37	49.00	2.08	6 > 1~4
總計	405	43.16	6.50		283	43.60	5.64	6 > 3~5

2. 圖畫詞彙表達測驗

學齡前不同年齡兒童的圖畫詞彙表達測驗 A 式與 B 式總分平均數、標準差及組間差異，如表 4 所示。以年齡為受試者間變項，每半歲為一區間，結果顯示圖畫詞彙表達測驗的得分因兒童年齡不同而有所差異（A 式 $F(7, 468) = 47.63, p < .001$ ；B 式 $F(7, 536) = 54.64, p < .001$ ），3~6 歲半間隨著年齡增長，得分也隨之增加，以 Scheffe 法進行事後比較，也大致支持此一年齡發展趨勢（ $ps < .05$ ）。

測驗學刊，第 72 輯第 1 期

表 4 學齡前不同年齡兒童在圖畫詞彙表達測驗 A、B 兩式的總分平均數、標準差及組間差異

年齡	A 式測驗				B 式測驗			
	N	平均數	標準差	組間差異比較	N	平均數	標準差	組間差異比較
3.0	68	6.09	3.26	4.5~6.5 > 3	45	5.02	2.77	4~6.5 > 3
3.5	59	6.19	3.32	5~6.5 > 3.5	58	5.86	3.05	4.5~6.5 > 3.5
4.0	46	8.22	4.13	5~6.5 > 4	74	7.58	3.39	5~6.5 > 4; 4 > 3
4.5	52	8.98	3.92	5.5~6.5 > 4.5; 4.5 > 3	95	9.67	3.21	5.5~6.5 > 4.5; 4.5 > 3、3.5
5.0	53	11.87	4.55	6.5 > 5; 5 > 3~4	72	11.71	3.20	6 > 5; 5 > 3~4
5.5	58	12.97	4.93	5.5 > 3~4.5	76	12.37	4.90	5.5 > 3~4.5
6.0	65	14.65	4.77	6 > 3~4.5	83	14.07	3.83	6 > 3~5
6.5	75	14.75	3.96	6.5 > 3~5	41	13.02	3.52	6.5 > 3~4.5
總計	476	10.64	5.37		544	10.17	4.67	

學齡階段不同年級兒童的圖畫詞彙表達測驗 A 式與 B 式總分平均數、標準差及組間差異，如表 5 所示。以年級為受試者間變項，每一年級為一區間，結果顯示圖畫詞彙表達測驗的得分因年級不同而有所差異（A 式 $F(5, 399) = 83.38, p < .001$ ；B 式 $F(5, 277) = 67.58, p < .001$ ），且隨著年齡增長，得分也隨之增加，以 Scheffe 法進行事後比較，也證實此一年齡發展趨勢（ $ps < .05$ ）。如同詞彙理解分數，學齡階段 A、B 兩式人數雖有差異，但在詞彙表達各分數範圍的分布情形大致相同，顯示兩組樣本在研究者所關心的詞彙表達變項上也無明顯差異。

為建立較為精確的常模，本研究學齡前以半歲為區間分組。然而，因部分學齡前年齡組別之間的詞彙理解和詞彙表達分數未呈現顯著的年齡效果，故嘗試將學齡前幼兒每半歲一組合併為每一歲一組，也就是 3 歲與 3 歲半合併、4 歲與 4 歲半合併、5 歲與 5 歲半合併、6 歲與 6 歲半合併。學齡前各年齡（3~6 歲半）兒童在圖畫詞彙理解測驗 A 式與 B 式總分平均數、標準差及組間差異，如表 6 所示。以年齡為受試者間變項，每一歲為一區間，結果顯示圖畫詞彙理解測驗得分因兒童年齡不同而有所差異（A 式 $F(3, 472) = 121.93, p < .001$ ；B 式 $F(3, 540) = 125.11, p < .001$ ），且隨著年齡增長，得分也隨之增加，再以 Scheffe 法進行事後比較，也證實此一年齡發展趨勢（ $ps < .05$ ）。

表 5 學齡階段不同年級兒童的圖畫詞彙表達測驗 A、B 兩式的總分平均數、標準差及組間差異

年級	A 式測驗				B 式測驗			
	N	平均數	標準差	組間差異比較	N	平均數	標準差	組間差異比較
一	86	16.24	4.58	2 > 1	45	14.96	3.81	2 > 1
二	60	19.28	4.22	3 > 2; 2 > 1	62	18.37	3.26	3 > 2; 2 > 1
三	48	21.92	3.97	5、6 > 3; 3 > 1、2	55	21.85	3.64	5、6 > 3; 3 > 1、2
四	72	23.50	3.69	5、6 > 4; 4 > 1、2	44	22.57	3.48	6 > 4; 4 > 1、2
五	72	26.49	3.53	5 > 1	40	25.03	3.48	5 > 1
六	67	26.81	3.33	6 > 1	37	26.81	3.20	6 > 1
總計	405	22.22	5.58		283	21.20	5.15	6 > 3~5

表 6 學齡前各年齡組兒童在圖畫詞彙理解測驗 A、B 兩式的總分平均數、標準差及組間差異（每一歲為一區間）

年齡	A 式測驗				B 式測驗			
	N	平均數	標準差	組間差異比較	N	平均數	標準差	組間差異比較
3	127	18.69	6.46	4~6 > 3	103	20.29	5.75	4~6 > 3
4	98	23.56	6.30	5~6 > 4	169	26.14	5.65	5~6 > 4
5	111	29.79	6.31	6 > 5	148	30.95	5.59	6 > 5
6	140	32.54	6.43	6 > 3~5	124	33.29	5.08	6 > 3~5
總計	476	26.36	8.47		544	27.97	7.18	6 > 5; 5 > 3~4

學齡前不同年齡兒童的圖畫詞彙表達測驗 A 式與 B 式總分平均數、標準差及組間差異，如表 7 所示。以年齡為受試者間變項，每一歲為一區間，結果顯示圖畫詞彙表達測驗的得分因兒童年齡不同而有所差異（A 式 $F(3, 472) = 110.47, p < .001$ ；B 式 $F(3, 540) = 117.82, p < .001$ ），3~6 歲半間隨著年齡增長，得分也隨之增加，以 Scheffe 法進行事後比較，也大致支持此一年齡發展趨勢（ $ps < .05$ ）。

測驗學刊，第 72 輯第 1 期

表 7 學齡前不同年齡兒童在圖畫詞彙表達測驗 A、B 兩式的總分平均數、標準差及組間差異（每一歲為一區間）

年齡	A 式測驗				B 式測驗			
	N	平均數	標準差	組間差異比較	N	平均數	標準差	組間差異比較
3	127	6.13	3.28	4~6>3	103	5.50	2.94	4~6>3
4	98	8.62	4.02	5~6>4	169	8.76	3.44	5~6>4
5	111	12.44	4.77	6>5	148	12.05	4.16	6>5
6	140	14.70	4.34	6>3~5	124	13.73	3.75	6>3~5
總計	476	10.64	5.37		544	10.17	4.67	6>5; 5>3~4

三、試題反應理論（IRT）分析結果

（一）圖畫詞彙理解測驗

研究者於刪題後進行 IRT 分析，其效度概念包含資料和模式適配度，圖畫詞彙理解測驗 A 式題本的題目難度介於 -2.122~2.270 之間，而適配度指標 infit MNSQ 值介於 0.75~1.50 之間，其中有 3 題的 infit MNSQ 值超出 1.30，其餘 49 題均介於 0.70~1.30，受試者能力與試題難度之配對情形如圖 1 所示，整體題目適配度大致良好。

圖畫詞彙理解測驗 B 式題本之題目難度介於 -2.273~2.410 之間。適配度指標 infit MNSQ 值介於 0.85~1.30 之間，受試者能力與試題難度之配對情形如圖 2 所示，所有題目的模式符合度大致良好。

（二）圖畫詞彙表達測驗

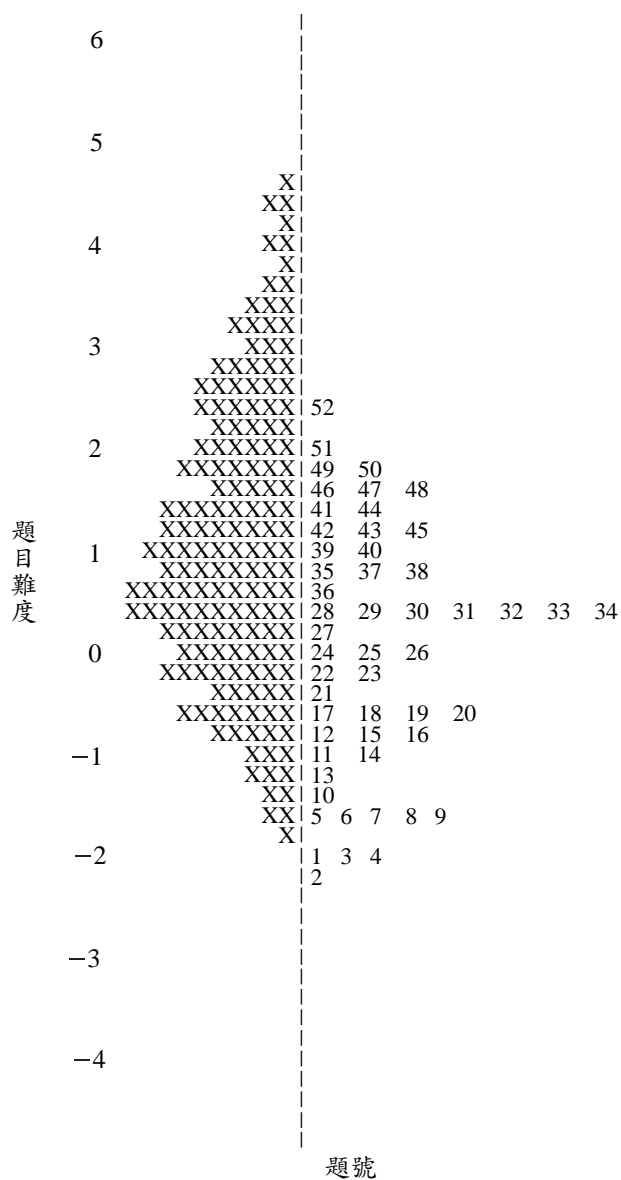
研究者於刪題後進行 IRT 分析，其效度概念包含資料和模式適配度，圖畫詞彙表達測驗 A 式題本的題目難度介於 -4.283~6.893 之間，適配度指標 infit MNSQ 值介於 0.74~1.19 之間，受試者能力與試題難度之配對情形如圖 3 所示，所有題目的模式符合度良好。

圖畫詞彙表達測驗 B 式題本的題目難度介於 -5.976~5.956 之間，適配度指標 infit MNSQ 值介於 0.74~1.38 之間，其受試者能力與試題難度之配對情形如圖 4 所示。

劉惠美等

「華語兒童圖畫詞彙理解與表達測驗」之編製與應用

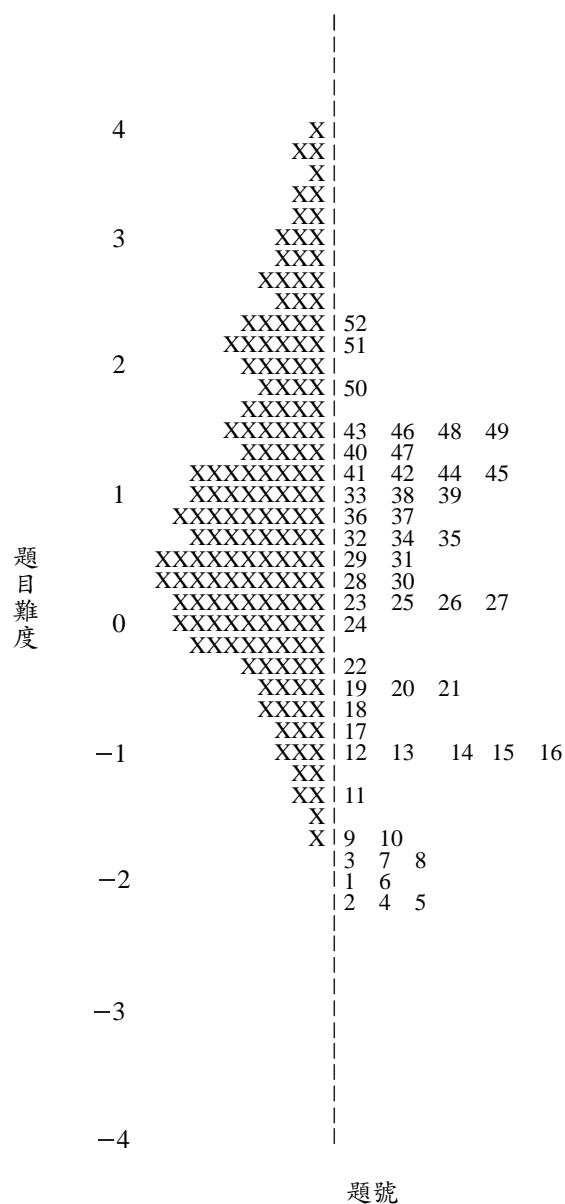
圖 1 圖畫詞彙理解測驗 A 式受試者能力與試題難度 Wright Map



註：每個 X 代表 5.3 位受試者。

測驗學刊，第 72 輯第 1 期

圖 2 圖畫詞彙理解測驗 B 式受試者能力與試題難度 Wright Map

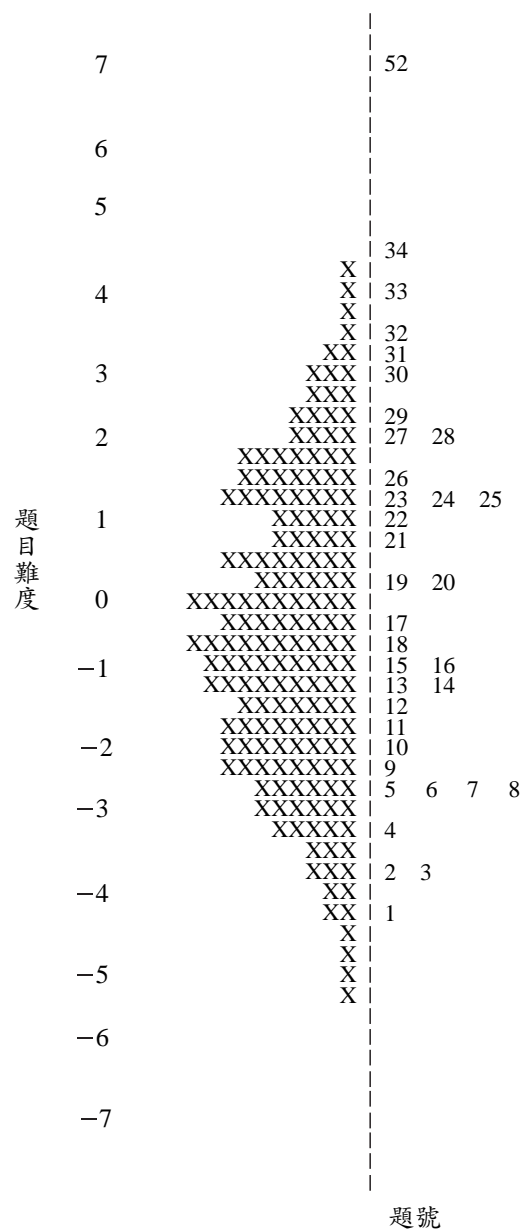


註：每個 X 代表 5.0 位受試者。

劉惠美等

「華語兒童圖畫詞彙理解與表達測驗」之編製與應用

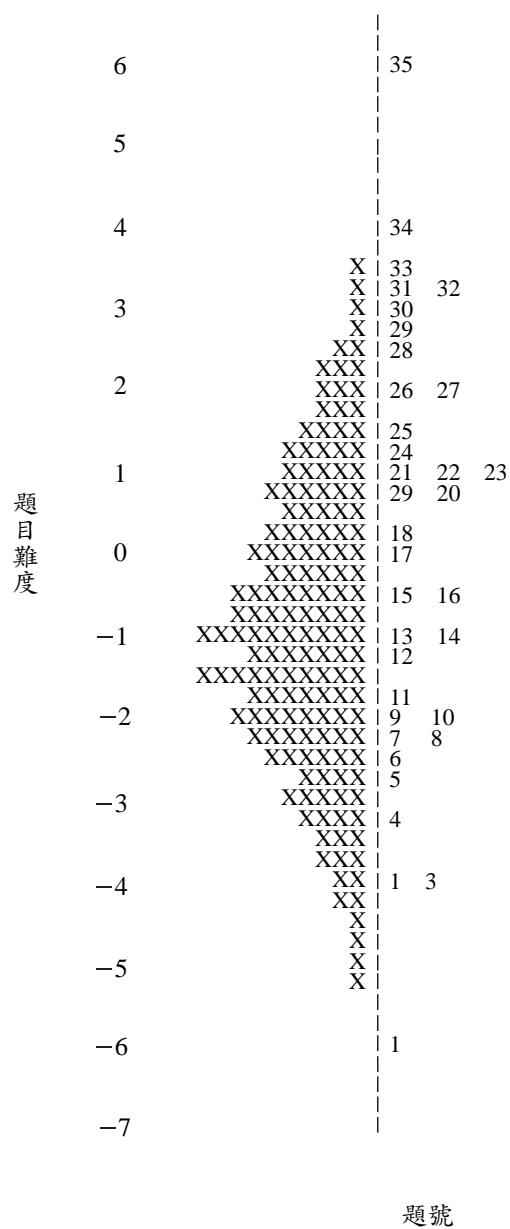
圖 3 圖畫詞彙表達測驗 A 式受試者能力與試題難度 Wright Map



註：每個 X 代表 5.1 位受試者。

測驗學刊，第 72 輯第 1 期

圖 4 圖畫詞彙表達測驗 B 式受試者能力與試題難度 Wright Map



註：每個 X 代表 5.2 位受試者。

四、測驗題目之內容分析結果

圖畫詞彙理解測驗 A、B 兩式題數各為 52 題，A 式題本詞彙之詞性為名詞 24 題、動詞 12 題、形容詞 16 題，類別共 16 類；B 式測驗題本詞彙包括名詞 21 題、動詞 22 題、形容詞 9 題，涵蓋 16 類語意類別（如表 8 所示）。

圖畫詞彙表達測驗 A、B 兩式題數各為 35 題：A 式測驗題本詞彙包括名詞 30 題、動詞 2 題、形容詞 3 題，涵蓋 15 類語意類別；B 式測驗題本詞彙包括名詞 26 題、動詞 5 題、形容詞 4 題，涵蓋 14 類語意類別（如表 9 所示）。

表 8 圖畫詞彙理解測驗的詞彙類別和題數

詞彙語意類別	A 式題數	B 式題數
日用品	3	2
休閒相關（玩具、樂器、運動用品）	1	2
工具、機器、儀器與其相關各部分	1	1
交通相關（交通工具、設施）	1	1
身體部位	1	1
與人有關的術語	1	1
動物	3	2
植物及其相關部位	4	2
氣候、室外景物及地理名詞	2	2
建築物	0	1
數學名詞	2	3
數量	0	2
狀態修飾	17	8
動作	11	21
食物	2	0
衣著	1	0
職業與角色	1	0
抽象概念	1	2
事件	0	1

測驗學刊，第 72 輯第 1 期

表 9 圖畫詞彙表達測驗詞彙類別和題數

詞彙類別	A 式題數	B 式題數
日用品	3	3
休閒相關（玩具、樂器、運動用品）	1	5
工具、機器、儀器與其相關各部分	10	2
交通相關（交通工具、設施）	1	1
身體部位	1	2
動物	2	2
植物及其相關部位	3	2
氣候、室外景物及地理名詞	0	1
建築物	0	1
場所	2	0
數學名詞	2	5
數量	2	3
狀態修飾	1	2
動作	2	4
衣著	3	0
職業與角色	1	2
節日	1	0

五、常模建立

以圖畫詞彙理解測驗的總分平均數來看，測驗總分會因兒童年齡不同而有所差異，並考量測驗所需時間及測驗難度，設定不同年齡層兒童適合的題目。由兒童的表現（能力值）去對應圖畫詞彙理解測驗的題目難度，從符合各年齡組之能力值平均數的 -2.5 個標準差至 2 個標準差難度的題目進行切割，3~5 歲半以每半歲為一組，6 歲之後以每一歲為一組，每年齡組兒童施測題目數為 25~38 題不等。以 EXCEL 提供的 PERCENTRANK 函數將各年齡兒童的個別總分轉換成占全體成績的百分比，並建立由總分對應之百分等級常模資料，以作為個別兒童在同年齡兒童常模對照下的詞彙理解能力表現。

圖畫詞彙表達測驗也以相同的方式進行各年齡層的試題切割，3~12 歲以每一歲為一組，適合兒童的題目為 12~21 題不等，因題目數量不多，因此各年齡組都施測全部共 35 題的題目。後續並以 PERCENTRANK 函數將各年齡兒童的個別總分轉換成占全體成績的百分比，並建立由總分對應之百分等

級常模資料，以作為個別兒童在同年齡兒童常模對照下的詞彙表達能力表現。

肆、討論與建議

兒童口語詞彙能力是溝通、學習的基礎，與其後來的閱讀表現與學業成就密切相關，是相當重要的基礎語文能力。有鑑於國內現有可用以評估學前和學齡兒童口語詞彙測驗工具之不足，本研究開發可適用於 3~12 歲、符合現今臺灣生活、且便於施測的兒童詞彙理解與表達的標準化評估工具。歷經預試、大規模正式施測和題目增刪，抽樣全臺灣北、中、南、東各區共 1,708 名不同家庭社經背景的學前 3~6 歲及國小一~六年級兒童參與正式施測，經過內部一致性、折半信度、同時效標關聯效度、區辨效度分析，結果顯示圖畫詞彙理解測驗和圖畫詞彙表達測驗之 A、B 兩式都具有良好的信效度，可做為平行複本。本工具之常模資料是以 PPVT-R（陸莉、劉鴻香修訂，1998）作為效標參照，同時效度分析結果顯示，圖畫詞彙理解測驗 A、B 式的總分與 PPVT-R 原始分數之相關達高度正相關，圖畫詞彙表達測驗 A、B 式的總分與 PPVT-R 原始分數之相關亦達中度以上的正相關，顯示本項自編圖畫詞彙理解測驗和圖畫詞彙表達測驗具有良好的效標關聯效度，可以用以測量兒童的詞彙理解與表達能力。此外，透過 IRT 分析，也證實本測驗題目適配度佳，顯示未來可以本測驗所測得之 IRT 能力值去了解個別兒童的詞彙理解或詞彙表達的成長情形。

本研究所編製的「華語兒童圖畫詞彙理解與表達測驗」適用年齡跨學前 3 歲~國小六年級兒童，測驗內容的詞彙類別涵蓋日常用品、休閒相關用品、工具、交通工具、身體部位、動物、植物、氣候、數學用語、動作、修飾語等，包含了不同詞性的詞彙，例如：名詞、動詞、形容詞等。正式版本之題項（見表 6 和表 7）涵蓋了過去文獻指出兒童詞彙的主要詞性類別（即名詞、動詞、形容詞等）之實詞，其中不論是理解或表達測驗，大致都以名詞的數量占比最高，動詞次之，符合過去兒童語言發展的趨勢，頗能符應兒童詞彙發展的趨勢（涂嘉新、陳姿蘭，2020；劉惠美、陳昱君，2015；Ham-an et al., 2017）。為了避免每套測驗的題目太多，本研究透過預試程序的試題刪修，保留有較佳心理計量與代表性的詞彙選項，依據題目難易度與適配度指標（MNSQ）刪除過於簡單的（幾乎所有年齡層兒童都會的）及適配度

不佳的題目，並依各難度分層去保留有不同詞性的詞彙。在表達題目選擇再加上考量兒童能清楚理解並具體看圖命名的詞彙，以更能量測兒童詞彙能力的發展。研究者將理解和表達測驗又各自分為 A、B 兩式，圖畫詞彙理解測驗每式各 52 題，圖畫詞彙表達測驗每式各 35 題，此一平行複本之設計，提供了臨床使用的便利性，例如：針對同一群兒童所實施的前測和後測之語言相關教學介入設計，即能運用此一工具進行前後測的測量與比較，以了解介入成效。此外，本測驗之特色之一，即適用之兒童年齡範圍大，包括學前 3~6 歲及國小一~六年級兒童，因此所設計之詞彙題目必須涵蓋不同難易程度的詞彙，採 A、B 兩式測驗之方式，分別設定適合不同年齡兒童施測的題組，也就是有不同的測試起點題目，可提供不同年齡兒童較為適合的施測起點，且可避免從頭施測起，而造成施測時間過長及較年幼受測者因題目難度所導致的作答壓力。此外，從各年齡常模樣本在各式測驗的平均得分大致呈現隨年齡增加而進步的趨勢，顯示本工具可以有效區分跨學前到學齡階段不同年齡兒童的詞彙理解和詞彙表達能力。然而，可能受限於每個版本的題數頗為精簡之緣故，部分年齡組別之間的詞彙理解或表達分數未呈現顯著的年齡效果，但應仍屬可以接受的範圍。

在施測的方式上，「華語兒童圖畫詞彙理解與表達測驗」適用於平板或單機版，所有的詞彙理解和詞彙表達題目皆以圖畫的方式呈現，詞彙理解由受測者在四格圖畫中指認出正確答案，詞彙表達由受測者針對所提供的圖畫進行命名，兩者的施測方法均屬簡便，易於學前及學齡兒童作答。另一方面，電腦化施測是當代測驗的趨勢，尤其在近三年新冠肺炎疫情嚴峻之期，更凸顯電腦施測的重要性。此測驗以電腦化施測方式，設定標準化的施測流程，配合受測者年齡，呈現適合其年齡施測的詞彙題組。施測成績之計算採自動化方式，在施測結束後，可立即計分並產出個人之施測結果報告。再根據正式施測結果所建立之常模對照表，可進一步篩檢出詞彙理解或是詞彙表達能力落後於同儕的兒童，也可用來觀察同一群兒童在不同年齡階段之詞彙理解和表達表現的進程，提供研究人員及教育臨床工作人員一項可以有效評估華語兒童詞彙能力的工具，有其臨床應用價值。本研究圖畫詞彙理解測驗已開發離線版電腦化施測，計分原則與紙本測驗相同，電腦程式會針對兒童作答結果自動計算成績和產出成績報表（張鑑如，2024）。考量學前幼兒對電腦操作較不熟悉，建議以個別施測方式進行較為合適；學齡兒童則可使用團體施測進行。

整體而言，本項「華語兒童圖畫詞彙理解與表達測驗」是華語兒童詞彙評估的一大進展，提供 PPVT-R（陸莉、劉鴻香修訂，1998）之外，另一個具有信效度且適合評估現今兒童的圖畫詞彙理解測驗。另一方面，此測驗也超越現有臺灣華語詞彙表達工具僅能評估學前 3~6 歲兒童的限制，例如：「華語兒童理解與表達詞彙測驗」（REVT）（黃瑞珍等人，2010）。本測驗的實用價值包括：題目涵蓋不同詞性及語意詞彙，以簡明且易施測的圖片指認和命名的方式，能便捷、快速和有效的施測；可以施測的兒童年齡層廣（3~12 歲）；同時可以測量詞彙理解與表達能力的發展情形。然而，本測驗仍存在一些使用限制，例如：本研究僅以 PPVT-R（陸莉、劉鴻香修訂，1998）為效標參照測驗，因受限於施測時間，並未同時以國內其他常用之兒童語言發展測驗作為效標參照，例如：「華語嬰幼兒溝通發展量表」（劉惠美、曹峰銘，2010）或「修訂學前兒童語言障礙評量表」（林寶貴等人，2008）等，因此效標效度的指標較為有限，為本研究的限制之一。另一方面，測驗所涵蓋的詞彙類別頗多，但題目還是很有限，無法得知受測兒童是否可以理解和說出非測驗上所測試的詞彙，需輔以其他口語詞彙評估方式，例如：透過語言樣本分析，方能了解兒童可表達出詞彙的總量和多樣性（洪宜芳、張鑑如，2019；Bernstein Ratner & MacWhinney, 2016; Hung & Chang, 2023; Spaulding et al., 2012）。此外，詞彙的理解能力包括廣度和深度，透過本測驗亦無法得知受測兒童對特定詞彙理解的深度，需輔以其他測驗，例如：「下定義測驗」（Binder et al., 2017; Ordonez et al., 2002），方可補強此部分之不足。

謝誌

本研究感謝國家科學及技術委員會「幼兒發展資料庫先期研究計畫」（NSC 100-2410-H-003-058-MY2）、「幼兒發展調查資料庫建置計畫」（MOST 103-2420-H-003-032-MY3）、「幼兒發展調查資料庫建置計畫第二期」（MOST 106-2420-H-003-014-SS3）、「幼兒發展調查資料庫建置計畫第三期」（MOST 109-2740-H-003-001-SS3），以及「幼兒發展調查資料庫建置計畫第四期」（NSTC 112-2740-H-003-001-SS5）提供經費補助與支持。同時，衷心感謝計畫研究團隊、工作人員，以及參與本研究的家長、兒童和教師無私的協助與付出。本文承蒙審查委員提供寶貴建議和細心指正，謹致誠摯謝意。

參考文獻

中文部分

- 余永吉（2006）。臺灣學齡前兒童口語詞彙資料庫之發展〔未出版之博士論文〕。國立成功大學。
- [Sher, Y. J. (2006). *Development of Lexicon Database from Speech Corpus for Taiwanese Pre-school Children* [Unpublished Doctoral Dissertation]. National Cheng Kung University.]
- 林寶貴、黃玉枝、黃桂君、宣崇慧（2008）。修訂學前兒童語言障礙評量表。國立臺灣師範大學特殊教育中心。
- [Lin, B. G., Huang, Y. C., Huang, K. C., & Xuan, C. H. (2008). *Preschool Language Disorder Scale Revised*. Special Education Center, National Taiwan Normal University.]
- 洪宜芳、張鑑如（2019）。三～五歲幼兒無字圖畫書敘述之縱貫性研究：基本語言指標。《華語文教學研究》，16（4），99-128。https://doi.org/10.6393/JCLT.201912_16(4).0004
- [Hung, Y. F., & Chang, C. J. (2019). A longitudinal study of non-word picture book narrations in three- to five-year-olds: Basic language measures. *Journal of Chinese Language Teaching*, 16(4), 99-128. https://doi.org/10.6393/JCLT.201912_16(4).0004]
- 涂嘉新、陳姿蘭（2020）。三至六歲幼兒故事敘說之詞類分析研究。《台灣教育研究期刊》，1（4），39-63。
- [Tu, C. H., & Chen, T. L. (2020). A study on three to six year-olds' use of part of speech in narrating story. *Journal of Taiwan Education Studies*, 1(4), 39-63.]
- 張鑑如（2024）。幼兒發展調查資料庫建置計畫第三期。行政院國家科學及技術委員會補助專案研究計畫成果報告。計畫編號 MOST109-2740-H-003-001-SS3。
- [Chang, C. J. (2024). *Kids in Taiwan: National Longitudinal Study of Child Development and Care, Phase III*. National Science and Technology Council-Funded Research Project Final Report (MOST109-2740-H-003-001-SS3).]
- 教育部（2002）。國小學童常用字詞調查報告書【原始數據】。http://language.moe.gov.tw/001/Upload/files/SITE_CONTENT/M0001/PRIMARY/shindex.htm?open

劉惠美等

「華語兒童圖畫詞彙理解與表達測驗」之編製與應用

[Ministry of Education. (2002). *The report of frequently used words and vocabulary in elementary school*. http://language.moe.gov.tw/001/Upload/files/SITE_CONTENT/M0001/PRIMARY/shindex.htm?open]

教育部（1998）。八十七年口語統計詞頻表【原始數據】。 https://language.moe.gov.tw/001/Upload/files/SITE_CONTENT/M0001/87ORAL/ch2-2.htm?open

[Ministry of Education. (1998). *1998 Spoken Vocabulary Frequency Table*. https://language.moe.gov.tw/001/Upload/files/SITE_CONTENT/M0001/87ORAL/ch2-2.htm?open]

陸莉、劉鴻香（修訂）（1998）。修訂畢保德圖畫詞彙測驗（PPVT-R）。心理。

[Lu, L., & Liu, H. S. (Eds.) (1994). *Peabody Picture Vocabulary Test-Revised*. Psychological Publishing Co., Ltd.]

傅仰止、章英華、杜素豪、廖培珊（2015）。臺灣社會變遷基本調查計畫第六期第五次調查計畫執行報告（C00310r）。中央研究院人文社會科學研究中心調查研究專題中心學術調查研究資料庫。 <https://doi.org/10.6141/TW-SRDA-C00310r>

[Fu, Y. C., Chang, Y. H., Tu, S. H., & Liao, P. S. (2015). *Taiwan Social Change Survey (Round 6, Year 5) (C00310r)* [Data file]. Survey Research Data Archive, Academia Sinica. <https://doi.org/10.6141/TW-SRDA-C00310r>]

黃瑞珍、簡欣瑜、朱麗璇、盧璐（2010）。華語兒童理解與表達詞彙測驗（REVT）。心理。

[Huang, R. J., Jian, S. Y., Jhu, L. S., & Lu, L. (2010). *Receptive and Expressive Vocabulary Test (REVT)*. Psychological Publishing Co., Ltd.]

劉惠美、曹峰銘（2010）。華語嬰幼兒溝通發展量表（MCDI-T）。心理。

[Liu, H. M., & Tsao, F. M. (2010). *Mandarin-Chinese Communicative Development Inventory (Taiwan) (MCDI-T)*. Psychological Publishing Co., Ltd.]

劉惠美、陳昱君（2015）。華語嬰幼兒表達性詞彙的語意內容及詞類組成之發展。 *教育心理學報*，47（2），217-242。 <https://doi.org/10.6251/BEP.20150205>

[Liu, H. M., & Chen, Y. C. (2015). Developmental changes in the content and composition of early expressive vocabulary in Mandarin-speaking infants and toddlers. *Bulletin of Educational Psychology*, 47(2), 217-242. <https://doi.org/10.6251/BEP.20150205>]

英文部分

Bernstein Ratner, N., & MacWhinney, B. (2016). Your laptop to the rescue: Using the Child

- Language Data Exchange System Archive and CLAN Utilities to improve child language sample analysis. *Seminars in Speech & Language*, 37(2), 74-84. <https://doi.org/10.1055/s-0036-1580742>
- Binder, K. S., Cote, N. G., Lee, C., Bessette, E., & Vu, H. (2017). Beyond breadth: The contributions of vocabulary depth to reading comprehension among skilled readers. *Journal of Research in Reading*, 40(3), 333-343. <https://doi.org/10.1111/1467-9817.12069>
- Bleses, D., Makranksy, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, 37(6), 1461-1476. <https://doi.org/10.1017/S0142716416000060>
- Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child Language Teaching and Therapy*, 18(1), 1-21. <https://doi.org/10.1191/0265659002ct224oa>
- Chang, C., & Huang, C. (2016). Mother-child talk during joint book reading in two social classes in Taiwan: Interaction strategies and information types. *Applied Psycholinguistics*, 37(2), 387-410. <https://doi.org/10.1017/S0142716415000041>
- Chang, C., & Luo, Y. (2020). A longitudinal study of maternal interaction strategies during joint book reading in Taiwan. *Journal of Child Language*, 47(2), 401-417.
- Chang, C., & McCabe, A. (2013). Evaluation in Mandarin Chinese children's personal narratives. In A. McCabe, & C. Chang (Eds.), *Chinese language narration: Culture, cognition, and emotion* (pp. 33-56). John Benjamins.
- Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, 22(3), 497-529. <http://doi.org/10.1017/S0305000900009934>
- Dockrell, J. E. (2001). Assessing language skills in preschool children. *Child Psychology & Psychiatry Review*, 6(2), 74-85. <http://doi.org/10.1017/S1360641701002532>
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Pearson.
- Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised*. American Guidance Service.
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: Users' guide and technical manual*. Paul H. Brookes.
- Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-Million-Word gap has serious consequences. *Child Development*, 90(3), 985-992. <http://doi.org/10.1111/cdev.13128>
- Haman, E., Łuniewska, M., Hansen, P., Simonsen, H. G., Chiat, S., Bjekić, J., Blažienė, A.,

- Chyl, K., Dabašinskienė, I., de Abreu, P. E., Gagarina, N., Gavarró, A., Håkansson, G., Harel, E., Holm, E., Kapalková, S., Kunnari, S., Levorato, C., Lindgren, J., Mieszowska, K. ... Armon-Lotem, S. ((2017). Noun and verb knowledge in monolingual preschool children across 17 languages: Data from Cross-linguistic Lexical Tasks (LITMUS-CLT). *Clinical Linguistics & Phonetics*, 31(11-12), 818-843. <https://doi.org/10.1080/02699206.2017.1308553>
- Hoff, E. (2015). Language development. In M. H. Bornstein, & M. E. Lamb (Eds.), *Developmental science: An advanced textbook* (7th ed., pp. 443-488). Psychology Press.
- Hung, Y. F., & Chang, C. J. (2023). Developing early lexical composition in Mandarin-speaking children: A longitudinal study. *Journal of Child Language*. <https://doi.org/10.1017/S0305000922000654>
- MacWhinney, B. (2015). Language development. In L. S. Liben, & U. Müller (Eds.), *Handbook of child psychology and developmental science* (Vol. 2, Cognitive Process, pp. 296-338). John Wiley & Sons.
- National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. <https://www.nichd.nih.gov/publications/pubs/nrp/smallbook>
- NICHD Early Child Care Research Network. (2005). *Child care and child development: Results from the NICHD study of early child care and youth development*. Guilford Press.
- Ogura, T., Dale, P. S., Yamashita, Y., Murase, T., & Mahieu, A. (2006). The use of nouns and verbs by Japanese children and their caregivers in book-reading and toy-playing contexts. *Journal of Child Language*, 33(1), 1-29. <https://doi.org/10.1017/s0305000905007270>
- Ordonez, C. L., Carlo, M. S., Snow, C. E., & McLaughlin, B. (2002). Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? *Journal of Educational Psychology*, 94, 719-728. <https://doi.org/10.1037/0022-0663.94.4.719>
- Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading & Writing*, 23(2), 189-208. <https://doi.org/10.1007/s11145-008-9159-1>
- Owens, R. E. (2007). *Language development: An introduction* (7th ed.). Allyn & Bacon.
- Pérez-Pereira, M., & Cruz, R. (2018). A longitudinal study of vocabulary size and composition in low risk preterm children. *First Language*, 38(1), 72-94. <https://doi.org/10.1177/0142723717730484>

- Protopapas, A., Mouzaki, A., Sideridis, G., Kotsolakou, A., & Simos, P. (2013). The role of vocabulary in the context of the simple view of reading. *Reading & Writing Quarterly*, 29(2), 168-202. <https://doi.org/10.1080/10573569.2013.758569>
- Rowe, M. L., Raudenbush, S., & Goldin-Meadow, S. (2012). The pace vocabulary growth helps predict later vocabulary skill. *Child Development*, 83(2), 508-525. <https://doi.org/10.1111/j.1467-8624.2011.01710.x>
- Sénéchal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D. K. Dickinson, & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp.173-182). Guilford Press.
- Spaulding, T. J., Szulga, M. S., & Figueroa, C. (2012). Using norm-referenced tests to determine severity of language impairment in children: Disconnect between US policy makers and test developers. *Language, Speech, and Hearing Services in Schools*, 43(2), 176-190. [https://doi.org/10.1044/0161-1461\(2011/10-0103\)](https://doi.org/10.1044/0161-1461(2011/10-0103))
- Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speaker's early vocabularies. *Developmental Psychology*, 32(3), 492-504. <https://doi.org/10.1037/0012-1649.32.3.492>
- Tardif, T., Fletcher, P., Liang, W., & Kaciroti, N. (2009). Early vocabulary development in Mandarin (Putonghua) and Cantonese. *Journal of Child Language*, 36(5), 1115-1144. <https://doi.org/10.1017/S0305000908009185>
- Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the “noun bias” in context: A comparison of English and Mandarin. *Child Development*, 70(3), 620-635. <https://doi.org/10.1111/1467-8624.00045>
- Williams, K. T. (2007). *Expressive Vocabulary Test* (2nd ed.). Pearson.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.