

# 兒童華語文能力測驗技術報告—2024（3）

---

## 聽力、閱讀測驗信效度

# 目錄

壹、前言.....	4
貳、測驗介紹.....	5
一、測驗說明.....	5
(一) 能力描述.....	5
(二) 測驗題型.....	6
(三) 通過門檻.....	7
二、等級對照.....	8
三、測驗標準化流程.....	10
參、測驗效能分析.....	11
一、信度分析.....	11
(一) 內部一致性信度.....	11
(二) 測驗訊息量.....	12
二、建構效度分析.....	14
肆、結論.....	18
伍、文獻.....	19
陸、附件.....	20
附件 1 兒童測驗正式考試測驗訊息量及測量誤差.....	20
附件 2 兒童測驗正式考試驗證性因素分析結果.....	25

## 表目錄

表 1 兒童測驗基本能力描述 .....	6
表 2 兒童聽力測驗題型 .....	7
表 3 兒童閱讀測驗題型 .....	7
表 4 兒童測驗通過分數 .....	8
表 5 測驗標準化流程說明 .....	10
表 6 兒童測驗各等級聽力、閱讀測驗庫李信度係數 .....	12
表 7 兒童測驗各等級門檻測驗訊息量、測量誤差及信度 .....	14
表 8 兒童測驗試題適配分布 .....	15
表 9 兒童測驗整體模式適配度指標摘要表 .....	17

## 圖目錄

圖 1 聽力與閱讀測驗等級對照圖 .....	9
圖 2 測驗標準化流程 .....	10
圖 3 成長級兒童聽力測驗測驗訊息量 .....	13
圖 4 成長級兒童聽力測驗測量誤差 .....	13
圖 5 萌芽級聽力測驗單因素模式 .....	16
圖 6 萌芽級閱讀測驗單因素模式 .....	16

# 壹、前言

「華語文能力測驗」為一套專為母語非華語者所研發的標準化語言能力測驗，旨在測知華語學習者在實際日常生活中的語言使用能力，故不以任何特定教材為命題依據。華語文能力測驗的測驗類別包括華語文聽力測驗、華語文閱讀測驗、華語文口語測驗、華語文寫作測驗以及兒童華語文能力測驗，測驗內容主要針對各種日常生活情境所設計，題材真實多元，提供語言學習者能夠衡量其語言能力的國際評量工具。

「兒童華語文能力測驗」(Children's Chinese Competency Certification，以下簡稱兒童測驗)是一套專為七至十二歲母語非華語之兒童學習者所設計的標準化語言能力測驗，由「國家華語測驗推動工作委員會」(以下稱為本會)專責研發。兒童測驗旨在幫助兒童學習者了解自己的華語程度、激勵他們學習的熱忱，故不以任何特定教材為命題依據，測驗題材取自真實的生活情境，豐富有變化，測驗內容涵蓋所有與兒童經驗相關之主題，如日常生活與休閒活動、交通運輸、人物特徵、身體與健康、學校用語、時間與空間概念以及氣候與季節等。測驗實施方式為紙筆測驗，聽力測驗與閱讀測驗合併施測。

本報告首先簡介兒童華語文能力測驗之各項測驗能力指標與測驗內容，再針對 113 年兒童測驗各等級通過門檻、製卷、與成績公布之標準化流程進行概述。最後分別闡述本年度正式考試之信度及效度分析結果，使用資料為 113 年 1 月至 8 月之測驗結果。

## 貳、測驗介紹

### 一、測驗說明

兒童測驗共分為三個等級，由易至難依序為萌芽級（Sprouting）、成長級（Seedling）與茁壯級（Blossoming）。測驗以「歐洲共同語文參考架構」（Common European Framework of Reference for Languages；簡稱 CEFR）之 A1 級（Breakthrough）和 A2 級（Waystage）（Council of Europe, 2001）的語言能力描述及定義為架構，所評量的語言能力皆落在初級語言使用者（Basic User）的範圍之內。其中，成長級和茁壯級分別根據 CEFR 之 A1 級和 A2 級發展而來；與此同時，為了幫助初學華語、年齡較小的兒童建立學習語言的自信，順利銜接學習歷程，本會參考 ACTFL 能力大綱初等（Novice）之能力說明，同時也參考 CEFR 之 Pre-A1 能力描述，發展出比 A1 等級更簡單的萌芽級（Pre-A1）測驗。以下將就兒童測驗三等級的能力描述、測驗題型與題數、以及通過門檻三方面進行介紹。

#### （一）能力描述

兒童測驗是為七至十二歲母語非華語之兒童所設計的聽讀測驗，通過者所需具備的基本能力如表1所示。兒童測驗分為萌芽級、成長級和茁壯級三個等級，三等級的基本能力都以兒童生活經驗相關之語言使用為範疇，著重在「日常生活的一般簡易溝通能力」。

表 1 兒童測驗基本能力描述

通過等級	聽力測驗能力描述	閱讀測驗能力描述
萌芽級	在有充分視覺輔助的情況下，能從速度緩慢、發音清晰且重複的話語中，理解切身相關的基本詞語。	在有充分視覺輔助的情況下，能辨別出簡單句子中的基本詞語。
成長級	<ul style="list-style-type: none"> <li>在話語速度慢、發音清晰時，能理解與日常生活有關的簡短談話，例如與家人的互動、學校生活、購物、穿著、飲食、交通等等。</li> <li>在聆聽簡短、簡單、發音清晰的宣佈及說明時，能理解其中的要點，例如自我介紹、電話留言、課堂上的宣佈或天氣預報等。</li> </ul>	<ul style="list-style-type: none"> <li>在有視覺輔助的前提下，能理解簡單敘事短文的大意及找出重要訊息。</li> <li>能從日常生活的簡易書寫材料，例如菜單、車票、明信片及標示中，讀出基本訊息，例如：姓名、日期、時間、價錢、地點等。</li> </ul>
茁壯級	<ul style="list-style-type: none"> <li>能理解有關居家生活、興趣嗜好、旅遊、休閒活動、同儕間的互動等主題的談話。</li> <li>在聆聽簡短、簡單、發音清晰的說明或錄製片段時，能掌握主旨及重要資訊，例如：介紹、宣佈、廣播、新聞等。</li> </ul>	<ul style="list-style-type: none"> <li>能理解主題具體、與個人生活經驗相關的簡單敘事短文。</li> <li>能辨識日常生活中不同書寫材料的功能，且能從書寫材料，例如留言、便條、信件、學校公告、廣告或海報中，讀出重要的訊息，例如：活動時間、地點、辦法、適用對象、注意事項等。</li> </ul>

## (二) 測驗題型

兒童測驗目標受測者為七至十二歲的兒童，研發題型時，首先考量兒童受測者的認知能力及語言能力尚在發展中，在一次測驗中需藉由多元的作答方式與題型的轉換，提升兒童的注意力，因此測驗中包含多種題型，題型設計也採用兒童熟悉且較直觀的測驗形式，以降低作答時的焦慮情緒，以期能客觀評估受測者真實的語言能力。其次，兒童測驗三等級所希望測得的聽力及閱讀能力，都始於初級語言使用者的範疇中，而初級使用者理解文本的同時往往需要視覺輔助，因此

無論是聽力測驗或是閱讀測驗都採用了多種圖文相輔的題型。同時，由於兒童生活經驗與成人相比，較為有限，故測驗內容所包含的主題、情境、任務皆與兒童生活經驗相關，避免兒童因為缺乏充足的背景知識而影響其作答。

兒童測驗採聽力測驗與閱讀測驗合併施測，萌芽級為聽力測驗 20 題、閱讀測驗 20 題，測驗時間約 40 分鐘，成長級與茁壯級皆為聽力測驗 25 題，閱讀測驗 25 題，成長級測驗時間約為 50 分鐘，茁壯級約為 60 分鐘。測驗方式分為三選項之單選題、連連看形式呈現的配合題以及是非題三大類，聽力測驗與閱讀測驗各等級的題型請見表 2 與表 3：

表 2 兒童聽力測驗題型

測驗等級/題型	聽力選圖 (單選題)	看圖回答 (單選題)	聽力 連連看	會話理解 (單選題)	段落理解 (單選題)
萌芽級	✓	✓	✓		
成長級	✓	✓	✓		✓
茁壯級	✓			✓	✓

表 3 兒童閱讀測驗題型

測驗等級/ 題型	閱讀選圖 (單選題)	看圖辨義 (是非題)	看圖回答 (單選題)	圖文 連連看	閱讀材料 (單選題)	短文理解 (是非題)	短文理解 (單選題)
萌芽級	✓	✓	✓	✓			
成長級				✓	✓	✓	
茁壯級					✓	✓	✓

### (三) 通過門檻

兒童測驗三等級中，成長級與茁壯級欲測得的考生能力與成人聽力測驗和閱讀測驗的入門級、基礎級門檻相當，皆依據 CEFR 之 A1 級與 A2 級發展而來；唯測驗目標對象不同，前者針對兒童設計，而後者則是成人。因此，兒童測驗經過一系列等化研究與標準設定程序後，除了兒童測驗三等級試題難度皆能在同一量尺上相互比較外，也將兒童測驗試題難度參數連結至 TOCFL 難度參數的量尺上（國家華語測驗推動工作委員會，2014）。也因此，研發人員依據等化研究結

果及 TOCFL 入門基礎級標準設定的結果，制定了兒童測驗成長級與茁壯級之通過門檻（國家華語測驗推動工作委員會，2015a；2015b），同時也根據 TOCFL 準備級之標準設定結果制定了兒童測驗萌芽級的通過門檻（國家華語測驗推動工作委員會，2017）。兒童聽力與閱讀測驗訂有量尺分數轉換方式（國家華語測驗推動工作委員會，2014），三個等級量尺分數範圍皆為 0-100 分，聽讀合計後，量尺分數滿分均為 200 分。在證書核發規範方面，考量兒童測驗採取聽力測驗與閱讀測驗合併施測的方式，且兒童測驗對考生採鼓勵態度，因此以各等級兩項測驗通過門檻分數的總和作為證書核發的標準，不另外設定聽力或閱讀測驗單項門檻，萌芽、成長與茁壯級測驗通過分數如表 4 所示：

表 4 兒童測驗通過分數

測驗等級	通過分數
萌芽級	101
成長級	116
茁壯級	118

## 二、等級對照

圖 1 為兒童測驗聽力能力及閱讀能力的分級方式及與歐洲共同語文參考架構（CEFR）、美國外語教學協會指標（ACTFL Proficiency Guidelines）之對應關係<sup>1</sup>。

<sup>1</sup> 資料來源為本會於 2012 至 2017 年進行之標準設定系列研究成果。

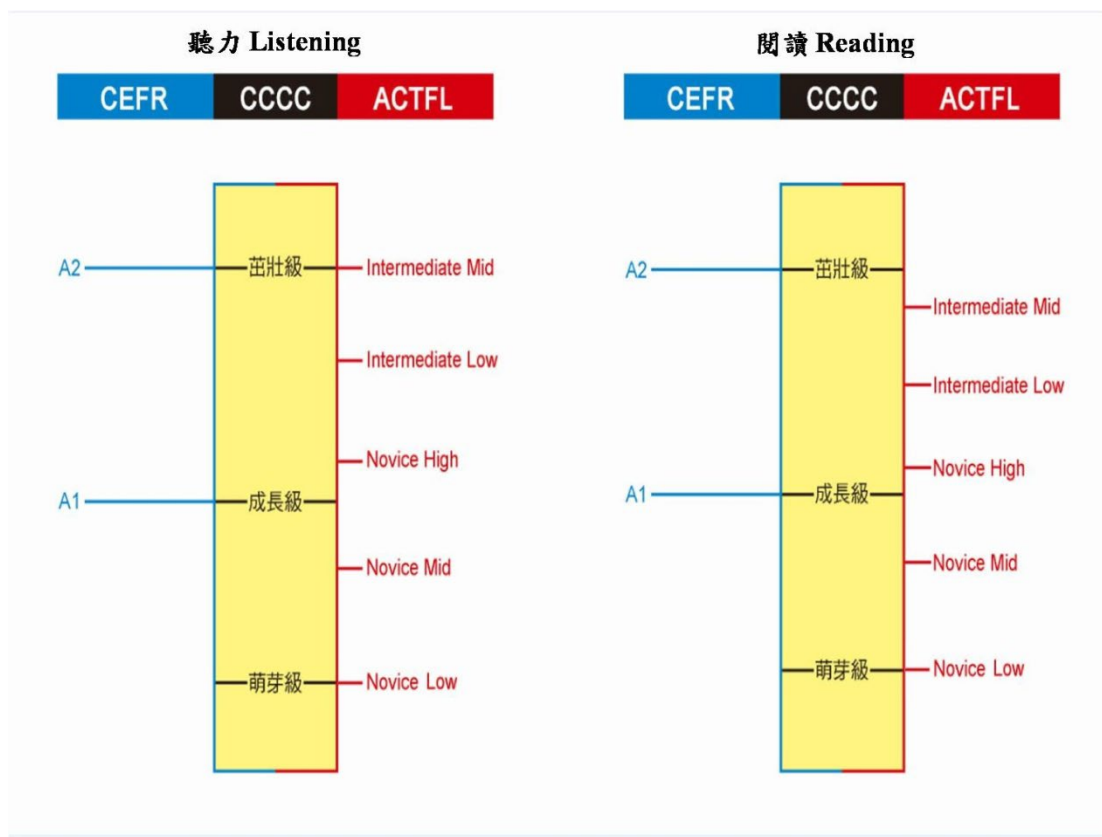


圖 1 聽力與閱讀測驗等級對照圖

### 三、測驗標準化流程

華語文能力測驗標準化流程如圖 2 所示，「測驗標準化」各步驟說明如表 5。

共包含：試題收集、修審、題庫輸入、組預試卷、試題分析、組合正式卷、成績檢核與口寫評分、考後結果分析等八個步驟，加上對外舉辦的預試考試、正式考試與考後成績公佈，完成整套「測驗標準化流程」。

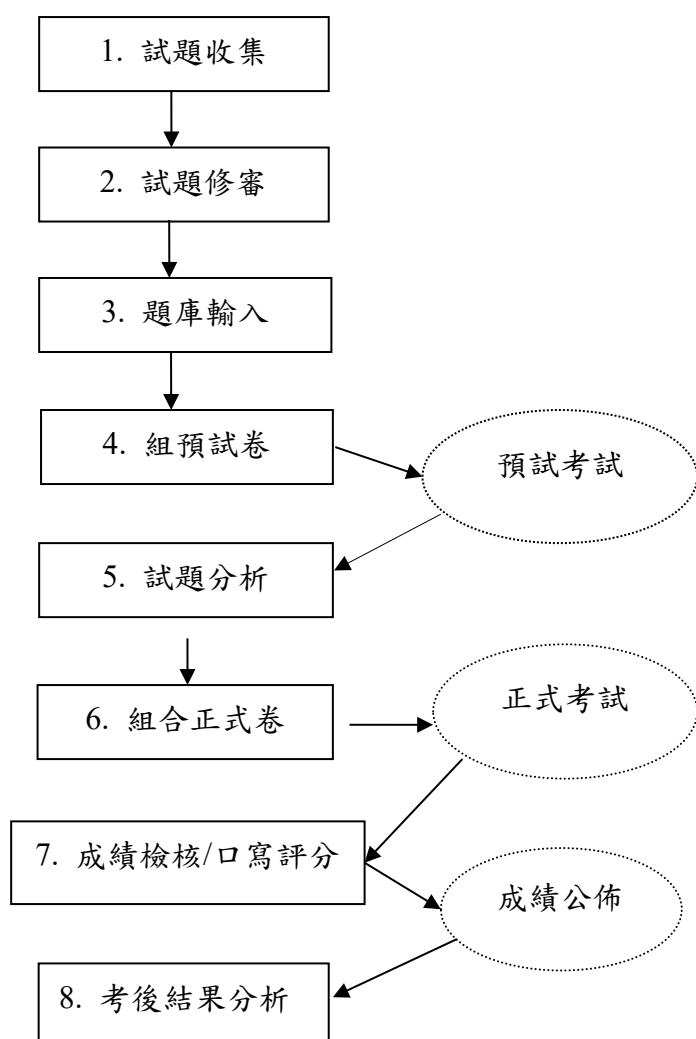


圖 2 測驗標準化流程

表 5 測驗標準化流程說明

項目	說明
1. 試題收集	辦理命題研習；與命題教師進行試題設計溝通
2. 試題修審	組修審稿、繪圖、三階段審查(會內初審、專家審查、會內複審)聯繫審查專家；依據審查意見修稿。
3. 題庫輸入	音檔、影片檔錄製與檢核；試題校對。
4. 組預試卷	紙筆測驗：製作繁體版與繁簡版題本；多國語題本。 電腦測驗：題目由題庫產出電子卷包，放上考試系統進行繁、簡版題目檢核。
5. 試題分析/ 統計數值輸入	成績檢核；試題統計分析；試題討論(確定保留的題目將標記為入庫題，並輸入統計分析數值，以利後續組正式卷)。
6. 組合正式卷	紙筆測驗：同「組預試卷」。 電腦測驗：測驗包含 CAT、CBT、快篩，檢核同「組預試卷」。
7. 成績檢核	各測驗成績報表檢核：聽讀口寫、兒童測驗與統計分析分別針對考後各項成績報表進行檢核，確認成績無誤，對外公告成績。 考後各類報表彙整。
8. 考後結果分析	進行下列項目分析：正式考試測驗成績分布及通過比例分析；測驗信效度分析

## 參、測驗效能分析

兒童測驗組卷方式是依據試題反應理論(Item response theory; 以下簡稱 IRT) 而來的, IRT 之一重要假設為單向度假設 (uni-dimensionality); 亦即測驗中所有題目皆測量相同潛在特質 (或能力), 當考生並非靠所欲測知的單一特質 (或能力) 作答時, 單向度 IRT 分析所獲得的試題參數及考生能力估計值將是偏誤的。上述測驗藉由 Winsteps<sup>2</sup>進行單向度 IRT 分析測驗相關資料, 分析結果可區分為信度分析與效度分析。

### 一、信度分析

以下從內部一致性及測驗訊息量 (test information,  $TI(\theta)$ ) 兩個面向評估兒童聽力測驗與兒童閱讀測驗的信度。

#### (一) 內部一致性信度

測驗的信度是指測驗結果的一致性和穩定性, 亦即測驗題目是否一致測量到相同的潛在特質。內部一致性信度是常見的測驗信度評估指標之一, 它是指同一份測驗中所有試題的相關程度有多高。如果所有試題皆測量到相同的目標特質, 那麼每道試題測量結果的相關性應該很高 (陳柏熹, 2011), 因此, 內部一致性指標可以顯示內容取樣及內容異質性的誤差, 以反映內容取樣對於測驗分數的影響。各級測驗之內部一致性信度將以庫李 20 號公式 (Kuder-Richardson) 為指標。

由表 6 顯示, 除了萌芽級、茁壯級聽力測驗外, 其餘聽力與閱讀測驗信度係數達到 .80 以上, 信度大致良好。萌芽級聽力測驗信度較低的原因應與考生同質性較高有關, 該卷考生在 20 道試題的平均答對題數為 19 題, 答對題數的標準差為 1.80, 成績分布為負偏態; 茁壯級也有類似情形, 25 題平均答對題數為 22.9 題, 標準差為 2.30。

---

<sup>2</sup> Linacre, J.M. (2009). Winsteps® (Version 3.68.2) [Computer Software]. Beaverton, Oregon: Winsteps.com.

表 6 兒童測驗各等級聽力、閱讀測驗庫李信度係數

測驗類型	萌芽級	成長級	茁壯級
聽力測驗	.75	.87	.67
閱讀測驗	.80	.81	.90

## (二) 測驗訊息量

測驗訊息量表示測量精準度，訊息量越高，表示試題對於考生能力的測量精準度越高。較詳細的理論及公式請見華測會出版之《華語文能力測驗技術報告 2013-1 聽力測驗信效度》第四章第一節。

圖 3、圖 4 為成長級聽力測驗測驗訊息量及測量誤差，其中，測量誤差均轉換為量尺分數。因篇幅有限，其他測驗類型及等級之測驗訊息量與測量誤差圖請見附件 1。由圖 3、圖 4 可知，成長級聽力測驗在通過門檻提供的測量相較於其他分數而言，具有較高的測驗訊息量與較低的測量誤差，並在通過門檻具有良好的信度係數，顯示測驗信度良好。

兒童測驗各等級正式考試通過門檻的測驗訊息量、測量誤差及轉換後的信度係數如表 7 所示，各等級各項測驗結果相仿，以成長級為例，成長級聽力測驗與成長級閱讀測驗門檻的測驗訊息量分別為 5.58 和 5.40，測量誤差分別為 5.29 和 4.30，轉換後信度係數各為 .92 和 .89。其他各等數值請參見表 7，測驗訊息量與測量誤差圖請見附件 1。

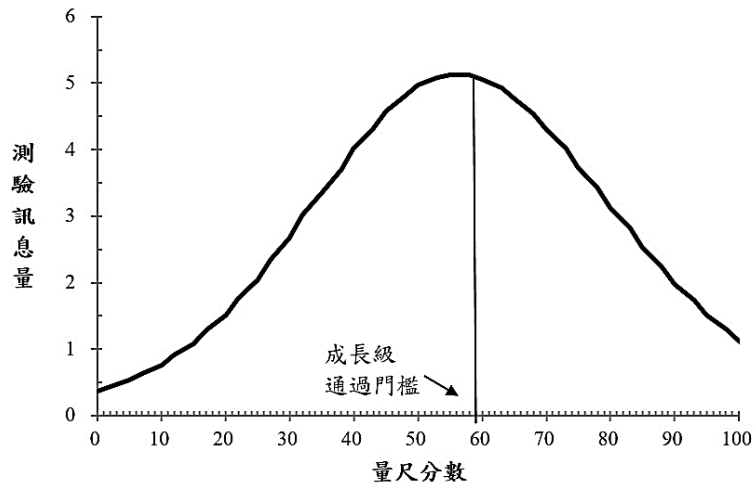


圖 3 成長級兒童聽力測驗測驗訊息量

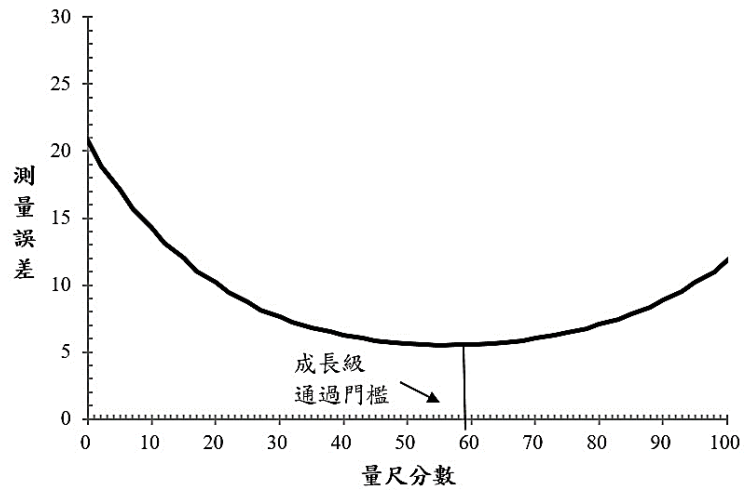


圖 4 成長級兒童聽力測驗測量誤差

表 7 兒童測驗各等級門檻測驗訊息量、測量誤差及信度

測驗類型	等級	測驗訊息量	測量誤差	轉換後信度係數
聽力	萌芽級	4.19	6.10	.82
	成長級	5.58	5.29	.92
	茁壯級	5.37	5.40	.87
閱讀	萌芽級	4.24	4.86	.87
	成長級	5.40	4.30	.89
	茁壯級	5.00	4.47	.93

註：因成長級聽力測驗與茁壯級閱讀測驗考生能力值標準差較大，分別為 1.49 及 1.74，故轉換後信度係數略高。

## 二、效度分析

為評估試題所測量的能力是否與測驗發展所訂定的架構內容相吻合，且是否測量到所欲測量的能力，兒童測驗效度分析藉由試題分析、因素分析，評估測驗的建構效度。

由 IRT 試題分析與驗證性因素分析結果可知，華語文能力測驗各項測驗各等級皆具有一定之建構效度。

### 1. IRT 試題分析

由於預試時考生答題動機可能不如正式考試，加上正式考試考生整體能力未必與試題難度相當，考生整體能力有時高於試題難度，適配度容易有偏低的現象，因此在正式考試採取較為寬鬆的標準，以 Winsteps 進行單向度 IRT 分析，採用 MNSQ 介於 0.6 到 1.4 以及 ZSTD 介於-3.0 到 3.0 的標準評估試題是否與單向度 IRT 模式適配，亦即 MNSQ 以及 ZSTD 超出設定範圍者，表示題目不符合單向度 IRT 模式。兒童測驗各等級試題適配分布如表 8 所示。在兒童測驗部分，除了萌芽級聽力測驗適配率較低，百分比為 75%外，其餘測驗類型與等級的適配率都達到 80%以上，顯示大部分試題都測量到相同構念的兒童華語文聽力或閱讀能力，具有建構效度。萌芽級聽力測驗適配情形較不理想原因，推測應與測驗難度較低，且應試者平均能力值與試題平均難度差異較大（3.7logits）有關。在應試者能力遠高於試題難度的情形下，若不小心答錯，模式預期的殘差就會很大，進而造成試題不適配。

表 8 兒童測驗試題適配分布

測驗類型	測驗等級	總題數/向度	適配題數/向度	適配率
聽力測驗	萌芽級	20	15	75%
	成長級	25	20	80%
	茁壯級	25	23	92%
閱讀測驗	萌芽級	20	17	85%
	成長級	25	24	96%
	茁壯級	25	24	96%

## 2. 驗證性因素分析

此節使用 Mplus<sup>3</sup>進行驗證性因素分析，使用之估計方法與相關評估指標請參考華測會出版之《華語文能力測驗技術報告 2013-1 聽力測驗信效度》第四章第二節。兒童測驗各等級測驗之因素結構皆採單因素模式，亦即聽力理解能力、閱讀理解能力。

綜合基本適配度與整體適配度之分析結果，可得出以下結論，兒童測驗之各等級具有一定的建構效度，各個試題可測得一致之聽力或閱讀理解能力。由於篇幅有限，在此僅呈現萌芽級聽力測驗和兒童閱讀測驗結果，其餘正式考試各等級各題因素負荷量等參數請見附件 2。

圖 5 為萌芽級聽力測驗單因素模式圖，在基本適配指標部分，單因素模式驗證性因素分析結果顯示，試題因素負荷量介於 0.36 至 0.89 之間，平均值為 0.65，95%的試題因素負荷量達.40 以上；各題因素負荷量統計考驗均達顯著水準 ( $p<.05$ )。圖 6 為萌芽級閱讀測驗單因素模式圖，單因素模式驗證性因素分析的結果顯示，試題因素負荷量介於 0.19 至 0.92 之間，平均值為 0.67，95%的試題因素負荷量達.40 以上；除 R\_Sp\_02 外，其餘試題因素負荷量統計考驗均達顯著水準 ( $p<.05$ )。

<sup>3</sup> Muthén, L.K. and Muthén, B.O. (2012). Mplus® (Version 7.0) [Computer Software]. Los Angeles, CA: Muthén & Muthén.

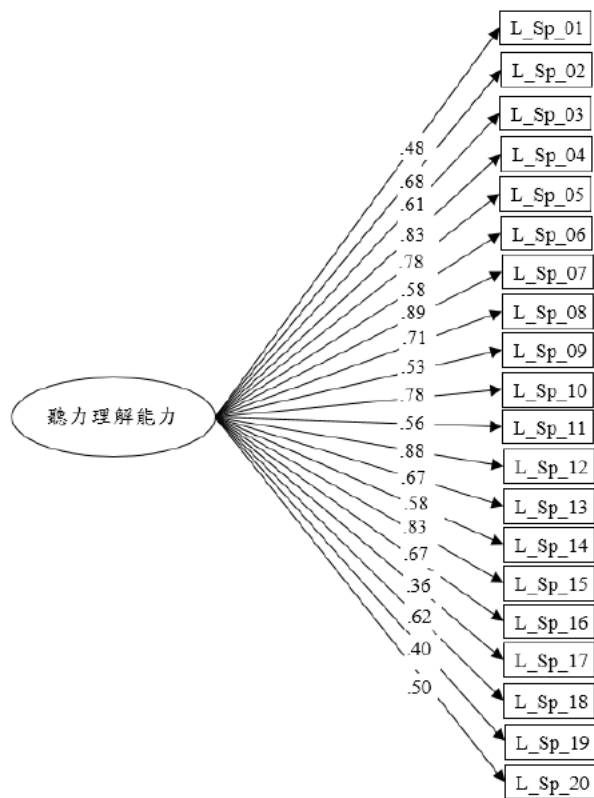


圖 5 萌芽級聽力測驗單因素模式

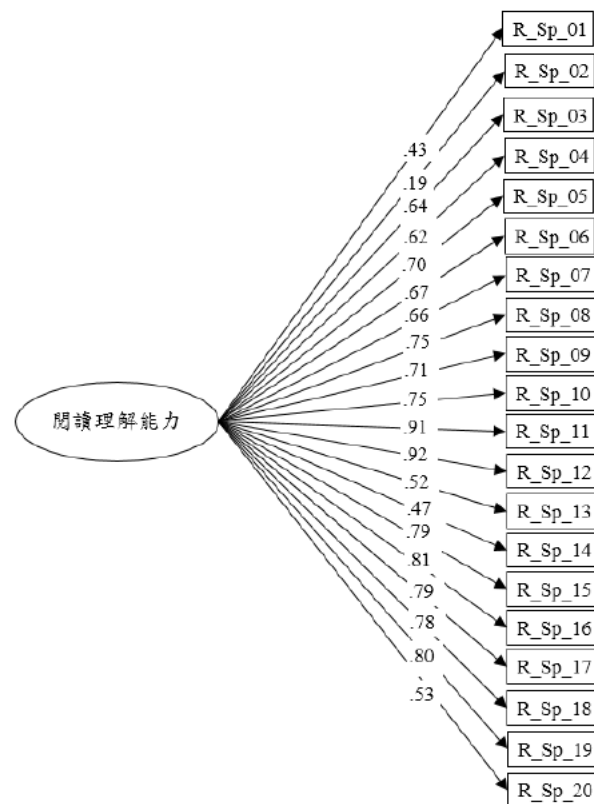


圖 6 萌芽級閱讀測驗單因素模式

透過整體模式適配度分析，可評鑑整個模式與觀察資料的適合程度。首先，經由卡方自由度比 ( $\chi^2/df$ ) 來評估模式適配度，其原因為卡方分配易受到自由度與估計參數數量的影響，當自由度越大或所需估計的參數個數越多時，卡方值越大，造成假設模式適配度不佳的可能性越大，此時卡方值越不能反映理論模式與觀察資料的適配程度 (邱皓政, 2011)。由表 9 可知，兒童測驗萌芽級、成長級、茁壯級的卡方自由度比均小於 3，表示模式適配度良好。絕對適配度評估的分析結果，平均概似平方誤根係數 (root mean square error of approximation, 以下簡稱 RMSEA) 均小於 .08，顯示皆符合絕對適配度指標。至於增值適配度評估部分，成長級聽力和茁壯級閱讀測驗的比較適配指標 (comparative-fit index, 簡稱 CFI) 和非規範適配指標 (non-normed fit index, 簡稱 NNFI) 都大於 .90，符合增值適配度指標，而其他等級或測驗類型的 CFI 和 NNFI 則大於 .80 而低於 .90，不過卡方自由度比和 RMSEA 均符合指標，顯示單因素模式整體適配度仍尚稱合理，未來將持續觀察或調整因素結構模型。

表 9 兒童測驗整體模式適配度指標摘要表

測驗類型	檢驗模式	$\chi^2/df$	RMSEA	CFI	NNFI
兒童聽力	萌芽級單因素模式	1.64	0.041	0.866	0.850
	成長級單因素模式	1.35	0.036	0.945	0.940
	茁壯級單因素模式	1.09	0.025	0.835	0.820
兒童閱讀	萌芽級單因素模式	2.39	0.060	0.880	0.866
	成長級單因素模式	2.03	0.062	0.823	0.806
	茁壯級單因素模式	1.22	0.039	0.964	0.961

## 肆、結論

本文主要針對兒童聽力測驗與兒童閱讀測驗之能力描述、測驗題型題數、通過門檻等方面進行概述，並說明測驗研發、施測和成績公布之標準化流程。同時檢視本年度正式考試信度與效度之測驗效能評估。

在測驗信度分析方面，藉由內部一致性指標來表示試題間之相關性，以確認兒童測驗的試題是否皆測量到相同潛在特質，並確認其程度為何，以及透過測驗訊息量來表徵測量精準度，結果顯示各等級內部一致性信度大致良好；也透過測驗訊息量檢視通過門檻的測量精準度，結果顯示各等級通過門檻皆具有較高的測驗訊息量與較低的測量誤差，顯示信度良好。

在測驗效度分析部分，本測驗的所有試題在施測前，皆通過專家審查來評估試題內容於各等級中的適切性以確保試題品質，具備內容效度的證據。接著，經由受測者作答反應分別進行試題分析及驗證性因素分析，確認由受測者反應所建構出的測驗架構是否與測驗發展之初所制訂的相同，以提供測驗之建構效度證據。試題分析的模式適配百分比結果顯示，大部分試題都測量到相同構念的能力；而驗證性因素分析結果，各等級單因素模式的整體適配度良好或尚稱合理，整體來說具有一定之建構效度。

經由信度及效度分析結果，顯示兒童測驗各等級之聽力測驗與閱讀測驗結果在測量一致性與構念代表性方面表現大致良好，能在一定程度上反映所欲評量的潛在能力構念，發揮了測驗效能。

## 伍、參考文獻

- 邱皓政 (2011)。結構方程模式：LISREL / SIMPLIS 原理與應用。雙葉書廊。
- 國家華語測驗推動工作委員會 (2015a)。華語文能力測驗技術報告2013—  
(1) 聽力測驗信效度。國家華語測驗推動工作委員會。
- 國家華語測驗推動工作委員會 (2015b)。華語文能力測驗技術報告 2013—  
(1) 閱讀測驗信效度。國家華語測驗推動工作委員。
- 國家華語測驗推動工作委員會 (2022)。兒童華語文能力測驗技術報告—2018  
(2) 聽力測驗、閱讀測驗信效度。國家華語測驗推動工作委員會。
- 國家華語測驗推動工作委員會 (主編) (2017)。華語文能力測驗106年成果報告[未出版成果報告]。國家華語測驗推動工作委員會。
- 陳柏熹 (2011)。心理與教育測驗：測驗編製理論與實務。精策教育。
- Linacre, J.M. (2009). Winsteps® (Version 3.68.2) [Computer Software]. Beaverton, Oregon: Winsteps.com.
- Muthén, L.K. and Muthén, B.O. (2012). Mplus® (Version 7.0) [Computer Software]. Los Angeles, CA: Muthén & Muthén.

## 陸、附件

### 附件 1 兒童測驗正式考試測驗訊息量及測量誤差

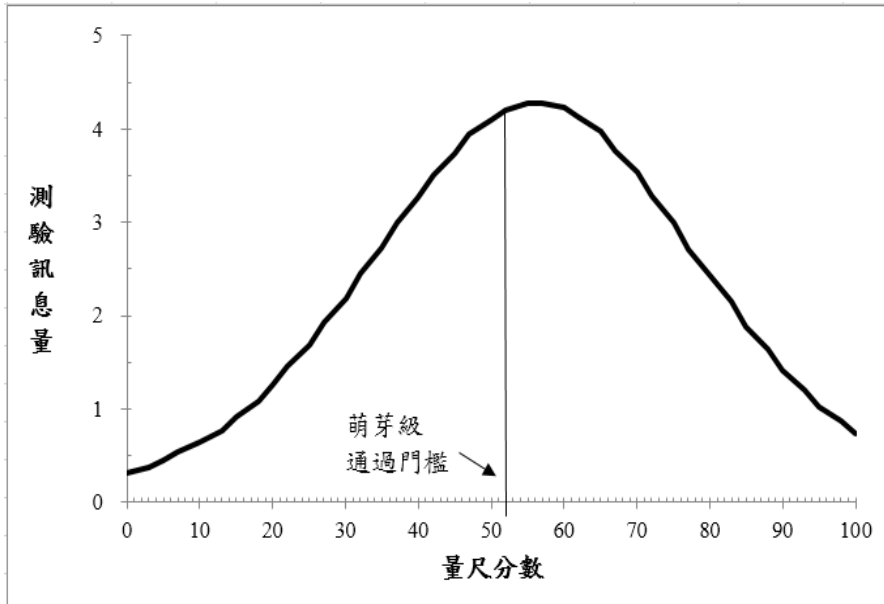


圖 1-1 萌芽級兒童聽力測驗測驗訊息量

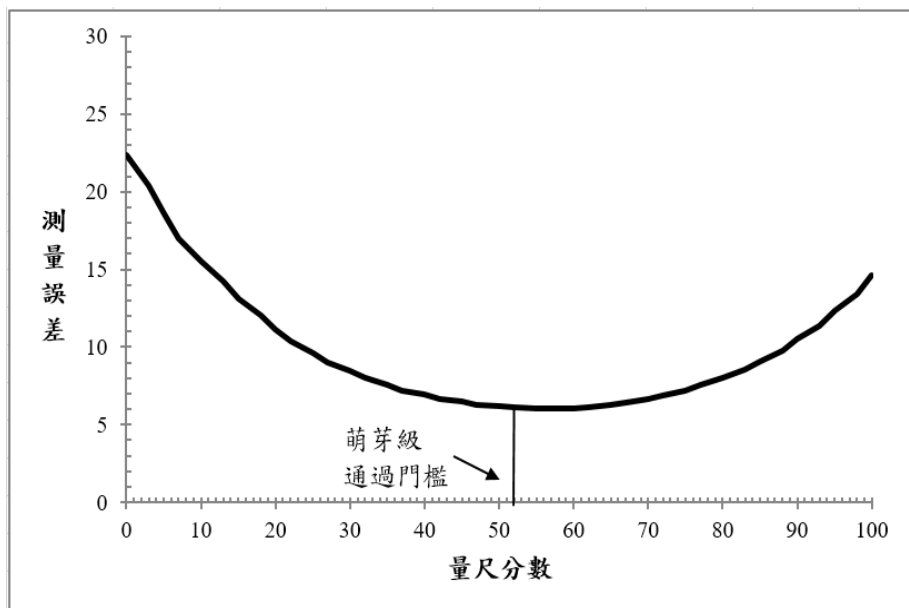


圖 1-2 萌芽級兒童聽力測驗測量誤差

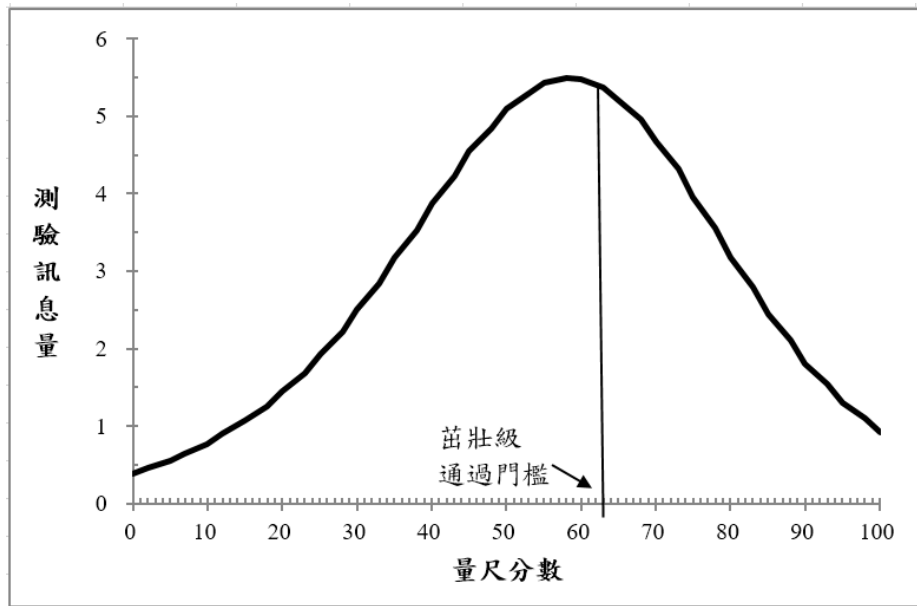


圖 1-3 苗壯級兒童聽力測驗測驗訊息量

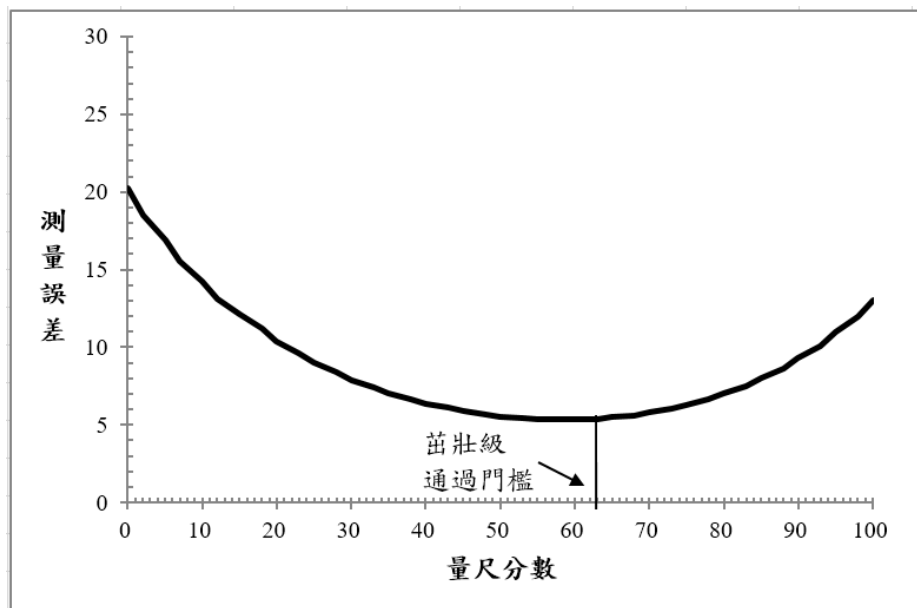


圖 1-4 苗壯級兒童聽力測驗測量誤差

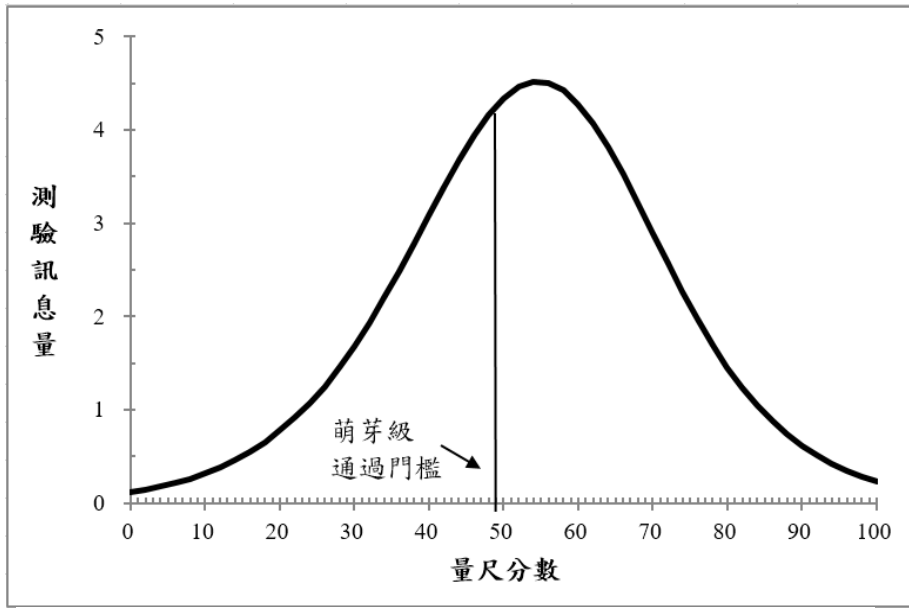


圖 1-5 萌芽級兒童閱讀測驗測驗訊息量

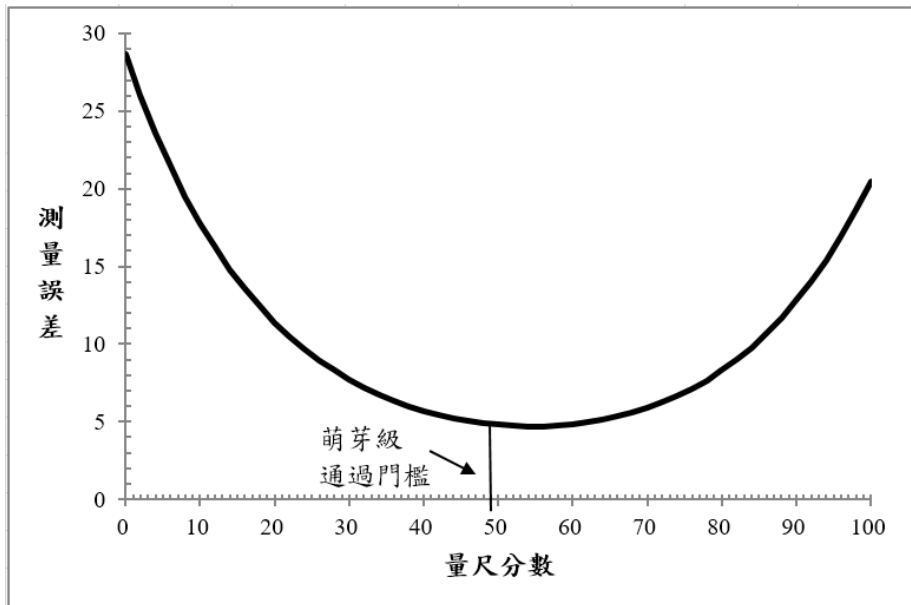


圖 1-6 萌芽級兒童閱讀測驗測量誤差

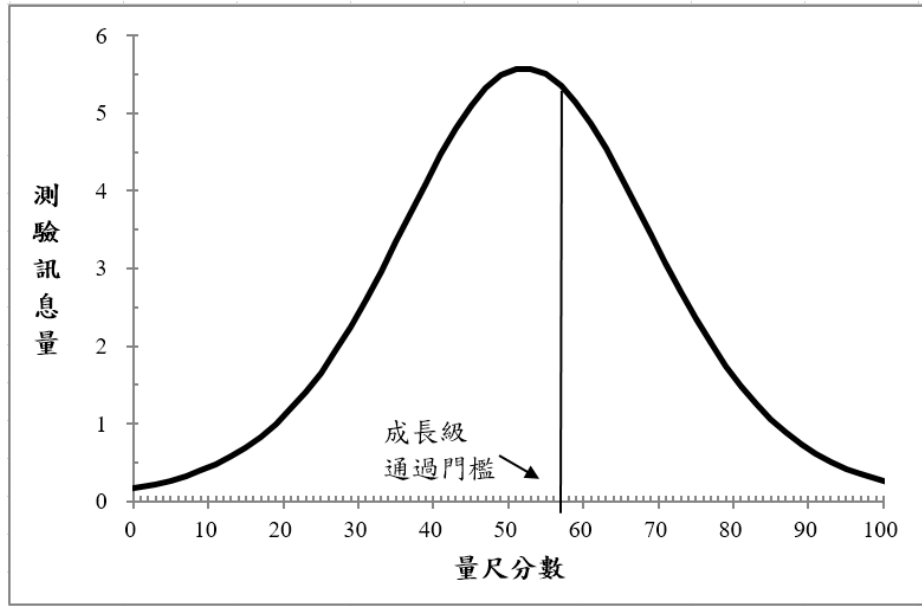


圖 1-7 成長級兒童閱讀測驗測驗訊息量

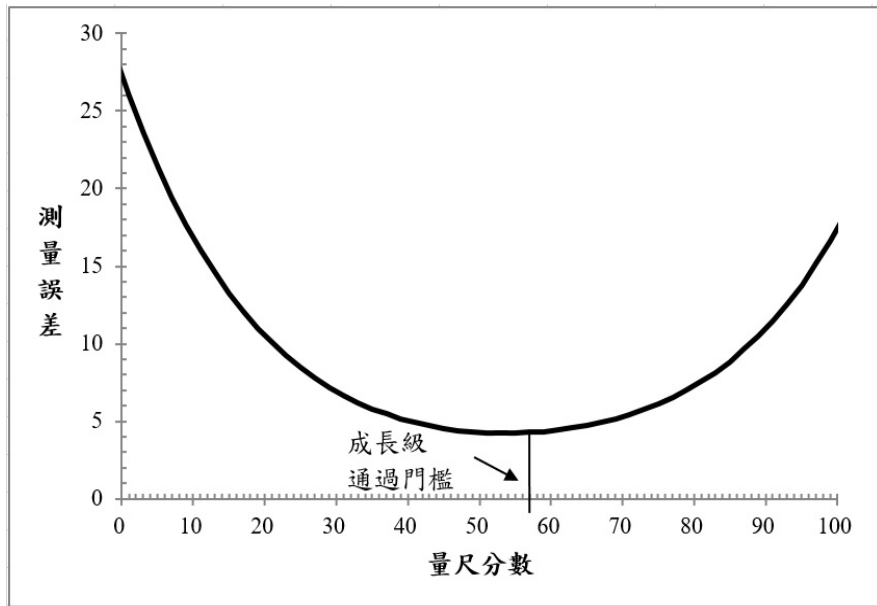


圖 1-8 成長級兒童閱讀測驗測量誤差

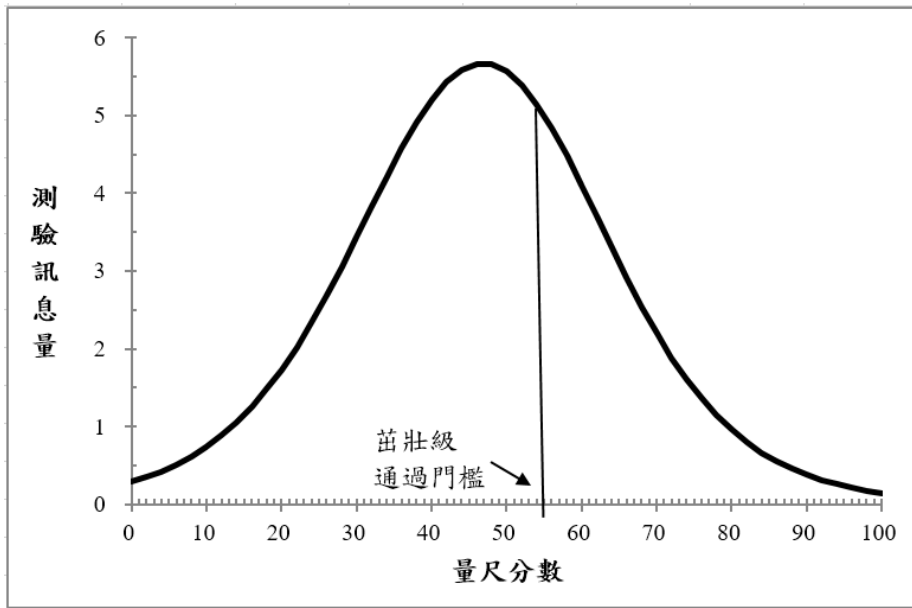


圖 1-9 茁壯級兒童閱讀測驗測驗訊息量

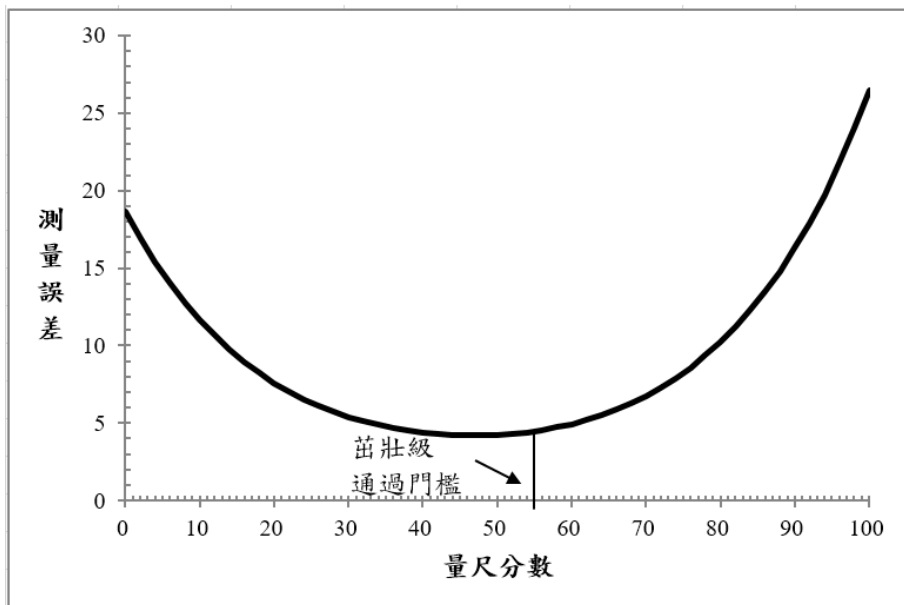


圖 1-10 茁壯級兒童閱讀測驗測量誤差

附件 2 兒童測驗正式考試驗證性因素分析結果

表 2-1 萌芽級聽力測驗各題因素負荷量摘要表

試題編號	因素負荷量	標準誤	P 值
L_Sp_01	0.48	0.06	0.00
L_Sp_02	0.68	0.08	0.00
L_Sp_03	0.61	0.13	0.00
L_Sp_04	0.83	0.06	0.00
L_Sp_05	0.78	0.06	0.00
L_Sp_06	0.58	0.10	0.00
L_Sp_07	0.89	0.04	0.00
L_Sp_08	0.71	0.06	0.00
L_Sp_09	0.53	0.09	0.00
L_Sp_10	0.78	0.07	0.00
L_Sp_11	0.56	0.08	0.00
L_Sp_12	0.88	0.05	0.00
L_Sp_13	0.67	0.08	0.00
L_Sp_14	0.58	0.09	0.00
L_Sp_15	0.83	0.06	0.00
L_Sp_16	0.67	0.08	0.00
L_Sp_17	0.36	0.08	0.00
L_Sp_18	0.62	0.08	0.00
L_Sp_19	0.40	0.08	0.00
L_Sp_20	0.50	0.07	0.00

表 2-2 成長級聽力測驗各題因素負荷量摘要表

試題編號	因素負荷量	標準誤	P 值
L_Se_01	0.75	0.11	0.00
L_Se_02	0.85	0.07	0.00
L_Se_03	0.73	0.07	0.00
L_Se_04	0.78	0.08	0.00
L_Se_05	0.67	0.09	0.00
L_Se_06	0.84	0.04	0.00
L_Se_07	0.74	0.07	0.00
L_Se_08	0.84	0.06	0.00
L_Se_09	0.58	0.08	0.00
L_Se_10	0.75	0.06	0.00
L_Se_11	0.75	0.06	0.00
L_Se_12	0.93	0.04	0.00
L_Se_13	0.86	0.04	0.00
L_Se_14	0.54	0.07	0.00
L_Se_15	0.92	0.03	0.00
L_Se_16	0.76	0.05	0.00
L_Se_17	0.77	0.06	0.00
L_Se_18	0.64	0.14	0.00
L_Se_19	0.71	0.07	0.00
L_Se_20	0.61	0.12	0.00
L_Se_21	0.50	0.08	0.00
L_Se_22	0.48	0.09	0.00
L_Se_23	0.57	0.08	0.00
L_Se_24	0.46	0.08	0.00
L_Se_25	0.68	0.07	0.00

表 2-3 茁壯級聽力測驗各題因素負荷量摘要表

試題編號	因素負荷量	標準誤	P 值	因素負荷量低於.3 可能原因
L_B1_01	0.54	0.20	0.01	
L_B1_02	0.70	0.23	0.00	
L_B1_03	0.99	0.08	0.00	
L_B1_04	0.93	0.21	0.00	
L_B1_05	0.41	0.15	0.01	
L_B1_06	0.70	0.23	0.00	
L_B1_07	0.82	0.19	0.00	
L_B1_08	0.30	0.12	0.01	
L_B1_09	0.19	0.13	0.14	正答描述較不直接
L_B1_10	0.31	0.13	0.02	
L_B1_11	0.60	0.11	0.00	
L_B1_12	0.27	0.14	0.06	對話句式結構較複雜，非典型茁壯級考法
L_B1_13	0.70	0.13	0.00	
L_B1_14	0.47	0.13	0.00	
L_B1_15	0.63	0.13	0.00	
L_B1_16	0.83	0.18	0.00	
L_B1_17	0.54	0.14	0.00	
L_B1_18	0.04	0.11	0.70	內容訊息量較多，考點偏細節
L_B1_19	0.22	0.12	0.07	內容訊息量較多，考點偏細節
L_B1_20	0.63	0.14	0.00	
L_B1_21	0.74	0.11	0.00	
L_B1_22	0.37	0.13	0.01	
L_B1_23	0.99	0.11	0.00	
L_B1_24	0.33	0.13	0.01	
L_B1_25	0.36	0.13	0.01	

表 2-4 萌芽級閱讀測驗各題因素負荷量摘要表

試題編號	因素負荷量	標準誤	P 值	因素負荷量低於 0.3 可能原因
R_Sp_01	0.43	0.08	0.00	
R_Sp_02	0.19	0.12	0.12	單純考數字，且在 1-5 以內，難度也相對簡單
R_Sp_03	0.64	0.08	0.00	
R_Sp_04	0.62	0.06	0.00	
R_Sp_05	0.70	0.06	0.00	
R_Sp_06	0.67	0.06	0.00	
R_Sp_07	0.66	0.06	0.00	
R_Sp_08	0.75	0.06	0.00	
R_Sp_09	0.71	0.07	0.00	
R_Sp_10	0.75	0.05	0.00	
R_Sp_11	0.91	0.04	0.00	
R_Sp_12	0.92	0.04	0.00	
R_Sp_13	0.52	0.10	0.00	
R_Sp_14	0.47	0.09	0.00	
R_Sp_15	0.79	0.04	0.00	
R_Sp_16	0.81	0.04	0.00	
R_Sp_17	0.79	0.06	0.00	
R_Sp_18	0.78	0.06	0.00	
R_Sp_19	0.80	0.07	0.00	
R_Sp_20	0.53	0.08	0.00	

表 2-5 成長級閱讀測驗各題因素負荷量摘要表

試題編號	因素負荷量	標準誤	P 值	因素負荷量低於 0.3 可能原因
R_Se_01	0.48	0.07	0.00	
R_Se_02	0.62	0.06	0.00	
R_Se_03	0.73	0.07	0.00	
R_Se_04	0.60	0.06	0.00	
R_Se_05	0.71	0.06	0.00	
R_Se_06	0.78	0.06	0.00	
R_Se_07	0.74	0.05	0.00	
R_Se_08	0.80	0.05	0.00	
R_Se_09	0.66	0.05	0.00	
R_Se_10	0.54	0.06	0.00	
R_Se_11	0.57	0.06	0.00	
R_Se_12	0.54	0.08	0.00	
R_Se_13	0.56	0.07	0.00	
R_Se_14	0.48	0.07	0.00	
R_Se_15	0.38	0.08	0.00	
R_Se_16	0.60	0.06	0.00	
R_Se_17	0.39	0.08	0.00	
R_Se_18	0.28	0.08	0.00	選項非簡單句式，考法和其他題較不同
R_Se_19	0.65	0.06	0.00	
R_Se_20	0.68	0.05	0.00	
R_Se_21	0.48	0.07	0.00	
R_Se_22	0.36	0.07	0.00	
R_Se_23	0.38	0.08	0.00	
R_Se_24	0.16	0.08	0.04	無法直接從文本抓取答案，非典型成長級考法
R_Se_25	0.44	0.07	0.00	

表 2-6 茁壯級閱讀測驗各題因素負荷量摘要表

試題編號	因素負荷量	標準誤	P 值	因素負荷量低於 0.3 可能原因
R_B1_01	0.57	0.09	0.00	
R_B1_02	0.77	0.08	0.00	
R_B1_03	0.79	0.07	0.00	
R_B1_04	0.84	0.07	0.00	
R_B1_05	0.82	0.06	0.00	
R_B1_06	0.75	0.07	0.00	
R_B1_07	0.78	0.08	0.00	
R_B1_08	0.79	0.06	0.00	
R_B1_09	0.72	0.07	0.00	
R_B1_10	0.64	0.08	0.00	
R_B1_11	0.76	0.08	0.00	
R_B1_12	0.65	0.08	0.00	
R_B1_13	0.62	0.13	0.00	
R_B1_14	0.73	0.08	0.00	
R_B1_15	0.20	0.13	0.12	需要對文本情節有較完整的理解，和其他題考法較不同
R_B1_16	0.67	0.10	0.00	
R_B1_17	0.76	0.07	0.00	
R_B1_18	0.76	0.09	0.00	
R_B1_19	0.73	0.07	0.00	
R_B1_20	0.62	0.08	0.00	
R_B1_21	0.73	0.07	0.00	
R_B1_22	0.64	0.08	0.00	
R_B1_23	0.78	0.06	0.00	
R_B1_24	0.77	0.06	0.00	
R_B1_25	0.77	0.06	0.00	