

# 建構英語文素養評量指引： TASAL 標準本位大型評量

回顧國際相關大型評量計畫中，包含「美國國家教育進展評量」(NAEP)、「國際學生能力評量計畫」(PISA)、「促進國際閱讀素養研究」(PIRLS)，甚至是2025年即將推動PISA 2025外語大型評量(PISA 2025 FLA)等，施測結果之所以能獲得學者與研究者的信賴，其基礎在於建立透明且具體陳述之評量架構、效度評估證據、評比結果報告等內容，讓世界各國能瞭解資料收集與測驗結果是如何分析、運用的來龍去脈。

基於此立場，本書目的在說明國家課程轉換時，如何將課程綱要所界定之教育目標，以形成評量指引。作者期望透過要點說明，建立臺灣自行發展的英語文素養大型評量有效性基礎。這本書或許只是個開端，卻是銜接後續透過「實徵資料結果導引教育決策、回饋國家課程綱要內容」的重要一步。

建構英語文素養評量指引：TASAL 標準本位大型評量

謝進昌 著

國家教育研究院

國家教育研究院  
National Academy for Educational Research



ISBN 978-626345120-9  
00230  
9 786263 451209  
GPN : 4711200014

謝進昌 著

# 建構英語文素養評量指引： TASAL 標準本位大型評量

謝進昌

112 年 2 月

本書通過雙向匿名學術審查

---

## 前言與致謝

---

本書完成有賴於國家教育研究院經費補助，作者想表達誠摯感謝，而其內容為作者推動院內研究成果與相關研究發展基礎，涉及計畫編號包含：NAER-107-12-B-1-07-06-1-08、NAER-2019-041-A-1-1-E1-07。本書內容所涉及相關研究，其推動是有賴於群體合作，才得以完成：一是有賴於研究期間，許多學者、專家對於相關研究概念發展意見與回饋，以及參與研究的英語文教師對於命題、修審題協助與投入等；二是本書所涉及學生實徵英語文表現，其資料蒐集也有賴於研究院共事同仁協助與投入，包含有行政協調、抽樣與大規模調查和評量作業、同仁間對於參數估計、次級資料分析概念激盪及磨合等。由於曾參與人員高達數百名，在有限篇幅下，作者較難逐一條列，但若沒有這些伙伴的參與、投入，是難以完成如此大規模研究成果。作者再次表達對於相關人員貢獻的謝意。

作者對於本書主要貢獻在於建立國家課程轉換間英語文素養評量架構，與透過學生英語文實徵表現資料進行效度評估，並利用有系統程序，進行內容撰寫與說明。本書部分內容來自於前述研究所延伸之未出版研究報告（謝進昌，2021a），考量相關報告為未出版內容，同時，缺乏針對英語文素養評量架構或效度評估議題，有系統進行探究與說明，不容易讓讀者一窺議題的全貌，因而，發起出版本書必要性；此外，作者也曾就相關研究內容，進行延伸發表、學術研討會參與或一般性命題技巧專書出版等，包含謝進昌（2020，2021b，2022a，2022b），但由於相關文章聚焦認知、情意面向的標準設定效度評估，其內容核心有別於本書焦點，故兩者間具有交互輔助，

來佐證研究與評量結果有效性之功能。另一方面，作者須強調本書僅運用所參與臺灣學生成就長期追蹤評量計畫（Taiwan Assessment of Student Achievement: Longitudinal Study, TASAL）之第四學習階段學生英語文素養表現資料進行論述。在此前提下，使得本書相關模式界定、試題與學生參數估計與次級資料分析等，僅能在此範圍內發展，其結論也會略有別於以整體五個學科（如使用多向度模式）或官方統一進行資料分析結果。最後，本書相關結果與結論並不代表任何政策決策，僅為作者個人學術觀點。

在回顧過去國內外相關大型評量計畫，包含如美國國家教育進展評量（National Assessment of Educational Progress, NAEP）、國際學生能力評量計劃（Programme for International Student Assessment, PISA）、促進國際閱讀素養研究（Progress in International Reading Literacy Study, PIRLS）、甚至是 2025 年即將推動的 PISA 2025 外語大型評量（PISA 2025 Foreign Language Assessment, PISA 2025 FLA）等，其結果能獲得學者與研究者信賴，基礎是建立在透明與具體陳述之評量架構、效度評估證據、評比結果報告等內容，讓全球讀者能一窺資料收集發展與測驗結果解釋、運用的來龍去脈。作者撰寫本書的出發點，也是基於此立場，期望透過具體要點說明，建立 TASAL 英語文素養評量資料使用有效性基礎。本書內容也許只是個開端，但卻是後續銜接「實徵資料結果導引教育決策、回饋國家課程綱要內容」基礎的重要一步。

---

## 摘要

---

十二年國民基本教育課程重視學生核心素養，而在政策推動下，國家課程革新對於學生英語文素養表現的影響與轉換，為重要關注焦點。在此脈絡下，國家教育研究院推動臺灣學生成就長期追蹤評量計畫（Taiwan Assessment of Student Achievement: Longitudinal Study, TASAL），目的之一在描述與追蹤同一群具代表性學生的英語文素養表現、探究影響其成長因子以及透過其實徵表現資料回饋國家課程綱要內容。本書目的在說明國家課程轉換間，如何將課程綱要所界定教育目標，轉化建構成評量指引，進而透過第四學習階段（國中）七、八年級學生正式實徵表現資料，以佐證相關英語文素養評量與調查工具有效性之評估基礎。

本書經敘說評量指引建構方式與內容，後續再透過三個面向來提供支持工具效度的證據，包含過程、內部與外部證據。就支持效度的過程證據而言，本書現行聚焦能提供「基本」支持效度過程證據的內容論述，包含有英語文理解評量架構、學習動機與策略使用理論說明和專家回饋建議、標準化評量與調查工具發展流程說明（包含如英語文理解題庫發展、組卷、施測與資料處理、抽樣及權重、參數估計和次級資料分析等程序），而在書寫這些過程說明內容時，其陳述要點大致依循現行各個國際大型評比所釋出的閱讀評量架構、技術報告，以力求具體、透明及詳盡陳述 TASAL 英語文素養評量工具建構和發展。

就支持效度的內部證據而言，本書透過接受九年一貫課綱最後一屆七年級（2019年）及其升至八年級（2020年）學生實徵表現資料分析結果來佐證，以檢核英語文理解評量與學習動機、策略使用調查工具之內部一致性信度、試題品質、模式適配度等。結果顯示各項評量與調查工具，皆具有接近 0.80 或以上之作答內部一致性，是具有不錯的信度水準；另一方面，英語文理解評量試題整體具有難易適中與高鑑別度，其中，除了少數試題未符合學者建議適配度準則外，其餘皆具有良好試題適配度，而試題難度分布也大致符合理論與邏輯原則，結果顯示愈高年級的試題、愈高層次理解認知成分的試題、具備較多抽象內容之第二、第三類群文本試題等，皆具有相對較高平均試題難度。就學生學習動機和策略使用調查工具而言，本書分別透過三因素、與四因素驗證性因素分析，結果顯示假定模式及實徵資料間具有良好適配度，後續，本書再搭配試題適配度檢核，結果顯示除了內在動機有兩個題項略超出學者建議試題適配度準則外，其餘試題皆具有良好試題適配度。整體而言，相關分析結果可提供支持工具有效性之內部證據。

就支持效度的外部證據而言，本書檢視「與相關構念之效標關聯以及各構念間聚斂、區別關聯」。就學生英語文理解表現與其學習動機、策略使用間之效標關聯而言，整體結果顯示在不分年級資料來源下，兩兩構念間 Pearson 積差相關係數，自 0.28 至 0.81 不等，實徵分析結果顯示出兩兩構念間是具有不錯效標

關聯效度。另一方面，若進一步檢視各主構念內的次構念、與跨主構念間的次構念所顯示之聚斂、與區別關聯，結果顯示各個次構念與英語文理解認知表現、或學習動機與策略使用各個次構念關聯，大致介於小至大程度效果量（積差相關係數介於 0.28 至 0.66 間）；此外，學習動機內的三項次構念、策略使用內的四項次構念間之兩兩關聯，大致屬於大程度效果量（積差相關係數是介於 0.61 至 0.81 間），顯示出本書所發展工具具有一定聚斂、區別關聯效度。

在檢視不同背景學生於英語文理解、學習動機與策略使用表現之差異時，結果顯示除了學習動機可能因年級增長而逐漸面臨較大壓力與學習負荷量，而未出現顯著差異之外，其餘包含英語文理解認知表現、策略使用等面向，皆顯示出學習表現會隨著學生年級增加而呈現顯著的成長。另一方面，本書透過單因子迴歸分析，經顯著考驗，檢視不同背景學生於英語文理解、學習動機與策略使用表現差異，結果顯示除了學習動機內有部分次構念與部分年級，並未隨學生背景差異出現顯著區別外，其餘皆一致顯示出男生在英語文理解、與記憶、認知、推論連結、理解監控策略使用方面，皆明顯低於女生；同時，在過去一年內，有接受學習扶助（補救教學者）於相關構念的平均表現，同樣明顯低於未接受學習扶助（補救教學者），相關結果是符合過去文獻與實徵研究結果。整體而言，不論就效標關聯效度、抑或是檢視不同背景學生平均表現差異，皆切合過去文獻與實徵研究結論，符合理論、邏輯與實徵結果方向，能適時提供支持本書所發展工具有效性之外部證據。

本書目的在於建立與提供支持 TASAL 英語文素養評量工具與資料有效性證據，以利後續決策者在透過相關評量結果進行詮釋與使用時，參酌的重要基礎。據此，本書於文末提出幾點建議方向，包含持續性長期推動效度證據累積與評估、108 課綱學生實徵表現結果回饋英語文國家課程綱要、大型評量次級資料分析提供教育政策發展建議、銜接第五學習階段英語文評量架構與抽樣設計等，以期相關大型評量資料之收集和使用能作出實質貢獻。

關鍵詞：十二年國民基本教育課程、大型評量、可能值、英語文素養、效度評估、標準本位評量

# 目次 Contents

前言與致謝.....	ii
摘要.....	iv
表次.....	x
圖次.....	xii
<b>第一章 國家課程綱要轉換與政策需求.....</b>	<b>2</b>
第一節 國家課程綱要轉換脈絡.....	2
第二節 國家課程內涵轉化於評量設計啟示.....	3
第三節 大型評量建置與政策需求.....	6
<b>第二章 英語文素養評量指引.....</b>	<b>10</b>
第一節 英語文素養與社會認知觀點.....	10
第二節 英語文理解認知成分.....	13
第三節 文本類型與接收性語言活動.....	33
第四節 源自文本與評量因素.....	40
第五節 轉化「教育目標」為「標準本位評量」.....	42
<b>第三章 英語文學習動機與策略使用.....</b>	<b>63</b>
第一節 社會文化脈絡的情境觀點.....	63
第二節 英語文學習動機.....	64
第三節 英語文理解策略使用.....	67
<b>第四章 效度評估.....</b>	<b>71</b>
第一節 效度評估證據來源.....	71
第二節 支持效度之過程證據評估.....	77
第三節 支持效度之內部證據評估.....	111
第四節 支持效度之外部證據評估.....	116

<b>第五章 結論與建議</b> .....	134
第一節 結論 .....	134
第二節 建議 .....	137
第三節 詞彙說明 .....	141
<b>參考文獻</b> .....	147
<b>附錄</b> .....	166
附錄一 教育目標（學習表現）與評量指引（認知成分）漸近複雜對應 .....	166
附錄二 教育目標（學習內容）與評量指引（相關名詞）來源與衍伸 .....	171
附錄三 語言知能文本範例 .....	174
附錄四 描述類文本範例 .....	176
附錄五 教學指引類文本範例 .....	181
附錄六 人際互動類文本範例 .....	183
附錄七 說明文範例 .....	193
附錄八 議論文範例 .....	201
附錄九 敘事與文學範例 .....	204
附錄十 英語文學習動機試題題項與適配度 .....	211
附錄十一 英語文策略使用試題題項與適配度 .....	213
附錄十二 正式評量試題參數與適配度（七、八年級） .....	217
附錄十三 預試評量試題通過率與鑑別度（九年級）（區塊 21 至 27） .....	228

# 表次

表 2-1	PIRLS、PISA 與 NAEP 閱讀理解認知成分 .....	31
表 2-2	PISA 2018 閱讀與 PISA 2025 FLA 外語評量閱讀認知成分對應.....	32
表 2-3	轉化教育目標為評量指引之詞彙使用 .....	48
表 2-4	學生能流暢運作語言基本知能表現描述（語言知能文本） .....	52
表 2-5	學生篇章閱讀理解（讀而寫）表現描述 .....	53
表 2-6	學生篇章聽力理解（聽而說）表現描述 .....	58
表 4-1	效度評估證據來源、對應與項目 .....	76
表 4-2	評量架構（初版）審核專家與建議重點 .....	81
表 4-3	學生問卷審核專家與建議重點 .....	83
表 4-4	七年級學生評量試題題本組卷設計 .....	85
表 4-5	八年級學生評量試題題本組卷設計 .....	86
表 4-6	九年級學生評量試題題本組卷設計 .....	88
表 4-7	各年級閱讀、聽力理解試題之分布 .....	90
表 4-8	各年級不同理解認知成分試題之分布 .....	91
表 4-9	各年級不同文本與篇章類型試題之分布 .....	92
表 4-10	閱讀與聽力理解篇章文本平均字數 .....	94
表 4-11	各年級不同類型文本平均字數 .....	94
表 4-12	英語文學習動機問卷題項跨年級組卷 .....	96
表 4-13	英語文理解策略使用問卷題項跨年級組卷 .....	97
表 4-14	本書所涉及學生背景特徵資料來源 .....	99
表 4-15	實際參與英語文評量（或學科問卷）學生背景分布 .....	103
表 4-16	透過 Jackknife 重覆抽樣以估計抽樣誤差示意 .....	109
表 4-17	本文評量與調查工具內部一致性信度 .....	112
表 4-18	閱讀與聽力理解試題平均試題難度 .....	113
表 4-19	不同理解認知成分之平均試題難度 .....	114
表 4-20	不同文本類型之平均試題難度 .....	114

表 4-21	七與八年級學生英語文理解可能值與學習動機、策略使用 WLE 能力相關係數.....	118
表 4-22	七、八年級學生英語文理解表現可能值描述.....	120
表 4-23	七、八年級學生學習動機與策略使用描述.....	120
表 4-24	不同背景七年級學生英語文理解表現可能值單因子迴歸分析.....	121
表 4-25	不同背景（七升至）八年級學生英語文理解表現可能值單因子迴歸分析.....	122
表 4-26	不同背景七年級學生英語文閱讀可能值單因子迴歸分析.....	122
表 4-27	不同背景（七升至）八年級學生英語文閱讀可能值單因子迴歸分析.....	123
表 4-28	不同背景七年級學生英語文聽力可能值單因子迴歸分析.....	123
表 4-29	不同背景（七升至）八年級學生英語文聽力可能值單因子迴歸分析.....	124
表 4-30	不同背景七年級學生內在動機單因子迴歸分析.....	125
表 4-31	不同背景（七升至）八年級學生內在動機單因子迴歸分析.....	125
表 4-32	不同背景七年級學生自我概念單因子迴歸分析.....	126
表 4-33	不同背景（七升至）八年級學生自我概念單因子迴歸分析.....	126
表 4-34	不同背景七年級學生社會動機單因子迴歸分析.....	127
表 4-35	不同背景七年級學生記憶策略使用單因子迴歸分析.....	128
表 4-36	不同背景（七升至）八年級學生記憶策略使用單因子迴歸分析.....	128
表 4-37	不同背景七年級學生認知策略使用單因子迴歸分析.....	129
表 4-38	不同背景（七升至）八年級學生認知策略使用單因子迴歸分析.....	130
表 4-39	不同背景七年級學生推論連結策略使用單因子迴歸分析.....	131
表 4-40	不同背景（七升至）八年級學生推論連結策略使用單因子迴歸分析.....	131
表 4-41	不同背景七年級學生理解監控策略使用單因子迴歸分析.....	132
表 4-42	不同背景（七升至）八年級學生理解監控策略使用單因子迴歸分析.....	132

# 圖次

圖 2-1	素養形成之學習者、情境與脈絡因素 .....	13
圖 2-2	閱讀理解成分 .....	16
圖 2-3	聽力理解成分 .....	17
圖 2-4	L134 各命題間關係概念圖.....	19
圖 2-5	L128 各命題間關係概念圖.....	22
圖 2-6	R018 產出推論與統整命題，以形成主旨 .....	25
圖 2-7	自我調節理解歷程示意圖 .....	27
圖 2-8	英語文標準本位大型評量建構模式 .....	44
圖 4-1	評量構念的建構理論觀點.....	73
圖 4-2	標準化評量與調查工具發展流程圖 .....	79
圖 4-3	TASAL 長期追蹤施測示意圖 .....	80
圖 4-4	權重計算示意圖 .....	102
圖 4-5	部分計分下的試題特徵曲線（3 個反應類別） .....	107
圖 4-6	部分計分下的 Thurstonian 閾值之累積機率曲線（3 個反應類別） .....	107



# 第一章 國家課程綱要轉換與政策需求

## 第一節 國家課程綱要轉換脈絡

十二年國民基本教育課程綱要總綱（以下簡稱 108 課綱）於民國 103 年 11 月 28 日已經正式發布，並訂於 108 學年度起，自國民中小學、及高級中學一年級起逐年實施（教育部，2017），以分年級、漸進方式取代自 90 學年開始推動、100 學年度修訂發布的國民中小學九年一貫課程綱要及高中九九課綱（以下簡稱九年一貫課綱）（教育部國民及學前教育署，無日期）。108 學年實施的課程綱要願景為重視成就每一個孩子，建立「適性揚才」、「終身學習」為培養孩子的目標，同時，結合自發、互動、共好理念，期使課程發展能以生命主體為起點，透過培養學習者核心素養，讓孩子潛能得以適性發展，進而能運用所學知識、善盡責任，逐步成為學會學習的終身學習者（教育部，2021）。為達到此目標，教育部除了逐步建構起各教育階段專業支持系統及網絡，具體化各類專業支持系統的定位、功能以及其與課程發展相關環節的連結外，在考量課程綱要的轉化及落實，同步調整、建立相關配套整合，例如，入學考試制度、師資培育制度與課程改變、教師專業學習社群發展、跨單位整合平台與協作機制的有效運作等（國家教育研究院，2014）。然而，在推動前述國家課程主軸革新發展及其配套措施建置下，引起關切的是在 108 課綱實施後，國家課程革新與轉換對於學校學生素養表現可能產生的影響及評估，同時，如何累積有效的學生實徵表現證據，以為下波國家課程改革建立基礎。

在回顧過去文獻與研究，針對國家課程改革對於實務現場影響，許多學者曾就九年一貫課程現況作出剖析與探討，例如，方德隆（2010）、吳俊憲與黃政傑（2010）、黃純敏等人（2004）；此外，學者也曾就國際或國內各大型教育評量、調查結果，來回應臺灣學生成就、素養表現概況，例如，國家教育研究院所推行臺灣學生學習成就評量計畫（Taiwan Assessment of Student Achievement, TASA）（蕭儒棠等人，2017）、經濟合作與發展組織（Organisation for Economic Co-operation and Development, OECD）所推動國際學生能力評量計畫（Programme for International

Student Assessment, PISA) (余曉清、林煥祥, 2017)、國際教育成就評量學會 (International Association for the Evaluation of Educational Achievement, IEA) 主辦促進國際閱讀素養研究 (Progress in International Reading Literacy Study, PIRLS) (柯華葳等人, 2017) 等。

在回顧相關研究, 本書發現國內鮮少針對國家課程轉換期間執行之影響研究, 多是以課程推動後評估或調查為取向, 以致較難掌握課程轉換當下之學生成長與變化影響概況。其中原因可能在於若欲符合該議題研究, 必須完全配合國家課程政策推動時程提前規劃, 具有相當難度且須高度仰賴主管機關行政協助; 另一方面, 國際大型教育評量雖然具有回應各國家課程實施成效之目的, 但因其評量架構必須符合國際規範, 難以完全切合單一國家課程內容。在此背景下, 為評估國家課程轉換對於學生素養表現影響概況, 國家教育研究院成立臺灣學生成就長期追蹤評量計畫 (Taiwan Assessment of Student Achievement: Longitudinal Study, TASAL), 目的在描述與追蹤同一群具代表性學生的素養表現, 以及透過實徵資料探究影響學生成長變化因子、回饋國家課程綱要內容發展 (國家教育研究院, 無日期)。

在上述國家課程轉換脈絡與背景下, 本書目的在建立 TASAL 第四學習階段英語文素養評量架構, 同時, 就現行所獲得學生英語文素養實徵表現資料, 建立支持相關評量與調查工具有效性之證據評估, 以期為後續相關評量結果的詮釋與使用, 建立穩固基礎。

## 第二節 國家課程內涵轉化於評量設計啟示

在上述脈絡背景下, 有關九年一貫課程綱要與 108 課程綱要內容重點差異與比較分析, 國家教育研究院 (2018b) 所撰寫英語文課程手冊, 已經完整且詳實分析英語文課程綱要中九年一貫課綱與 108 課綱之差異。其內容主要依照課綱整體 (包含基本理念、課程目標、時間分配、核心素養、學習重點、選修課程綱要、實施要點等) 以及學習重點 (如學習表現與能力指標對應) 細部比較並進行說明。其中, 考量本書是以大型評量為論述目標, 並非以教學為導向, 因此, 以下僅就相關部分內容進行重點闡述, 以利讀者了解相關背景與發展。

108 課綱的研修有幾項重大變革，是有別於九年一貫課綱。首先，就總綱層面而言，108 課綱提出核心素養（core competence）概念，著重三面九項具體內涵；進而，課綱發展者將此內涵貫穿著十二年國民基本教育，結合英語文基本理念與課程目標，研擬出利用「學習表現」與「學習內容」來陳述各學習階段學生英語文教學與評量之教育目標、方向，而其內涵除了英語文本身學習外，更包含科技資訊與媒體素養，強調著跨領域能力的培養（國家教育研究院，2018a）。整體而言，108 課綱是以核心素養作為組織各學習階段學生所需學習目標、內容以及建立跨領域課程之統整主軸。

就學習重點而言，除了九年一貫課綱強調的語言能力、文化與習俗外，108 課綱新增「培養以英語文進行邏輯思考、判斷與創造力」，例如，9-IV-3 能根據上下文語境釐清不同訊息間因果關係（國家教育研究院，2018a，頁 16）。同時，課綱發展者更突顯學生學習興趣與態度以及學習方法與策略之重要性，例如，7-IV-2 善用相關主題之背景知識，以利閱讀或聽力理解、6-IV-4 樂於接觸課外的英語文多元素材，如歌曲、英語學習雜誌、漫畫、短片、廣播、網路等（國家教育研究院，2018a）。整體而言，108 英語文課程綱要強調著培養學生透過英語文學習培養高層次思考能力、跨領域多元識讀能力、透過有效策略進行英語文學習以及樂於以語文進行學習能力。

學生素養的養成逐漸受到諸多國家的重視，以作為國家課程改革的理論背景，而十二年國民基本教育課程發展的理念也同樣建立在類似的素養發展基礎上（教育部，2021）。為了避免受陷於西方教育實用主義所延伸定義的素養概念，同時，適時融入臺灣本土的教育傳統、價值與文化觀點，及兼顧課程發展之縱向連結、橫向統整與銜接九年一貫課綱目的，108 課綱所秉持的核心素養架構是歷經探究歐盟理事會所建構「歐洲語言共同參考架構」（Common European Framework of Reference for Language, CEFR）所強調語言功能性（Council of Europe, 2001）、經濟合作與發展組織（OECD）所委託「素養界定與選擇」研究案所界定工具溝通與互動、異質社群間互動及自主行動等素養範疇，以確保個人成功實現與社會良好運作（Organisation for Economic Co-operation and Development [OECD], 2005; Rychen & Salganik, 2003）及歐洲議會延伸發展「支持終身學習的核心素養」一文所提出的八大核心素養，強調包含如學習基礎、學習如何學習和高層次能力等培養（European Commission, 2007）。同時，課綱發展者在參考國內國家科學及技術委員會委託「界定與選擇國民核心素養」研究案以界定臺灣國民核心素養四維架構（洪裕宏，2008；陳伯璋等人，

2007），以及後續國家教育研究院所委託研究團隊進行串結與統整各教育階段核心素養架構與內涵（蔡清田，2011；蔡清田等人，2012；蔡清田等人，2011），以正式提出自主行動、溝通互動、社會參與等三面向及其延伸九大項核心素養。最後，課綱發展者再搭配各學科領域前導研究建議（張武昌等人，2013）與課程理論，來形塑現行英語文領域綱要內容。

108 課綱核心素養係指一個人為適應生活及未來挑戰，所應該培養、具備的知識、能力與態度（教育部，2021），而此觀點是有別於九年一貫課綱內容所重視的十大基本能力。核心素養內涵更強調著培養每位學生成為整體的人為目標，而非片段的知識及技能的學習，進而，在融入核心素養概念後，更具體突顯知識、能力與態度整體培養和結合的重要性。此外，核心素養更跳脫以學科知識為唯一的學習範疇，更重視其與情境結合並於生活中力行實踐的概念，強調著終身學習意涵，著重歷程、方法與策略的學習（國家教育研究院課程及教學研究中心核心素養工作圈，2015）。對上述概念，本書在進一步釐清說明。九年一貫課綱所強調「基本能力」，其理念是在能力本位（competency based）基礎，重視培養學生以聽、說、讀、寫及其所延伸發展的十項基本能力，然而，過往發展似乎過於重視學生能力培養，而容易忽略社會、文化等情境脈絡重要性，以致於過度強調單點能力的養成，而無法形塑出全人素養。在此基礎，核心素養的提出強調學生在面對未來生活，能結合生活與所學知識、技能與具備相應態度，迎接可能的挑戰（吳璧純、詹志禹，2018）。簡單來說，若將上述概念具體落實在英語文領域學習層面，其內涵重視學生能透過英語文學習，具備多元文化與觀點、思考能力、學習策略使用及掌握語言學習的溝通本質等（葉錫南，2016）。

除了從九年一貫課綱所強調的能力指標，108 課綱更改調整為學習表現（與學習內容）的規範外，相關組織編排、條目書寫有增、刪或修訂，更重要的基本理念與內涵轉變。整體而言，九年一貫課綱強調培育學生帶得走的基本能力，而 108 英語文課程綱要重視以學生為中心，培育成為全人的素養，同時，期望兼具知識、態度與策略學習能力，以回應現行或未來生活、學習挑戰。

### 第三節 大型評量建置與政策需求

大型評量（large-scale assessment）的建置具有評估、啟發與促進整個國家教育系統功能，而其出發點在透過建立具有國家代表性學生在某個課程所界定範圍內實徵表現概況，以了解現行政策推動與影響。由於大型評量具有嚴謹性功能，能有系統的進行學生實徵資料蒐集與分析，使得學生整體表現與相關議題次級資料分析，往往得以輔助教育決策者了解政策推動現況、結果與影響，以利後續擬訂或調整政策方向（Rutkowski et al., 2020）。在國際大型評量興起的期間，同時也伴隨著「以證據本位教育決策」（evidence-based education）理念的興起，許多學者、研究者紛紛訴諸決策的基礎得建立在有效的證據基礎下（Davies, 1999; Slavin, 2017），此理念也使得國內外大型評量計畫與成果更顯重要，包含如臺灣學生學習成就評量計畫（TASA）、美國國家教育進展評量（National Assessment of Educational Progress, NAEP）、國際學生能力評量計劃（PISA）、促進國際閱讀素養研究（PIRLS）等。這些計畫目的與發起理念雖略有差異，但共同目的皆一致兼持著回饋政策需求與重視大規模、有代表性資料蒐集的必要性。

在上述背景，可了解到大型評量雖然具有許多優點與重要性，然而，受限實務與邏輯可行性，本書仍得定位、聚焦評量方向，才得以有效回應原始建置目的。在此基礎，本書對於 TASAL 英語文素養評量面向，大致可定位如下條列要點。同時，各要點的詳細說明，請見本書各章節。

#### 一、轉化國家課程綱要所界定教育目標為評量指引

本書對於評量工具的發展，首要建立評量架構與指引，而所蒐集評量資料可回應學生在國家課程綱要表現概況。在此脈絡與目的下，最常見迷思在於是否可「直接」使用國家課程所界定學習表現或學習內容，來進行試題發展對應與評量架構與指引的建立？

國家課程綱要所界定之內容係為「教育目標、理念」，使得相關學習表現、學習內容的界定與陳述，會傾向使用較廣泛通用（transversal）與具有指引性（orientational）語詞來說明，以便多數教師能了解國家培育學生的目標與方向（Anderson et al., 2001）。然而，一般評量指引的書寫重視使用具體、特定（specific）

的描述，才能有效將各個發展的評量試題進行實質、有效的對應。舉例來說，以學習表現「3-IV-8 能了解短文、簡訊、書信的主要內容」為例（國家教育研究院，2018a），該指標所陳述的「了解」，是一通用性語詞，概念上表示就大方向而言，國家希望培育學生了解短文、簡訊、書信內容的能力；然而，就過往學者（如 Kintsch, 1988）所提理解歷程模式而言，「理解或了解」（comprehension）一詞是同時隱含著多個層次的概念（如解碼、字面意理解、隱含意理解、抑或是評估性理解等，詳見後續章節說明）。因此，使用者若直接將學習表現作為試題撰寫的對應，往往會讓彼此的對應過於籠統，容易落入各自解讀與失去試題對應功能。

讀者也可以用上述類似概念來檢視課程綱要所界定的「學習內容」指標，例如，不同體裁的簡易文章（Ae-IV-5）。就本指標的描述而言，其概念同樣包含許多類型文本，其中，幾乎多數篇章命題皆可對應到本指標。據此，就概念的層面而言，本書依據 Bloom 等人（1956）及其後續建議（Anderson et al., 2001），透過「動詞 名詞」使用來轉化學習表現、學習內容等教育目標，進一步建立能具體表徵學生「認知歷程與知識內容」之評量架構與指引，詳細可見第二章的說明。

## 二、評量架構以學生認知、知識層次為基礎與考量學習階段間表現銜接

108 課綱內容界定是以「學習階段」為單位（國家教育研究院，2018a），而其內並未具體區分出學習階段內各年級學生學習表現或內容差異。然而，在考量配合原始 TASAL 逐年級長期追蹤抽樣設計，本書得考量各年級間評量試題差異與整體（國中階段）結果是否能回應該學習階段表現、與上、下學習階段國家課程內容銜接；另一方面，在回顧各大型評量計畫，相關評量架構多是建立在理解認知與知識內容等面向（詳見第二章），因此，本書在區分各年級評量內容差異時，是透過組卷方式來進行調節，以區分出不同年級適合認知與知識內容評量試題比重，例如，低年級學生重視較低層次理解認知成分與相對較簡易文本內容，會置入較多相近試題，而愈高年級學生，會融入較高層次理解認知成分與相對較複雜文本內容、試題。最後，本書考量課程綱要內容安排具有跨學習階段銜接性，同時，評量試題存在跨年級垂直等化（vertical equating）共同題需求，以連結跨年級間學生表現（Kolen & Brennan, 2004）。整體而言，本書評量標的雖然聚焦於第四學習階段，其間，相關評量架構發展仍得考量跨學習階段銜接需求，以回應國家課程對於第三、四與五學習階段內容界定。

### 三、以總結性評量發展為方向，聚焦接收性語言活動的理解評量

本書源自於 TASAL 之英語文素養評量，而其設計為透過長期追蹤蒐集學生表現資料，其中，現階段實施是以第四學習階段七至九年級學生為焦點，並抽取年級內具代表性大規模樣本為施測標的。然而，在考量可能成本限制下，相關題型僅能以選擇、簡答式建構題等適合大規模計分方式進行評量，較無法採用適合課室內小規模樣本進行之實作評量等（詳見第四章）。此外，TASAL 英語文素養評量方向以發展適合該年級文本內容之總結性評量為目標，並聚焦英語文聽力、閱讀等接收性語言溝通活動（receptive language activities）理解，未涉及口說、（兩或多人）互動、寫作等面向評量。

### 四、平衡不同層次認知與語言知識之真實情境命題發展

若簡要略舉新舊國家課程綱要差異，其核心在於 108 課綱重視培養學生的「核心素養」，而此概念與九年一貫課綱所重視「基本能力」略有不同。若依照教育部回應表示「核心素養承續過去課程綱要的基本能力，但涵蓋更寬廣與豐富的教育內涵，其中，學校教育不再只是以學科知識作為學習的唯一範疇，而是彰顯學習者的主體性，重視學習者能夠運用所學於生活情境中」（教育部全球資訊網，2016）。在以此概念省思英語文課程綱要理念時，學生核心素養的培養，是應具備「語言溝通互動的功能性，以凸顯其獲取新知的工具性角色；同時，能引導學生具備獨立思考、處理與運用訊息的能力等」（國家教育研究院，2018a，頁 3）。簡言之，現行教育目標希望學生不單純只是「知道」（knowing）英語文基本知識（如字詞彙、語法等），更期待學生能依照所學得知識，進行高層次理解與思考，以因應日常生活與未來學習的任務與挑戰。整體而言，將理念落實為英語文素養評量發展時，本書認為相關試題得重視平衡不同層次學生理解認知與語言知識應用，而其佈題情境是以符合日常生活、社會與文化情境脈絡為導向。詳細說明請見第二章與第四章內容。

### 五、兼顧學生語言學習動機與策略使用

108 課綱核心素養係指一個人為適應生活及未來挑戰，所應具備的知識、能力與態度（教育部，2021），而此內涵除反映出培養學生認知能力外，國家課程綱要也重視培養學生學習動機與策略使用。此概念即呼應英語文素養理念注重「學習者的情意因素、及其發展自主學習與終身學習英語文的能力」（國家教育研究院，

2018a, 頁 3)。在落實相關情意面向評量上方面，本書配合 TASAL 大規模抽樣調查，因此，較難採用如一般課室內以行為觀察、晤談等深度評估方式，來檢核學生動機行為或策略使用。整體而言，本書透過學生自陳問卷發展，來測量學生對於英語文學習動機與理解策略使用傾向。詳細內容請見第三章。

## 六、銜接下波國家課程綱要發展方向

TASAL 英語文素養評量目的之一為透過學生實徵表現結果來回饋現行 108 英語文課程綱要。在此背景下，相關評量指引的建立與轉化，大多得使用與參照既有英語文課程綱要學習表現與學習內容的界定。然而，若就導引未來國家課程綱要發展目的而言，本書認為評量架構的建立應同時著眼於回應未來或下波國家課程綱要發展需求之評量資料蒐集，然而，相關指標或需求往往尚未、或未特別條列於現行綱要內容。舉例來說，伴隨著數位科技快速發展，在透過數位科技進行溝通活動下，學生藉由英語文進行如數位網頁與社交軟體之溝通和閱讀理解，更顯得普遍存在且重要。誠如 Britt 等人（2018）所提，學生對於這類多元文本的理解，係需要在既有認知理解基礎上，展現出更高或複雜層次表現。整體而言，對於英語文素養評量架構的發展，本書會適當融入未來可能的發展方向與指引，其中，是以學生在數位科技溝通情境下，考量其所需理解認知成分與知識作為建構、銜接方向。詳細內容請見第二章與第五章。

## 第二章 英語文素養評量指引

### 第一節 英語文素養與社會認知觀點

TASAL 英語文素養評量目的在於蒐集學生英語文素養實徵表現結果，以回饋英語文國家課程綱要調整，並從中探究可能影響學生成長表現之因子（國家教育研究院，無日期）。本書對於英語文素養評量架構的建立，主要呼應 108 英語文課程綱要內涵與理念（國家教育研究院，2018a），並經課程內容與內涵轉化，以建構各項評量指引。在說明評量架構組成前，本書先從英語文素養與社會知識觀點，進行簡要闡述。

#### 一、英語文素養

在第一章的說明中，108 英語文課程綱要理念係重視在真實生活或學術情境中，培養學生語言溝通、獨立思考能力、跨領域識讀能力以獲取新知，並透過有效策略進行英語文學習、或自主學習及樂於使用語文進行日常學習等（國家教育研究院，2018a），以回應未來生活中各種任務與挑戰。在此理念與精神下，英語文素養可定義為「在真實生活情境中，學生具備運用英語文溝通知能、富含學習動機與使用適當策略等有效能力，用以進行學習或回應現在、未來生活中的各項任務與挑戰」。整體而言，學習者得同時具備語言知識、相關認知與非認知能力等，以回應來自不同任務情境與社會文化脈絡挑戰。在此基礎下，本書主要關注英語文溝通知能、學習動機與有效策略使用，為建構素養評量指引之焦點。

在英語文溝通知能方面，它係指學生能透過聽、說、讀、寫與混合四種知能，於日常生活各項任務活動中進行有效溝通（國家教育研究院，2018a）。為達到有效溝通目的，學生得具備一定程度的英語文溝通知能（communicative competence）與一般性知能（general competence）（OECD, 2021）。就英語文溝通知能而言，它有別於抽象語言知識的學習，重視學生具有因應不同的情境脈絡、適當進行溝通的能力（Hymes, 1972）。若依據 Canale（1983）、Canale 與 Swain（1980）建

議，語言溝通知能大致能區分出語言知能（linguistic competence）、社會語言知能（sociolinguistic competence）、策略知能（strategic competence）與篇章知能（discourse competence）或擴大稱為實用知能（pragmatic competence）（林明佳，2018；OECD, 2021, p. 28）。

語言知能係指涉及語音、語詞、語法等建構語言的知識，重視學生能精準掌握語法和語言複雜性（complexity）、廣度（breadth）等（如字彙）（Council of Europe, 2020, p. 130; OECD, 2021, p. 28）；社會語言知能係強調在社會情境脈絡下，學生進行適當溝通所需知能，包含如，社會規範覺識、適當談話禮節等知識掌握，這項知能的形塑往往會受到國家文化或傳統語言學習習慣影響而不同；策略知能係指在語言知識外，學生能透過其他策略來進行有效溝通的知能；最後，篇章知能為注重學生能透過命題間和句子間關係，來達成溝通的知能，而在 OECD 對於語言知能的界定中，可擴大稱為實用知能（pragmatic competence），係注重語言使用原則的知能，包含篇章如何用於架構與組織的知能（篇章知能）、訊息如何被使用於執行溝通功能的知能（functional competence）、訊息如何被傳遞以達到互動、交流機制的知能（design competence）等（Council of Europe, 2020, p. 137）。最後，一般性知能與前者的概念略有不同，它的內涵並非涉及語言知能，而是重視學生對於這個世界所擁有、了解的知識與經驗等，包含如其他學科知識、生活經驗與技能、自身人格特質及學習態度等。由於一般性知能往往不是英語文素養評量所關注者，其作用往往作為試題撰寫者在進行命題時，會控制其干擾影響的參考來源（OECD, 2021, p.28）。

## 二、社會認知觀點

誠如 van Dijk（1988）所述「大部分的篇章都得在社會脈絡下進行詮釋」（p. 107），它代表著文本的意義不僅是源自於文本本身，更源自於社會群體參與者所共同分享的詮釋方式。其中，最常見的社會認知理解形式，例如，觀點、態度與思想意識等評估標準，即是建立在社會群體或文化所形塑規範、目標或價值基礎上。另一方面，除了語言理解認知發展外，非認知因素發展也同樣依循此脈絡，學者們認為語言學習者的動機觸發、投入、策略使用等，會與情境、社會脈絡間，形成一個動態、交互影響發展模式（Al-Hoorie, 2017; Dörnyei & Ryan, 2015）。其中，有關非認知面向探討，詳細請見第三章內容說明。

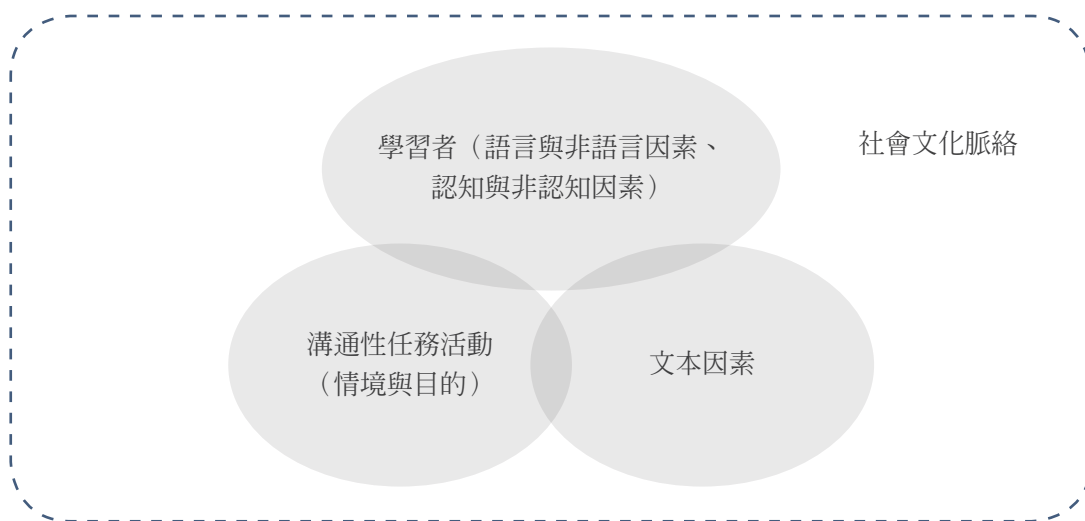
重視社會文化脈絡對於語言認知、非認知發展，或可稱為素養發展理念，現今已經逐漸成為趨勢，而此概念若對照學者在探究學習者是如何理解（comprehension）內涵、或相關模式發展時（Alvermann & Moje, 2019; Britt et al., 2018; Pearson & Cervetti, 2015; RAND Reading Study Group [RRSG], 2002），概念可謂不謀而合。相關概念皆指向語言理解會受到學生本身認知、語言發展、人際互動、社會與文化脈絡等交互影響（Gee, 2019）。當將此發展模式融入非認知層面，以界定素養形成時，也同樣承襲類似概念，學者認為素養本身不僅是一項學習項目，也是一項學習其他知識、獲取其他領域能力的途徑。其中，影響素養學習的因素，會隨著時空演變，大致可歸納出如圖 2-1 所示各個面向，分別是源自於學習者因素（learner factors）、文本因素（text factors）與語言溝通活動任務（communicative task）所形成的情境（situation）、與社會文化脈絡（context）等相互交織、動態運作影響彼此。

就源自學習者因素方面，學習者能否順利理解文本或建構其意義，往往會受到如自身先備知識及經驗、理解認知能力、語言知識、動機與策略等語言、非語言和認知、非認知因素影響；此外，就情境與脈絡因素而言，本書將情境定義為學習者當下所接觸局部、小範圍任務活動，而脈絡係指整體、大範圍的環境，為涉及廣泛時間與空間之概念，通常是一個動態、具變化性的社會文化型態（Alvermann & Moje, 2019）。

在情境因素方面，其影響源於學習者所直接接觸的文本因素或所投入的溝通活動型態。就文本因素而言，學習者會受到自身所設定目標、文本類型、文本長度、字彙與語法複雜度等，甚至是評量命題形式影響（RRSG, 2002）。就溝通活動型態方面，係涉及學習者所處實體空間內進行的溝通活動，例如，在學校、家庭、職場所進行閱讀、聽力、視聽、口說、寫作、交流互動等活動，而其參與溝通互動目的、情境條件與線索等，都會影響溝通形成和結果。

就脈絡因素而言，其屬於一動態、變化性的型態，而學習者可能是在有覺識或無覺識脈絡發展狀態下進行語言溝通與理解，其形態容易受到大環境變動影響，例如，身處環境之於外語學習自我意象與自信形成、數位科技發展使溝通模式與文化改變等（Al-Hoorie, 2017; Britt et al., 2018; Dörnyei & Ryan, 2015）。整體而言，在社會認知發展模式下，源自學習者、溝通情境與文本、社會文化脈絡等之因素，是彼此交織、動態運作著、同時影響學生素養的形成，是為本書評量架構建立的主軸核心。

圖 2-1 素養形成之學習者、情境與脈絡因素



## 第二節 英語文理解認知成分

社會認知模式認為影響理解或素養形成的因素，是由學習者、情境與脈絡等多個因素交織而成。若轉化成 108 課綱以評量學生核心素養為目標時，就學習者因素而言，本書關注的評量架構焦點大致可區分為學習者語言理解認知與學習動機、策略使用等非認知面向。在本節，本書先就理解認知成分進行說明。

### 一、英語文理解認知歷程

文本理解的認知運作歷程（cognitive process in text comprehension）是一個複雜的運作機制，至今仍不斷被探究。在經過數十年研究發展後，學者們大多逐漸認同它是一個結合由下而上（bottom-up）（Gough, 1972）與由上而下（top-down）（Goodman, 1982）的互動運作歷程模式（interactive model）。在此模式下，學習者在接收到訊息時，會就自身背景知識，不斷進行交互參照、補償、建構與統整、同化與調適修訂等動態、循環歷程，來形成其意義的理解與認知（Field, 2013; Goodman, 1988; Kintsch, 1988, 1998, 2019; Rumelhart, 1980, 2013; Stanovich, 1980; van den Broek & Helder, 2017; Verhoeven & Perfetti, 2008）。

面對複雜的理解歷程，其相應的評量工具發展往往得配合實務與邏輯可行性來作出轉化、調整，且其在某程度下仍不得不有所侷限，例如，本書評量僅能測量出學習者於該命題當下理解結果之外在表徵，並無法擷取學習者內隱的動態、連續性運作歷程。因此，本書傾向探討不同層次認知成分（cognitive component）組成說明，而非探究其實質運作機制。有關文本理解歷程的認知成分，為方便解釋內容，本書就兩個層面進行說明，一是在形成句子層次的命題理解（proposition comprehension at sentence level）；另一是篇章理解（discourse comprehension）。

### （一）在形成句子層次的命題理解認知成分

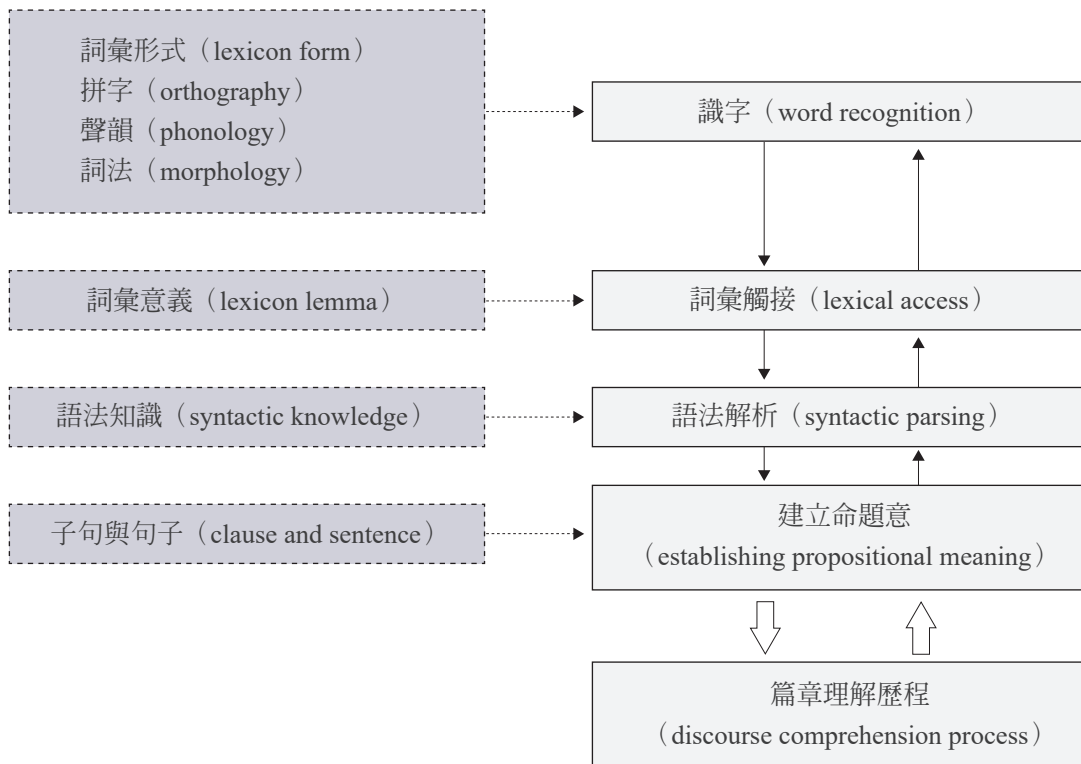
在形成句子層次的命題理解，其運作始於學習者對於字形或聲音的感官接收、解碼、比對與轉譯等過程（Gagne et al., 1993），而此階段大致包含有訊息解碼（decoding）、字詞辨識（word identification）與詞彙理解（vocabulary comprehension）等。早先 LaBerge 與 Samuels（1974）認為學習者會從自身的視覺（visual）、聲韻（phonological）、情節（episodic）與語義（semantic）記憶中，依序將所接收的形素（graphemic）或音素（phoneme）特徵，轉譯為字，再透過字的集合形成如名詞片語、動詞片語，並依照文法規則，主動建構句子意義（Samuels, 2013）。有別於重視由下而上解碼順序性，Rumelhart（2013）認為在接收到訊息後，學習者會同時結合和運作拼字知識、詞彙知識、語法知識、語義知識與脈絡情境訊息等，針對訊息進行適當理解。學者們對於運作機制的解釋雖然具有其差異性，但就實際應用於評量而言，OECD（2021）FLA 外語評量對於閱讀與聽力理解，如圖 2-2、2-3 所示，分別引用與修訂自 Khalifa 與 Weir（2009）閱讀、Field（2013）聽力認知歷程模式，則大多共享著與過去學者們所建議之相似共同認知成分。

就閱讀理解認知成分而言，如圖 2-2 所示。Khalifa 與 Weir（2009）認為學習者在接收到字面訊息後，會先比對文字的外形與內在拼字結構表徵來辨識文字（word recognition），再進入詞彙觸接（lexical access），透過外在形式（如拼字、聲韻、詞彙結構）與語義訊息結合，來理解詞彙意義。後續，經搭配語法結構解析（syntactic parsing）或通稱文法之運用，學習者就能透過更多字彙的集合，形成更大單位，如片語，來理解句子或子句中命題意義。

在聽力理解認知成分方面，其運作依循著類似訊息處理觀點，如圖 2-3 所示。若依照 Field (2013) 所描述認知成分，相關概念大多雷同於閱讀理解成分，是自解碼、詞彙搜尋 (lexical search)、語法解析至命題理解等。然而，有別於文字訊息，聽力理解重視透過聆聽所接收的輸入訊息為聲音線索 (acoustic cues) 或口語表述 (utterance)，使得學習者在運作各成分時，產生些微有別於閱讀歷程之差異。在初始解碼時，學習者會透過聲韻知識進行訊息解碼，例如子音、母音等音素、音節 (syllable)，進而藉由聲韻比對、重音線索、關聯訊息間激發 (spreading activation) 來搜尋詞彙意義，同時搭配語法結構進行命題意義理解。

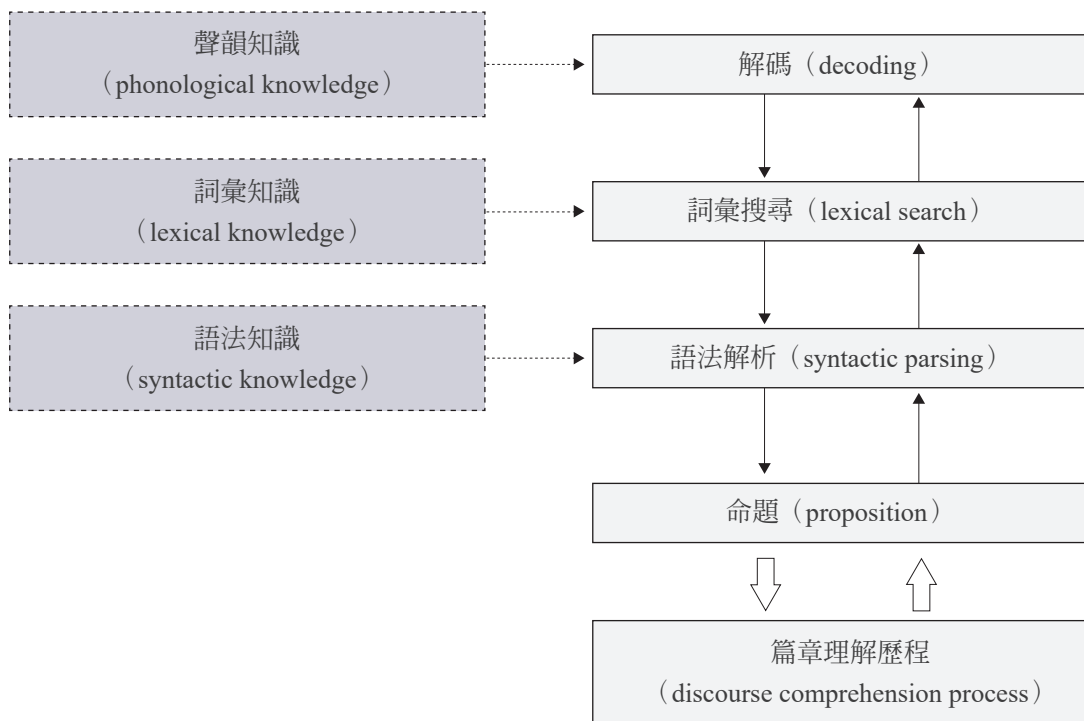
熟練學習者往往能透過有效識字、語法解析流暢性等，釋放出更多注意力資源 (attentional resource) 來運作其他複雜理解層次 (Samuels, 2013) 或透過知識補償 (Stanovich, 1980) 以促進理解。在上述基礎，轉化 108 英語文課程綱要內容上，本層次認知成分與學習表現對應，大致可條列如附錄一所示。就建構英語文基本知識而言，相關表現大多是學生得學習的重要基本元素，因此，各學習階段幾乎都有相對應的學習表現，其中，較低學習階段多重視基本的形素、音素習得，例如，1-II-1 能聽辨 26 個字母、1-II-2 能聽辨英語的子音、母音及其基本的組合、3-II-1 能辨識 26 個印刷體大小寫字母等。此外，隨著年級上升，課程綱要除了重視形素與音素對應學習之字母拼讀規則熟悉外，例如 5-II-4、5-III-9、5-III-10、5-IV-5 能運用字母拼讀規則讀出及拼寫英文字詞，也注重學生學習表現能透過聲音唸讀與字母外形，來形成詞彙的心理表徵，然而，其內容多尚未涉及詞彙意義辨識 (Snow et al., 1998)。後續，在習得表徵的技巧後，課綱對於不同年級學習者，也逐步加強常用詞彙量擴充與基本句型句子理解，例如，1-IV-1 能聽懂課堂中所學的字詞、1-IV-2 能聽懂常用的教室用語及日常生活用語、1-IV-3 能聽懂基本或重要句型的句子、IV-2 能辨識課堂中所學的字詞、3-IV-3 能看懂簡易的英文標示、3-IV-5 能看懂簡易的生活用語、3-IV-6 能看懂基本的句型等 (國家教育研究院, 2018a)。

圖 2-2 閱讀理解成分



註：引用與修訂自 Khalifa 與 Weir (2009, p. 43) 、OECD (2021, p. 42) 。

圖 2-3 聽力理解成分



註：引用與修訂自 Field (2013, p. 97)、OECD (2021, p. 48)。

## (二) 篇章理解認知成分

在前述層次，學習者形成理解的成分重視字母、聲韻與字彙組成，以及字彙、語法到單一句子建構，而學習者對此層次理解大致僅相當於訊息輸入表面層次 (surface level) (Kintsch, 1988, 2019) 或解碼、字彙理解層次 (Gagne et al., 1993)。然而，隨著句子累積形成具連貫意義的篇章時，學習者對於篇章之理解，會轉而受到命題意義 (propositional meaning) 與修辭格式 (rhetorical format) 或文本結構 (text grammar) 影響。雖然學者們對於篇章理解成分的區分各有其見解，但其層次大多彼此具有概念銜接與承續性，其中，尤以 Kintsch (1988, 2019) 所提文本模式 (text-based model)、情境模式 (situation model) 較受學者們關注。

另一方面，隨著數位科技興起，多元文本或數位訊息傳遞媒介也逐漸轉為理解議題探究焦點，學者提出如文件模式（document model）、任務模式（task model）、脈絡模式（content model）等概念（Britt & Rouet, 2012; Britt et al., 2013; Britt et al., 2018），來詮釋學習者在面臨多元文本或複雜訊息下，可能會產生衝突（conflicting information）、多元觀點（multi-perspectives）、社會文化之脈絡意義。

在上述模式基礎，本節將透過四個層次來說明篇章理解成分如下。

### 1. 直接找出明顯訊息

命題（proposition）為一有意義語言陳述的基本單位（idea unit）（註：本節所指命題，中文譯詞容易與試題撰寫意義雷同）。在特定文本結構下，文本作者會透過多個關係性命題（relational propositions）進行語意連結或結合，以組成複雜、連貫的篇章。因此，命題之意義建構可謂建立篇章字面意義理解的重要基礎（Kintsch & van Dijk, 1978; Mann & Thompson, 1986）。舉例來說，*I am hungry. Let's go to the Fuji Gardens.*，其中，「命題」為可以判斷其正確與否的語言意義單位，例如，*I*、*am hungry* 等主詞與述詞組成一意義單位（Mann & Thompson, 1986, p. 60）。其中，上述兩個句子中之命題具有 Mann 與 Thompson（1986）所界定的問題與解決（problem and solution）連貫關係，後面命題適度回應前面命題意義。

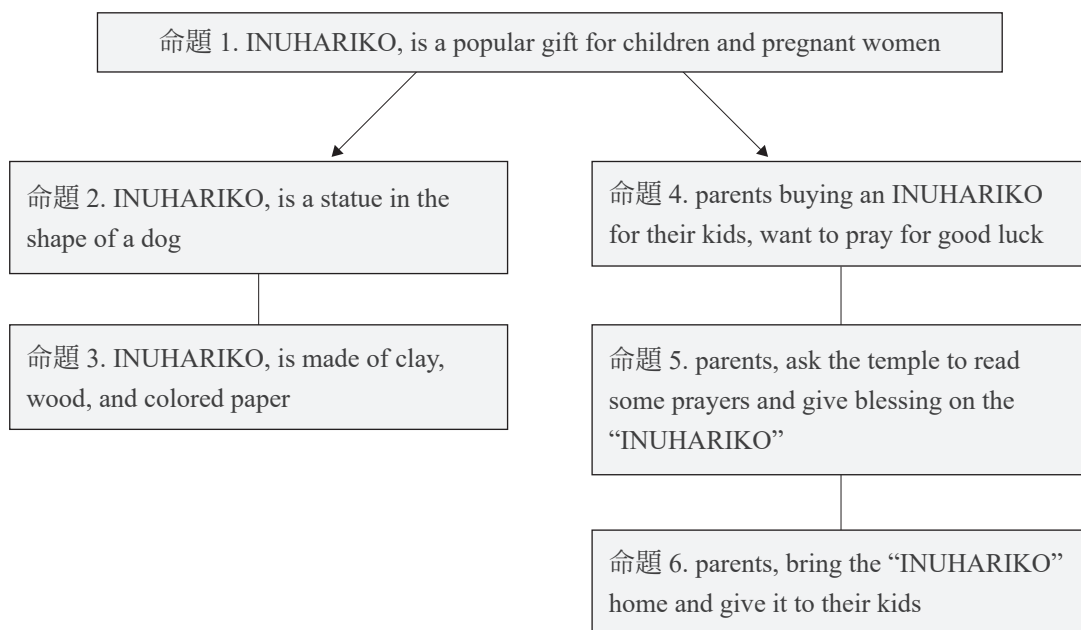
本書例舉附錄六 L134 進行說明，部分內容如下所示（註：更多具體範例說明可參考謝進昌（2022a），同時，本節所舉範例也曾同步刊載於該書，但是以本書為主要理論與發展基礎）。以首句為例，本書擷取出的第一個命題為「*INUHARIKO, is a popular gift for children and pregnant women*」，是由主詞「*INUHARIKO*」與述語「*is a popular gift for children and pregnant women*」組成，形成一個可判讀其正確或錯誤之意義基本單位。

*In Japan, INUHARIKO is a popular gift for children and pregnant women. INUHARIKO is a statue in the shape of a dog, and it's usually made of clay, wood, and colored paper. When parents want to pray for good luck, they can buy an INUHARIKO for their kids and ask the*

temple to read some prayers and give blessing on the “INUHARIKO”. Then, they would bring the “INUHARIKO” home and give it to their kids. (附錄六，L134)

依上述概念，本書進一步擷取文內各個命題，在經過組織，形成如下列各命題關係概念圖 2-4。命題 1 是「INUHARIKO, is a popular gift for children and pregnant women」，命題 2 為「INUHARIKO, is a statue in the shape of a dog」，而命題 3 是「INUHARIKO, is made of clay, wood, and colored paper」。其中，命題 2、3 補充說明 INUHARIKO 的外形特徵與製作材質；命題 4、5、6 為說明家長購買 INUHARIKO 的動機與行動。

圖 2-4 L134 各命題間關係概念圖



在本層次，可否「直接找出文中明顯訊息」涉及學習者能否從文中「直接找出、擷取出」明顯陳述的訊息（或命題意）來回應題目。若以下列 L134 第 1 題 What is an “INUHARIKO”？為例，學習者經直接擷取「命題 1 和 2 的意義或訊息」來回應題目，就可得到正確答案為 INUHARIKO is a dog-like statue for children。整體而言，學習者除了具備基本語言知識外，只要掌握篇章中明示（explicit）命題意義及理解其（局部）（local）關係，就能順利擷取出正確訊息。

What is an “INUHARIKO”？

- (1) A dog-like statue for children. \*
- (2) A cat-like statue for rich people.
- (3) A kid-like statue for pregnant women.

（附錄六，L134，第 1 題）

對於這類字面清楚陳述之訊息的擷取（locate explicitly stated information），幾乎是所有語言學習者平時生活常見的任務。然而，不同能力學習者會隨著文本或任務複雜度差異而展現不同程度擷取與辨識區別能力（Leu et al., 2013）。成功的學習者大致須具備能立即或自動理解明顯陳述，且不需要過多推論、解釋之關鍵訊息。此成分大致需要學習者聚焦在文本模式（textbase）中微觀結構（microstructure）內微觀命題（micropropositions）意義理解（Kintsch, 1988, 2019; Kintsch & van Dijk, 1978; Perfetti & Adlof, 2012）。有關評量學習者直接找出訊息能力如附錄一所示。相關評量指標包含有「直接找出人事物、時間、地點、物品等外在顯著特徵」、「直接找出明顯陳述的想法」、「直接找出與特定目標有關的明顯訊息」等。此外，在擬訂各評量指標之動詞詞彙使用上，本書多採用「找出、擷取」等動詞（詞彙說明請見第五章第三節）來界定或突顯本認知成分的特色。

## 2. 能理解字面文意

當學習者能連結多個微觀命題關係時，Kintsch 與 van Dijk（1978）、Kintsch（1988, 2019）認為他們所建構之篇章意義會形成更大單位命題，稱

為鉅觀命題 (macropropositions)。其表徵著一個更大、概化的 (generalized) 有意義語言集合單位或概念，以協助學生理解整體篇章文意。其中，不論是可直接找出訊息或是能理解文本字面意義 (literal comprehension)，這兩個層次理解成分多屬於 Kintsch 所稱文本模式 (textbase model)，而兩者的不同點往往得視命題理解是存在於微觀結構或鉅觀結構中。

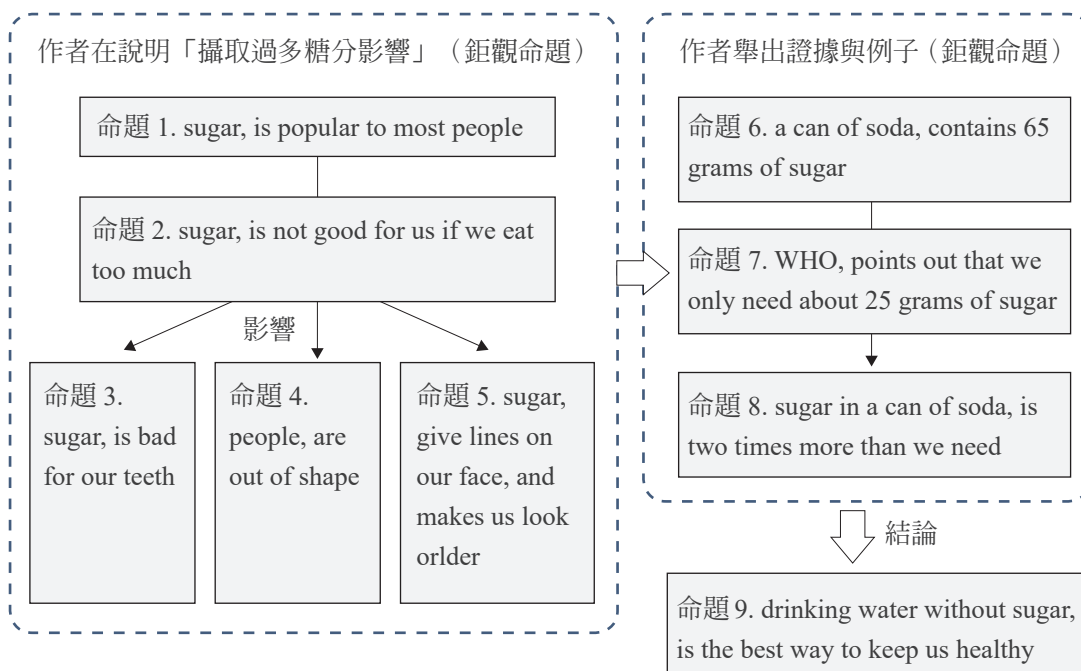
能理解字面文意係指學習者能理解文本字面清楚陳述的內容 (Pearson & Johnson, 1978)。在本層次，學者尤其重視學生能透過多個關係性命題意義的連結 (connect)，來建立整體 (global) 連貫性，以形成字面要點和大意之理解。然而，由於篇章中某些命題關係可能存在隱含、暗示性命題 (implicit propositions)，代表著學習者在建立語意連貫間，存在某些缺失訊息，須待學習者透過既有背景知識或經驗來重新建構其意義。以前述 L134 的命題 1 與命題 4、5 為例，兩者關係就暗示著 Why is “INUHARIKO” a popular gift for pregnant women? 的原因，其隱含著 Parents believe INUHARIKO can protect their children. 或 INUHARIKO can give good luck or blessing to their children. 等意義。

這類暗示性隱含大多屬於 Perkins 等人 (1983) 所界定之常識性知識 (makes sense epistemology)，而學習者並不需要投入太多認知空間，可直接透過既有知識、經驗進行直觀理解或推論。這概念與 Kintsch (1998) 認為擷取式過程 (retrieval process) 雷同，其中，學習者只要透過橋接式推論 (bridging inference)，直接連結兩或多個命題意關係，即能理解文意。整體而言，學習者若理解字面文意，除了能結合多個明顯陳述訊息以形成大意，更能進行直接推論 (make simple inferences) 得出表層中未清楚陳述 (常識性暗示) 之文本意涵，以達到整體字面意義理解 (West & Stanovich, 2000)。

本書例舉附錄七 L128 進行說明。依照前述概念，本書可將篇章內容拆解成命題與其關係概念圖 2-5。圖中各命題屬於微觀命題，其中，本書若連結命題 1 至 5 意義，大致可直接推論出「作者在說明攝取過多糖分的影響」，可視為一概化的鉅觀命題。同理，本書若連結命題 6 至 8，也可直接推論出「文本作者舉出證據與例子來說明」，也是另一鉅觀命題。若進一步再

連結這兩個鉅觀命題與結論關係，就可以建立出整體（global）連貫之字面大意。

圖 2-5 L128 各命題間關係概念圖



以下列 L128 第 2 題而言，學習者只要從命題 9 的意義或全文大意中，連結自身既有常識或背景知識判斷，就可以直接推論出正確答案為 Iced water。此外，就 L128 第 1 題而言，學習者橋接或連結兩個鉅觀命題關係，就可以進一步理解本文大意在於說明「It is bad to eat too much sugar.」的事理。

Tina agrees with the speaker. What kind of drink may Tina have every day?

- (1) Chocolate milk.
- (2) Apple juice.
- (3) Iced water. \*

（附錄七，L128，第 2 題）

What is the video mainly about?

- (1) It is important to take care of our teeth.
- (2) It is popular to drink soda every day.
- (3) It is bad to eat too much sugar. \*

(附錄七，L128，第 1 題)

就線上閱讀環境而言，有關學習者對於相關數位文本理解，學者認為由於數位網頁情境通常具有跳躍式、不具固定文本架構等特徵，往往會增加學習者認知負荷，並多半會提升學生理解困難。其中，若相對應本層次，最常見的評量指標為「能從搜尋引擎檢索結果，找到適當連結並做出推論」（Henry, 2006）。整體而言，在建立評量指標上，本書使用「連結、結合、直接推論」等動詞詞彙來界定或突顯學生表現是透過多個關係性命題意義來結合，以形成整體理解過程（詞彙說明請見第五章第三節）。在此基礎，如附錄一所示，本層次評量指標大致包含「從上下文，直接推論出特定字詞、句子意義或主詞與代名詞關聯」、「直接推論文本文字面大意與重點」、「直接推論出或直接比較具體訊息間明顯關聯」、「直接描述或蒐集文本明顯陳述的訊息、想法或特徵」、「就關鍵字檢索出（數位網頁）結果，直接推論出適當數位連結標的」等。

在回應 108 英語文課程綱要時，其相關指標包含能找出訊息、理解字面文意，多是一般學習者在接收篇章訊息時，須熟悉的基本課題，而其對應的表現內容如附錄一所示，大致包含 1-IV-4 能聽懂日常生活對話的主要內容、1-IV-6 能聽懂簡易故事及短劇的主要內容、1-IV-8 能聽懂簡易影片的主要內容、1-IV-11 能聽懂公共場所廣播的內容（如捷運、車站、機場廣播）、3-IV-8 能了解短文、簡訊、書信的主要內容、3-IV-9 能了解故事的主要內容與情節、3-IV-10 能辨識簡易故事的要素（如背景、人物、事件和結局）、3-IV-13 能了解短劇的主要內容與情節等（國家教育研究院，2018a）。

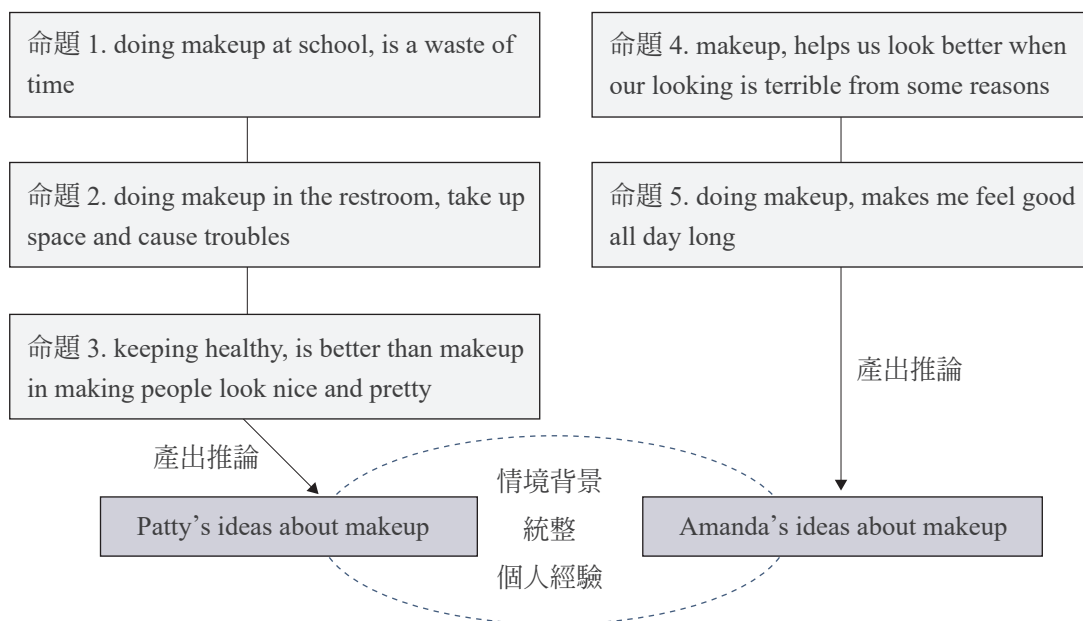
### 3. 能理解隱含文意

文本理解是學習者對於文本內容建構心理表徵（mental representation）的過程，而在 Kintsch（1988, 1998, 2019）所提出建構統整模式（construction-integration model），認為學習者是以命題（proposition）來分析文本，而在其執行理解的過程，學習者會以不斷執行新命題的建構及參照式推論，如此循環、交錯，直到學習者了解文本意涵為止。因此，學習者在進行理解時，除了理解文本內容外，還會融入自身背景經驗及知識，對於文本內容進行更深層推論與精緻化，而此便稱為情境模式（situation model）（Kintsch, 1988, 1998, 2019）。

在情境模式中，學習者對於文本之理解往往跳脫文本字面明顯訊息，其歷程的運作會涉及複雜推論。學習者得連結其背景知識與經驗，來產出推論（generate inference）（Graesser et al., 1994）或精緻化推論（elaborate inference），例如，因果、預測性推論等，係經過統整相關想法與訊息（integrate ideas and information）後，以理解文本內隱含關聯與其深層抽象意涵（figurative meaning）。自這個層次開始至下一認知層次，學習者所需推論知識，會逐步涉及 Perkins 等人（1983）所指的批判式知識（critical epistemology）。學習者得逐漸融入如邏輯的推理、系統性規則、科學實徵等類型命題知識，思考文本內未說明之其他相似或抽象概念、假設或衝突訊息等，以進行文本暗示（或隱含）（implicit）意涵之解讀、推論。

本書例舉附錄八 R018 為例，而各命題間關係概念如圖 2-6 所示。其中，圖中命題多為經過精緻化之相對鉅觀命題，是略有別於前述 L128 與 L134 字面可直接擷取之微觀命題。本書若透過命題 1 至 3，可推論、產出 Patty 對於化妝品的想法，是秉持著反對使用化妝品的立場；此外，本書若透過命題 4 至 5，同理可推論、產出 Amanda 對於化妝品的想法，是秉持著贊成使用化妝品的立場。在此基礎上，如同 R018 第 1 題所示，學習者經統整 Patty、Amanda 想法與立場後，連結個人經驗與掌握篇章情境訊息（如社交軟體訊息交流），就可以推論出本文主旨在 Sharing ideas about makeup.。

圖 2-6 R018 產出推論與統整命題，以形成主旨



What are Patty and Amanda talking about?

- (1) Doing makeup is important.
- (2) Telling stories about makeup.
- (3) Sharing ideas about makeup. \*
- (4) Changing our lives with makeup.

(附錄八，R018，第 1 題)

有關評量指標，如附錄一所示。本書使用「統整、推論、應用」等動詞詞彙來界定或突顯學生理解得結合自背景知識，經推論與統整全體文本意涵與理解過程（詞彙說明請見第五章第三節）。相關評量指標大致包含「統整文本要點以推論出主旨、共通性原則」、「推論出文本抽象、隱含的意涵或片段訊息間關聯」、「解釋文本抽象、隱含的意涵或片段訊息間關聯並從文內提出相對支持證據」、「應用或說明文本概念、通則或規則（至其他情境），以得出結果」、「比較文本複雜訊息或片段、抽象的想

法、概念之相似或相異性」。另一方面，考量本層次評量文本往往會涉及較複雜、多元訊息、甚至是語文表達手法，而對於這類上位概念的理解，學習者必須從整體上位組織架構著手，才得以理解，是有其理解難度，例如，「區辨出文本結構或寫作手法等語文特徵，如比喻」，因此，在本書，這類指標也列為本層次評量重點。

在回應 108 英語文課程綱要方面，如附錄一所示，本層次強調學生不僅須理解文本字面文意，更注重能否理解深層隱含文意，使其呼應學習表現，除了前一層次指標外，另外包含如 1-IV-7 能辨識簡短說明或敘述的情境及主旨、3-IV-11 能藉圖畫、標題書名等作合理的猜測、3-IV-12 能熟悉重要的閱讀技巧，如擷取大意、猜測字義、推敲文意、預測後續文意及情節發展等涉及高層次推論或猜測、主旨統整等層面（國家教育研究院，2018a）。

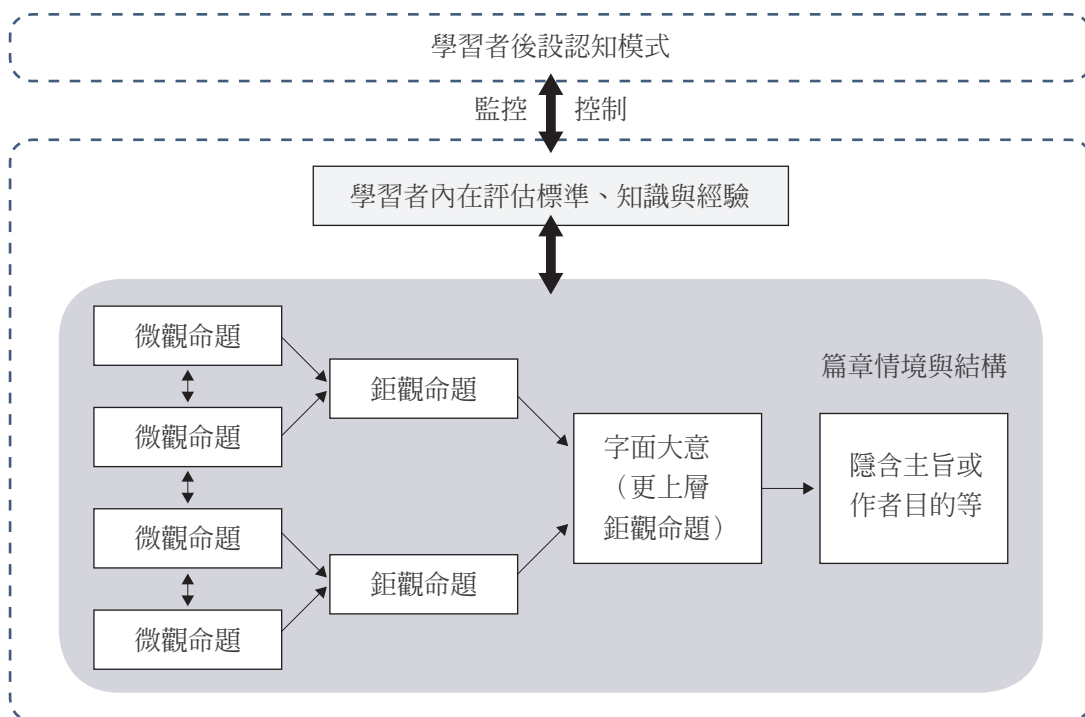
#### 4. 在文本理解上，能透過自身觀點進行評估與省思進而認識自我

在前述四個層面，相關概念皆聚焦在文本內（within text）理解，直接呼應 Kintsch（1988, 1998, 2019）所述文本與情境模式。學習者得透過相關理解成分，來掌握原來文本作者所欲表達的字面或隱含內容。在本層次，相關概念重視能在既有文本理解基礎下，超越文本理解（beyond text comprehension），學習者能運用連結背景知識，透過自身觀點（perspective），來詮釋文本作者的想法或立場，進而，評估、省思文本內容或形式（Wiggins & McTighe, 2005）。

在回顧認知理解之理論發展，本層次大致涉及學習者理解監控（comprehension monitor）與自我調節（self-regulated comprehension）層面。就理解監控之發展而言，Baker（1979, 2005）、Baker 與 Brown（1984）認為理解監控涉及學習者對於自身理解過程的評估（evaluation）與調節（regulation），前者係指學習者可對自身理解進行正確、有效性之確認，而調節在於透過校正、調整等程序以確保理解的有效性與持續運作，以達到重新建構針對文本意義理解與發展自身理解的詮釋。另一方面，Hacker（2004）認為自我調節理解（self-regulated comprehension）屬於認知與後設認知模式同時運作之機制。其中，如圖 2-7 所示，學習者在認知層次中，會存在一個既有理解架構，例如 Kintsch（1988）所述文本模式、情境模式與既有評估

準則（standard for evaluation），而學習者會透過這些既有架構，提供後設認知模式對於既有認知概況，來進行監控或調節、控制（control），使得整體理解過程會呈現出一遞迴式認知 後設認知交互往來、運作的歷程。

圖 2-7 自我調節理解歷程示意圖



註：修訂自 Hacker (2004, p. 169)。

省思與評估（reflection and evaluation）為檢視學習者在文本以外的知識與經驗，是否能運用與連結背景知識，以反思文本內容、形式。相關元素涉及如理解文本作者之目的、觀點、甚至是寫作風格與手法；此外，隨著訊息傳遞途徑的多元化，評估文本內容可信度、品質及面對不一致訊息衝突時的因應，也逐漸成為進階學習者所必須擁有的能力（OECD, 2019）。這概念也大致呼應 108 英語文課程綱要所新增之邏輯思考、判斷與創造力向度，目的為培養學生透過語言進行高層次思考，例如，9-V-6 能綜合現有訊息，預測可能的發展、9-V-7 能評估不同資訊，提出合理的判斷或建議（國家教育

研究院，2018a）。另一方面，隨著線上數位環境漸趨普及下，訊息除了更為複雜、多元外，也更容易存在偏誤，使得線上文本訊息評估能力重要性更甚以往。其中，同一主題的不同文本，隨著來源、時間差異，都可能產生不一致結果或論點，因此，有競爭力的學習者必須能有效運用其評估能力，掌握訊息相關性、可信度，以做出預測性（predictive）、評估性判斷（evaluative judgment）（Rich, 2002）或適時因應衝突訊息（Bråten et al., 2009; Bråten et al., 2011）。進而，在超出情境模式基礎下，學習者得應用額外且更複雜的心理表徵來因應線上閱讀環境的挑戰。

有關範例試題，本書例舉附錄八 R018 第 3 題為例。如下所示，該試題評量重點在施測於學習者是否能用自身理解，經分析、評估 Patty 或 Amanda 對於化妝品使用觀點，以提出證據來支持自身的立場。在實際作答歷程中，學習者不僅得具備字面與隱含文意理解，還得交互遞迴監控、控制確認自己理解正確性、及其與自身觀點適配性等，以提出自身立場說明。

What do you think about doing makeup? Use Patty's or Amanda's idea to write down what you think. (請盡量以英文回答，完成下列句子)

參考句型：I agree with \_\_\_\_\_ (Patty or Amanda). I think

\_\_\_\_\_.

(附錄八，R018，第 3 題)

相關的評量指標，如附錄一所示。本書使用「分析、評估、省思」等動詞詞彙來界定或突顯學習者在理解文本基礎下，透過自身觀點進行評估或省思之過程（詞彙說明請見第五章第三節）。整體而言，評量指標大致包含「分析、評估文本論點、觀點合理性等，並提出自身立場或觀點說明」、「說明作者目的、使用的文本寫作風格、手法或語文特徵、結構」、「省思作者論點、觀點或文中寓意」與「融合運用各理解歷程成分，進行探究實作以解決問題」。其中，本書認為探究實作是融合了各理解歷程成分的運用，才得以進行，而相關推展往往得仰賴實作評量（performance assessment），才足以突顯與評估學生表現結果。然而，此點在大型評量模式下，往往容易受限於成本，較難完成，因此，本書僅為求完整性列出該指標，但尚未推動與蒐集相關學生表現結果。

## 二、各大型評量理解認知成分概念銜接與比較

有關理解認知成分或稱為思考歷程，學者通常以閱讀理解（reading comprehension）來稱呼，然而，相關概念並不僅限於閱讀，更通用於聽力或其他媒介傳遞之訊息（Kintsch & van Dijk, 1978）。根據 Snow 在主持 RRSB（2002）時提及對於閱讀理解內涵的想法，他們認為閱讀理解本身不僅是一項學習課程，也是一項學習其他知識或獲取其他能力的途徑。簡單來說，它即是「學習如何透過英語文去理解」（learn to read）與「學習如何透過英語文的理解，再去學習其他領域內容」（read to learn）。在此基礎下，許多國際大型評量計畫目的，多透過類似概念來進行學生閱讀評量設計與發展。

在回顧過去國際大型評量閱讀理解歷程，許多學者已經針對彼此內涵進行內容與概念比較，例如，Binkley 與 Kelly（2003）比較 NAEP 2002 與 PIRLS 2001 四年級學生閱讀理解歷程、International Activities Program（2010）比較 NAEP 2009 與 PISA 2009 閱讀評量架構、Mullis 等人（2006, p. 101）比較 PIRLS 2006 與 PISA 閱讀理解成分相似性。整體結果也大多顯示彼此間認知成分與內涵，係分享著相似發展概念。在此基礎下，本書企圖透過另一種途徑來詮釋彼此間相似性，而其概念大致可整理如表 2-1、2-2 所示。表中內容可發現各國際閱讀理解評量所提出之理解成分或歷程內容雖然具些許差異性，然而，若是以「詞類」來進行檢視，可發現相關評比皆重視「動詞或搭配副詞」的使用，來突顯出不同層次之理解成分差異，例如，使用「找出」來說明學習者對於篇章明顯訊息的擷取，而使用「詮釋、統整、推論」來說明學習者對於篇章字面文意或隱含文意的理解。這個概念就如同 Bloom 等人（1956）在發展教育目標分類時，使用「動詞」來詮釋認知向度（Anderson et al., 2001），認為「動詞」的使用會反映或描繪出人類如何（How）思考與思維歷程。

在檢視不同國際評比閱讀理解成分時，結果如表 2-1 所示。PIRLS 2016 將理解成分區分為四個層次，其中，第一層次的「聚焦與直接找出文中明確陳述的訊息」（Mullis & Martin, 2015），所使用的動詞詞彙為「聚焦與直接找出」，來表徵學生能找出文中明顯陳述的訊息，而其概念是類似於 PISA 2018 閱讀的「找出訊息、或訊息檢索與擷取」（OECD, 2019）與美國 NAEP 2019 閱讀的「指出、回憶」（National Assessment Governing Board [NAGB], 2019）；此外，若進一步往高層次理解成分進行比對，可發現 PIRLS 2016 使用「詮釋、整合、推斷、運用、建構、比較、評估」

等來說明隱含文意理解、與評估省思歷程。其中，若對照 PISA 2018 閱讀架構所使用的「統整、產生推論、省思與評估」與美國 NAEP 2019 閱讀評量架構所使用的「統整與解釋、複雜推論、批判與評估」等詞彙，彼此間也大多具有類似概念。另一方面，在比較外語與母語閱讀理解成分而言，如表 2-2 所示，OECD (2021, pp. 45-46) 提出 PISA 2025 外語評量與 PISA 2018 母語閱讀理解成分對應表。此對應表正呼應學者認為理解認知成分具有跨語言可轉換性，學習者能透過其原始習得母語閱讀技巧，轉換至外語閱讀脈絡中 (Cummins, 1979, 1991)。其中，在流暢運作語言基本知能上，PISA 2025 外語評量所涉及的識字、詞彙觸接與語法解析 (Khalifa & Weir, 2009)，就相當於 PISA 2018 所評量之閱讀流暢性，重視學習者的字彙與句子建構知識掌握力。此外，若再往更高層次理解成分進行檢視，PISA 2025 外語評量對於篇章理解成分，是為形成命題意與推論、建構心智模式、文本內或文本間表徵，而此概念就相當於 PISA 2018 對於學生篇章理解成分的界定，能理解字面文意並經統整、推論以理解全文內涵 (OECD, 2019)。

在上述論述中，可發現不同國際評比間閱讀理解成分雖然具有概念一致性或互通性，然而，由於學習者的理解歷程多屬於一個連續狀態，為評量與說明方便，將其類別化，使得各層次界線間確是存在相當模糊的區別空間，例如，PIRLS 2016 將其區分為直接找出、字面意理解（直接推論）兩個層次 (Mullis & Martin, 2015)，而 NAEP 則是將這兩個層次合併為指出與回憶，同時融入能辨識與直接推論的能力 (NAGB, 2019)。除了透過動詞使用來表徵理解成分外，各國際評比基於不同目的或立場，使用的文本類型也多具備些許差異性，使得學生所應理解的「內容」，也會有所不同。相關概念若依據 Bloom 等人 (1956) 建議，可透過「名詞使用」來界定 (Anderson et al., 2001)，對此，本書將在第二章第三節進行說明。

表 2-1 PIRLS、PISA 與 NAEP 閱讀理解認知成分

理解認知成分	PIRLS 2016	PISA 2018	NAEP 2019
能流暢運作語言基本知能	--	閱讀流暢性 (reading fluency) 準確與自動閱讀字彙與文本 / 為理解單一文本整體意義所進行字彙與文本運作歷程—	-- 字彙評量融入篇章中 (NAGB, 2019, p. 13)
能直接找出明顯訊息	聚焦與直接找出文中明確陳述的訊息	找出訊息— 針對單一文本內容進行訊息檢索與擷取 / 搜尋與選擇相關資料	指出 / 回憶 (locate/recall) 能辨識文本明顯訊息與直接推論文本內與文本間訊息
能理解字面意	透過文內陳述的訊息，直接進行推論	理解— 能表徵字面文意與統整、產生推論	
能理解隱含意	詮釋整合段落內或段落間訊息，推斷出訊息間關係、運用自己知識去理解與建構文章中細節及更完整意思	評估與省思— 連結背景知識省思與評估文本內容、格式，並評估文本品質與可信度，及覺察與處理衝突	統整與解釋 (integrate/interpret) 能針對文本內、文本間訊息進行複雜推論
在文本理解上，能透過自身觀點進行評估			批判與評估 (critique/evaluate) 能就文本內容進行批判性思考
能省思認識自我	比較與評估文本中訊息可信度、作者觀點等		

註：1. 彙整自 Mullis 與 Martin (2015)、NAGB (2019)、OECD (2019)。

2. 本表內容表徵各國際評比間閱讀理解歷程具有概念一致性，並非代表本書評量架構是直接引用自任一評比。
3. 虛線表徵各理解成分是一連續且區分界線存在相當模糊空間。因此，理解歷程可依照評量工具發展者目的與需求，進行拆解或合併以形成各成分。

表 2-2 PISA 2018 閱讀與 PISA 2025 FLA 外語評量閱讀認知成分對應

理解認知成分	PISA 2025 外語評量	PISA 2018
能流暢運作語言基本知能	識字 詞彙觸接 語法解析	閱讀流暢性（reading fluency）－ 準確與自動閱讀字彙與文本／為 理解單一文本整體意義所進行字 彙與文本運作歷程
能直接找出訊息	融入其他層次進行評 量	找出訊息－ 針對單一文本內容進行訊息檢索 與擷取／搜尋與選擇相關資料
能理解字面意	建立命題意	理解－ 能表徵字面文意與統整、產生推 論
能理解隱含意	推論 建立心智模式 建立文本層次表徵 建立文本間表徵	
在文本理解上，能透 過自身觀點進行評估	非本次外語評量焦點	評估與省思－ 連結背景知識省思與評估文本內 容、格式，並評估文本品質與可 信度，及覺察與處理衝突
能省思認識自我		

註：1. 引用自 OECD（2021, pp. 45-46）。

2. 虛線表徵各理解成分具連續性且區分界線存在相當模糊空間。因此，理解歷程可依照評量工具發展者目的與需求，進行拆解或合併以形成各成分。
3. 本表內容表徵各國際評比間閱讀理解歷程具有概念一致性，並非代表本書評量架構直接引用自任一評比。此外，PISA 2025 架構說明是於 2021 年釋出，本書是於原研究評量架構建立後回溯，以期完整說明內容。

### 第三節 文本類型與接收性語言活動

文本 (text) 係指稱任何由英語文傳遞的訊息或素材，為通用性詞彙 (OECD, 2021)，而篇章 (discourse) 係指在特定情境脈絡下，由閱讀文字或口語表述所組成的內容。相關特徵往往具備如，社會情境脈絡、較完整字義與語意連貫性、具備訊息傳遞性或隱含作者目的等 (de Beaugrande & Dressler, 1981; Lee, 2001; Östman & Virtanen, 2011)。文本因素或特徵 (text features) 包含許多面向，而最主要的特徵為文本類型 (text types)。

文本類型界定大致可呼應 108 英語文課程綱要所述「學習內容、體裁」，然而，在考量現行英語文課程綱要並未具體、明確界定文本類型下，加諸學者對於文本類型的區分，從修辭 (rhetoric)、語言與結構特徵 (linguistic/structural features)、脈絡與功能 (context and function) 等面向切入，分類十分多元 (Biber, 1989; Lee, 2001; Östman & Virtanen, 2011; Werlich, 1976)，因此，使得現今並未有統一文本分類方式。

就參照國際大型閱讀評量，評量文本具體分類也未有統一類群，例如，PIRLS 區分為說明文 (informational texts)、故事體 (literary texts) (註：中文譯名來自 PIRLS 臺灣國家報告)，其中，前者大致指稱以傳遞訊息為目的之文本，包含如說明文、議論文等，而後者係指敘事類小說 (narrative fiction) (柯華葳等人，2017)。其中，PIRLS 因考量各個參與國家課程與文化差異，較難納入如詩、劇本等文本；此外，NAEP 主要將文本區分為兩類，一為文學類文本 (literary texts)，包含小說 (fiction)、非小說文學 (literary nonfiction) 和詩 (poetry)，另一類為訊息類文本 (informational texts)，包含說明文 (exposition)、議論與說服文本 (argumentation and persuasive text)、過程類文本與文件 (procedural text and documents) (NAGB, 2019)。

本書在考量 TASAL 英語文素養評量目的與實務需求等因素，認為 TASAL 英語文評量文本分類，一是得呼應國家課程綱要內涵，重視英語文實用功能與社會文化、文學內涵；此外，由於不同文本類型多隱含不同作者目的、語言結構、內容或寫作風格，使得文本分類類型不宜過於單一，以影響評量試題發展的廣度，加諸，文本分類命名得兼顧命題者認同與可理解性，以利於溝通與傳播。在此基礎上，本書界定出一加六類文本。第一類為評量學習者語言知能文本，由於這類文本不似篇

章有連貫語意與情境脈絡，單純為評量學生基本的「語言知識」而設計，因此，本書將此類文本獨立出來。此外，其餘六類是參考 OECD（2019, pp. 46-48）所界定的 PISA 2018 閱讀六種文本類型，而此分類同時也是 OECD（2021）PISA 2025 外語評量（FLA）所使用的文本分類，包含描述類文本（descriptive texts）、教學指引類文本（instructional texts）、人際互動類文本（transactional texts）、說明文（expository texts）、議論文（argumentative texts）、敘事與文學（narrative and literary texts）。其中，為便於解釋，本書是以類群方式來進行各文本類型說明。

## 一、語言知能文本

篇章的理解往往得銜接在社會情境脈絡下進行解讀，才會具有較完整語意或字句連貫性，同時，帶有訊息傳遞性、作者目的或情境等（de Beaugrande & Dressler, 1981）。然而，考量語言知能文本並不具備篇章（discourse）特徵。因此，本書在設計評量試題時，將這類文本獨立出來。本類文本目的在於評量學生的基本英語文知識程度，相關試題特徵大多是由簡短、或基本的形素、音素、音節、字彙（詞）、單一句子或圖片所組成。

現行英語文國家課程綱要是將英語文學習界定為外語學習（國家教育研究院，2018a），而學生往往得在具備相對成熟理解認知發展後，才開始學習，使得語言知能文本在學校小考或段考評量試題中相當常見，多用於評量學生對英語文基本知能的流暢性與熟悉度。然而，本書考量這類文本所延伸之評量命題，因容易受到學校所選用教科書差異影響，其試題難度會直接視相關字彙或句型有教過或沒教過而有別，因此，在選材上，命題教師多會選擇基本且適合全臺灣該年段學生為考量，例如，若以評量七年級學生，會選擇完成國小階段學習應具備字彙或基本句型。

在回應英語文國家課程綱要，如附錄二所示。語言知能文本大多源自於回應學習內容，例如，Ab-IV-3 字母拼讀規則、Ac-IV-1 簡易的英文標示、Ac-IV-2 常見的教室用語、Ac-IV-3 常見的生活用語、Ac-IV-4 國中階段所學字詞（能聽、讀、說、寫最基本的 1,200 字詞）等（國家教育研究院，2018a）。此外，若以「名詞使用」而言，本類文本內容大多涉及有關「字母、音素、字母拼讀、字詞（彙）、片語、句子」等；同時，進一步搭配前述第二章第二節有關理解認知歷程之動詞使用，例如，「辨識、認讀」，經結合兩者，就會形成附錄一各項評量指標（詳細轉化方式，請見第二章

第五節)。有關語言知能文本範例，請見附錄三，若以 L121 為例，本命題為評量學生能使用字母拼讀規則辨識出字詞：聽到語音 [ʃɪp]，能選出對應的字彙；而 L120 為評量學生聽到字彙，能選出對應的圖片。整體而言，語言知能文本單純為評量學生基本語言知識而設計，其篇幅較短且相對較不具備真實生活情境脈絡。

## 二、第一類群篇章

第一類群篇章之目的在於傳遞訊息 (to convey information) 與通知 (to inform)，而隨訊息型態差異，大致區分為描述類文本、教學指引類文本、人際互動類文本。

描述類文本目的在描述 (describe) 特定訊息，而內容往往是經過有系統與具組織方式形成 (Berge et al., 2016)。描述類文本訊息多半會涉及物件空間中的屬性 (properties of object in space)，而相關型態可以客觀的訊息內容來呈現，例如景點、物品等外在特徵 (characteristic)、功能 (function)，或者以主觀的描述來呈現，例如，空間中物品關聯、品質等。有關描述類文本，相關篇章的評量試題內容大多傾向於設計在回應「什麼」(what) 之題項 (OECD, 2019)，以便於評量學生是否能理解相關訊息的描述內容。同時，在對應英語文課程綱要建議體裁 (或 OECD 所稱物件) 概念上，這類文本大致可呼應如地圖、菜單、時刻表等 (國家教育研究院, 2018a)。有關描述類文本範例，內容可見附錄四。若以 R071 篇章為例，該文本內容為描述藥品特徵、用藥時間等，此外，L153 則以聽力型式來描述公車出發時間。

教學指引類文本之目的在於提供指引 (guideline) 與方向 (orientation)，教導如何做或進行的過程或程序性 (procedural) 訊息，而內容往往也經過系統組織 (如條列) 所形成。教學指引類文本的訊息型態可以客觀的任務活動內容來呈現，例如，食譜、科學實驗程序、遊戲規則、程序性操作手冊等、抑或是以主觀的人類行為描述來呈現，例如，活出幸福的人生守則。本類篇章所延伸評量試題內容除了具備「什麼」(what) 的訊息，往往還具備「做什麼、怎麼做」(what to do or how to do something) 之相關導引或指引方向的訊息 (OECD, 2019)，以便於評量學生是否能理解相關程序或指引的訊息內容。另一方面，在對應英語文課程綱要所建議體裁，就概念層面而言，可條列如附錄二所示。這類文本大致呼應食譜、公告或學習內容所界定，如 Ae-IV-3 公共場所廣播 (如捷運、車站、機場廣播)、Ae-IV-2 常見的圖表 (國家教育研究院, 2018a)。其中，圖表雖然常見於程序性文本，然而，在考量其他如

說明或議論文本中，圖示或表格也具有簡要、具體呈現資料數據或證據的功能，因此，本書暫時將其條列在教學指引類群文本，然而，此舉並不代表在實際評量時，圖表僅會出現在這類文本。最後，有關教學指引類文本範例，詳細內容請見附錄五。其中，R065 篇章為描述紙卡片製作的程序步驟，而相關內容除了有過程性訊息描述外，另整體架構是以有組織、條列式步驟程序，來導引讀者了解卡片的製作程序。

人際互動類文本係為達到某個特定目的，所進行個人與個人、或團體間之訊息交流（exchange）與互動（interaction）。尤其在隨著數位科技進步，例如，社交軟體的廣泛運用，這類文本在日常生活中更顯得普遍（Berge et al., 2016）。人際互動類文本的訊息型態可以單純傳遞客觀訊息內容來呈現，例如約定兩方的時間、或以傳遞情感為目的所建立、促進人際關係之相關日常生活對話或交流內容等（OECD, 2019）。此外，凡日常生活中常見的廣告、傳單等文本，有鑑於該類文本某程度是在吸引群眾注意力目的下，所進行個人與群體間訊息傳遞，以期與群眾建立初步關係或進行後續參與、購買等互動回應，因此，本書將其界定為人際互動類文本。在對應英語文課程綱要體裁上，就概念層面而言，相關結果如附錄二所示。這類文本大致可呼應信件、會話、廣告或學習內容所界定，如 Ae-IV-4 簡易賀卡、書信、電子郵件（國家教育研究院，2018a）。有關人際互動類文本範例，詳細內容請見附錄六。其中，L132 以聽力型式來呈現失物招領處的兩人對談內容、L134 為朋友間生活對談；此外，R115 以閱讀型式來呈現在氣候變遷背景下，作者所提出的雙贏策略，然而，其原始出發點則為廣告推銷。最後，R142 為父子間的書信內容，富有情感交流意味。

### 三、第二類群篇章

第二類群篇章除了具備第一類群篇章之訊息傳遞與知會（to inform）目的外，相關文本與評量試題設計方向，往往隱含著作者想進一步解釋（to explain）、探索（to explore）、說服（to persuade or convince）等目的。相關篇章包含說明文、議論文。前者內容多涉及解釋事物如何運作、探索現象發生理由等；後者內容多涉及在特定立場下，作者提出證據、論點等，來表達其觀點或主張，以尋求外部支持或說服他人。就概念層面，這兩類文本具有一定程度的相似性，然而，其間主要差異在於溝通目的。前者重視透過解釋的過程，從（社會）科學層面來累積或發展知識（to

develop knowledge)；後者重視透過相關論述過程，來吸引或影響他人 (to influence audience) (Andrews & Hertzberg, 2009; Berge et al., 2016)。

說明文目的在於透過(複合性)概念 (composite concepts) 或心理構念 (mental construct) 來解釋事物，藉以傳遞(科學或社會科學)知識或促進知識的理解為主旨。說明文的訊息有許多型態，例如，從自身觀點或經驗來解釋某特定概念、提供定義 (definition) 與解釋相關概念意涵、論述多個概念間或和其從屬元素間關聯等 (OECD, 2019)。此外，說明文的文本組織或架構大多具備邏輯連貫性，而內容可能涉及如推論的 (deductive)、統計的 (statistical)、因果的 (causal) 語言與情境脈絡之基礎內容，同時，相關陳述的科學現象多具有共識性且不受質疑 (Osborne & Patterson, 2011; Thagard, 2008)。有關說明文的範例，詳細內容請見附錄七。其中，R079 為藉由太陽東升西落、月滿圓缺與四季變化現象，來解釋自然現象的循環不止，而 R126 閱透過身體部位來說明生活中有趣的英語慣用語 (idioms)；此外，L128 以聽力型式藉由實例與專家建議，來闡述生活中糖分攝取的壞處。

議論文目的在於透過主張 (proposition)、聲明 (claim)、立場與觀點等呈現與論述，企圖吸引或說服他人。議論文相關訊息型態大多包含如對於某個事件的評論、思想與價值訊息的傳達等。此外，文本內容主題如果屬於科學性議論 (scientific argumentation)，其架構組織往往會透過如某個論述主張、證據 (evidence) 的建立，來支持某個知識體系，而相關結果多會依照某個約定過程來驗證其主張是否有效，例如，對氣候變遷辯證 (OECD, 2019)。因此，這類科學性議論文同樣具備邏輯連貫性，然而，相較於前述說明文，其論述的現象往往具有多種可能的解釋或觀點，故其結果只是暫時性結論 (Osborne & Patterson, 2011)，因此，本書也暫將其歸於議論文。有關議論文範例，詳細內容請見附錄八。其中，R018 為閱讀型式呈現兩人對於使用化妝品的評論及其相應支持的立場與觀點。

在評量試題發展方向上，OECD (2019) 認為說明文側重回應「如何」(how) 的試題，而議論文則側重回應「為什麼」(why) 的試題，以便於評量學生除了是理解相關訊息的描述外，對於隱含的目的、邏輯、因果等理解。另一方面，誠如前文所述，不同類型文本間界線的區分，係存在相當大的模糊空間，因此，本書認為兩類文本除了基本什麼 (what) 題項外，詢問「如何」與「為什麼」的題項，也相

當常見於彼此，例如，說明文可解釋現象發生原因（why 的層面）、議論文可辯證某現象如何形成（how 的層面）。在對應英語文課程綱要建議體裁，就概念層面而言，如附錄二所示。第二類群篇章文本大致可呼應學習內容所界定 Ae-IV-8 簡易故事及短文的大意（國家教育研究院，2018a）中的短文大意。

#### 四、第三類群篇章

第三類群篇章除了具備訊息傳遞的基本功能外，同時還隱含著作者為了啟發與愉樂（to entertain and delight）、吸取文學經驗（for literary experience）等目的而撰寫的文章。本類包含有敘事與文學。如同 OECD（2019, p. 48）所述，就寫作型態與架構層面而言，敘事（narrative text）與文學（literary text）二者間的區分界限，是愈來愈模糊，兩者皆可能涉及事實與虛構內容。基於此與為了說明與理解的便利性，本書將兩者視為同一類群文本，一併進行說明。

敘事與文學撰寫目的大多富有啟發與愉樂的意涵，而相關內容同樣具有特定架構與組織型態。然而，有別於描述類文本，敘事文本多半涉及物件「時間」中的屬性（properties of object in time），例如，某個事件何時（when）、以哪種時間順序關係（in what sequence）發生等，來呈現相關內容（OECD, 2019）。敘事文本擁有許多的型態，其中之一是透過敘事者的主觀觀點，來呈現某個事件、人物的記錄；另一種是透過報導者角度來呈現的新聞故事或事件，抑或是以小說、故事型態來呈現，多涉及有時間順序關係之主要、或從屬事件的組成（sequences of events）（Grabe, 2002, pp. 252-253）、情節（plot）與背景（setting）等（OECD, 2019; Rumelhart, 1975; Stein & Glenn, 1979）。這類文本的評量試題發展，除了設計基本「什麼」（what）訊息題項外，多會重視如「為什麼」（why）人物行為或行動原因，甚至藉由文本中所富含作者經驗、情感與思想，詢問讀者在理解內文後，所觸發的自我省思，進而促進自我理解或認識（self-insight, to have self knowledge）（Berge et al., 2016）。有關敘事與文學範例，詳細內容請見附錄九。其中，L086 以聽力型態為描述四個傳統節日，而內容主要涉及各個節日由來的歷史故事與意涵；L139 是以時間順序手法，來描述 mango 歷史發展由來；最後，R135 則是以富含友情與智慧內涵的極短故事內容來呈現。

在對應英語文課程綱要建議體裁上，就概念層面而言，結果如附錄二所示。敘事與文學類文本大致可呼應如故事、劇本、詩、歌曲等或學習內容所界定，如 Ae-

IV-1 簡易歌謠、韻文、短文、故事及短劇、Ae-IV-6 簡易故事的背景、人物、事件和結局、Ae-IV-7 敘述者的觀點、態度及寫作目的、Ae-IV-8 簡易故事及短文的大意（國家教育研究院，2018a）。

整體而言，對於文本分類，誠如 OECD（2019, p. 46）所述，不同文本類型間界線是存在相當模糊的空間，再加諸真實生活情境中，更常見混合著不同類型內容於各式篇章中。因此，實務層面是較難以區分彼此的差異。然而，在考量文本分類目的係在於提供評量發展者於試題撰寫時導引方向，一是可兼顧不同類型評量內容，另一是平衡各類型試題發展而不過於侷限單一類型試題。因此，本書認為文本分類有其必要性，但並不過於拘泥於釐清不同類型內容的差異，來設計相關評量工具。

## 五、接收性語言活動

就文本傳遞媒介而言，其概念大致呼應 108 英語文課程綱要的「聽、讀、綜合應用」（國家教育研究院，2018a），而本書為順應評量目的，係將該文本訊息傳遞媒介，區分為閱讀、聽力與視聽（audio-visual）、數位科技（digital activity）等溝通活動。就閱讀活動而言，它代表學習者接收文本訊息，多透過純文字與（或）圖片，例如，紙本書籍、文章等；就聽力活動而言，它代表學習者接收文本訊息，多透過聲音傳遞，例如，廣播；此外，有關視聽活動，它係指學習者得同時透過閱讀與聽力媒介進行溝通活動，而最常見的媒介包含如影片、電視等；最後，就透過數位科技活動傳遞訊息而言，它代表學習者接收文本訊息，多透過數位網路媒介來呈現，而其顯著特徵為利用超連結來連結不同網頁，內容則多混合各式文字、圖片、多媒體影像或語音等。

在回應實際發展評量工具層面，本書聚焦的接受性語言活動，是以閱讀與聽力活動發展為主，此外，對於透過視聽、數位網頁等傳遞訊息之媒介，在考量現行英語文國家課程綱要並未條列過多學習表現或內容，多僅止於如，Ae-IV-4 簡易賀卡、書信、電子郵件、1-IV-8 能聽懂簡易影片的主要內容等中電子郵件、影片，使得相關評量試題發展僅止於少數題量。因此，在進行相關分析或結果詮釋時，本書並不會特別針對相關活動進行說明或獨立分析，而是將視聽相關評量試題歸為聽力活動、數位網頁相關試題歸為閱讀活動進行分析結果解釋。

## 第四節 源自文本與評量因素

在文本類型之外，其他可能影響學習者表現的因素，包含字彙與語法使用、篇章長度（length）、評量試題格式（item format），甚至是講者腔調、語速等特徵，以下簡要說明。

### 一、字彙與語法使用

對於字彙與語法使用，本書建議評量試題發展團隊教師依照 108 英語文課程綱要中附錄字彙表（2,000 字）與附錄英語文基礎文法句構表進行工具發展（國家教育研究院，2018a）。其中，若相關文本涉及可能超出範圍的字彙時，命題教師可依照情況差異，提供不同的因應方式。情況一是當該字彙對於學生的理解，會產生直接影響時，則建議團隊教師提供小字典或改以其他適合字彙來替代，例如，附錄四範例 R071 提供 pill 與 appearance 等詞彙小字典。相關考量在於這些詞彙會直接影響學生對於題項的理解；情況二是當該字彙意可從文內直接推論而得、或者不影響理解時，相關試題則可適當保留而不使用替代字彙，而此點就如同 OECD（2021, p. 73）推動 PISA 2025 外語評量時，認為檢視學生對於陌生詞彙的應對，雖有其需求，然而，這類字彙往往不會直接影響學生作答，因此，多可保留。此外，考量不同年級與水準學生英語文能力，相關字彙深度與廣度安排也會隨之產生區別；另一方面，由於 TASAL 透過長期追蹤設計，逐年評量七、八與九年級學生表現，然而，在現行英語文國家課程綱要，相關字彙與語法並未有年級間區分，使得雖然評量試題發展有提供建議準則與規範，但實際進行工具設計，某程度的試題發展仍得仰賴教師教學經驗及其對於不同年級、水準學生可能的英語文表現的理解，以建構出適當篇章與評量試題。

### 二、篇章長度、腔調與語速

篇章長度、文本類型與試題發展者對於內容陳述連貫性需求，彼此具有相關性。對於第四學習階段學生而言，相關篇章發展原則大致是以不超過 400 字，同時，篇章長度安排會隨著不同年級或水準學生能力差異而有所變化，其中，是以愈高年級、或愈高水準表現者，相關篇章平均設計長度相對愈長。其中，本書所發展的篇章長度結果，請見第四章第二節說明。

有關評量中聽力試題之講者的腔調與語速，在考量現行第四學習階段課程綱要並未涉及辨識不同腔調 / 語言背景講者內容。因此，本書對於聽力評量發展是以標準美國腔與一般語速來進行相關語音篇章建構。

### 三、試題格式

有關評量試題格式 (item format)，本書採用兩種各國際閱讀評比最常運用的題型，一是選擇題 (multiple-choice items)、另一是建構反應題 (constructed-response items)。有關選擇題型，其題項數量會依閱讀或聽力活動而不同，閱讀試題為四選一選擇題，而聽力試題為三選一選擇題，而此差異安排多為配合國中學生對於題目選項數量熟悉度，例如，參考國中教育會考之題項數量 (國立臺灣師範大學心理與教育測驗研究發展中心，無日期)。其中，聽力試題播音次數為配合 TASAL 線上評量系統設計，學習者在統一作答總時間受限下，皆可無限次數重播。另一方面，由於學習者對於選擇題的作答，屬於被動式反應 (passive response)，相關試題與題項內容皆是由命題教師所提供，而學習者僅是從中 (被動) 選擇最佳答案，使得在對應理解認知成分上，本書傾向將選項內容視為文本的延伸，而非學生者所主動建構或表述。因此，當需要評量學生表達意見或觀點之理解成分時 (例如，在文本理解上，透過自身觀點進行評估)，多半得透過建構反應題型才足以評量出學習者的理解表現。

有關建構反應試題，大致可區分為兩種形式。一是單題建構題、另一是兩階段建構題。就前者而言，以附錄九 R135 第 4 題為例，下述為單題建構題之範例。學生得在理解題意後，透過幾個句子陳述進行作答。此外，本書對於「參考句型」的提供會視學生年級或題目難易度，來考量是否給予學生相關鷹架 (scaffolding) 的支持，以避免過於艱澀而影響評量試題效度。

What lesson can you learn from the story?

參考句型：I can learn that \_\_\_\_\_.

(附錄九，R135，第 4 題)

由於上述題型對於較低年級學生而言，大多未具備明確、逐步導引其回應高層次理解認知成分的功能，例如，學生除了表達其理解外，同時進一步表達出其分析、評估或省思自身生活等。因此，本書使用兩階段建構題來進行相關理解成分的評量。下述為兩階段建構題範例，其中，第一題與第二題間存在連續作答相關性，學生得

在理解與回應第一題後，再就第二題題意進行舉例佐證或說明，如此，以便同時導引出學生在理解上，能進一步表述自己分析或省思的內容。整體而言，兩階段建構題大多具備輔助學生作答的鷹架角色與功能，會視學生年級與試題欲展現難易程度，視情況使用。

Q4-1: What lesson do you learn from the story?

Q4-2: Please give an example from your life that can show the lesson you learn from the story. (請以你的個人生活經驗作為一個例子，來呼應你從故事中所學。)

(附錄九，R135，第4題)

不論哪一種建構題型，本書大致依照學生年級與能力（尤其是七年級）適時提供參考句型，或在不影響作答前提下，透過簡要中文來說明或導引題意。最後，有關計分準則的設計，相關內容大抵依循 NAGB (2019, p. 44) 建議，聚焦於學習者作答整體內容回應理解多寡程度，而非考量語法或拼字正確性，同時，作答者的反應訊息必須是基於理解文本內容前提下，才得以獲得分數。

## 第五節 轉化「教育目標」為「標準本位評量」

108 課程綱要內容，訂立出核心素養與相關領域學習表現、學習內容，它指引出決策者對於英語文教育理念與目標的界定（國家教育研究院，2018a）。對於國家教育目標的描述，國家課程綱要撰寫者往往是透過抽象、一般（general）、通用性（transversal）語詞來表達對於教師在課室教學所進行意向行動（intentional act）之期望，以便於進行日常溝通與交流，達到宣傳推廣目的。然而，在將相關指標或內容轉化為具體評量工具開發上，相關內容的解讀也容易落於各自表述與詮釋；同時，在考量評量工具發展所具備實務與邏輯可行限制下，評量工具開發者大多得進行主題聚焦與將相關內容具體化。據此，在本節，本書將針對標準本位大型評量建構模式、如何轉化以廣泛性語詞陳述之教育目標為具體評量指引、學生表現標準描述三項，進行簡要說明。

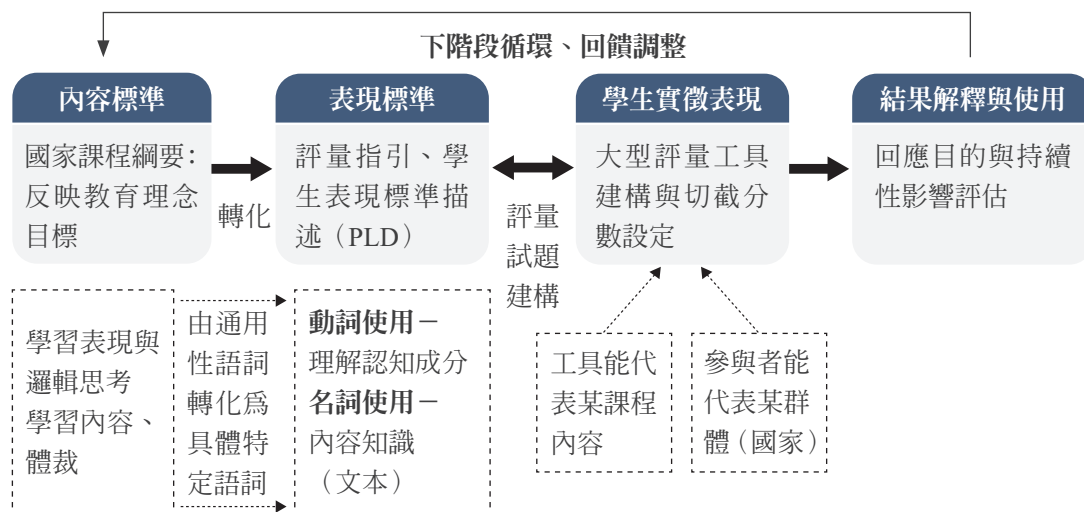
## 一、標準本位大型評量建構模式

TASAL 英語文評量發展基礎係屬於標準本位大型評量（standard-based large-scale assessment），而它是同時具有「標準本位」與「大型評量」內涵。其中，標準本位評量（standard-based assessment）係指相關評量工具的發展對於學生學習表現的評估、分數解釋與使用，係透過「標準」（standards）來進行詮釋與建構發展程序、內容。此概念一如圖 2-8 所示，英語文國家課程綱要所界定內容可視為學者所稱內容或課程標準（content standards or curriculum standards），代表著學習者被期望應該具備的知識與能力（Glatthorn et al., 1998; Hambleton et al., 2012），而相關標準是一種方向性、期望性的內容指引。就課綱具體條列的目標而言，現行課程綱要是以學習表現、學習內容來表徵，例如，3-IV-7 能了解對話的主要內容、3-IV-16 能閱讀不同體裁、不同主題的簡易文章等。內容標準的陳述往往會使用較通用的語詞來呈現，而其優點是容易溝通，但缺點是容易造成各方解讀與見解不一致。因此，在評量工具發展上，本書得進行指標轉化，將課綱所呈現通用性目標內容，進一步轉化或拆解為具體、條列的評量指引，以利評量試題發展與後續蒐集學生實徵表現資料與解釋。

在上述基礎，相關內容則會進一步形成表現標準（performance standards），它係指在內容標準基礎上，不同表現水準的學習者（例如，基礎、精熟與進階學習者、或水準一、二、三不等學習者），能展現出的知識與能力程度（Linn & Herman, 1997），為一具體階層性或基於學生實徵表現結果的內容描述（詳細請見本節後述）。後續，本書再透過具有代表性學生的實徵表現資料、與各水準切截（cutoff score）或通過分數（passing score）（如達到水準二通過標準為某個分數）搭配，來進行學生評量結果解釋或議題回應。另一方面，大型評量（large-scale assessment）係指評量結果具有足以代表橫跨大範圍課程內容與某個特定群體學習者的代表性，若以 TASAL 而言，相關課程內容多指課程綱要所界定內容，而特定群體係為可代表整個國家層次（country-level）的學生。簡而言之，相關評量工具與試題發展是具有課程內容代表性，而學生作答反應是具有某群體樣本代表性（說明請見第四章第一節）。

在轉化內容標準為表現標準上，本書使用 Bloom 及其後續發展建議概念來落實（Anderson et al., 2001; Bloom et al., 1956），茲說明如下。

圖 2-8 英語文標準本位大型評量建構模式



註：修訂自謝進昌 (2022b)。

## 二、轉化教育目標為大型評量指引

在一般評量情境下，由於內容標準的陳述多存在些許模糊、隱含意的空間，使得相關指標內容大多不適合「直接」作為評量工具或試題發展指引。因此，Bloom 早先在提出使用評量架構 (framework) 或分類 (taxonomy) 來建立評量試題發展的指引 (Bloom et al., 1956) 時，就提出相關概念來建構評量指標。架構係指某個概念或現象相關的一系列分類類別，其中，Anderson 等人 (2001, p. 4) 所提出的分類 (taxonomy) 概念為建立架構的一種途徑，其主體是由「認知歷程」與「知識內容」等雙向度所組成，而前者多透過動詞 (verb) 使用、後者多透過名詞 (noun) 使用來表徵與描述。進一步，本書得搭配副詞、形容詞等來修飾不同水準間學生認知表現或所具備內容知識差異，以能使具體回應試題或受測學生所能展現出的評量指引或後續表現水準。簡要說明如下。

### (一) 動詞詞彙使用與階層性

在「動詞」詞彙使用發展方面，如表 2-3 所示，它反映的是學習者展現出的認知、理解或思考成分 (Anderson et al., 2001)。細究現行英語文課程

綱要學習表現，如附錄一所示。課綱撰寫者使用「辨識、聽懂、看懂、了解」等詞語，但就如 Kintsch (2019) 指出「了解 (understanding) 與理解 (comprehension) 是日常用語 有用但不精確」 (p. 178)。在此基礎，本書得進一步轉化或拆解相關語詞，才得進行具體評量指標與表現標準描述的發展。在回顧本書第二章第二節所述，許多學者曾提出理解歷程或成分界定，其間彼此的命名、分類觀點也許不盡相同，但多半具有一致、互通性，使得其運用的「動詞詞彙與認知層次對應」也多具有相似性。

在較低層次或水準階段，學者大多使用如「辨識、找出、擷取、回憶」等動詞詞彙（詞彙的操作型定義說明請見第五章第三節）來表徵學習者「對於基本語言知能、或顯著訊息的直接擷取過程」（國家教育研究院，2018a；Mullis & Martin, 2015; NAGB, 2019; OECD, 2019）。此外，在形成字面文意理解發展方面，其常用動詞詞彙為「連結、結合」，來表徵學習者連結多個微觀命題關係或進行橋接式推論，以建構更大單位鉅觀命題之過程（Kintsch, 1988, 2019），而相關延伸的詞彙，本書的界定包含有「歸類、排序、收集」等詞彙，來表達學習者對於多個明顯、可觀察訊息、命題意或直觀常識性推論的接收運作歷程。

在理解隱含文意過程方面，學者較常使用動詞詞彙有「推論、統整」等，來表徵學習者在結合自身背景知識與經驗下，經推論文本暗示的文意、統整所有想法與訊息，理解整體文本意涵過程（Mullis & Martin, 2015; NAGB, 2019; OECD, 2019）。相關延伸詞彙包含「應用、區分、解釋」等，以表徵學習者進行之統整，或在既有整體組織架構下延伸既有抽象想法、概念以進行理解歷程運作。此外，關於「比較、對照」，其定義為偵測物件或想法間相似或相異性（Anderson et al., 2001），然而，因該物件組成可能是抽象概念、也可能是具體外在特徵，使得該詞彙可同時適用於評量學生字面文意、與隱含文意理解歷程。

就第五、六水準，在學習者於文本理解上，能否透過自身觀點進行評估與省思方面，其常用詞彙包含「評估、省思」等（Mullis & Martin, 2015; NAGB, 2019; OECD, 2019）。相關延伸詞彙有「分析、預測」等，以表徵學習者能主動、持續性進行評估、批判式推論或省思回饋認識自我。

整體而言，學習者理解歷程屬於一連續性狀態，若單就動詞詞彙使用並不足以反映整體歷程與層次，加諸各動詞詞彙定義間係存在模糊與重疊空間。因此，本書進一步搭配名詞使用，呈現學生所需內容知識，以精緻化各階層內容。

## （二）名詞詞彙使用

在「名詞」詞彙（詳細請見第五章第三節）使用發展方面，如表 2-3 所示，它反映學習者在面對生活情境時所展現出的內容知識（Anderson et al., 2001），例如，字母拼讀、常用語、不同體裁與主題文本之重點、情節、觀點、想法、寓意等。相關概念如附錄二所示，主要在呼應現行英語文課程綱要「學習內容」（國家教育研究院，2018a）。

現行英語文課程綱要對於內容界定，是以體裁與主題來區分，然而，隨著數位科技進展，其型態較難囊括所有且不易具體彰顯各類型對於評量命題導引功能。因此，如同第二章第三節所述，本書係參考 OECD（2019）對於文本類型界定，使用一加六類文本，包含語言知能文本、描述文本、教學指引文本、人際互動文本、說明文、議論文、敘事與文學等。其中，如附錄二所示，不同名詞詞彙是不同類型文本內容的顯著特徵，例如，語言知能文本主要特徵是由形素、音素、音節、字彙（詞）、單一句子所組成，故，相關名詞使用重視其對應特徵；就第一類群文本而言，其主要目的大多為傳遞訊息與情感交流，同時，具備其他類型文本基本功能與要件，使得其相關延伸名詞包含有「訊息、細節、腔調、情緒、態度、要點、主旨、文本架構、作者目的、寫作風格」等，大多是一般文本組成基本元素及文本相關上位組織與架構；在第二類群文本方面，相關內容除了保有前述文本特徵外，說明與議論文內容通常會特別包含如「論點、概念、證據、實例、因果」等；而第三類群敘事與文學相關詞彙通常會特別涉及「人物、事件、行動、特質、背景、情節」等（NAGB, 2019; Rumelhart, 1975; Stein & Glenn, 1979）。其中，本書得強調各個名詞詞彙的界定，並非某一類型文本所獨有，而本書僅是透過各類型文本一般會呈現主要顯著特徵，作為名詞使用發想來源，以利建構評量指標。

有關名詞階層性安排，如表 2-3 所示，由於名詞本身並不具認知層次差異性。它得搭配（慣用）動詞，才足以突顯出階層性，例如，「字義」多半得從文本作者所提供上、下文脈絡線索進行推論；「概念」為抽象或通用想法（generic ideas）所概括（generalize）形成的主題；此外，「情節」（plot）為具次序性事件（events）或行動（actions）所形成，屬於相對複雜內容；「觀點」（perspective）代表思考或評估事情的立場、主張；其他諸如作者目的、文本結構、寫作手法等，多半涉及整體上位的文本組織結構，學習者大多必須啟動較高層次理解認知成分（如分析、區分），才足以達到實質性理解（Anderson et al., 2001; International Literacy Association [ILA], n.d.; NAGB, 2019）。

有關形容詞或副詞的使用，本書主要用於突顯各個文本內容訊息複雜性（text complexity）與不同階層理解認知成分差異性。有關形容詞的使用，相關詞彙包含例如，「明顯的、抽象的」訊息或文本的「單一或多元性」、「文本內或跨文本」等，來表徵學習者對於所接觸文本訊息的複雜程度；有關副詞的使用，相關詞彙包含如「直接地」推論，來代表學習者推論過程屬於橋接式、常識性推論，以和需要較複雜推論的過程呈現區別。

整體而言，對於表 2-3 詞彙使用，本書以虛線來作為各層次間間隔表徵，而這代表著各個詞彙的使用界線，存在相當大的模糊空間。本書目的並非在於重新定義或劃分各個詞彙概念與差異，單純為具體、清楚傳遞本書所界定評量指引、學生表現標準描述內容以及對於轉化英語文課程綱要學習表現、學習內容與相關詞彙運用的規則、方向，以便試題發展者、讀者能具體掌握、理解，達到有效溝通目的。

表 2-3 轉化教育目標為評量指引之詞彙使用

層次	名稱	邏輯思考與動詞使用 (學習表現)	形容詞 (或副詞)	名詞使用 (學習內容)
一	能流暢運作語言基本知能	認讀、聽辨、辨識 (recognize)、	基本的 (basic)、	字母 (letter)、
		找出 (locate)、	日常生活的	音素 (phoneme)、
		擷取 (retrieve)、	(daily used)、	字母拼讀 (phonics)、
二	能直接找出訊息	回憶 (recall)、	簡易的 (easy)、	字詞 (彙)
		再製或再現 (reproduce)、	明顯的 (obvious)	(vocabulary)、
		使用 (use)		片語 (phrase)、
				句子 (sentence)、
				細節 (detail)、
				訊息 (information)
三	能理解字面意	連結 (link)、	(未) 明確描述的	字義 (meaning of a
		結合 (combine)、	((not) explicitly	word)、
		歸類 (categorize)、	stated)、	人物 (character)、
		排序 (sort)、	簡短的 (simple and	事件 (event)、
		收集 (gather)、	short)、	行動 (action)、
		描述 (describe)	單一的 (single)、	背景 (setting)、
			具體的 (concrete)、	想法 (idea)、
			直接的	證據 (evidence)、
			(straightforward)、	實例 (example)、
			文本內 (intratextual)	理由 (reason)、
				原因 (cause)、
				要點 (main idea)

(續下頁)

層次	名稱	邏輯思考與動詞使用 (學習表現)	形容詞 (或副詞)	名詞使用 (學習內容)
四	能理解隱含意	推論 (infer)、 比較 (compare)、 對照 (contrast)、 應用 (apply)、 說明 (interpret)、 解釋 (explain)、 統整 (integrate)、 區辨 (distinguish)	間接的 (indirect)、 邏輯連貫的 (coherent)、 複雜的 (complex)、 隱含的 (implicit)、 抽象的 (abstract)、 多元的 (multiple)、 文本間 (intertextual)、 超越文本 (beyond the text)	特質 (trait)、 情緒 (emotion)、 態度 (attitude)、 意圖 (intention)、 情節 (plot)、 因果 (cause and effect)、 論點 (argument)、 概念 (concept)、 主旨 (theme)、 慣用語 (idiom)、 文本架構 (text structure)、 作者目的 (author's purpose)、 寫作風格或手法 (author's craft)、 作者觀點 (author's point of view)
五	在文本理解上，能透過自身觀點進行評估	分析 (analyze)、 評估 (鑑) (evaluate)、 預測 (predict)、 省思 (reflect)		
六	能省思認識自我			

註：1. 虛線表徵各理解成分是一連續且界線存在相當模糊空間，因此，各個詞彙使用並非是單向組合或表徵，可能涉及跨層次使用。同時，鄰近層次可視學生年級與評量需求進行合併。

2. 名詞本身不具認知層次差異性，多半得搭配動詞，才足以突顯出階層性。
3. 相關動詞與名詞詞彙的操作型定義請見第五章第三節。

### 三、評量指標與學生表現標準描述

在表 2-3 所界定動詞、名詞與其他相關副詞、形容詞詞彙基礎下，本書可透過詞彙搭配與組織，來建構各個試題發展的評量指引、與學生在各個水準的表現標準描述 (performance level description, PLD)。

有關評量架構，誠如 Bloom 所述，它係指某個構念或現象相關的一系列分類類別（Anderson et al., 2001; Bloom et al., 1956）。在本書，主要以理解認知成分、內容知識（透過文本類型）為主軸，再納入字彙與語法使用、篇章長度、評量命題格式、講者腔調、語速等安排，而這些指標使用目的皆在於指引、導引評量工具與試題發展。其中，就涉及動詞、名詞詞彙的使用方面，僅界定各理解認知層次相關評量指引。如附錄一所示，在能流暢運作語言基本知能方面，本書運用「使用」、「字母拼讀規則」等動、名詞組成，以形成指引：使用字母拼讀規則辨識出字詞。在透過同樣概念基礎下，本書就可組織、形成不同層次的評量指引，例如，直接「找出」明顯陳述的「想法或訊息」、直接「推論」文本字面「大意、重點」、「推論」出文本抽象、隱含的「意涵」或片段「訊息間關聯」等。然而，本書認為若區分出過多細節性的評量指引，也容易造成命題者設計試題的困擾，因此，本書是將具有「共通性特徵」詞彙，經統整後於單一評量指引來呈現。

在前述基礎上，本書已經說明如何透過「動詞與名詞詞彙使用」，將內容標準轉化為試題發展的評量指標。後續，本書依照類似概念，將相關內容轉化為學生表現標準描述。表現標準係代表在內容標準基礎下，學生預期能達到、展現出的表現程度（Hambleton et al., 2012）。在過去相關的研究應用，學者往往會將學生評量的表現結果，區分為數個水準或層次來進行描述，因此，又稱表現水準描述或其他相關概念之類似詞彙，例如，成就水準描述（achievement level description, ALD）（NAGB, 2019）、熟練水準描述（proficiency level description, PLD）等（OECD, 2019, 2021）。

就內容呈現方式而言，學生表現標準描述某程度上是類似於上述評量指引的組成，屬於一體的兩面，因為，評量指引多作為試題發展使用，而學生表現標準描述則是於評量後，透過學生於該試題實徵結果，來進行表現描述，作為分數解釋使用。其中，兩者主要差異僅在於相關文字內容呈現的具體與細緻化程度，會依照使用者需求調整、或搭配有實徵表現範例試題（數據）結果來呈現。本書所界定學生表現標準描述如表 2-4、2-5 與 2-6 所示，而其係依照閱讀、聽力與文本類型進行區分，以表徵學生在不同語言活動、文本類型所能展現出的理解或能達到的能力描述。

就水準一的學生而言，結果如表 2-4 所示。相關表現標準描述包含如學生「能識讀字母」、「能使用字母拼讀規則辨識出字詞」、「能識讀基本詞彙、日常用語、基本片語、適合該學習階段（文法）結構的句子」等。有關對應範例試題，本書例

舉附錄三 L120：當學習者聽到 supermarket，能選出對應的詞彙圖片，其中，學習者大致得在本評量工具，得到 190.95 分（Thurstonian 閾值（Thurstonian threshold））（估計說明請見第四章，各試題對應分數請見各試題所在附錄）才能有較高機率（設定為 62%）正確作答該試題。

就水準二學生而言，在閱讀理解時，結果如表 2-5 所示。學生表現標準描述包含如學生「能直接找出文本內明確陳述的想法」、「能直接找出人事時地物等外在顯著特徵」、「能直接找出與特定目標有關的明顯訊息」等。其中，對於不同文本類型之描述內容，由於各文本多會涉及一般性訊息的擷取、或找出明顯人事時地物特徵等，因此，在本水準，各文本類型內的學生表現標準描述也大多傾向是完全一致或相同。本水準相關範例試題，本書例舉如附錄四文本 R071 第 1 題（題號 R071\_1）閱讀理解：學習者若能直接擷取藥品特徵訊息，就能正確作答，而學習者大致得在本評量工具，得到 497.46 分，才足以具備理解相關試題能力。另一方面，在聽力理解時，結果如表 2-6 所示。水準二學習者同樣得具備類似前述閱讀之表現標準描述，而相關對應範例試題如附錄六聽力理解文本 L134 第 1 題（題號 L134\_1）：學習者必須具備從文本內，直接擷取 INUHARIKO 特徵的能力。學生在本評量工具，必須得到 466.18 分，才足以具有較高機率正確作答該試題。

在隨著學習者所能展現出的理解認知層次或能運用得內容知識愈趨高階，相關表現標準描述也會愈趨複雜。相關表現水準內容所使用的動詞（表徵理解認知）、名詞（表徵內容知識），也會作出相對的層次變化來因應，例如，水準二為「直接找出」明顯特徵，轉變為水準三使用的「直接推論」訊息間明顯關聯，再到水準四的「推論」出抽象或複雜訊息關聯。因此，在水準三、水準四的層次，學生得分別具備理解字面意、與隱含文意的能力。相關表現標準描述包含如「能直接連結文本內多個訊息，直接推論出字面大意」，範例試題如文本 R115 第 1 題（題號 R115\_1）（水準三，第一類群篇章閱讀理解）、聽力文本 L153 第 1 題（題號 L153\_1）（水準三，第一類群篇章聽力理解）、或能統整文本主要概念、論點，推論出主旨、通則，範例試題如文本 R079 第 2 題（題號 R079\_2）（水準四，第二類群篇章閱讀理解）。

在水準五、水準六的層次，學生得超越文本，對於作者文意進行再次理解或監控理解，展現出「分析、評估、省思」等認知成分、或用自己的話，說明出較上位的文本組織架構（如作者寫作手法）。水準五、六相關表現標準描述包含如「能分

析文本主要論點、概念並提出自身立場或支持證據的說明（屬於第二類群文本內容特色）」、「能分析、評估敘事者主要論點、觀點合理性並提出自身立場說明，例如，新聞事件觀點評估等（屬於第三類群文本內容特色）」、「能省思作者論點、觀點或立場，經覺察可能的一致或衝突，以促進對自我的認識」等。

整體而言，學生表現標準描述反映出在現行 108 英語文課程綱要內容基礎下，經透過本書所發展英語文素養評量工具進行測量時，不同表現水準學生能展現或所具備英語文理解表現概況。同時，本書期望透過範例題的輔助，使讀者能更具體了解與掌握該各個水準學生能達到的具體表現樣貌。

表 2-4 學生能流暢運作語言基本知能表現描述（語言知能文本）

表現水準	名稱	表現描述
一	能流暢運作語言基本知能	<p>當文本 (text) 涉及英語文基本知識，包含如字母、音素、基本字詞 (彙) 與片語、日常用語或適合該學習階段學生學習的句子時，學生：</p> <p>在透過閱讀與 (或) 聽力，接收訊息時：</p> <ul style="list-style-type: none"> <li>— 能識讀字母</li> <li>— 能聽辨語音 (子音、母音) 與其組合</li> <li>— 能使用字母拼讀規則辨識出字詞 (彙)，例如，聽到英語語音 [ʃɪp] 或 [dʒɪp]，選出對應的字彙 [L121]</li> <li>— 能識讀與 (或) 聽辨基本或適合該學習階段字詞 (彙) [L120]</li> <li>— 能識讀與 (或) 聽辨日常用語、基本片語、適合該學習階段 (文法) 結構的句子 [L109][L093]</li> </ul>

註：1. 為力求表現描述內容完整呈現，本表納入銜接不同學習階段學生可能的表現。

2. 相關內容與範例試題同步刊載於謝進昌 (2022a)，但是以本書為主要理論與發展的參照來源。

表 2-5 學生篇章閱讀理解（讀而寫）表現描述

表現 水準	名稱	表現描述		
		第一類群篇章 (描述、教學指引、 人際互動)	第二類群篇章 (說明、議論文)	第三類群篇章 (敘事與文學)
二	能直接找出明顯訊息	<p>當涉及單一、簡短的描述、教學指引、人際互動文本，學生：</p> <ul style="list-style-type: none"> <li>— 能直接找出文本內明確陳述的想法</li> <li>— 能直接找出人事時地物等外在顯著特徵 [R071_1][R071_2][R142_2]</li> <li>— 能直接找出與特定目標有關的明顯訊息</li> </ul>	<p>當涉及單一、簡短的說明、議論文本，學生：</p> <ul style="list-style-type: none"> <li>— 能直接找出文本內明確陳述的想法</li> <li>— 能直接找出人事時地物等外在顯著特徵</li> <li>— 能直接找出與特定目標有關的明顯訊息</li> </ul>	<p>當涉及單一、簡短的敘事文本，學生：</p> <ul style="list-style-type: none"> <li>— 能直接找出文本內明確陳述的想法</li> <li>— 能直接找出人事時地物等外在顯著特徵</li> <li>— 能直接找出與特定目標有關的明顯訊息</li> </ul>

(續下頁)

表現 水準	名稱	表現描述		
		第一類群篇章 (描述、教學指引、 人際互動)	第二類群篇章 (說明、議論文)	第三類群篇章 (敘事與文學)
三	能理解字面意	<p>當涉及相對複雜的描述、教學指引、人際互動文本，學生：</p> <ul style="list-style-type: none"> <li>–能直接連結文本內多個訊息，直接推論出字面大意 [R115_1][R142_1]</li> <li>–能直接推論文本訊息間關聯，例如常識性推論、上下文字義間直接推論、局部連貫推論等 [R065_2][R115_2]</li> <li>–能直接比較文本具體訊息、特徵相似或相異性 [R071_3][R142_3][R115_3]</li> <li>–能直接描述或蒐集文本具體陳述訊息內容 [R065_1]</li> </ul>	<p>當涉及相對複雜的說明、議論文本，學生：</p> <ul style="list-style-type: none"> <li>–能直接連結文本內想法、證據、現象等，直接推論出字面大意</li> <li>–能直接推論文本內想法、證據、現象等內容關聯，例如常識性推論、原因、理由、上下文字義間直接推論、局部連貫推論等 [R018_2]</li> <li>–能直接比較具體想法、證據、現象等內容相似或相異性 [R079_1][R126_1]</li> <li>–能直接描述或蒐集文本具體陳述的想法、證據或現象內容</li> </ul>	<p>當涉及相對複雜的敘事文本，學生：</p> <ul style="list-style-type: none"> <li>–能直接連結文本內人物、事件、行動等內容，直接推論出字面大意</li> <li>–能直接推論文本內人物、行動、事件等內容關聯，例如人物明示意图、行事理由、上下文字義間直接推論、局部連貫推論等 [R135_1][R135_2][R135_3]</li> <li>–能直接比較人物具體特徵、事件、行動等內容相似或相異性</li> <li>–能直接描述或蒐集文本具體陳述的人物特徵、行動、事件等內容</li> </ul>

(續下頁)

表現 水準	名稱	表現描述		
		第一類群篇章 (描述、教學指引、 人際互動)	第二類群篇章 (說明、議論文)	第三類群篇章 (敘事與文學)
四	能理解隱含意	<p>當涉及複雜或多元的描述、教學指引、人際互動文本，學生：</p> <ul style="list-style-type: none"> <li>—能統整文本要點，推論出文本主旨、通則</li> <li>—能從文本片段、複雜或抽象訊息中，推論結果或關聯，例如，作者目的、全文整體連貫之字詞推論、暗示的關聯等 [R115_4]</li> <li>—能從片段、複雜或抽象訊息中，解釋重要訊息間關聯或規則並從文內提出相對的證據 [R142_5][R115_5]</li> <li>—能區辨出文本所運用語文、圖表組織特徵，例如區辨事實與意見、辨識文本結構（註2）</li> </ul>	<p>當涉及複雜或多元的說明、議論文本，學生：</p> <ul style="list-style-type: none"> <li>—能統整文本主要概念、論點，推論出主旨、通則 [R079_2][R018_1]</li> <li>—能推論文本抽象概念、觀點、論點或彼此關聯，例如，作者目的、因果關係、全文整體連貫之字詞推論、相似現象的比擬等 [R079_3][R126_2]</li> <li>—能解釋文本抽象概念、觀點、論點或彼此關聯並從文內提出相對的證據</li> <li>—能區辨出文本所運用語文、圖表組織特徵，例如區辨事實與意見、辨識出類比、文本結構（註2）</li> </ul>	<p>當涉及複雜或多元的敘事文本、或文學內容，學生：</p> <ul style="list-style-type: none"> <li>—能統整文本情節，推論出主旨、寓意</li> <li>—能推論文本隱含意涵，例如，作者目的、人物特質、情節安排、事件因果、全文整體連貫之字詞推論等</li> <li>—能解釋文本隱含意涵並從文內提出相對的證據，例如，人物特質、行動影響等</li> <li>—能區辨出文本所運用語文特徵或修辭，例如區分事實與意見、辨識出隱喻、語調、意象、文本結構（註2）</li> </ul>

(續下頁)

表現 水準	名稱	表現描述		
		第一類群篇章 (描述、教學指引、 人際互動)	第二類群篇章 (說明、議論文)	第三類群篇章 (敘事與文學)
		<ul style="list-style-type: none"> <li>—能比較片段、複雜或跨文本中，抽象訊息相似或相異性</li> <li>—能應用或說明文本大意、規則，以得出結果或類推至其他情境，例如，習得概念應用等 [R142_4]</li> </ul>	<ul style="list-style-type: none"> <li>—能比較片段、複雜或跨文本中，抽象想法、概念、論點相似或相異性</li> <li>—能應用或說明文本概念、論點或通則，以得出結果或類推至其他情境 [R079_4][R126_3]</li> </ul>	<ul style="list-style-type: none"> <li>—能比較片段、複雜或跨文本中，人物抽象意圖、特質、情節內容相似或相異性</li> <li>—能應用或說明文本觀點、意涵，以得出結果或類推至其他情境</li> </ul>
五	在文本理解上，能透過自身觀點進行評估	<p>當涉及複雜、多元文本的描述、教學指引、人際互動文本，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能分析、評估作者目的、規則適用性並提出自身立場說明</li> <li>—能說明文本所運用語文、圖表組織特徵、結構或作者目的（註2）</li> </ul>	<p>當涉及複雜或多元的說明、議論文本，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能分析、評估文本主要論點、觀點合理性並提出自身立場（或證據）說明 [R018_3]</li> <li>—能說明作者目的、文本所運用語文、圖表組織特徵或提出符合文意與語法結構之主旨、標題（註2） [R126_4]</li> </ul>	<p>當涉及複雜或多元的敘事文本、或文學內容，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能分析、評估敘事者主要論點、觀點合理性並提出自身立場說明，例如，新聞事件觀點評估</li> <li>—能說明作者目的、文本寫作風格、手法或結構（註2）</li> </ul>

（續下頁）

表現 水準	名稱	表現描述		
		第一類群篇章 (描述、教學指引、 人際互動)	第二類群篇章 (說明、議論文)	第三類群篇章 (敘事與文學)
六	能省 思認 識自 我	<p>當涉及複雜、多元文本的描述、教學指引、人際互動文本，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能省思作者所呈現訊息或立場，經覺察可能的一致或衝突，以促進對自我的認識</li> <li>—融合運用各理解歷程成分，進行探究實作以解決問題（註3）</li> </ul>	<p>當涉及複雜或多元的說明、議論文本，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能省思作者論點、觀點或立場，經覺察可能的一致或衝突，以促進對自我的認識</li> <li>—融合運用各理解歷程成分，進行探究實作以解決問題（註3）</li> </ul>	<p>當涉及複雜或多元的敘事文本、或文學內容，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能省思作者目的或文中寓意，覺察可能一致或衝突，以促進對自我的認識 [R135_4]</li> <li>—融合運用各理解歷程成分，進行探究實作以解決問題（註3）</li> </ul>

註：1. 為力求表現描述內容完整呈現，本表也納入銜接不同學習階段學生可能表現。

2. 第五學習階段（高中職）才會涉及正式的文本結構或篇章組織。
3. 學習者融合運用各理解歷程之複合性表現的評量，實際得透過實作途徑，才足以評量出學生表現結果，而此並非本書現行目標，但為求描述完整性，故同步條列出。
4. 相關內容與範例試題也同步刊載於謝進昌（2022a），但是以本書為主要理論與發展的參照來源。

表 2-6 學生篇章聽力理解（聽而說）表現描述

表現 水準	名稱	表現描述		
		第一類群篇章 (描述、教學指引、 人際互動)	第二類群篇章 (說明、議論文)	第三類群篇章 (敘事與文學)
二	能直接找出明顯訊息	<p>當涉及單一、簡短的描述、教學指引、人際互動文本，學生：</p> <ul style="list-style-type: none"> <li>—能直接找出文本內明確陳述的想法</li> <li>—能直接找出人事時地物等外在顯著特徵 [L134_1]</li> <li>—能直接找出與特定目標有關的明顯訊息，例如題幹明顯指示的結果等</li> </ul>	<p>當涉及單一、簡短の説明、議論文本，學生：</p> <ul style="list-style-type: none"> <li>—能直接找出文本內明確陳述的想法</li> <li>—能直接找出人事時地物等外在顯著特徵</li> <li>—能直接找出與特定目標有關的明顯訊息</li> </ul>	<p>當涉及單一、簡短的敘事文本，學生：</p> <ul style="list-style-type: none"> <li>—能直接找出文本內明確陳述的想法</li> <li>—能直接找出人事時地物等外在顯著特徵</li> <li>—能直接找出與特定目標有關的明顯訊息 [L086_4]</li> </ul>

(續下頁)

表現 水準	名稱	表現描述		
		第一類群篇章 (描述、教學指引、 人際互動)	第二類群篇章 (說明、議論文)	第三類群篇章 (敘事與文學)
三	能理解字面意	<p>當涉及相對複雜的描述、教學指引、人際互動文本，學生：</p> <ul style="list-style-type: none"> <li>—能直接連結文本內多個訊息，直接推論出字面大意 [L153_1]</li> <li>—能直接推論文本訊息間關聯，例如，常識性推論、交談者外顯意圖或態度等 [L132_1] [L132_2][L134_2] [L153_3]</li> <li>—能直接比較文本具體訊息、特徵相似或相異性 [L134_3] [L153_2]</li> <li>—能直接描述或蒐集文本具體陳述訊息內容</li> </ul>	<p>當涉及相對複雜的說明、議論文本，學生：</p> <ul style="list-style-type: none"> <li>—能直接連結文本內想法、證據、現象等，直接推論出字面大意 [L128_1]</li> <li>—能直接推論文本內想法、證據、現象等內容關聯，例如，常識性推論、原因、理由等 [L128_2] [L128_3]</li> <li>—能直接比較文本具體想法、證據、現象等內容相似或相異性</li> <li>—能直接描述或蒐集文本具體陳述的想法、證據或現象內容</li> </ul>	<p>當涉及相對複雜的敘事文本，學生：</p> <ul style="list-style-type: none"> <li>—能直接連結文本內人物、事件、行動等內容，直接推論出字面大意</li> <li>—能直接推論文本內人物、行動、事件等內容關聯，例如，人物明示意图、行事理由等</li> <li>—能直接比較人物具體特徵、事件、行動等內容相似或相異性 [L086_1][L086_3] [L139_2]</li> <li>—能直接描述或蒐集文本具體陳述的人物特徵、行動、事件等內容</li> </ul>

(續下頁)

表現 水準	名稱	表現描述		
		第一類群篇章 (描述、教學指引、 人際互動)	第二類群篇章 (說明、議論文)	第三類群篇章 (敘事與文學)
四	能理解隱含意	<p>當涉及複雜或多元的描述、教學指引、人際互動文本，學生：</p> <ul style="list-style-type: none"> <li>—能統整文本要點，推論出文本主旨、通則</li> <li>—能從文本片段、複雜或抽象訊息中，推論出結果或關聯，例如，作者目的、暗示的關聯等</li> <li>—能從片段、複雜或抽象訊息中，解釋重要訊息間關聯或規則並從文內提出相對的證據</li> <li>—能區辨出文本所運用語文、圖表組織特徵，例如區分事實與意見、辨識文本結構（註2）</li> <li>—能比較片段、複雜或跨文本中，抽象訊息相似或相異性</li> <li>—能應用或說明文本大意、規則，以得出結果或類推至其他情境</li> </ul>	<p>當涉及複雜或多元的說明、議論文本，學生：</p> <ul style="list-style-type: none"> <li>—能統整文本主要概念、論點，推論出主旨、通則</li> <li>—能推論文本抽象概念、觀點、論點或彼此關聯，例如因果、作者目的等</li> <li>—能解釋文本抽象概念、觀點、論點或彼此關聯並從文內提出相對的證據</li> <li>—能區辨出文本所運用語文、圖表組織特徵，例如區辨事實與意見、辨識出類比、文本結構（註2）</li> <li>—能比較片段、複雜或跨文本中，抽象想法、概念、論點相似或相異性</li> <li>—能應用或說明文本概念、論點或通則，以得出結果或類推至其他情境</li> </ul>	<p>當涉及複雜或多元的敘事文本、或文學內容，學生：</p> <ul style="list-style-type: none"> <li>—能統整文本情節，推論出主旨、寓意</li> <li>—能推論文本隱含意涵，例如，作者目的、人物特質、情節安排、事件因果等 [L086 2][L139 1]</li> <li>—能解釋文本隱含意涵並從文內提出相對的證據，例如，人物特質、行動影響等</li> <li>—能區辨出文本所運用語文特徵或修辭，例如區辨事實與意見、辨識出隱喻、語調、意象、<u>文本結構</u>（註2）</li> <li>—能比較片段、複雜或跨文本中，人物抽象意圖、特質、情節內容相似或相異性</li> <li>—能應用或說明文本觀點、意涵，以得出結果或類推至其他情境</li> </ul>

（續下頁）

表現 水準	名稱	表現描述		
		第一類群篇章 (描述、教學指引、 人際互動)	第二類群篇章 (說明、議論文)	第三類群篇章 (敘事與文學)
五	在文本理解上，能透過自身觀點進行評估	<p>當涉及複雜、多元文本的描述、教學指引、人際互動文本，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能分析、評估作者目的、規則適用性並提出自身立場說明</li> <li>—能說明文本所運用語文、圖表組織特徵、結構或作者目的(註2)</li> </ul>	<p>當涉及複雜或多元的說明、議論文本，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能分析、評估文本主要論點、觀點合理性並提出自身立場(或證據)說明(註2)</li> <li>—能說明作者目的、文本所運用語文、圖表組織特徵</li> </ul>	<p>當涉及複雜或多元的敘事文本、或文學內容，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能分析、評估敘事者主要論點、觀點合理性並提出自身立場說明(註2)</li> <li>—能說明作者目的、文本寫作風格、手法或結構(註2)</li> </ul>
六	能省思認識自我	<p>當涉及複雜、多元文本的描述、教學指引、人際互動文本，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能省思作者所呈現訊息或立場，經覺察可能的一致或衝突，以促進對自我的認識</li> <li>—融合運用各理解歷程成分，進行探究實作以解決問題(註3)</li> </ul>	<p>當涉及複雜或多元的說明、議論文本，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能省思作者論點、觀點或立場，經覺察可能的一致或衝突，以促進對自我的認識(註2)</li> <li>—融合運用各理解歷程成分，進行探究實作以解決問題(註3)</li> </ul>	<p>當文本涉及複雜或多元之人物故事、敘事內容，在全文理解基礎下，學生：</p> <ul style="list-style-type: none"> <li>—能省思作者目的或文本寓意，覺察可能的一致或衝突，以促進對自我的認識(註2)</li> <li>—融合運用各理解歷程成分，進行探究實作以解決問題(註3)</li> </ul>

(續下頁)

- 註：1. 為力求表現描述內容完整的呈現，本表也納入銜接不同學習階段中學生可能的表現。同時，受限成本，現行尚未有口說命題。
2. 第五學習階段（高中職）才會涉及聽力理解、較高層次的表述能力與文本結構或篇章組織。
  3. 學習者融合運用各理解歷程之複合性表現的評量，實際得透過實作途徑，才足以評量出學生表現結果，而此並非現行目標，但為求描述完整性，故同步條列出。
  4. 相關內容與範例試題也同步刊載於謝進昌（2022a），但是以本書為主要理論與發展的參照來源。

## 第三章 英語文學習動機與策略使用

108 課綱所定義核心素養，係指學生為適應生活及未來挑戰，所應具備的知識、能力與態度（教育部，2021）。除了學習學科知識外，相關概念更重視其與情境結合並於生活中實踐力行的概念，並重視學習歷程與相關方法、策略的學習（國家教育研究院課程及教學研究中心核心素養工作圈，2015）。在此理念與精神下，學習者必須同時具備語言知識與相關認知、非認知能力等，才足以回應來自不同任務情境與社會文化脈絡挑戰。在前面第二章中，本書透過形成素養之語言理解認知層面為主體，進行論述，而在本章中，將以形成素養之非認知因素為主體進行說明。

### 第一節 社會文化脈絡的情境觀點

心理學者對於「動機」（motivation）的探究，主要關注人如何形成信念、價值與思想，而就促進學習動機（motive to learn）探究而言，學者自然會關注學習者如何理解、使用策略與技能或學習者自身特質等；然而，就情境（situative view）理論觀點而言，它重視整合認知與互動觀點，認為學習的意義會隨著社會實務（social practice）變動，其中，不同社會和文化脈絡則會賦予學習對於學習者不同意義（Nolen et al., 2015）。

社會文化脈絡對於「動機」形成的重要性，在於賦予理論不同詮釋方向，而這可從相關理論歷史發展與變化之中一窺端倪。就自我決定理論（self-determination theory）而言，相關學者重視學習者如何建立三個心理需求，分別是能控制自我行動意向的自主性（autonomy）、建立歸屬或與他人關係連結性（relatedness）、自身所具備或達到的能力覺知（competence）（Deci & Ryan, 1985; Ryan & Deci, 2000），而脈絡代表著環境能否提供學習者獲得這些心理需求的支持條件、滿足與否對於學習者動機影響（Nolen et al., 2015）；在成就目標導向（achievement goal orientation）理論下，脈絡代表著個體因社會化影響所形成的各種成就目標，例如，趨向精熟、趨向表現目標，導引學習者投入或迴避參與相關情境活動（Wentzel, 1996）。

在期望價值理論，Eccles 等人（1983）認為動機是個人期望成功（expectancy for success）與主觀任務價值（subjective task values）判斷結果。期望成功代表著學習者對於相信自己在某個特定任務所能表現出最佳程度的信念，而它是具有領域、時間特定性（domain and time specific）（Eccles & Wigfield, 2020），代表在不同領域範疇、時間等條件下，學習者對於自我能力概念覺知會隨著脈絡條件差異而不同；主觀任務價值代表著學習者對於任務所賦予的價值，帶有主觀性，即使相同水平難度的任務，不同個體也具備不同主觀價值，而形成其價值來源可能歸於內在興趣（intrinsic value）、成就價值（attainment value）、效用價值（utility value）與代價（cost），其中，在成就價值部分，它會隨著個人對於任務活動重要性認同（identity）程度而具差異性，其涉及在社會文化脈絡下，個體如何形成自我機制、社會與個人認同等層面，而最常被提及的就是性別角色與學科表現下，個體對於自身角色是否適合該任務活動意象的認同形成（Eccles & Wigfield, 2020）。

整體而言，社會文化脈絡之於動機發展，扮演著重要角色，而此概念銜接本書第二章第一節所論述，素養是學習者同時具備語言知識與相關認知、非認知能力等，以回應不同任務情境挑戰之能力，且素養是由學習者、情境與脈絡所交織影響而形成。

## 第二節 英語文學習動機

在 108 英語文課程綱要中，有關學生學習興趣與態度方面，相關概念強調學生自主學習或樂於學習，而細究其所陳述的學習表現，例如，6-II-1 能專注於教師的說明與演示、6-III-1 具有好奇心，主動向教師或同學提出問題、6-IV-1 樂於參與課堂中各類練習活動，不畏犯錯等（國家教育研究院，2018a）。在此基礎，本書同樣以動詞與名詞的概念，來擷取重要詞彙，大致可擷取出「能專注、好奇心、主動參與、樂於參與、與他人互動、接受挑戰」等重要元素成分。進一步，在對照期望價值與成就歸因理論時，多數成分也大致呼應學習者之所以投入學習，多是源自於內在興趣與價值及主動參與、展現與他人良好溝通、互動之社會動機（Eccles & Wigfield, 2002, 2020; Wentzel, 1996）。

在回顧歷史發展，語言學習動機（motivation）已經被學者探究超過六十多年（Al-Hoorie, 2017; Dörnyei & Ryan, 2015），很難匯聚出統一定義或概念。其中，若就期望價值理論來詮釋，它係指學習者行動價值（value）、信念（belief）與目標（goal）（Eccles & Wigfield, 2002），而動機往往會隨著時間、任務領域範疇差異而不同。其中，相關構念及其延伸概念發展也十分多元，例如後文介紹之 Schiefele 等人（2012），因此，本書很難在單一有限章節篇幅內完整論述。因此，本書考量在能呼應英語文課程綱要內涵下，同時能兼具實際配合進行大型評量與調查的可能性，聚焦內在動機、自我概念與社會動機等三個構面，分別代表內在價值、能力信念與社會目標。

### 一、內在動機

就學習內在動機而言，相關理論是建立在 Deci 與 Ryan（1985）的自我決定理論（self-determination theory），認為內在動機（intrinsic motivation）係指學生對於學科知識的主動攝取，多源自於本身對學科的興趣與喜悅，能從中獲得自身滿足，為人類自然的基本需求，而學生涉入學習的程度乃完全出自於自我的決定（self-determined）。其中，內在動機的元素包含如好奇（curiosity）、尋找新經驗或訊息、樂於參與或沉浸於學習等內涵（Guthrie & Klauda, 2016; Schiefele et al., 2012）；和 Eccles 與 Wigfield（2002, 2020）。若在對照期望價值理論，學者認為內在動機反映出學習者對於學科學習任務的價值判斷，是源自於內在價值（intrinsic value），而對於相關活動的投入往往會伴隨有喜悅、興趣與喜歡等傾向。整體而言，自我決定理解與期望價值理論對於內在動機的詮釋，是具有一定程度的相似性，其中，在相關評量與調查工具的發展上，本書依循上述概念與定義，經擷取重要成分元素（例如興趣、好奇、樂於參與等）來進行工具發展。

### 二、學科表現自我概念

學習者對於自身能力信念（competence belief）覺知，是動機理論學者廣泛探究的一環。就自我決定理論而言，學者認為人類先天具有對於自身能力覺知的需求，因此，會主動的投入自己認為能勝任的任務活動中（Deci & Ryan, 1985; Ryan & Deci, 2000）。此外，期望價值理論學者認為學習者對於學科自我概念（academic self concept）會形塑個人對於成功的期望（expectancy for success），進而在特定領域，建立自己對於某項學習任務能表現出最佳能力程度的信念（Eccles & Wigfield, 2002）。

另一方面，就學科自我概念、與自我效能而言，學者對於不同詞彙概念，是認為有其差異性與相似性。Eccles 與 Wigfield（2020）在更名其理論為「情境觀的期望價值理論」（situated expectancy-value theory）時，認為學習者對於成功之期望（expectancies for success）會隨著時間、任務特徵而改變，因此，相關概念是具有任務特定（task specific）性質，而其內涵則略有別、但相當近似於學科自我概念所重視的能力覺知領域特定性（domain specific）。另一方面，學習者對於能力信念也展現在自我效能（self-efficacy）範疇，Bandura（1997）在社會認知理論（social cognitive theory）中，認為個人、行為與社會、環境屬於交互影響的關係，而此互動會影響學生對於未來學科學習所擁有或期望達到（capability）的信心、能力覺知（perceived competence），其中，學者認為具有較高自我效能的學習者，通常較會主動攝取知識、使用有效學習策略，並透過目標設定進行自主學習，善於監控自身學習、理解與目標達成程度（Schunk & Bursuck, 2016）。在概念層次，學科的自我概念與自我效能共享著許多類似元素，然而，Bong 與 Skaalvik（2003）、Marsh 等人（2019）在區分這兩者觀點的差異時，就曾以領域特定性（domain specificity）程度、能力比較的參照來源、面向涵蓋性程度、時間趨向、調查題項是透過描述或評估性進行測量等進行區分，認為兩者在概念與觀點層次，是存在差異性，然而，若實際要進行兩者界線的劃分、甚至是評量工具發展與結果詮釋，確是處於相對未定論，很難具體區分彼此。

在上述的定義與概念層次，儘管不同學者對於「能力信念」界定所使用概念詞彙，存在些許歧異看法，然而，在相關調查工具發展上，本書在基於大型評量對於特定構念的調查與試題發展，是有其篇幅、作答時間限制性等因素考量下，會傾向調查學生對自身整體英語文能力覺知，故以學科表現自我概念來定義。同時，學生形成自我對於英語文表現概念或衡量來源，本書所界定元素包含有參照架構（來自於社會比較或內在標準）、因果歸因、重要他人的評價（如同儕、教師、或家人）、過去精熟經驗、心理重要性等，用來導引相關調查工具的發展（Bandura, 1997; Bong & Skaalvik, 2003; Marsh et al., 2019; Schunk, 2012）。

### 三、社會動機

在自我決定理論中，學者認為人類具有建立歸屬及與他人建立關聯的需求（Deci & Ryan, 1985; Ryan & Deci, 2000）。在此基礎下，學者認為語言（學習）不僅具有工

具性功能，甚至扮演著促使自身融入社會群體一份子的途徑，用以追求自身的社會目標（social goal）（Wentzel & Wigfield, 1998）。社會目標係指學生個體企圖在某個情境下完成的任務活動，而相對於其他成就目標，其所促發的價值來源在於自身是否能被社會群體（如同儕、老師或家人）所接受，並認同他們學習行為是適當的、好的，進而增強自身對於任務活動參與，因此，學習者所展現行為往往是能被社會所認同或接受的反應（Wentzel, 1997）。

在上述概念基礎下，本書對於社會動機調查工具的發展，大致可延伸出兩個層面。一是學生追求融入社會群體，另一是遵從社會規範。前者代表學生想透過學習，成為社會群體一員的過程，後者係指學生服從外在規範與期望來進行學習（宋曜廷等人，2003；Baker & Wigfield, 1999; Guthrie & Klauda, 2016; Schiefele et al., 2012）。然而，在考量大規模調查，相關作答時間的限制性與成本考量，本書目前僅就前者追求融入社會群體為題項設計焦點，而相關元素包含有活動參與融入群體（如同儕、老師或其他社群）、獲得樂趣等來進行題項命題。最後，值得一提的是就臺灣學生作為外語學習者角度而言，英語文之於學生的重要性，不一定著眼於增進和西方人社交互動為目標，往往更強調學生能在學業、職場獲得好表現或好工作之發展（黃淑真，2007），據此，雖然本書目前尚未納入外在動機（extrinsic motivation）調查，然而，這是未來值得進一步建構與發展面向。

### 第三節 英語文理解策略使用

如何運用學習、理解策略來提升學生英語文學習表現，早已經是過去許多語文研究的焦點。學者對於策略（strategy）定義與分類更是多元且複雜，現行很難有統一的詮釋方式（Ardasheva et al., 2017; Cohen & Macaro, 2007; Dörnyei & Ryan, 2015; Jeon & Yamashita, 2014; Plonsky, 2011）。在此現況下，本書企圖從相對通用性的定義著手，再進行聚焦，認為學習策略包含任何能促進習得了解或新知識技能轉移之思想、行為、信念或情緒（Weinstein et al., 2000, p. 727）。Weinstein 等人（2000）的定義似乎包含學習者任何可能的內、外在成分元素，然而，基於通用定義也容易衍生出一些在工具發展，過於發散的疑慮，因此，在回應本書第二章英語文素養的論述時，便是聚焦哪些策略的使用能促進學生英語文學習，以利學習者進行有效語言溝通活動。

108 課程綱要重視學生對於策略使用的學習，以便培養學生能成為自主學習者。在此基礎，本書對於學習策略使用相關工具的建構，主要為呼應第二章理解認知成分所發展或可簡略稱為「促進理解的策略」著手，同時，依照相關概念大致可界定與區分出「促進字彙理解策略、認知策略（促進字面文意理解策略）、推論連結策略、理解監控策略」（謝進昌，2021b）等。說明如下。

### 一、促進字彙理解策略

記憶策略係指在語言溝通過程中，學習者能透過有效的記憶方式以促進自身對於相關語言知識的理解。其中，在對照語言知識學習面向，本書重視字彙的理解、記憶保留與擴張，以提升學生運作語言知識的流暢性。在早期研究中，字彙學習策略多重視透過持續性重述（rehearsal）方式來增進字彙長期記憶，然而，單純使用重述會流於單調、缺乏效率性。因此，後續研究者逐漸提出其他能輔助字彙意象建立的表徵方式，其中，Schmitt（1997, 2008）認為除了透過字彙外形與意義（form-meaning）連結，來增進學生理解外，甚至連結原有字彙與其他相似詞彙理解方式，也都是能有效促進學生對於原有字彙的理解、擴大字彙量與長期記憶儲存。相關概念它包含如連結應用字的外形（字根、字首）、聲音、圖像、情境脈絡與字詞，使學習者產生心理表徵（Gambrell & Bales, 1986）、針對字詞進行有意義的歸類、使用某些有意義型態或規則辨識字詞等（RRSG, 2002）。

在對照現行 108 英語文國家課程綱要內容，相關概念大抵能呼應的學習表現，包含如 3-V-12 能利用字詞結構、上下文意、句型結構及篇章組織推測字詞意義或句子內容，重視「字詞結構」對應（國家教育研究院，2018a）。

### 二、認知策略

認知策略係指在語言溝通過程中，學習者能透過有效的認知途徑進行字面訊息處理，以促進文本文意的理解。在呼應理解認知成分發展上，相關能協助學習者理解的策略包含如分辨重點與細節、概括出大意、重點標示等。在回顧過去研究發展，學者們認為摘取重點、作筆記、標示重點或畫底線等，皆是有效組織文內訊息的技巧，能幫助學習者理解主題（Jeon & Yamashita, 2014; Oxford, 2002）；此外，對於文本結構教學、句型與語法規則掌握等，也是過往研究關注焦點，而結果顯示對於學習者進行語言知識結構教學，多能協助學習者有效促進自身文意理解能力（RRSG,

2002)。另一方面，在銜接現行英語文國家課程綱要上，相關概念大致可呼應的學習表現，包含有 3-IV-14 能快速閱讀了解文章重點，並有效應用於廣泛閱讀中、3-IV-12 能熟悉重要的閱讀技巧，如擷取大意、猜測字義、推敲文意、預測後續文意等（國家教育研究院，2018a）。

### 三、推論連結策略

推論連結策略係指在語言溝通過程中，學習者透過連結自身的背景經驗與知識，在經過訊息推論與統整等，以促進對於文本內涵深層理解，而相關能協助學習者推論的策略，包含如，透過上下文脈絡、語言線索或其他顯著視覺表徵等進行推論、反覆重讀尋找關聯等。在回顧過去研究發展，學者們發現若學習者能預先啟動先備知識、內容線索或者透過圖表、標題等明顯的視覺表徵協助，往往可有效促進自身對文本整體理解（Coiro & Dobler, 2007; Guo et al., 2020; RRSg, 2002）。在銜接英語文國家課程綱要內容上，相關概念大致可呼應的學習表現，包含 3-IV-11 能藉圖畫、標題、書名等作合理的猜測、3-IV-12 能熟悉重要的閱讀技巧，如擷取大意、猜測字義、推敲文意、預測後續文意等、3-V-14 能從圖畫、圖示或上下文，猜測字義或推論文意等。

### 四、理解監控策略

理解監控策略係指在語言溝通過程中，學習者使用能促進自身理解連貫性與意義性掌握度的相關策略，以因應（在對照自身理解架構與評估準則時）可能面臨的衝突或矛盾，包含對自身理解的再次檢驗、監控自己是否真的明白、調整評估與修正可能的理解錯誤等方式（Baker, 2005; Baker & Brown, 1984）。此外，在回顧過往研究，學者們認為學習者若能真正意識到自己理解狀況，採取因應的策略、作出調整，是有助於自身達成實質性理解，而相關策略的重要性也已經獲得許多理論與研究支持（Cain et al., 2004; Coiro & Dobler, 2007; Oxford, 2002; RRSg, 2002）。在銜接英語文國家課程綱要內容上，相關概念大致可呼應的學習表現，包含如 7-IV-5 能訂定英文學習計畫，檢視自我學習過程，並隨時改進、7-V-3 能思考及釐清課文內容及找尋相關資料，強化學習成效、7-V-8 能檢視自我學習過程，並隨時改進等。

最後，本書得強調理解技能（comprehension skill）或理解認知成分、與理解策略（strategy）間，是屬於兩個類似但彼此具有差異性的概念。其中，理解技能通常

是指學習者於回應某個溝通任務時，所展現出的內在、無意識的自主運作能力；然而，理解策略是屬於在特定目的下，學習者在有意識下，所採取的行動，以協助其理解（Manoli & Papadopoulou, 2012）。由於當學生習得某個策略久了，通常會逐漸將該策略內化為自然或自動化，使得學生在使用技能或策略時，是屬於在有、或無意識狀態，相關界線會逐漸變得模糊而難以區分。在此基礎，本書對於相應工具發展，是在考量實務與邏輯可行性下，透過客觀評量來測量學生技能或理解認知成分，同時，透過自陳量表來測量學習者自覺自身策略使用傾向。

## 第四章 效度評估

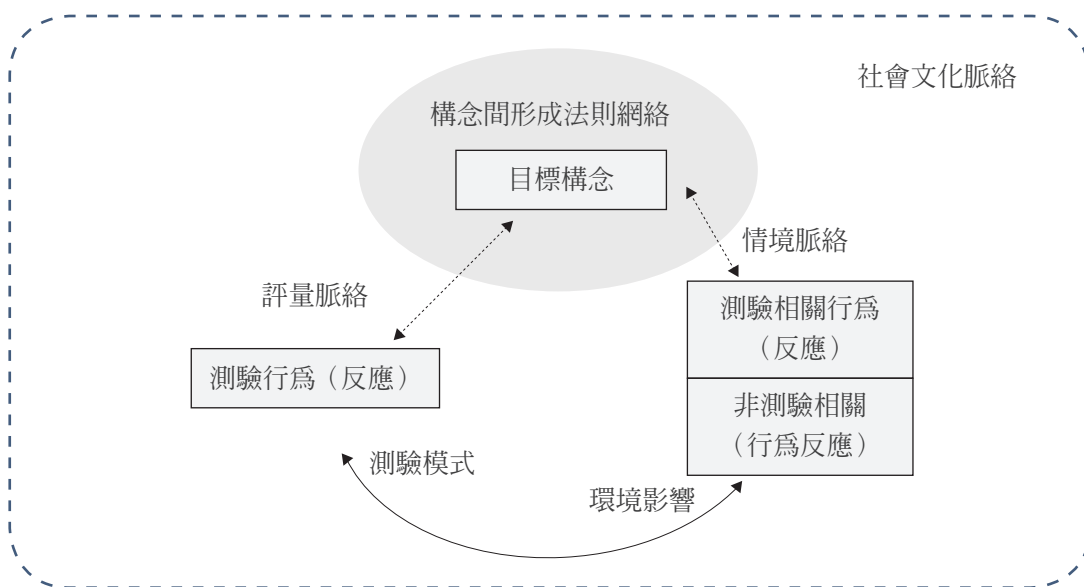
### 第一節 效度評估證據來源

在評量發展的歷史演變中，效度（validity）始終占有重要角色。學者們對於效度定義與見解雖然十分多元，但多數學者大多認同效度評估（validity evaluation）途徑是建立在實徵性證據（empirical evidence）與理論原理（theoretical rationale）之整合性評估基礎上，而其關注的面向不僅止於評量工具本身，更重視使用者對於測驗分數的詮釋、使用的充分性（adequacy）、意義性（meaningfulness）、適切性（appropriateness）（American Educational Research Association [AERA] et al., 2014; Kane, 2013; Messick, 1989; Shepard, 1993）。當本書以此概念來檢視大型評量資料建置目的時，相對應的需求與概念可謂不謀而合。其中，大型評量建立目的在於透過實徵數據回饋教育系統，藉以導引教育決策者進行政策制訂或調整。就 TASAL 英語文素養評量而言，其首要目的在於透過學生實徵表現結果，來回饋英語文課程綱要，而在此基礎下，本書特別關注相關測驗分數使用之適切性和仰賴資料分析結果所下的決策，可能產生之正面或反向影響。另一方面，大型評量的建立不僅涉及評量或調查工具的發展，更與該評量脈絡下，複雜資料蒐集與分析程序息息相關，使得在建立支持評量結果有效性證據的方式，就更顯得多元而不單侷限於某個層面。相關方式可能是透過實徵表現數據資料、或透過主觀合理性論述等方式進行舉證支持，同時，更涉及這些證據支持評量工具本身詮釋合理性、抑或是支持相關測驗分數使用合理性程度等。整體而言，支持評量工具有效性的證據來源及其詮釋，是與效度評估的充分與適切性程度，有著密切相關。以下茲就構念的建構理論觀、效度評估證據來源的多元性、與本書效度評估證據來源等面向，進一步說明。

## 一、構念的建構理論觀點

在回應支持評量工具有效性證據來源前，本書先透過 Messick (1989, p. 28) 從建構理論觀點 (constructivist view) 來詮釋構念建構模式，以了解效度評估所可能涉及的前置理論基礎與方向。相關概念如圖 4-1 所示。建構理論學者認為學習者目標構念 (target construct) (例如，研究者目標在於關注英語文理解) 是透過學習者所接受測量的行為或反應所表徵，其中，相關測驗的發展往往是在特定目的評量脈絡下，透過有系統性與標準化測驗模式所形成。以本書所關注的英語文理解評量而言，相關測驗模式可能涉及語言活動界定、文本選擇、題項建構、參數分析模式等，而評量工具發展者會使用透過此過程所蒐集到的測量結果，來表徵學習者對於目標構念的表現概況，以進一步落實、達成預期目的。另一方面，學者認為觀察實際測驗行為、假設理論的構念與構念間關係，可能存在模糊詮釋空間，如此，就會形成所謂構念間法則網絡 (nomological network) (Cronbach & Meehl, 1955)，並藉此來代表秉持相似理論，但係由其他測驗相關行為 (反應) 所表徵不同構念與目標構念間關聯網絡。在真實情境下，學習者所能展現出構念與其法則網絡，往往並非單獨存在，而它是鑲嵌在某個評量、情境脈絡下，形成學者所述構念是在某個脈絡基礎下，進行詮釋 (construct in context) (Messick, 1989)。此概念正呼應本書第二章中社會認知模式學者所述，它是交織著情境 (situation) 與更廣博之社會文化脈絡影響。整體而言，評量學者對於構念的建構理論觀點，某程度是與社會認知模式學者秉持共同相似的觀點，因此，本書對於相關理論轉化與評量工具建構，是具有相當一致性。

圖 4-1 評量構念的建構理論觀點



註：1. 引用與部分修訂自 Messick (1989, p. 28)。

2. 若以 Messick (1989, p. 30) 描述的建構－實在論 (constructive-realist) 觀點，認為學習者的本體特質是無法被測量的，使得其模式會多出「人格結構與特質」一環，但由於本文重點並非釐清兩者差異，故簡化使用原建構理論模式來詮釋。

## 二、效度評估證據來源的多元性

在了解建構理論學者對於構念建構模式的觀點後，效度評估方向與證據來源往往隱含著如何建立證據才能對映 (mapping) 支持模式內的複雜關聯。舉例來說，工具發展者可以蒐集學生實際的測驗行為，來表徵目標構念的程度 (即圖內左側關聯虛線)、不同構念間的關聯程度 (即表徵構念間法則網絡)、測驗分數使用在社會文化脈絡下受到的影響程度等證據。在此概念，若對照過去效度的歷史發展，不同領域學者早已經提出許多支持效度證據來源框架或面向。在早期，Cronbach 與 Meehl (1955) 是從心理測驗發展角度，提出內容 (content)、構念 (construct)、同時 (concurrent) 與預測 (predictive) 等四類效度 (或合併後兩類為效標關聯 (criterion-related) 效度)。後續，Messick (1989) 在回應 Cronbach 與 Meehl 所提效度分類概念時，Messick 是進一步提出整合性效度理論，認為效度是單一整合 (unified) 概念，

稱為建構效度（construct validity），而工具發展者得配合測驗詮釋、測驗使用、證據基礎、影響基礎等四個面向的指引，藉由蒐集多元證據來源（項目）來支持評量工具有效性。Messick 所提概念除了關注評量工具本身發展的有效性外，更強調測驗分數詮釋價值、使用與影響評估，如此，才得以形成更全面性的效度評估證據框架。其中，為便於實務應用，Lissitz 與 Samuelsen（2007）進一步建議工具發展者可以關注評量內部、外部與理論、實務面等面向的檢視，以期使效度評估更容易達到溝通與使用目的。

在實務應用層面而言，AERA 等（2014）也曾提出一些評估工具有效性的證據來源建議，而相關概念大抵是依循 Messick（1995）所提理論而略做調整，包含測驗內容（test content）、反應歷程（response process）、內部結構（internal structure）、其他變項關聯（relations to other variables）與測驗後效影響（consequences of testing）等。在同時期，相關實務應用研究也大致依循相同軌跡進行評估，但使用略有差異詞彙來進行語言評量效度評估，例如，Khalifa 與 Weir（2009）、Geranpayeh 與 Taylor（2013）對於英語文評量效度評估證據，是從認知（cognitive）、脈絡（context）、計分（scoring）、效標關聯（criterion related）、後效影響（consequential）等五個層面著手來檢核劍橋大學英語認證（Cambridge English Qualifications）測驗效度。另一方面，對於在評量工具發展中進行標準設定議題，學者則提出支持標準設定有效性之證據面向，可以包含過程、內部與外部證據（procedural, internal and external evidence for validity）（Kane, 1994, 2001）。

### 三、本書效度評估證據來源

在上述簡要回顧效度歷史發展脈絡時，讀者可以發現「效度」是一個複雜、多元的動態演變過程，然而，為了滿足實際或實務上推動效度評估的可行性，研究者多得在邏輯可行的基礎下，提出多個不同來源或類型的證據來支持工具的有效性，甚至，這個證據收集過程還可能涉及持續性、長時間的累積程序，才得以瞭解當下或未來在不斷隨著時間演變，相關測驗與評量工具分數詮釋、使用的有效性程度。

對於評量工具的「效度評估」，相關證據具備有來源多元性、類型多元性、與持續性累積等特徵。其中，證據來源的多元性係指評估效度的證據可以來自於各種不同的層面，而一如前述，學者們所提出的評估層面，是具有多元與複雜性，其

中，不同面向的界定也存在著相當重疊性，加諸，考量本書目的並非釐清不同學者所界定概念差異，因此，若單純就可理解性與溝通方便而言，本書是引用 Lissitz 與 Samuelsen (2007) 所建議可從評量工具本身的內部、與其他變項關聯之外部證據著手，同時，本書再融合並調整 Kane (1994, 2001) 提出有關標準設定有效性評估之證據面向，包含過程證據 (procedural evidence for validity)，以及納入 Messick (1989) 所提關注評量結果使用所產生後效影響 (consequential evidence for validity) 層面 (Pant et al., 2009)。其中，Kane 所提過程證據來源多半來自標準設定程序內容，然而，一如 Cizek 與 Bunch (2007)、Hambleton (2001)、Hambleton 等人 (2012) 認為標準設定有效性不僅涉及標準或通過分數本身，更和評量目的釐清、評量工具發展等歷程有關，因此，本書直接使用相關詞彙並延伸其概念至整個評量工具建置與資料分析過程以建立證據。

證據類型多元性係指工具發展者對於證據的論述，可來自於不同類型，例如，基於理論或原理的詮釋性文字論述、提供實徵事實性數據等 (Kane, 2013)，來說明工具的有效性；最後，持續性累積係指工具發展者在邏輯可行性與實務限制性，在有限時間、人力成本、研究者 / 學生可負荷量等因素下，考量單次研究、評量或結果呈現並無法一次到位，以蒐集所有來源證據進行呈現，因此，認為支持效度評估證據會隨著時間演變、累積產生變化，目前結果都僅止於暫時性結論。

在上述特徵與概念基礎下，本書對於進行工具效度評估的證據來源框架，大致如表 4-1 所示，可區分為過程、內部、外部、後效影響等四個層面。支持效度的內部證據係指本書關注評量工具發展與資料分析本身的有效性 (Lissitz & Samuelsen, 2007)，其中，若將評量工具發展與資料建立過程描述獨立出來，本書是透過檢視「英語文素養與相關構念理論、架構發展論述」（一如第二、三章內容）、「英語文素養評量與相關構念調查工具發展流程說明」、「資料建置過程說明」（一如第四章中對於學生抽樣、參數估計程序說明）等來提供支持效度之過程證據。另一方面，本書透過分析學生實徵表現數據，來檢核「學生作答內部一致性程度」、「試題難度與鑑別度適切性」、「試題與模式適配度程度」等來提供支持評量工具有效性的內部證據。整體而言，本類群證據大致可呼應過去學者所提測驗內容、認知、反應歷程、內部結構、計分、脈絡等層面 (AERA et al., 2014; Geranpayeh & Taylor, 2013; Khalifa & Weir, 2009; Messick, 1995)。

在支持工具有效性的外部證據方面，其係指本書關注評量結果與其他變項的關聯、或基於外在目的下，測驗所展現效益性、影響性等外部（external）層面（Lissitz & Samuelsen, 2007）。在某程度上，外部效度證據大致雷同於過去學者所提效標關聯、與其他變項關聯等，其中，本書特別將基於標準設定或評量分數使用所產生的後效影響（consequence）獨立出來（Hambleton, 2001; Hambleton et al., 2012; Lissitz & Samuelsen, 2007; Messick, 1989）進行說明，然而，由於評估的後效影響證據並非本書現行論述標的，僅為求論述完整性，條列於此。整體而言，本書現行檢核「學生英語文理解、學習動機與策略使用間關聯」與「不同背景（年級、性別、有無接受補救教學）學生英語文理解、學習動機與策略使用表現差異」是否符合過去理論或實徵預期程度，來提供支持工具效度的外部證據。

表 4-1 效度評估證據來源、對應與項目

證據來源	概念對應面向	項目
內部—過程	測驗內容、認知、反應歷程、內部結構、計分、脈絡	<ul style="list-style-type: none"> <li>– 英語文素養與相關構念理論、架構發展合理性（見第二、第三章）</li> <li>– 英語文素養評量與相關構念調查工具發展流程合理性</li> <li>– 資料建置過程合理性（如抽樣、參數估計程序）</li> </ul>
		<ul style="list-style-type: none"> <li>– 學生作答內部一致性程度</li> <li>– 試題難度與鑑別度適切性</li> <li>– 試題與模式適配度程度</li> </ul>
外部—後效影響	相關構念（或效標）間關聯、測驗分數使用後效影響	<ul style="list-style-type: none"> <li>– 學生英語文理解、學習動機與策略使用間關聯符合預期程度</li> <li>– 不同背景（年級、性別、有無接受補救教學）學生英語文理解、學習動機與策略使用表現差異符合預期程度</li> </ul>
		<ul style="list-style-type: none"> <li>– 測驗結果使用（如搭配標準設定）對於英語文課程綱要回饋或教育系統、政策訂立影響程度</li> </ul>

註：後效影響證據非本書現行論述焦點，然為求完整性，條列於本表。此外，標準設定中切截分數有效性評估已部分發表於謝進昌（2022b），為避免重複，並未重述於本書。

## 第二節 支持效度之過程證據評估

有關英語文素養大型評量工具的發展，本書採用標準化流程來推動，而相關重要元素包含大型評量設計與目的釐清、轉化「教育目標」為評量指引、與學生表現標準描述、題庫建置與線上施測系統、代表性抽樣與施測作業說明、正式施測與參數估計、標準設定與切截分數有效性評估、效度評估與研究目的回應等面向。某些元素已經於本書前面章節，進行說明，而某些是於本節進行陳述。

### 一、整體評量與調查工具發展標準化流程

本書英語文素養大型評量發展標準化流程，大致可繪製如圖 4-2 所示。就大型評量設計與目的而言，相關概念發想大致延伸自先前臺灣學生成就評量資料庫（TASA）目的與教育部因應 108 課綱轉換之影響評估需求，原始總計畫主持人企圖透過蒐集具臺灣代表性學生於五個主要學科的素養表現實徵數據，來回饋國家課程綱要，其中，本書主要聚焦第四學習階段（國中）英語文素養評量。然而，有別於 TASA 針對各學習階段橫斷面學生（如四、六、八與十一年級）實施評量與調查，TASAL 係使用長期追蹤模式，針對選取同批學生進行長期性評量與追蹤調查，其示意如圖 4-3 所示。在 107 學年，TASAL 會正式針對最後一屆接受九年一貫課綱之七年級學生進行正式評量；至 108 學年時，除了同時追蹤同一群最後一屆接受九年一貫課綱之七年級升至八年級學生外，另外，TASAL 會針對第一屆接受 108 課綱之七年級學生進行正式評量；後續，待時間來到 109、110 學年，TASAL 同樣持續追蹤該批接受不同課程綱要學生至其結束該階段學習。TASAL 特別設計長期追蹤研究，目的在於檢核國家課程轉換前後，不同世代學生當下表現與其隨年級成長表現，可能受到課程轉換影響程度，並從中探究可能影響因素或課綱調整面向（國家教育研究院，無日期）。其中，由於新型冠狀病毒對臺灣產生之直接影響，教育部於 2021 年 5 月宣布取消學校實體課程，使得 TASAL 同步停止最後一屆接受九年一貫課綱九年級的正式施測，因此，本書現行僅透過 2019 年七年級、與 2020 年八年級學生表現實徵資料，來進行結果呈現與說明。

在釐清評量目的與設計後，一如本書第二章所述，本書經回顧文獻，企圖透過 Bloom 等人（1956）建議，以認知成分（動詞使用）與知識內容（名詞使用）方式

(Anderson et al., 2001)，轉化英語文課程綱要所界定教育目標為評量架構，後續，在經過數名學科專家及教師諮詢、修訂，預先確立初步評量架構與同步建立學生表現標準描述草案。在此基礎下，本書同時期透過滾雪球方式，亦即由審題教師推薦優良命題教師，經說明徵求意願與目的、方向後，逐年建置合作的英語文命、審題人才庫，共同努力透過遞迴式程序，進行試題編製、預試、修審等作業，發展第四學習階段七、八與九年級學生英語文理解題庫。另一方面，有關學生英語文學習動機與策略使用問卷發展，同樣依循此標準化流程進行。

在題庫建置後，配合 TASAL 整體研究推動，選取七年級具臺灣代表性學生樣本，在經過施測學校辦理講習與說明後，進行正式英語文評量與問卷調查。其中，後續八與九年級學生之追蹤參與者，同樣在徵詢被抽樣學校是否有意願繼續參與情況後，持續追蹤原抽樣班級學生升至新年級之正式評量表現。TASAL 完成正式評量與問卷調查後，會經資料整理與檢核，提供各學科與問卷作答反應資料予各學科研究者進行相關分析。本書在取得原始學生英語文表現與問卷資料，係透過 R 平台套件發展相關程式語法，進行各項參數估計與次級資料分析。同時，考量本書目的為透過標準 (standard) 詮釋學生表現，相關內涵屬於標準參照模式 (standard-based assessment)，因此，本書陸續於 2020 年 4 月 8 日、2020 年 5 月 20 日由專家判讀取徑，分別針對英語文理解認知表現 (謝進昌, 2022b)、策略使用 (謝進昌, 2021b) 進行切截分數的標準設定 (standard setting)，為詮釋學生分數表現提供參照依據，其中，由於上述標準設定結果之效度評估，已於相關期刊發表，故於本書僅略提及，不做詳細具體說明。

在取得正式學生實徵表現資料與確立標準設定程序後，本書再次調整學生表現標準描述，而一如本書第二章第五節所述，相關學生表現標準描述係已經多次調整與修訂後，較完整之現行版本。後續，本書同樣得配合更新的學生表現數據，來進行相關文件遞迴式調整與修正，以提供有效性評估。整體而言，本書係配合 TASAL 所蒐集各年度學生於正式評量之英語文表現實徵資料，並透過文獻回顧、理論與評量架構建構、標準設定等程序，針對本書所需長期累積支持效度之各項實徵證據與政策議題回饋、影響評估等進行回應。

圖 4-2 標準化評量與調查工具發展流程圖

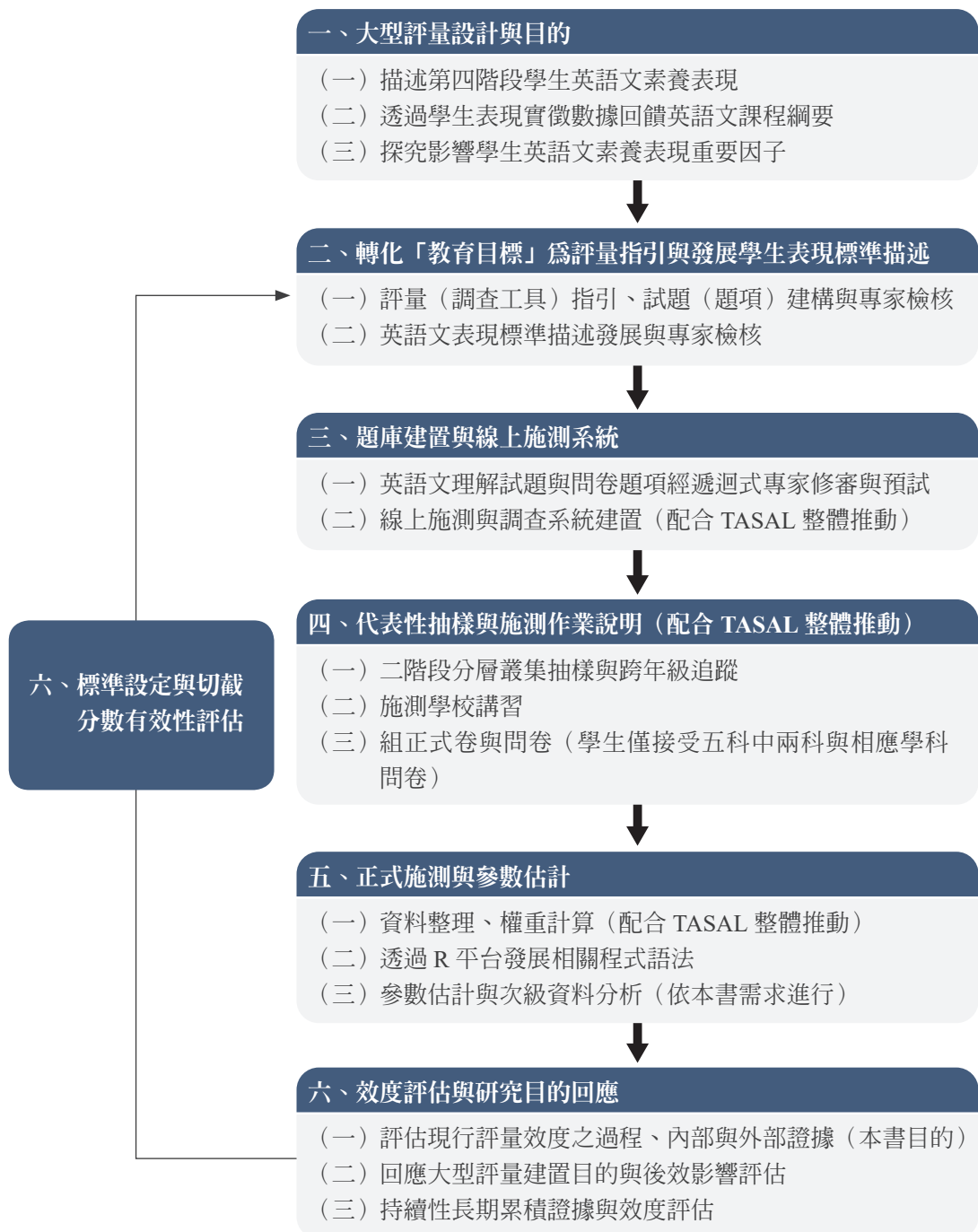
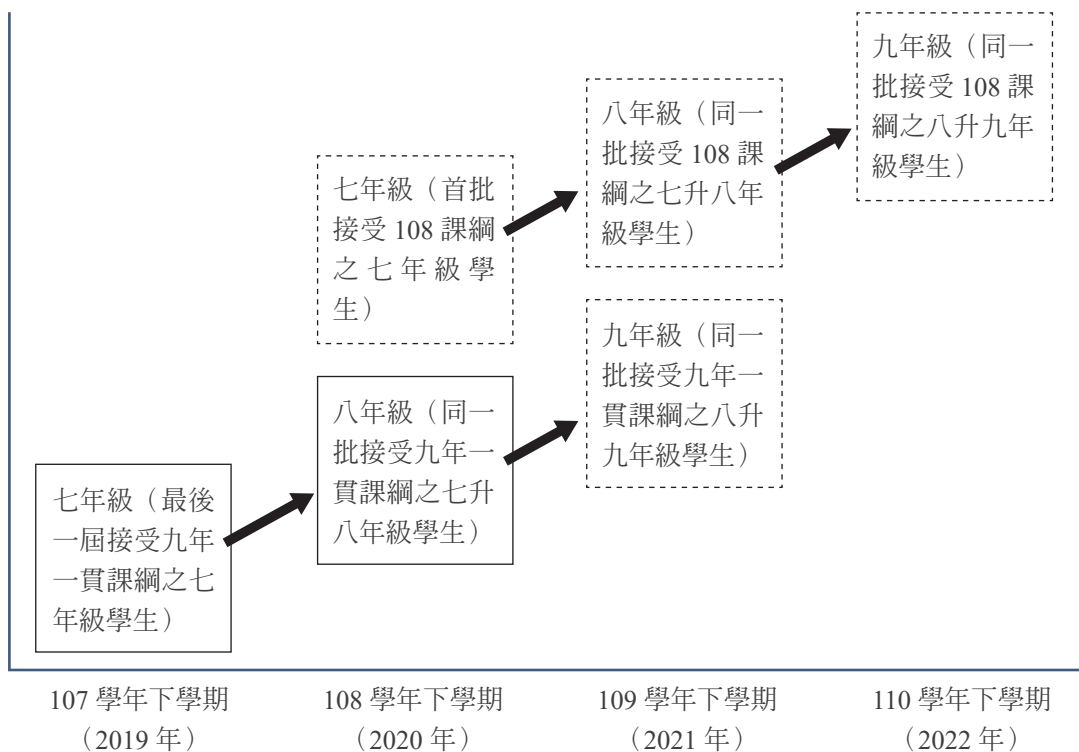


圖 4-3 TASAL 長期追蹤施測示意圖



註：1. 2021 年 5 月施測時，因 COVID-19 疫情影響，教育部宣布全國各級學校停止到校上課，故，該次九年級學生評量中途停止。

2. 實線框為本書進行效度評估之學生實徵表現數據來源。

## 二、評量與調查架構的專家審核過程

本書在初步確立英語文素養評量架構後，於 2018 年 3 月起，陸續透過紙本審查方式，邀請 5 位相關領域專家學者，針對評量架構初版進行修訂與提供意見。相關匿名名單與建議概念重點條列如表 4-2 所示。本次所邀專長學者包含 3 位英語文領域專業、2 位涉及閱讀理解領域專業，而相應建議重點有英語文素養定義與評量架構內容釐清、命題與認知成分的對應、素養評量納入情意與策略調查面向、高層次理解成分評量可行性、腔調 / 語言背景、短劇的命題方式、題型納入建構反應試題等，多數建議已融合、調整於本書架構內容。

在前述評量架構基礎上，英語文理解試題的撰寫與修訂，初始是由幾位英語文核心成員進行，然而，隨著時間遞移，本書透過滾雪球推薦方式（即由曾參與命審題教師推薦新進教師），以全臺灣為目標徵詢教師意願，邀請有意願參與之國中、高中英語文教師。這些新進命題教師會經研究方向與評量架構說明、培訓後，進行命題，而教師所擬訂每個試題皆會通過英語文學科專員檢核，覆經資深英語文教授、教師審查與修訂，以作為預試候選題目。為建立初版評量架構，TASAL 於 2017 年 6、9、12 月進行三次預試，以確立可搭配本書所初步建立評量架構進行試題說明，除此之外，隨著整體 TASAL 計畫推動，陸續於 2018 年 5、10、12 月、2019 年 10 月、2020 年 10 月等，配合進行多次小規模試題預試，後續，除透過檢核試題之古典難度與鑑別度，作為試題修訂調整依據外，另邀請英語文外師覆核內容，逐年建立正式英語文評量題庫。

表 4-2 評量架構（初版）審核專家與建議重點

姓名	專長領域	建議概念重點
陳 **	閱讀理解	<ul style="list-style-type: none"> <li>— 應釐清「多元識讀」概念定義與評估方式</li> <li>— 「解碼與字彙」理解的評估可考量「詞彙意義習得策略」相關題型</li> <li>— 「省思評估」是重要語文素養成分，建議以非選擇式建構題來評估</li> <li>— 除現有文本類型，考量以任務為導向的「多文本」來評估學生表現</li> </ul>
陳 **	英語文	<ul style="list-style-type: none"> <li>— 「1-V-15 能聽懂不同腔調 / 語言背景英語使用者的主要內容」新舊課綱差別之一，宜跳脫美籍人士錄音之框架，由口語表達清楚之不同語言背景人士錄音，即可符合此學習表現</li> <li>— 「1-V-7 能聽懂英語短劇主要內容」也可以將廣播劇片段列為聽力評量內容，改變傳統以對話做為評量方式</li> </ul>
詹 **	英語文	<ul style="list-style-type: none"> <li>— 評量的內容架構與文本類型應清楚表列</li> <li>— 其他文內編號與格式應統一一致</li> </ul>

（續下頁）

姓名	專長領域	建議概念重點
柯 **	閱讀理解	<ul style="list-style-type: none"> <li>— 應具體釐清英語文素養定義以及英語文素養評量架構，而非單純文獻回顧</li> <li>— 英語文素養包含情意面是很好方向</li> <li>— 清楚定義所謂的英語文素養</li> <li>— 英語文素養評量架構圖包括認知、情意動機、策略等</li> <li>— 國際評比（PIRLS）四年級已有詮釋整合，省思、線上文本；雖英語是外語，但看出題者如何出題而不是先假設學生沒有能力回應</li> </ul>
張 **	英語文	<ul style="list-style-type: none"> <li>— 技術層面應如何建立各年級學習表現客觀機制是關鍵</li> <li>— 問卷調查學生策略和動機可行，現行雖呈現架構，然問卷題目宜參酌外語領域學者專家意見</li> </ul>

註：1. 專家姓名皆以匿名表徵。

2. 考量篇幅與時空轉換，現行評量架構大多歷經數次調整，當時專家某些建議可能無法於現行本書內容完整理解。因此，本書透過概念重點呈現，以表徵當時建議方向。

有關英語文學習動機、策略使用等問卷發展，本書在初步確立英語文問卷調查架構與題項內容後，同樣於 2018 年 3 月起，陸續透過紙本審查方式，邀請 9 位相關領域專家學者，針對問卷架構與內容進行修訂與提供意見。相關匿名名單與建議概念重點，條列如表 4-3 所示。本次所邀專長學者包含 5 位英語文領域專業、3 位涉及教育心理或閱讀理解領域專業、1 位有關教育測驗與評量領域專業，而相應建議重點有構念定義釐清、題項內容、用詞與格式、題目新增等。本書經確立問卷題項內容後，再邀請國語文專長之教師與中下識讀學生進行試填，以確認文字適切性與可讀性，以作為預試候選問卷試題。同時，在配合 TASAL 整體計畫推動，英語文學生問卷於 2018 年 5、10 月進行預試，而相關工具是於首次預試後，檢核學生填答數據與結果，經微幅調整與修訂後，再次進行第二次預試，以建立正式學生英語文動機、策略等問卷調查內容。

表 4-3 學生問卷審核專家與建議重點

姓名	專長領域	建議概念重點
高 **	英語文	—題項內容、用詞、新增題目等建議，例如，使用 good 及 excellent 為範例能表示記憶策略可幫助學生用簡單的英文字詞去背誦較難、但相同字義的字詞
莊 **	英語文	—題項內容、用詞、新增構念與題目等建議，例如，自我效能某題目可修訂為：老師或同學覺得我的英語文能力很好
黃 **	英語文	—題項內容、用詞等建議，例如，某記憶策略題目部分內容改為：……剛新學到的字詞……
林 **	英語文	—構念定義釐清、題項內容、用詞等建議，例如可透過期望價值理論來詮釋相關英語文學習動機與自我概念定義
蘇 **	教育心理、 閱讀理解	—構念分類釐清、題項內容、用詞等建議，例如建議將後設認知改為理解監控策略
劉 **	教育心理、 閱讀理解	—題項內容、用詞、新增題目等建議，例如，新增「學習英語文時，我會一邊讀一邊確認自己哪裡不懂」
黃 **	教育測驗 與評量	—題項內容、用詞、格式等建議，例如，某題項刪除建議
林 **	英語文	—題項內容、用詞、題意釐清等建議，例如，某題項所述「英語文相關問題」可具體舉例說明
蔡 **	閱讀理解	—題項內容、用詞、題意釐清等建議，例如，修訂某題項語句連貫性為「我會試圖把單字的發音和單字代表的圖像連結起來」

註：1. 專家姓名皆以匿名表徵。

2. 考量篇幅與時空轉換，現行問卷內容歷經些許調整，當時專家某些建議可能無法於現行本書內容完整理解。因此，本書透過概念重點呈現，以表徵當時建議方向。
3. 本書所涉及問卷調查構念主要為學習動機與策略使用，然而，原始英語文問卷內容另包含學生英語文學習先備經驗、行為投入、學習活動與素材等構念，考量這些內容非本書現行論述方向，故在此暫省略。

### 三、英語文理解試題與學生問卷題項組卷、分布

以下茲就英語文理解試題組卷、題目分布、問卷題項配置等，進行說明。

#### (一) 英語文理解試題組卷

爲了涵蓋大範圍國家課程綱要所界定內容，本書一方面須發展廣泛（*extensive*）且同時平衡（*balanced*）兼顧各內容面向之英語文評量題庫，以回應內容標準，其中，若以評量七年級學生工具爲例，本書發展了 182 題英語文理解試題，才得以有效涵蓋課綱所界定英語文部分內容（如閱讀與聽力）範疇；然而，另一方面，在有限評量時間與邏輯可行性下，學校學生往往僅能接受部分測驗試題。因此，TASAL 透過測驗試題矩陣抽樣設計（*matrix sampling of test items*）方式進行組卷（Childs & Jaciw, 2002; Gonzalez & Rutkowski, 2010），才得以兼顧理論與實務可行性。本書對於英語文素養評量主要採用矩陣抽樣中部分平衡不完整區塊設計（*partial balanced incomplete block design, PBIB*），而其概念是由單一題目組成試題區塊（*item blocks*），再由試題區塊組成題本（*booklets*），而不同題本間是透過螺旋（*rotated*）方式以試題區塊組成，使得跨題本間存在共同區塊，且各試題區塊出現次數是相等的，但因成對試題區塊（*paired item blocks*）於題本中出現次數不同，故稱其爲部分平衡。

本書各年級英語文理解實際組卷設計結果分別如表 4-4、4-5、4-6 所示。就正式評量七年級試題組卷而言，總計開發 182 題，其中包含閱讀 93 題、聽力 89 題。依照試題特性均衡，本書可區分爲 13 個區塊（E01-E13），而每個區塊包含 14 個試題。本書以 2 個區塊組成 1 個題本，例如，題本 1 是由區塊 1、2 組成，題本 2 爲區塊 2、3 組成，以至題本 13 爲區塊 13、1 組成。同時，爲平衡題本內區塊位置影響學生作答的可能，故將兩兩區塊互換位置組成另外 13 題本，總計組成 26 題本。每位七年級學生僅接受其中 1 個題本，共 28 題。此組卷概念除部分平衡各區塊出現次數，同時，各題本間皆有 1 個共同區塊進行連結，以利後續透過等化（*equating*）進行參數估計。最後，考量特殊生（聽障生）無法接受聽力理解評量，因此，本書透過區塊 1、2、3、4 中的閱讀理解試題來組成題本 27。

有關八、與九年級學生所評量試題的組卷，相關題本試題的組成概念，大致是依照上述螺旋方式進行組卷。然而，在考量跨年級垂直等化的需求，因此，八年級學生組卷會從評量七年級學生試題題庫中，選取區塊2、4、6、8、10、12作為垂直定錨共同區塊，並搭配新開發區塊14至20區塊（E14-E20）（每區塊有16題），進行八年級學生題本組卷；沿用相同做法，九年級學生的評量題卷同樣會從八年級學生評量題庫中，選取區塊14至20，並搭配新開發區塊21至27（E21-E27）（除E25為16題，其餘各有17題），進行評量九年級學生題本組卷。

表 4-4 七年級學生評量試題題本組卷設計

題本	第一個區塊	第二個區塊
題本 01	E01	E02
題本 02	E02	E03
題本 03	E03	E04
題本 04	E04	E05
題本 05	E05	E06
題本 06	E06	E07
題本 07	E07	E08
題本 08	E08	E09
題本 09	E09	E10
題本 10	E10	E11
題本 11	E11	E12
題本 12	E12	E13
題本 13	E13	E01
題本 14	E02	E01
題本 15	E03	E02

（續下頁）

題本	第一個區塊	第二個區塊
題本 16	E04	E03
題本 17	E05	E04
題本 18	E06	E05
題本 19	E07	E06
題本 20	E08	E07
題本 21	E09	E08
題本 22	E10	E09
題本 23	E11	E10
題本 24	E12	E11
題本 25	E13	E12
題本 26	E01	E13
題本 27 (聽障生)	E1 (閱讀)、E2 (閱讀)	E3 (閱讀)、E4 (閱讀)

表 4-5 八年級學生評量試題題本組卷設計

題本	第一個區塊	第二個區塊
題本 01	E02	E14
題本 02	E04	E14
題本 03	E04	E15
題本 04	E06	E15
題本 05	E06	E16
題本 06	E08	E16
題本 07	E08	E17

(續下頁)

題本	第一個區塊	第二個區塊
題本 08	E10	E17
題本 09	E10	E18
題本 10	E12	E18
題本 11	E12	E19
題本 12	E19	E20
題本 13	E02	E20
題本 14	E14	E02
題本 15	E14	E04
題本 16	E15	E04
題本 17	E15	E06
題本 18	E16	E06
題本 19	E16	E08
題本 20	E17	E08
題本 21	E17	E10
題本 22	E18	E10
題本 23	E18	E12
題本 24	E19	E12
題本 25	E20	E19
題本 26	E20	E02
題本 27 (聽障生)	E2 (閱讀)、E4 (閱讀)	E14 (閱讀)、E16 (閱讀)

表 4-6 九年級學生評量試題題本組卷設計

題本	第一個區塊	第二個區塊
題本 01	E14	E21
題本 02	E15	E21
題本 03	E15	E22
題本 04	E16	E22
題本 05	E16	E23
題本 06	E17	E23
題本 07	E17	E24
題本 08	E18	E24
題本 09	E18	E25
題本 10	E19	E25
題本 11	E19	E26
題本 12	E20	E26
題本 13	E20	E27
題本 14	E14	E27
題本 15	E21	E14
題本 16	E21	E15
題本 17	E22	E15
題本 18	E22	E16
題本 19	E23	E16
題本 20	E23	E17
題本 21	E24	E17

(續下頁)

題本	第一個區塊	第二個區塊
題本 22	E24	E18
題本 23	E25	E18
題本 24	E25	E19
題本 25	E26	E19
題本 26	E26	E20
題本 27	E27	E20
題本 28	E27	E14
題本 29 (聽障生)	E14 (閱讀)、E16 (閱讀)	E22 (閱讀)、E24 (閱讀)

## (二) 英語文理解試題分布

在前述，本書已經說明如何透過部分平衡不完整區塊設計的組卷方式，來回應英語文國家課程綱要所界定內容標準。在此，進一步說明實際發展的各個年級英語文理解試題分布概況。

### 1. 不同語言活動所設計試題的分布

就試題所對應的語言活動而言，相關英語文理解試題分布結果如表 4-7 所示。評量七年級學生者，分別計有 93 題 (51.1%) 閱讀理解試題、89 題 (48.9%) 聽力理解試題；此外，評量八年級學生者，分別計有 102 題 (52.0%) 閱讀理解試題、94 題 (48.0%) 聽力理解試題。其中，在八年級的組卷上，閱讀理解計有 46 題 (占七年級 93 題閱讀之 49.5%、占八年級 102 題閱讀之 45.1%)、聽力理解計有 38 題 (占七年級 89 題聽力之 42.7%、占八年級 94 題聽力之 40.4%) 是取自於七年級題卷的垂直共同題。此外，評量九年級學生的題卷，計有 118 題 (51.3%) 閱讀理解試題、112 題 (48.7%) 聽力理解試題，同樣，其內各有 56 題是來自八年級題卷，作為垂直定錨之閱讀、聽力理解試題，分別占八年級整體閱讀題數 54.9%、聽力 59.6% 與九年級整體閱讀 47.5%、聽力 50.0%。就試題情境設計為視

聽影片與數位網頁而言，由於這類情境的試題並非本書評量主軸，因此，相關試題的比重並不高，總計僅發展出 30 題視聽影片試題、與 15 題數位網頁試題，其中，考量這類試題略有別於單純閱讀或聽力篇章之理解特徵，因此，特別針對其發展進行說明。

整體而言，不同語言活動試題分布，各自皆占近五成比重，而作為跨年級之垂直定錨試題也有四成以上比重，是具有相當平衡性，同時，此分佈也大致符合一般大型評量在設定跨年度定錨題比重水準。

表 4-7 各年級閱讀、聽力理解試題之分布

年級	閱讀理解	聽力理解	總題數
七年級	93 (51.1%)	89 (48.9%)	182 (100%)
八年級	102 (52.0%) (46題取自七年級)	94 (48.0%) (38題取自七年級)	196 (100%)
九年級	118 (51.3%) (56題取自八年級)	112 (48.7%) (56題取自八年級)	230 (100%)
第四學習階段	211 (51.2%)	201 (48.8%)	412 (100%)

註：第四學習階段分別發展有 30 題涉及視聽影片試題、15 題涉及數位網頁篇章試題。其中，七年級有 15 題視聽影片、15 題數位網頁；八年級有 13 題視聽影片、15 題數位網頁；九年級有 13 題視聽影片，無數位網頁試題。

## 2. 不同認知成分所設計試題的分布

就試題所對應理解認知成分而言，相關試題分布結果如表 4-8 所示。本書對於相關試題分佈配置大致考量幾個需求因素，一是考量未來在垂直銜接低、與高學習階段（如銜接六年級或十年級）的試題需求；另一則是依照學生能力表現，會逐年級增加高層次認知成分的試題比重。如此，試題才得以適當銜接未來跨學習階段設計與發展。

就七年級學生試題評量重點分布而言，評量學生對於篇章字面文意理解的試題占多數，計有 84 題（46.2%），其次為評量學生對於篇章隱含意理解試題，計有 45 題（24.7%）、直接找到文中明顯陳述訊息試題，計有

28 題 (15.4%)。此外，本書另配置 21 題 (11.5%) 在評量學生是否能流暢運作語言基本知能，而主要考量在於低學習階段學生 (如五與六年級) 評量內容，多半關注英語文基本知識的熟悉度。因此，相關試題配置可作為未來作為垂直定錨共同題使用，來銜接第三學習階段學生英語文表現。另一方面，隨著 TASAL 持續推展與學生年級增長，本書所分配於評量較高層次認知成分的試題比重，也會隨之增加，例如，評量學生隱含文意理解的試題比重，是自七年級的 24.7%、八年級的 32.2%、逐步增加至九年級的 38.3%。

有關評量學生評估與省思能力試題的配置，本書在檢視附錄一英語文國家課程綱要學習表現與認知成分對應，可發現大多數相關評量指引皆屬於第五學習階段 (高中職)，加諸，相關試題設計多數得仰賴學生在理解文意後，透過英語文來表達 (expression)、陳述或寫出來，才足以反映學生認知歷程，而在考量此部分對於低年級學生是具有難度，使得本書目前所配置題數比重會相對偏少。同時，在考量兼顧未來垂直銜接第五學習階段學生表現的需求上，本書在評量九年級學生仍配置 7.4% 試題於評量評估與省思層次能力，以期望相關試題比重是具有與第五學習階段學生表現銜接的可行性。

表 4-8 各年級不同理解認知成分試題之分布

年級	流暢運作 語言基本 知能	直接找到 訊息	理解 字面意	理解 隱含意	能透過 自身觀點 進行評估	能省思 認識自我	總題數
七年級	21 (11.5%)	28 (15.4%)	84 (46.2%)	45 (24.7%)	4 (2.2%)	--	182 (100%)
八年級	--	17 (8.7%)	104 (53.1%)	65 (32.2%)	9 (4.6%)	1 (0.5%)	196 (100%)
九年級	--	15 (6.5%)	110 (47.8%)	88 (38.3%)	14 (6.1%)	3 (1.3%)	230 (100%)
第四學 習階段	21 (5.1%)	43 (10.4%)	194 (47.1%)	133 (32.3%)	18 (4.4%)	3 (0.7%)	412 (100%)

### 3. 不同文本與篇章類型試題分布

就不同文本與篇章類型試題分布而言，相關試題分布結果如表 4-9 所示。以評量七年級學生試題分布為例，結果顯示以評量第一類群篇章居多，計有 126 題（69.2%），而其內容主要包含描述、教學指引、人際互動等篇章，大多為學生一般生活或學術情境常見篇章內容；試題占比其次為敘事與文學與語言知能文本，各計有 21 題（11.5%）；最後則是第二類群篇章，其內容包含說明文與議論文。由於第二類群文本內容大多涉及抽象概念或意見陳述，使得學生對於相關題目理解難度，會傾向較高，使得本配所配置比重會相對較低，計有 14 題（7.7%）。另一方面，在考量隨著學生年級的增長，第一類群篇章題目持續維持高比重，然而，第二類與第三類群篇章試題的題數安排，則會跟隨年級的增長而提高比重，使得第二類篇章試題比重分別是八年級的 17.3%、九年級的 25.2%，而第三類群篇章試題比重分別為八年級的 16.3%、九年級的 23.5%。

表 4-9 各年級不同文本與篇章類型試題之分布

年級	語言知能文本	第一類群篇章			第二類群篇章		第三類群篇章	總題數
		描述類	教學指引類	人際互動類	說明文	議論文	敘事與文學	
七年級	21 (11.5%)	126 (69.2%)			14 (7.7%)		21 (11.5%)	182 (100%)
		32 (17.6%)	14 (7.7%)	80 (44.0%)	11 (6.0%)	3 (1.6%)		
八年級	--	130 (66.3%)			34 (17.3%)		32 (16.3%)	196 (100%)
		46 (23.5%)	7 (3.6%)	77 (39.3%)	20 (10.2%)	14 (7.1%)		

(續下頁)

年級	語言知能文本	第一類群篇章			第二類群篇章		第三類群篇章	總題數
		描述類	教學指引類	人際互動類	說明文	議論文	敘事與文學	
九年級	--	118 (51.3%)			58 (25.2%)		54	230 (100%)
		29 (12.6%)	9 (3.9%)	80 (34.8%)	21 (9.1%)	37 (16.1%)	(23.5%)	
第四階段	21 (5.1%)	244 (59.2%)			72 (17.5%)		75	412 (100%)
		61 (14.8%)	23 (5.6%)	160 (38.8%)	32 (7.8%)	40 (9.7%)	(18.2%)	

#### 4. 文本平均字數

就文本長度而言，雖然本書在附件範例，聽力文本是以播放秒數（seconds）來表徵，然而，為力求閱讀與聽力文本具有可比較基礎，本書在此統一以平均字數（words）來表徵。

在閱讀與聽力理解篇章字數分析方面，結果如表 4-10 所示。評量七年級學生閱讀理解文本的平均字數為 109.3、聽力理解文本平均字數為 104.8，而整體文本平均字數為 107.1。其中，隨著學生評量年級的增長，不論是閱讀或聽力理解文本平均字數，皆隨之增加，是符合愈高年級學生須理解愈長文本長度之要求。此外，少數視聽影片與數位網頁篇章，考量其本身的情境特性，平均文本長度皆會高於一般閱讀或聽力理解篇章，相關文本的文本字數分別為 222.9、209.7。

就不同類型篇章文本長度而言，結果如表 4-11 所示。就評量七年級學生文本為例，第一類群篇章平均字數為 111.6、第二類群篇章平均字數為 123.7、第三類群篇章平均字數為 170.0，而語言知能文本平均字數為 6.3，且隨著評量學生年級的增加，不同類群文本長度也會隨著學生年級能力而

增長。整體而言，文本長度大致視該評量學生年級能力而選擇相應適合的長度進行評量。

表 4-10 閱讀與聽力理解篇章文本平均字數

年級	閱讀理解	聽力理解	整體平均字數
七年級	109.3	104.8	107.1
八年級	148.4	140.2	144.4
九年級	203.9	143.4	174.4
第四學習階段	162.2	126.3	144.7

註：第四學習階段內有 30 題視聽影片、15 題數位網頁，其文本平均字數分別為 222.9、209.7。其中，七年級視聽影片文本平均字數為 209.7、數位網頁文本平均字數為 203.9；八年級視聽影片文本平均字數為 209.7、數位網頁文本平均字數為 149.8；九年級視聽影片文本平均字數為 241.8、無數位網頁試題。

表 4-11 各年級不同類型文本平均字數

年級	語言知能文本	第一類群篇章			第二類群篇章		第三類群篇章	整體平均字數
		描述類	教學指引類	人際互動類	說明文	議論文	敘事與文學	
七年級	6.3	111.6			123.7		170.0	107.1
		118.9	99.6	110.7	119.3	140.0		
八年級	--	130.8			144.2		200.2	144.4
		133.7	70.4	134.4	124.1	172.9		
九年級	--	130.2			207.2		236.0	174.4
		118.2	93.3	138.7	187.5	218.4		
第四學習階段	6.3	120.6			191.0		217.5	144.7
		118.6	97.1	124.7	164.0	212.5		

### (三) 學生問卷題項配置

英語文學生問卷除了調查本書所關注之英語文學習動機、理解策略使用與相關學生背景特徵外，題項內容另包含學生學習英語文先備經驗、對於英語文各類學習投入時間、英語文學習素材及活動參與、TASAL 配合其他學科調查所設計一般性問卷等。然而，本書考量某些構念與背景題項並非本書現行論述焦點，故，在此僅討論本書關注焦點構念與所需背景變項，大致包含英語文學習動機、策略使用與性別、有無接受補救教學等。

有關學生英語文學習動機題項，結果如附錄十與表 4-12 所示。本書對於英語文學習內三個構念，分別各自發展出 7 個題項，並採用四點量表進行計分，自非常同意、同意、不同意至非常不同意等表徵選項 1 至 4。本書再經過反向計分，使得學生作答總分分數愈高，代表其所秉持正向學習動機傾向會愈高。此外，在考量學生總作答時間可負荷（實際調查還包含一般性學生問卷題目，非僅止於學生學科問卷）、題數所涉及測量精準性（以至少四個題項為考量）與跨年級調查必要性等因素後，八年級學生僅調查內在動機、學科自我概念等構念、與其下部分題項，而九年級學生僅調查內在動機、社會動機與其下部分題項。

有關調查學生英語文理解策略使用，鑑於相關內容作者已部分發表於謝進昌（2021b），因此，本書僅進行簡略說明，而該工具相關內容如附錄十一與表 4-13 所示。理解策略的四個構念分別各自發展出 6、6、8、10 個題項，採用四點量表計分，自常常如此、有時候、很少如此至從來沒有等表徵選項 1 至 4，本書再經過反向計分，使得學生作答總分分數愈高，代表學生在理解時，使用該策略的傾向或頻率會愈高。同樣，在考量測量精準性與學生總作答時間後，對於八年級、與九年級學生調查，仍維持調查學生在四個構念表現，然而，各個構念題項數目僅選擇原始所發展全部總題目內的 5 個題項，來進行調查。

對於學生相關背景特徵調查，本書目前僅涉及包含性別、有無接受補救教學等變項。結果如表 4-14 所示，學生性別資料來自抽樣階段於收集學生名冊時，由教師代為確認學生為男生或女生，而相關資料會經重新編碼後，將男生編為 1，女生編為 0。此外，有關學生是否參與學習扶助（補救教學）的

調查，本書係透過問卷分別調查學生在七年級、與八年級接受評量當下，在「過去一年內，你是否曾經參加學校舉辦的英語文學習扶助（補救教學）課程？」，而相關選項包含「沒有參加、一週 1 節、一週 2 節、一週 3 節、一週 4 節」等。其中，本書將學生勾選含一週 1 節以上者，合併為「有參加補救教學」，編碼為 1，而沒有參加補救教學者，編碼為 0。

表 4-12 英語文學習動機問卷題項跨年級組卷

向度	題號	元素	七年級	八年級	九年級
內在動機	P1	樂於參與	V	V	V
	P2	樂於參與	V	-	-
	P3	樂於參與	V	V	V
	P4	樂於參與	V	-	-
	P5	沉浸學習	V	V	V
	P6	好奇 / 尋求新經驗	V	V	V
	P7	好奇 / 尋求新經驗	V	-	-
英語文自我概念	E1	過去經驗	V	V	-
	E2	自我肯定	V	-	-
	E3	自我肯定	V	-	-
	E4	參照比較	V	V	-
	E5	參照比較	V	-	-
	E6	自我肯定	V	V	-
	E7	他人肯定	V	V	-

(續下頁)

向度	題號	元素	七年級	八年級	九年級
社會動機	A1	同儕	V	-	V
	A2	同儕	V	-	-
	A3	同儕	V	-	V
	A4	同儕	V	-	V
	A5	社群	V	-	-
	A6	師生	V	-	V
	A7	社群	V	-	V

註：1. 因疫情影響，影響 2021 年九年級學生正式調查，但為求說明完整性，本書仍條列於此表。

2. V 代表該題項被分派調查該年級學生，而各個題項內容請見附錄十。

表 4-13 英語文理解策略使用問卷題項跨年級組卷

向度	題號	元素	七年級	八年級	九年級
記憶策略	M1	圖像化	V	V	V
	M2	聲韻	V	-	-
	M3	脈絡化	V	V	V
	M4	搭配詞 / 同義詞	V	V	V
	M5	字根 / 首 / 尾	V	V	V
	M6	音節	V	V	V
認知策略	C1	快速瀏覽	V	V	V
	C2	重點標示 / 畫底線	V	V	V
	C3	摘取大意	V	V	V
	C4	文法規則	V	V	V
	C5	語調 / 氣	V	-	-
	C6	文本架構	V	V	V

(續下頁)

向度	題號	元素	七年級	八年級	九年級
推論連結 策略	I1	上下文脈絡	V	V	V
	I2	圖表 / 標題	V	V	V
	I3	上下文脈絡	V	-	-
	I4	大意理解	V	-	-
	I5	關鍵字詞	V	V	V
	I6	肢體動作 / 表情	V	-	-
	I7	背景知識 / 經驗	V	V	V
	I8	概念關係連結	V	V	V
理解監控 策略	S1	監控	V	-	-
	S2	調整	V	V	V
	S3	檢驗	V	-	-
	S4	檢驗	V	-	-
	S5	檢驗	V	V	V
	S6	監控	V	V	V
	S7	監控	V	-	-
	S8	調整	V	V	V
	S9	調整	V	-	-
	S10	評估	V	V	V

註：1. 因疫情影響，影響 2021 年九年級學生正式調查，但為求說明完整性，本書仍條列於此表。相關工具發展已發表於謝進昌（2021b）。

2. V 代表該題項被分派調查該年級學生，而各題項內容請見附錄十一。

表 4-14 本書所涉及學生背景特徵資料來源

特徵	來源	選項	重新編碼
性別（註 1）	在確認被抽樣班級內學生名冊時，由教師代為填寫	①男生 ②女生	①重編為 1 – 男生 ②重編為 0 – 女生
七、八年級是否參與補救教學（註 2）	問卷題項：過去一年內，你是否曾經參加學校舉辦的英語文學習扶助（補救教學）課程？	①沒有參加 ② 1 週 1 節 ③ 1 週 2 節 ④ 1 週 3 節 ⑤ 1 週 4 節	①重編為 0 – 沒有接受補救教學 ②至⑤重編為 1 – 有接受補救教學

註：1. 有關學生性別資料，TASAL 除了透過問卷調查外，另請學校教師代為填寫相關學生基本資料，以重覆確認正確性，本書係採用教師填答資料來進行相關分析。

2. 有關學生是否參與學習扶助（補救教學），透過問卷調查題項分別於學生七、八年級時，請學生填寫。

#### 四、代表性抽樣、權重與參與者人數

對於選取參與英語文評量的學生，相關抽樣方式與權重計算係配合 TASAL 整體計畫推動來進行。相關抽樣框架與概念也大致雷同於 Trends in International Mathematics and Science Study (TIMSS) 與 TASA 抽樣程序與權重計算（黃馨瑩，2017；LaRoche et al., 2020），唯有部分內容，如跨年級學生橫斷面權重，會因應需求有所調整。簡要說明如下。

##### （一）抽樣與追蹤

對於接受舊課綱最後一屆（107 學年）七年級學生的抽樣方式，TASAL 採用兩階段分層叢集抽樣（stratified two-stage cluster sampling）。其概念是在第一階段就各分層（stratum）內，進行標的學校（target schools）選取，而在第二階段，會就所選取學校，進行學校內標的班級（target classes）選取。在第一階段，相關分層係包含顯性（explicit）與隱性（implicit）分層，其中，顯性分層係依照教育部統計處（無日期）107 學年度偏遠地區學校界定，區分為一般學校與偏遠學校（包含偏遠、特偏與極偏）；隱性分層係由一向臺灣

師範大學心理與測驗中心申請國中教育會考數學平均成績，將學校由高至低區分為六個層級，另一則是依照侯佩君等人（2008）所提對於全臺灣鄉鎮市都市化等級建議，並將五至七層合併（並納入離島），共區分為五個層級。在確立分層後，相關實際抽樣係針對各顯性分層內，依據學校規模大小成比例之系統性抽樣（systematic sampling with probability proportional to sample size）抽出學校。

在進行學校抽樣時，相關操作概念是先給予一個隨機亂數起始點，再依照所計算出之固定間隔區間（interval），將已經依照上述隱性分層由高至低、學校規模由大至小排序的學校進行系統性抽樣，以選取出整體所需學校數。在第二階段，TASAL 會在被選取學校內，隨機抽取某一班級全部學生進行施測，其中，若被選取學校無意參與評量，則會改以有意願且在前述排序位置相鄰近（即學校規模大小類似）的替代學校來進行意願調查，直至選取到受測學校。由於實際 TASAL 抽樣作業並非本書作者所執行，僅能從理論、概念層次進行描述。詳細與具體完整說明可見 TASAL 官方網站。

## （二）權重

有關權重計算，本書主要討論學生橫斷面權重計算，而其目標係用以回推 2019 年首次抽樣的七年級學生所映照橫斷面母群概況。簡要說明如下。

有關橫斷面權重，相關概念涉及三個元素，分別是學校、班級與學生基礎權重（base weight）。學校基礎權重（school base weight）係為該分層內，該學校被選中機率，乘以選取次數之倒數。相關概念如公式 1 所示， $M_h$  為第  $h = 1, \dots, H$  分層內的總學生數， $n_h$  為第  $h = 1, \dots, H$  分層內的抽樣學校數， $M_{hi}$  為第  $h = 1, \dots, H$  分層內之第  $i$  間（ $i = 1, \dots, n_h$ ）學校內學生數。此外，班級基礎權重（class base weight）如公式 2 所示，係為該學校內班級總數  $C_i$ ，除以被選取班級數  $c_i$ 。有關學生基礎權重（student base weight），相關概念如公式 3 所示，由於被選取班級學生皆參與施測，因此，學生基礎權重通常為 1，即全部班級學生  $S_i$  等於實際參與學生  $s_i$ 。若實際施測時，有未到考缺席者，則會依照公式 3 概念，就實際參與學生數進行權重校正。有關學生總權重（student total weight），相關概念如公式 4 所示，為學校、班級與學生基礎（或其校正）權重的總乘積。此外，由於 TASAL 並未出現被選取學校或班級為非母群界定

者或停招等情況，因此，本書即省略學校、班級因缺失而進行校正權重計算之說明。另一方面，由於每位學生僅接受 TASAL 五科中的兩科測驗，因此，若是以學科（如本書英語文）為焦點，進行分析所使用計算，上述學生總權重須再乘以 2.5（即 10 種學科測驗組合中，各科皆出現 4 次），才可得出正式使用學生總權重。

有關八年級橫斷面學生權重，相關計算概念雷同於前述，而目標為回推 2019 年七年級母群表現。其概念如圖 4-4 所示，實心圓係代表 2019 年七年級母群而虛線圓代表 2020 年這群學生升至八年級概況，從圖中可發現實際權重計算細節會涉及升至八年級時轉入與轉出學生，使得權重計算會衍伸出不同類型，一是將轉出（學）生視為未到考，進行學生校正權重計算；另一是將轉出（學）生視為補充樣本，以分享該班相同權重來回推，進行學生校正權重計算。本書擇後者來進行權重校正計算，如此，八年級整體加權結果得以直接映照或回推 2019 年七年級學生母群概況。

$$W_{SCH} = \frac{M_h}{n_h \times M_{hi}} \quad (1)$$

$$W_{CLA} = \frac{C_i}{c_i} \quad (2)$$

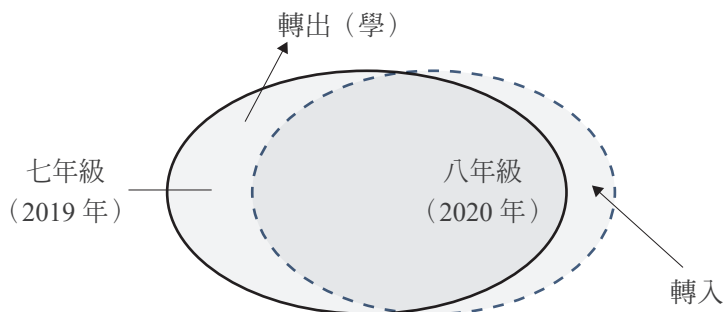
$$W_{STU} = \frac{S_i}{s_i} \quad (3)$$

$$W_{TOT} = W_{SCH} \times W_{CLA} \times W_{STU} \quad (4)$$

前述的權重概念，係作回推整體母群表現概況之用，稱為擴張式權重（expansion weight），然而，在進行顯著性考驗時，為避免樣本數膨脹，會使用相對式權重（relative weight）（Groves et al., 2011）或 TIMSS 稱為 Houseweight（黃馨瑩，2017；LaRoche et al., 2020），而其計算概念為「學生總權重除以母群總人數，再乘以實際樣本人數」。整體而言，本書七、八年級學生之橫斷面權重，是以上述概念進行計算，而其差異僅在於缺席或轉入、轉出學生所衍伸之校正權重界定差異。此外，值得一提的是，本書所使用資料目前是以橫斷面分別檢視七與八年級學生英語文表現，尚未涉及同一群七升八年級學生之縱貫性表現，故，在此暫不說明縱貫性權重估計概念。最後，

由於實際權重計算並非本書作者執行，僅從概念層次進行描述，詳細與具體完整說明請見 TASAL 官方網站。

圖 4-4 權重計算示意圖



### (三) 參與者人數與分布

在上述抽樣概念下，TASAL 初步抽出 256 校，經詢問被選取學校參與意願，針對拒絕校數，再選取替代校數。最後，本書排除 1 所因實驗教學成效評估目的而納入追蹤的學校後，初步取得接受舊課綱最後一屆七年級（107 學年）正式參與數為 245 所學校、269 班，共計 6,970 名學生。該群學生經升至八年級時，有 244 名學生轉入、212 名轉出，後共計 7,002 名學生。由於學生僅參與五科中兩科測驗，因此，被分派參與英語文素養評量者，七年級計有 2,793 人，升至八年級計有 2,893 人。

在排除全程未到考或轉出學生（即權重為 0 者）後，相關人數背景分布如表 4-15 所示，其中，在七年級階段，女生計有 1,309 人（47.9%）、男生計有 1,423 人（52.1%），而在升至八年級時，隨著納入轉入學生，性別比率雖然略有變化，但屬於極小程度變動；此外，七年級有接受補救教學者，計有 462 人（16.9%），而升至八年級，再次進行問卷調查後，整體有接受補救教學者，計有 473 人（17.3%）。

表 4-15 實際參與英語文評量（或學科問卷）學生背景分布

特徵	水平	七年級		八年級	
		人數	占總和比率 (%)	人數	占總和比率 (%)
性別	女生	1,309	47.9	1,301	47.6
	男生	1,423	52.1	1,434	52.4
有無接受 補救教學	無	2,266	82.9	2,243	82.0
	有	462	16.9	473	17.3
	遺漏	4	0.1	19	0.7
總和		2,732	100	2,735	100

## 五、參數估計與次級資料分析

有關學生能力表現估計，一是涉及認知層面，為英語文理解表現可能值估計，另一涉及學習動機與策略使用，為透過加權概似估計值（weighted likelihood estimation, WLE）估計。此外，進一步說明相關測驗、統計分析軟體使用。分述如下。

### （一）可能值估計

有關七、八年級學生橫斷面英語文理解表現能力的估計，本書使用可能值（plausible value, PV），而相關估計概念主要源自於 Beaton（1987）、Mislevy 等人（1992）使用於 NAEP 估計群體學生能力表現（group-level ability estimate）。隨著近年估計模式、參數估計方法演變，可能值實際估計概念也經過些許調整，本書主要在多向度隨機係數多項洛基模式（multidimensional random coefficients multinomial logit model, MRCMLM）基礎下（Adams & Wu, 2007; Adams et al., 1997），透過下列程序，搭配 R 平台中 Test Analysis Modules（TAM）套件進行參數估計（Robitzsch et al., 2020），以獲得學生表現可能值。

1. 在取得學生英語文表現資料後，本書以設定接受九年一貫課綱七年級學生整體平均能力值為 0 下，自由估計試題參數。在實際估計時，為求參數穩定與未來估計需求（係非本書現行目標）考量，本書同時結合 2019 年最後

一屆接受九年一貫課綱與 2020 年第一屆接受 108 課綱七年級學生英語文作答資料，透過多群組模式進行試題參數估計。

2. 建立條件變項 (conditioned variables)。本書將原始學生問卷中，有關「性別、特殊生、原住民、新移民子女、地區類型、題本、學校代碼」等，經虛擬編碼 (dummy coding) 或效應編碼 (effect coding) 設定，視為主要條件變項，同時，將學生問卷中其餘題項，逐題經虛擬編碼並透過主成分分析，擷取能解釋 85% 變異量之主成分分數，作為次要條件變項。
3. 由固定步驟 1 取得試題參數後，本書再結合步驟 2 之主要與次要條件變項，透過潛在迴歸分析模式 (Adams & Wu, 2007; Adams et al., 1997)，進行學生可能值估計，並從後驗分配中，隨機抽取 10 個可能值。
4. 為建立標準化量尺分數 (standardized scale score)，以進行分數詮釋與說明，本書透過公式 5，將七年級學生可能值 logit 分數，轉換為平均數 500、標準差 100 之標準化量尺分數，其中， $S_{ij}$  代表第  $i$  個學生在第  $j$  可能值所轉換量尺分數，而  $\theta_{ij}$  為該位學生可能值， $\bar{\theta}_j$ 、 $SD_{\theta_j}$  分別為所抽取第  $j$  個可能值全部學生加權平均值與標準差。

$$S_{ij} = 500 + 100 \times \frac{\theta_{ij} - \bar{\theta}_j}{SD_{\theta_j}} \quad (5)$$

5. 透過上述步驟 1 至 4，估計出七年級學生可能值，而同樣仿照上述步驟，估計而得八年級學生可能值，差別只在透過前述步驟 2 所直接獲得八年級相關垂直定錨試題之試題參數，同時須多納入七年級學生貝氏期望後驗能力值 (Bayesian expected a posterior, EAP) 作為主要條件變項 (Bock & Mislevy, 1982)，經固定試題參數，以潛在迴歸分析模式進行，擷取八年級學生 10 個可能值與進行量尺轉換。
6. 由於本書涉及不同語言活動，包含閱讀與聽力理解，故，學生在這兩個向度的可能值估計，相關程序與前述大致雷同，差別僅在於其使用組間多向度估計模式 (between-item multidimensionality) (Adams & Wu, 2007)，界定單群體兩個向度，進行相關試題參數、能力估計與量尺轉換。

## (二) 加權概似估計值估計

有別於前述學生可能值，主要用以表徵群體能力表現，而本書對於學生英語文學習動機、策略使用等，是以能力點估計表徵（ability point estimate），使用加權概似估計值（Warm, 1989）。相關能力的計算同樣是基於前述概化之多向度隨機係數多項洛基模式基礎上（Adams & Wu, 2007; Adams et al., 1997），在界定為多向度部分計分模式（multidimensional partial credit model）（Masters & Wright, 1997; Yao & Schwarz, 2006），透過學生於相關問卷作答反應，進行反向計分後，將 1、2、3、4，重新編碼為 0、1、2、3，再經加權後，進行學生 WLE 能力值估計。在建立標準化量尺分數時，本書透過公式 6，逐一將七年級學生於學習動機內三個構念、與策略使用內四個構念，逐一轉換為平均數 10、標準差 2 之量尺分數，其中， $S_{ij}$  代表第  $i$  個學生在第  $k$  個構念所轉換量尺分數，而  $\theta_{ik}$  為該  $i$  位學生在第  $k$  個構念 WLE 能力值， $\bar{\theta}_k$ 、 $SD_{\theta_k}$  分別為第  $k$  個構念內全部學生加權平均值與標準差。另一方面，關於八年級學生動機與策略使用 WLE 能力值估計，有鑑於八年級學生僅有調查部分構念與使用部分題項，因此，本書重新界定七年級向度與固定相關使用試題參數，來進行八年級學生 WLE 能力值估計與量尺轉換。

$$S_{ik} = 10 + 2 \times \frac{\theta_{ik} - \bar{\theta}_k}{SD_{\theta_k}} \quad (6)$$

## (三) 軟體使用與次級資料分析

有關本書各項參數估計與統計分析，由於涉及大型評量能力值估計與次級資料分析技術，目前尚無簡便且適合 TASAL 資料型態之慣用軟體。因此，本書透過 R 平台內適合套件，經撰寫相關程式語法進行分析，簡要說明如下。

### 1. TAM 套件

有關各項學生能力值與試題參數估計，本書透過前述步驟，撰寫相應之語法與搭配 TAM 套件（Robitzsch et al., 2020），進行學生 PV、WLE、EAP 能力值與相關試題難度參數、Thurstonian 閾值、參數適配度、信度指標等估計。

有關表徵試題難度的方式，大致有兩個面向。一是如公式 7 所示，為平均試題難度  $\bar{\delta}$ ，而其估計是在前述多向度隨機係數多項洛基模式下所得分數點之平均難度參數（Adams & Wu, 2007; Adams et al., 1997）。其中，學生反應資料若為多元計分（如 0、1、2 分），相關概念如公式 7 與圖 4-5 所示。 $m$  代表該試題最高分數點（如 2 分），即會有  $m + 1$  個反應類別（如 0、1、2 類別），而整體平均試題難度  $\bar{\delta}$  為 1 至  $m$  個分數點試題難度參數（如  $\delta_1$ 、 $\delta_2$ ）之平均值，分別代表反應類別如 0 與 1 間、1 與 2 間之試題難度參數。此外，如公式 8 所示， $\tau_m$  代表著整體平均試題難度參數與其他分數點試題難度間的距離，稱為試題階梯參數（step parameter）。

$$\bar{\delta} = \sum_{p=1}^m \frac{\delta_p}{m} \quad (7)$$

$$\tau_m = \bar{\delta} - \delta_p \quad (8)$$

第二種表徵試題難度方式為 Thurstonian 閾值，相關概念如圖 4-6 所示。每個曲線反映出學生若要達到某個分數點所須累積機率（cumulative probability）。一般界定若 0.5 機率下，學生要獲得各分數點（如 1、2 分），其得分所需能力值大致與前述試題難度參數雷同，而本書主要依照 PISA 界定，設定在 0.62 機率下，以估計學生為獲得 1 分所需得分難度  $\gamma_1$ ，為獲得 2 分之得分難度  $\gamma_2$  等。後續，再透過前述公式 5 量尺轉換，來表徵學生在每個試題獲得各分數點所需能力值。

圖 4-5 部分計分下的試題特徵曲線 (3 個反應類別)

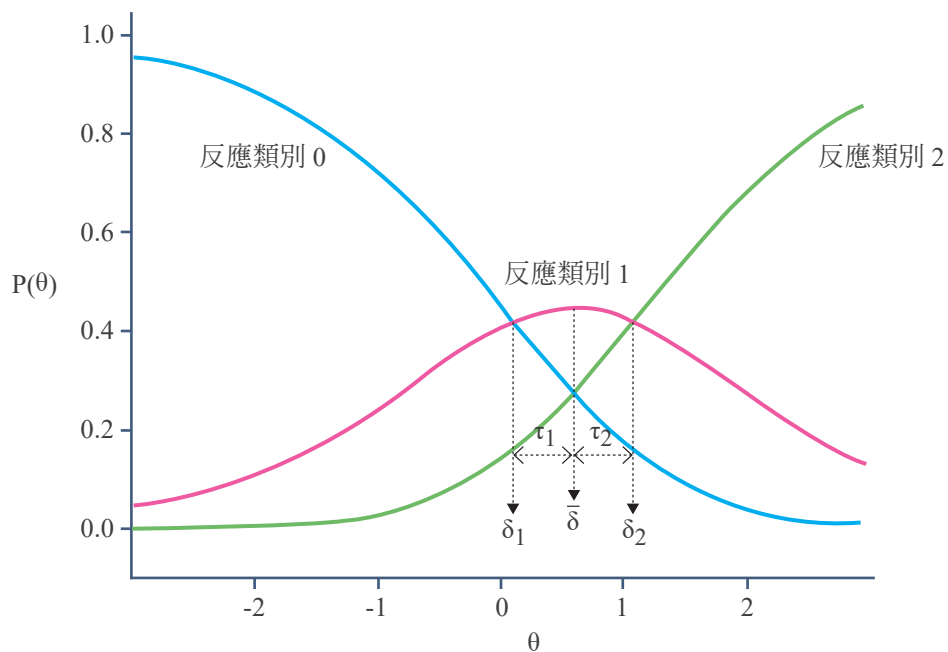
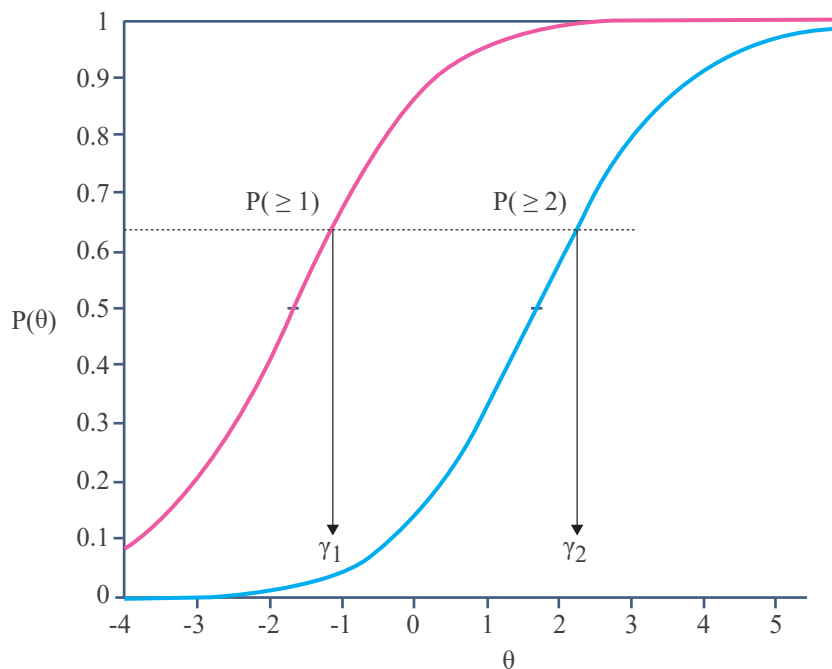


圖 4-6 部分計分下的 Thurstonian 閾值之累積機率曲線 (3 個反應類別)



有關參數適配度指標，相關概念建立在標準化殘差基礎，學者所提出未加權均方統計量（unweighted mean square）或稱 outfit（outlier sensitive mean square），與加權均方統計量（weighted mean square）或稱 infit（information-weighted fit mean square）（Adams & Wu, 2007; Wright & Masters, 1982）。其中，前者未加權均方統計量（outfit）對於極端值較敏感，後者加權均方統計量（infit）則透過訊息量來作加權來因應極端值影響。依據 Wright 與 Linacre（1994）的建議，當適配值介於 0.5 至 1.5 間時，代表著該試題仍能對測量結果產生貢獻，屬於可接受的範圍。有關於信度指標，就大型評量學生表現可能值估計而言，信度概念是以潛在能力值的變異數相對於群體學生觀察分數變異數來表徵（Adams, 2005），而本書以 EAP 信度來表徵英語文理解評量作答一致性，並透過 WLE 信度來表徵學生在學習動機與策略使用填答一致性。

## 2. survey 與 mitools 套件

有關本書所涉及相關大型評量資料次級分析，本書利用 survey 與 mitools 套件（Lumley, 2019, 2020），以因應複雜抽樣設計（complex sampling）與多元補插（multiple imputation）中各項統計分析之參數估計與顯著考驗。

有關大型評量抽樣誤差估計概念，相關執行方式有許多種類型，而本書採用 Jackknife 重複抽樣技術（Shao & Tu, 1995），以估計參數抽樣誤差（sampling error）。相關概念如表 4-16 所示。若就 TASAL 抽樣設計而言，其概念是將全部 245 個學校，將特徵類似的兩學校（或若不足數，擇一大校拆解）組成爲一個虛擬分層（pseudo-stratum）或稱 Jackknife zone（JKZONE）（LaRoche et al., 2020），如此，將產生出 123 個分層。其中，各分層內一個學校標示爲 0、另一個學校標示爲 1，稱爲 Jackknife replicate（JKREP）（LaRoche et al., 2020），而其功能是作爲運作重複抽樣選取權重標示之用。在實際執行重覆抽樣中，第一次計算重複抽樣權重（replicate weight）時，其概念是先針對分層 1 的學校，將標示爲 1 的學生權重設爲 2，另一間標示 0 的學生權重設爲 0，其餘非分層 1 的學校學生權重，同樣維持爲 1；同理，第二次計算重複抽樣權重時，同樣針對分層 1 內的學校，經交換權重，將標示 1 的學生權重設爲 0，另一間標示爲 0 的學生權重設

為 2，其餘同樣維持權重為 1。以同樣概念，不斷進行重複權重計算與設定，就 123 個分層而言，會產生 246 組重複抽樣權重。在搭配前述學生總權重基礎下，整體估計所得參數抽樣誤差如公式 9 所示， $t_0$  代表透過全部樣本計算所得參數估計值，而  $t_{si}$  代表使用每次重覆抽樣權重下，計算所得參數估計值，其中， $s$  為 1 至 123 個分層、 $i$  為各分層內標示數，而兩者所得參數估計值變異數（imputation variance）之開根號，就是抽樣誤差。實際執行過程中，本書透過 survey 套件（Lumley, 2020）來估計各項參數抽樣誤差。

$$se_{jrr}(t_0) = \sqrt{var_{jrr}(t_0)} = \frac{1}{2} \sum_{s=1}^{123} \sum_{i=1}^2 (t_{si} - t_0)^2 \quad (9)$$

表 4-16 透過 Jackknife 重覆抽樣以估計抽樣誤差示意

虛擬分層	學校重複抽樣標示	分層 1		分層 2		分層 3		...		分層 123	
		(1)	(2)	(3)	(4)	(5)	(6)	...	...	(245)	(246)
1	0	<u>2</u>	<u>0</u>	1	1	1	1	...	...	1	1
	1	<u>0</u>	<u>2</u>	1	1	1	1	...	...	1	1
2	0	1	1	<u>2</u>	<u>0</u>	1	1	...	...	1	1
	1	1	1	<u>0</u>	<u>2</u>	1	1	...	...	1	1
3	0	1	1	1	1	<u>2</u>	<u>0</u>	...	...	1	1
	1	1	1	1	1	<u>0</u>	<u>2</u>	...	...	1	1
...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...
123	0	1	1	1	1	1	1	...	...	<u>2</u>	<u>0</u>
	1	1	1	1	1	1	1	...	...	<u>0</u>	<u>2</u>

在前述組卷設計說明上，讀者可發現每位學生實際僅接受某一個題本，並透過題本間共同試題，來補插推估學生在作答全部試題時可能表現（Rubin, 1987），而其間對於學生能力估計技術會涉及學生可能值估計，使得每次學生理解表現參數估計也得考量可能值補插變異數或其標準誤（total standard error）所得測量誤差（measurement error），才可進行參數估計值顯著考驗（Mislevy et al., 1992）。

在使用 10 個可能值來表徵學生能力表現基礎下，本書計算整體參數估計值  $t_0$  的過程，其概念如公式 10 所示，而它是由 10 個可能值所得參數估計值  $t_m$ （ $m$  為 1 至 10 個可能值）之平均值來表徵。此外，源自於 10 個可能值補插所得參數估計值變異數，相關計算如公式 11 所示，概念係估計每個可能值所得估計參數值  $t_m$  與整體參數估計值  $t_0$  之變異數，並經過校正因子  $1 + m^{-1}$  來計算（Rubin, 1987, p. 78），取得整體參數估計值補插變異數。另一方面，在運用學生可能值進行參數估計脈絡下，本書對於各參數估計值抽樣誤差估計，相關概念雷同於前述公式 9，而其差異在於本書還得透過公式 13，分別考量 10 個可能值下，估計參數估計值反複抽樣變異數  $var_{jrr}(t_m)$ ，再經過公式 12 計算整體參數估計值平均抽樣變異數  $var_{jrr}(t_0)$ ，才能得到整體參數估計抽樣變異數。經結合補插變異數、與抽樣變異數後，各項參數估計值全部標準誤，就如公式 14 所示，是由整體變異數經取平方根所得。

$$t_0 = \frac{1}{10} \sum_{m=1}^{10} t_m \quad (10)$$

$$var_{impu}(t_0) = \frac{11}{10} \sum_{m=1}^{10} \frac{(t_m - t_0)^2}{10 - 1} \quad (11)$$

$$var_{jrr}(t_0) = \frac{1}{10} \sum_{m=1}^{10} var_{jrr}(t_m) \quad (12)$$

$$var_{jrr}(t_m) = \frac{1}{2} \sum_{s=1}^{123} \sum_{i=1}^2 (t_{msi} - t_m)^2 \quad (13)$$

$$se_{tot}(t_0) = \sqrt{var_{jrr}(t_0) + var_{impu}(t_0)} \quad (14)$$

在上述基礎下，針對參數估計值顯著考驗，本書透過 Wald 檢測與校正自由度來進行參數顯著考驗（Rubin, 1987）。實際執行過程中，多搭配使用 survey 與 mitools 套件（Lumley, 2019, 2020），進行各項統計分析之參數估計與顯著考驗。

### 第三節 支持效度之內部證據評估

在提供支持效度之內部證據上，本書透過分析各評量與調查工具內部一致性信度、檢核試題參數品質與模式適配度等數據來進行。以下茲就英語文理解與學習動機、策略使用等面向，簡要說明。

#### 一、英語文理解信度、試題適配度與平均難度

有關英語文理解評量，七年級學生總計正式施測 182 題、八年級學生正式施測 196 題，兩年級總計共正式施測 294 題。其中，學生整體作答 EAP 內部一致性信度值，結果如表 4-17 所示。七、八年級學生作答一致性信度值分別為 0.88、0.91，而其下閱讀理解 EAP 內部一致性信度分別為 0.87、0.89、聽力理解內部一致性信度分別為 0.87、0.90。

就試題參數與適配度而言，如附錄十二所示，七年級學生平均通過率為 0.51、古典平均鑑別度（即高分組通過率與低分組通過率差距）為 0.41，而八年級學生平均通過率為 0.47、古典平均鑑別度為 0.46。此外，由於新冠疫情爆發，影響 2021 年九年級正式評量，因此，本書（審核通過前）未獲得正式有效資料。其中，本書預計用於九年級正式評量之新命試題共計 118 題，預試結果如附錄十三所示，而相關試題平均通過率為 0.53、古典平均鑑別度為 0.51。整體而言，就古典試題參數而言，整體試題屬於難易適中，且具有不錯的高、低學生能力群的鑑別力。另一方面，就試題適配度而言，結果如附錄十二所示。七、八年級各試題加權均方適配度值 *infit* 整體分別介於 0.7210 至 1.4198、0.6122 至 1.5845 間，僅八年級所接受 R018\_3 (1.5845) 略微超過建議準則外，其餘試題皆符合 Linacre (2005)、Wright 與 Linacre (1994) 對於「認知試題評量」所建議介於 0.5 至 1.5 間可接受之模式適配度範圍。此外，就未加權均方值 *outfit* 而言，七、八年級各試題適配度整體分別介於 0.5865 至 1.8480、

0.5704 至 1.6793 間，經檢視，共有七題略微超出建議範圍，分別是七年級的 R066\_4 (1.5315)、L084\_2 (1.8480)、R082\_4 (1.5551) 與八年級的 L084\_2 (1.6339)、C002\_1 (1.5523)、L146\_2 (1.6793)、R123\_2 (1.6024)、L144\_1 (1.6791)，而其對應整體平均試題難度分別是 1.1577 (R066\_4)、1.2930 (L084\_2)、1.3883 (R082\_4)、1.8799 (C002\_1)、1.3353 (L146\_2)、1.7557 (R123\_2)、1.4258 (L144\_1)，使得本書認為可能是由於試題偏難，存在較多極端作答反應造成。此外，進一步檢視各試題的特徵曲線 (item characteristic curve)，發現施測七與八年級之全部二元計分試題，皆未出現非次序性 (disordinal) 現象，而在九個多元計分試題中，僅有 C002\_6 的 ICC 曲線出現此現象。然而，由於相關試題總題數偏少，同時多分散在不同區塊中，對於測量精準度之影響應實屬微小。最後，本書須強調由於在此書撰寫期間，已與謝進昌 (2022b) 針對切截分數有效性進行期刊投稿發表前相比，獲得更豐富資料。因此，對於英語文理解試題參數估計部分，本書係透過「多群組模式」同時合併最後一屆接受九年一貫課綱七年級與第一屆接受 108 課綱七年級學生作答反應資料來進行，因此，相關試題參數與適配值會微略不同於先前已發表之數據。然而，整體分析結果的方向多屬一致，皆指出相關試題具有可接受試題適配度。

表 4-17 本文評量與調查工具內部一致性信度

工具	構念	七年級	八年級
英語文理解	整體英語文理解	0.88	0.91
	閱讀理解	0.87	0.89
	聽力理解	0.87	0.90
英語文學習動機	內在動機	0.91	0.87
	自我概念	0.92	0.87
	社會動機	0.90	--
英語文理解策略 使用	記憶策略	0.82	0.77
	認知策略	0.85	0.82
	推論連結策略	0.89	0.79
	理解監控策略	0.91	0.84

註：英語文理解評量為 EAP 信度係數；學習動機、策略使用為 WLE 信度係數。

在平均試題難度方面，本書透過檢視不同語言活動、理解認知成分、文本類型來說明。就閱讀與聽力理解試題平均難度而言，結果如表 4-18 所示，七年級閱讀理解平均難度為 0.3043，而聽力理解為 -0.6386，略較閱讀試題簡易，然而，隨著學生評量的年級增長，相關試題難度安排也隨之增加。至八年級時，閱讀理解試題平均難度提升至 0.6116，而聽力理解試題平均難度為 -0.0033，大致符合評量高年級學生需求。

就理解認知成分試題平均難度而言，結果如表 4-19 所示。七年級涉及評量學生流暢運作語言基本知能試題，其平均難度為 -1.9204，而直接找到訊息試題平均難度為 -1.0785，是較前者為難。後續，本書在分別檢視理解字面意、隱含意、透過自身觀點進行評估與詮釋等，各層次試題平均難度會隨著成分水準上升而增加，自 -0.0323、0.7827 提升至 2.3700。若以同樣概念檢視評量八年級學生不同理解認知成分試題之平均難度分配，結果也大致呈現為理解成分愈高，相關平均試題難度會隨之增加現象，大致符合理論與邏輯性。

不同文本類型試題的平均難度，結果如表 4-20 所示。七年級第一類群篇章試題平均難度為 0.0269，而第二與第三類群篇章試題平均難度分別為 0.1876、0.2750。可見學生於作答第二、三類群篇章試題時較難，而相關原因可能在於這類篇章多涉及抽象概念與想法，同時作答時往往得連結自身學科知識與生活經驗，才得以回應題目，使得題目難度變高。以同樣概念檢視評量八年級學生不同篇章試題平均難度時，結果也展現出類似現象，同時，隨著評量學生年級增長，各類群篇章試題平均難度也隨之增加，使得試題組成大致符合評量不同年級學生能力需求。

表 4-18 閱讀與聽力理解試題平均試題難度

	閱讀理解	聽力理解	整體平均試題難度
七年級	0.3043	-0.6386	-0.1568
八年級	0.6116	-0.0033	0.3167

註：七年級視聽影片平均試題難度為 -0.2240、數位網頁平均試題難度為 0.7716；八年級視聽影片平均試題難度為 -0.0843、數位網頁平均試題難度為 0.7709；九年級尚未有正式評量結果。

表 4-19 不同理解認知成分之平均試題難度

年級	流暢運作 語言基本 知能	直接找 到訊息	理解字 面文意	理解隱 含文意	能透過 自身觀點 進行評估	能省思 認識自我	整體平 均試題 難度
七年級	-1.9204	-1.0785	-0.0323	0.7827	2.3700	--	-0.1568
八年級	--	-0.6879	-0.0055	0.7680	2.4101	2.7264	0.3167

註：九年級現行未有正式結果。

表 4-20 不同文本類型之平均試題難度

年 級	語言知 能文本	第一類群篇章			第二類群篇章		第三類 群篇章	整體平均 試題難度
		描述類	教學指 引類	人際互 動類	說明文	議論文	敘事與 文學	
七 年 級	-1.9204	0.0269			0.1876		0.2750	-0.1568
		0.0017	0.4636	-0.0394	-0.1091	1.2754		
八 年 級	--	0.1343			0.7284		0.6203	0.3167
		0.2277	0.3323	0.0604	0.4956	1.0610		

註：九年級現行未有正式結果。

整體而言，本書經檢視英語文理解評量信度、試題適配度與平均難度，結果多顯示相關評量工具擁有不錯作答內部一致性、可接受模式適配度，且相關試題平均難度會隨著評量學生的年級增長、理解認知成分提升、與篇章類型而具區別性，而此結果大多是符合過去理論與邏輯性。

## 二、英語文學習動機信度與適配度

有關學生英語文學習動機問卷填答一致性，相關 WLE 信度分析結果如表 4-17 所示。七與八年級學生內在動機作答一致性信度分別為 0.91、0.87，而學科自我概念作答一致性信度分別為 0.92、0.87。其中，由於社會動機現行僅調查七年級學生，其作答一致性信度為 0.90。整體而言，相關結果皆具有 0.8 以上的程度。

就模式適配度評估而言，本書就七年級學生作答資料，以三因素結構進行驗證性因素分析。學習動機模式適配度分析結果分別是  $\chi^2(186) = 1431.276$  ( $p < .001$ )、CFI = .952、TLI = .946、RMSEA = .050 與 SRMR = .032，其中，除了卡方值因參與者人數過多達到顯著外，其餘結果顯示不論是絕對適配指標、抑或是相對適配指標，皆符合 Hu 與 Bentler (1999) 及 McDonald 與 Ho (2002) 等人對於模式適配指標的建議。此外，本書若就試題 outfit 與 infit 指標進行檢視，結果如附錄十所示，三個構念內各題項加權均方適配度值 infit 整體介於 0.6649 至 2.1612，而 outfit 整體介於 0.5930 至 2.3117 間，其中，以 P7 的 2.1612 (infit)、2.3117 (outfit) 與 P4 的 0.5930 (outfit) 超出學者所建議準則，若暫時排除，則其餘各題項皆符合 Wright 與 Linacre (1994) 針對「評定量表」所建議之 0.6 至 1.4 適配準則範圍。

整體而言，本書經檢視英語文學習動機問卷量表信度、驗證性因素分析、試題適配度後，結果多顯示相關調查工具擁有不錯作答內部一致性、可接受試題適配度與符合原始界定的三個因素的模式架構。

## 三、英語文理解策略使用信度與適配度

有關學生英語文理解策略使用問卷填答一致性，相關 WLE 信度分析結果如表 4-17 所示。七與八年級學生記憶策略使用作答一致性信度分別為 0.82、0.77，認知策略信度分別為 0.85、0.82，推論連結策略分別為 0.89、0.79，理解監控策略分別為 0.91、0.84。其中，如表 4-13 所示，由於用以評量八年級學生策略使用之間卷題項數量，與評量七年級學生時相較之下一致為少，因此，學生作答內部一致性信度會連帶受到些微影響而下降。

就模式適配度評估而言，本書就七年級學生作答資料，以四因素結構進行驗證性因素分析。策略使用模式適配度分析結果分別是絕對適配指標的  $\chi^2(399) = 1720.837$  ( $p < .001$ )、RMSEA = .035、SRMR = .029，相對適配指標的 CFI = .958、

TLI = .954 (謝進昌, 2021b)。就結果而言, 除了卡方值容易因參與者人數過多而顯著外, 其餘指標皆符合 Hu 與 Bentler (1999) 及 McDonald 與 Ho (2002) 等人對於模式適配指標的建議。另一方面, 本書透過試題 *infit* 與 *outfit* 指標進行試題與模式適配度檢視, 結果如附錄十一所示。四個構念內各題項加權均方適配度值 *infit* 整體介於 0.7723 至 1.2793 間, 而未加權均方適配度值 *outfit* 整體介於 0.7242 至 1.2945 間, 皆符合 Wright 與 Linacre (1994) 針對「評定量表」所建議 0.6 至 1.4 適配準則範圍內。

整體而言, 本書經檢視英語文理解策略使用問卷量表信度、驗證性因素分析、試題適配度, 結果多顯示相關調查工具擁有不錯作答內部一致性, 各構念多具有近 0.80 或以上程度、可接受試題適配度與符合原始界定的四個因素的模式架構。

#### 第四節 支持效度之外部證據評估

有關支持效度之外部證據, 本書透過「與相關構念之效標關聯及各構念間聚斂、區別關聯」來說明, 可分為兩個層面。一是檢視各年級學生英語文理解與其學習動機、策略使用間關聯符合預期程度; 二是檢視不同背景學生英語文理解、學習動機與策略使用表現差異符合預期程度。茲就分析結果, 分述如下。

##### 一、各年級學生英語文理解、學習動機與策略使用關聯

在回顧過往文獻與實徵研究, 本書發現有關學生英語文理解表現與其學習動機、策略使用之關聯主題, 早已經被學者探究許久。就探討英語文為外語學習者動機而言, Al-Hoorie (2017) 從歷史發展角度分析, 指出其已經有超過 60 年相關動機研究歷史, 而近年回顧研究也指出學習者動機與其英語文表現存在正向關聯 (Vonkova et al., 2021), 同時, 相關結果亦多能獲得相關後設分析研究支持 (Masgoret & Gardner, 2003); 此外, 有關理解策略使用與學生英語文理解表現, 兩者可謂一體兩面, 學生自陳策略使用頻率, 往往與其於客觀英語文理解表現存在正向關聯性, 而此點也大致獲得相關後設分析研究支持, 顯示兩者約具中等關聯性 (Park, 2010)。

有關七與八年級學生英語文理解與其動機、策略使用關聯, 本書首先將學生「英語文理解表現」視為目標構念, 相關分析結果如表 4-21 所示。若就矩陣下三角七年

級學生分析，結果顯示各次構念與學生英語文理解之 Pearson 積差相關是自最小 0.30（與社會動機關聯）至最大 0.51（與自我概念關聯）之間不等；而矩陣上三角八年級結果，則是自最小 0.28（與記憶策略使用關聯）至最大 0.51（與自我概念關聯）間不等。整體而言，學生學習動機、策略使用與學生自身英語文理解表現間，具有 Cohen（1988）所界定小至大程度效果量程度關聯性，而此也大致符合過去相關實徵研究結果。另一方面，有關學生學習動機與策略使用間之次構念關聯，本書若將彼此視為目標構念，相關結果也多顯示出兩兩跨主構念間之次構念具有 Cohen 建議之中等至大程度關聯，其中，在不分年級資料來源下，相關數據之相關係數大多介於 0.45（自我概念與記憶策略）至 0.66（社會動機與理解監控策略）。

若就主構念內各次構念關聯而言，本書分別檢視學習動機內三個次構念、策略使用內四個次構念關聯，在不分年級資料來源下，其結果顯示學習動機內三個次構念關聯，其相關係數介於 0.61（自我概念與社會動機）至 0.71（內在動機與社會動機）間；此外，就策略使用內四個次構念內關聯，在不分年級資料來源下，其相關係數介於 0.63（記憶與推論連結策略）至 0.81（推論連結與理解監控策略）間，大致介於 Cohen（1988）建議之大程度效果量。

整體而言，若將學生英語文理解表現視為目標構念，分別以檢視其學習動機與策略使用調查量表作答，彼此間有小至大等效果量關聯性，具一定程度效標關聯效度；此外，若就聚斂與區別效度角度而言，分別檢視量表主構念內次構念與跨主構念間次構念之關聯，結果顯示主構念內各次構念關聯程度，是相對大於跨主構念間各次構念之相關係數，大致具備一定聚斂與區別關聯效度。相關結果符合理論與邏輯原則，顯示本書所發展認知評量與問卷調查工具效度，是具有基本外部證據支持。

表 4-21 七與八年級學生英語文理解可能值與學習動機、策略使用 WLE 能力相關係數

	英語文理解	內在動機	自我概念	社會動機	記憶策略	認知策略	推論連結策略	理解監控策略
英語文理解	1	0.36 (0.020)	0.51 (0.018)	--	0.28 (0.020)	0.32 (0.019)	0.45 (0.019)	0.39 (0.018)
內在動機	0.35 (0.021)	1	0.69 (0.014)	--	0.51 (0.021)	0.55 (0.019)	0.51 (0.021)	0.61 (0.018)
自我概念	0.51 (0.020)	0.65 (0.014)	1	--	0.45 (0.023)	0.49 (0.018)	0.51 (0.020)	0.59 (0.015)
社會動機	0.30 (0.023)	0.71 (0.014)	0.61 (0.016)	1	--	--	--	--
記憶策略	0.36 (0.020)	0.50 (0.020)	0.47 (0.021)	0.57 (0.019)	1	0.75 (0.014)	0.63 (0.017)	0.67 (0.015)
認知策略	0.38 (0.021)	0.57 (0.017)	0.49 (0.020)	0.63 (0.016)	0.68 (0.018)	1	0.69 (0.016)	0.73 (0.015)
推論連結策略	0.47 (0.018)	0.53 (0.019)	0.52 (0.019)	0.60 (0.016)	0.68 (0.017)	0.78 (0.012)	1	0.81 (0.010)
理解監控策略	0.43 (0.020)	0.62 (0.015)	0.55 (0.019)	0.66 (0.016)	0.69 (0.014)	0.79 (0.010)	0.80 (0.010)	1

註：1. 相關係數矩陣之下三角為七年級學生資料分析結果，而上三角為（七升）八年級學生資料分析結果。

2. 八年級學生未調查社會動機。

3. ( ) 內數值為標準誤。同時，表中所有數值經顯著考驗，皆達  $p < .01$  顯著水準，故，不特別標示。

## 二、不同背景學生英語文理解、學習動機與策略使用表現差異

在回顧過往英語文文獻與實徵研究，「性別」是較常為學者所探究的背景變項。性別落差（gender gap）議題更常見於實徵研究中（Henry, 2011; Lietz, 2006; Reilly et al., 2019），而相關結果多指出女生在相關英語文表現或相關動機、策略使用等結果，

相較於男生來得高，例如，Dörnyei 與 Csizér（2002）所做大型長期追蹤研究中顯示，在外語學習上女生較男生有顯著更高的動機，而此點也在相關文獻回顧研究中獲得支持（Vonkova et al., 2021）。此外，在學習策略使用上，女生也較男生更常使用廣泛語言學習策略（Oxford et al., 1988），進而使得女生展現出較佳的英語文理解表現（van der Slik et al., 2015）。除了性別外，本書同步納入學生年級、是否有接受補救教學，來評估不同背景特徵之於英語文理解評量與學習動機、策略使用調查，以檢核工具是否存在預期的區別效果。簡要說明如下。

### （一）年級與英語文理解、學習動機與策略使用

就跨年級英語文理解表現描述統計而言，相關結果如表 4-22 所示。本書在設定七年級學生整體英語文理解表現為平均 500 分基礎下，同群學生升至八年級時，其理解平均表現變為 520.57，若相對於標準誤計算所得 95% 信賴區間，兩年度未出現重疊區間，顯示八年級學生英語文理解整體平均表現，是顯著較其七年級時平均表現為高。此外，本書分別檢視閱讀、聽力理解等次構念表現時，結果也可觀察出八年級學生之平均表現較其於七年級時顯著地高。

就學生學習動機與策略使用而言，分析結果如表 4-23 所示。首先，就學習動機而言，在設定七年級學生填答平均值為 10 之單位基準下，同群學生升至八年級時，其自陳的內在動機與自我概念等次構念平均表現分別為 9.87、9.86，若對照標準誤所形成的 95% 信賴區間，兩年度出現重疊區，顯示出其與七年級時並無顯著差異存在。此外，就策略使用表現而言，八年級學生在四項策略使用平均表現，分別為 10.16、10.27、10.61、10.31，若對照標準誤所形成 95% 信賴區間，結果顯示除了「記憶策略」出現微略重疊區外，其餘策略於前年度平均表現皆未出現重疊區域，顯示出八年級學生於相關策略表現是顯著高於其七年級時表現。

整體而言，就學生自然成長發展而言，大多認同隨著年級增長，學生表現也會隨之進步。就本書分析結果，除了學習動機可能因學生年級增長而逐漸面臨較大壓力與學習負荷量，未出現顯著差異外，其餘結果都顯示學生在英語文理解認知表現、策略使用等面向，皆會隨著年級增加，出現顯著增長情況。此結果大致符合理論預期，也部分提供了支持工具有效性之外部證據。

表 4-22 七、八年級學生英語文理解表現可能值描述

變項	107 學年七年級			108 學年（七升至）八年級		
	平均值	標準誤	95% CI	平均值	標準誤	95% CI
整體						
英語文理解	500	2.81	[494.49, 505.51]	520.57	3.24	[514.22, 526.92]
語言活動						
閱讀理解	500	3.11	[493.90, 506.10]	516.78	3.57	[509.78, 523.78]
聽力理解	500	2.82	[494.47, 505.53]	517.37	3.03	[511.43, 523.31]

表 4-23 七、八年級學生學習動機與策略使用描述

變項	107 學年七年級			108 學年（七升至）八年級		
	平均值	標準誤	95% CI	平均值	標準誤	95% CI
學習動機						
內在動機	10	0.06	[9.88, 10.12]	9.87	0.06	[9.75, 9.98]
自我概念	10	0.05	[9.91, 10.09]	9.86	0.05	[9.76, 9.95]
社會動機	10	0.06	[9.89, 10.11]	--	--	--
策略使用						
記憶策略	10	0.05	[9.90, 10.10]	10.16	0.05	[10.07, 10.25]
認知策略	10	0.05	[9.90, 10.10]	10.27	0.05	[10.17, 10.37]
推論連結策略	10	0.05	[9.90, 10.10]	10.61	0.04	[10.53, 10.69]
理解監控策略	10	0.05	[9.90, 10.10]	10.31	0.05	[10.22, 10.40]

## (二) 性別、有無接受補救教學與英語文理解

有關不同性別、有無接受補救教學之七與八年級學生於橫斷面英語文理解表現差異分析，本書透過單因子迴歸分析，經計算抽樣誤差與補插誤差後，進行參數估計。分析結果分別如表 4-24、表 4-25 所示。就七年級學生分析而言，結果顯示女生英語文理解平均表現為 509.25，與男生的平均數差異為 -17.99 分（標準誤 = 4.42），經顯著性考驗， $t$  值為 -4.07 ( $p < .001$ )，可見男生英語文理解平均表現顯著地較女生為低；此外，本書以此概念來檢視學生其他背景變項的分析結果，也可發現在過去一年，有接受補救教學者的平均表現是顯著較無接受補救教學者來得低 ( $t = -12.70, p < .001$ )。

就八年級學生分析而言，結果如表 4-25 所示。結果顯示女生英語文理解平均表現為 532.46，與男生的平均數差異為 -23.01（標準誤 = 4.85），經顯著性考驗， $t$  值為 -4.75 ( $p < .001$ )，顯示男生整體英語文理解平均表現是顯著較女生為低。此外，就過去一年，有無接受補救教學而言，結果顯示有接受補救教學者是顯著較無接受補救教學者平均表現來得低 ( $t = -13.39, p < .001$ )。

表 4-24 不同背景七年級學生英語文理解表現可能值單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距 (女生)	509.25	3.53	[502.32, 516.18]		
斜率	-17.99	4.42	[-26.70, -9.29]	-4.07	< .001
補救教學					
截距 (無接受)	510.54	3.08	[504.50, 516.59]		
斜率	-73.29	5.77	[-84.61, -61.97]	-12.70	< .001

註：B = 未標準化迴歸係數值；SE = 迴歸係數標準誤；95% CI = 95% 信賴區間； $t$  = 斜率之顯著考驗統計量  $t$  值； $p$  =  $t$  統計量之顯著考驗機率值。

表 4-25 不同背景（七升至）八年級學生英語文理解表現可能值單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	532.46	4.05	[524.52, 540.41]		
斜率	-23.01	4.85	[-32.52, -13.50]	-4.75	< .001
補救教學					
截距（無接受）	531.96	3.50	[525.09, 538.83]		
斜率	-75.88	5.67	[-87.03, -64.73]	-13.39	< .001

註：符號意涵同表 4-24。

除檢視整體英語文素養表現外，本書進一步檢視不同背景學生於不同語言活動（次構念）之理解表現差異。就閱讀理解而言，七、八年級學生分析結果分別如表 4-26、4-27 所示。就七年級分析而言，結果顯示七年級男生閱讀理解的平均表現，顯著地較女生為低（ $t = -4.05, p < .001$ ），同時，在過去一年，有接受補救教學者閱讀理解平均表現是顯著較無接受補救教學者來得低（ $t = -11.49, p < .001$ ）；此外，就八年級分析而言，結果顯示八年級男學生閱讀理解平均表現，顯著低於女生（ $t = -4.66, p < .001$ ），而有接受補救教學者閱讀理解平均表現則顯著低於未接受補救教學者（ $t = -13.28, p < .001$ ）。

表 4-26 不同背景七年級學生英語文閱讀可能值單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	508.68	3.50	[501.83, 515.54]		
斜率	-16.89	4.17	[-25.07, -8.72]	-4.05	< .001
補救教學					
截距（無接受）	510.38	3.33	[503.86, 516.90]		
斜率	-72.09	6.27	[-84.40, -59.78]	-11.49	< .001

註：符號意涵同表 4-24。

表 4-27 不同背景（七升至）八年級學生英語文閱讀可能值單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	528.48	4.18	[520.28, 536.68]		
斜率	-22.63	4.85	[-32.14, -13.12]	-4.66	< .001
補救教學					
截距（無接受）	528.86	3.86	[521.28, 536.44]		
斜率	-80.55	6.07	[-92.54, -68.56]	-13.28	< .001

註：符號意涵同表 4-24。

就聽力理解而言，七、八年級學生分析結果分別如表 4-28、4-29 所示。就七年級分析而言，結果顯示七年級男生聽力理解平均表現，顯著地較女生為低（ $t = -4.37, p < .001$ ），同時，在過去一年，有接受補救教學者聽力理解平均表現是顯著較無接受補救教學者來得低（ $t = -11.86, p < .001$ ）；此外，就八年級分析而言，結果顯示八年級男學生聽力理解平均表現，顯著低於女生（ $t = -5.00, p < .001$ ），而有接受補救教學者，其聽力理解平均表現則顯著低於未接受補救教學者（ $t = -11.90, p < .001$ ）。

表 4-28 不同背景七年級學生英語文聽力可能值單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	509.92	3.76	[502.54, 517.30]		
斜率	-19.30	4.42	[-27.97, -10.63]	-4.37	< .001
補救教學					
截距（無接受）	510.12	3.15	[503.96, 516.29]		
斜率	-70.25	5.93	[-81.87, -58.63]	-11.86	< .001

註：符號意涵同表 4-24。

表 4-29 不同背景（七升至）八年級學生英語文聽力可能值單因子迴歸分析

變項	B	SE	95%CI	t	p
性別					
截距（女生）	528.81	3.82	[521.32, 536.31]		
斜率	-22.13	4.43	[-30.81, -13.45]	-5.00	< .001
補救教學					
截距（無接受）	527.33	3.16	[521.13, 533.53]		
斜率	-66.13	5.56	[-77.05, -55.21]	-11.90	< .001

註：符號意涵同表 4-24。

整體而言，在檢視不同性別、有無接受補救教學等背景學生於英語文整體與次構念之閱讀、聽力理解表現時，相關結果大致符合過往文獻與實徵研究結果，顯示男生較女生具有相對較低平均表現，而有接受補救教學者之平均表現也相對較無接受補救教學者為低。相關結果可提供部分支持本書所發展工具是有效度之外部證據。

### （三）性別、有無接受補救教學與學習動機

在認知理解表現之外，本書進一步檢視學生背景與其英語文學習動機之差異分析。透過單因子迴歸分析，經估計抽樣誤差與補插誤差後，學習動機各個次構念分析結果分別如表 4-30、4-31 所示。

就七年級學生內在動機分析而言，表 4-30 分析結果顯示女生內在動機平均表現為 10.08，與男生的平均數差異為 -0.16 分（標準誤 = 0.09），經顯著性考驗，t 值為 -1.83（ $p = .07$ ），顯示男生與女生自陳的內在動機平均表現並無顯著差異；同時，本書以此概念來檢視學生其他背景特徵，發現在過去一年，有接受補救教學者內在動機平均表現顯著較無接受補救教學者來得低（ $t = -2.46, p = .015$ ）。

就八年級學生內在動機分析而言，結果如表 4-31 所示。結果顯示女生所自陳學習英語文的內在動機平均表現為 10.04，與男生的平均數差異為 -0.33 分（標準誤 = 0.09），經顯著性考驗，t 值為 -3.61（ $p = .001$ ），顯示男生自陳的

內在動機平均表現顯著低於女生；此外，在過去一年，有接受補救教學者的內在動機平均表現是顯著較無接受補救教學者來得低（ $t = -5.24, p < .001$ ）。

表 4-30 不同背景七年級學生內在動機單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	10.08	0.08	[9.93, 10.24]		
斜率	-0.16	0.09	[-0.32, 0.01]	-1.83	.07
補救教學					
截距（無接受）	10.04	0.07	[9.91, 10.17]		
斜率	-0.30	0.12	[-0.55, -0.06]	-2.46	.015

註：符號意涵同表 4-24。

表 4-31 不同背景（七升至）八年級學生內在動機單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	10.04	0.07	[9.90, 10.18]		
斜率	-0.33	0.09	[-0.51, -0.15]	-3.61	.001
補救教學					
截距（無接受）	9.95	0.06	[9.83, 10.07]		
斜率	-0.58	0.11	[-0.79, -0.36]	-5.24	< .001

註：符號意涵同表 4-24。

就七年級學生自我概念分析而言，表 4-32 結果顯示女生自陳的自我概念平均表現為 10.02，與男生的平均數差異為 -0.03 分（標準誤 = 0.09），經顯著性考驗， $t$  值為 -0.39（ $p = .70$ ），顯示不同性別學生間所自陳的自我概念並無顯著差異；此外，本書在檢視有、無接受補救教學時，結果顯示過去一年，有接受補救教學者所自陳的自我概念平均表現顯著較無接受補救教學者來得低（ $t = -5.80, p < .001$ ）。

就八年級學生自我概念分析而言，結果如表 4-33 所示。結果顯示女生所自陳學習英語文的自我概念平均表現為 9.90，與男生的平均數差異為 -0.08 分（標準誤 = 0.09），經顯著性考驗， $t$  值為 -0.86（ $p = .001$ ），顯示男生與女生自陳的自我概念平均表現並無顯著差異；此外，在過去一年，有接受補救教學者自陳的自我概念平均表現顯著較無接受補救教學者來得低（ $t = -5.43, p < .001$ ）。

表 4-32 不同背景七年級學生自我概念單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	10.02	0.06	[9.89, 10.14]		
斜率	-0.03	0.09	[-0.20, 0.13]	-0.39	.70
補救教學					
截距（無接受）	10.11	0.05	[10.01, 10.20]		
斜率	-0.74	0.13	[-0.98, -0.49]	-5.80	< .001

註：符號意涵同表 4-24。

表 4-33 不同背景（七升至）八年級學生自我概念單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	9.90	0.07	[9.77, 10.03]		
斜率	-0.08	0.09	[-0.25, 0.10]	-0.86	.389
補救教學					
截距（無接受）	9.96	0.05	[9.86, 10.06]		
斜率	-0.71	0.13	[-0.96, -0.45]	-5.43	< .001

註：符號意涵同表 4-24。

就學生自陳的社會動機而言，僅調查七年級學生，分析結果如表 4-34 所示。結果顯示女生自陳的社會動機平均表現為 10.14，與男生的平均數差異為 -0.27 分（標準誤 = 0.09），經顯著性考驗， $t$  值為 -3.07（ $p = .003$ ），顯示

男生所自陳透過英語文進行社會交流的動機顯著地較女生為低；此外，就有無接受補救教學而言，結果顯示在過去一年，不論有無接受補救教學，其二者間並無顯著差異（ $t = -1.90, p = .06$ ）。

表 4-34 不同背景七年級學生社會動機單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	10.14	0.07	[10.01, 10.28]		
斜率	-0.27	0.09	[-0.45, -0.10]	-3.07	.003
補救教學					
截距（無接受）	10.04	0.06	[9.92, 10.15]		
斜率	-0.25	0.13	[-0.50, 0.01]	-1.90	.06

註：符號意涵同表 4-24。

整體而言，在檢視不同性別、有無接受補救教學等背景學生於學習動機內各次構念之自陳表現時，若單就學生填答平均數而言，係符合過去實徵研究與理論邏輯，以男生較女生具有相對較低平均表現、有接受補救教學者平均表現也相對較無接受補救教學者為低。然而，若進一步考量背景顯著考驗分析，僅有部分年級於部分次構念結果，符合過往文獻與實徵研究結果。本書認為在考量相關調查工具屬於學生自陳問卷，測量準確性較容易受當下填答環境影響，使得本結果僅能大略提供部分支持工具效度的外部證據。

#### （四）性別、有無接受補救教學與理解策略使用

進一步檢視學生背景與其英語文理解策略使用之差異分析時，本書透過單因子迴歸分析，經估計抽樣誤差與補插誤差後，各次構念分析結果分別如表 4-35、4-36 所示。

就七年級學生「記憶策略」使用分析而言。表 4-35 結果顯示女生自陳的記憶策略使用平均表現為 10.30，與男生自陳的平均數差異為 -0.59 分（標準誤 = 0.08），經顯著性考驗， $t$  值為 -7.02（ $p < .001$ ），顯示男生自陳的記憶策略使用平均表現顯著地較女生為低；同時，就有、無接受補救教學經驗而言，

結果顯示在過去一年，有接受補救教學者相對無接受補救教學者是使用顯著較低之記憶策略，其  $t$  值為  $-2.98$  ( $p = .004$ )。

就八年級學生自陳的「記憶策略」使用差異分析而言，結果如表 4-36 所示。結果顯示女生所自陳記憶策略使用平均表現為 10.34，與男生自陳的平均數差異為  $-0.34$  分（標準誤 = 0.09），經顯著性考驗， $t$  值為  $-4.01$  ( $p < .001$ )，顯示男生自陳的記憶策略使用平均表現是顯著低於女生；此外，在過去一年，有接受補救教學者自陳的記憶策略使用平均表現是顯著較無接受補救教學者來得低 ( $t = -4.18, p < .001$ )。

表 4-35 不同背景七年級學生記憶策略使用單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距 (女生)	10.30	0.06	[10.19, 10.42]		
斜率	-0.59	0.08	[-0.75, -0.42]	-7.02	< .001
補救教學					
截距 (無接受)	10.05	0.05	[9.95, 10.15]		
斜率	-0.37	0.12	[-0.61, -0.13]	-2.98	.004

註：符號意涵同表 4-24。

表 4-36 不同背景（七升至）八年級學生記憶策略使用單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距 (女生)	10.34	0.06	[10.22, 10.46]		
斜率	-0.34	0.09	[-0.51, -0.18]	-4.01	< .001
補救教學					
截距 (無接受)	10.23	0.05	[10.14, 10.32]		
斜率	-0.46	0.11	[-0.67, -0.24]	-4.18	< .001

註：符號意涵同表 4-24。

就七年級學生「認知策略」使用分析而言，表 4-37 結果顯示女生自陳的認知策略使用平均表現為 10.25，與男生的平均數差異為 -0.50 分（標準誤 = 0.08），經顯著性考驗， $t$  值為 -6.19 ( $p < .001$ )，顯示男生自陳的認知策略使用平均表現顯著地較女生為低；此外，就有、無補救教學經驗而言，結果顯示在過去一年，有接受補救教學者是顯著相對較無接受補救教學者自陳使用較低的認知策略，其  $t$  值為 -3.56 ( $p < .001$ )。

就八年級學生「認知策略」使用差異分析而言，結果如表 4-38 所示。結果顯示女生自陳的認知策略使用平均表現為 10.50，與男生的平均數差異為 -0.45 分（標準誤 = 0.10），經顯著性考驗， $t$  值為 -4.76 ( $p < .001$ )，顯示男生自陳的認知策略使用平均表現顯著地較女生為低；此外，就有、無補救教學經驗而言，結果顯示在過去一年，有接受補救教學者是顯著相對較無接受補救教學者自陳使用較低的認知策略，其  $t$  值為 -5.25 ( $p < .001$ )。

表 4-37 不同背景七年級學生認知策略使用單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距 (女生)	10.25	0.06	[10.13, 10.38]		
斜率	-0.50	0.08	[-0.65, -0.34]	-6.19	< .001
補救教學					
截距 (無接受)	10.07	0.05	[9.96, 10.17]		
斜率	-0.46	0.13	[-0.71, -0.21]	-3.56	< .001

註：符號意涵同表 4-24。

表 4-38 不同背景（七升至）八年級學生認知策略使用單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	10.50	0.07	[10.37, 10.63]		
斜率	-0.45	0.10	[-0.64, -0.27]	-4.76	< .001
補救教學					
截距（無接受）	10.37	0.05	[10.26, 10.47]		
斜率	-0.66	0.13	[-0.91, -0.42]	-5.25	< .001

註：符號意涵同表 4-24。

就七年級學生「推論連結策略」使用分析方面，結果如表 4-39 所示。結果顯示七年級女生自陳的推論連結策略使用平均表現為 10.27，與男生的平均數差異為 -0.52 分（標準誤 = 0.08），經顯著性考驗，t 值為 -6.33（ $p < .001$ ），顯示男生自陳的推論連結策略使用平均表現是顯著較女生為低；此外，就有、無接受補救教學經驗而言，結果顯示在過去一年，有接受補救教學者與無接受補救教學者相對自陳使用較低推論連結策略，其 t 值為 -5.68（ $p < .001$ ）。

就八年級學生自陳「推論連結策略」使用差異分析而言，結果如表 4-40 所示。結果顯示八年級女生自陳的推論連結策略使用平均表現為 10.91，與男生的平均數差異為 -0.58 分（標準誤 = 0.08），經顯著性考驗，t 值為 -6.96（ $p < .001$ ），顯示男生自陳的推論連結策略使用平均表現是顯著較女生為低；此外，就有、無接受補救教學經驗而言，結果顯示在過去一年，有接受補救教學者與無接受補救教學者相對自陳使用較低推論連結策略，其 t 值為 -9.25（ $p < .001$ ）。

表 4-39 不同背景七年級學生推論連結策略使用單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距 (女生)	10.27	0.06	[10.14, 10.39]		
斜率	-0.52	0.08	[-0.68, -0.36]	-6.33	< .001
補救教學					
截距 (無接受)	10.10	0.06	[9.99, 10.21]		
斜率	-0.69	0.12	[-0.92, -0.45]	-5.68	< .001

註：符號意涵同表 4-24。

表 4-40 不同背景 (七升至) 八年級學生推論連結策略使用單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距 (女生)	10.91	0.05	[10.81, 11.00]		
斜率	-0.58	0.08	[-0.74, -0.41]	-6.96	< .001
補救教學					
截距 (無接受)	10.76	0.04	[10.67, 10.84]		
斜率	-1.00	0.11	[-1.21, -0.79]	-9.25	< .001

註：符號意涵同表 4-24。

在七年級學生自陳「理解監控」策略使用分析方面，結果如表 4-41 所示。結果顯示七年級女生自陳的理解監控策略使用平均表現為 10.24，與男生的平均數差異為 -0.47 分 (標準誤 = 0.08)，經顯著性考驗，t 值為 -5.54 ( $p < .001$ )，顯示男生自陳的理解監控策略使用平均表現顯著較女生為低；此外，就有、無接受補救教學經驗而言，結果顯示在過去一年，有接受補救教學者與無接受補救教學者相對自陳使用較低理解監控策略，其 t 值為 -4.07 ( $p < .001$ )。

就八年級學生自陳「理解監控」策略使用差異分析而言，如表 4-42 所示。結果顯示八年級女生自陳的理解監控策略使用平均表現為 10.50，與男生的平均

數差異為 -0.38 分（標準誤 = 0.08），經顯著性考驗，t 值為 -4.68（ $p < .001$ ），顯示男生自陳的理解監控策略使用平均表現是顯著較女生為低；此外，就有、無接受補救教學經驗而言，結果顯示在過去一年，有接受補救教學者與無接受補救教學者相對自陳使用較低理解監控策略，其 t 值為 -6.48（ $p < .001$ ）。

表 4-41 不同背景七年級學生理解監控策略使用單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	10.24	0.07	[10.10, 10.38]		
斜率	-0.47	0.08	[-0.63, -0.30]	-5.54	< .001
補救教學					
截距（無接受）	10.07	0.06	[9.96, 10.19]		
斜率	-0.52	0.13	[-0.76, -0.27]	-4.07	< .001

註：符號意涵同表 4-24。

表 4-42 不同背景（七升至）八年級學生理解監控策略使用單因子迴歸分析

變項	B	SE	95% CI	t	p
性別					
截距（女生）	10.50	0.06	[10.40, 10.61]		
斜率	-0.38	0.08	[-0.53, -0.22]	-4.68	< .001
補救教學					
截距（無接受）	10.42	0.05	[10.32, 10.51]		
斜率	-0.74	0.11	[-0.96, -0.51]	-6.48	< .001

註：符號意涵同表 4-24。

在檢視不同性別、有無接受補救教學等背景學生於自陳英語文理解策略使用的各個次構念表現時，相關結果大致符合過往文獻與實徵研究結果，顯示男生較女生自陳使用相對較低策略頻率，而有接受補救教學者也與無接受補救教學者相對自陳使用較低策略頻率。另一方面，如同本書第三章第二節末所強

調，理解認知與理解策略使用，屬於兩個類似但是具差異性概念，前者為學習者所展現的內在、無意識自動運作能力，後者屬於在特定目的下，學習者有意識所採取的行動，以協助其理解，而當某個策略習得久了，逐漸就會內化以自動使用，進而影響自身理解認知表現。因此，依據本節背景差異分析，不論是在英語文理解認知表現、抑或是策略使用，結果皆一致顯示出類似結論。整體而言，本節分析結果係提供支持相關工具有效度之外部證據。

## 第五章 結論與建議

### 第一節 結論

教育部自 108 學年度起，正式逐年推動十二年國民基本教育課程綱要，而在新、舊國家課程綱要轉換期間，為評估此轉換對於學生素養表現影響，國家教育研究院成立臺灣學生成就長期追蹤評量計畫（TASAL）。該計畫目的在描述與追蹤同一群具代表性臺灣學生素養表現及探究影響學生成長變化因子，以回饋現行或未來國家課程綱要內容調整（國家教育研究院，無日期）。在此脈絡下，本書目的在透過 2019、2020 年七、八年級學生參與英語文學生實徵表現數據，經搭配理論陳述、評量架構建立，來提供支持英語文素養評量與相關調查工具有效性之證據。以下茲就三個面向的效度證據，簡要綜整結論。

#### 一、支持效度的過程證據

誠如本書第四章第一節所述，「效度」是表徵工具發展品質與內涵評估方向，而能提供支持工具有效性證據的來源，十分多元與複雜，再加上評量工具的效度往往會隨著時間遞移而略有變動，很難在單一書章篇幅與單一時間點內完全釐清。據此，本書現行聚焦於提供「基本」支持工具效度之過程證據內容，包含有英語文理解評量架構說明、學習動機與策略使用理論陳述和專家建議、標準化評量與調查工具發展流程說明（例如題庫發展與範例題、組卷、施測與資料處理、抽樣及權重、參數估計和次級資料分析等程序）。

本書在說明這些過程內容時，相關陳述要點指引大致參考現行國際各大型評比所釋出學科評量架構、技術報告等，例如，PISA 閱讀、外語評量 FLA 評量架構（OECD, 2019, 2021）、IEA 推動 PIRLS、美國 NAEP 閱讀對於大型評量工具發展過程重點描述（Mullis & Martin, 2015; NAGB, 2019）等。然而，在有限人力且考量某些任務並非作者實質推動下（例如，TASAL 整體行政施測與講習、抽樣、資料匯整，故，於本書僅能略提概念），是賴群體之力推動，詳盡程度尚無法完全比擬前述所

釋出相關報告內容。然而，本書期望能以英語文素養評量發展為主軸，透過具體、透明與詳盡的說明，力求使讀者了解第四階段學生英語文素養評量資料形成過程，為後續政策決策者或大型評量次級資料使用者能在此資料基礎下，建立政策或評估相關可能決策影響等建立穩固基礎。

## 二、支持效度的內部證據

有關支持效度之內部證據，本書經檢核英語文理解評量與學習動機、策略使用調查工具內部一致性信度、試題與題項參數品質、模式適配度等後，顯示就學生作答內部一致性信度而言，不論是英語文理解評量工具、抑或是英語文學習動機、英語文學習動機問卷填答，結果皆顯示出學生作答具有近 0.80 或以上之作答的內部一致性程度，具不錯的信度水準。

就英語文理解試題品質而言，七年級學生正式施測 182 題、八年級學生正式施測 196 題，經排除垂直等化共同題，兩年級總計正式施測 294 題，再加上目前只有經過預試尚未正式施測九年級之新命題 118 題，總計共發展 412 題以評估第四學習階段學生英語文理解表現。在經過古典試題分析，整體試題屬於難易適中且具有不錯的區別高、低能力群之鑑別力。此外，就試題適配度分析而言，結果顯示除了 1 題加權均方值 *infit* 與 7 題未加權均方值 *outfit*，未達學者建議準則外，其餘各年級正式評量試題適配度皆能符合 Linacre (2005)、Wright 與 Linacre (1994) 針對「認知試題評量」所建議之介於 0.5 至 1.5 間可接受適配度準則範圍。最後，有關試題難度分布，結果顯示當學生年級愈高，相關高層次理解認知成分與具備相對較抽象內容之第二、第三類群文本等配置，也具有相對較高的平均試題難度，而此結果也大致符合過去理論與邏輯性。

就學生學習動機與策略使用題項品質而言，本書分別透過三因素與四因素驗證性因素分析，結果顯示模式與資料適配度大多能符合 Hu 與 Bentler (1999)、McDonald 與 Ho (2002) 等人對於絕對、相對適配指標的建議準則；此外，本書進一步透過試題適配度分析，結果顯示除了內在動機下 P7、P4 等兩個題項超出學者建議準則外，其餘各題項皆能符合 Wright 與 Linacre (1994) 針對「評定量表」所建議的介於 0.6 至 1.4 適配準則範圍。整體而言，英語文理解評量與相關學習動機、策略使用等調查工具擁有可接受作答內部一致性、試題品質和適配度，得以提供支持工具有效性的內部證據。

### 三、支持效度的外部證據

有關支持效度的外部證據，本書引用「與相關構念之效標關聯及各構念間聚斂、區別關聯」來說明。就學生英語文理解表現與其學習動機、策略使用間之效標關聯而言，整體結果顯示在不分年級資料來源下，兩兩構念間 Pearson 積差相關係數，自 0.28（八年級學生英語文理解與記憶策略使用）至 0.81（八年級學生推論理解策略與理解監控策略使用）不等，實徵分析結果顯示出兩兩構念間具有正向相關，而效果量強度大致橫跨 Cohen（1988）所建議小（0.1 起）、中（0.3 起）與大（0.5 起）程度之評估準則，整體而言，相關評量與調查工具間具有不錯效標關聯效度。另一方面，進一步檢視各主構念內的次構念、跨主構念間的次構念所顯示之聚斂與區別關聯，結果顯示各個次構念與英語文理解認知表現、或學習動機與策略使用各個次構念關聯，大致介於小至大程度效果量（不分年級資料，積差相關係數介於 0.28 至 0.66 間），此外，學習動機內的三項次構念、策略使用內四項次構念之兩兩關聯，大致為大程度效果量（不分年級資料，積差相關係數是介於 0.61 至 0.81 間），顯示出具有一定聚斂與區別關聯效度。

此外，就跨年級學生於英語文理解表現、學習動機與策略使用平均表現檢視，經顯著性考驗，結果顯示除了學習動機可能因年級增長而逐漸面臨較大壓力與學習負荷量，而未出現顯著差異之外，其餘含英語文理解認知表現、策略使用等面向，皆顯示出學習表現會隨著學生年級增加顯著增長。另外，本書透過單因子迴歸分析，經顯著考驗，檢視不同背景學生於英語文理解、學習動機與策略使用表現差異，結果顯示除了學習動機內有部分次構念與部分年級，並未隨學生背景差異出現顯著區別外，其餘皆一致顯示出男生在英語文理解、記憶、認知、推論連結、理解監控策略使用方面，皆明顯低於女生，同時，在過去一年內，有另接受補救教學者於相關構念的平均表現，同樣明顯低於未接受補救教學者，符合過去文獻與實徵研究結果。

整體而言，不論就效標關聯效度、抑或是檢視不同背景學生平均表現差異，皆切合過去文獻與實徵研究結論，大致符合理論、邏輯與實徵結果方向，能適時提供支持本書所使用工具有效性之外部證據。

## 第二節 建議

### 一、持續性長期推動效度證據累積與評估

誠如過去學者們所認為「效度並非全有或全無的狀態」，而是涉及長期與持續性證據累積的過程，其關注面向不僅是評量工具本身，更重視測驗分數詮釋與使用（Messick, 1989）。就 TASAL 英語文素養評量發展而言，相關實徵結果與資料使用目的之一為回饋現行 108 英語文課程綱要，而任何依據相關結果所作決策都可能對於學校、教師、甚至是教育系統產生（正面或負面）影響。因此，後效影響評估遂成為未來效度評估關注的重點，值得未來進一步累積發展。

評量標準建立與設定對於本書評量工具的發展，占有重要角色。對此，讀者可將國家課程綱要視為內容或課程標準，而本書評量工具擁有標準本位大型評量內涵，其發展涉及將「內容標準」轉化為「表現標準」與設定出具體「切截、通過分數」過程，如此，後續研究者才得以透過相關結果（如本書第二章第五節所提學生表現標準描述、各水準切截分數與相關實徵表現數據等）來詮釋臺灣學生英語文表現概況與進一步形成回饋英語文國家課程綱要的建議內容。因標準設定中切截分數有效性係關係著測驗分數詮釋或使用，相關效度評估自然成為重要環節，另外，由於該程序已有部分刊載於謝進昌（2021b, 2022b），本書就未有過多著墨，建議讀者相互搭配本書與相關文章內容，可得更完整了解現行效度評估結果。

### 二、透過學生實徵表現結果回饋英語文國家課程綱要

本書英語文素養評量資料與結果的運用，首要目的為回應英語文國家課程綱要內容，而就實際參與研究經驗與過程中所建立評量架構、學生表現標準描述等素材，建議未來可針對下列方向，搭配接受 108 課綱學生實徵數據分析結果，來進行回應與回饋。

#### （一）建立國家課程綱要內容關鍵詞彙解釋或解讀的建議方向

誠如本書所述，國家課程綱要所界定內容為教育目標，是一通用性、方向性的指引內容，而相關說明詞彙就容易落入讀者各自解讀的情況。若以學習表現「3-IV-8 能了解短文、簡訊、書信的主要內容」為例，何謂「了解」？若依

據學者對於閱讀或聽力「理解」(comprehension)解釋，其存在不同理解成分或層次，因此，若能搭配具體解釋或單純提供建議解讀方向，學校教師會較容易了解；此外，就學習內容「B-V-12 故事及短文的主旨或大意」為例，若依據馬行誼(2014, 頁 55)的說明，「大意」與「主旨」是存在內涵形成之差異性的。整體而言，本書建議可依據 Bloom 等人(1956)及其後續演變(Anderson et al., 2001)，就國家課程綱要內容中關鍵「動詞」、「名詞」詞彙提供解釋或解讀建議方向，以便教師能清楚掌握學習表現、學習內容的內涵。此外，本書從課程內容轉化為評量指引、甚至是具體試題的過程，也大致採用此建議，也許未來讀者在發展英語文素養評量架構與試題亦可參考使用。

## (二) 透過「理解成分」回饋英語文課綱中學習表現、邏輯思考內容

誠如本書附錄一、二所建立「學習表現；邏輯思考、判斷與創造力」與「理解認知成分」的對應內容，讀者可觀察到現行英語文國家課程綱要對於第四學習階段學生的表現，除了基本語言知識外，整體建議方向為期望學生能掌握篇章直接陳述明顯訊息與理解字面文意、隱含文意等，其中，對於較高理解成分，例如，詮釋與評估等，多半建議於第五學習階段(高中職)學習。然而，本書若單純就學生實徵表現結果來檢視，例如，附錄八範例試題 R018 問題 3、附錄九 R135 問題 4 等，可發現部分七或八年級學生已經具備「在文本理解上，能透過自身觀點進行評估與詮釋、甚至是省思」之能力。這些零星試題學生表現證據，也許不足以一窺群體表現全貌，有待更完整群體數據檢核。因此，本書建議未來可透過學生實徵表現數據與搭配標準設定結果，來檢視不同水準學生表現概況，以評估第四學習階段學生是否適合或單純補充建議納入較高層次理解認知成分的學習。

## (三) 透過「文本類型」回饋英語文課綱中學習內容

在一般生活或學術情境下，文本(text)或篇章(discourse)是學生學習英語文所直接接觸的材料，其中，若課程綱要內容能提供文本或篇章歸類建議方向，可有效導引教科書編製、教師掌握選材與評量要點，避免過度集中某類型文本素材情況。有別於國語文課程綱要內容，現行英語文課程綱要對於文本或篇章類型建議，大致是以故事、短文、書信、賀卡、簡訊、電子郵件、或體裁等形式、情境表徵，尚未有系統性地從「目的、功能」等面向提供文

本或篇章歸類建議。除了從理論與概念層面提供文本分類建議外，透過實徵學生表現數據也能提供不同視野角度。若就本書分析結果，可發現不同類型文本試題平均難度有著差異性，以第二、三類群文本具有相對第一類群文本平均較難試題，相關原因可能在於這類文本較容易連結抽象概念、想法的陳述與涉及學生背景知識、經驗多寡，使得相關評量試題也進而相對容易與高層次理解認知成分連結。整體而言，本書所使用的文本類型是從評量工具發展和語用角度出發，係未來可參考的一個面向，建議可搭配學生實徵表現結果與理論，經探究以具體提出文本或篇章分類建議。

#### （四）透過「語言溝通活動」回饋英語文課綱中學習內容

本書現行所涉及語言溝通活動，主要以閱讀、聽力等接收性語言活動理解，且可再細分納入視聽理解（如影片）、數位網頁等，然而，由於後兩者的試題比重並不高，僅涉及少數篇章試題，加諸，現行英語文課程綱要內容對此的著墨，也僅止於少數或幾乎無學習表現、學習內容特別進行說明（例如，1-IV-8 能聽懂簡易影片的主要內容、或涉及電子郵件相關指標），使得整體評量結果未能反映這些特別溝通活動表現概況。然而，本書從文內視聽影片、數位網頁試題分析結果發現，數位網頁文本平均試題難度高於閱讀、聽力試題，加上隨著數位科技興起，學生透過視聽影片、數位網頁進行學習、溝通的頻率提高，使其逐漸成為趨勢，是有進一步探究與發展需求性。整體而言，由於本書僅有少數涉及數位網頁試題學生表現證據，也許不足以一窺表現全貌，考量未來發展趨勢下，將來有待更完整評量設計，蒐集學生於數位網頁（或多文本）活動理解評量數據，以提供「未來下一波」英語文國家課程綱要修訂的建議。

### 三、透過大型評量次級資料分析提供教育政策發展建議

透過科學證據來建立導引教育決策的理念，逐漸受到國內外學者的推廣與認同，然而，何謂有效證據及其來源，始終仍是爭議的焦點。由於隨機分派實驗研究設計（randomized controlled trial, RCT）結果具有因果推論可行性，多能獲得學者認同，然而，受限於主題與邏輯可行性，隨機分派實驗研究設計往往難以廣泛運用於教育領域，使得如何透過其他方式或管道來建立有效且可接受科學證據，成為重要探究議題。就大型評量資料使用而言，透過觀察資料型態（observational data）來進行

Shadish 等人（2002）所謂因果關係的解釋（causal explanation）方式，逐漸受到學者運用，而其多半藉由統計分析技術，例如，傾向分數（propensity score）、工具變項取徑（instrumental variable approach）來因應仿照實驗研究中隨機分派效果，以達到兩組別均衡性（關秉寅、李敦義，2008；Schneider et al., 2007）。據此，本書對於學生背景特徵分析，目的在於建立效度證據，並非實質探究與解釋因素間因果關係，加諸，大型評量次級資料分析結果，往往具有回饋教育政策發展功能，未來若能搭配相關解釋因果關係統計技術使用，其結果將具有相當政策回應價值，值得未來進一步關注。

#### 四、銜接第五學習階段英語文評量架構與抽樣設計

本書所提英語文素養評量架構，相關主軸是由理解認知成分、文本類型（表徵內容知識）等雙主軸向度所建構，而這些向度具有跨學習階段通用性，甚至在架構發展之際，本書已經融入銜接下一學習階段命題考量。據此，未來進一步推動第五學習階段學生英語文素養評量資料收集時，值得後續延伸使用本評量架構，唯，研究者得特別關注與納入第五學習階段特有的學習表現內容或面向，例如，普通高中與技術型高中篇章情境與試題差異、較高層次理解認知成分之試題比重、涉及文本結構的理解、不同腔調／語言背景英語使用者談話內容理解試題發展等。

就抽樣設計方面，TASAL 第四學習階段評量使用長期追蹤方式，針對同群學生自七年級追蹤至九年級，其優點雖然提供探究影響學生素養成長表現相關因素之因果解釋可能，然而，在計畫推展期間，本書作者從實際執行獲得經驗中，也發現幾點待考量之風險與議題，值得未來相關研究者深入著墨，分別是「新冠病毒疫情影響與參與者流失」、「目的優先序與成本」。說明如下。

參與者的流失（attrition）一向是長期追蹤設計研究者關注的焦點之一，而其影響會隨著時間遞移或環境變化加大。因此，研究者往往會透過初始加重參與者人數的抽取、權重校正等方式來因應，然而，不可控制的大環境因素是統計技術無法觸及者。一如本書所述，2021年5月時，因為新冠病毒疫情影響，致使學校實體課程被迫停止，TASAL 即使是以線上評量施測，其執行仍得考慮無關干擾因子影響，例如，居家遠距評量無法掌握學生作答環境穩定性、試題保密性等，使得九年級資料蒐集未能如預期推動。在現行仍受新冠病毒疫情干擾脈絡下，未來的發展也許可從線上評量系統更新或更替抽樣方式來因應。

就目的優先序與成本而言，誠如本書所述，TASAL 目的之一是透過學生實徵表現回饋國家課程綱要內容，另一是探究影響學生成長表現因子，以回饋教育政策制訂等。就檢視現行國家課程綱要內容，相關學習表現與內容界定是以「學習階段」來區分，並未涉及年級間差異，因此，若挑選具該學習階段代表性的年級（例如，完整接受完八年級課程學生）進行橫斷面代表性抽樣，其評量表現結果同樣能回應第一個目的，同時，也可適時避免現行大環境新冠病毒疫情不確定性干擾（如透過延期施測）與支出相對較少時間、人力成本完成資料收集；然而，在回應第二個目的時，橫斷面代表性抽樣設計反會降低探究因素間因果解釋可行性。整體而言，不論是透過長期追蹤、抑或是橫斷面抽樣設計，各有其優劣勢，然而，其決定關鍵在於目的優先序與成本評估等政策考量，建議未來研究者適時納入大環境變動因素，來進行下一學習階段抽樣與評量設計。

### 第三節 詞彙說明

有關評量指引建置與學生表現標準描述中，本書所使用動詞、名詞之詞彙說明，主要參考與修訂自 Anderson 等人（2001）、NAEP 2019 閱讀評量架構（NAGB, 2019, pp. 59-63）、國際素養協會（International Literacy Association, ILA）網站（前身為國際閱讀協會（International Reading Association, IRA））與相關學者所提供詞彙說明等，而實際詞彙定義某程度會因應本書需求與概念，進行調整。逐一就動詞、名詞詞彙使用，簡要說明如下。

#### 一、動詞詞彙的使用

記得（remember）：能從長期記憶中擷取（retrieve）相關知識，而相關詞彙包含辨識（recognize, identify）、找出（locate）、回憶（recall）等（Anderson et al., 2001, p. 69），同時，本書再納入再現或再製（reproduce），以表徵學習者對於長期記憶中語言知識的擷取、重現過程。

使用（use）：在 Anderson 等人（2001）所提的分類架構中，使用（use）一詞屬於應用範疇，然而，本書為強調學習者所運用知識，將其歸於基本技能（相對於其他複雜程序），例如，使用字母拼讀。

結合 (combine)：連結 (link) 兩個或多個微觀命題關係的過程 (Kintsch, 1988, 2019)。

分類 (classify)：能辨識某個物件屬於某個類別的過程，而相關詞彙包含有歸類 (categorizing)、舉例 (exemplifying)，前者是由多個具體實例，找出原則的過程；後者為從原則中，找出特定例子的過程 (Anderson et al., 2001, p. 72)。此外，若對照英語文課程綱要學習表現，相關詞彙包含排序 (sort)，係指針對不同類別特徵，有規則進行排序。

描述 (describe)：透過文字或聲音來呈現 (present)、傳遞 (convey) 訊息的過程。在轉化為理解評量中，學習者得從文內收集 (gather) (直接擷取多個) 訊息，透過表達 (expression) 來傳遞其對文意之理解，才足以證明其具備相關能力。

摘要 (summarize)：以共通性 (general) 主題來重新表徵既有文本內容的過程，而其相似詞彙包含如概括 (generalize) (Anderson et al., 2001, p. 73)。此外，依照馬行誼 (2014) 觀點，認為摘要是主觀、濃縮的摘取文中重點 (main point) 的過程 (頁 55)，而程序會涉及刪除不必要訊息 (如分辨要點與細節)、重新概括與表徵等，有別於能統整全文觀點之主旨 (theme)。

推論 (infer)：從已知前提或訊息，經判斷程序得到邏輯關聯與結論的過程，在此過程，學習者往往會運用到比較 (compare)、對照 (contrast) 等程序 (Anderson et al., 2001; NAGB, 2019)。在理解認知歷程發展中，「推論」一詞包含許多層次意涵，可以自較低層次的橋接式推論 (bridging inference)、連結式推論 (connecting inference)、參照式推論 (referential inference)、簡單推論 (make simple inference) (Kintsch, 2019; Mullis & Martin, 2015; NAGB, 2019)，來代表學習者推論所需知識，是基於一般常識性、邏輯或僅涉及局部脈絡基礎；若再往高層次發展，往往會涉及複雜推論，學習者得連結自身背景知識、經驗，以形成更完整的語意連貫性，例如，透過自身背景知識來填補文內未明示訊息，所進行的精緻化推論 (elaborative inference)、解釋推論 (explanatory inference)，甚至是在覺察衝突或不一致時，能主動對訊息進行評估 (evaluative inference)、批判式推論 (critical inference)、作出未來預期之預測推論 (predictive inference) 等過程 (Graesser et al., 1994; OECD, 2019; van den Broek & Helder, 2017)。本書所使用「解釋、評估、預測」等詞彙，都是基於此前提，

學習者得在理解文意基礎上，透過表達來解釋文內不明訊息或進一步提供推論後之評估、批判、預期性相關內容，才得足以支持其具備相關能力。

比較 (compare)：偵測兩個或多個想法、物件、事件的相似、相異性的過程，相關詞彙為對照 (contrast) (Anderson et al., 2001, p. 75)。由於其運作涉及推論過程，因此學習者所比較與對照的物件，也可能因內容屬性、複雜性而具不同層次，例如，常識性範圍內的比較、抑或是抽象概念比較。

應用 (apply)：使用某個步驟、程序來執行任務的過程 (Anderson et al., 2001, p. 77)，而本書用以表徵學習者在理解文意後，使用文本所呈現原則、規則、步驟、概念等，進行延伸運用之過程。

說明 (interpret)：轉換不同訊息表徵的過程 (Anderson et al., 2001, p. 70)，例如，改述、換句話說 (paraphrase)、用另一種表達方式來轉譯 (translate) 文本內容等。

分析 (analyze)：將素材拆解成各組成成分 (constituent parts) 的過程，同時，能確定各組成成分間關聯、各成分與整體結構 (structure) 之關聯 (Anderson et al., 2001, pp. 79-80)，例如，能從文內舉證或說明某組成成分與整體結論、作者目的、文本結構、寫作手法、上層目標等關聯，而相關概念類似 Graesser 等人 (1994) 認為形成整體連貫性推論 (global coherence inference) 之過程。

區分 (differentiate)：就關聯性或重要性，來區分涉及整體結構中各組成成分 (constituent parts) 的過程 (Anderson et al., 2001, pp. 79-80)。其屬於分析的相關詞彙，運作過程涉及文本上位結構性組織 (structural organization) 與整體 (whole)，因此，有別於分辨文內細節與要點，多運用於分析文本整體或結構組成成分，例如，區分事實與觀點 (Schell, 1967)。

統整 (integrate)：將部分 (parts) 統合、綜合，以形成整體 (whole) 的過程，而相關詞彙包含統合、綜合 (synthesize)、組織 (organize) (Anderson et al., 2001, pp. 81, 276)。Kintsch (1988, 1998, 2019) 建構統整模式認為學習者以命題來分析文本，而在其執行理解的過程，會不斷建構新命題及參照式推論，直到學習者了解文本整體意涵。因此，學習者在此過程，除了理解文本內容外，會融入自身背景經驗及知識，對於文本內容進行複雜推論，以形成整體理解。

評估 (evaluate)：依據某個準則與標準進行判斷，而相關詞彙包含判斷 (judge)、依據內在準則進行檢核 (check)、以外在標準進行批判 (critique)，其中，判斷不一定會涉及評估，往往得連結某個標準或準則，才會啟動評估過程 (Anderson et al., 2001, p. 83)。此外，van den Broek 等人 (2011) 認為讀者對於評估文意連貫性標準 (standard for coherence)，同時涉及其有意識 (顯性) (explicit)、無意識 (隱性) (implicit) 覺察，而連貫性類型如，透過訊息參照、因果、時空或邏輯標準等 (Graesser et al., 1994)。

省思 (reflect)：涉及主動、持續與謹慎思考任何支持的信念或知識與進一步形成的結論 (Dewey, 1933, p. 118)。Kintsch (2019) 建構之統整模式認為理解是一不斷遞迴過程，而學習者得主動、持續有意識覺察 (aware) 歷程的運作與評估，才能達到實質的整體理解。

## 二、名詞詞彙

音素 (phoneme)：組成聲音的最小單位 (ILA, n.d.)。

字母拼讀 (phonics)：重視聲音與形素 (graphemes) 間系統性關聯的教學取徑 (ILA, n.d.)。

細節 (detail)：提供支持文內重點 (main point) 的訊息，例如，支持作者態度、重點形成的訊息 (NAGB, 2019, p. 60)。

文句脈絡線索 (context clues)：作者提供用來支持讀者連結背景知識與文本內容之暗示、線索，以進行字彙意義 (meaning of a word) 或訊息理解 (ILA, n.d.)。

情節 (plot)：具次序性事件 (events) 或行動 (actions) (NAGB, 2019, p. 62)。

場景 (setting)：事件發生的時間與地點 (NAGB, 2019, p. 63)。

證據 (evidence)：能支持文本內主張 (claim)、論點 (argument) 的內容，包含如數據資料、引用官方來源或能支持論證所提供的事實性說明 (ILA, n.d.)。

觀點 (perspective)：某件事情被思考或評估的立場、主張 (NAGB, 2019, p. 62)。

概念 (concept)：由抽象或通用想法 (generic ideas) 所概括形成的主題，分享著共同屬性 (ILA, n.d.)。

大意 (main idea) 與主旨 (theme)：依據馬行誼 (2014, 頁 55) 論述，認為大意是客觀、簡要的呈現作者寫作的內容，然而，主旨係指閱讀者依照文章結構意義關係，經統整、抽繹出能概括全文的主要觀點。若對應理解認知歷程，大意需要學習者直接推論或結合命題，涉及理解字面文意成分，而主旨涉及學習者統整與概括過程，屬隱含意理解層次。

人物動機或意圖 (character's motivation, intention)：促使人物以特定方式行動或導致某情境結果的一系列條件 (NAGB, 2019, p. 62)。

氣氛與情緒 (mood and emotion)：讀者藉由文本傳遞訊息所使用語氣或其他特徵，所理解出的氣氛 (atmosphere)、情緒 (emotion) 或態度 (attitude)。

衝突 (conflict)：文本內所呈現人物對立的想法、情緒、立場或其所傳遞與學習者內在標準不一致性訊息所引起之衝突 (NAGB, 2019; OECD, 2019)。

口音、腔調 (accent)：具區別性的語言發音，屬於方言一環 (ILA, n.d.)。

特質 (trait)：人所具備的獨特、區別性特徵 (NAGB, 2019, p. 63)。

慣用語 (idiom)：由片語、一段陳述所組成，其意涵往往為非字面文意所組成 (Merriam-Webster, n.d.)。

事實與意見 (fact and opinion)：事實為真實存在的事物，具可驗證性，例如，被定義 (defined)、實徵 (empirical) 證明的結果；意見是基於信念所形成，即使融入事實性內容，但仍具有推測、判斷性的結果 (Schell, 1967)。

作者觀點 (author's point of view)：文本作品被作者傳遞或呈現的觀點 (NAGB, 2019, p. 62)。

作者目的 (author's purpose)：作者寫作特定理由，代表作者期望建立的目標 (NAGB, 2019, p. 62)。

文本結構 (text structure)：依照目的所形成不同組織文本訊息的方式，例如，比較與對照 (compare and contrast)、問題與因應 (problem and solution)、因果 (cause and effect)、時序 (sequences) 等。其相關詞彙包含如故事結構 (story grammar, story structure)，包含如人物 (characters)、場景 (setting) 等安排 (ILA, n.d.; Rumelhart, 1975)。

寫作風格或手法（author's craft or writing style）：為傳遞某類訊息，作者所運用的語言技巧以進行寫作或展現出方式特徵，包含如訊息傳遞的語氣（tone）、意象（imagery）、隱喻（metaphor）、舉例（example）、使用平行結構（parallel structure）、邏輯論證（logical arguments）、情緒訴求（emotional appeal）等（NAGB, 2019）。

## 參考文獻

- 方德隆（2010）。九年一貫課程綱要總綱實施現況相關研究之後設分析。**課程研究**，5（2），63-108。
- 余曉清、林煥祥（主編）（2017）。**PISA 2015 臺灣學生的表現**。心理。
- 吳俊憲、黃政傑（2010）。中小學課程政策改革之研究：九年一貫課程的回顧與前瞻。**課程研究**，5（2），47-62。
- 吳璧純、詹志禹（2018）。從能力本位到素養導向教育的演進、發展及反思。**教育研究與發展期刊**，14（2），35-64。<http://doi.org/10.3966/181665042018061402002>
- 宋曜廷、劉佩雲、簡馨瑩（2003）。閱讀動機量表的修訂及相關因素研究。**測驗學刊**，50（1），47-72。<http://doi.org/10.7108/PT.200306.0047>
- 林明佳（2018）。從「能力」到「素養」的英語教學：溝通式教學觀的情境與活用。**教育研究月刊**，294，72-90。<http://doi.org/10.3966/168063602018100294005>
- 侯佩君、杜素豪、廖培珊、洪永泰、章英華（2008）。台灣鄉鎮市區類型之研究：「台灣社會變遷基本調查」第五期計畫之抽樣分層效果分析。**調查研究－方法與應用**，23，7-32。<http://doi.org/10.7014/TCYCFYYY.200804.0007>
- 柯華葳、張郁雯、詹益綾、丘嘉慧（2017）。**PIRLS 2016 臺灣四年級學生閱讀素養國家報告**。國立中央大學學習與教學研究所。
- 洪裕宏（2008）。**界定與選擇國民核心素養：概念參考架構與理論基礎研究**（專題研究計畫成果報告，NSC 95-2511-S-010-001）。行政院國家科學委員會；國立陽明大學。
- 馬行誼（2014）。閱讀摘要教學的幾點省思。**語教新視野**，1，51-65。[http://doi.org/10.6751/NVLA.201403\\_\(1\).0004](http://doi.org/10.6751/NVLA.201403_(1).0004)
- 國立臺灣師範大學心理與教育測驗研究發展中心（無日期）。**國中教育會考**。<https://cap.rcpet.edu.tw/background.html>
- 國家教育研究院（無日期）。**臺灣學生成就長期追蹤評量計畫**。<https://tasal.naer.edu.tw/>
- 國家教育研究院（2014）。**十二年國民基本教育課程發展建議書**。<https://ws.moe.edu.tw/001/Upload/23/refile/8006/51082/5dcfe708-8956-4324-867e-564fdecbe001.pdf>

- 國家教育研究院（2018a）。十二年國民基本教育課程綱要國民中小學暨普通型高級中等學校：語文領域—英語文。<https://www.naer.edu.tw/upload/1/16/doc/812/>(發布版)國民中小學暨普通型高級中等學校 - 語文領域 - 英語文課程綱要 .pdf
- 國家教育研究院（2018b）。十二年國民基本教育課程綱要國民中小學暨普通型高級中等學校：語文領域—英語文課程手冊。<https://www.naer.edu.tw/upload/1/16/doc/2017/>語文領域 - 英語文課程手冊（定稿版）.pdf
- 國家教育研究院課程及教學研究中心核心素養工作圈（2015）。十二年國民基本教育領域課程綱要：核心素養發展手冊。<https://ws.moe.edu.tw/001/Upload/23/refile/8006/51358/9df0910c-56e0-433a-8f80-05a50efeca72.pdf>
- 張武昌、楊承淑、葉錫南、游毓玲、陳秋蘭、林于仙（2013）。十二年國民基本教育外國語文領域綱要內容之前導研究（NAER-102-06-A-1-02-02-1-11）。國家教育研究院。
- 教育部（2017，5月10日）。十二年國民基本教育課程綱要總綱生效令。<https://edu.law.moe.gov.tw/LawContent.aspx?id=GL001630#lawmenu>
- 教育部（2021）。十二年國民基本教育課程綱要總綱。[https://www.naer.edu.tw/upload/1/16/doc/288/\(111學年度實施\)十二年國教課程綱要總綱.pdf](https://www.naer.edu.tw/upload/1/16/doc/288/(111學年度實施)十二年國教課程綱要總綱.pdf)
- 教育部全球資訊網（2016，12月1日）。總綱中提出「核心素養」，這和原有九年一貫課程強調培養孩子帶著走的「基本能力」有什麼不同？[https://www.edu.tw/News\\_Content.aspx?n=BA5E856472F10901&sms=5588FE86FEB94225&s=E62B1B1EBC8AE952](https://www.edu.tw/News_Content.aspx?n=BA5E856472F10901&sms=5588FE86FEB94225&s=E62B1B1EBC8AE952)
- 教育部國民及學前教育署（無日期）。97年國民中小學九年一貫課程綱要。<https://www.k12ea.gov.tw/Tw/Common/SinglePage?filter=F34E9417-8374-4A02-97CC-0DEDFD13514F>
- 教育部統計處（無日期）。各級學校地理資訊及地區別統計查詢。<https://stats.moe.gov.tw/EduGis/>
- 陳伯璋、張新仁、蔡清田、潘慧玲（2007）。界定與選擇國民核心素養：概念參考架構與理論基礎研究—子計畫三：全方位的國民核心素養之教育研究（專題研究計畫成果報告，NSC 95-2511-S-003-001）。行政院國家科學委員會；致理管理學院教育研究所。

- 黃純敏、程玉秀、林秀珍（2004）。國中英語教師課程意識研究：九年一貫課程之回應。**英語教學**，**28**（4），21-44。
- 黃淑真（2007）。外語學習動機理論的發展與教學研究的回顧。**英語教學**，**31**（3），101-124。
- 黃馨瑩（2017）。抽樣設計與權重。載於蕭儒棠、曾建銘、謝佩蓉、黃馨瑩、吳慧珉、陳冠銘、蔡明學、謝名娟、謝進昌（編著），**大型教育調查研究實務：以 TASA 為例**（頁 129-168）。國家教育研究院。
- 葉錫南（2016）。由十二年國教課綱草案展望未來中小學之英語文教科書。**教科書研究**，**9**（3），143-168。[http://doi.org/10.6481/JTR.201612\\_9\(3\).05](http://doi.org/10.6481/JTR.201612_9(3).05)
- 蔡清田（2011）。從課程學理基礎與核心素養論 K-12 年級課程綱要。**課程研究**，**6**（2），63-83。
- 蔡清田、洪若烈、陳延興、徐彬瀚、葉雅卿（2012）。**K-12 各教育階段核心素養與各領域課程統整研究**（NAER-100-16-A-1-02-00-2-03）。國家教育研究院；國立中正大學課程研究所。
- 蔡清田、陳延興、梁玲銀、葉雅卿（2011）。**K-12 中小學課程綱要的核心素養與各領域之連貫體系研究**（NAER-99-12-A-1-05-00-2-11）。國家教育研究院；國立中正大學課程研究所。
- 蕭儒棠、曾建銘、謝佩蓉、黃馨瑩、吳慧珉、陳冠銘、蔡明學、謝名娟、謝進昌（2017）。**大型教育調查研究實務：以 TASA 為例**。國家教育研究院。
- 謝進昌（2020，10月16日）。**英語文標準本位大型教育評量建置與效度評估研究**（論文發表）。2020 中國測驗學會年會暨學術推廣工作坊，新北市，臺灣。
- 謝進昌（2021a）。**第四學習階段英語文素養長期追蹤**（NAER-2019-041-A-1-1-E1-07）。國家教育研究院。
- 謝進昌（2021b）。以「補充性表現水平描述輔助自陳式測量構念」之延伸 Angoff 標準設定研究。**教育心理學報**，**53**（2），307-334。[http://doi.org/10.6251/BEP.202112\\_53\(2\).0003](http://doi.org/10.6251/BEP.202112_53(2).0003)
- 謝進昌（2022a）。**英語文素養導向評量命題技巧**。國家教育研究院。
- 謝進昌（2022b）。「混合專家與學生實徵表現導向」大型教育評量標準設定之效度評估研究（出版中）。**教育科學研究期刊**。國立臺灣師範大學。

- 關秉寅、李敦義 (2008)。補習數學有用嗎？一個「反事實」的分析。《臺灣社會學刊》，41，97-148。
- Adams, R. J. (2005). Reliability as a measurement design effect. *Studies in Educational Evaluation, 31*(2-3), 162-172. <https://doi.org/10.1016/j.stueduc.2005.05.008>
- Adams, R. J., & Wu, M. L. (2007). The mixed-coefficients multinomial logit model: A generalized form of the Rasch model. In M. von Davier & C. H. Carstensen (Eds.), *Multivariate and mixture distribution Rasch models: Extensions and applications* (pp. 57-75). Springer. [https://doi.org/10.1007/978-0-387-49839-3\\_4](https://doi.org/10.1007/978-0-387-49839-3_4)
- Adams, R. J., Wilson, M., & Wang, W.-C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement, 21*(1), 1-23. <http://doi.org/10.1177/0146621697211001>
- Al-Hoorie, A. H. (2017). Sixty years of language motivation research: Looking back and looking forward. *SAGE Open, 7*(1), 1-11. <https://doi.org/10.1177/2158244017701976>
- Alvermann, D. E., & Moje, E. B. (2019). A relational model of adolescent literacy instruction: Disrupting the discourse of “every teacher a teacher of reading”. In D. E. Alvermann, N. J. Unrau, M. Sailors, & R. B. Ruddell (Eds.), *Theoretical models and processes of literacy* (7th ed., pp. 362-380). Routledge.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. AERA Publications.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Andrews, R., & Hertzberg, F. (2009). Introduction: Special issue on argumentation in education in Scandinavia and England. *Argumentation, 23*, 433-436. <https://doi.org/10.1007/s10503-009-9168-5>
- Ardasheva, Y., Wang, Z., Adesope, O. O., & Valentine, J. C. (2017). Exploring effectiveness and moderators of language learning strategy instruction on second language and self-regulated learning outcomes. *Review of Educational Research, 87*(3), 544-582. <https://doi.org/10.3102/0034654316689135>

- Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Reading Behavior*, *11*(4), 365-374. <https://doi.org/10.1080/10862967909547342>
- Baker, L. (2005). Developmental differences in metacognition: Implications for metacognitively oriented reading instruction. In S. E. Israel, C. C. Block, K. L. Bauserman, & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction, and professional development* (pp. 61-79). Lawrence Erlbaum Associates.
- Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of Reading Research* (pp. 353-394). Longman.
- Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly*, *34*(4), 452-477. <https://doi.org/10.1598/RRQ.34.4.4>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Beaton, A. E. (1987). *Implementing the new design: The NAEP 1983-84 technical report* (No. 15-TR-20). National Assessment of Educational Progress.
- Berge, K. L., Evensen, L. S., & Thygesen, R. (2016). The wheel of writing: A model of the writing domain for the teaching and assessing of writing as a key competency. *The Curriculum Journal*, *27*(2), 172-189. <https://doi.org/10.1080/09585176.2015.1129980>
- Biber, D. (1989). A typology of English texts. *Linguistics*, *27*(1), 3-43. <https://doi.org/10.1515/ling.1989.27.1.3>
- Binkley, M., & Kelly, D. (2003). *A content comparison of the NAEP and PIRLS fourth-grade reading assessments* (NCES No. 2003-10). U.S. Department of Education, National Center for Education Statistics.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (Eds.). (1956). *Taxonomy of educational objectives: Handbook I: Cognitive domain*. David McKay.
- Bock, R. D., & Mislevy, R. J. (1982). Adaptive EAP estimation of ability in a microcomputer environment. *Applied Psychological Measurement*, *6*(4), 431-444. <https://doi.org/10.1177/014662168200600405>

- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1-40. <https://doi.org/10.1023/A:1021302408382>
- Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly*, *44*(1), 6-28. <https://doi.org/10.1598/RRQ.44.1.1>
- Bråten, I., Strømsø, H. I., & Salmerón, L. (2011). Trust and mistrust when students read multiple information sources about climate change. *Learning and Instruction*, *21*(2), 180-192. <https://doi.org/10.1016/j.learninstruc.2010.02.002>
- Britt, M. A., & Rouet, J.-F. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (pp. 276-314). Cambridge University Press. <https://doi.org/10.1017/CBO9781139048224.017>
- Britt, M. A., Rouet, J.-F., & Braasch, J. L. G. (2013). Documents as Entities: Extending the situation model theory of comprehension. In M. A. Britt, S. R. Goldman, & J.-F. Rouet (Eds.), *Reading—From words to multiple texts* (pp. 160-179). Routledge. <https://doi.org/10.4324/9780203131268-15>
- Britt, M. A., Rouet, J.-F., & Durik, A. M. (2018). *Literacy beyond text comprehension: A theory of purposeful reading*. Routledge.
- Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31-42. <https://doi.org/10.1037/0022-0663.96.1.31>
- Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 2-27). Longman.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1-47. <https://doi.org/10.1093/applin/I.1.1>

- Childs, R. A., & Jaciw, A. P. (2002). Matrix sampling of items in large-scale assessments. *Practical Assessment, Research, and Evaluation*, 8(16), 1-9. <https://doi.org/10.7275/gwvh-4z51>
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Sage. <https://doi.org/10.4135/9781412985918>
- Cohen, A. D., & Macaro, E. (2007). *Language learner strategies: Thirty years of research and practice*. Oxford University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the internet. *Reading Research Quarterly*, 42(2), 214-257. <https://doi.org/10.1598/RRQ.42.2.2>
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment—Companion volume*. <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. <https://doi.org/10.1037/h0040957>
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222-251. <http://doi.org/10.3102/00346543049002222>
- Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70-89). Cambridge University Press. <https://doi.org/10.1017/CBO9780511620652.006>
- Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, 47(2), 108-121. <https://doi.org/10.1111/1467-8527.00106>
- De Beaugrande, R.-A., & Dressler, W. U. (1981). *Introduction to text linguistics*. Longman.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum.

- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. D.C. Heath and Company.
- Dörnyei, Z., & Csizér, K. (2002). Some dynamics of language attitudes and motivation: Results of a longitudinal nationwide survey. *Applied Linguistics*, 23(4), 421-462. <https://doi.org/10.1093/applin/23.4.421>
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75-146). W. H. Freeman.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109-132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>
- Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61, Article 101859. <https://doi.org/10.1016/j.cedpsych.2020.101859>
- European Commission. (2007). *Key competences for lifelong learning: European reference framework*.
- Field, J. (2013). Cognitive validity. In A. Geranpayeh & L. Taylor (Eds.), *Studies in language testing No. 35. Examining listening: Research and practice in assessing second language listening* (pp. 77-151). Cambridge University Press.
- Gagne, E. D., Yekovich, C. W., & Yekovich, F. R. (1993). *The cognitive psychology of school learning* (2nd ed.). HarperCollins College Publishers.
- Gambrell, L. B., & Bales, R. J. (1986). Mental imagery and the comprehension-monitoring performance of fourth- and fifth-grade poor readers. *Reading Research Quarterly*, 21(4), 454-464. <https://doi.org/10.2307/747616>
- Gee, J. P. (2019). Reading as situated language: A sociocognitive perspective. In D. E. Alvermann, N. J. Unrau, M. Sailors, & R. B. Ruddell (Eds.), *Theoretical models and processes of literacy* (7th ed., pp. 105-117). Routledge.

- Geranpayeh, A., & Taylor, L. (Eds.). (2013). *Studies in language testing, No. 35. Examining listening: Research and practice in assessing second language listening*. Cambridge University Press.
- Glatthorn, A. A., Bragaw, D., Dawkins, K., & Parker, J. (1998). *Performance assessment and standards-based curricula: The achievement cycle*. Eye on Education.
- Gonzalez, E., & Rutkowski, L. (2010). Principles of multiple matrix booklet designs and parameter recovery in large-scale assessments. In M. von Davier & D. Hastedt (Eds.), *IERI monograph series: Issues and methodologies in large-scale assessments* (Vol. 3, pp. 125-156). IEA-ETS Research Institute.
- Goodman, K. (1982). *Language and literacy: The selected writings of Kenneth S. Goodman*. Routledge; Kegan Paul.
- Goodman, K. (1988). The reading process. In P. L. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 11-21). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524513.005>
- Gough, P. B. (1972). One second of reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading* (pp. 331-358). MIT Press.
- Grabe, W. (2002). Narrative and expository macro-genres. In A. M. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 249-267). Lawrence Erlbaum Associates.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371-395. <https://doi.org/10.1037/0033-295X.101.3.371>
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology* (2nd ed.). John Wiley & Sons.
- Guo, D., Zhang, S., Wright, K. L., & McTigue, E. M. (2020). Do you get the picture? A meta-analysis of the effect of graphics on reading comprehension. *AERA Open*, *6*(1), 1-20. <https://doi.org/10.1177/2332858420901696>
- Guthrie, J. T., & Klauda, S. L. (2016). Engagement and motivational processes in reading. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 41-53). Routledge.

- Hacker, D. J. (2004). Self-regulated comprehension during normal reading. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 755-779). International Reading Association.
- Hambleton, R. K. (2001). Setting performance standards on educational assessments and criteria for evaluating the process. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 89-116). Lawrence Erlbaum Associates.
- Hambleton, R. K., Pitoniak, M. J., & Copella, J. M. (2012). Essential steps in setting performance standards on educational tests and strategies for assessing the reliability of results. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 47-76). Routledge.
- Henry, A. (2011). Gender differences in L2 motivation: A reassessment. In S. A. Davies (Ed.), *Gender gap: Causes, experiences and effects* (pp. 81-101). Nova Science Publishers.
- Henry, L. A. (2006). SEARCHing for an answer: The critical role of new literacies while reading on the internet. *The Reading Teacher*, 59(7), 614-627. <https://doi.org/10.1598/RT.59.7.1>
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Hymes, D. H. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269-293). Penguin.
- International Activities Program. (2010). *Comparison of the PISA 2009 and NAEP 2009 reading assessments*. U.S. Department of Education, National Center for Education Statistics.
- International Literacy Association. (n.d.). *Literacy glossary*. <https://www.literacyworldwide.org/get-resources/literacy-glossary>
- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64(1), 160-212. <https://doi.org/10.1111/lang.12034>
- Kane, M. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, 64(3), 425-461. <https://doi.org/10.3102/00346543064003425>

- Kane, M. T. (2001). So much remains the same: Conception and status of validation in setting standards. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 53-88). Lawrence Erlbaum Associates.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1-73. <https://doi.org/10.1111/jedm.12000>
- Khalifa, H., & Weir, C. J. (2009). *Studies in language testing No. 29. Examining reading: Research and practice in assessing second language reading*. Cambridge University Press.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163-182. <https://doi.org/10.1037/0033-295X.95.2.163>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Kintsch, W. (2019). Revisiting the construction—Integration model of text comprehension and its implications for instruction. In D. E. Alvermann, N. J. Unrau, M. Sailors, & R. B. Ruddell (Eds.), *Theoretical models and processes of literacy* (7th ed., pp. 178-203). Routledge.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394. <https://doi.org/10.1037/0033-295X.85.5.363>
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd ed.). Springer.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293-323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)
- LaRoche, S., Joncas, M., & Foy, P. (2020). Sample design in TIMSS 2019. In M. O. Martin, M. von Davier, & I. V. S. Mullis (Eds.), *Methods and procedures: TIMSS 2019 technical report* (pp. 3.1-3.33). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College; International Association for the Evaluation of Educational Achievement. <https://timssandpirls.bc.edu/timss2019/methods/chapter-3.html>

- Lee, D. Y.-W. (2001). Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology*, 5(3), 37-72. <https://doi.org/10.125/44565>
- Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., & Henry, L. A. (2013). New literacies: A dual-level theory of the changing nature of literacy, instruction, and assessment. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 1150-1181). International Reading Association.
- Lietz, P. (2006). A meta-analysis of gender differences in reading achievement at the secondary school level. *Studies in Educational Evaluation*, 32(4), 317-344. <https://doi.org/10.1016/j.stueduc.2006.10.002>
- Linacre, J. M. (2005). *A user's guide to Winsteps Ministep Rasch-Model computer programs*. IL MESA Press.
- Linn, R. L., & Herman, J. L. (1997). *A policymaker's guide to standards-led assessment*. Education Commission of the States.
- Lissitz, R. W., & Samuelson, K. (2007). A suggested change in terminology and emphasis regarding validity and education. *Educational Researcher*, 36(8), 437-448. <https://doi.org/10.3102/0013189X07311286>
- Lumley, T. (2019). *Package 'mitools': Tools for multiple imputation of missing data*. Retrieved January 18, 2021, from <https://cran.r-project.org/web/packages/mitools/mitools.pdf>
- Lumley, T. (2020). *Package 'survey': Analysis of complex survey samples*. Retrieved January 18, 2021, from <https://cran.r-project.org/web/packages/survey/survey.pdf>
- Mann, W. C., & Thompson, S. A. (1986). Relational propositions in discourse. *Discourse Processes*, 9(1), 57-90. <https://doi.org/10.1080/01638538609544632>
- Manoli, P., & Papadopoulou, M. (2012). Reading strategies versus reading skills: Two faces of the same coin. *Procedia - Social and Behavioral Sciences*, 46, 817-821. <https://doi.org/10.1016/j.sbspro.2012.05.205>
- Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of

- lurking jingle-jangle fallacies. *Journal of Educational Psychology*, *111*(2), 331-353. <https://doi.org/10.1037/edu0000281>
- Masgoret, A.-M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, *53*(1), 123-163. <https://doi.org/10.1111/1467-9922.00212>
- Masters, G. N., & Wright, B. D. (1997). The partial credit model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 101-121). Springer. [https://doi.org/10.1007/978-1-4757-2691-6\\_6](https://doi.org/10.1007/978-1-4757-2691-6_6)
- McDonald, R. P., & Ho, M.-H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, *7*(1), 64-82. <http://doi.org/10.1037/1082-989X.7.1.64>
- Merriam-Webster. (n.d.). Idiom. In *Merriam-Webster.com dictionary*. <https://www.merriam-webster.com/dictionary/idiom>
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). Macmillan Publishing; American Council on Education.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741-749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Mislevy, R. J., Beaton, A. E., Kaplan, B., & Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, *29*(2), 133-161. <https://doi.org/10.1111/j.1745-3984.1992.tb00371.x>
- Mullis, I. V. S., Kennedy, A. M., Martin, M. O., & Sainsbury, M. (2006). *PIRLS 2006 assessment framework and specifications* (2nd ed.). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College; International Association for the Evaluation of Educational Achievement. <https://timssandpirls.bc.edu/PDF/P06Framework.pdf>
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 assessment framework* (2nd ed.). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College; International Association for the Evaluation of Educational Achievement. <https://timssandpirls.bc.edu/pirls2016/framework.html>

- National Assessment Governing Board. (2019). *Reading framework for the 2019 National Assessment of Educational Progress*. U.S. Department of Education.
- Nolen, S. B., Horn, I. S., & Ward, C. J. (2015). Situating motivation. *Educational Psychologist, 50*(3), 234-247. <https://doi.org/10.1080/00461520.2015.1075399>
- Organisation for Economic Co-operation and Development. (2005). *The definition and selection of key competencies*.
- Organisation for Economic Co-operation and Development. (2019). *PISA 2018 assessment and analytical framework*. <https://doi.org/10.1787/b25efab8-en>
- Organisation for Economic Co-operation and Development. (2021). *PISA 2025 foreign language assessment framework*.
- Osborne, J. F., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education, 95*(4), 627-638. <https://doi.org/10.1002/sce.20438>
- Östman, J.-O., & Virtanen, T. (2011). Text and discourse linguistics. In J. Zienkowski, J.-O. Östman, & J. Verschueren (Eds.), *Handbook of pragmatics highlights: Vol. 8. Discursive pragmatics* (pp. 266-285). John Benjamins Publishing Company. <https://doi.org/10.1075/hoph.8.15ost>
- Oxford, R. L. (2002). Language learning strategies in a nutshell: Update and ESL suggestions. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 124-132). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667190.018>
- Oxford, R. L., Nyikos, M., & Ehrman, M. (1988). Vive la difference? Reflections on sex differences in use of language learning strategies. *Foreign Language Annals, 21*(4), 321-329. <https://doi.org/10.1111/j.1944-9720.1988.tb01076.x>
- Pant, H. A., Rupp, A. A., Tiffin-Richards, S. P., & Köller, O. (2009). Validity issues in standard-setting studies. *Studies in Educational Evaluation, 35*(2-3), 95-101. <https://doi.org/10.1016/j.stueduc.2009.10.008>
- Park, Y.-H. (2010). A relationship between reading comprehension and reading strategy use: Meta-analysis. *English Teaching, 65*(3), 3-22. <https://doi.org/10.15858/engtea.65.3.201009.3>

- Pearson, P. D., & Cervetti, G. N. (2015). Fifty years of reading comprehension theory and practice. In P. D. Pearson & E. H. Hiebert (Eds.), *Research-based practices for teaching common core literacy* (pp. 1-24). Teachers College Press.
- Pearson, P. D., & Johnson, D. D. (1978). *Teaching reading comprehension*. Holt, Rinehart and Winston.
- Perfetti, C., & Adlof, S. M. (2012). Reading comprehension: A conceptual framework from word meaning to text meaning. In J. P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 3-20). Rowman & Littlefield Education.
- Perkins, D. N., Allen, R., & Hafner, J. (1983). Difficulties in everyday reasoning. In W. Maxwell (Ed.), *Thinking: The expanding frontier* (pp. 177-189). The Franklin Institute Press.
- Plonsky, L. (2011). The effectiveness of second language strategy instruction: A meta-analysis. *Language Learning*, 61(4), 993-1038. <https://doi.org/10.1111/j.1467-9922.2011.00663.x>
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. RAND. Retrieved February 1, 2018, from [https://www.rand.org/content/dam/rand/pubs/monograph\\_reports/2005/MR1465.pdf](https://www.rand.org/content/dam/rand/pubs/monograph_reports/2005/MR1465.pdf)
- Reilly, D., Neumann, D. L., & Andrews, G. (2019). Gender differences in reading and writing achievement: Evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*, 74(4), 445-458. <https://doi.org/10.1037/amp0000356>
- Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the web. *Journal of the American Society for Information Science and Technology*, 53(2), 145-161. <https://doi.org/10.1002/asi.10017>
- Robitzsch, A., Kiefer, T., & Wu, M. (2020). *Package 'TAM': Test analysis modules*. Retrieved September 18, 2020, from <https://cran.r-project.org/web/packages/TAM/TAM.pdf>
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.
- Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211-236). Academic Press. <https://doi.org/10.1016/B978-0-12-108550-6.50013-6>

- Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 33-58). Lawrence Erlbaum Associates.
- Rumelhart, D. E. (2013). Toward an interactive model of reading. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 719-747). International Reading Association.
- Rutkowski, D., Thompson, G., & Rutkowski, L. (2020). Understanding the policy influence of international large-scale assessments in education. In H. Wagemaker (Ed.), *Reliability and validity of international large-scale assessment: Understanding IEA's comparative studies of student achievement* (pp. 261-277). Springer. [https://doi.org/10.1007/978-3-030-53081-5\\_15](https://doi.org/10.1007/978-3-030-53081-5_15)
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Rychen, D. S., & Salganik, L. H. (Eds.). (2003). *Key competencies for a successful life and a well-functioning society*. Hogrefe & Huber Publishers.
- Samuels, S. J. (2013). Toward a theory of automatic information processing in reading, revisited. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 698-718). International Reading Association.
- Schell, L. M. (1967). Distinguishing fact from opinion. *Journal of Reading*, 11(1), 5-9.
- Schiefele, U., Schaffner, E., Möller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly*, 47(4), 427-463. <https://doi.org/10.1002/RRQ.030>
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 199-227). Cambridge University Press.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363. <https://doi.org/10.1177/1362168808089921>

- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs*. American Educational Research Association.
- Schunk, D. H. (2012). Social cognitive theory. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol. 1. Theories, constructs, and critical issues* (pp. 101-123). American Psychological Association. <https://doi.org/10.1037/13273-005>
- Schunk, D. H., & Bursuck, W. D. (2016). Self-efficacy, agency, and volition: Student beliefs and reading motivation. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 54-66). Routledge.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Company.
- Shao, J., & Tu, D. (1995). *The jackknife and bootstrap*. Springer.
- Shepard, L. A. (1993). Evaluating test validity. In L. Darling-Hammond (Ed.), *Review of research in education* (Vol. 19, pp. 405-450). American Educational Research Association. <https://doi.org/10.2307/1167347>
- Slavin, R. E. (2017). Evidence-based reform in education. *Journal of Education for Students Placed at Risk*, 22(3), 178-184. <https://doi.org/10.1080/10824669.2017.1334560>
- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. National Academy Press. <https://doi.org/10.17226/6023>
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16(1), 32-71. <https://doi.org/10.2307/747348>
- Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing, Vol. 2. Advances in discourse processing* (pp. 53-120). Ablex Publishing Corporation.
- Thagard, P. (2008). Explanatory coherence. In J. E. Adler & L. J. Rips (Eds.), *Reasoning: Studies of human inference and its foundations* (pp. 471-513). Cambridge University Press. <https://doi.org/10.1017/CBO9780511814273.026>

- van den Broek, P., Bohn-Gettler, C. M., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 123-140). Information Age Publishing.
- van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, *54*(5-6), 360-372. <https://doi.org/10.1080/0163853X.2017.1306677>
- van der Slik, F. W. P., van Hout, R. W. N. M., & Schepens, J. J. (2015) The gender gap in second language acquisition: Gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *PLOS ONE*, *10*(11), Article e0142056. <https://doi.org/10.1371/journal.pone.0142056>
- van Dijk, T. A. (1988). *News as discourse*. Lawrence Erlbaum Associates.
- Verhoeven, L., & Perfetti, C. (2008). Advances in text comprehension: Model, process and development. *Applied Cognitive Psychology*, *22*(3), 293-301. <https://doi.org/10.1002/acp.1417>
- Vonkova, H., Jones, J., Moore, A., Altinkalp, I., & Selcuk, H. (2021). A review of recent research in EFL motivation: Research trends, emerging methodologies, and diversity of researched populations. *System*, *103*, Article 102622. <https://doi.org/10.1016/j.system.2021.102622>
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*(3), 427-450. <https://doi.org/10.1007/BF02294627>
- Weinstein, C. E., Husman, J., & Dierking, D. R. (2000). Self-regulation interventions with a focus on learning strategies. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 727-747). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50051-2>
- Wentzel, K. R. (1996). Social and academic motivation in middle school: Concurrent and long-term relations to academic effort. *The Journal of Early Adolescence*, *16*(4), 390-406. <https://doi.org/10.1177/0272431696016004002>
- Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology*, *89*(3), 411-419. <https://doi.org/10.1037/0022-0663.89.3.411>

- Wentzel, K. R., & Wigfield, A. (1998). Academic and social motivational influences on students' academic performance. *Educational Psychology Review*, *10*(2), 155-175. <https://doi.org/10.1023/A:1022137619834>
- Werlich, E. (1976). *A text grammar of English*. Quelle & Meyer.
- West, R. F., & Stanovich, K. E. (2000). Automatic contextual facilitation in readers of three ages. In K. E. Stanovich (Ed.), *Progress in understanding reading: Scientific foundations and new frontiers* (pp. 13-20). The Guilford Press.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Association for Supervision and Curriculum Development.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, *8*(3), 370-371.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.
- Yao, L., & Schwarz, R. D. (2006). A multidimensional partial credit model with associated item and test statistics: An application to mixed-format test. *Applied Psychological Measurement*, *30*(6), 469-492. <https://doi.org/10.1177/0146621605284537>

## 附錄

### 附錄一 教育目標（學習表現）與評量指引（認知成分） 漸近複雜對應

教育目標 (學習表現；邏輯思考、判斷與創造力)	理解認 知成分	評量指引 (認知成分) 細項
1-II-1 能聽辨 26 個字母 (註 3)		辨識字母或聽辨子音、母音與其組合 (註 3)
1-III-1 能聽辨英語的子音、母音及其不同的組合 (註 3)		
3-II-1 能辨識 26 個印刷體大小寫字母 (註 3)		
3-IV-1 能辨識連續書寫體大小寫字母		
5-IV-5 能運用字母拼讀規則讀出及拼寫英文字詞		
1-III-4 能聽辨課堂中所學的片語、句子及其重音 (註 3)	能流暢 運作語 言基本 知能 (L1)	使用字母拼讀規則 辨識出字詞
1-IV-1 能聽懂課堂中所學的字詞		
1-IV-2 能聽懂常用的教室用語及日常生活用語		
1-IV-3 能聽懂基本或重要句型的句子		
3-IV-2 能辨識課堂中所學的字詞		
3-IV-3 能看懂簡易的英文標示		聽辨、識讀字彙、 片語、常用語、基 本句型之句子
3-IV-5 能看懂簡易的生活用語		
3-IV-6 能看懂基本的句型		
9-III-1 能夠將所學字詞做簡易歸類 (註 3)		
1-IV-10 能了解歌謠、韻文的節奏與音韻		

(續下頁)

教育目標 (學習表現；邏輯思考、判斷與創造力)	理解認 知成分	評量指引 (認知成分)細項
1-IV-4 能聽懂日常生活對話的主要內容 1-IV-5 能聽懂簡易歌謠和韻文的主要內容 1-IV-6 能聽懂簡易故事及短劇的主要內容 1-IV-8 能聽懂簡易影片的主要內容 1-IV-11 能聽懂公共場所廣播的內容，如捷運、車站、機場廣播	能直接 找出明 顯訊息 (L2)	直接找出人事物、 時間、地點、物品 等外在顯著特徵 ----- 直接找出明顯陳述 的想法 ----- 直接找出與特定目 標有關的明顯訊息
1-IV-9 能辨識句子語調所表達的情緒和態度 (註 4) 3-IV-7 能了解對話的主要內容 3-IV-8 能了解短文、簡訊、書信的主要內容 3-IV-9 能了解故事的主要內容與情節 3-IV-10 能辨識簡易故事的要素，如背景、人物、 事件和結局 3-IV-13 能了解短劇的主要內容與情節等 3-IV-14 能快速閱讀了解文章重點，並有效應用於 廣泛閱讀中 (註 5) 3-IV-16 能閱讀不同體裁、不同主題的簡易文章 (註 1) 8-IV-3 能了解國內外風土民情及主要節慶習俗， 並加以比較 (註 1) 9-III-2 能將事件依故事的發展加以排序 (註 3) 9-IV-2 能把二至三項訊息加以比較、歸類、排序	能理解 字面文 意 (L3)	從上下文，直接推 論出特定字詞、句 子意義或主詞與代 名詞關聯 ----- 直接推論文本字面 大意、重點 ----- 直接推論出或直接 比較具體訊息間關 聯 (例如，特徵相 似或相異性、事件 發生順序等) ----- 直接描述或蒐集文 本具體陳述的訊息、 想法或特徵 ----- 就關鍵字檢索出 (數 位網頁) 結果，直 接推論出適當數位 連結標的

(續下頁)

教育目標 (學習表現；邏輯思考、判斷與創造力)	理解認 知成分	評量指引 (認知成分) 細項
<p>1-IV-7 能辨識簡短說明或敘述的情境及主旨 (註 5)</p> <p>1-V-15 能聽懂不同腔調 / 語言背景英語使用者談話的主要內容 (註 6)</p> <p>3-IV-11 能藉圖畫、標題書名等作合理的猜測</p> <p>3-IV-12 能熟悉重要的閱讀技巧, 如擷取大意、猜測字義、推敲文意、預測後續文意及情節發展 (註 8)</p> <p>3-V-12 能利用字詞結構、上下文意、句型結構及篇章組織推測字詞意義或句子內容 (註 8)</p> <p>3-V-14 能從圖畫、圖示或上下文, 猜測字義或推論文意 (註 8)</p> <p>9-IV-1 能綜合相關資訊作合理的猜測</p> <p>9-IV-3 能根據上下文語境釐清不同訊息間的因果關係</p> <p>9-IV-4 能依上下文所提供的文字線索 (如 in my opinion、maybe) 分辨客觀事實與主觀意見 (註 7)</p> <p>3-V-7 能了解短文、書信的內容及文本結構 (註 6、7)</p>	<p>能理解 隱含文 意 (L4)</p>	<p>統整文本要點以推論出主旨、共通性原則</p> <p>-----</p> <p>比較文本複雜訊息或片段、抽象的想法、概念之相似或相異性</p> <p>-----</p> <p>推論出文本抽象、隱含的意涵或片段訊息間關聯 (例如, 人物特質、因果脈絡、作者目的等)</p> <p>-----</p> <p>解釋文本抽象、隱含的意涵或片段訊息間關聯並從文內提出相對支持證據</p> <p>-----</p> <p>應用或說明文本概念、通則或規則 (至其他情境), 以得出結果</p> <p>-----</p> <p>區辨出文本結構或寫作手法等語文特徵 (例如, 區辨事實與意見、比喻) (註 6、7)</p>

(續下頁)

教育目標 (學習表現；邏輯思考、判斷與創造力)		理解認 知成分	評量指引 (認知成分) 細項
3-IV-15	能分析及判斷文章內容，了解敘述者的觀點、態度及寫作目的	在文本理解上，能透過自身觀點進行評估(L5)	分析、評估文本論點、觀點合理性等，並提出自身立場或觀點說明 ----- 說明作者目的、使用的文本寫作風格、手法或語文特徵、結構(註6、7)
8-IV-4	能了解、尊重不同之文化習俗(註9)		
8-V-3	能從多元文化觀點，了解並尊重不同的文化及習俗(註6、9)		
9-V-4	能分析、歸納多項訊息的共通點或結論(註6)		
9-V-5	能將習得的原則類推到新情境中，解決問題(註6)		
9-V-6	能綜合現有訊息，預測可能的發展(註6)	能省思認識自我(L6)	省思作者論點、觀點或文中寓意 ----- 融合運用各理解歷程成分，進行探究實作以解決問題(註10)
9-V-7	能評估不同資訊，提出合理的判斷或建議(註6)		
9-V-8	能整合資訊，合理規劃並發揮創意完成任務(註6、10)		

註：1. 學生理解認知為一漸近複雜歷程，因此，在轉化教育目標與評量指引對應(如 1-IV-4 代表對應學生具備從流暢運作語言基本知能、能直接找出訊息、到能理解字面文意)時，透過虛線來表徵各層級間並未有明顯區隔，僅能就字面字詞來進行判讀與轉化；此外，若該學習表現或相關指標內容未明顯呈現出能辨識屬較高層次之詞彙(如分析、評估)時，大致以能理解字面文意來作對應，例如，3-IV-16 能閱讀不同體裁、不同主題的簡易文章。最後，若該學習表現或相關指標同時呈現兩個(或以上)可辨識詞彙時，是以較高者為判讀依據，例如，9-V-6 能綜合現有訊息，預測可能的發展。

2. 為力求評量架構完整性與銜接不同學習階段表現，本文也列出其他學習階段指標，然而，各學習階段若有重覆指標內容，本表以列出第四學習階段為原則。
3. 這些指標為銜接國小學習階段。
4. 一般「情緒與態度」多半透過常識性推論即可理解，因此，本書將其對應於字面文意理解。
5. 「大意」是客觀、簡要呈現作者寫作內容，而「主旨」係涉及讀者依文章結構意義關係，經統整、抽出全文主要觀點(馬行誼，2014，頁 55)。因此，本書將大意或重點對應於字面意理解，而主旨對應於隱含意理解。
6. 這些指標為銜接高中階段。
7. 這些指標涉及文本架構、寫作手法等語文特徵，知識層面較高階，即使涉及相對較低層次認知成分(如辨識)，其整體水準也會提高。
8. 指標同時涉及多個認知成分，是以較高層次來對應，例如，猜測字義可能是上下文直接橋接式推論，屬字面意理解，而推論文意可能涉及隱含、推論產生過程，屬隱含意理解。

(續下頁)

9. 尊重一詞隱含著省思、同理等層面，加諸涉及文化習俗知識，因此，本書將其對應於該水準。
10. 學習者得融合運用各理解歷程之複合性能力，以解決問題。實際得透過實作評量，才足以反映學生表現結果，而為求架構完整性，故條列出。
11. 本表部分內容同步刊載於謝進昌（2022a），但是以本書為主要理論與發展的參照來源。

## 附錄二 教育目標（學習內容）與評量指引（相關名詞） 來源與衍伸

教育目標 (學習內容)	文本 類型	延伸名詞 (註 9)
Ab-III-1 子音、母音及其組合 (註 4)	語言知 能文本	字母 (letter)、 音素 (phoneme)、 字母拼讀 (phonics)、 字詞 (彙) (vocabulary)、 片語 (phrase)、 句子 (sentence)、 慣用語 (idiom)、 文法 (grammar) 等
Ad-III-2 簡易、常用的句型結構 (註 3、4)		
Aa-IV-1 連續體大小寫字母的辨識及書寫		
Ab-IV-1 句子的發音、重音及語調		
Ab-IV-2 歌謠、韻文的節奏與音韻		
Ab-IV-3 字母拼讀規則 (含字母拼讀的精 熟能力、字彙拼寫的輔助)		
Ac-IV-1 簡易的英文標示		
Ac-IV-2 常見的教室用語		
Ac-IV-3 常見的生活用語		
Ac-IV-4 國中階段所學字詞 (能聽、讀、 說、寫最基本的 1,200 字詞) (註 3)		
Ac-V-3 高中階段所學字詞 (字頻最高的 4,500 字詞) (註 3、5)		

(續下頁)

教育目標 (學習內容)	文本 類型	延伸名詞 (註 9)
Ae-IV-2 常見的圖表	描述類 篇章	細節 (detail) 、 訊息 (information) 、 脈絡線索 (context clue) 、 字義 (meaning of a word) 、
Ae-IV-3 公共場所廣播 (如捷運、車站、 機場廣播)		特徵 (characteristic) 、
Ae-IV-4 簡易賀卡、書信、電子郵件	教學 指引類 篇章	要點 (main idea) 、 情緒 (emotion) 、
Ae-IV-5 不同體裁、不同主題之簡易文章 (註 1)		態度 (attitude) 、 腔調 (accent) 、
Ab-V-3 不同腔調 / 語言背景英語使用者的 談話 (註 3、5)	人際 互動類 篇章	主旨 (theme) 、 文本架構 (text structure) 、
Ad-IV-1 國中階段所學的文法句型 (註 3)		作者目的 (author's purpose) 、 寫作風格或手法 (author's craft) 、 作者觀點 (author's point of view) 等 (註 9)
Ad-V-1 高中階段所學的結構 (註 3、5)	說明文	論點 (argument) 、 概念 (concept) 、
Ab-V-1 句子語調所表達的情緒和態度 (註 3、5)		證據 (evidence) 、 實例 (example) 、 現象 (phenomenon) 、 因果 (cause and effect) 等 (註 9)
Ae-V-5 教學廣播節目的內容 (註 5)	議論文	
Ae-V-6 學習雜誌、漫畫 (註 5)		
Ae-V-7 新聞報導 (註 5)		
Ae-V-8 工具書 (如百科全書) 或其他線 上資源 (註 5)		
B-V-8 短文、書信的內容及文本結構 (註 5、6)		
C-V-4 國際議題 (如全球暖化、人工智 慧、氣候變遷等) (註 5、8)		

(續下頁)

教育目標 (學習內容)	文本 類型	延伸名詞 (註 9)
Ae-III-2 繪本故事、兒童短劇 (註 4)		
Ae-IV-8 簡易故事及短文的大意		
Ae-IV-1 簡易歌謠、韻文、短文、故事及短劇		人物 (character)、
Ae-IV-7 敘述者的觀點、態度、及寫作目的		事件 (event)、
Ae-IV-6 簡易故事的背景、人物、事件和結局	敘事與文學	行動 (action)、
Ae-V-10 故事及短劇的內容與情節 (註 5)		背景 (setting)、
C-IV-1 國內外節慶習俗 (註 7)		特質 (trait)、
C-IV-2 國內外風土民情 (註 7)		情節 (plot)、
C-IV-3 文化習俗的了解及尊重 (註 7)		寓意 (insight)
C-V-1 多元文化觀點、不同文化及習俗的尊重 (註 5、7)		等 (註 9)

註：1. 本表為就概念層次，轉化教育目標學習內容與評量指引中文本類型對應，透過虛線來表徵教育目標指標間並非有明顯區隔，僅能就字面字詞來進行判讀與轉化，實際仍得依據評量文本內容，才能判讀其類型；此外，若該學習內容未明顯呈現能辨識文本類型之詞彙時，表示其對應於各篇章類型中，例如，Ae-IV-5 不同體裁、不同主題之簡易文章。

2. 為力求評量架構完整性與銜接不同學習階段學習內容，本書也列出其他學習階段指標，然而，各學習階段若有重覆指標內容，本表以列出第四學習階段為原則。
3. 除了國小或為銜接六年級英語文表現，會針對如單一字彙、常用語等進行片段、知識性評量外，其餘高年級學習階段中，這些指標會融入篇章中進行評量，不特別針對單一文法、語音或結構等進行片段、知識性評量，例如，學生得在篇章理解中，了解說者的情緒、態度。
4. 這些指標為銜接國小學習階段。
5. 這些指標為銜接高中階段。
6. 這些指標涉及文本架構、寫作手法等語文特徵，知識層面較高階，即使涉及相對較低層次認知成分（如辨識），其整體水準也會提高。
7. 節慶習俗往往與人物或歷史文化有關聯，故，條列於敘事。
8. 本指標所列議題，大多涉及概念解釋或說服，因此，本書將其條列於說明或議論文。
9. 第一類群文本具備其他文本的基本特徵，因此，將相關名詞統一置放於本類群，例如，訊息、特徵等為一般通用性語詞，而大意、主旨、字義、文本架構、作者目的等，也存在於各類型文本。此外，某些名詞使用擷取與參照自國際各大型閱讀評量對於表現標準描述 (PLD) 內容。

### 附錄三 語言知能文本範例

#### L121

文本編號	溝通活動	文本類型	施測年級	領域	長度（秒數）
L121	聽力	語言知能文本	七	個人	00:01

聽到：ship

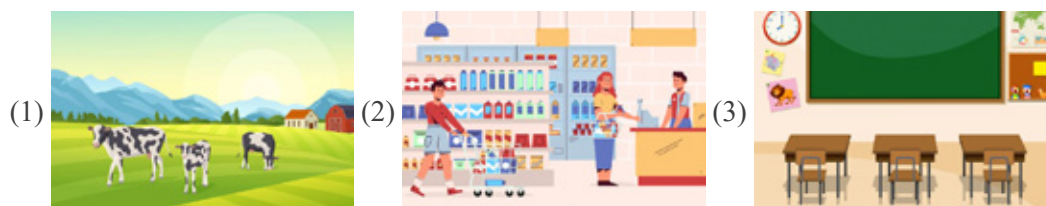
- (1) chip
- (2) drip
- (3) ship

指標	使用字母拼讀規則辨識出字詞							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：-2.8019
	人數	-	26	7	379	-	2	閾值：252.36
選項分析	高分組	-	2	0	183	-	1	Outfit：0.8418
	低分組	-	20	6	114	-	0	Infit：0.9692

#### L120

文本編號	溝通活動	文本類型	施測年級	領域	長度（秒數）
L120	聽力	語言知能文本	七	個人	00:01

聽到：supermarket



指標	聽辨、識讀字彙、片語、常用語、基本句型之句子							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：-3.3752
	人數	-	13	390	10	-	3	閾值：190.95
選項分析	高分組	-	3	174	0	-	0	Outfit：1.0572
	低分組	-	9	100	10	-	3	Infit：0.9328

**L109**

文本編號	溝通活動	文本類型	施測年級	領域	長度 (秒數)
L109	聽力	語言知能文本	七	個人	00:03

聽到：I have three pencil boxes.

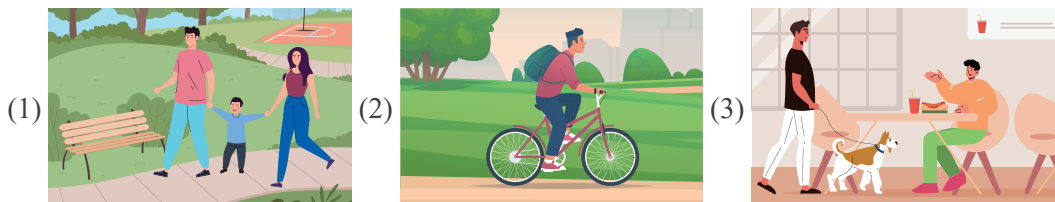


指標	聽辨、識讀字彙、片語、常用語、基本句型之句子							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3*	4	其他	難度：-2.3528
	人數	-	41	8	364	-	3	閾值：300.47
	高分組	-	0	0	177	-	0	Outfit：0.6698
	低分組	-	34	5	80	-	3	Infit：0.9347

**L093**

文本編號	溝通活動	文本類型	施測年級	領域	長度 (秒數)
L093	聽力	語言知能文本	七	個人	00:06

聽到：On weekends, Joe loves going for a walk with his family in the park.



指標	聽辨、識讀字彙、片語、常用語、基本句型之句子							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1*	2	3	4	其他	難度：-2.2870
	人數	-	365	31	19	-	2	閾值：307.51
	高分組	-	150	0	0	-	0	Outfit：0.6890
	低分組	-	105	25	18	-	1	Infit：0.9218

註：本附錄相關內容與範例試題也同步刊載於謝進昌（2022a），但是以本書為主要理論與發展的參照來源，同時，本書數據資料為 2019 年七年級學生表現。

## 附錄四 描述類文本範例

### R071

文本編號	溝通活動	文本類型	施測年級	領域	長度（字數）
R071	閱讀	描述文本	七、八	個人	66

# Cha Cha Hospital

Name: Tommy Chen Date: May 20, 2018

Days: For 3 days

Usage: 1 pill each time after meals

Pills appearance: round, pink

- Morning
- Noon
- Evening
- Bed time

Usage: 1 pill each time

Pills appearance: triangle, yellow

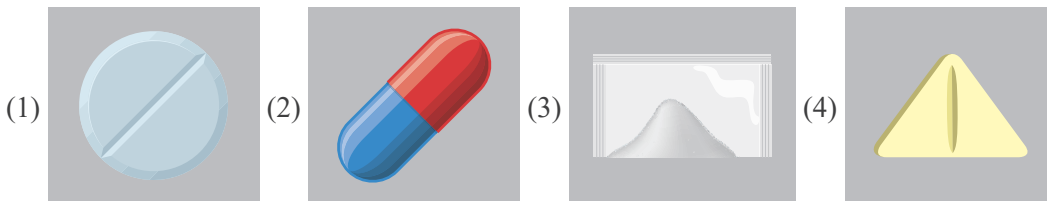
- Morning
- Noon
- Evening
- Bed time

Warning: Do not eat with milk, tea or coffee  
Any questions, please call (02)2266-7304

Address: No. 2, Sanshu Rd., Sanxia Dist., New Taipei City

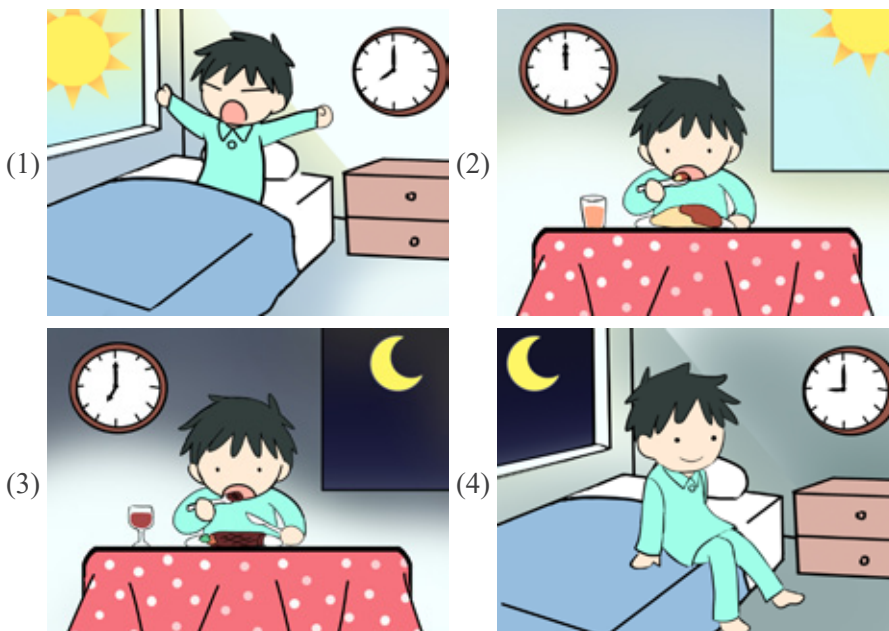
小字典：pill 藥、appearance 外表

1. Which pills does Tommy eat after breakfast for three days?



指標	直接找出人事物、時間、地點、物品等外在顯著特徵	試題參數						
古典理論 (CTT)	選項或得分	0	1*	2	3	4	其他	難度：-0.5136
	人數	-	236	73	38	61	3	閾值：497.46
選項分析	高分組	-	136	4	1	18	0	Outfit：0.8783
	低分組	-	47	54	31	34	1	Infit：0.9147
古典理論 (CTT)	選項或得分	0	1*	2	3	4	其他	
	人數	-	253	72	20	74	2	
選項分析	高分組	-	141	1	2	24	0	Outfit：0.9742
	低分組	-	53	52	15	30	1	Infit：0.9743

2. Which time of the day should Tommy take the yellow pill? 小字典：should 應該



指標	直接找出人事物、時間、地點、物品等外在顯著特徵							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3	4*	其他	難度：-0.8211
	人數	-	63	49	32	257	10	閾值：464.53
	高分組	-	3	4	2	148	2	Outfit：0.8023
	低分組	-	42	32	25	62	6	Infit：0.9079
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3	4*	其他	
	人數	-	59	52	30	277	3	
	高分組	-	4	3	0	161	0	Outfit：0.7542
	低分組	-	44	39	22	44	2	Infit：0.8643

3. From the list, which of the following is true?

- (1) Tommy takes the yellow pills once a day.
- (2) Cha Cha hospital is on Sanxia Road.
- (3) Tommy can take the pills with milk tea.
- (4) Mr. Chen is Tommy's doctor.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1*	2	3	4	其他	難度：-0.1475
	人數	-	211	90	50	57	3	閾值：536.69
	高分組	-	131	16	2	9	1	Outfit：0.9958
	低分組	-	39	56	38	33	1	Infit：0.9812
古典理論 (CTT) 選項分析	選項或得分	0	1*	2	3	4	其他	
	人數	-	233	98	43	45	2	
	高分組	-	141	21	1	5	0	Outfit：0.8731
	低分組	-	39	55	33	23	1	Infit：0.9114

**L153**

文本編號	溝通活動	文本類型	施測年級	領域	長度（秒數）
L153	聽力	描述文本	八	公眾	00:58

聽到：

People don't waste too much time in traffic when going to Juifen now. The number 965 bus is a new way for people to go directly to Juifen. From Banqiao to Juifen only takes 90 minutes! It costs 80 dollars from Banqiao to Juifen one way. The return trip costs 80 dollars as well. Bus 965 has 26 stops for travelers to go more places along the way. On the weekdays, the first bus leaves at 6 o'clock in the morning and the last bus is 9 o'clock at night. On weekends, buses start running from 9 o'clock in the morning, until 9 o'clock at night. Foreigners may even book the seats online. It has Chinese, English, Japanese, and Korean language service.

聽到：

1. What is this talk mainly about?

- (1) The news about a new way to go to Juifen.
- (2) The cheaper bus ticket for foreign visitors.
- (3) Four languages can be heard on Bus 965.

指標	直接推論文本字面大意、重點							試題參數
	選項或得分	0	1*	2	3	4	其他	
古典理論 (CTT)	選項或得分	0	1*	2	3	4	其他	難度：-0.2092
	人數	-	246	90	86	-	6	閾值：530.08
選項分析	高分組	-	102	11	22	-	1	Outfit：1.1104
	低分組	-	64	46	44	-	4	Infit：1.0895

聽到：

2. What can we learn from this talk?

- (1) Bus 965 can get to Juifen in 80 minutes.
- (2) Bus 965 only runs on Saturday and Sunday.
- (3) Banqiao is the first stop on the Bus 965 line.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：0.1402
	人數	-	118	88	209	-	13	閾值：567.50
選項分析	高分組	-	23	9	101	-	3	Outfit：1.0332
	低分組	-	49	55	47	-	7	Infit：1.0379

聽到：

3. Emily comes from the UK. She will have dinner with her Taiwanese friend at 18:30 in Banqiao. She doesn't want be late. When should she take the bus back from Juifen to Banqiao?

- (1) 16:50.
- (2) 17:20.
- (3) 17:40.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1*	2	3	4	其他	難度：0.3792
	人數	-	202	138	82	-	6	閾值：593.10
選項分析	高分組	-	87	27	21	-	1	Outfit：1.1294
	低分組	-	54	65	34	-	5	Infit：1.1022

註：本附錄相關內容與範例試題也同步刊載於謝進昌（2022a），但是以本書為主要理論與發展的參照來源，同時，本書數據資料為 2019 年七年級與 2020 年八年級學生表現。

## 附錄五 教學指引類文本範例

## R065

文本編號	溝通活動	文本類型	施測年級	領域	長度（字數）
R065	閱讀	教學指引文本	七	個人	123

Here are some steps about making paper and cards.

What you will need:

Paper, water, a bowl, a blender, a plate, a screen, some seeds, scissors and markers.

What you will do:

Step 1. Put pieces of used paper into a bowl of water for one night.

Step 2. Put the pieces into the blender. Fill the blender halfway with water. Blend until the paper and water mix like soup.

Step 3. Pour the “paper soup” into the plate.

Step 4. Put the screen into the “paper soup”. Take the screen up, so some “wet paper” will be on the screen.

Step 5. Leave the paper dry for at least 24 hours. Cut the paper in different shapes and color them with markers.

小字典：blender 攪拌機、screen 篩子

1. The following photos are out of order. Please write the correct order of the photos.



（上面圖片非正確順序，請把正確的順序填入下面空白的作答框格內。）

指標	直接描述或蒐集文本具體陳述的訊息、想法或特徵	試題參數						
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3	4	其他	難度：0.5325
	人數	256	148	-	-	-	11	閾值：609.51
	高分組	86	133	-	-	-	2	Outfit：0.8207
	低分組	95	4	-	-	-	5	Infit：0.8696
學生建構題作 答反應（範例）	1分	學生反應能依序正確寫出 BCDAE、bcdae						
	0分	其他不正確次序，如 ACDBE、bdace 等						

2. Kathy's birthday is on October 19. Carol wants to give her a handmade birthday card.

When is the good day for Carol to start to make the card?

- (1) October 17.
- (2) October 18.
- (3) October 19.
- (4) October 20.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
	選項或得分	0	1*	2	3	4	其他	
古典理論 (CTT)	選項或得分	0	1*	2	3	4	其他	難度：0.5180
	人數	-	164	111	85	48	7	闕值：607.96
選項分析	高分組	-	134	51	19	16	1	Outfit：0.8693
	低分組	-	11	30	38	22	3	Infit：0.8931

註：本附錄相關內容與範例試題也同步刊載於謝進昌（2022a），但是以本書為主要理論與發展的參照來源，同時，本書數據資料為 2019 年七年級學生表現。

## 附錄六 人際互動類文本範例

## L132

文本編號	溝通活動	文本類型	施測年級	領域	長度（秒數）
L132	聽力	人際互動文本	八	公眾	01:02

聽到：

M: Good morning. How may I help you, ma'am?

W: I lost my umbrella on the train the other day. I wonder if anyone has picked it up and sent it here.

M: What color is your umbrella?

W: It's blue.

M: Let me check. Hmm... it seems we have a couple of blue umbrellas here. Could you please give me more details about it?

W: Sure. It's dark blue, with some bears and little U.K. flags on it. I bought it in the U.K. just a few weeks ago. It's quite new.

M: Great! I think we've got it here. Just one more question: Do you remember where you lost it?

W: I'm not sure where exactly, but it should be somewhere between Old Town and Main Station, where I travel from home to work every day.

M: Great! It's reported to have been picked up there, too. Please sign your name here and you can have your umbrella back.

聽到：

1. What is the woman doing?

(1) Buying an umbrella in the U.K.

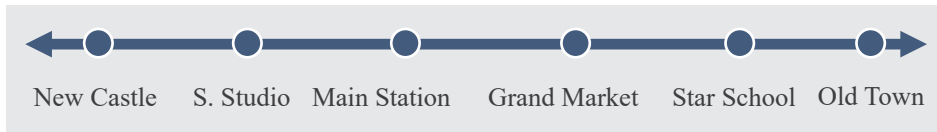
(2) Traveling to work on the train.

(3) Asking for help at Lost and Found.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：-0.1218
	人數	-	80	128	211	-	3	閾值：539.43
選項分析	高分組	-	8	9	129	-	1	Outfit：0.7791
	低分組	-	54	75	29	-	1	Infit：0.8398

聽到：

2. Look at the map. Where did the woman most likely lose her umbrella?



- (1) New Castle.
- (2) Grand Market.
- (3) S. Studio.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
	選項或得分	0	1	2*	3	4	其他	
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：-0.3739
	人數	-	66	106	247	-	3	閾值：512.43
選項分析	高分組	-	6	7	134	-	0	Outfit：0.8886
	低分組	-	36	68	54	-	1	Infit：0.9536

### L134

文本編號	溝通活動	文本類型	施測年級	領域	長度（秒數）
L134	聽力	人際互動文本	八	個人	01:40

聽到：

M: Hello! Fiona, what's that in your hand?

W: Hi, Sam. It's an "INUHARIKO". I'm poor at Japanese. Sorry for my poor pronunciation.

M: It looks like a cat. Is it a cat?

W: No, it's a dog: a dog-like statue. "INU" means "dog" in Japanese.

M: What's that for?

W: In Japan, INUHARIKO is a popular gift for children and pregnant women.

INUHARIKO is a statue in the shape of a dog, and it's usually made of clay, wood, and colored paper. When parents want to pray for good luck, they can buy an INUHARIKO for their kids and ask the temple to read some prayers and give blessing on the "INUHARIKO". Then, they would bring the "INUHARIKO" home and give it to their kids.

M: So the “INUHARIKO” can protect kids and even unborn babies?

W: Yes. The Japanese believe that the “INUHARIKO” will make their kids healthy and strong and keep them away from danger.

M: Then how about the paper fans and coins placed next to the INUHARIKO.

W: Oh, yeah. Paper fans mean that the kid’s future will be wide open, full of chances. It’s just like a fully opened paper fan. And coins mean...(being interrupted)

M: Let me guess, coins mean having money and being rich.

W: Wow. I have to say you are so smart!

M: Thanks! I think so, too.

聽到：

1. What is an “INUHARIKO”?

(1) A dog-like statue for children.

(2) A cat-like statue for rich people.

(3) A kid-like statue for pregnant women.

指標	直接找出人事物、時間、地點、物品等顯著外在特徵							試題參數
古典理論 (CTT)	選項或得分	0	1*	2	3	4	其他	難度：-0.8056
	人數	-	287	61	69	-	4	閾值：466.18
選項分析	高分組	-	142	5	11	-	1	Outfit：0.9947
	低分組	-	74	40	35	-	3	Infit：1.0189

聽到：

2. Why are there often paper fans beside the “INUHARIKO”?

(1) Fans mean joy and happiness.

(2) Fans mean money and good luck.

(3) Fans mean great and hopeful futures.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：0.9622
	人數	-	90	169	152	-	10	閾值：655.54
選項分析	高分組	-	18	55	83	-	3	Outfit：1.3058
	低分組	-	48	72	27	-	5	Infit：1.1935

聽到：

3. Which is true about the dialogue?

- (1) People in love often give an INUHARIKO to each other.
- (2) Parents believe INUHARIKOs can protect their children.
- (3) The man knows a lot about what an INUHARIKO is for.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：-0.2696
	人數	-	75	249	92	-	5	閾值：523.61
選項分析	高分組	-	7	137	13	-	2	Outfit：0.9671
	低分組	-	45	56	48	-	3	Infit：1.0034

## R115

文本編號	溝通活動	文本類型	施測年級	領域	長度（字數）
R115	閱讀	人際互動文本	九	職場	227

Dear Sir/Madame,

Feb. 8, 2019

Hello! Thank you for reading this letter and giving us a chance to save the earth. We are planning to open another meat-free restaurant in Cashmere City this fall. Because of global warming, our world is facing great danger. We believe that eating vegetables will help to reduce CO2 emissions.

We have noticed that almost every supermarket in the city throws away a large number of fruits and vegetables every day. To us, this is a great waste because these fruits and vegetables are still edible even though there might be some black spots on them. Therefore, below is a piece of advice for you:

We would like to buy all the edible fruits and vegetables from your supermarket every day at a lower price. If possible, 40% off would be good. For example, if five apples are sold for \$3 at regular price, we would like to pay \$1.80. This way, these unsold fruits and vegetables will have a second chance to be eaten. We will use these fruits and vegetables to make delicious dishes for our customers. This will be a win-win situation.

If you agree, please contact Mr. Brown or Ms. Hill at 0987-999911 or email us at: eatgreen@qmail.com. We will get back to you as soon as we can.

Best regards,

Betty Wang,

Owner of Eat Green Restaurant

小字典：global warming 全球暖化、emission 排放、edible 可食用的

1. What is the letter mainly about?

- (1) Making delicious dishes.
- (2) Growing apples on farms.
- (3) Asking for a business deal.
- (4) Opening a new supermarket.

指標	直接推論文本字面大意、重點							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	古典通過率：
	人數	-	9	8	29	7	0	0.55
選項分析	高分組	-	1	0	10	1	0	古典鑑別度：
	低分組	-	7	8	12	3	0	0.43

2. Eat Green Restaurant sent an order to Dada Supermarket. Dada Supermarket agrees to give Eat Green Restaurant 40% off regular price. How much can Eat Green Restaurant save?

Fruits/Vegetables	Regular Price	Amount
bananas	\$20/kg	3kg
beans	\$10/kg	4kg
peaches	\$50/kg	2kg

- (1) 200.
- (2) 120.
- (3) 100.
- (4) 80.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3	4*	其他	古典通過率：
	人數	-	7	22	8	15	1	0.28
選項分析	高分組	-	0	6	0	6	0	古典鑑別度：
	低分組	-	4	12	7	6	1	0.30

3. Which is true?

- (1) Mr. Brown works as a clerk at a local supermarket.
- (2) The letter asks people to grow more trees to save the earth.
- (3) Ms. Hill works as a farmer who grows fruits and vegetables.
- (4) Betty Wang plans to open a restaurant in Cashmere City this fall.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3	4*	其他	古典通過率：
	人數	-	10	10	16	17	0	0.32
選項分析	高分組	-	2	0	0	10	0	古典鑑別度：
	低分組	-	7	7	12	4	0	0.70

4. Who will most likely receive the letter? 小字典：likely 最有可能地

- (1) Jenny Chang. She eats vegetables only.
- (2) Mr. Walker. He owns three supermarkets.
- (3) Mrs. Jobs. She grows apples on her farm.
- (4) James Hank. He wants to open a restaurant.

指標	推論出文本抽象、隱含的意涵或片段訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	古典通過率：
	人數	-	13	15	12	13	0	0.28
選項分析	高分組	-	1	9	1	1	0	古典鑑別度：
	低分組	-	8	4	10	8	0	0.62

5. From the reading, why does the author says “ This will be a win-win situation”? Please use examples from the reading to answer the question in your own words.

(為何作者提到「This will be a win-win situation」？請依據文本內容或例子，用你自己的話來解釋。) 請用英文作答。

指標	解釋文本抽象、隱含的意涵或片段訊息間關聯 並從文內提出相對支持證據							試題參數
	選項或得分 人數	0	1	2	3	4	其他	
古典理論 (CTT)	古典平均得分：							
	0.15							
選項分析	古典鑑別度：							
	0.50							

## R142

文本編號	溝通活動	文本類型	施測年級	領域	長度（字數）
R142	閱讀	人際互動文本	九	個人	238

Dear Dad,

After the four-day basketball camp this week, I'm still standing strong, and I still like basketball. This has been a wonderful week so far.

On Monday, they placed me into a group with nine other kids. We got to choose our own team name, the Rocks, because each of us wanted to be as strong as a rock with a lot of power. I met so many kids who love to play basketball, and we became good friends. I'm happy to be part of the team. We also met our coach, Max, a tall, strong guy with glasses. He used to play basketball for HBL. Cool, isn't it?

We kept on practicing basics such as dribbling, passing, shooting, and so on. It was really boring, and all of my teammates thought so, too. However, Coach Max gave us a **pep talk** about how important they were, and said that we would be the best team at the camp if we practiced harder.

Today, we will work on our free throws. I'm super nervous about it although you have been teaching me and practicing with me. I don't know why I still have trouble with free throws, but maybe Coach Max will show me how.

I have to go now. My teammates are calling. I'll see you this Saturday!

Love,

Martin

P.S. Thanks for the stamps you gave me, so I could write this letter to you.

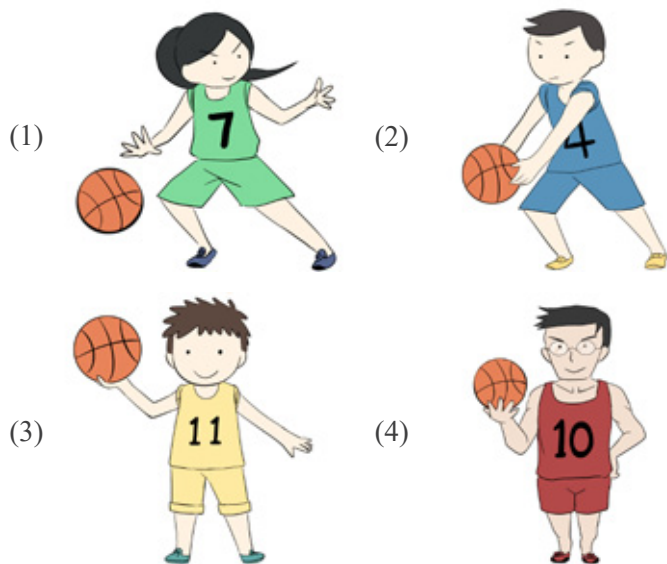
小字典：basics 基本技巧、coach 教練

1. Why is Martin writing to his dad?

- (1) He wants to tell his dad about the basketball camp.
- (2) He needs to ask for some money to spend at the camp.
- (3) One of his friends is in big trouble and needs help.
- (4) He wants to tell his dad that he can't stand it anymore.

指標	直接推論文本字面大意、重點							試題參數
	選項或得分	0	1*	2	3	4	其他	
古典理論 (CTT)	選項或得分	0	1*	2	3	4	其他	古典通過率： 0.67
	人數	-	34	9	5	3	0	
選項分析	高分組	-	16	0	0	0	0	古典鑑別度： 0.64
	低分組	-	8	8	5	1	0	

2. What may Coach Max look like?



指標	直接找出人事物、時間、地點、物品等顯著外在特徵							試題參數
	選項或得分	0	1	2	3	4*	其他	
古典理論 (CTT)	選項或得分	0	1	2	3	4*	其他	古典通過率： 0.69
	人數	-	7	7	2	35	0	
選項分析	高分組	-	0	0	0	16	0	古典鑑別度： 0.55
	低分組	-	5	6	1	10	0	

3. What do we know from this letter?

- (1) Martin's teammates practice free throws very well.
- (2) Martin is not afraid of the difficult basketball training.
- (3) Martin will use the stamps to send letters to his mom.
- (4) Martin did not like basketball before, but now he does.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	古典通過率：
	人數	-	12	24	12	3	0	0.47
選項分析	高分組	-	4	11	1	0	0	古典鑑別度：
	低分組	-	2	8	10	2	0	0.33

4. Which is **NOT** a “**pep talk**”?

- (1) You're a fireball. You may feel tired, but keep trying. You're smart. Don't stop.
- (2) Please shout out loud with me: I'm good. I'm special. I can do anything. I'm No. 1.
- (3) I know it's hard, but just keep going. Don't give up! You'll thank yourself in the end.
- (4) Going to bed and getting up early can help your skin to look healthy. Keep it in mind.

指標	應用或說明文本概念、通則或規則 (至其他情境)，以得出結果							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3	4*	其他	古典通過率：
	人數	-	5	12	11	23	0	0.45
選項分析	高分組	-	0	0	3	13	0	古典鑑別度：
	低分組	-	4	9	7	2	0	0.72

5. What is Kevin's relationship (相處關係) with his father? Explain with examples from the reading by using your own word.

參考句型：Kevin has a \_\_\_\_\_ relationship with his father because  
\_\_\_\_\_. I think \_\_\_\_\_.

指標	解釋文本抽象、隱含的意涵或片段訊息間關聯 並從文內提出相對支持證據							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3	4	其他	古典平均得分：
	人數	19	22	0	-	-	10	0.43
	高分組	1	15	0	-	-	0	古典鑑別度：
	低分組	13	2	0	-	-	7	0.85
學生建構題 作答反應 (範例)	2分	學生能從文中舉出適當例子支持其想法 -Kevin has a great relationship with his father because he would share his dad what he happened in the camp. -Kevin has a good relationship with his father because his father has been teaching him and practicing with him before.						
	1分	學生反應僅能佐證其理解文意，但未能從文中舉出適當例子 -Kevin has a good relationship with his father.						
	0分	其他無關反應或所舉例子不符合文意 -His father is his coach.						

註：本附錄相關內容與範例試題也同步刊載於謝進昌（2022a），但是以本書為主要理論與發展的參照來源，同時，本書數據資料為 2020 年八年級與九年級預試學生表現。

## 附錄七 說明文範例

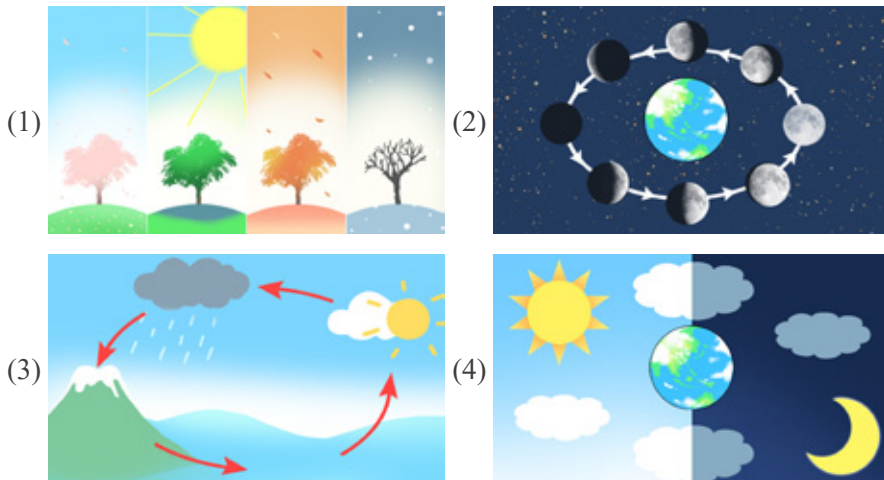
### R079

文本編號	溝通活動	文本類型	施測年級	領域	長度（字數）
R079	閱讀	說明文	七、八	學術	117

The Sun comes up in the morning, and goes down in the evening day by day. We can see the Moon above us changing from the New Moon to the Full Moon. Then the four seasons of a year go one by one, year after year. The things in nature in a special order happen again and again, and we call **these** cycles.

小字典：nature 大自然、order 順序、cycle 循環

1. Which poster is the writer NOT talking about?



指標	直接推論出或直接比較具體訊息間關聯							試題參數
	選項或得分	0	1	2	3*	4	其他	
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：0.3550
	人數	-	73	74	182	87	1	閾值：590.51
選項分析	高分組	-	16	19	123	26	1	Outfit：1.0058
	低分組	-	41	33	21	40	0	Infit：0.9856
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	
	人數	-	66	77	209	63	3	
選項分析	高分組	-	10	18	107	4	2	Outfit：0.9996
	低分組	-	41	33	36	36	0	Infit：0.9966

2. What does **these** mean?

- (1) Days and nights.
- (2) Repeated things in nature.
- (3) Spring, summer, fall, and winter.
- (4) The Sun, the Moon, and the stars.

指標	統整文本要點以推論出主旨、共通性原則							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：0.7228
	人數	-	63	141	121	88	4	闕值：629.91
選項分析	高分組	-	16	90	45	33	1	Outfit：1.2303
	低分組	-	34	30	40	28	3	Infit：1.0794
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	
	人數	-	57	165	124	67	5	
選項分析	高分組	-	16	85	30	7	3	Outfit：1.2051
	低分組	-	19	43	52	31	1	Infit：1.1186

Now, look at us. Our grandparents gave birth to our parents, and then our parents gave birth to us children. Maybe someday we will have our babies. We can also see other living things around us, like chickens. Hens lay eggs and hatch them out into young babies. Then, they turn into adult chickens.

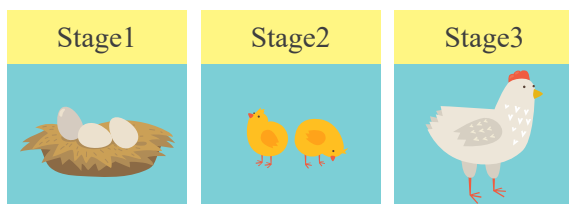
小字典：adult 成年動物、hatch 孵化、lay eggs 產卵

3. What is the writer talking about?

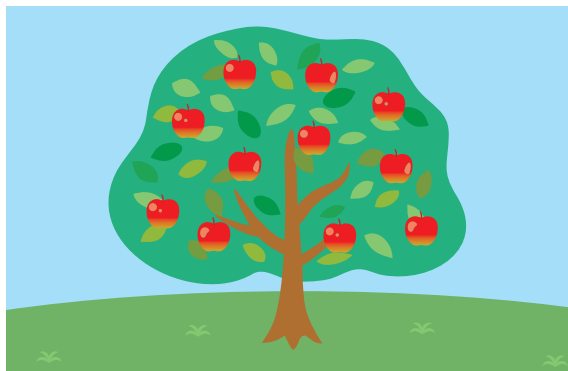
- (1) Living things go through a life cycle as their parents do.
- (2) Living things take good care of their parents and babies.
- (3) Living things have rules between the young and the old.
- (4) Living things believe that their babies follow their parents' rules.

指標	推論出文本抽象、隱含的意涵或片段訊息間關聯	試題參數						
古典理論 (CTT) 選項分析	選項或得分	0	1*	2	3	4	其他	難度：0.8047
	人數	-	141	125	88	59	4	闕值：638.68
	高分組	-	91	42	33	17	2	Outfit：1.1280
	低分組	-	24	52	32	26	1	Infit：1.0326
古典理論 (CTT) 選項分析	選項或得分	0	1*	2	3	4	其他	
	人數	-	155	105	90	61	7	
	高分組	-	86	22	22	8	3	Outfit：1.0066
	低分組	-	28	45	40	32	1	Infit：0.9930

4. Here are three stages of a chicken's life cycle.



Please think about a life cycle of apples. Which stage is the following picture at?



指標	應用或說明文本概念、通則或規則 (至其他情境)，以得出結果							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3	4	其他	難度：0.5961
	人數	184	155	-	-	-	78	闕值：538.81 / 669.57
	高分組	60	111	-	-	-	14	Outfit：0.9118
	低分組	84	9	-	-	-	42	Infit：0.9431
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3	4	其他	
	人數	184	147	-	-	-	87	
	高分組	41	91	-	-	-	9	Outfit：0.9580
	低分組	76	17	-	-	-	53	Infit：0.9340
學生建構題 作答反應 (範例)	1 分	學生反應能佐證其理解文意，例如，說明出正確 stage3 或說明出蘋果成長相關循環概念 - Stage3. - 先是一顆種子，然後成了幼苗，最終成了樹。 - The seed, and grow up, then be an apple.						
	0 分	其他無關反應						

## R126

文本編號	溝通活動	文本類型	施測年級	領域	長度(字數)
R126	閱讀	說明文	八	學術	190

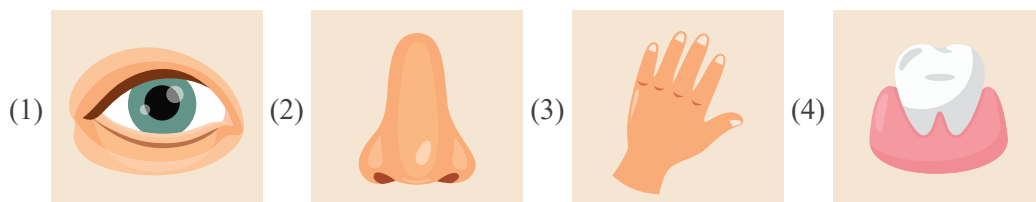
Sometimes when we need help, we may say, “Give me a hand, please.” Or on the weekend, we go to a shopping center and pay thousands of dollars for a new comfortable sofa. The next day, when we go to work, we may tell our friends in the office, “My goodness! The fashionable sofa cost me an arm and a leg.” The two idioms about our body parts are used to talk about the matters in our daily life. Using idioms like these can be fun in some ways.

Here are more examples below. Last Sunday, I went to a movie with friends. The movie had such a sad ending that I cried my eyes out. “Cry someone’s eyes out” means crying very hard. It also speaks out my true feelings. Another example is that my cousin Kelly loves everything sweet, from chocolate, ice cream, doughnuts to cakes. People around Kelly always say she has a sweet tooth.

Aren't **they** interesting? If we read the idioms more and learn them more, we can use them more easily. Then, we will be an old hand at using English in our lives.

小字典：based on 以為基礎、idiom 成語；慣用語

1. Which body part does the writer **NOT** talk about in the reading?



指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：0.1568
	人數	-	59	205	61	99	8	閾值：569.29
選項分析	高分組	-	5	137	3	14	2	Outfit：0.8240
	低分組	-	42	22	40	50	5	Infit：0.8702

2. What does “**they**” mean in the reading?

- (1) Body parts.
- (2) Sweet things.
- (3) These idioms.
- (4) Kelly’s family and friends.

指標	推論出文本抽象、隱含的意涵或片段訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：0.7028
	人數	-	118	92	154	59	9	閾值：627.75
選項分析	高分組	-	46	11	96	5	3	Outfit：1.0033
	低分組	-	29	57	28	40	5	Infit：0.9764

3. Jim has a rich experience of fixing cars and does his job well. We may say,

“Jim \_\_\_\_\_.”

- (1) is an old hand
- (2) gives me a hand
- (3) cries his eyes out
- (4) costs me an arm and a leg

指標	應用或說明文本概念、通則或規則 (至其他情境)，以得出結果							試題參數
古典理論 (CTT)	選項或得分	0	1*	2	3	4	其他	難度：0.9247
	人數	-	158	114	76	73	11	闕值：651.52
選項分析	高分組	-	102	29	7	20	3	Outfit：1.0035
	低分組	-	24	56	42	31	6	Infit：0.9739

4. What is the reading mainly about? (Please give a best title for this reading.)

指標	說明作者目的、使用的文本寫作風格、 手法或語文特徵、結構							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3	4	其他	難度：3.2643
	人數	263	35	8	-	-	126	闕值：832.75 / 943.66
選項分析	高分組	112	32	8	-	-	9	Outfit：0.5959
	低分組	80	0	0	-	-	79	Infit：0.9792
學生建構題 作答反應 (範例)	2分	學生反應符合文意與標題所需相關語言特徵 - I think the best title for the reading is “Interesting Idioms About Body Parts”. - Fun body parts in learning idioms.						
	1分	學生反應僅符合文意 - It is talk about body parts idiom. - Our body parts can be the idioms, and they have a lot of meaning.						
	0分	其他無關反應 - Interesting body part - Our body						

## L128

文本編號	溝通活動	文本類型	施測年級	領域	長度（秒數）
L128	聽力	說明文	八	學術	01:16

影片：<https://youtu.be/nLAzsJry7AM>

場景	腳本
1.	Sugar is sweet to taste, easy to get, and also cheap to buy, so most people cannot say no to it.
2.	But sugar is not really good for us if we eat too much.
3.	For example, it is bad for our teeth.
4.	Also, many people are out of shape because they put weight or get heavy.
5.	What's worse, sugar can even give lines on our faces, and makes us look older.
6.	Now let's take a look at this.
7.	Here is a can of soda, and there are 65 grams sugar in it.
8.	You may think 65 grams of sugar isn't that much.
9.	In fact, the WHO points out that we only need about 25 grams of sugar every day.
10.	So, 65 grams of sugar is two times more than we need.
11.	Drinking water without sugar is the best way to keep us healthy. Let's give it a try and start your healthy life.

1. What is the video mainly about?

- (1) It is important to take care of our teeth.
- (2) It is popular to drink soda every day.
- (3) It is bad to eat too much sugar.

指標	直接推論文本字面大意、重點							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：-0.9264
	人數	-	73	59	294	-	2	閾值：453.25
選項分析	高分組	-	4	0	131	-	1	Outfit：0.7233
	低分組	-	50	49	59	-	0	Infit：0.8494

2. Tina agrees with the speaker. What kind of drink may Tina have every day?

- (1) Chocolate milk.
- (2) Apple juice.
- (3) Iced water.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：-0.8099
	人數	-	70	71	278	-	9	閾值：465.73
選項分析	高分組	-	8	3	124	-	1	Outfit：0.8969
	低分組	-	39	52	63	-	4	Infit：0.9221

3. Joy wants to buy some drink. From the video, which of the following will he pick?

	Nutrition Facts		Nutrition Facts		Nutrition Facts
(1)	Fat ..... 0.8 g	(2)	Fat ..... 0.7 g	(3)	Fat ..... 0.85 g
	Sugar ..... 12.0 g		Sugar ..... 8.0 g		Sugar ..... 10.0 g
	Vitamin A ..... 0.7 g		Vitamin A ..... 0.7 g		Vitamin A ..... 0.65 g
	Protein ..... 7.0 g		Protein ..... 7.1 g		Protein ..... 6.9 g

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：-0.4234
	人數	-	67	236	121	-	4	閾值：507.13
選項分析	高分組	-	10	112	13	-	1	Outfit：0.9093
	低分組	-	30	50	76	-	2	Infit：0.9456

註：本附錄相關內容與範例試題也同步刊載於謝進昌（2022a），但是以本書為主要理論與發展的參照來源，同時，本書數據資料為 2019 年七年級與 2020 年八年級學生表現。

## 附錄八 議論文範例

## R018

文本編號	溝通活動	文本類型	施測年級	領域	長度（字數）
R018	閱讀	議論文	七、八	個人	158



小字典：makeup 化妝；化妝品、lipstick 口紅、mirror 鏡子

1. What are Patty and Amanda talking about?

- (1) Doing makeup is important.
- (2) Telling stories about makeup.
- (3) Sharing ideas about makeup.
- (4) Changing our lives with makeup.

指標	統整文本要點以推論出主旨、共通性原則							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3*	4	其他	難度：0.7921
	人數	-	100	111	143	62	5	閾值：637.32
	高分組	-	21	21	65	15	1	Outfit：1.2534
	低分組	-	51	65	48	30	4	Infit：1.1403
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3*	4	其他	
	人數	-	89	85	172	66	5	
	高分組	-	13	18	85	12	2	Outfit：1.1464
	低分組	-	41	44	38	36	2	Infit：1.1084

2. Amanda says, “It lights up my face and makes me feel good.” What does “It” mean?

- (1) Looking terrible.
- (2) Putting on makeup.
- (3) Going to bed too late.
- (4) Looking into the mirror.

指標	從上下文，直接推論出特定字詞、 句子意義或主詞與代名詞關聯							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1	2*	3	4	其他	難度：-0.2274
	人數	-	63	207	64	77	10	閾值：528.12
	高分組	-	4	100	3	14	2	Outfit：0.8855
	低分組	-	47	52	45	47	7	Infit：0.9251
古典理論 (CTT) 選項分析	選項或得分	0	1	2*	3	4	其他	
	人數	-	45	267	44	53	8	
	高分組	-	1	123	1	3	2	Outfit：0.8631
	低分組	-	33	59	34	32	3	Infit：0.9006

3. What do you think about doing makeup? Use Patty’s or Amanda’s idea to write down what you think.

(請盡量以英文回答，完成下列句子)

參考句：I agree with \_\_\_\_\_ (Patty or Amanda). I think \_\_\_\_\_.

指標	分析、評估文本論點、觀點合理性等， 並提出自身立場或觀點說明							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3	4	其他	難度：3.2614
	人數	247	90	2	-	-	82	閾值：684.02 / 1117.75
	高分組	37	79	2	-	-	3	Outfit：0.5865
	低分組	132	2	0	-	-	64	Infit：0.7210
古典理論 (CTT) 選項分析	選項或得分	0	1	2	3	4	其他	
	人數	184	162	4	-	-	67	
	高分組	19	104	3	-	-	4	Outfit：1.3717
	低分組	101	12	0	-	-	48	Infit：1.5845
學生建構題 作答反應 (範例)	2分	學生所選擇立場與其作答反應一致，同時能分析，提出自身想法 - I agree with Patty. I think makeup is not a good thing. Although it can help us become pretty, but when you are at school, learning is important. - I agree with Amanda. I think doing a little makeup does help people look better, but it is always not a good choice to do a things too many times.						
	1分	學生所選擇立場與其作答反應一致 - Patty, because if you are healthy, you will look pretty and nice, so you didn' t have to makeup. - I agree with Amanda. I think sometime when you look clean and good, other people will think you are a good guy.						
	0分	學生所選擇立場與其作答反應不一致、其他無關反應 - Amanda, I think her idea is good. - I agree with Patty. I think her makeup is very beautiful.						

註：本附錄相關內容與範例試題也同步刊載於謝進昌（2022a），但是以本書為主要理論與發展的參照來源，同時，本書數據資料為 2019 年七年級與 2020 年八年級學生表現。

## 附錄九 敘事與文學範例

### L086

文本編號	溝通活動	文本類型	施測年級	領域	長度（秒數）
L086	聽力（視聽）	敘事與文學	七	學術	02:37

影片：<https://youtu.be/wKDesMokcJI>

腳本：

Children can't wait for the winter vacation to start! There are many holidays to celebrate in the winter.

#### **Christmas**

Christmas is celebrated on December 25th. It honors the birth of Jesus Christ. People decorate Christmas trees, or sing Christmas carols. Family and friends often exchange presents. Kids still believe Santa Claus will bring them gifts.

#### **Hanukkah**

Hanukkah is a Jewish holiday. People usually celebrate it near the end of November or in December. Thousands of years ago, a ruler asked soldiers to attack a Jewish temple. The Jews got together to fight back and they saved the temple. They wanted to keep the temple bright, so they used candles and the candles burned for eight days. That is why Hanukkah is eight days long, and people light up a candle each night to honor what happened. On Hanukkah, people give chocolate coins or money to children.

#### **Kwanzaa**

Kwanzaa is celebrated from December 26th through January 1st. It is a 7-day celebration. The holiday honors African culture and history. Each of the seven days of Kwanzaa focuses on a different idea. For example, there's a day for people to stay together with family and friends. There's a day for people to work together for the whole neighborhood. They also put out books and African items to help children learn about African culture. Family members often make gifts and give them to each other. The last day of Kwanzaa is always on January 1st.

#### **New Year**

People all over the world celebrate New Year in different ways. They also do different things for good luck on that day. In Mexico and some countries in South America, people

wear yellow or red underwear on New Year's Eve for good luck. In Japan, it's good luck to watch the first sunrise of the year. In Taiwan, parents give children red envelopes for good luck in Chinese New Year. Now, you have a list of holidays to celebrate in your winter vacation. Have a good time!

1. What do Christmas and Kwanzaa have in common?

- (1) Their stories are both about God.
- (2) The holidays end on the first day of a year.
- (3) People send gifts to their families and friends.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：-0.0950
	人數	-	72	121	220	-	1	閾值：542.30
選項分析	高分組	-	34	27	124	-	1	Outfit：1.1483
	低分組	-	27	66	47	-	0	Infit：1.1092

2. Why people celebrate Hanukkah?

- (1) People need to do eight things during the holidays.
- (2) People want to show that they believe in eight gods.
- (3) People want to remember something happened in old days.

指標	推論出文本抽象、隱含的意涵或片段訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：0.5981
	人數	-	124	105	162	-	23	閾值：616.53
選項分析	高分組	-	53	37	89	-	7	Outfit：1.1688
	低分組	-	41	52	35	-	12	Infit：1.1162

3. What is the similar part between Chinese New Year and Hanukkah?

- (1) The holidays last eight days.
- (2) People give money to children.
- (3) Their stories are both about monsters.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：0.0783
	人數	-	131	194	78	-	11	閾值：560.85
選項分析	高分組	-	52	112	19	-	3	Outfit：1.2531
	低分組	-	47	46	40	-	7	Infit：1.1766

4. Which is true about the new year celebrations in South America?

- (1) People light up candles at home for good luck.
- (2) People get up early to watch sunrise for good luck.
- (3) People wear yellow or red underwear for good luck.

指標	直接找出與特定目標有關的明顯訊息							試題參數
古典理論 (CTT)	選項或得分	0	1	2	3*	4	其他	難度：-0.0243
	人數	-	105	101	203	-	5	閾值：549.87
選項分析	高分組	-	27	25	133	-	1	Outfit：0.9599
	低分組	-	45	48	43	-	4	Infit：0.9720

## L139

文本編號	溝通活動	文本類型	施測年級	領域	長度（秒數）
L139	聽力	敘事與文學	八	學術	01:01

聽到：

Which fruit is your favorite? When it comes to summer fruits, many people may think of mango. Mango was first found in India. The history of the mango goes back more than 4,000 years ago when Indians first started to grow it. The word “mango” comes from one of the official Indian languages. In the 16th century, the Portuguese brought mango to countries all over the world. However, mango was brought to Taiwan by Dutch people. Now, one popular kind of mango in Taiwan is called “Aiwen”. Aiwen mango is mostly grown in Tainan. There are many mango products, such as mango ice cream, mango cakes, dried mango, and so on. People in Taiwan, not only Taiwanese but also foreigners, all love Aiwen mango.

聽到：

1. What is the main purpose of this talk?

- (1) To help sell more Aiwon mango in Taiwan.
- (2) To let people know more about mango.
- (3) To introduce different mango products.

指標	推論出文本抽象、隱含的意涵或片段訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：-0.2828
	人數	-	110	239	53	-	1	閾值：522.19
選項分析	高分組	-	14	129	10	-	1	Outfit：0.9545
	低分組	-	59	47	16	-	0	Infit：0.9900

聽到：

2. What do we know from the talk?

- (1) Dutch people are good at growing mango.
- (2) The word “mango” comes from India.
- (3) Sailors from Portugal brought mango to Taiwan.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：-0.1374
	人數	-	68	221	102	-	11	閾值：537.77
選項分析	高分組	-	12	119	22	-	1	Outfit：1.0442
	低分組	-	33	37	45	-	7	Infit：1.0355

## R135

文本編號	溝通活動	文本類型	年級	領域	長度（字數）
R135	閱讀	敘事與文學	八	個人	250

One spring morning, the mother pig on Mr. Arable’s farm delivers a litter of piglets, a group of baby pigs. Mr. Arable finds one of them a runt, which means the smallest and weakest pig, so he decides to kill it. After his daughter, Fern, asks him to let her keep it, she names it Wilbur.

Wilbur is sold to Mr. Zuckerman after a while. Although Fern tries to visit Wilbur in his **pen** very often, Wilbur gets lonelier each day. Finally, a warm and kind voice tells him that she'll be his friend. The next day, he meets Charlotte, a grey spider. One day, an old sheep tells Wilbur that he'll be killed for his meat, but Charlotte promises to save him. As summer passes, Charlotte comes up with **a good plan**. She weaves words into her web, thinking that Mr. Zuckerman won't kill a famous pig. When all the animals are sleeping, she begins working hard. The next morning, people see the words "SOME PIG" in her web; they think there must be something special about Wilbur. Later, more and more people visit the farm. Then Charlotte writes another word "TERRIFIC," meaning "great", in her web. People get excited about this wonderful surprise.

More surprisingly, at the county fair, Charlotte writes another word "HUMBLE," meaning "not proud". This makes Wilbur more and more famous, so Mr. Zuckerman decides to keep Wilbur on his farm forever.

Thanks to Charlotte the spider, Wilbur has a chance to continue his life...

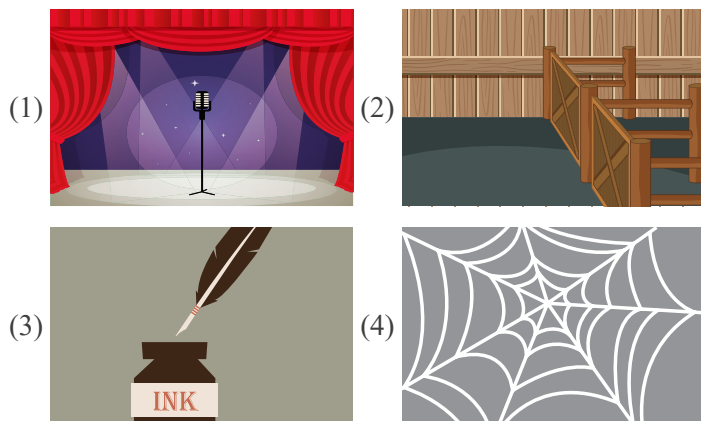
小字典：deliver 生出；分娩、weave 編織、county fair 縣或郡的市集

1. Why does Mr. Arable want to kill Wilbur?

- (1) It is winter.
- (2) He is a runt.
- (3) He is very sick.
- (4) He eats too much.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
	選項或得分	0	1	2*	3	4	其他	
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：0.3132
	人數	-	75	181	112	62	9	闕值：586.02
選項分析	高分組	-	5	90	29	11	3	Outfit：1.2596
	低分組	-	51	41	41	29	3	Infit：1.1525

2. What is a **pen** in this story?



指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1	2*	3	4	其他	難度：0.4268
	人數	-	49	174	135	71	10	閾值：598.19
選項分析	高分組	-	4	87	26	17	4	Outfit：1.0499
	低分組	-	30	38	62	32	3	Infit：1.0331

3. Why does Charlotte start weaving words into her web?

- (1) To save Wilbur.
- (2) To surprise people.
- (3) To keep herself busy.
- (4) To teach Mr. Zuckerman.

指標	直接推論出或直接比較具體訊息間關聯							試題參數
古典理論 (CTT)	選項或得分	0	1*	2	3	4	其他	難度：0.2104
	人數	-	192	90	80	66	11	閾值：575.01
選項分析	高分組	-	109	12	6	7	4	Outfit：0.9116
	低分組	-	28	57	42	34	4	Infit：0.9290

4. What lesson can you learn from the story?

參考句：I can learn that \_\_\_\_\_.

指標	省思作者論點、觀點或文中寓意							試題參數
古典理論 (CTT) 選項分析	選項或得分	0	1*	2	3	4	其他	難度：2.7264
	人數	246	48	-	-	-	145	闕值：775.87 / 884.96
	高分組	78	41	-	-	-	19	Outfit：0.6392
	低分組	96	1	-	-	-	68	Infit：0.8160
學生建構題 作答反應 (範例)	2分	學生反應能突顯出其理解隱含文意，同時能從生活經驗中舉例呼應省思 - 暫無						
	1分	學生反應能突顯出其理解隱含文意 - I can learn that friends are important in life. - Friendship.						
	0分	其他無關反應 - Don't kill the pig. - I can learn that people need to be nice.						
備註	九年級時，本試題題幹改為： Q4-1: What lesson do you learn from the story? Q4-2: Please give an example from your life that can show the lesson you learn from the story. (請以你的個人生活經驗作為一個例子，來呼應你從故事中所學。) 請盡量以英文作答。							

註：本附錄相關內容與範例試題也同步刊載於謝進昌(2022a)，但是以本書為主要理論與發展的參照來源，同時，本書數據資料為2019年七年級與2020年八年級學生表現。

## 附錄十 英語文學習動機試題題項與適配度

英語文學習動機題項與適配度（以七年級學生作答資料分析）

向度	題號	元素	題項內容	Outfit	Infit
內在動機	P1	樂於參與	我喜歡學習英語文	0.7580	0.8044
	P2	樂於參與	我覺得英語文相關的學習活動很有趣	0.6283	0.7079
	P3	樂於參與	上英語文課程讓我感到很愉快	0.9630	0.9635
	P4	樂於參與	我喜歡接觸英語文相關的事物或活動	0.5930	0.6649
	P5	沉浸學習	我很享受英語文課程的學習	0.6247	0.6955
	P6	好奇 / 尋求新經驗	我喜歡透過英語文去學習有趣或新奇的事物	0.7026	0.7674
	P7	好奇 / 尋求新經驗	我喜歡了解有趣或新奇的異國文化	2.3117	2.1612
自我概念	E1	過去經驗	我在英語文方面表現不錯	0.8431	0.8847
	E2	自我肯定	我可以解決英語文學習上的難題	0.8540	0.9446
	E3	自我肯定	我可以理解英語文法句型與使用方式	1.0155	1.1043
	E4	參照比較	和班上許多同學比起來，英語文對我來說是比較容易的	0.7980	0.8714
	E5	參照比較	和其他科目比起來，我覺得學英語文比較容易	0.8262	0.8845
	E6	自我肯定	英語文是我厲害的科目之一	0.8979	0.9251
	E7	他人肯定	老師或同學覺得我的英語文能力很好	1.0433	1.0412

(續下頁)

向度	題號	元素	題項內容	Outfit	Infit
社會 動機	A1	同儕	我喜歡和朋友一起學習英語文	0.8586	0.9045
	A2	同儕	我喜歡和朋友交換一些英語文學習的訊息	0.7285	0.7871
	A3	同儕	我會和朋友談論我目前正在用英語文學習的內容	0.7681	0.8439
	A4	同儕	我喜歡和朋友討論英語文的相關作業	0.9780	1.0218
	A5	社群	我希望透過英語文學習增加與外國朋友的交流機會	0.9673	1.0338
	A6	師生	我喜歡向老師或朋友詢問英語文相關問題	0.8630	0.9163
	A7	社群	我希望透過英語文學習增加在社群網站（例如，臉書、部落格）中與朋友互動的機會	1.2421	1.2519

註：1. 原始選項為①非常同意、②同意、③不同意、④非常不同意。

2. 適配度值源自於七年級學生作答反應。

## 附錄十一 英語文策略使用試題題項與適配度

英語文策略使用問卷題項與適配度（以七年級學生作答資料分析）

向度	題號	元素	題項內容	Outfit	Infit
記憶策略	M1	圖像化	在學英文單字時，我會試圖把單字的發音和單字代表的圖像連結起來（例如，念 refrigerator（冰箱）時，腦袋會聯想到冰箱的樣子）	0.9458	0.9728
	M2	聲韻	我會用相似的發音（例如，old, cold, hold 都有 o 的長母音）來記憶新的英文單字	0.9391	0.9654
	M3	脈絡化	我會將剛剛新學到的英文單字，轉換為自己熟悉或想像的場景，以幫助記憶	0.7928	0.8495
	M4	搭配詞 / 同義詞	在背單字時，我會將意思有關聯的英文單字都放在一起以幫助記憶（例如，good 與 excellent 都表示「好」的意思）	0.7820	0.8303
	M5	字根 / 首 / 尾	當碰到不熟悉的英文單字時，我會用字首、字根、字尾的概念，把英文單字拆成幾個部分再背起來（例如，player 是 play 與 er 組成）	0.9340	0.9755
	M6	音節	當碰到不熟悉的英文單字時，我會根據音節將英文單字拆成幾個部分再背起來（例如，sunny 會分成 sun-ny 兩個音節）	1.0371	1.0712

（續下頁）

向度	題號	元素	題項內容	Outfit	Infit
認知策略	C1	快速瀏覽	讀英文時，我會先大致讀過整篇文章，了解文章的大意後再仔細閱讀	1.1232	1.0766
	C2	重點標示 / 畫底線	上英文課時，我會將上課聽到或讀到的內容重點標示出來（例如，畫底線）	1.2945	1.2726
	C3	摘取大意	我會把英文文章中重要的內容或訊息簡短地記錄下來	0.9983	1.0245
	C4	文法規則	讀英文時，我會試著找出英文片語或句型的用法與規則來幫助我理解文章大意	0.7952	0.8474
	C5	語調 / 氣	在聽別人說英語時，我會利用不同的方式（例如，語調、或停頓語氣），來了解對方所說的內容	0.8806	0.9155
	C6	文本架構	讀英文時，我會先分析文章內容的架構（例如，文章長度、主要人物）後，再仔細閱讀	0.8345	0.8685

(續下頁)

向度	題號	元素	題項內容	Outfit	Infit
推論 連結 策略	I1	上下文脈絡	我會從上下文或整段內容來理解文章在說甚麼	0.8454	0.8579
	I2	圖表 / 標題	我會從英文文章中插圖、表格或標題來推測文章大意	1.1030	1.0947
	I3	上下文脈絡	用英語交談或閱讀時，我會利用上下文已經知道的字（或其他線索），來推測新字或詞的意思	0.7815	0.8231
	I4	大意理解	讀英文時，我會先了解大意而不會碰到每個不懂的單字都去查字典	1.2806	1.2793
	I5	關鍵字詞	進行英語交談時，我會根據聽到的一些關鍵字詞來了解大意	0.8089	0.8447
	I6	肢體動作 / 表情	進行英語交談時，我會注意對方的臉部表情或肢體語言（例如，手勢）來幫助了解他想表達的訊息	1.1562	1.1640
	I7	背景知識 / 經驗	在聽或讀英文時，如果碰到不懂的內容時，我會試著從自己過去的經驗與知識來推測	0.7855	0.8439
	I8	概念關係 連結	我會來回閱讀文章內容，試著找出文章裡不同概念之間的關係	0.8531	0.9072

(續下頁)

向度	題號	元素	題項內容	Outfit	Infit
理解 監控 策略	S1	監控	我會留意自己英語文學習的進步情況	0.9940	0.9910
	S2	調整	我會記下自己常犯的英語文錯誤並改進，以幫助學習	0.8420	0.8648
	S3	檢驗	我會試著用自己理解的方式，重新檢查所聽或讀到的英語文內容，以幫助我更加了解	0.7242	0.7723
	S4	檢驗	如果我找到的資訊對於同一件事情有不同的說法（例如，有人說吃蛋會導致高血壓，有些人則說兩者沒有關係），我會再次確認這些資訊的來源是否正確	1.2496	1.2319
	S5	檢驗	當碰到英文內容與我知道的不一致時，我會再次確認內容是否正確	0.7686	0.8208
	S6	監控	學習英語文時，我會一邊讀一邊確認自己哪裡不懂	0.7245	0.7811
	S7	監控	在聽和讀英語文時，當我發現自己不專心時，我會再集中注意力來聽和閱讀內容	0.9796	1.0192
	S8	調整	當碰到英語文相關問題時，我會發展出自己理解或解決問題的方法	0.7440	0.8056
	S9	調整	學習英語文遇到不懂的地方時，我會想辦法去解決	0.7709	0.8299
	S10	評估	我會小心地評估與選擇我找到的相關英文素材（例如，網路文章、雜誌），才進行閱讀	1.2248	1.2617

註：1. 引自謝進昌（2021b）。

2. 原始選項為①常常如此、②有時候、③很少如此、④從來沒有。

## 附錄十二 正式評量試題參數與適配度（七、八年級）

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
C005_2_1	E01	-0.0951	--	--	--	1.2657	1.1617	0.5290	0.3329	--	--	--	--
C005_2_2	E01	-0.1448	--	--	--	0.9370	0.9561	0.4903	0.5044	--	--	--	--
L086_1	E01	-0.0950	--	--	--	1.1483	1.1092	0.5314	0.3310	--	--	--	--
L086_2	E01	0.5981	--	--	--	1.1688	1.1162	0.3913	0.2285	--	--	--	--
L086_3	E01	0.0782	--	--	--	1.2531	1.1766	0.4686	0.2736	--	--	--	--
L086_4	E01	-0.0243	--	--	--	0.9599	0.9720	0.4903	0.4079	--	--	--	--
L090	E01	-0.6556	--	--	--	0.8536	0.9193	0.5725	0.6174	--	--	--	--
L121	E01	-2.8019	--	--	--	0.8418	0.9692	0.9155	0.1696	--	--	--	--
R059_1	E01	0.7349	--	--	--	1.0025	0.9621	0.3585	0.4500	--	--	--	--
R059_2	E01	-0.2372	--	--	--	0.9748	0.9863	0.5354	0.4404	--	--	--	--
R059_3	E01	0.0417	--	--	--	0.9437	0.9429	0.5094	0.4567	--	--	--	--
R059_4	E01	0.5243	--	--	--	1.0436	1.0228	0.3892	0.3895	--	--	--	--
R084_1	E01	0.0456	--	--	--	0.9316	0.9438	0.4693	0.5420	--	--	--	--
R084_2	E01	-2.2501	--	--	--	1.1460	1.0455	0.8703	0.1816	--	--	--	--
C001_1	E02	0.1975	--	--	--	0.9882	0.9889	0.4504	0.4287	0.8634	0.8890	0.5319	0.6886
C001_2	E02	-0.6761	--	--	--	0.9656	0.9798	0.5908	0.5494	0.8205	0.9073	0.6324	0.6118
C001_3	E02	-0.4158	--	--	--	1.2419	1.1338	0.5690	0.3007	1.3903	1.2110	0.5564	0.2765
C004_1	E02	0.7561	--	--	--	0.9812	0.9694	0.3414	0.3996	1.0266	1.0032	0.4755	0.5793
C004_2	E02	0.5391	--	--	--	0.9899	0.9724	0.3729	0.4218	0.8435	0.8608	0.4583	0.6150
C004_3	E02	-0.0604	--	--	--	0.8451	0.8872	0.4988	0.5920	0.7863	0.8422	0.5882	0.6927
C004_4	E02	-0.1174	--	--	--	1.0049	1.0231	0.5375	0.4038	0.9996	1.0412	0.5784	0.4664
C004_5	E02	0.1185	--	--	--	1.0177	1.0143	0.4576	0.4461	1.1018	1.0551	0.5172	0.4203
L022_1	E02	-0.6086	--	--	--	0.8849	0.9372	0.6223	0.5389	0.7644	0.8623	0.7157	0.5631
L022_2	E02	-0.7494	--	--	--	1.3662	1.1508	0.6441	0.2676	1.4954	1.2798	0.6201	0.2462
R069	E02	0.7606	--	--	--	1.1294	1.0388	0.3428	0.2502	1.1794	1.0912	0.4069	0.3367
R090_1	E02	0.4540	--	--	--	0.8858	0.9084	0.3759	0.4951	0.7639	0.7983	0.4779	0.7667
R090_2	E02	0.2910	--	--	--	1.2449	1.1830	0.4397	0.2394	1.2077	1.1658	0.5000	0.3511
R090_5	E02	1.2026	-0.6680	0.6680	--	0.6602	0.7598	0.2695	0.5622	0.5704	0.6122	0.3407	0.6767

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
C005_1_2	E03	-0.5294	--	--	--	0.8397	0.9025	0.5962	0.6016	--	--	--	--
C005_1_3	E03	-0.5373	--	--	--	0.8027	0.8715	0.6221	0.6598	--	--	--	--
L015_1	E03	-0.3232	--	--	--	0.8312	0.8798	0.5673	0.6123	--	--	--	--
L015_2	E03	-1.5435	--	--	--	0.9600	0.9705	0.7837	0.4158	--	--	--	--
L015_3	E03	-1.0403	--	--	--	0.8904	0.9515	0.7067	0.4777	--	--	--	--
L094	E03	-1.0555	--	--	--	0.7747	0.8874	0.6755	0.6408	--	--	--	--
L109	E03	-2.3528	--	--	--	0.6698	0.9347	0.8750	0.3443	--	--	--	--
L114	E03	-0.4627	--	--	--	1.3765	1.2535	0.5962	0.1384	--	--	--	--
L120	E03	-3.3752	--	--	--	1.0572	0.9328	0.9375	0.1634	--	--	--	--
L122	E03	-3.4805	--	--	--	0.6561	0.8429	0.9495	0.1419	--	--	--	--
R066_1	E03	0.5476	--	--	--	0.8781	0.8868	0.3756	0.4579	--	--	--	--
R066_2	E03	0.0092	--	--	--	0.9252	0.9448	0.4695	0.4519	--	--	--	--
R066_3	E03	1.6444	--	--	--	1.4898	1.2362	0.2066	0.0765	--	--	--	--
R066_4	E03	1.1577	--	--	--	1.5315	1.2847	0.2887	0.0741	--	--	--	--
L051_1	E04	-1.0041	--	--	--	0.9144	0.9785	0.6881	0.4949	1.0785	1.0696	0.6643	0.5616
L051_2	E04	-0.1706	--	--	--	0.9874	0.9902	0.4952	0.5290	0.9052	0.9503	0.5198	0.5460
L051_3	E04	-1.0738	--	--	--	1.0163	1.0130	0.7024	0.2947	0.9114	0.9633	0.7296	0.3452
L084_1	E04	0.6811	--	--	--	1.2662	1.1619	0.3000	0.1558	1.1494	1.0990	0.3380	0.2949
L084_2	E04	1.2930	--	--	--	1.8480	1.4198	0.2571	-0.0980	1.6339	1.3556	0.2424	-0.0868
L084_3	E04	-0.7591	--	--	--	0.9775	1.0072	0.6167	0.4414	1.0678	1.1087	0.5967	0.4537
L084_4	E04	-0.0834	--	--	--	0.9822	0.9971	0.4929	0.4535	1.0183	1.0268	0.4825	0.4532
R046_1	E04	0.6930	--	--	--	1.0650	1.0208	0.3140	0.3738	0.9931	0.9832	0.3776	0.4308
R046_2	E04	0.7109	--	--	--	1.1340	1.0649	0.3488	0.3609	0.9845	0.9759	0.3636	0.4586
R046_3	E04	1.0181	--	--	--	1.1722	1.0448	0.2628	0.2456	0.9775	0.9722	0.3380	0.4723
R051_1	E04	0.9642	--	--	--	1.1314	1.0567	0.2930	0.3063	1.0903	1.0332	0.2937	0.4396
R051_2	E04	-0.3483	--	--	--	1.0824	1.0710	0.5465	0.4163	1.0962	1.0570	0.6224	0.3332
R051_3	E04	3.2244	--	--	--	0.6614	0.8960	0.0628	0.1487	0.9803	1.3330	0.0909	0.2788
R051_4	E04	1.1056	--	--	--	1.1041	1.0421	0.2767	0.3126	1.0249	1.0054	0.2960	0.4291
L034_1	E05	-0.7740	--	--	--	1.0310	1.0243	0.5875	0.4643	--	--	--	--

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
L034_2	E05	0.0001	--	--	--	0.8392	0.8645	0.4748	0.6655	--	--	--	--
L034_3	E05	-0.5064	--	--	--	1.1629	1.0870	0.5803	0.4044	--	--	--	--
L091	E05	-1.8597	--	--	--	0.7697	0.9496	0.8106	0.4430	--	--	--	--
L093	E05	-2.2870	--	--	--	0.6890	0.9218	0.8753	0.2953	--	--	--	--
L096	E05	-0.6051	--	--	--	0.8234	0.8952	0.5635	0.5579	--	--	--	--
L108	E05	-0.5752	--	--	--	0.7986	0.8754	0.5995	0.6114	--	--	--	--
L123	E05	-3.1887	--	--	--	1.1290	1.0039	0.9400	0.1141	--	--	--	--
R007_1	E05	0.0667	--	--	--	0.8526	0.8728	0.4436	0.6658	--	--	--	--
R007_2	E05	0.3435	--	--	--	0.9394	0.9420	0.4173	0.4785	--	--	--	--
R057_1	E05	-0.6642	--	--	--	1.0662	1.0541	0.5899	0.4040	--	--	--	--
R057_2	E05	-0.8211	--	--	--	0.7608	0.8600	0.6283	0.6447	--	--	--	--
R057_3	E05	0.1254	--	--	--	0.9506	0.9477	0.4556	0.4982	--	--	--	--
R057_4	E05	0.3951	--	--	--	1.0382	0.9991	0.3789	0.3851	--	--	--	--
L074	E06	0.2431	--	--	--	1.0313	1.0226	0.4842	0.4538	1.0519	1.0184	0.4228	0.5071
L075	E06	-0.6764	--	--	--	0.7735	0.8565	0.6058	0.6437	0.7672	0.8633	0.6888	0.6120
L085_1	E06	-0.5733	--	--	--	1.0005	1.0308	0.5572	0.4823	0.8462	0.9240	0.6722	0.5193
L085_2	E06	0.3487	--	--	--	1.1487	1.0999	0.4258	0.3649	1.1354	1.1015	0.4988	0.4143
L085_3	E06	0.5750	--	--	--	1.2043	1.1320	0.3942	0.3331	1.2178	1.1467	0.4394	0.4362
L085_4	E06	0.0503	--	--	--	0.8979	0.9131	0.4453	0.5329	0.8472	0.8946	0.5606	0.5591
L085_5	E06	0.1449	--	--	--	0.9973	1.0066	0.4599	0.4520	0.9404	0.9655	0.5392	0.5559
R071_1	E06	-0.5136	--	--	--	0.8783	0.9147	0.5742	0.5739	0.9742	0.9743	0.6010	0.4883
R071_2	E06	-0.8211	--	--	--	0.8023	0.9079	0.6253	0.5596	0.7542	0.8643	0.6580	0.6669
R071_3	E06	-0.1475	--	--	--	0.9958	0.9812	0.5134	0.5904	0.8731	0.9114	0.5534	0.5810
R082_1	E06	1.1998	--	--	--	0.9929	0.9341	0.2311	0.3570	1.0744	1.0195	0.3468	0.5309
R082_2	E06	1.5579	--	--	--	1.1425	1.0059	0.2214	0.2576	1.3530	1.1807	0.2874	0.3815
R082_3	E06	0.1893	--	--	--	1.0067	1.0081	0.4404	0.4637	0.9186	0.9450	0.4584	0.6268
R082_4	E06	1.3883	--	--	--	1.5551	1.2324	0.2287	0.1040	1.3977	1.1938	0.2708	0.1658
L042_1	E07	-1.3147	--	--	--	0.9967	1.0190	0.7354	0.4312	--	--	--	--
L042_2	E07	0.4638	--	--	--	1.2907	1.2087	0.4345	0.2696	--	--	--	--

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
L042_3	E07	0.3946	--	--	--	1.1477	1.1003	0.4150	0.3131	--	--	--	--
L042_4	E07	1.1866	--	--	--	1.3381	1.1474	0.2791	0.1978	--	--	--	--
L092	E07	-0.4157	--	--	--	0.8673	0.9203	0.5922	0.5341	--	--	--	--
L095	E07	-1.1555	--	--	--	0.8422	0.9393	0.6966	0.4414	--	--	--	--
L106	E07	-0.1682	--	--	--	0.9823	0.9938	0.5218	0.4478	--	--	--	--
L107	E07	-1.9820	--	--	--	0.8449	0.9953	0.8350	0.2944	--	--	--	--
L124	E07	-3.1672	--	--	--	1.2744	1.0894	0.9320	0.1235	--	--	--	--
R006	E07	0.5391	--	--	--	1.1541	1.0684	0.4150	0.3539	--	--	--	--
R015_1	E07	-0.1527	--	--	--	0.7923	0.8377	0.5000	0.6724	--	--	--	--
R015_2	E07	0.2808	--	--	--	0.9260	0.9391	0.4539	0.4668	--	--	--	--
R062_1	E07	1.0162	--	--	--	0.9493	0.9389	0.2816	0.4152	--	--	--	--
R062_2	E07	0.3853	--	--	--	1.0103	0.9798	0.4175	0.4777	--	--	--	--
C002_1	E08	1.8799	--	--	--	1.4683	1.1652	0.1599	0.0961	1.5523	1.1646	0.1981	0.0824
C002_2	E08	0.7005	--	--	--	1.2628	1.1649	0.3508	0.2469	1.0830	1.0625	0.3647	0.3627
C002_3	E08	0.1718	--	--	--	0.8560	0.8713	0.4248	0.6430	0.8244	0.8609	0.4807	0.7491
C002_4	E08	-0.2978	--	--	--	0.8832	0.9266	0.5513	0.5634	0.8527	0.9224	0.6039	0.5771
C002_5	E08	1.1457	--	--	--	1.1915	1.0700	0.2864	0.2990	1.3476	1.1801	0.3116	0.3525
C002_6	E08	1.7917	0.3490	-0.3490	--	1.0195	1.0527	0.2578	0.4834	0.8908	1.0174	0.2778	0.6115
L077_1	E08	-0.1827	--	--	--	1.0540	1.0423	0.5274	0.4415	1.1106	1.0921	0.5580	0.3859
L077_2	E08	-0.3506	--	--	--	0.9379	0.9742	0.5656	0.4981	0.9027	0.9370	0.6111	0.5977
L078_1	E08	-0.3574	--	--	--	1.1580	1.0619	0.5394	0.4026	1.2081	1.0860	0.6063	0.3676
L078_2	E08	-0.3097	--	--	--	0.9366	0.9674	0.5656	0.5094	0.8948	0.9570	0.5725	0.5078
R068	E08	-0.8343	--	--	--	0.8425	0.9151	0.6420	0.5719	0.9274	0.9673	0.6498	0.5445
R088_1	E08	0.5228	--	--	--	1.0436	1.0132	0.3795	0.4634	0.9465	0.9488	0.3454	0.5420
R088_2	E08	1.0340	--	--	--	0.9446	0.9267	0.2649	0.4239	1.0481	1.0322	0.3502	0.4237
R088_3	E08	0.7174	--	--	--	1.3254	1.2054	0.3389	0.1603	1.1701	1.1387	0.3261	0.1895
L006_1	E09	-1.0467	--	--	--	0.8250	0.9084	0.6871	0.5700	--	--	--	--
L006_2	E09	-0.4330	--	--	--	0.8146	0.8628	0.5600	0.6859	--	--	--	--
L006_3	E09	-1.2630	--	--	--	0.7856	0.9181	0.7365	0.4876	--	--	--	--

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
L125	E09	-2.4254	--	--	--	1.0192	0.9912	0.8753	0.2125	--	--	--	--
R070_1	E09	0.4449	--	--	--	1.1440	1.0956	0.4047	0.4394	--	--	--	--
R070_2	E09	-1.6932	--	--	--	0.7825	0.9607	0.7882	0.4052	--	--	--	--
R070_3	E09	-2.0274	--	--	--	0.7561	0.9917	0.8424	0.2963	--	--	--	--
R077_1	E09	0.8579	--	--	--	1.0783	1.0353	0.3365	0.3531	--	--	--	--
R077_2	E09	0.4443	--	--	--	1.1050	1.0623	0.3882	0.2874	--	--	--	--
R077_3	E09	0.8213	--	--	--	0.8899	0.8845	0.3247	0.4684	--	--	--	--
R081_1	E09	0.1896	--	--	--	1.0260	1.0026	0.4329	0.4815	--	--	--	--
R081_2	E09	0.2569	--	--	--	1.2578	1.1786	0.4424	0.1944	--	--	--	--
R081_3	E09	0.4754	--	--	--	1.0704	1.0256	0.3647	0.3740	--	--	--	--
R081_4	E09	0.9813	--	--	--	0.9196	0.8998	0.3082	0.4639	--	--	--	--
L079_1	E10	0.5831	--	--	--	1.3751	1.2405	0.3729	0.2072	1.2188	1.1552	0.4101	0.3359
L079_2	E10	0.3772	--	--	--	1.1096	1.0686	0.4181	0.3830	1.0956	1.0623	0.4460	0.4667
L079_3	E10	0.5522	--	--	--	0.9977	0.9871	0.3729	0.4565	1.0134	1.0053	0.4484	0.5057
L087_1	E10	0.4832	--	--	--	0.9105	0.9074	0.3919	0.6250	0.9581	0.9691	0.5156	0.6752
L087_2	E10	0.3883	--	--	--	1.2190	1.1521	0.3777	0.2751	1.1387	1.1101	0.4484	0.3513
L087_3	E10	-0.4496	--	--	--	0.8958	0.9221	0.5629	0.6401	0.8393	0.9115	0.6139	0.6418
L088_1	E10	0.6889	--	--	--	1.3183	1.2148	0.3088	0.1915	1.2271	1.1667	0.3261	0.1341
L088_2	E10	-0.6672	--	--	--	1.0794	1.0469	0.6247	0.3473	1.0794	1.0527	0.6835	0.3277
R018_1	E10	0.7921	--	--	--	1.2534	1.1403	0.3397	0.2860	1.1464	1.1084	0.4125	0.4178
R018_2	E10	-0.2274	--	--	--	0.8855	0.9251	0.4917	0.5504	0.8631	0.9005	0.6403	0.5797
R018_3	E10	3.2614	-2.0050	2.0050	--	0.5865	0.7210	0.2233	0.6647	1.3717	1.5845	0.4077	0.7716
R028_1	E10	0.7942	--	--	--	1.1018	1.0522	0.3207	0.2869	1.0454	1.0288	0.3813	0.4335
R028_2	E10	0.2808	--	--	--	1.2344	1.1697	0.4252	0.3351	1.2526	1.1566	0.4197	0.2743
R028_3	E10	0.8761	--	--	--	1.1833	1.0839	0.3017	0.2908	1.0877	1.0400	0.3861	0.4921
L083	E11	-1.9057	--	--	--	0.8551	0.9565	0.8310	0.3190	--	--	--	--
L098	E11	-0.5141	--	--	--	0.8976	0.9415	0.5714	0.5750	--	--	--	--
L112	E11	-2.1683	--	--	--	0.8998	1.0472	0.8595	0.2712	--	--	--	--
L115	E11	-1.2691	--	--	--	0.8264	0.9139	0.7167	0.4826	--	--	--	--

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
L118	E11	-3.0627	--	--	--	0.9327	1.0837	0.9262	0.1383	--	--	--	--
L119	E11	-0.3983	--	--	--	0.8916	0.9310	0.5476	0.5889	--	--	--	--
R012_1	E11	-1.1361	--	--	--	0.8917	0.9532	0.6571	0.5613	--	--	--	--
R012_2	E11	-0.4884	--	--	--	0.8360	0.8952	0.5952	0.5952	--	--	--	--
R034_2	E11	-0.0099	--	--	--	0.9027	0.9201	0.4738	0.6097	--	--	--	--
R034_3	E11	0.6487	--	--	--	0.9460	0.9292	0.3238	0.4445	--	--	--	--
R055_1	E11	0.5782	--	--	--	1.3459	1.2364	0.4000	0.1494	--	--	--	--
R055_2	E11	0.9007	--	--	--	0.8303	0.8784	0.2857	0.5093	--	--	--	--
R055_3	E11	0.1726	--	--	--	1.1347	1.0998	0.4452	0.3009	--	--	--	--
R055_4	E11	0.6024	--	--	--	1.0487	1.0056	0.3452	0.4264	--	--	--	--
L082_1	E12	-0.8389	--	--	--	0.9872	0.9944	0.6043	0.5415	0.9796	0.9910	0.6555	0.4744
L082_2	E12	0.3753	--	--	--	1.1977	1.1168	0.4388	0.2436	1.0083	1.0112	0.5263	0.5345
L082_3	E12	0.5979	--	--	--	1.2517	1.1791	0.3525	0.2743	1.2104	1.1684	0.5096	0.4147
L082_4	E12	1.1528	--	--	--	1.3892	1.1355	0.2902	0.2937	1.2001	1.0995	0.2751	0.3101
L089_1	E12	-1.5323	--	--	--	0.8737	0.9768	0.7578	0.4511	0.8555	1.0219	0.7775	0.4237
L089_2	E12	0.9701	--	--	--	1.0858	1.0413	0.3094	0.3357	0.9938	0.9842	0.3110	0.4588
L089_3	E12	1.0425	--	--	--	1.4839	1.2778	0.2998	0.1798	1.2392	1.1652	0.3445	0.1714
R076_1	E12	-0.0596	--	--	--	1.1772	1.1418	0.5060	0.3241	1.0396	1.0417	0.4785	0.4913
R076_2	E12	-0.2768	--	--	--	0.8950	0.9288	0.5492	0.5590	0.8898	0.9245	0.5335	0.6040
R076_3	E12	0.6958	--	--	--	1.1364	1.0658	0.3237	0.2655	1.0238	0.9928	0.4258	0.6226
R079_1	E12	0.3549	--	--	--	1.0058	0.9856	0.4365	0.5093	0.9996	0.9966	0.5000	0.5123
R079_2	E12	0.7228	--	--	--	1.2303	1.0794	0.3381	0.2643	1.2051	1.1186	0.3947	0.3083
R079_3	E12	0.8047	--	--	--	1.1280	1.0326	0.3381	0.3141	1.0066	0.9930	0.3708	0.4181
R079_4	E12	0.5961	-0.2398	0.2398	--	0.9118	0.9431	0.3717	0.5333	0.9580	0.9340	0.3517	0.5290
L012_1	E13	-0.9610	--	--	--	0.9133	0.9512	0.6627	0.5571	--	--	--	--
L012_2	E13	-1.9802	--	--	--	0.7821	0.9160	0.8289	0.4779	--	--	--	--
L040_1	E13	-1.5062	--	--	--	0.9318	0.9690	0.7759	0.4429	--	--	--	--
L040_2	E13	-1.2137	--	--	--	0.9481	1.0344	0.7205	0.4774	--	--	--	--
L040_3	E13	-0.7390	--	--	--	0.8430	0.8988	0.6024	0.6352	--	--	--	--

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
L040_4	E13	-2.1672	--	--	--	0.9421	0.9736	0.8361	0.3354	--	--	--	--
L063	E13	1.0706	--	--	--	1.0367	0.9411	0.2964	0.3100	--	--	--	--
L117	E13	-1.9879	--	--	--	1.0586	1.0537	0.7928	0.2811	--	--	--	--
R011_1	E13	-2.6307	--	--	--	0.6288	0.8857	0.8988	0.3416	--	--	--	--
R011_2	E13	-1.2547	--	--	--	0.7722	0.8876	0.7133	0.6618	--	--	--	--
R065_1	E13	0.5325	--	--	--	0.8207	0.8696	0.3566	0.5633	--	--	--	--
R065_2	E13	0.5180	--	--	--	0.8693	0.8931	0.3952	0.5006	--	--	--	--
R073_1	E13	-0.9224	--	--	--	0.7993	0.8678	0.6675	0.6080	--	--	--	--
R073_2	E13	-1.6409	--	--	--	0.7234	0.8507	0.8024	0.5690	--	--	--	--
L128_1	E14	-0.9264	--	--	--	--	--	--	--	0.7233	0.8494	0.6869	0.5898
L128_3	E14	-0.8099	--	--	--	--	--	--	--	0.8969	0.9221	0.6495	0.5130
L128_4	E14	-0.4234	--	--	--	--	--	--	--	0.9093	0.9456	0.5514	0.5071
L149_1	E14	0.6289	--	--	--	--	--	--	--	1.3763	1.2891	0.3551	0.0512
L149_2	E14	0.9667	--	--	--	--	--	--	--	1.2260	1.1315	0.3178	0.2954
L153_1	E14	-0.2092	--	--	--	--	--	--	--	1.1104	1.0895	0.5748	0.3449
L153_2	E14	0.1402	--	--	--	--	--	--	--	1.0332	1.0379	0.4883	0.4452
L153_3	E14	0.3792	--	--	--	--	--	--	--	1.1294	1.1022	0.4720	0.2979
R017_1	E14	-0.0620	--	--	--	--	--	--	--	0.7980	0.8412	0.5034	0.6938
R135_1	E14	0.3132	--	--	--	--	--	--	--	1.2596	1.1525	0.4123	0.4037
R135_2	E14	0.4268	--	--	--	--	--	--	--	1.0499	1.0331	0.3964	0.4001
R135_3	E14	0.2104	--	--	--	--	--	--	--	0.9116	0.9289	0.4374	0.6202
R135_4	E14	2.7264	-0.0379	0.0379	--	--	--	--	--	0.6392	0.8160	0.1093	0.2910
R161_1	E14	0.1533	--	--	--	--	--	--	--	0.7734	0.8113	0.4601	0.7386
R161_2	E14	-0.6632	--	--	--	--	--	--	--	0.9620	1.0005	0.5581	0.5217
R161_4	E14	-0.5779	--	--	--	--	--	--	--	0.7008	0.8056	0.6014	0.7794
L129_1	E15	0.0609	--	--	--	--	--	--	--	0.7600	0.8241	0.5012	0.7349
L129_3	E15	-0.8343	--	--	--	--	--	--	--	1.1460	1.0650	0.6556	0.4143
L129_4	E15	-0.7935	--	--	--	--	--	--	--	0.8428	0.9375	0.6983	0.5221
L134_1	E15	-0.8056	--	--	--	--	--	--	--	0.9947	1.0189	0.6817	0.4062

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
L134_2	E15	0.9622	--	--	--	--	--	--	--	1.3058	1.1935	0.3610	0.3444
L134_3	E15	-0.2696	--	--	--	--	--	--	--	0.9671	1.0034	0.5914	0.4932
L146_1	E15	0.7648	--	--	--	--	--	--	--	1.0318	1.0216	0.3682	0.4393
L146_2	E15	1.3353	--	--	--	--	--	--	--	1.6793	1.3556	0.2850	0.0836
R126_1	E15	0.1568	--	--	--	--	--	--	--	0.8240	0.8702	0.4745	0.7126
R126_2	E15	0.7028	--	--	--	--	--	--	--	1.0033	0.9764	0.3565	0.4202
R126_3	E15	0.9247	--	--	--	--	--	--	--	1.0035	0.9739	0.3657	0.4826
R126_5	E15	3.2643	-0.0560	0.0560	--	--	--	--	--	0.5959	0.9792	0.1181	0.2981
R159	E15	0.1056	--	--	--	--	--	--	--	0.8331	0.8817	0.5069	0.6121
R160_2	E15	-0.4029	--	--	--	--	--	--	--	0.7375	0.8224	0.6019	0.7677
R160_3	E15	0.0406	--	--	--	--	--	--	--	0.9362	0.9547	0.5278	0.6248
R160_4	E15	-0.3590	--	--	--	--	--	--	--	0.7828	0.8567	0.5856	0.6991
L017	E16	-0.9903	--	--	--	--	--	--	--	0.7527	0.8918	0.7156	0.5698
L035_1	E16	-0.0603	--	--	--	--	--	--	--	0.8838	0.9184	0.5047	0.5700
L035_2	E16	-0.1402	--	--	--	--	--	--	--	0.8468	0.8969	0.5261	0.6491
L132_1	E16	-0.1218	--	--	--	--	--	--	--	0.7791	0.8398	0.5000	0.6952
L132_2	E16	-0.3739	--	--	--	--	--	--	--	0.8886	0.9536	0.5853	0.5719
L145_1	E16	-0.5029	--	--	--	--	--	--	--	1.0541	1.0624	0.6445	0.4216
L145_2	E16	0.8262	--	--	--	--	--	--	--	1.3182	1.1927	0.3791	0.2889
L145_3	E16	0.4475	--	--	--	--	--	--	--	1.3776	1.3035	0.4289	0.1056
R094_1	E16	-0.1621	--	--	--	--	--	--	--	0.8278	0.8866	0.5497	0.6114
R094_2	E16	1.1854	--	--	--	--	--	--	--	0.9170	0.9106	0.2887	0.3916
R094_4	E16	3.1403	-1.1212	1.1212	--	--	--	--	--	1.2452	1.1908	0.2286	0.3626
R106_1	E16	-0.0106	--	--	--	--	--	--	--	0.8949	0.9361	0.5058	0.6234
R106_2	E16	0.2741	--	--	--	--	--	--	--	0.8311	0.8699	0.4249	0.6233
R106_3	E16	0.7752	--	--	--	--	--	--	--	1.2158	1.1422	0.3857	0.3564
R106_4	E16	0.9820	--	--	--	--	--	--	--	1.0887	1.0275	0.3210	0.3261
R165	E16	-0.8627	--	--	--	--	--	--	--	0.7921	0.9111	0.6744	0.5923
L032	E17	0.4503	--	--	--	--	--	--	--	1.0190	1.0028	0.4044	0.5245

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
L058	E17	-1.2185	--	--	--	--	--	--	--	0.9852	0.9952	0.7700	0.3162
L137_1	E17	0.1566	--	--	--	--	--	--	--	0.8739	0.8976	0.5085	0.6657
L137_2	E17	0.2411	--	--	--	--	--	--	--	0.8629	0.8954	0.4722	0.6521
L137_3	E17	-1.2744	--	--	--	--	--	--	--	0.8326	0.9653	0.7724	0.4150
L138_1	E17	0.4958	--	--	--	--	--	--	--	0.8817	0.8931	0.4358	0.6171
L138_2	E17	0.4998	--	--	--	--	--	--	--	0.9422	0.9407	0.3898	0.5417
L138_3	E17	0.6719	--	--	--	--	--	--	--	1.1915	1.1489	0.3220	0.2912
R053_1	E17	0.4110	--	--	--	--	--	--	--	1.0756	1.0579	0.3986	0.3812
R053_2	E17	1.5908	--	--	--	--	--	--	--	1.3461	1.1453	0.2335	0.2450
R053_3	E17	1.3686	--	--	--	--	--	--	--	1.3659	1.2284	0.2453	0.1958
R053_4	E17	1.1897	--	--	--	--	--	--	--	1.2431	1.1380	0.2948	0.2584
R092_1	E17	0.6547	--	--	--	--	--	--	--	0.9346	0.9369	0.3656	0.5347
R092_2	E17	0.1815	--	--	--	--	--	--	--	0.9834	0.9796	0.4623	0.4959
R092_3	E17	0.6495	--	--	--	--	--	--	--	0.9981	0.9874	0.3561	0.4854
R092_5	E17	2.9213	-1.0706	0.0711	0.9995	--	--	--	--	0.7454	0.8864	0.2005	0.4833
L037_1	E18	-0.5216	--	--	--	--	--	--	--	0.9376	0.9899	0.6063	0.5100
L037_2	E18	0.3842	--	--	--	--	--	--	--	1.2336	1.1962	0.4469	0.2784
L131_1	E18	-0.2631	--	--	--	--	--	--	--	1.0040	1.0313	0.5580	0.4738
L131_2	E18	-0.0440	--	--	--	--	--	--	--	0.9236	0.9477	0.5531	0.5150
L131_3	E18	-0.5315	--	--	--	--	--	--	--	0.8742	0.9401	0.6232	0.4829
L151_1	E18	0.1887	--	--	--	--	--	--	--	1.2463	1.1991	0.4855	0.2829
L151_2	E18	0.1787	--	--	--	--	--	--	--	1.0327	1.0329	0.4976	0.4486
L151_3	E18	-0.1442	--	--	--	--	--	--	--	0.9695	0.9967	0.5459	0.4795
R100_1	E18	0.2398	--	--	--	--	--	--	--	0.8967	0.9198	0.4614	0.6006
R100_2	E18	1.3773	--	--	--	--	--	--	--	1.3334	1.1925	0.2705	0.2106
R100_3	E18	0.5850	--	--	--	--	--	--	--	1.0370	1.0222	0.3961	0.3998
R125_1	E18	0.5849	--	--	--	--	--	--	--	1.0675	1.0484	0.3937	0.3303
R125_2	E18	0.6270	--	--	--	--	--	--	--	0.9872	0.9717	0.3454	0.4636
R125_4	E18	0.8562	--	--	--	--	--	--	--	0.8671	0.9037	0.3744	0.6343

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
R163_1	E18	0.3652	--	--	--	--	--	--	--	0.9705	0.9693	0.4251	0.5629
R163_2	E18	0.0021	--	--	--	--	--	--	--	0.7974	0.8360	0.5048	0.7991
L027_1	E19	-1.1468	--	--	--	--	--	--	--	0.9011	0.9773	0.7260	0.5247
L027_2	E19	-0.6064	--	--	--	--	--	--	--	0.9854	0.9865	0.6514	0.5192
L043_1	E19	-0.5615	--	--	--	--	--	--	--	0.9095	0.9540	0.6394	0.5931
L043_2	E19	0.0177	--	--	--	--	--	--	--	0.8328	0.8694	0.4904	0.6620
L130_1	E19	-0.6635	--	--	--	--	--	--	--	0.9410	1.0090	0.6226	0.4732
L130_2	E19	0.9833	--	--	--	--	--	--	--	0.9456	0.9216	0.2861	0.4365
L130_3	E19	0.9572	--	--	--	--	--	--	--	1.1524	1.0738	0.3389	0.3299
L130_4	E19	0.8123	--	--	--	--	--	--	--	0.9736	0.9405	0.3702	0.4130
R123_1	E19	0.7238	--	--	--	--	--	--	--	1.1117	1.0700	0.3606	0.3566
R123_2	E19	1.7557	--	--	--	--	--	--	--	1.6024	1.2877	0.2163	0.0456
R123_3	E19	1.3562	--	--	--	--	--	--	--	1.1451	1.0932	0.2452	0.2116
R123_4	E19	1.0337	--	--	--	--	--	--	--	1.1392	1.0774	0.2861	0.2251
R123_5	E19	0.5500	--	--	--	--	--	--	--	1.1373	1.1018	0.3798	0.2645
R162_1	E19	-0.6827	--	--	--	--	--	--	--	1.0608	1.0016	0.7067	0.5047
R162_2	E19	-1.0369	--	--	--	--	--	--	--	0.7705	0.8838	0.7163	0.6259
R164	E19	-0.7516	--	--	--	--	--	--	--	0.8851	0.9333	0.6875	0.4981
L031	E20	-0.6299	--	--	--	--	--	--	--	0.7875	0.8766	0.6526	0.6530
L139_1	E20	-0.2828	--	--	--	--	--	--	--	0.9545	0.9900	0.5931	0.4524
L139_2	E20	-0.1374	--	--	--	--	--	--	--	1.0442	1.0355	0.5484	0.4694
L144_1	E20	1.4258	--	--	--	--	--	--	--	1.6791	1.4452	0.2630	0.0483
L144_2	E20	0.5389	--	--	--	--	--	--	--	1.1346	1.0707	0.4119	0.3122
L144_3	E20	0.6006	--	--	--	--	--	--	--	1.1635	1.1185	0.4218	0.3628
L152_1	E20	-0.2555	--	--	--	--	--	--	--	1.0912	1.0511	0.6129	0.4053
L152_2	E20	0.2553	--	--	--	--	--	--	--	0.9266	0.9352	0.4591	0.4845
R095_1	E20	-0.5799	--	--	--	--	--	--	--	1.3089	1.1590	0.6129	0.2420
R095_2	E20	0.2356	--	--	--	--	--	--	--	0.8919	0.9193	0.4690	0.5777
R095_3	E20	0.4847	--	--	--	--	--	--	--	1.1619	1.1184	0.4342	0.2859

(續下頁)

題號	區塊號	平均試題 難度值 ( $\bar{\delta}$ )	step1 ( $\tau_1$ )	step2 ( $\tau_2$ )	step3 ( $\tau_3$ )	七年級				八年級			
						Outfit	Infit	通過率	鑑別度	Outfit	Infit	通過率	鑑別度
R099_1	E20	0.6225	--	--	--	--	--	--	--	0.8486	0.8671	0.4367	0.6009
R099_2	E20	0.6408	--	--	--	--	--	--	--	0.9275	0.9308	0.3871	0.4953
R099_3	E20	1.2609	--	--	--	--	--	--	--	1.2742	1.1456	0.2829	0.2568
R099_4	E20	2.0287	-0.6547	0.6547	--	--	--	--	--	0.6373	0.7837	0.3226	0.7078
R166	E20	-0.1641	--	--	--	--	--	--	--	0.8806	0.9065	0.5484	0.5853

註：相關評量指標，請見第四章說明。

### 附錄十三 預試評量試題通過率與鑑別度（九年級）（區塊 21 至 27）

題號	區塊	預試時間 (年月)	通過率	高分組通過率	低分組通過率	鑑別度
L156_1	E21	202010	0.79	0.96	0.50	0.46
L156_2	E21	202010	0.72	0.93	0.45	0.48
L156_3	E21	202010	0.74	0.98	0.45	0.53
L156_4	E21	202010	0.31	0.59	0.05	0.54
L157_1	E21	202010	0.53	1.00	0.30	0.70
L157_2	E21	202010	0.53	0.67	0.37	0.30
L157_3	E21	202010	0.87	1.00	0.77	0.23
L157_4	E21	202010	0.85	1.00	0.73	0.27
R111_1	E21	202010	0.58	0.83	0.34	0.49
R111_2	E21	202010	0.48	0.79	0.22	0.57
R111_3	E21	202010	0.65	0.97	0.34	0.63
R111_4	E21	202010	0.67	1.00	0.27	0.73
R111_5	E21	202010	0.55	0.88	0.23	0.65
R114_1	E21	202010	0.48	0.94	0.19	0.75
R114_2	E21	202010	0.61	1.00	0.29	0.71
R114_3	E21	202010	0.54	0.81	0.29	0.52
R114_4	E21	202010	0.43	0.69	0.19	0.50
L160_1	E22	202010	0.8	1.00	0.52	0.48
L160_3	E22	202010	0.77	1.00	0.48	0.52
L160_4	E22	202010	0.58	0.84	0.33	0.51
L167_1	E22	202010	0.91	1.00	0.74	0.26
L167_2	E22	202010	0.56	0.70	0.43	0.27
L167_3	E22	202010	0.69	0.94	0.36	0.58
L167_4	E22	202010	0.89	1.00	0.70	0.30
L182	E22	202010	0.82	1.00	0.63	0.37

(續下頁)

題號	區塊	預試時間 (年月)	通過率	高分組通過率	低分組通過率	鑑別度
R120_1	E22	202010	0.37	0.79	0.04	0.75
R120_2	E22	202010	0.57	0.95	0.30	0.65
R120_3	E22	202010	0.43	0.74	0.30	0.44
R120_4	E22	202010	0.39	0.53	0.26	0.27
R140_1	E22	202010	0.25	0.53	0.04	0.49
R140_2	E22	202010	0.57	0.89	0.26	0.63
R140_3	E22	202010	0.37	0.84	0.04	0.80
R140_4	E22	202010	0.49	0.84	0.17	0.67
R140_5	E22	202010	0.53	0.84	0.13	0.71
L143_1	E23	201910	0.375	0.68	0.21	0.47
L143_2	E23	201910	0.335	0.37	0.25	0.12
L143_3	E23	201910	0.337	0.46	0.23	0.23
L158_1	E23	202010	0.71	0.95	0.48	0.47
L158_2	E23	202010	0.69	1.00	0.39	0.61
L158_3	E23	202010	0.76	0.95	0.57	0.38
L165_1	E23	202010	0.43	0.75	0.18	0.57
L165_2	E23	202010	0.49	0.75	0.18	0.57
R142_1	E23	202010	0.67	1.00	0.36	0.64
R142_2	E23	202010	0.69	1.00	0.45	0.55
R142_3	E23	202010	0.47	0.69	0.36	0.33
R142_4	E23	202010	0.45	0.81	0.09	0.72
R142_5	E23	202010	0.43	0.94	0.09	0.85
R143_1	E23	202010	0.59	0.94	0.38	0.56
R143_2	E23	202010	0.52	0.94	0.29	0.65
R143_3	E23	202010	0.59	1.00	0.19	0.81

(續下頁)

題號	區塊	預試時間 (年月)	通過率	高分組通過率	低分組通過率	鑑別度
R143_4	E23	202010	0.46	0.94	0.10	0.84
L163_1	E24	202010	0.67	1.00	0.19	0.81
L163_2	E24	202010	0.74	1.00	0.43	0.57
L163_3	E24	202010	0.76	0.94	0.52	0.42
L164_1	E24	202010	0.71	0.84	0.59	0.25
L164_2	E24	202010	0.42	0.56	0.31	0.25
L164_3	E24	202010	0.82	0.95	0.59	0.36
L164_4	E24	202010	0.81	0.95	0.56	0.39
L181	E24	202010	0.56	0.90	0.17	0.73
R141_1	E24	202010	0.37	0.38	0.36	0.02
R141_2	E24	202010	0.51	0.94	0.27	0.67
R141_3	E24	202010	0.39	0.87	0.09	0.78
R141_4	E24	202010	0.41	0.63	0.27	0.36
R141_5	E24	202010	0.29	0.75	0.00	0.75
R141_6	E24	202010	0.22	0.50	0.05	0.45
R146_1	E24	202010	0.43	0.68	0.11	0.57
R146_2	E24	202010	0.63	1.00	0.16	0.84
R146_3	E24	202010	0.41	0.79	0.05	0.74
L148_1	E25	202010	0.52	0.81	0.22	0.59
L148_2	E25	202010	0.35	0.57	0.17	0.40
L148_3	E25	202010	0.29	0.43	0.17	0.26
L174	E25	202010	0.57	0.94	0.35	0.59
L178_1	E25	202010	0.47	0.75	0.37	0.38
L178_2	E25	202010	0.64	0.83	0.53	0.30
L178_3	E25	202010	0.17	0.33	0.07	0.26
L178_4	E25	202010	0.43	0.92	0.17	0.75

(續下頁)

題號	區塊	預試時間 (年月)	通過率	高分組通過率	低分組通過率	鑑別度
R139_1	E25	202010	0.61	0.94	0.43	0.51
R139_2	E25	202010	0.65	0.94	0.38	0.56
R139_3	E25	202010	0.46	0.63	0.38	0.25
R139_4	E25	202010	0.24	0.37	0.05	0.32
R156_1	E25	202010	0.28	0.83	0.10	0.73
R156_2	E25	202010	0.09	0.33	0.03	0.30
R156_3	E25	202010	0.42	0.75	0.27	0.48
R156_4	E25	202010	0.28	0.58	0.13	0.45
L162_1	E26	202010	0.45	0.77	0.20	0.57
L162_2	E26	202010	0.59	0.85	0.28	0.57
L162_3	E26	202010	0.65	0.87	0.39	0.48
L179_1	E26	202010	0.31	0.21	0.37	-0.16
L179_2	E26	202010	0.57	0.79	0.42	0.37
L179_3	E26	202010	0.57	1.00	0.21	0.79
L179_4	E26	202010	0.51	0.84	0.26	0.58
L179_5	E26	202010	0.673	1.11	0.11	1.00
R110_1	E26	201910	0.284	0.48	0.20	0.28
R110_2	E26	201910	0.442	0.73	0.22	0.52
R110_3	E26	201910	0.306	0.29	0.26	0.03
R110_4	E26	201910	0.45	0.74	0.28	0.46
R115_1	E26	202010	0.55	0.83	0.40	0.43
R115_2	E26	202010	0.28	0.50	0.20	0.30
R115_3	E26	202010	0.32	0.83	0.13	0.70
R115_4	E26	202010	0.28	0.75	0.13	0.62
R115_5	E26	202010	0.15	0.50	0.00	0.50
L161_1	E27	202010	0.77	0.98	0.50	0.48

(續下頁)

題號	區塊	預試時間 (年月)	通過率	高分組通過率	低分組通過率	鑑別度
L161_2	E27	202010	0.66	0.91	0.40	0.51
L161_3	E27	202010	0.52	0.89	0.25	0.64
L168_1	E27	202010	0.82	0.98	0.50	0.48
L168_3	E27	202010	0.75	0.95	0.43	0.52
L168_4	E27	202010	0.54	0.74	0.27	0.47
L172	E27	202010	0.65	0.87	0.38	0.49
L173	E27	202010	0.71	0.94	0.45	0.49
R128_1	E27	202010	0.59	0.87	0.40	0.47
R128_2	E27	202010	0.55	0.87	0.45	0.42
R128_3	E27	202010	0.49	0.81	0.35	0.46
R128_4	E27	202010	0.41	0.88	0.00	0.88
R128_5	E27	202010	0.59	1.19	0.00	1.19
R132_1	E27	202010	0.39	0.56	0.40	0.16
R132_2	E27	202010	0.43	0.75	0.25	0.50
R132_3	E27	202010	0.51	0.75	0.45	0.30
R132_4	E27	202010	0.24	0.31	0.25	0.06

註：由於新型冠狀病毒肺炎疫情影響，干擾九年級正式施測，故，本書現行僅呈現預試參數估計值。此外，預試後，命題團隊會針對鑑別度不佳試題，進行檢視與修訂。



## 電子出版品

---

### 建構英語文素養評量指引：TASAL 標準本位大型評量

發行人：林崇熙

作者：謝進昌

設計審核：策略溝通辦公室

美編設計：加斌有限公司

出版機關：國家教育研究院

地址：237201 新北市三峽區三樹路 2 號

電話：(02) 7740-7890

傳真：(02) 7740-7064

網址：<https://www.naer.edu.tw>

出版日期：112 年 2 月

定價：新臺幣 230 元

GPN：4711200014

ISBN：9786263451209

本書通過雙向匿名學術審查

本書著作財產權為國家教育研究院所有，請進行合理引用。

若欲直接使用本書內容，須徵求同意及書面授權。