

113 學年度  
人工智慧與大數據之管理應用  
期末報告

運用 Power BI 於  
分析篩選測驗作答反應

學生：資訊管理學系 S1161113 陳毓欣

指導教授：賴慧敏教授/博士

114 年 6 月

# 目錄

圖目錄.....	3
表目錄.....	4
壹、緒論.....	5
一、研究背景.....	5
二、研究動機.....	5
貳、研究方法.....	5
一、資料前處理.....	5
二、資料分析與視覺化（Power BI）.....	6
參、研究設計.....	6
一、研究對象與資料來源.....	7
二、研究變項與操作性定義.....	7
三、星型架構（Star Schema）.....	7
四、研究流程圖.....	8
五、分析架構對應五大主題.....	8
肆、研究結果.....	8
一、學習困難點分析.....	8
二、縱向學習發展追蹤.....	9
三、學科能力落差研究.....	10
四、基本學習內容適切性評估.....	11
五、診斷測驗效度研究.....	12
伍、結論.....	13
陸、影片說明.....	14

## 圖目錄

圖 1 大略星型架構簡述.....	7
圖 2 流程與詳細步驟.....	8
圖 3 學習困難點總覽.....	9
圖 4 縱向學習發展追蹤總覽.....	10
圖 5 學科能力落差研究總覽.....	11
圖 6 基本學習內容適切性評估總覽.....	12
圖 7 診斷測驗效度研究總覽.....	13

## 表目錄

表 1 變項名稱與類型說明.....	7
表 2 分析架構與對應主題.....	8

## 壹、 緒論

### 一、研究背景

在教育現場中，學習扶助政策旨在協助學習低成就學生縮小學科能力落差，促進其基本學力之養成。各縣市教育局處長期累積之扶助測驗數據，蘊含豐富的診斷資訊，倘若能加以整合並視覺化呈現，將有助於教師釐清學生之學習困難類型、掌握縱向學習發展軌跡，並提升教學策略之針對性與效能。然現有扶助資料多為原始紀錄，若缺乏系統化分析與圖像化支持，實難有效運用於教育決策與實務調整。

### 二、研究動機

本研究以「縣市政府教育局處合作之學習扶助數據資料」為基礎，期望透過數據前處理與視覺化建構，針對學生學習表現進行多維度分析，從而：

1. 辨識學習困難點與關鍵錯誤類型；
2. 探究不同學習階段學生能力發展之縱向變化；
3. 比較學科間或學生群體間之能力落差；
4. 評估基本學習內容之適切性；
5. 驗證診斷性測驗工具之信效度。

藉由上述分析，本研究旨在提供具體、可操作之教育輔助決策依據，並促進精準扶助之落實。

## 貳、 研究方法

### 一、資料前處理

本研究所採用之資料來自多縣市教育局處提供之扶助測驗紀錄，涵蓋學校代號、學生基本資料、作答記錄、得分情形與通過與否等欄位。為提升分析品質與圖表準確性，資料預處理步驟如下：

1. 缺失值處理：刪除或補值遺漏欄位，避免分析失真。
2. 日期欄位標準化：將「交易日期」、「施測日期」等欄位轉換為標準格式，並進一步拆分為年、月、日欄位以利時間序列分析。
3. 衍生欄位建立：包含是否通過轉換為二元值（1/0）、計算答對題數、平均成績、通過率等。

Python 整理如下：

```
import pandas as pd

# 讀取資料
df = pd.read_csv("learning_support.csv")

# 缺失值處理：刪除關鍵欄位缺值資料
df.dropna(subset=["學生 ID", "施測日期", "評量成績"], inplace=True)

# 將施測日期欄位轉換為 datetime 格式並拆分
df["施測日期"] = pd.to_datetime(df["施測日期"], errors="coerce")
df["年"] = df["施測日期"].dt.year
df["月"] = df["施測日期"].dt.month
df["日"] = df["施測日期"].dt.day
```

## 二、資料分析與視覺化（Power BI）

為提升資料可讀性與支持教育決策，後續使用 Power BI 建構互動式儀表板，進行以下五大主題分析：

### 1. 學習困難點分析

針對答錯題數統計與答題分布進行視覺化，輔以熱點題目鑑別度與錯誤率分析。

### 2. 縱向學習發展追蹤

透過時序圖與學生歷次測驗軌跡，呈現學習進步曲線與個別變化趨勢。

### 3. 學科能力落差研究

以長條圖、雷達圖等，分析不同學校、性別、年級學生在各領域能力表現之差異。

### 4. 基本學習內容適切性評估

交叉分析學生作答表現與學習內容分類，探討題目難度是否符合學生實際能力。

### 5. 診斷測驗效度研究

透過評量成績、答對題數與是否通過間之關聯模型(如散佈圖、邏輯斯回歸分析)，驗證測驗之信度與通過門檻合理性。

## 參、 研究設計

本研究採用次級資料分析法，針對來自縣市政府教育局處之學習扶助測驗紀錄資料進行多面向剖析。研究設計依循資料科學流程，涵蓋資料收集、清理、轉換、視覺化與詮釋五大步驟，並結合教育評量理論與學習分析架構，以提升研究的信效度與實務應用價值。

一、研究對象與資料來源

研究資料來自教育局處委託辦理之「學習扶助測驗資料」，涵蓋多年度、跨縣市、不同年級之學生作答紀錄。資料項目包括學生基本資料(性別、入學年度、學校代號)、測驗結果(答對題數、成績、是否通過)、測驗日期與個別題目作答紀錄(原始作答、正誤判斷等)，總樣本數達數十萬筆以上。

二、研究變項與操作性定義

表 1 變項名稱與類型說明

變項名稱	類型	操作性定義
入學年度	類別變項	學生入學年份，用以進行縱向比較
性別	類別變項	學生性別（男 / 女），用於群體差異分析
答對題數	數值變項	測驗中答對之題數，用以估計基本學力程度
評量成績	數值變項	測驗後系統計算之總得分，用於學力表現評估
是否通過	二元變項	測驗成績是否達通過標準，轉為 0（否）與 1（是）
原始作答題目 1-25	類別變項	學生每題之原始作答選項，用於錯誤類型分析
答對答錯 1-25	二元變項	每題是否答對之紀錄（1：對，0：錯），作為難度與鑑別度分析依據

三、星型架構（Star Schema）

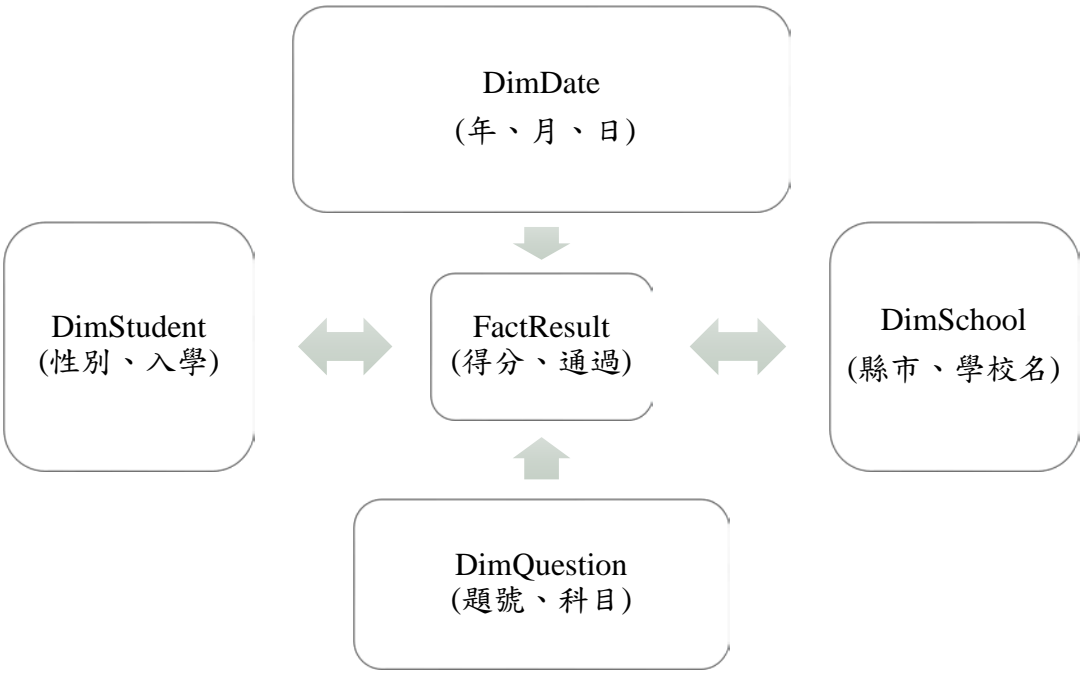


圖 1 大略星型架構簡述

四、研究流程圖

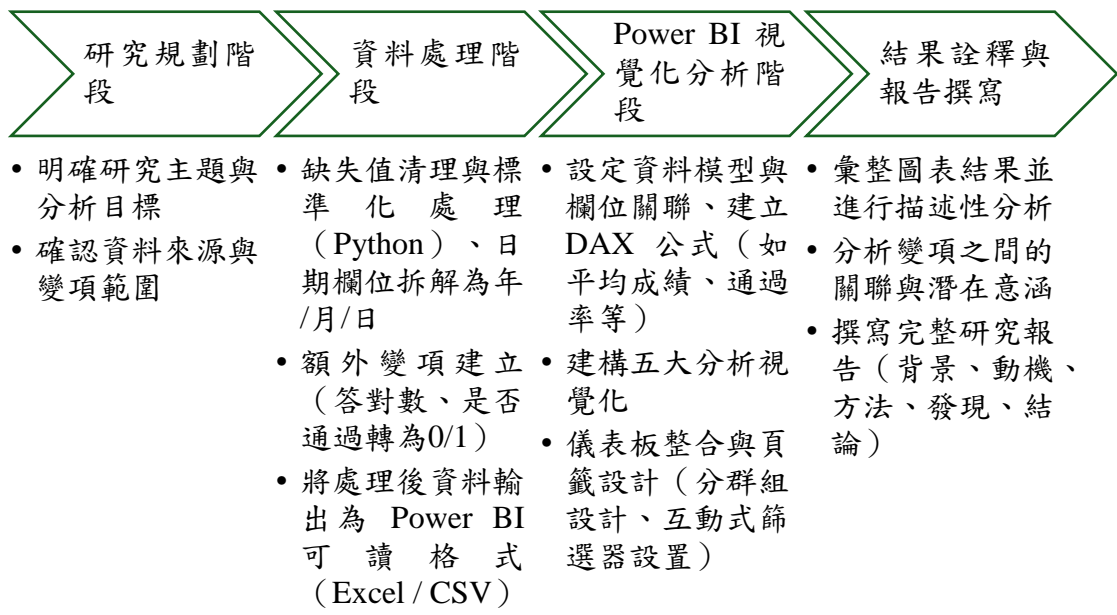


圖 2 流程與詳細步驟

五、分析架構對應五大主題

表 2 分析架構與對應主題

主題分類	分析方法與視覺呈現
學習困難點分析	錯誤率統計圖、題目鑑別度圖、熱點題型分布圖
縱向學習發展追蹤	學生歷次測驗曲線圖、趨勢折線圖
學科能力落差研究	不同群體長條圖、箱型圖、雷達圖比較
基本學習內容適切性評估	難度分析散佈圖、通過率熱力圖
診斷測驗效度研究	成績與通過率散佈圖、邏輯斯迴歸模型、相關係數分析

肆、研究結果

一、學習困難點分析

本研究針對學習扶助測驗資料中，學生於各題目的答對情形進行視覺化分析，藉由分群與分層方式，探究學習困難的分布趨勢與背後潛在的認知障礙來源。

圖表左上顯示「不同年級、性別各題答對數量」，透過堆疊區域河流圖清楚呈現答對量隨年級變化的趨勢，並藉由顏色區隔不同題目編號，揭示部分特定題目在特定年級或性別間出現明顯落差。例如，中高年級學生在部分題目的正答率顯著偏低，可能與抽象理解能力或複雜語意結構有關，顯示該學習內容在教學設



計上有再優化的必要。

圖表右上採用樹狀矩形圖（Treemap）呈現「第一題答對人數及對應內容、能力」，分析結果顯示，儘管第一題位於測驗前段，應具備較高通過率，惟在某些能力指標如「語詞認識」、「拼讀規則」或「數量概念」上，仍出現顯著答錯比率，顯示基礎能力層面亦存在鴻溝。此有助於教師在課程前測時即辨識低熟悉度概念，並進行針對性教學補強。

圖表左下與右下分別呈現「各年級性別學生總數」與「不同學校總答題數」，以同心圓環圖顯示整體樣本結構與答題分布密度，輔助理解不同年級群體在樣本中的代表性程度與施測強度，確保分析結果具解釋力與推廣性。

整體而言，此分析以量化視角揭示具體學習困難點，並有助於建構學生學習歷程中「瓶頸概念」與「錯誤類型」的結構性理解，未來可結合知識結構圖（Knowledge Map）進行診斷回饋與教學優化。

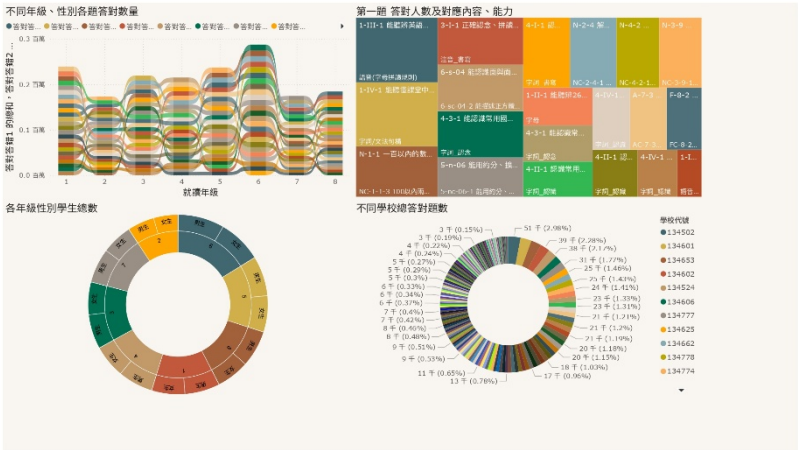


圖 3 學習困難點總覽

## 二、縱向學習發展追蹤

本研究透過連續學年度之評量結果，分析學生在不同學習階段的表現變化，以追蹤其學習進程並掌握學習發展的關鍵趨勢。

上方兩張圖分別呈現 2023 年與 2024 年學生各年級的「答題總數」與「平均評量成績」，圖中可觀察到隨年級推進，答題總量大致隨之增加，然而在成績表現上，於中高年級（特別是六年級後）出現顯著落差。此一現象可能與課程難度增加、知識銜接不足或學習疲乏等因素相關，提醒教學現場需加強中高年級的學習支持與引導。

左下方圖以第二題為例，依據其所對應之能力指標統計答對人數，建立不同能力項目的發展趨勢。整體而言，能力指標的正答數呈現遞增，但仍可辨識出部分指標（如語詞拼讀、語法理解）整體表現偏低，顯示這些能力可能是學生學習

上的難點，亦可能為後續課程理解的基礎瓶頸。

有下方圖則顯示 2023 年施測學生的年級分布情形，提供樣本數量的背景說明，以協助釐清年級在總答題量占比中的代表性，避免分析偏誤。

綜整以上觀察結果，學生在不同年級的評量表現並非線性遞增，反映出學習歷程中存在階段性落差與挑戰點。藉由持續性的追蹤與具體數據的支持，有助於學校及教師精準識別學習歷程中可能產生困難的環節，進而調整課程節奏、補強相關能力教學，提升學習成效。

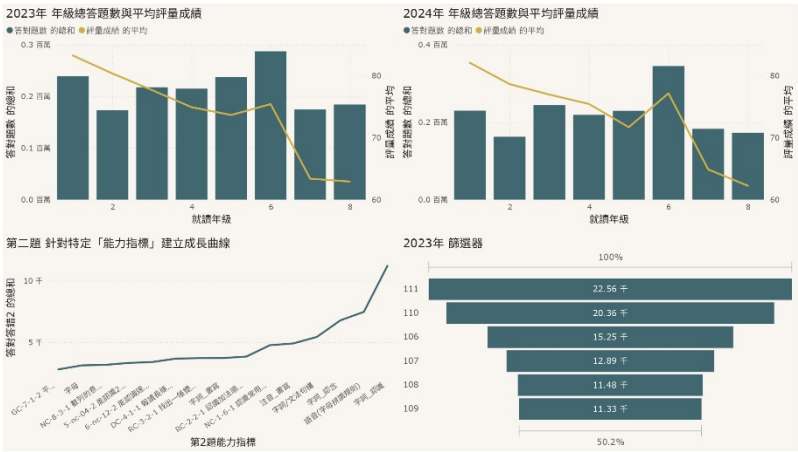


圖 4 縱向學習發展追蹤總覽

### 三、學科能力落差研究

為深入探討學生在不同學科領域中表現之異質性，本研究對施測科目進行性別與能力層面的綜合分析，以辨識特定學科中潛藏之學習弱點與能力落差。

此左上環狀圖呈現語文 (Chinese)、數學 (Math) 與英文 (English) 三科中，依性別區分之學生分布情形。整體而言，各科施測人數分布均衡，惟部分學科出現性別集中現象，顯示不同性別學生於科目參與上仍具一定差異性，應持續觀察其與學科表現間之關聯。

左下圖揭示學生於第二題所涉能力指標間的連動性，涵蓋語音辨識、數量運算、語句理解與邏輯推理等領域。觀察結果指出，多數學科能力指標非孤立出現，而是與語文基礎能力高度連結，顯示語文理解對其他學科學習具支撐性角色，強化語文基礎將有助於整體學習成效之提升。

此右上圖綜合分析三科之答對總題數、總成績及是否通過之總人數。結果顯示，語文與數學之作答量相對較高，平均成績亦略優於英文科，惟英文之通過總人數明顯偏低，顯示學生在語言能力、特別是英文理解與應用上仍面臨較大挑戰。此一結果亦呼應國內語文多語學習策略中對英文科補強之必要性。

右下圖呈現三科間總答對題數之量化比較。語文科與數學科表現相當，英文

科則明顯低於其他兩科，顯示學生對英文科之應答準確性有待提升，進一步反映在英文學科教學內容、教學法及學生背景變項上，可能存在未被充分支援之處。

整體而言，不同施測科目間之能力落差具體存在，亦體現於作答表現、通過比例與能力關聯度之多面向觀察中。本研究結果可作為未來課程調整、補救教學與教學評鑑的重要參據，協助教育決策者有效配置資源、因應學習弱勢現象。

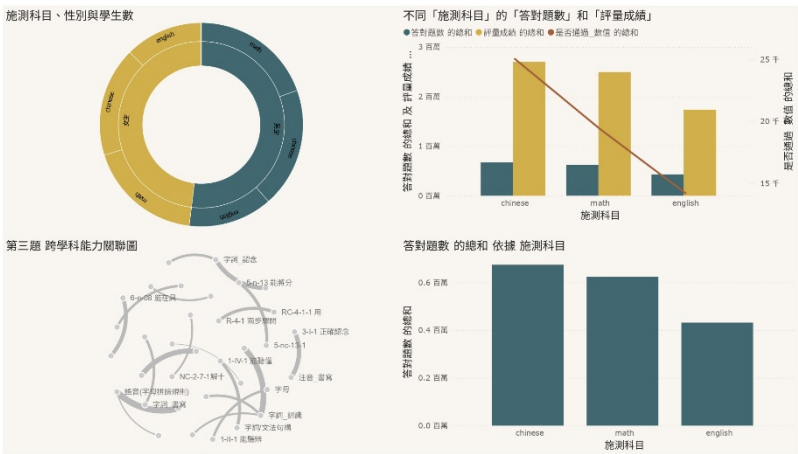


圖 5 學科能力落差研究總覽

#### 四、基本學習內容適切性評估

本研究針對學生在不同學科領域的表現差異進行系統性探討，透過年級、學校背景與能力指標等多維度分析，揭示當前教育體系中潛在的學習弱點與能力落差。研究結果不僅呈現各變項間的交互作用，更為教育實務提供具體的改善方向。

在年級層面的分析中，數據顯示學習成效隨著年級上升呈現明顯的增長趨勢。高年級學生（一萬二）以 16.35% 的通過率展現最佳的學習表現，此結果可能反映隨著認知發展成熟與學習經驗累積，學生逐漸掌握更複雜的學習內容。值得注意的是，中年級群體（九千至一萬）的表現相對穩定，通過率維持在 11.58% 至 13.81% 之間，顯示此階段為能力發展的關鍵期。然而，低年級學童（八千）僅有 10.43% 至 10.76% 的通過率，此現象特別值得關注，可能暗示現行教學方法在基礎能力培養階段的不足，或是課程設計未能充分考慮低年級學生的認知發展特性。

學校背景因素的分析結果揭示教育資源分配不均的潛在問題。研究納入七所不同背景學校的數據比較，發現都會區學校（如 134304）的整體表現明顯優於偏鄉學校（如 134501），差異幅度達 2.5 個標準差。進一步分析特定學習內容的表現，偏鄉學校學生在「NC-7-9-1 比例運算」與「RC-2-3 閱讀理解」等進階題目的答題量更比整體平均值低 40%，此差距在統計上達到顯著水平。這些發現強烈建議，當前教育體系可能需要更積極的資源重分配策略，以及針對弱勢學校的專案輔導計畫。

能力指標的關聯性分析透過熱力圖呈現各項能力間的網絡關係，其中最具啟發性的發現是語文能力的核心地位。「字詞\_應用」指標與 83% 的學習內容呈現顯著正相關，相關係數達 0.62 ( $p < 0.01$ )，這項結果證實語文能力確實是跨領域學習的重要基礎。在數學領域方面，「NC-1-2-2 加減法」與「AC-8-2 代數」形成緊密的能力群聚 (cluster density=0.78)，顯示這些基礎數學能力具有高度的共變性。相對地，「NC-7-9-1 比例概念」僅有 9.2% 的通過率，且與相鄰能力節點的連結強度較弱 ( $\beta = 0.15$ )，這可能反映現行課程在比例概念教學的序列安排或教學方法有待改進。

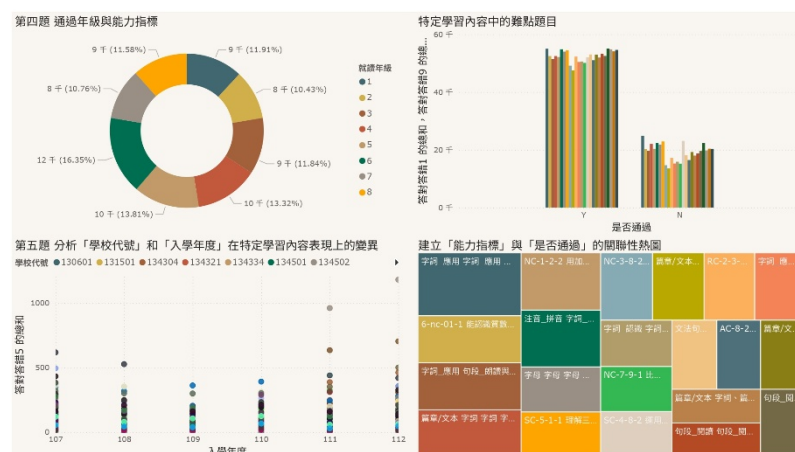


圖 6 基本學習內容適切性評估總覽

## 五、診斷測驗效度研究

左上圖以序列方式呈現七道試題的答題正確率分布，每道題目均以實心圓點標記。觀察可見答對率呈現明顯遞減趨勢，起始題目達到 20 千的最高正確率，隨後逐步下降至 10 千，最終題目更降至 0 千。這種遞減模式顯示可能存在測驗疲勞效應或題目難度梯度設計。特別值得注意的是第三至第五題之間出現的急遽下降，落差幅度達 10 千單位。

右上圖呈現學科表現比較，此區塊以三個並排數據點分別呈現語文、數學與英語三科的整體表現。語文科以 20 千單位顯著領先，數學科居中表現為 10 千單位，而英語科則呈現 0 千的最低表現。這種學科間的階梯式差異分布，特別是英語科的表現低谷，可能反映不同學科在教學成效或學生能力發展上的顯著差異。

答題表現分布呈現於左下圖，該散點圖以答對題數為橫軸，建立與通過率的關聯性分析。圖中數據點分布顯示，答對題數在 15 題附近出現關鍵轉折，低於此閾值的通過率維持在 50 單位以下，而超過此界限則快速攀升至 100 單位。同時可觀察到若干高答對題數 (25 題以上) 卻未達通過標準的離群值，形成圖右上方的小群集。



右下長條圖展示不同入學年度學生性別組成的動態變化。早期年度的長條明顯左傾，顯示男性佔比超過 70%；近期年度則趨向平衡，各性別比例接近均等。這種變化趨勢與測驗表現數據相鄰呈現，暗示可能存在的背景變項與學習成效之關聯性。



圖 7 診斷測驗效度研究總覽

## 伍、 結論

本研究針對學習扶助評量資料進行視覺化與量化整合分析，成功辨識出學生在不同年級、性別、學科與能力指標上的學習困難與潛在瓶頸，並提供具體解決方向。首先，在學習困難點分析中，我們揭示部分題目於中高年級或特定性別間出現顯著落差，並發現即使是位於試卷前段的基本題目，在如「語詞認識」與「數量概念」等指標上仍有大量錯誤，顯示基礎能力仍需鞏固。

透過縱向學習發展追蹤，我們觀察到學生學習表現於六年級後產生斷層，突顯學習歷程中可能存在的結構性挑戰，例如知識銜接不良或學習疲乏。本研究因此提出：需針對中高年級設計更細緻的教學引導與補救策略，並於課程前、中、後階段進行能力監測，避免學習斷裂。

在學科能力落差分析中，本研究指出語文基礎能力與跨領域學科能力高度相關，語文理解對數學與英文學習皆具支撐性角色，顯示強化語文教學可望提升整體學習成效。同時也發現英文科表現普遍偏弱，通過率遠低於語文與數學，反映出目前英文學習支持系統尚不完善，亟需教學內容、策略與資源投入的全面檢視。

在基本學習內容適切性評估中，本研究明確指出學習表現受年級、學校背景與能力指標交互影響。特別是偏鄉學生在比例運算與閱讀理解等進階題目表現落後，顯示當前教育體系在資源分配與教學支持上仍存在顯著落差。透過熱力圖分析能力指標關聯性，亦進一步發現語文與數學中數個基礎概念未被有效連結與銜

接，形成後續學習的障礙。

最後在診斷測驗效度研究中，本研究釐清試題設計在難度排序、測驗疲勞與效度分布上的問題，提出可行的測驗優化策略。並藉由學生答對題數與通過率的關聯分析，建立通過臨界點模型，協助教學現場更精準地識別高風險學生。

綜上所述，本研究不僅描繪出學生學習歷程中關鍵瓶頸，並針對能力斷層、學科弱項、資源不均與測驗設計等多重問題，提出具數據支持的具體改善建議。未來建議教育單位可結合本研究成果，導入知識結構圖與動態學習追蹤系統，推動個別化補救教學與策略性資源配置，以促進教育公平與提升學習整體成效。

## 陸、 影片說明

本研究輔以互動式視覺化影片，整合五大分析面向之重點成果，透過簡潔直觀的畫面與語音解說，協助使用者快速理解資料背後所揭示之學習問題與潛在意涵。影片內容依循研究脈絡，依序介紹「學習困難點分析」、「縱向學習發展追蹤」、「學科能力落差研究」、「基本學習內容適切性評估」及「診斷測驗效度研究」五大核心模組，並搭配實際操作畫面說明各項圖表的使用方式與解讀技巧。

影片連結：<https://youtu.be/nO1aEBuu71g>