



大東文化大学中国語学科 第30回學術シンポジウム

2025/11/30

生成式AI 應用於台灣華語文能力測驗(TOCFL) 初級試題編寫之研究

陳怡攸 政治大學華語文教學博士學位學程 博士生/ 淡江大學華語中心

李國印 淡江大學前瞻教育領導與科技管理博士班/ 淡江大學華語中心

謝奇懿 高雄師範大學華語文教學研究所 副教授



研究大綱

一、前言：研究背景與目的

二、文獻回顧

華語文能力測驗(TOCFL) 測驗介紹
聽讀測驗入門基礎級(Band A) 題型

三、生成試題-兩種生成模式

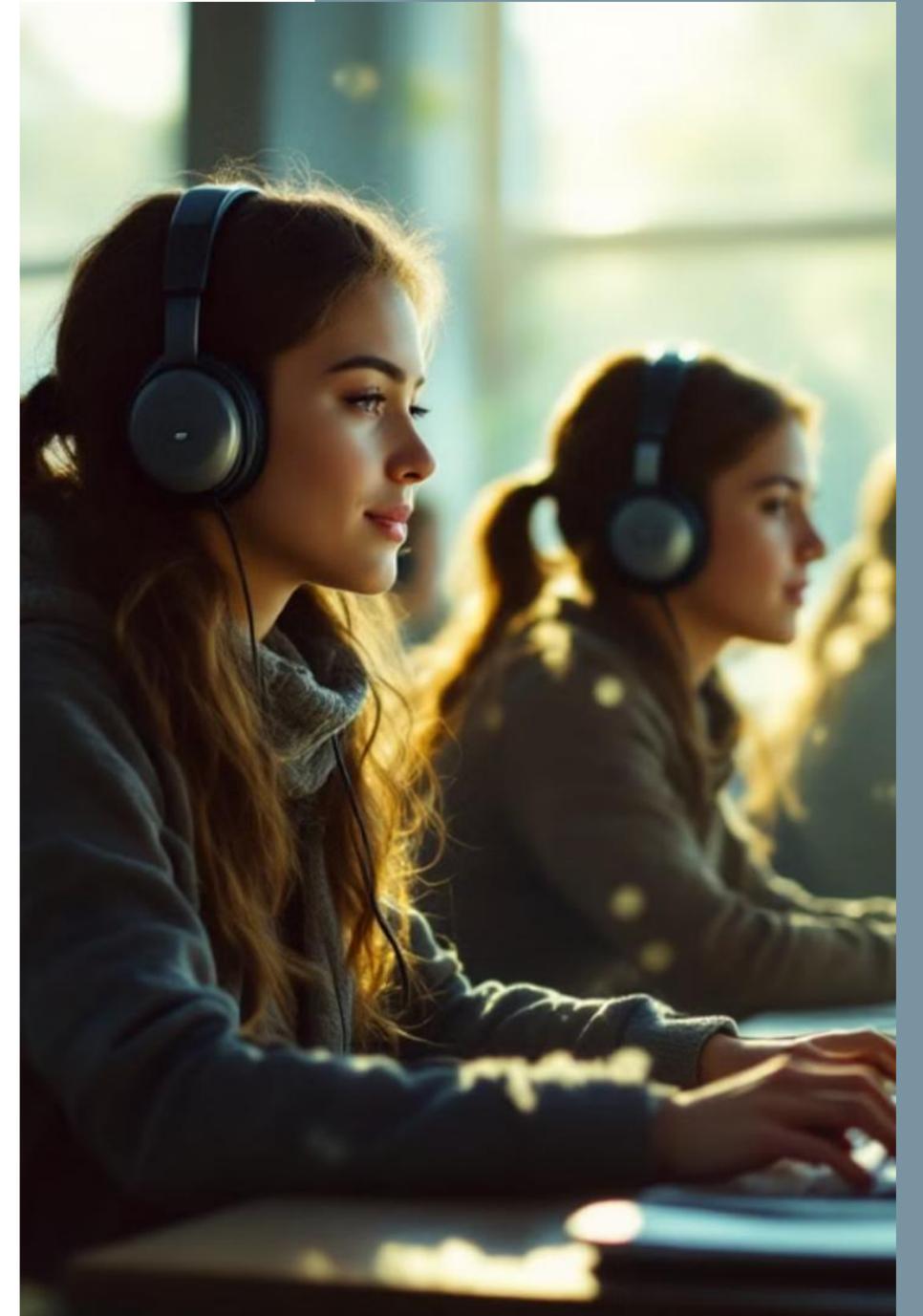
模式(一) 試題展示-聽力測驗、閱讀測驗生成與調整

模式(二) 試題展示-聽力測驗、閱讀測驗生成與調整

四、研究結果

五、研究研論

六、參考文獻



一、前言：研究背景與目的

淡江華語中心申請教育部計畫，內部開設題庫工作坊及編製試題：

召集教師→命題、審修試題→整理試題(加入繪圖)、組卷、製卷→施測→數據分析→修正試題

工作坊模式	工作坊內容	實施過程
聽力測驗工作坊-12週、 閱讀測驗工作坊-12週	召集教師，於每週討論會後進行命題、繳題，隔週共同審修試題。 工作坊帶領教師：陳怡攸 計畫指導：高師大謝奇懿副教授	1.每週開一次工作坊，帶領教師群討論命題原則及每個題型的出題重點。 2.華語教師一週後繳題（當週所討論的題型），於隔週討論及修改題目。 3.修改過的題目加入繪製圖檔後，再次進行討論。 4.試題進行審修後製卷(內審-華語中心老師，外審-外校華語文系所教授)。

▼開設工作坊、老師共同參與命題、討論試題



▼製作模擬試題、規劃施測時間

淡江大學華語中心
華語文能力測驗
模擬考

日期：2020/7/20 (一)
時間：上午場 9:30-12:00
下午場 13:30-16:00

TOCFL
BAND-A
MOCK
TEST

請扫描下方二维码
完成線上報名



報名或其他問題
請至 B1-007
洽陳思好老師

歡迎參加7月25日正式考試的同學報名!

▼進行班級施測、數據分析及修正試題



一、前言：研究背景與目的

華語中心編製TOCFL模擬試題的困境

1. 華語教師不足



- 華語教師擔負教學、行政等工作，缺少能長期參與命題、審修試題的老師。

2. 華語教師缺少訓練

- 華語教師缺少編寫「標準化能力測驗」試題的訓練和經驗。

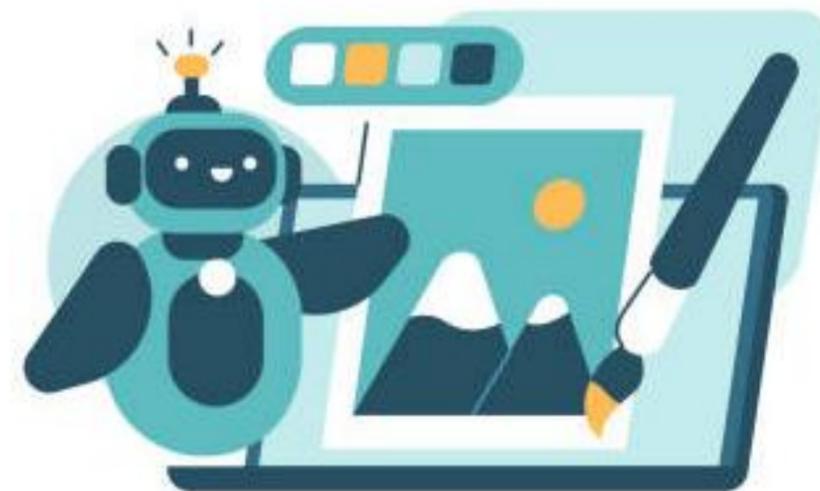
3. 耗時



- (培訓教師)、教師命題、審修試題、錄製聽力音檔、繪製試題圖片(**Band A聽讀測驗-約160張**)、組卷、製卷

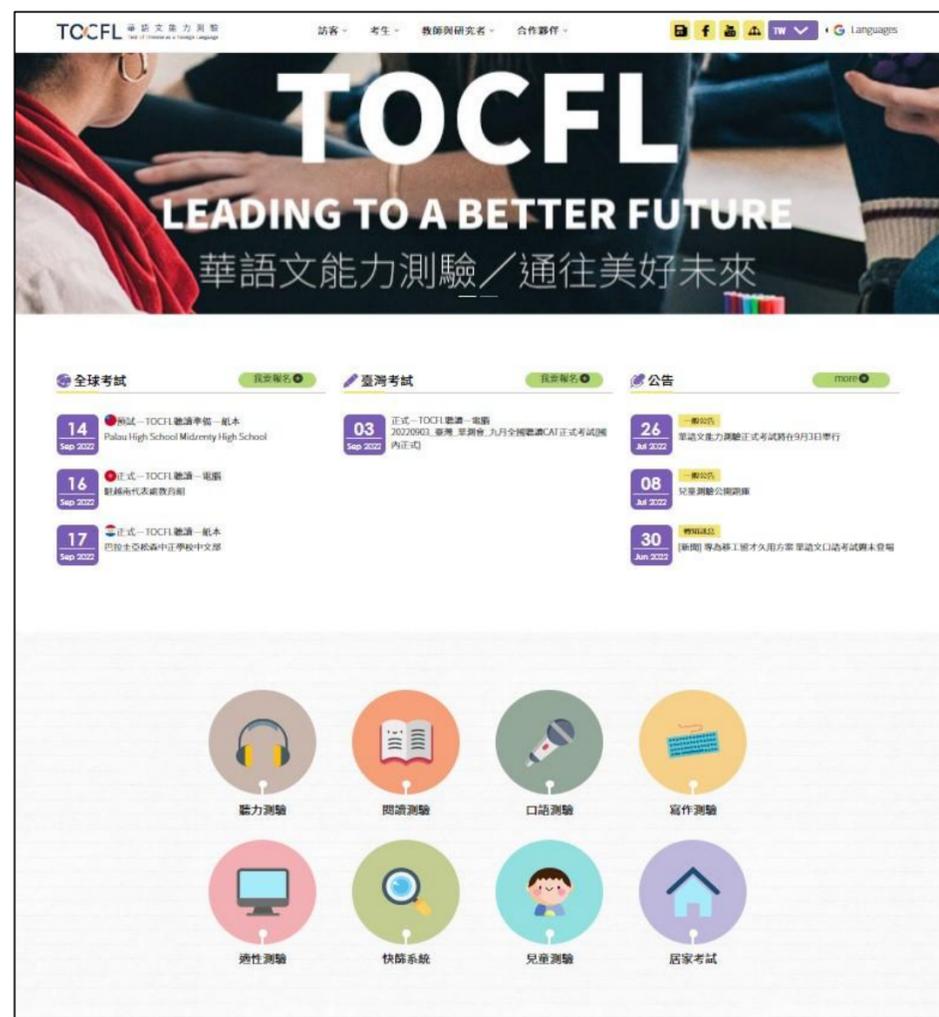
4. 缺少經費

- 命題、審修題教師(內部、外部人員)、繪圖人員、行政人員、製卷、施測等費用



本研究的目的旨在探討運用生成式AI在華測入門基礎級(Band A)聽力、閱讀測驗試題設計中的應用，是否具有可行性和有效性。

二、文獻回顧：測驗簡介



臺灣華語文能力測驗

Test of Chinese as Foreign Language
(簡稱TOCFL)

<https://tocfl.edu.tw/tocfl/>

簡介：華語文能力測驗(TOCFL)

③ 類型：標準化的語言能力測驗

能力測驗	標準化測驗
<ul style="list-style-type: none">-為了瞭解一個人的語言能力是否能符合某種規範或等級。-測驗重點為在一個受控制、一致且標準化流程下，去測量學習者能用語言完成什麼事情，滿足什麼樣的溝通需求。-這類型的測驗多是由正式的機構去執行。	<ul style="list-style-type: none">-根據測驗的編製標準化程序而成。✓ 試題編製✓ 測驗實施✓ 信度與效度驗證✓ 解釋分數

- 台灣的華語文能力測驗屬於標準化的語言能力測驗，旨在檢測華語學習者在日常生活中的語言使用能力。
- 測驗內容針對各種日常生活情境所設計，題材真實且多元化，內容涵蓋各種來源，例如：廣告、公告、信件、便條、天氣預報等不同體裁。所選短文也均衡涵蓋文學、科學、藝術、天文、歷史等不同領域主題。

(資料來源：華測會網站)

二、文獻回顧：測驗形式

臺灣華語文能力測驗

Test of Chinese as Foreign Language (簡稱TOCFL)

<https://tocfl.edu.tw/tocfl/>



CBT

須選擇適合的測驗等級

根據既有的試題安排

固定題數
聽力、閱讀各50題

考生可於固定時間內
隨意跳題作答或檢查

TOCFL

報考等級

題目安排

測驗題數

閱讀測驗
檢查機制

CAT

不須選擇適合的測驗等級

電腦即時估計考生能力並
安排適合的試題

約35題

無法返回檢查或修改答案

TOCFL 電腦測驗形式

- 電腦化測驗 Computer-Based Test (CBT)
(2006年開辦，但僅在台灣考場施測，大部分的海外考場仍採取紙筆形式的測驗，逐漸增加電腦施測的考場)
- 電腦化適性測驗 Computerized Aptitude Test (CAT)
(2017年推出，2022年台灣考場全數改為CAT測驗)

二、文獻回顧：測驗形式

臺灣華語文能力測驗在日本施測的情況

<https://tocfl.jp/>



台湾華語（中国語）能力検定試験 - TOCFL公式サイト
TOCFL 華語文能力測驗 Test of Chinese as Foreign Language

試験のお申し込み

TOCFLについて 受験について Q&A 問題集のご紹介 お問い合わせ先

TOCFL検定試験日
2026年 1月18日(日)

試験のお申し込み

アカウント作成方法
申し込み手順

申込期間：
2025年10月31日(金)~2025年12月3日(水)

① INFORMATION インフォメーション

2025.11.13 TOCFL受験者や取得希望者を対象に國家華語測驗推動工作委員會が日本で初の就職フェアを開催いたします。台湾華語・中国語人材を求める台湾企業9社が参加し、日台での就職機会拡大を目指します。

- 至2023年，華測考試在海外47個國家進行過施測(試題設有繁體字、簡體字兩種版本供考生選擇)。
- 2025年華測在日本施測的現況：
一年12個月皆舉辦考試，場次分散在9個城市(札幌、仙台、東京、橫濱、名古屋、大阪、廣島、福岡、沖繩)，以紙筆形式為主或為電腦測驗形式。
- 測驗證書用途
 - 1.申請台灣的獎學金
 - 2.來台灣留學
 - 3.來台灣就業
 - 4.美國的「雙語徽章(Seal of Biliteracy)」是一項官方認證，授予在英語及至少一種其他語言上展現熟練能力的高中畢業生。在美國採認華測作為雙語徽章標準的州別(華測作為評估中文能力的標準測驗之一)，目前有26州。

二、文獻回顧：測驗分級

測驗歷史背景

(1981年日本開辦中國語檢定試驗

1990年中國開辦HSK考試)

1996年美國建立外語教學ACTFL能力指標

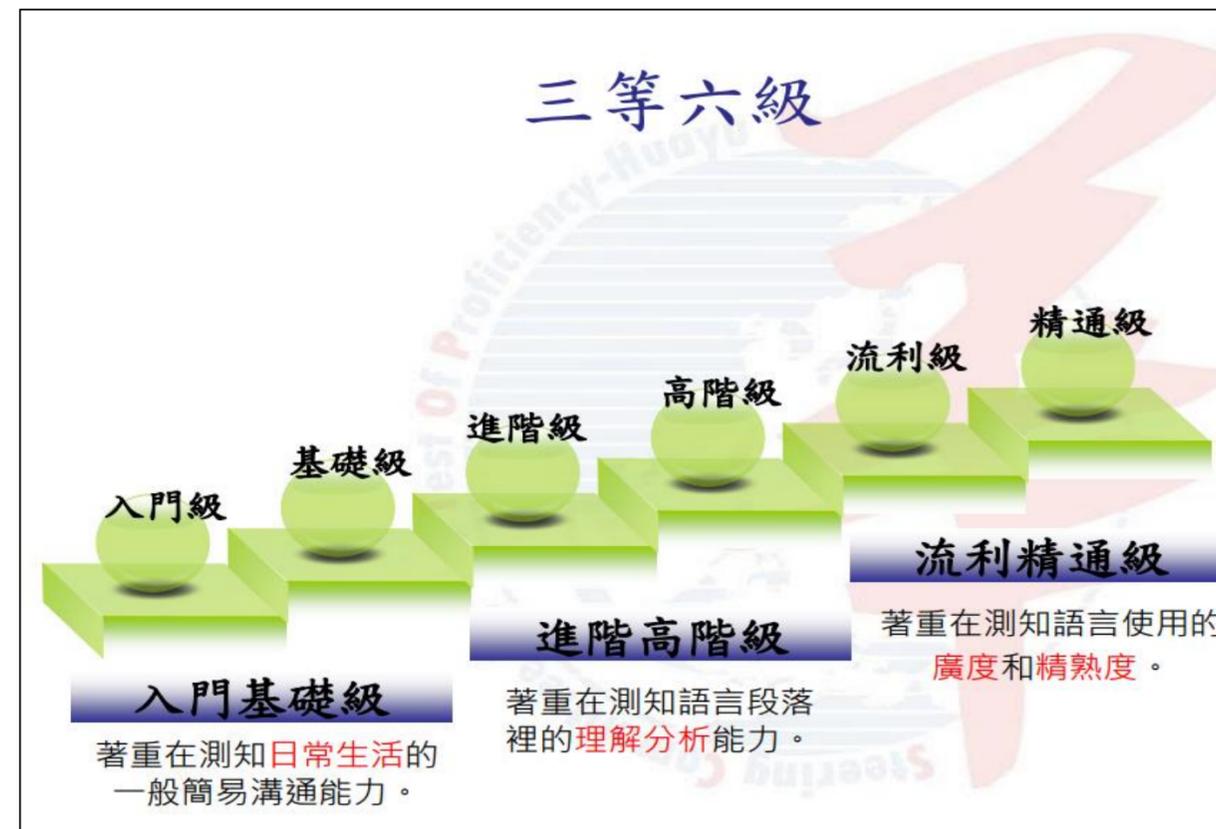
2001年歐洲建立歐洲共同語言參考架構(CEFR)能力指標

2003年台灣開辦華語文能力測驗(紙本測驗)

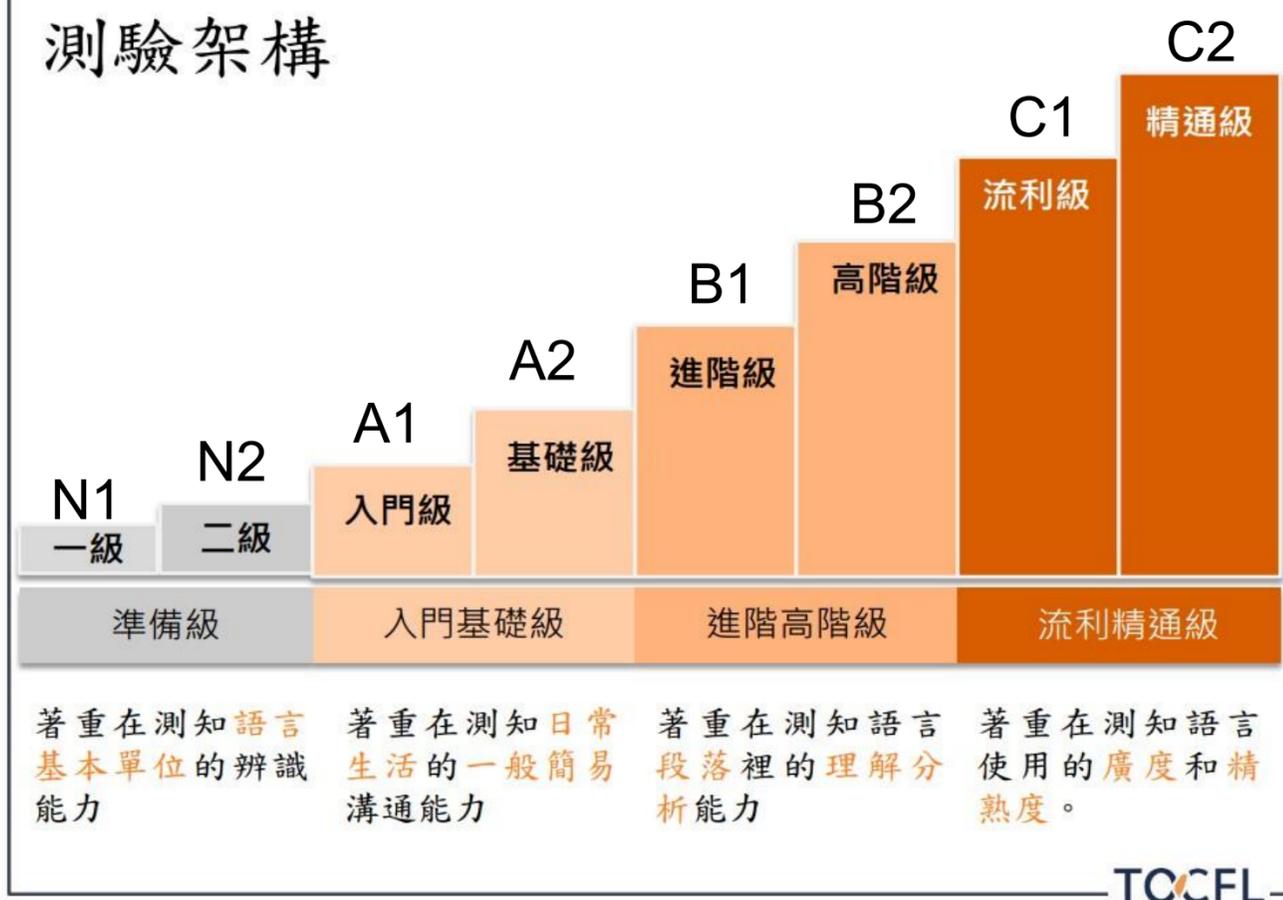
2006年開辦電腦化測驗(CBT)

2017年開辦電腦化適性測驗(CAT)

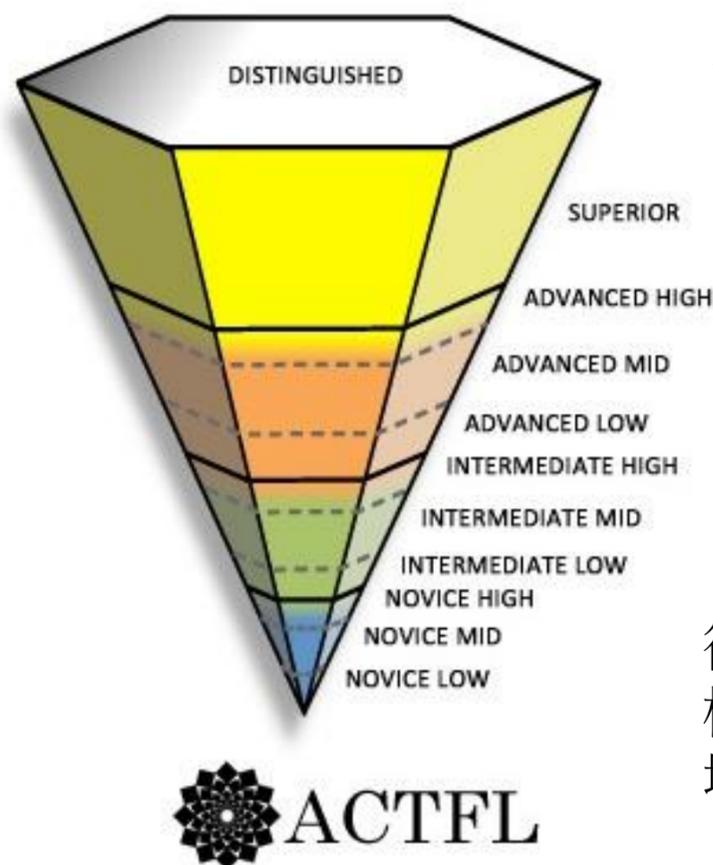
2020年教育部建置COCT系統、20201年提出TBCL指標



測驗架構



原參照CEFR指標，設有6個等級



後加入美國外語教師協會能力指標(ACTFL Proficiency Guidelines)，增為8個等級。

二、文獻回顧：測驗分級

測驗歷史背景

(1981年日本開辦中國語檢定試驗

1990年中國開辦HSK考試)

1996年美國建立外語教學ACTFL能力指標

2001年歐洲建立歐洲共同語言參考架構(CEFR)能力指標

2003年台灣開辦華語文能力測驗(紙本測驗)

2006年開辦電腦化測驗(CBT)

2017年開辦電腦化適性測驗(CAT)

2020年教育部建置COCT系統、2021年提出TBCL指標

參考來源：

日本中国語檢定試験(中檢) <https://www.chuken.gr.jp/>

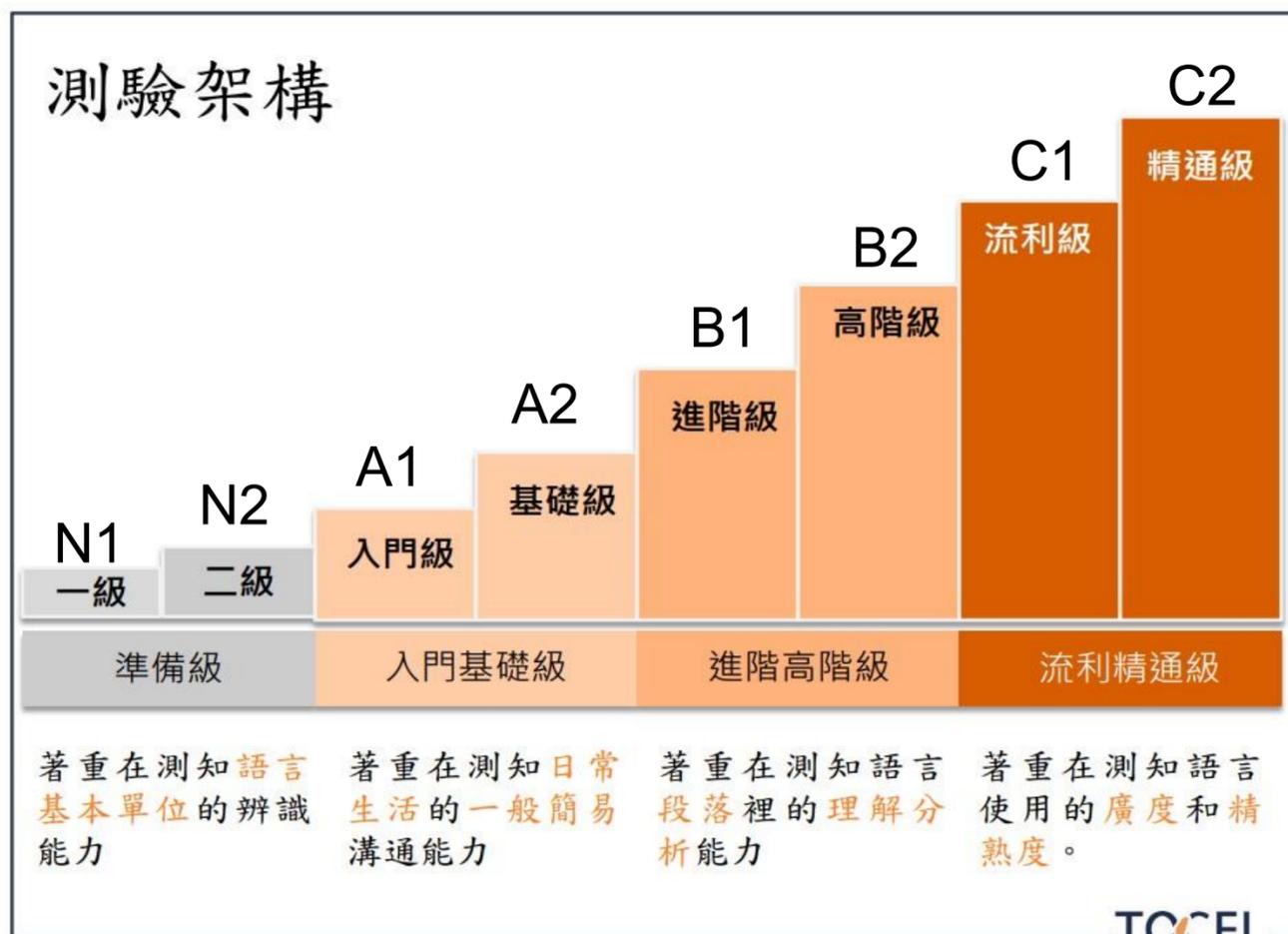
HSK 中文考試服務網 <https://www.chinesetest.cn/HSK>

<https://www.hskj.jp/>

難易度對應表

中国語の資格試験との難易度対応表です。試験の目的も出題形式も異なるため厳密な比較はできませんが、およその目安を記します。

中檢	TOCFL	HSK	TECC	CATTI國際	CEFR
1級	精通級	7~9級(高級) 2022年	900-1,000	A級	C2
準1級	流利級		750-899	B級	
2級	高階級	6級	550-749	C級	C1
3級	進階級	5級	400-549	D級	B2
4級	基礎級	4級	300-399	E級	B1
準4級	入門級	3級		-	A2
-	準備級	2級	200-299	-	A1
-	-	1級	0-199	-	-



二、文獻回顧： 測驗對應華語文能力基準與指標

- 臺灣華語文語料庫與能力基準整合應用系統 <https://coct.naer.edu.tw/>
(Corpus of Contemporary Taiwanese Mandarin, COCT)

華語文語料庫 與 能力基準整合應用系統 (試用版)

- 教材編輯輔助系統
- 作文錯別字自動批改系統(試用版)
- 語義場關聯詞查詢系統
- 國教院分詞系統
- ▼ 華語文分級標準查詢系統
- ▼ 索引典系統
- 詞表比較工具
- 語料庫覆蓋率統計系統

臺灣華語文能力基準 華語文教材原型 文件下載 歷年得獎作品

〈華語文語料庫與能力基準整合應用系統〉國家教育研究院版權所有 ©2023 National Academy for Educational Research, R.O.C. All rights reserved

- 臺灣華語文能力基準 TBCL (Taiwan Benchmarks for the Chinese Language)
- COCT系統
 - 1.教材編輯輔助系統
 - 2.作文錯別字自動批改系統(試用版)
 - 3.語義場關聯詞查詢系統
 - 4.國教院分詞系統
 - 5.華語文分級標準查詢系統
 - (1) 漢字分級標準檢索系統
 - (2) 詞語分級
 - (3) 語法點分級
 - (4) 基礎詞彙檢索系統
 - (5) 類詞綴檢索系統
 - 6.索引典系統
 - (1) 華英雙語索引典系統
 - (2) 國教院索引典系統
 - (3) 華語中介語索引典系統
 - 7.詞表比較工具
 - 8.語料庫覆蓋率統計系統

二、文獻回顧： 測驗對應華語文能力基準與指標

「臺灣華語文能力基準」(TBCL) 能力指標-聽力1-3級

第一級 (準備級)	第二級 (A1級)	第三級 (A2級)
整體描述： 能理解並運用與個人生活相關的簡單字詞。	整體描述： 能理解與個人及日常生活有關的字詞、短語及短句。能運用簡單字詞與句子與他人交流。	整體描述： 能理解公共、學校及職場等周遭場域的簡短對話、討論與指示。能以簡單完整的句子敘述經歷，與人進行對話及簡單討論。
聽力： 能聽懂以簡單字詞表達之個人生活相關訊息。	聽力： 能在語速緩慢、內容清晰的情況下，聽懂對話、指示、視聽媒體之簡單字詞、短語及短句。	聽力： 能在語速緩慢、內容清晰的情況下，聽懂以句子表達之對話、指示、非正式討論與視聽媒體之內容。
1-1 能聽懂簡單的數量名稱（數字、價錢、時間）。	2-1 能聽懂以短語及短句陳述之個人相關問題。	3-1 能聽懂以句子或簡短段落表達之熟悉主題的對話。
1-2 能聽懂以簡單字詞介紹的個人相關資訊（國籍、住址、學經歷、家庭狀況）。	2-2 能聽懂以短語及短句之宣布與指示。	3-2 能聽懂以句子或簡短段落敘述的通知或指示。
1-3 能聽懂生活常用的簡單詞彙及短語（打招呼、謝謝、對不起）。	2-3 能聽懂以短語及短句陳述之日常生活訊息。	3-3 能大致聽懂視聽媒體報導的事件。
	2-4 能聽懂視聽媒體之簡單字詞、短語及短句。	3-4 能聽懂以句子或簡短段落所表達之熟悉主題的非正式討論。

二、文獻回顧： 測驗對應華語文能力基準與指標

「臺灣華語文能力基準」(TBCL) 能力指標-閱讀1-3級

第一級 (準備級)	第二級 (A1級)	第三級 (A2級)
閱讀： 能讀懂與個人生活相關的簡單字詞與指示。	閱讀： 能讀懂以簡單句子描述的與個人相關的訊息與指示。	閱讀： 能理解以句子及簡短段落呈現與個人周遭場域有關的敘述及指示。
1-1 能讀懂簡單的數量名稱（數字、價錢、時間）。	2-1 能讀懂以短語或短句描寫的個人相關訊息與指示。	3-1 能理解以完整句子或簡短段落所書寫的私人信件。
1-2 能讀懂個人相關的簡單字詞（國籍、住址、學經歷、家庭狀況）。	2-2 能讀懂生活場域中的短語或短句指示（街道、飯館、火車站的路牌和指示標誌）。	3-2 能讀懂公共場合中以句子或簡短段落呈現之常用設施的操作與指示（如何用機器買票）。
1-3 能讀懂社交的問候語與感謝語（請、謝謝、對不起）。	2-3 能讀懂以短語或短句呈現的學習訊息。	3-3 能從句子或簡短段落之書面資料理解基本訊息（宣傳單、菜單、廣告、佈告欄或海報上的時間、地點或聯繫方式）。
		3-4 能理解以完整句子或簡短段落所呈現有關學習或工作的敘述。

二、文獻回顧：測驗編制理念、情境範疇

測驗編理念：

華測採取歐洲共同語言參考架構(CEFR)的理念—語言使用者/學習者為社群份子，在一定狀況下、特殊環境裡或某些特定行為範疇中，利用語言能力完成特定的任務或議題。

試題內容整體涵蓋「領域主題」、「語言任務」和「溝通策略」(體現語言功能)。



情境範疇	
1. 個人資料	10. 餐飲、烹飪
2. 日常、起居	11. 公共服務
3. 職業	12. 安全
4. 休閒、娛樂	13. 自然社會
5. 交通、旅遊	14. 社會
6. 社交、人際	15. 文化
7. 身體、醫療	16. 情緒、態度
8. 教育、學習	17. 科技
9. 購物、商店	

▪ TBCL 設定17類「情境分類」

- 工作事件交辦、職務調動
- 工作環境、工作權益、工安、工傷、勞資糾紛、勞工政策及其相關議題

• 技術報告《說情話境-華語文詞語情境分類》

二、文獻回顧：文本範疇、題型任務活動

測驗編理念：

華測採取歐洲共同語言參考架構(CEFR)的理念—語言使用者/學習者為社群份子，在一定狀況下、特殊環境裡或某些特定行為範疇中，利用語言能力完成特定的任務或議題。

試題內容整體涵蓋「領域主題」、「語言任務」和「溝通策略」(體現語言功能)。

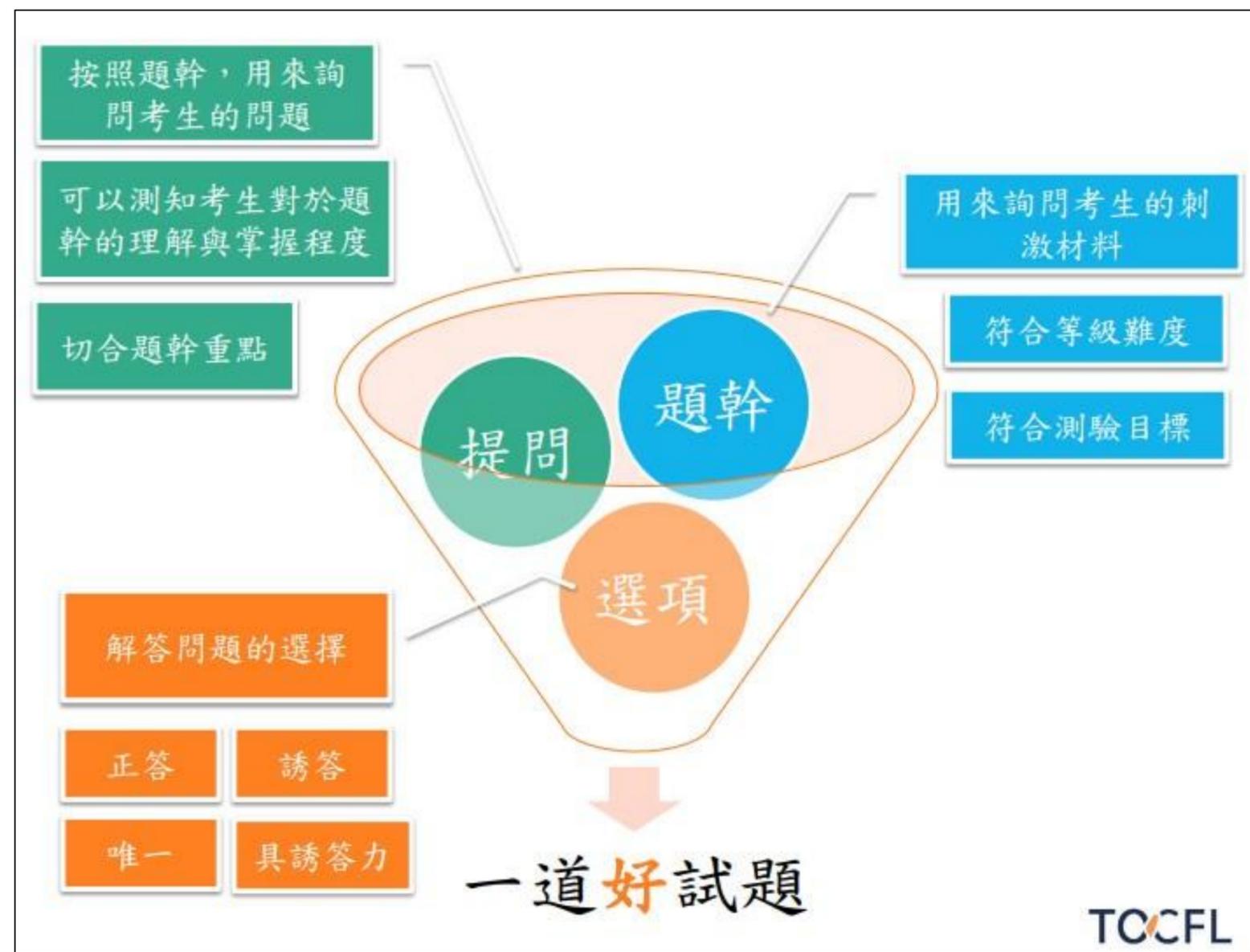
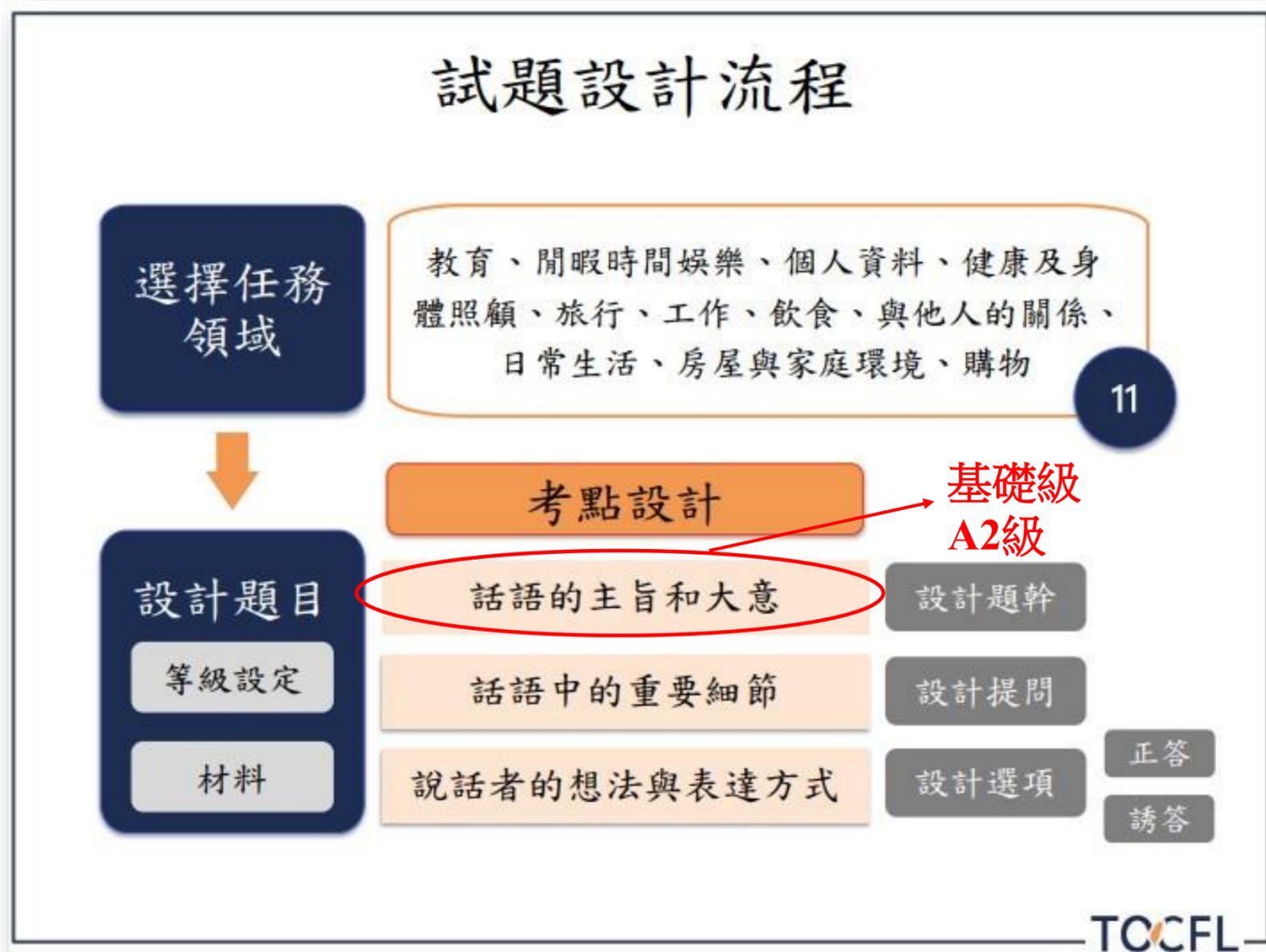
各等級文本範疇



任務活動

- 聽力活動
 - 了解母語者間的交談
 - 聽宣布及說明指導
 - 聽語音媒體與錄音
 - 做為現場聽眾一員的聽力表現
 - 整體性聽力理解
- 閱讀活動
 - 閱讀書信
 - 閱讀說明指導
 - 導向閱讀
 - 為資訊論證而閱讀
 - 整體性閱讀理解

二、文獻回顧：試題編寫流程及原則



- 一、題幹設計原則
- 二、提問設計原則
- 三、選項設計原則

- 四、公平性原則、
- 五、原創性原則

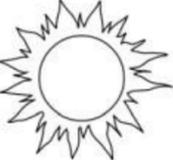
二、文獻回顧：試題編寫流程及原則

華測模擬試題編寫的訓練依據及參照華測會及臺灣華語文能力基準(TBCL)資源，試題概念考量命題理念、命題規範、主題領域、詞語分級表、語法點分級表、模擬試題、編寫範例等。

二、文獻回顧：聽力測驗題型

TOCFL CBT 入門基礎級聽力測驗具有圖片題幹或選項的題型：聽力(一)、(二)、(三)

資料來源：華測會官網模擬試題

一、看圖回答 10題 題幹圖片 1張*10題=10張	二、問答理解 15題 選項圖片 3張*15題=45張	三、對話理解 15題 選項圖片 3張*15題=45張
<p>例題如下 Example：</p> <p>你會看到一張圖片 You will see one picture：</p>  <p>接著，你會聽到一個問題和三個選項。 Then, you will hear one question and three possible answers.</p> <p>第一題：眼鏡在哪裡？</p> <ul style="list-style-type: none">(A) 在盒子裡面。(B) 在盒子下面。(C) 在盒子上面。	<p>你會看到三張圖片 You will see three pictures：</p> <p>(A)  (B)  (C) </p> <p>接著，你會聽到一問一答的對話。 Then, you will hear a dialogue between two speakers.</p> <p>第二題</p> <p>A：昨天下雨了嗎？ B：是啊，還下得很大呢！</p>	<p>你會看到三張圖片 You will see three pictures：</p> <p>(A)  (B)  (C) </p> <p>接著，你會聽到一段對話和一個問題。 Then, you will hear a dialogue and a question.</p> <p>第三題</p> <p>女：你什麼時候去開會？ 男：下個星期。 女：你要坐火車去嗎？ 男：坐火車太慢了，我坐飛機去。</p> <p>Q：請問這位先生怎麼去開會？</p>

二、文獻回顧：聽力測驗題型

TOCFL CBT入門基礎級聽力測驗僅有文字的題型：聽力(四)

資料來源：華測會官網模擬試題

四、閱讀理解10題
選項僅有文字，沒有圖片

例題如下 Example：

第四題

(男)：小美，妳的電話號碼是幾號？

(女)：上次不是才給你嗎？

(男)：抱歉，我不小心把那張紙給弄丟了。

(女)：把你的手機拿出來吧！記在你的手機裡，這樣才不會再弄丟。

(男)：小美為什麼要這位先生把手機拿出來？

- (A) 她想知道對方有沒有手機
- (B) 她想借對方的手機打電話
- (C) 她不相信對方把手機弄丟了
- (D) 她想把自己的電話號碼給對方

二、文獻回顧：閱讀測驗題型

TOCFL CBT 入門基礎級閱讀測驗具有圖片題幹或選項的題型：閱讀(一)、(二)、(三)

資料來源：華測會官網模擬試題

<p>一、單句理解 15題 題幹圖片 3張*15題=45張</p>	<p>二、看圖釋義 15題 圖片題+材料題(題數不等) 1張*15題=15張</p>	<p>三、選詞填空 5題 情境圖片 1張</p>
<p>例題如下 Example：</p> <p>你會看到一句話和三張圖片： You will see one sentence and three pictures.</p> <p>1. 她在電影院門口等人。</p> <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <p>(A)</p>  </div> <div style="text-align: center;"> <p>(B)</p>  </div> <div style="text-align: center;"> <p>(C)</p>  </div> </div>	<p>你會看到一張圖片和三個句子： You will see one picture and three sentences.</p> <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  </div> <div style="text-align: left;"> <p>2.</p> <p>(A) 她正在買豬肉。 (B) 她還沒買水果。 (C) 她已經買了菜了。</p> </div> </div> <div style="margin-top: 20px;">  <p>宿舍停水通知</p> <p>男生宿舍： 7月2日上午八點~下午六點</p> <p>女生宿舍： 7月3日上午八點~下午六點</p> <p>有任何問題請找： 03-3761-1080 張小姐 日和大學 總務處 2016.6.28</p> </div> <div style="margin-top: 20px; text-align: left;"> <p>(A) 日和大學全校停水兩天。 (B) 女生宿舍一共停水十二小時。 (C) 男生宿舍比女生宿舍早一天停水。</p> </div>	<p>你會看到一張圖片和五個句子： You will see one picture and five sentences.</p> <div style="text-align: center; margin-bottom: 20px;">  </div> <ol style="list-style-type: none"> 1. 張先生_____她吃法國菜。 (A)請 (B)送 (C)買 2. 他們_____喝酒。 (A)沒有 (B)只有 (C)正在 3. 餐桌上有_____。 (A)筷子 (B)碗 (C)盤子 4. 這家餐廳有_____人。 (A)太少 (B)很多 (C)多少 5. 這家餐廳的菜____好吃____便宜。 (A)也...也...(B)越...越...(C)又...又...

二、文獻回顧：閱讀測驗題型

TOCFL CBT 入門基礎級閱讀測驗僅有文字的題型：閱讀(四)、(五)

資料來源：華測會官網模擬試題

四、完成段落 10題 (共2篇短文/題組，每篇有6個短語，須填入5個/題)	五、閱讀理解 5題 (共5篇短文，200字以內)												
<p>例題如下 Example：</p> <p>你會看到一段短文和六個選項： You will see one short passage and six possible answers.</p> <p>我的室友跟我一樣大，<u> (1) </u>。因為我們都很喜歡旅行，所以去年我們<u> (2) </u>。那裡的天氣不冷也不熱，<u> (3) </u>。旅行的時候，我們看到很多漂亮的花朵，照了<u> (4) </u>，玩得很開心。暑假快到了，我們正在想<u> (5) </u>！</p> <table border="1" data-bbox="593 1352 1059 1684"><tbody><tr><td>(A)</td><td>可是常常下雪</td></tr><tr><td>(B)</td><td>都是二十歲</td></tr><tr><td>(C)</td><td>很多照片</td></tr><tr><td>(D)</td><td>也不常下雨</td></tr><tr><td>(E)</td><td>這次要去哪裡玩</td></tr><tr><td>(F)</td><td>一起去歐洲玩</td></tr></tbody></table>	(A)	可是常常下雪	(B)	都是二十歲	(C)	很多照片	(D)	也不常下雨	(E)	這次要去哪裡玩	(F)	一起去歐洲玩	<p>例題如下 Example：</p> <p>你會看到一段短文、一個問題和四個選項： You will see one short passage, one question and four possible answers.</p> <p>每天我要到許多地方去，也會遇到很多人。有些人喜歡叫我「左轉」、「右轉」、「停」；有些人會把髒東西留在我的車上。不過也有一些不錯的人，可以從他們身上學到很多東西，所以我也交了好幾個朋友。真是什麼樣的人都有啊！</p> <p>5. 寫文章的人可能是做什麼工作的？</p> <p>(A) 教書 (B) 賣汽車 (C) 開計程車 (D) 打掃房子</p>
(A)	可是常常下雪												
(B)	都是二十歲												
(C)	很多照片												
(D)	也不常下雨												
(E)	這次要去哪裡玩												
(F)	一起去歐洲玩												

三、生成試題 (模式一)

- 使用ChatGPT付費帳號：

淡江大學華語中心購買團隊公用帳號，提供模擬試題人員、數位組人員、高師大華語所研究生(與研究者合作編寫試題、測試模式(二)之模組)使用。

- 模式(一)

研究者利用華測會官網資料，於ChatGPT中輸入不同模擬試題的指示詞，並進行調修，探討生成式AI「仿寫」試題的情況。

- 使用材料：

(1)華測會官網模擬試題5輯，共有聽力250題、閱讀題250題 (在指示詞夾帶題本圖片的截圖)、(2)華測會八千詞表 (在指示詞夾帶Excel檔案)

- 提示詞：1.範例試題、2.題型選項、題幹、字數、詞表等級、3.領域主題

請以專業的華語教師編寫視題的角度，生成TOCFL華語文能力聽力測驗「看圖回答」的題目。以下是一題範例。

提問題目：這位先生怎麼了？

選項(A) 他餓了。

選項(B) 他生病了。

選項(C) 他生氣了。

這個題目有一張圖片，考生根據圖片的內容，在提問的題目以後，選出正確的答案。題目為三個選項，其中只有一個選項的句子是對的，並且生成一張圖片。另外兩個選項的句子跟提問的題目有關，但是必須是錯的，這個題目的正確答案是選項B。

每個句子的字數在10個字以內，詞彙的等級請用檔案裡準備一級、準備二級、入門級和基礎級的詞彙。這個題目的主題為「身體健康與照護」，請生成相似的另一個題目，要有一個提問、三個選項和一張圖片。

1.



請以專業的華語教師編寫試題的角度，生成TOCFL華語文能力聽力測驗「看圖回答」的題目。以下是一題範例。

提問題目：這位先生怎麼了？

選項(A) 他餓了。

選項(B) 他生病了。

選項(C) 他生氣了。

這個題目有一張圖片，根據圖片的內容，在提問題目以後，選擇正確的句子。選項為三個選項，其中只能有一個選項的句子是對的，並且生成一張圖片。另外兩個選項的句子跟提問的題目有關，但是必須是錯的。每個句子在10個字以內，請用以下檔案裡準備一級、準備二級、入門級和基礎級的詞彙。這個題目的主題為「身體健康與照護」，請生成相似的另一個題目。

< 2/2 >



我將使用這些表格資料中的詞彙來生成符合「TOCFL 華語文能力聽力測驗」的「看圖回答」題目，並符合「身體健康與照護」主題。以下是題目與選項的設計：

題目：

這位小朋友怎麼了？

選項：(A) 他感冒了。

(B) 他在運動。

(C) 他餓了。

根據題目和圖片的內容，請選擇正確的答案。

三、生成試題 (模式一)

- 華測會官網模擬試題 (第3輯第1題)
一、看圖回答

1.



題目：這位先生怎麼了？
選項(A) 他餓了。
選項(B) 他感冒了。
選項(C) 他生氣了。

- ChatGPT 生成的圖片及選項
生成的第一題

接下來，我會生成一張圖片，描繪一位小朋友坐在床上，用手摸著額頭，明顯感覺不舒服，看起來像是感冒了。 [-]



已完成此題的圖片設計。請參考上方的圖片，用於「看圖回答」題目。根據圖片，小朋友感覺不舒服，正符合選項 (A)「他感冒了」的描述。這題目測試考生對健康狀況的理解，並提供其他合理但錯誤的選項，以增加挑戰性。

如果還需要其他調整或題目，隨時告訴我！

題目：這位小朋友怎麼了？
選項(A) 他感冒了。
選項(B) 他在運動。
選項(C) 他餓了。

生成情況：

選項和模擬試題重複性高。

繼續輸入其他的指示詞：

- 1.請再生成一題。
- 2.請換其他選項，句子的詞彙等級是A1/ A2。
- 3.請把「...」改為「...」。(說明詞語或圖片內容)，
- 4.請把圖片改為... (說明圖片風格)。

三、生成試題 (模式一)

- 華測會官網模擬試題 (第3輯第1題)

一、看圖回答

1.



題目：這位先生怎麼了？
選項(A) 他餓了。
選項(B) 他感冒了。
選項(C) 他生氣了。

- ChatGPT 生成的圖片及選項 (每個模擬試題嘗試生成3~4題)

生成的第一題



題目：這位小朋友怎麼了？
選項(A) 他感冒了。
選項(B) 他在運動。
選項(C) 他餓了。

生成的其他題目



題目：這位男孩怎麼了？
選項(A) 他在喝水。
選項(B) 他生病了。
選項(C) 他在玩遊戲。



題目：這位先生的健康出了什麼問題？
選項(A) 他感冒了，需要吃藥。
選項(B) 他在健身房運動得很開心。
選項(C) 他生病了，所以要去上班。

生成情況：

- 有時會出現超綱詞(超過等級的詞)。
- 有時會產出語法、語序有誤，或有語義邏輯問題的句子。

三、生成試題 (模式一)

聽力大題第二大題~第三大題，調修指示詞的過程中增加更多說明。

- 提示詞：**1.範例試題**、**2.題型內容**、**3.領域主題**、**4.對話角色**、**對話情境**、**語用功能**、**5.單輪對話改為雙輪對話**

1.試題的組成說明：

例如：第二大題「對話的下面不能有提問的問題，對話女生的回答就是選項的答案」。

2.試題對話內容的說明：

- 1.「對話角色為....」例如：「對話的男生、女生是朋友」
- 2.「對話情境為...」例如：「女生幫忙男生解決問題」
- 3.單輪對話改為雙輪對話，例如「請生成跟這一題...(主題)類似的題目，但是對話要改成雙輪對話，也就是要有1.男生說話、2.女生說話、3.男生說話、4.女生說話，一共四句對話。對話的下面要有提問的問題和不同的選項。」

16.

(A)



(B)



(C)



請以華語教師編寫試題的角度，生成TOCFL華語文能力聽力測驗「問答理解」的題目。以下是一題範例。

對話男生：這東西要 100 元！明美，我想買這個，可是還少 30 元，怎麼辦？

對話女生：我先用信用卡幫你付錢吧！

選項(A) 幾個零錢

選項(B) 幾張鈔票

選項(C) 一張信用卡

這個題目有一個單輪對話，一個男生和一個女生一問一答。根據對話的內容，選擇正確的句子，選項為三個選項，每個選項都要生成一張圖片，一共有三張。其中只能有一個選項是對的，另外兩個選項要跟對話提到的內容有關，但是必須是錯的。對話的長度，每個句子在25字以內，請用簡單的初級詞彙。這個題目對話的男生、女生是朋友，女生幫了男生的忙。對話的情境是「幫忙別人」，請生成相似的另一個題目。



對話：

男生：我明天要搬家，可是只有我一個人，怎麼辦？

女生：沒問題，我明天可以來幫你搬東西。

問題：女生明天會怎麼幫男生？

選項：

A. 幫忙打掃房子。

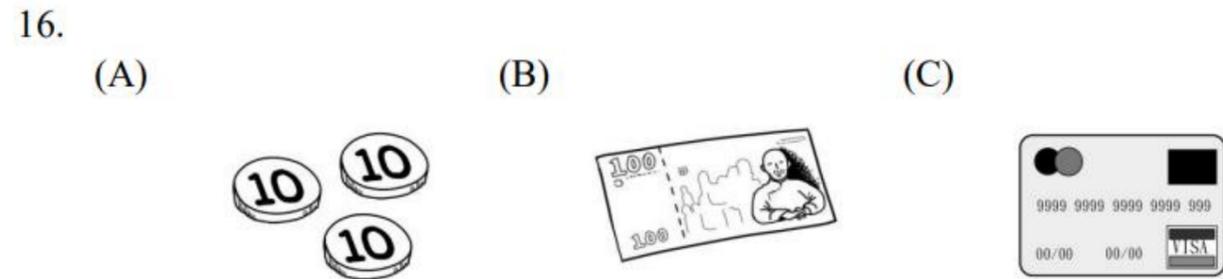
B. 幫忙搬東西。

C. 幫忙煮飯。

接下來將會生成每個選項的對應圖片。

三、生成試題 (模式一)

- 華測會官網模擬試題 (第3輯第16題)
- ## 二、問答理解

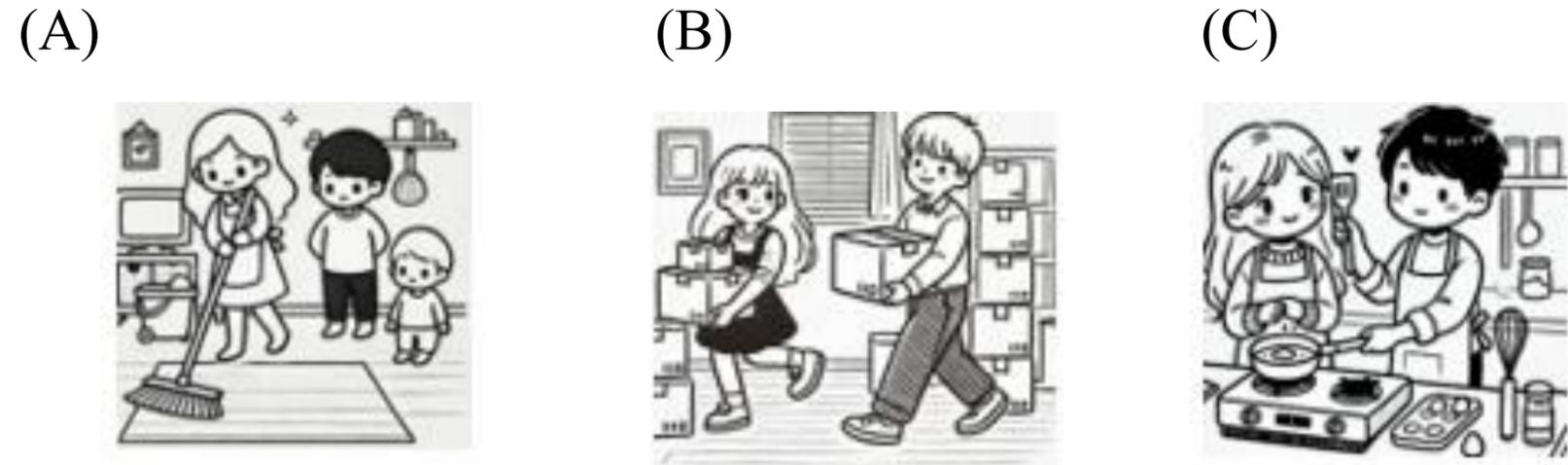


先生：這東西要100元！明美，我想買這個，可是還少30元，怎麼辦？

小姐：我先用信用卡幫你付錢吧！

- 選項(A) 幾個零錢
- 選項(B) 幾張鈔票
- 選項(C) 一張信用卡

- ChatGPT 生成的圖片及選項
- ## 生成的題目



先生：我明天要搬家，可是只有我一個人，怎麼辦？
小姐：沒問題，我明天可以來幫你搬東西。

- 選項(A) 幫忙打掃房子
- 選項(B) 幫忙搬東西
- 選項(C) 幫忙煮飯

生成情況：對話常出現指示詞情境中關鍵的詞語，例如：「對話情境為女生幫忙男生解決問題」。
生成的選項常和動作相關。

三、生成試題 (模式一)

- ChatGPT 生成的圖片及選項
生成的題目-第二大題問答理解

(A)

(B)

(C)



先生：我明天要搬家，可是只有我一個人，怎麼辦？
小姐：沒問題，我明天可以來幫你搬東西。

- 選項(A) 幫忙打掃房子
- 選項(B) 幫忙搬東西
- 選項(C) 幫忙煮飯

- ChatGPT 生成的圖片及選項
生成的題目用於第三大題對話理解

(A)

(B)

(C)



先生：我下個星期要搬家，可是只有我一個人。
小姐：搬家很麻煩，需要很多時間。
先生：對啊，而且我還沒收好東西。
小姐：別擔心，我可以來幫你一起搬東西。

問題：小姐會怎麼幫先生？

- 選項(A) 幫忙搬東西
- 選項(B) 幫忙修電腦
- 選項(C) 幫忙買東西

三、生成試題 (模式一)

閱讀測驗

- 華測會官網模擬試題 (第4輯第19題)

二、看圖釋義(1) 圖片題 (請選出符合圖片內容的描述)



- 選項(A) 這裡幾乎坐滿了學生。 →描述圖片場景
- 選項(B) 這些學生正在慶祝開學。 →描述人物動作
- 選項(C) 他們幾個人就要畢業了。 →推論描述人物的狀態

選項特點：每個選項都和圖片有所關聯，但是重點都不同，句子的句構也不同(V滿了、正在V、就要V了)。

- ChatGPT 生成的圖片及選項生成的題目



- 選項(A) 他們在草地上玩雪。
- 選項(B) 他們在公園裡野餐。
- 選項(C) 他們在房間裡烤肉。

生成情況：容易重複同樣或接近的句構，且以描述人物的動作(人在某地做某事)的句子為多。

三、生成試題 (模式一)

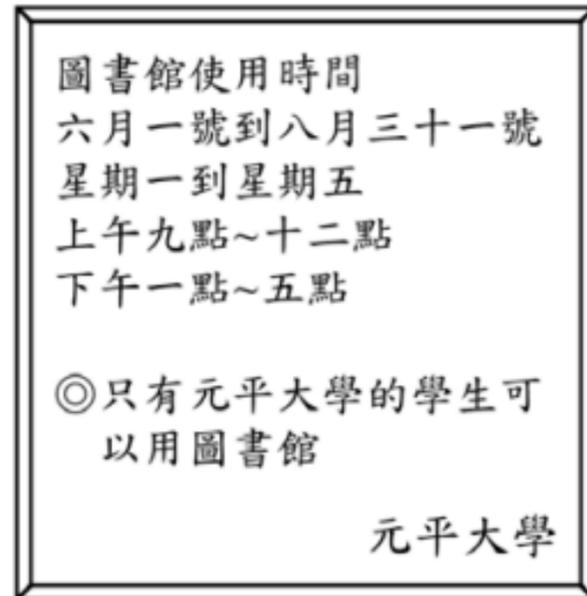
閱讀測驗

- 華測會官網模擬試題 (第4輯第26題、第5輯第22題)

二、看圖釋義(2) 材料題



- 選項(A) 這個活動能換到水果。
- 選項(B) 這些活動從中午開始。
- 選項(C) 活動的時間有一個禮拜。



- 選項(A) 誰都可以進去這間圖書館。
- 選項(B) 圖書館中午休息一個小時。
- 選項(C) 只有一個月可以進去圖書館。

- 提示詞：1.範例試題、2.材料題本的內容、3.材料題的類型及功能

請以專業的華語教師編寫視題的角度，生成TOCFL華語文能力閱讀測驗「看圖釋義」的題目。以下是一題範例。

題目的內容：一張圖書館的公告，上面寫著：

「圖書館使用時間
六月一號到八月三十一號
星期一到星期五
上午九點~十二點
下午一點~五點

只有元平大學的學生可以用圖書館」

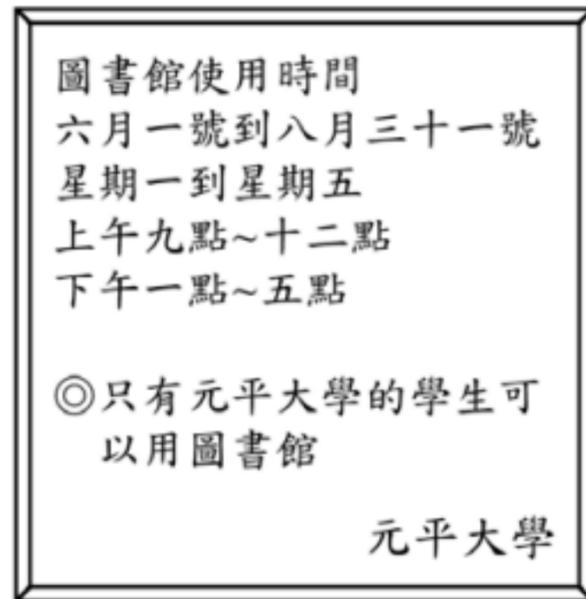
- 選項(A) 誰都可以進去這間圖書館。
- 選項(B) 圖書館中午休息一個小時。
- 選項(C) 只有一個月可以進去圖書館。

這個題目有一張公告，根據公告上寫的內容，選擇正確的選項。題目一共有三個選項，其中只有一個選項的句子是對的。另外兩個選項的句子跟公告的內容有關，但是必須是錯的，這個題目的正確答案是選項B。公告的字數在60個字以內，詞彙的等級請用檔案裡入門級和基礎級的詞彙。這個題目的主題為「公共場所的公告」，請生成相似的另一個題目，要有一個公告和三個不同的文字選項。

三、生成試題 (模式一)

閱讀測驗

- 華測會官網模擬試題 (第5輯第22題)
二、看圖釋義(2) 材料題

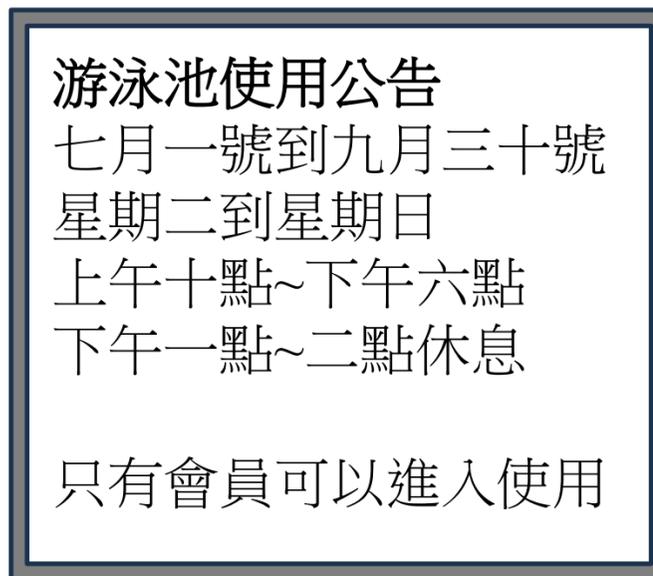


- 選項(A) 誰都可以進去這間圖書館。
- 選項(B) 圖書館中午休息一個小時。
- 選項(C) 只有一個月可以進去圖書館。

- ChatGPT 生成的圖片及選項
生成的題目嘗試了兩種情況：

1.生成文字

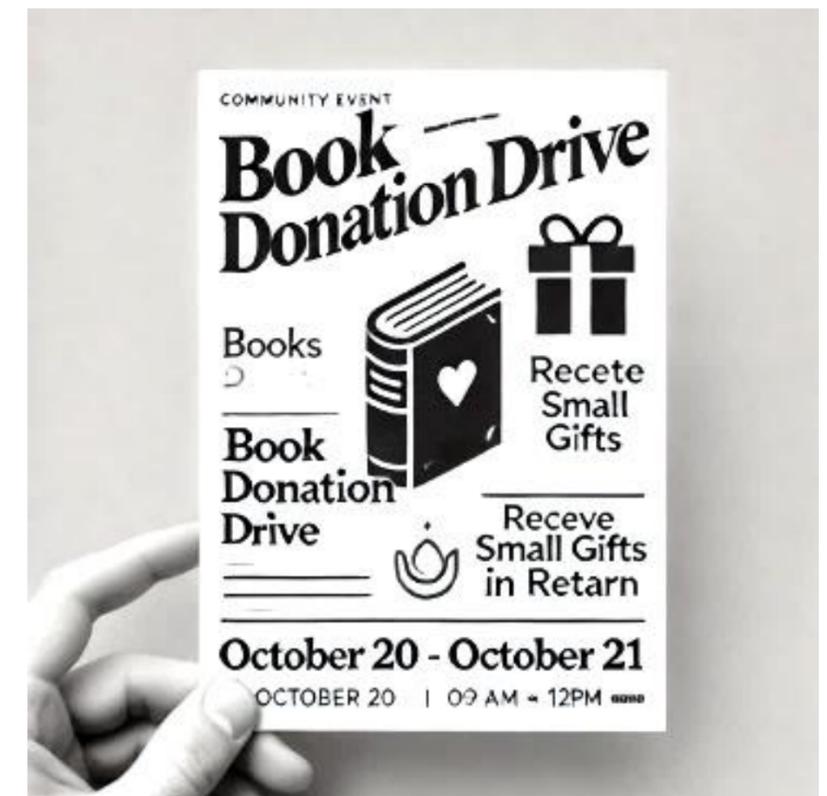
模仿性高，過於接近。無法直接生成圖片



- 選項(A) 任何人都可以進入游泳池。
- 選項(B) 游泳池每天從下午一點休息到二點。
- 選項(C) 游泳池的使用時間只有七月一個月。

2.生成材料類型的圖片

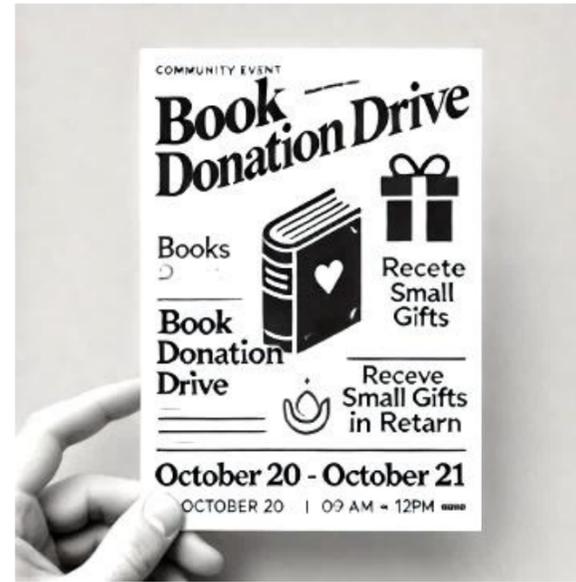
(指示詞：生成一張宣傳單)，圖片會包含英文、其他語言文字或無法辨識的文字內容



三、生成試題 (模式一)

- ChatGPT 生成的圖片及選項嘗試的做法：

- 1.讓ChatGPT將圖檔的文本內容轉為純文字內容，(指示詞：請幫我把圖片的內容翻譯成中文) 調修題幹和選項。
- 2.以修改過的文本內容重製題幹圖片。



學校活動公告
活動名稱：交換舊書
活動日期：10月20日~10月21日
活動時間：上午9點~中午12點
活動內容：帶一本舊書來，就能換一個小禮物。歡迎大家來參加！

請幫我設計下面學校活動公告的宣傳單插圖
學校活動公告
活動名稱：交換舊書
活動日期：10月20日~10月21日
活動時間：上午9點~中午12點
活動內容：帶一本舊書來，就能換一個小禮物。歡迎大家來參加！



這是為學校活動公告「交換舊書」設計的插圖，展示了戶外愉快的場景，包括舊書交換桌、小禮物和快樂的學生，並以溫暖的陽光和裝飾氣球作背景，非常適合用於活動宣傳單！希望您滿意！

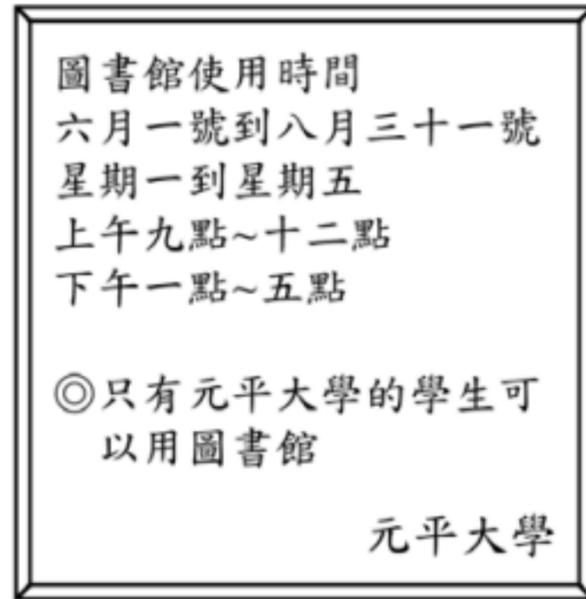
學校活動公告
活動名稱：交換舊書
活動日期：10月20日~10月21日
活動時間：上午9點~中午12點
活動內容：帶一本舊書來，就能換一個小禮物。歡迎大家來參加！



三、生成試題 (模式一)

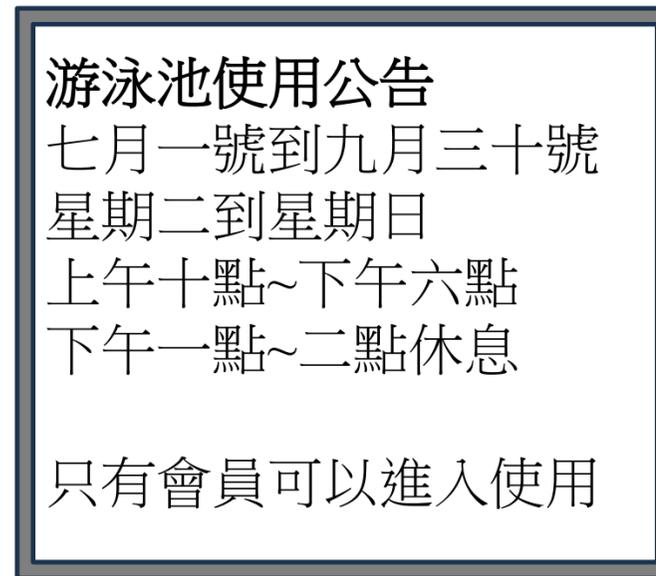
閱讀測驗

- 華測會官網模擬試題 (第5輯第22題)
二、看圖釋義(2) 材料題



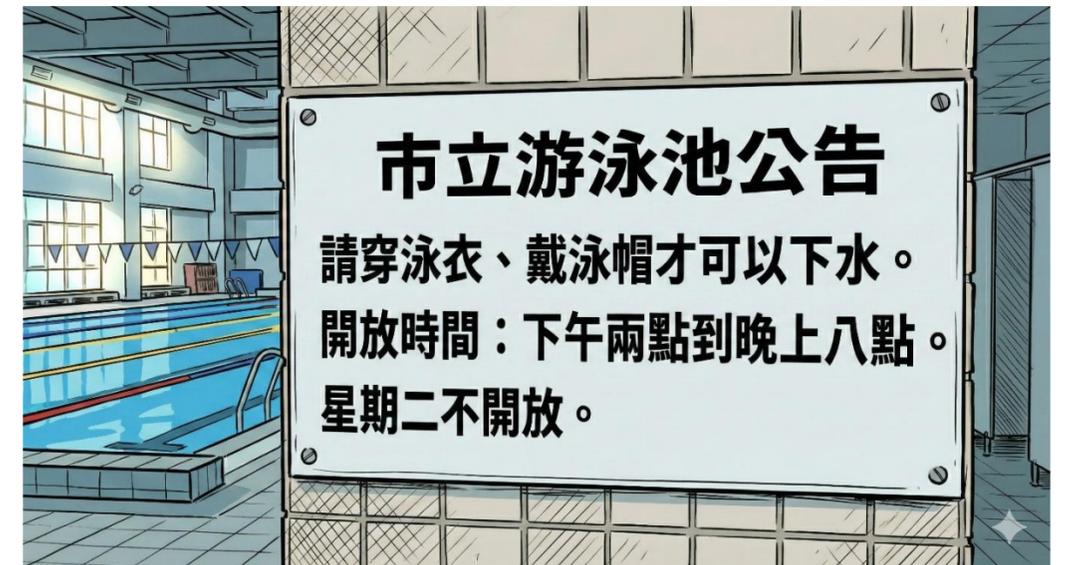
- 選項(A) 誰都可以進去這間圖書館。
- 選項(B) 圖書館中午休息一個小時。
- 選項(C) 只有一個月可以進去圖書館。

- 1.(2024年) ChatGPT 生成的材料題
生成文字：模仿性高，過於接近。
無法直接生成圖片



- 選項(A) 任何人都可以進入游泳池。
- 選項(B) 游泳池每天從下午一點休息到二點。
- 選項(C) 游泳池的使用時間只有七月一個月。

- 2.(2025年) Gemini 生成材料題
生成文字：可轉換其他內容
直接生成圖片中具有中文字的內容



- 選項(A) 每天下午都可以去游泳。
- 選項(B) 只要穿泳衣就可以下水了。
- 選項(C) 這個游泳池每個星期休息一天。

三、生成試題 (模式一)

- (2024年) ChatGPT 生成材料題

嘗試的做法：

- 1.讓ChatGPT將圖檔的文本內容轉為純文字內容，調修題幹和選項。
- 2.以修改過的文本內容重製題幹圖片。

學校活動公告

活動名稱：交換舊書

活動日期：10月20日～10月21日

活動時間：上午9點～中午12點

活動內容：帶一本舊書來，就能換一個小禮物。歡迎大家來參加！



- (2025年) Gemini 生成材料題

利用已有生成試題的內容來迭代指示詞，生成主題相同、內容不同的試題。

指示詞：生成一張學校活動的宣傳單，以下是一題範例。請生成類似風格，但是不同內容的活動，並提供三個合適的選項。



學校活動公告

活動名稱：
校園打掃日

活動日期：12月5日（星期五）

活動時間：上午8點～11點

活動內容：幫忙打掃學校旁邊的公園。學校會請參加的同學吃午餐。

歡迎大家來幫忙！

選項(A) 參加這個活動要自己買午餐。

選項(B) 這是星期五晚上的活動。

選項(C) 打掃的時間一共是三個小時。

三、生成試題 (模式一)

- (2025年) Gemini 生成材料題
利用已有生成試題的內容來迭代指示詞，生成主題相同、內容不同的試題

指示詞：生成一張學校活動的宣傳單，以下是一題範例。請生成類似風格，但是不同內容的活動，並提供三個合適的選項。



學校活動公告

活動名稱：
校園打掃日

活動日期：12月5日（星期五）
活動時間：上午8點~11點
活動內容：幫忙打掃學校旁邊的公園。學校會請參加的同學吃午餐。
歡迎大家來幫忙！

- 選項(A) 參加這個活動要自己買午餐。
- 選項(B) 這是星期五下午的活動。
- 選項(C) 打掃的時間一共是三個小時。

- (2025年) ChatGPT 生成材料題
利用同樣指示詞，生成圖片

漢字仍有錯誤或不清楚(動、幫、學、邊、會、園、請、參、歡等)，文字效果仍不佳。



校園打掃日

活動日期：12月5日
(星期五)
活動時間：上午8點~
11點
活動內容：幫忙打掃
學校旁邊的公園
學校會請參加的同學
吃午餐
歡迎大家來幫忙！

三、生成試題 (模式二)

• 模式(二)

設計者：本研究作者中的謝奇懿副教授（研究者三）以ChatGPT 建置測驗試題模型來生成試題。

模型概念：

- 研究者依據Band A聽力和閱讀題型的題型內容、命題原則、評量面向及彙整TBCL詞語分級表對應A1和A2詞語等級的檔案，設計整體的試題架構及指示(prompt)細節，共建置了九個模型
- 研究者授權給淡江大學華語中心，提供教師、中心人員及高師大參與生成試題的研究生使用，進行測試及校準。



三、生成試題 (模式二)

• 模式(二)

題型：聽力測驗第一大題-看圖回答

- 1.輸入「領域主題，程度」，例如「人際關係，A2」
- 2.ChatGPT 生成題幹圖片及聽力選項內容。(圖片和選項先後生成)
- 3.教師依生成的內容進行調修。
- 4.進行文本詞彙等級檢查。



生成調修後的內容

題目：他們正在做什麼？
選項 (A) 他們一起吃飯。
選項 (B) 他們在公園裡聊天。
選項 (C) 他們正在比賽跑步。



現在生成：看圖回答-圖文結合題目

請依照格式輸入主題及難度(A1或A2)，格式：主題，難度。如：“學校，A1”。

現在生成：看圖回答-圖文結合題目
請依照格式輸入主題及難度(A1或A2)



題目：

提問：他們兩個人正在做什麼？

1. 他們在一起吃飯。
2. 他們在公園裡聊天，其中一人正在給另一人水喝。
3. 他們正在跑步比賽。

題目完成，請問是否要進行文本詞彙等級檢查？



題目完成，請問是否要進行文本詞彙等級檢查？

人際關係，A2

人際關係，A2

選項2句子太長

選項2句子太長

三、生成試題 (模式二)

• 模式(二)

題型：聽力測驗第四大題-對話

- 1.輸入「領域主題，程度」，例如「身體及健康照護，A2」
- 2.ChatGPT 生成對話及選項內容。
- 3.教師依生成的內容進行調修。
- 4.進行文本詞彙等級檢查。

生成調修的內容

男：你看起來不太好，是不是感冒了？

女：對啊，我這幾天一直咳嗽，喉嚨也很痛，晚上還睡不好，覺得很不舒服。

男：那你去看了醫生嗎？醫生怎麼說？

女：有啊，醫生說是感冒，給我一些藥，還要我多喝水、休息，這幾天就會好了。

問題：這位女生為什麼不舒服？

選項(A) 她感冒了，需要多喝水和休息。

選項(B) 她吃了不乾淨的東西，肚子很不舒服。

選項(C) 她工作太忙，沒有時間睡覺。

選項(D) 她跌倒受傷了，需要去看醫生。

男：你看起來不太好，是不是感冒了？最近是不是太累了？
女：對啊，我這幾天一直咳嗽，喉嚨也很痛，晚上睡不著，覺得很不舒服。

男：那你有去看醫生嗎？醫生怎麼說？
女：有啊，醫生說是感冒，給我一些藥，還說要多喝水、休息，這樣比較容易康復。

問題：這位女生為什麼不舒服？
A. 她感冒了，需要多喝水和休息。
B. 她吃了不乾淨的東西，結果拉肚子了。
C. 她工作太忙，沒有時間睡覺。
D. 她跌倒受傷，需要去看醫生。

題目完成，請問是否要進行文本詞彙等級檢查？

康復、拉肚子是超綱詞

男：你看起來不太好，是不是感冒了？還是最近沒睡好？
女：對啊，我這幾天一直咳嗽，喉嚨也很痛，晚上還睡不好，覺得全身沒力氣。

男：那你有去看醫生嗎？醫生怎麼說？
女：有啊，醫生說是感冒，給了我一些藥，還要我多喝水和休息，過幾天就會好了。

問題：這位女生為什麼不舒服？
A. 她感冒了，需要多喝水和休息。
B. 她吃壞肚子了，肚子很不舒服。
C. 她工作太忙，沒有時間睡覺。
D. 她跌倒受傷，需要去看醫生。

題目完成，請問是否要進行文本詞彙等級檢查？



三、生成試題 (模式二)

• 模式(二)

題型：閱讀測驗第一大題-單句理解

- 1.輸入「領域主題，程度」，例如「職業，A2」
- 2.ChatGPT 生成題幹文字及圖片選項內容。(先生成選項，後生成圖片)
- 3.教師依生成的內容進行調修。
- 4.進行文本詞彙等級檢查。

一開始生成的內容：

題目：他是一名服務員，
在餐廳裡幫客人點餐。

選項/生成圖片

- (A) 一位服務員正在幫客人端菜
- (B) 一位教師在教室裡教學生
- (C) 一位醫生在醫院裡照顧病人

修改為

題目：服務員在餐廳裡幫客人點餐。
(A) 一位服務員拿著菜單和客人說話
(B) 一位服務員在廚房清洗碗盤
(C) 一位服務員在櫃台為客人結帳

 閱讀材料：
他是一名服務員，在餐廳裡幫客人點餐。

圖片設計：
A. 一位服務員正在幫客人端菜。
B. 一位老師在教室裡教學生。
C. 一位醫生在醫院照顧病人。

是否接受圖片設計，若是則開始生成圖片。

請更改圖片設計

 圖片設計：
A. 一位服務員拿著菜單和客人說話。
B. 一位服務員在廚房清洗碗盤。
C. 一位服務員正在收拾餐桌。

是否接受圖片設計，若是則開始生成圖片。

是



題目完成，請問是否要進行文本詞彙等級檢查？

請換收拾餐桌的選項

三、生成試題 (模式二)

• 模式(二)

題型：閱讀測驗第五大題-閱讀理解

- 1.輸入「領域主題，字數」，例如「颱風天放假，150字」
- 2.ChatGPT 生成一段短文內容。教師依生成的內容進行調修。
- 3.待確認短文後，生成題目及選項內容，教師進行調修。
- 4.進行文本詞彙等級檢查。

重新開始，颱風天放假，150字

🗨️ 請稍候，我將生成閱讀材料。

重新開始，颱風天放假，150字

請稍後，我將生成閱讀材料

今天早上，我在電視上看到颱風的消息。老師說，因為颱風很大，學校停課一天。媽媽告訴我不要出門，因為外面風很強，雨也很大。我們在家裡看電視、聊天，也玩了一些遊戲。中午，我和弟弟一起吃午餐，然後我們看了一部電影。下午，雨變小了，但風還是很大。晚上，我們早早吃晚餐，準備明天回學校上課。我希望以後也能安全度過每個颱風天。

是否接受這篇閱讀材料？

是否接受這篇閱讀材料？

情境改為上班的人不用上班，但是必須在家處理工作上的事

🗨️ 請稍候，我將修改閱讀材料。

情境改為上班的人不用上班，但是必須在家裡處理工作上的事

請稍後，我將修改閱讀材料

今天早上，我在電視上看到颱風的消息。公司通知大家不用去上班，但需要在家完成一些工作。外面風很大、雨很強，老闆提醒我們要注意安全。早上，我用電腦回覆客戶的郵件。中午，我和家人一起吃飯，休息了一會兒。下午，我開了一個線上會議，討論明天的計畫。雖然不必去公司，但工作還是很多，希望颱風快點結束，大家都能平安。

是否接受這篇閱讀材料？

是

🗨️ 請依格式輸入閱讀題組需要的題數，例如:3題。

請依格式輸入閱讀題組需要的題數↓

2題

🗨️ 題目

1. 根據文章，為什麼今天大家不用去上班？
 - A. 公司舉辦活動
 - B. 颱風來了
 - C. 老闆休假

2題

三、生成試題 (模式二)

• 模式(二)

題型：閱讀測驗第五大題-閱讀理解

- 1.輸入「領域主題，字數」，例如「颱風天放假，150字」
- 2.ChatGPT 生成一段短文內容。教師依生成的內容進行調修。
- 3.待確認短文後，生成題目及選項內容，教師進行調修。
- 4.進行文本詞彙等級檢查。

一開始生成的內容：

(以學校、主角為學生的情境為主)

今天早上，我在電視上看到颱風的消息。老師說，因為颱風很大，學校停課一天。媽媽告訴我不要出門，因為外面風很強，雨也很大。我們在家裡看電視、聊天，也玩了一些遊戲。中午，我和弟弟一起吃午餐，然後我們看了一部電影。下午，雨變小了，但風還是很大。晚上，我們很早吃了晚餐，準備明天回學校上課。我希望以後也能安全度過每個颱風天。

提供指示詞後修改的內容：

情境改為上班的人不用上班，但是必須在家裡處理工作上的事

今天早上，我在電視上看到颱風的消息。公司通知大家不用去上班，但需要在家做完一些工作。外面風很大，雨很強，老闆告訴我們要注意安全。早上，我用電腦回了客戶的郵件。中午，我和家人一起吃飯，休息了一會兒。下午，我開了一個線上會議，討論明天的計畫。雖然我不必去公司，但工作還是很多。希望颱風快點離開，大家都能平安。

題目：為什麼今天大家在家工作？

選項(A) 公司舉辦了一個活動。

選項(B) 颱風來了，不用去上班。

選項(C) 停電了，公司休息一天。

四、研究結果

4.1 生成試題的數量、有效性及可行性

模式(一)

1. 研究者必須輸入大量的指示詞。
2. 在生成試題後，ChatGPT會提供接續對話的選擇，例如：「可以再設計另一個題目嗎？」、「這些題目能再難一點嗎？」、「還要換其他的主題嗎？」等，讓研究者得以在同一個主題中生成3~4個題目。
3. 過程較費時，研究者必須不斷換模擬試題內容或主題，輸入例如「請再設計另一個題目，主題為...(依照分級詞表的主題情境)，句子的詞彙等級是...(A1、A2)」等。

模式(二)

1. 由於模型已經設定好格式、等級，ChatGPT會在生成試題後，詢問給予試題修正意見、是否接受選項的圖片設計或是閱讀材料。
2. 如果研究者提出修正或重啟開始，往往就會再生成其他題幹或選項內容，也會在題目完成後詢問是否繼續生成其他試題。

整體而言，運用模式(二)的架構在輸入指示詞上較為省力，但兩種模式都能持續性增加生成試題的數量，且試題都能經過教師和ChatGPT的協作修正達到品質較高、具多樣性的的題本內容。

四、研究結果

4.2 生成試題模式之情況與修正

一、生成文本的內容接近重複

1. 生成順利的情況常是句子內容較為簡短，句構簡單的題目，相對地也造成語句、詞彙的重複性高。
2. 模式(一)生成出來的題日常會相當接近所輸入的模擬例題，容易造成試題內容過於重覆的情況。
3. 模式(二)是在同一個題型，輸入同一個主題後，所生成的選項語句或是跟前面已生成過的題目相當接近。

二、詞彙超過等級

雖然兩種模式的提示設定中都已附上測驗的分級詞表檔案，並在試題生成設定時限定詞彙等級，但生成的試題中仍會出現少數或一部分超過等級的詞語。

三、反覆修正多次後容易產生耗弱情況，無法得到理想的文本

1. 兩種模式都會出現指示修正試題的詞彙或句子有時更不合適，反覆生成多次試題(約7~8次)後，反倒沒有第一次生成的試題內容好，有準確程度下降或耗弱的情況。
2. 有時會產生少數上下文不合邏輯，或不符中文語法規範的句子。
3. 過程中偶有最不理想的情況是整題不採用，必須重啟生成。

四、研究結果

4.2 生成試題模式之情況與修正

四、生成圖片的問題 (此為2024年的情況，2025年採取利用其他AI工具試驗，例如利用Gemini 來生成)

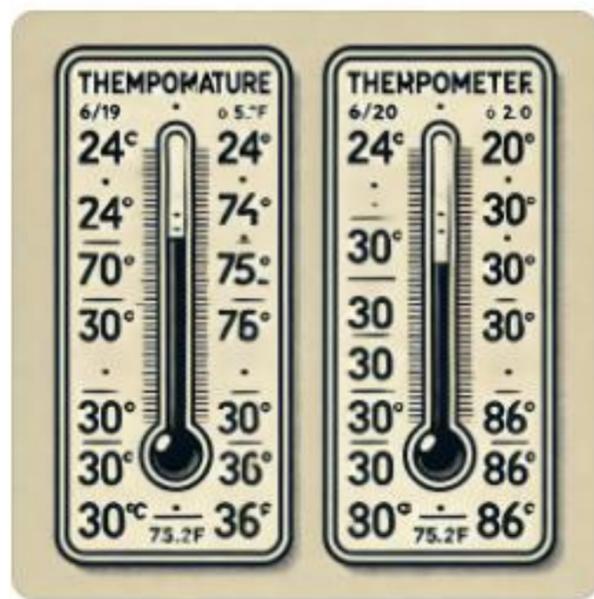
生成圖片不合適最常發生的兩個情況

1. 閱讀測驗第二大題材料文本的問題，圖片題幹和選項也可能產生此類問題，主要為圖片檔案中具有文字(例如宣傳、公告)、數字(例如時間、號碼、價錢、溫度)、方位(例如具有方向性的地圖、路線圖)等。
2. 圖片內容和生成選項的文字描述並不完全相同，在本研究中，模式(一)發生此情況的頻率較高，研究者調整修正指令後，圖片仍然不正確或不合適，或與題幹、選項的關聯程度差距更大。



(A) 郵局裡的工作人員在_____顧客。

- A. 打包
- B. 幫助
- C. 跳舞



生成選項的文字敘述：
樓上有兩個人跳舞，音樂聲很大，
老太太捂住耳朵。

生成的圖片結果：
老太太旁邊有個大音響，音樂聲很
大，樓上的人也捂住耳朵。

四、研究結果

4.3 研究者對兩種模式生成試題的研究結果

- **模式(一)**
- 華語文能力測驗試題中僅有初級試題具有多種圖文結合的題型，但其變化涉及了試題難度、圖片繪製選擇、圖文搭配、圖文配置等層面。
- 模式(一) 採取由研究者自由輸入大量的指示詞，實際上相當考驗華語教師本身對於華測題型和命題內涵的掌握程度，及對使用生成式AI與之互動的能力，因而影響試題文本生成的效率，產出質量不一或圖文不符合試題概念的內容。

- **目前仍在測試、校準的模式(二)**
- 研究者以對華測試題題型的解析、架構和ChatGPT大型語言模型的應用專業，所創建的聽力及閱讀各題型試題的生成模組。
- 本研究希望透過持續地生成及測試試題、輸入修正指令、訓練語言模型等的過程中，找到最適合生成式AI應用於試題的框架和結構。
- 研發期待即使未來是對華測題型不熟悉的教師或研究者，只需要在模組中輸入最少的指示詞和修正次數，就能獲得有效率及高準確度的試題輸出結果，以確保試題內容能具有一定的試題品質。



五、研究結論

研究結論：

1. 生成式AI特別能在短時間內生成多樣化的試題內容和圖片，提高了試題的產量，節省了人力和時間，並簡化了過去繁瑣的出題和繪圖流程，但在試題研發上仍有一些侷限和待調整之處，因此需要持續進行測試與調整。
2. 目前本研究的試題研發僅限於研究者的華語中心使用，研究限制為尚未針對試題的適用性擴大測驗施測及回收測驗結果和數據，以評估測驗的信效度。

後續研究：

生成式AI的發展趨勢有更多潛在的可能性，未來的研究預期生成式AI的發展性：

1. 技術上能進一步優化試題的自動化生成能力。
2. 能提供更加個性化、針對不同語言學習者需求的測驗內容。
3. 能運用自然語言處理技術來實現自動檢核內容或是即時回饋。
4. 能生成更準確、豐富的測驗和圖片內容，使測驗達到更多樣的應用方式。
5. 在應用實務上還需要更多專業團隊的合作，特別是在AI技術與華語教學實踐的過程中，如何整合求其最大化其效益，擴展在華語文能力測驗中的應用範圍。



六、參考文獻

1. 吳欣儒、林慶隆、白明弘、何莉玉、王冠孺、陳沛璇、許鈞惟、吳鑑城、李詩敏、丁彥平（編著）（2022）。華語文能力基準應用參考指引。新北市：教育部國家教育研究院。
2. 吳燦銘（2024）。聰明提問AI的技巧與實例：ChatGPT、Copilot、AgentGPT、AI繪圖一次滿足。新北市：碩博文化股份有限公司。
3. 周中天、張莉萍（2007）。華語文能力分級指標之建立。「東亞教育評論論壇：新興議題及挑戰」國際會議。臺灣師範大學，台北，台灣。
4. 高宛愉（2024）。第一線教師對於生成式AI在教育現場應用現況及態度。師友雙月刊，644，23-28。
5. 徐臺屏、王政忠（2023）。AI融入教學的可行策略與教學示例。師友雙月刊，644，29-34。
6. 陳右怡（2022）。AI未來四大方向：分散式AI、生成式AI、可信任AI、永續AI。電腦與通訊，192，33-41。
7. 陳怡攸（2020）。教學視角下華語文測驗題庫的編製與分析—以聽力測驗題庫為例。第十九屆台灣華語文教學年會暨國際學術研討會。中央大學，桃園，台灣。
8. 陳思妤（2020）。華語文線上測驗系統之建構與試行。第十九屆台灣華語文教學年會暨國際學術研討會。中央大學，桃園，台灣。
9. 張瓊文、朱晉杰（2023）。圖卡的力量（九）當圖卡遇見生成式AI（上）—圖片生成常見AIGC工具介紹。師友雙月刊，644，82-88。
10. 教育部國家教育研究院（編著）（2020）。遣辭用「據」—臺灣華語文能力第一套標準。新北市：教育部國家教育研究院。
11. 華語測驗推動工作委員會網站〈測驗類別〉（2025）。聽力測驗。取自<https://tocfl.edu.tw/index.php/test/listening>
12. 華語測驗推動工作委員會網站〈測驗類別〉（2025）。閱讀測驗。取自<https://tocfl.edu.tw/index.php/test/reading>

六、參考文獻

13. 華語測驗推動工作委員會網站〈相關下載-研習資料〉（2025）。華語文能力基準應用參考指引於測驗評量應用實作(2023年)。取自https://tocfl.edu.tw/assets/files/2023presentation_TBCL.pdf
14. 華語測驗推動工作委員會網站〈相關下載-研習資料〉（2025）。華語文能力測驗命題設計概述(2022年)。取自<https://tocfl.edu.tw/assets/files/2022presentation.pdf>
15. 華語測驗推動工作委員會網站〈相關下載-研習資料〉（2025）。華語文能力測驗命題趨勢研習講義(2016年)。取自<https://tocfl.edu.tw/assets/files/20160828.pdf>
16. 華語測驗推動工作委員會網站〈試題編寫範例〉（2025）。試題編寫範例。取自<https://tocfl.edu.tw/index.php/teach/examples>
17. 華語測驗推動工作委員會網站〈TBCL資源〉（2025）。TBCL資源。取自<https://tocfl.edu.tw/index.php/exam/resources>
18. 華語測驗推動工作委員會網站〈參考詞表〉（2025）。華語八千詞、華語八千詞調整說明。取自<https://tocfl.edu.tw/index.php/exam/download>
19. 華語測驗推動工作委員會網站〈模擬試題〉（2025）。聽力測驗、閱讀測驗入門基礎級題庫、第一輯~第五輯模擬試題。取自<https://tocfl.edu.tw/index.php/exam/test/page/1>
20. 曾文璇、蔡怡珊、林玲英、陳柏熹（2023）。使用華語文語料庫與能力基準整合應用系統設計TOCFL測驗試題。華文世界，132，115-123。
21. 蔡宜坦、施威銘研究室（2023）。ChatGPT 4 Turbo萬用手冊2024春季號：提示工程、超強外掛、My GPTs、OpenAI API、Midjourney、Copilot、Bard、Claude2。台北市：旗標科技股份有限公司。
22. 謝奇懿、周湘華（2020）。教學視角下華語聽讀模擬能力測驗建置之基礎、問題及試行。第十九屆台灣華語文教學年會暨國際學術研討會。中央大學，桃園，台灣。
23. 謝奇懿、陳怡攸、陳思妤、吳瑋庭（2022）。臺灣華語文能力測驗入門基礎級題型特點及命題難點探析—以閱讀能力測驗圖文結合題型為對象。章法論叢，14，97-133。
24. 藍珮君（2007）。基礎華語文能力測驗與歐洲語言共同參考架構的對應關係。台灣華語文教學，3，39-47。



感謝聆聽 *Thank you*

