



STUDENT ABILITY DIAGNOSTIC REPORT

學業能力 診斷報告

歷史 · HISTORY

能力估計 θ · 實際 / 預期對照 · 維度畫像 · 認知診斷 · 成績穩定性 · 個人化提升

研值 YANZ 測評研究院 · 2025—2026 學年第二學期期中學情診斷 · 2026-04-22

學生資料 · 報告導讀

這份報告屬於你

本報告依你在本次聯考數學科目的**逐題作答**，以項目反應理論 (IRT) 估計你的能力，並以認知診斷 (DINA) 還原你在各知識與能力維度上的掌握情況。它不只告訴你「考了多少分」，更回答「你的真實水平在哪裡、下一步往哪裡使勁」。

學校	研值實驗高級中學	本次總分	118 / 150
姓名	李思源	校內百分位	前 18%
班級	初一(3)班	能力估計 θ	+0.92
考號	0719	總體等級	4 / 5
測驗	2025—2026 學年第二學期期中學情診斷 · 歷史	全校均分	96.5 / 150
測驗日期	2026-04-22		

如何閱讀這份報告

- **實際 vs 預期**：「預期」是依你當前能力 θ 推算「你這種水平本應答對的機率」。 ■ 綠底 = 實際高於預期（超常發揮）； ■ 紅底 = 實際低於預期（失常，最值得回看）。
- **等級 1—5**：1 = 待加強，3 = 合格，5 = 優秀，由該維度的得分率與全體分布共同決定。
- **難度 E/M/H**：易 / 中 / 難，由全體考生的正確率自動校準。**Norm** = 全體考生在該題的得分率。
- **穩定性系數**：跨多次測驗的成績波動程度——數值越小越穩定。

方法：項目反應理論 2PL/3PL 能力估計 · 經典測量理論難度與鑑別度 · DINA 認知診斷 · 維度加權合成。演算法已在約 2,000 萬條真實聯考逐題作答上驗證。

總體表現

01 你的總體表現概覽

本次數學測驗你的總分為 118 / 150 (得分率 79%)，校內百分位約 前 18%，總體等級 4 / 5。下圖按維度給出你的能力等級——虛線為滿級參照。

<p>118/150</p> <p>總分</p>	<p>79%</p> <p>得分率</p>	<p>前 18%</p> <p>校內百分位</p>	<p>+0.92</p> <p>能力估計 θ</p>
---------------------------------	------------------------------	----------------------------------	---

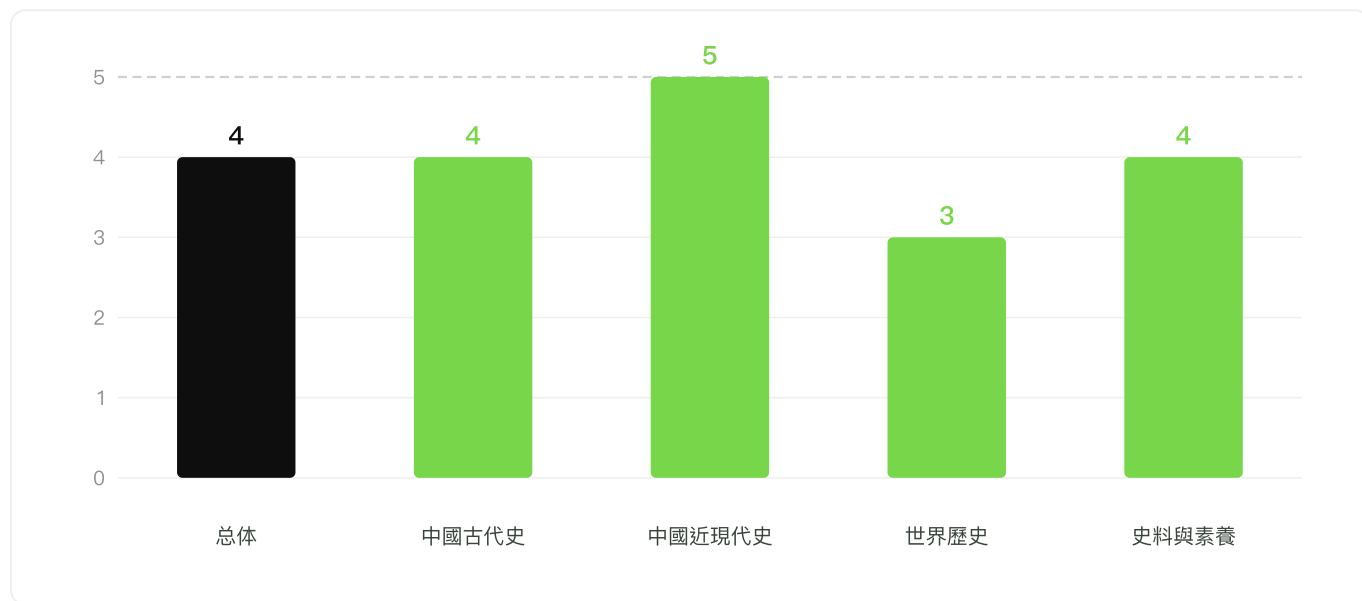


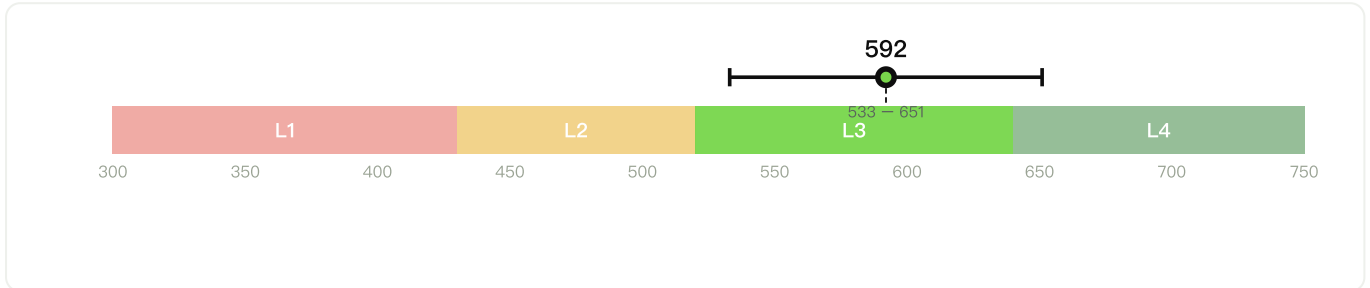
圖 1 · 各維度能力等級 (1-5)。深色為總體，綠色為各內容維度；虛線 = 滿級 5。 如何看圖：柱越高 = 該維度等級越高；虛線為滿級 5。

一句話畫像：你總體居於全年級前列，**中國近現代史與史料與素養**是你的強項；但**世界歷史**明顯拖後，且成績在歷次測驗間有起伏。把**世界歷史**補上、把**穩定性**提上來，是你提分最快的兩條路徑。

能力量表 · 達標水平 · 測量精度

你的能力量表分與達標水平

項目反應理論 (IRT) 把你的每一道作答轉換成統一的能力量表分 (均值 500、標準差 100) —— 它比卷面分更穩定，且可跨不同卷子比較。任何一次測驗都有誤差，因此我們同時給出測量標準誤 (SEM) 與 95% 置信區間：你的真實能力有 95% 的把握落在這個區間內。



如何看圖：色帶為四個達標水平，圓點是你的能力量表分，左右鬚線是 95% 置信區間——區間越窄，本次測量越精確。

達標水平

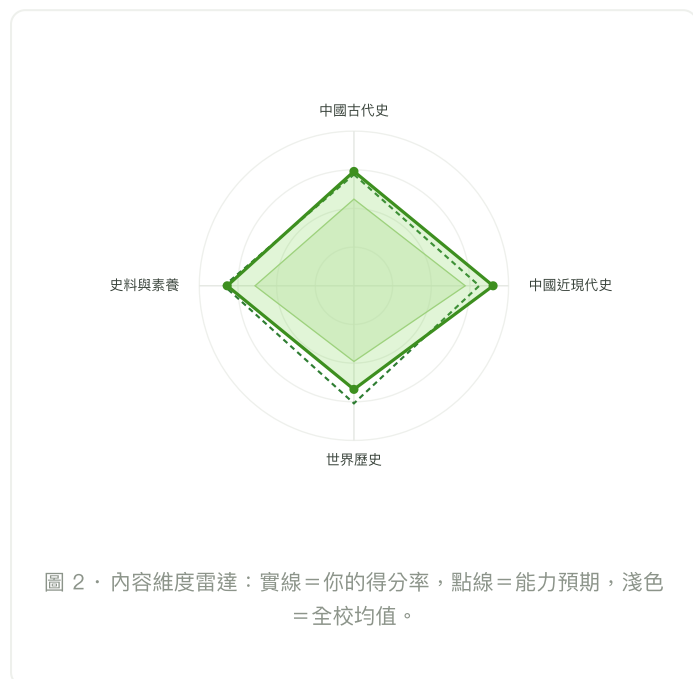
- L1 待夯實 基礎知識與技能尚不穩固，需要系統性補強後再進階。
- L2 接近達標 已掌握部分核心知識，能完成基礎任務；面對綜合與複雜問題仍不穩定。
- L3 達標·熟練 熟練掌握課程標準的核心要求，能穩定完成中等難度任務，並已能解決部分綜合 / 較難問題。
- L4 優秀·拔尖 在多數維度表現優異，能靈活遷移知識、穩定攻克高階與新情境問題。

測量精度：本次你的 SEM 約為 30 分 (較高精度)；置信區間主要落在 L3、上沿已觸及 L4，說明「達標·熟練」這一判定是穩健的，且你已處在邁向「優秀」的臨界帶。

維度畫像

02 各內容維度表現

雷達圖對比你在四個內容維度上的得分率、全校均值與能力預期。當「你的表現」超過「能力預期」，說明該維度發揮超常；反之則是高產出的提升點。



維度	得分 / 滿分	得分率	全校均	能力預期	等級
中國古代史	28/38	74%	56%	72%	4
中國近現代史	27/30	90%	72%	81%	5
世界歷史	18/27	67%	49%	76%	3
史料與素養	45/55	82%	64%	84%	4

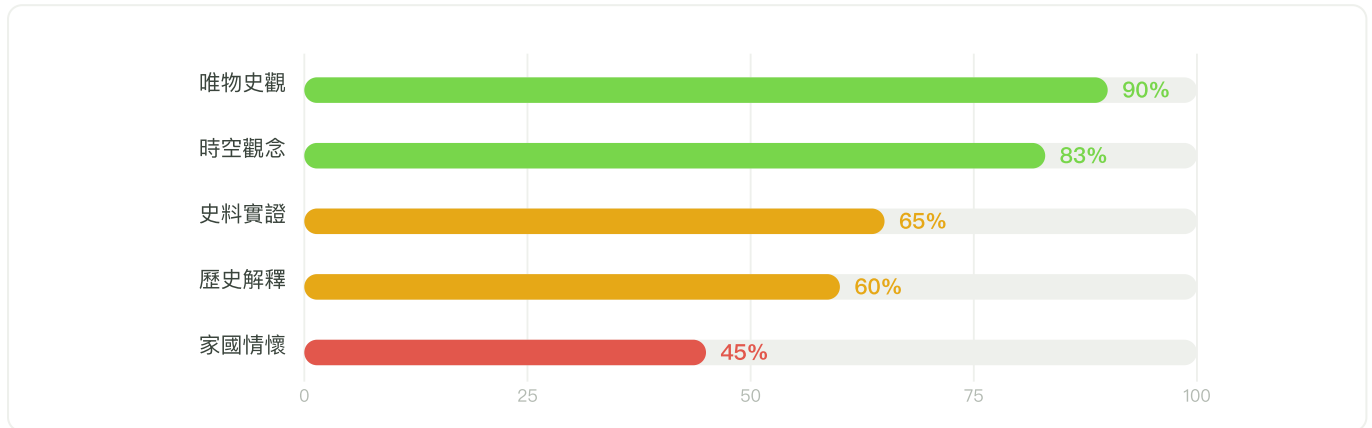
如何看圖：實線越往外 = 你越強；實線越過點線（能力預期）即為發揮超常。

中國近現代史顯著高於你的能力預期（+9 個百分點），是你的「拿手戲」；而**世界歷史**低於預期 9 個百分點——同樣的能力，本可換來更高的分數。

學科核心素養

03 歷史核心素養畫像

超越「分數」，本報告用認知診斷（DINA）估計你在本學科課程標準核心素養各維度上的掌握程度——它解釋「為什麼」：得失分往往源於某項核心素養的強弱，而非單個知識點。



各項學科核心素養的掌握程度（0—100%）。綠 = 已達成(≥75%)，黃 = 發展中(50—75%)，紅 = 待加強(<50%)。如何看圖：橫條越長 = 該項認知能力掌握越好；顏色為掌握程度分級。

你的多數核心素養已達成；相對薄弱的素養是後續重點培養方向——素養比單題更能遷移到新題型。

實際 VS 預期

04 逐題：實際表現與能力預期對照

下表逐題對照你的實際得分與能力預期。綠底 = 超常（高於預期），紅底 = 失常（低於預期）。失常題是你「會做卻沒做對」的題，回看它們的性價比最高。

選擇題（含多選，部分給分）

題號	難度	滿分	實際	預期
1 單選	易	5	5	5
2 單選	易	5	5	5
3 單選	中	5	5	5
4 單選	中	5	5	5
5 單選	中	5	0	5
6 單選	中	5	5	5
7 單選	難	5	5	3
8 單選	難	5	5	4
9 多選	中	6	6	4
10 多選	難	6	4	4
11 多選	難	6	3	5

填空題與解答題

題號	難度	滿分	實際	預期
12 填空	易	5	5	5
13 填空	中	5	5	5
14 填空	難	5	0	4
15 填空	難	5	5	2
16 解答	中	12	11	10
17 解答	中	12	8	11
18 解答	中	14	11	11
19 解答	難	17	14	11
20 解答	難	17	11	9

你有 7 道題超常、4 道題失常。失常集中在 第5題、第11題、第14題、第17題——這些是你「會做卻沒做對」的題，回看它們的性價比最高。

難度層級

05 不同難度層級上的表現

把題目按難度（易 / 中 / 難）分層，看你在每一層的強弱。★ 標出你當前的能力位置——它落在中—難之間：中檔題應當全取，難檔題已能穩定撕開口子。

表現較強（達到或超過預期）	難度	表現偏弱（低於預期）
Q7、Q8、Q10、Q15、Q19、Q20	難	Q11、Q14
Q3、Q4、Q6、Q9、Q13、Q16、Q18	中 ★ 你的能力位置	Q5、Q17
Q1、Q2、Q12	易	—

你的能力位置已逼近「難」檔。當務之急不是攻更難的題，而是堵住中檔的失分——中檔題每穩拿一題，提分都比死磕壓軸更確定。

成績穩定性

06 成績穩定性分析

穩定性反映你的數學成績在多次測驗間的波動：**穩定性系數越小越穩**。下圖橫軸為穩定性系數、縱軸為數學總分，按「成績高低 × 是否穩定」劃為四個區，★ 為你當前的位置（依據本學期 5 次測驗：109、124、103、121、118）。

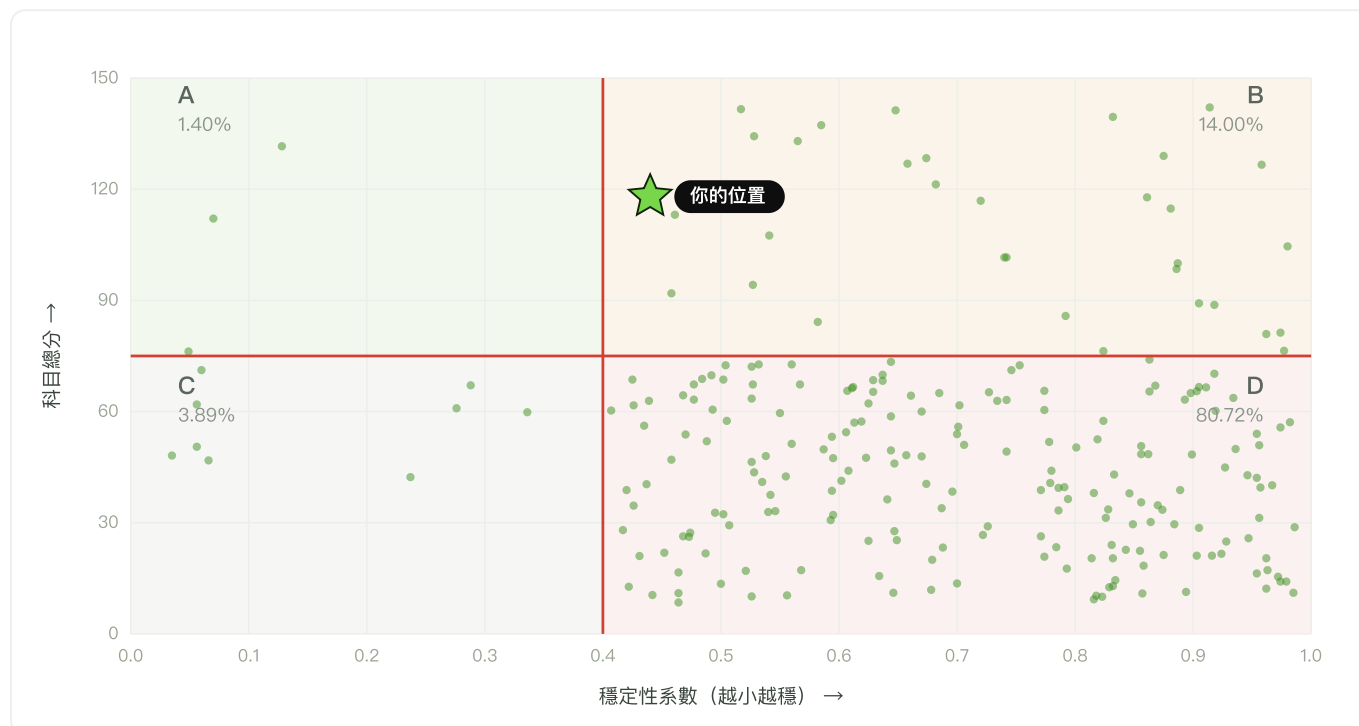


圖 4 · 成績穩定性四象限。每個點為一名考生；★ 為你。橫軸 = 穩定性系數(0—1)，縱軸 = 數學總分(0—150)。如何看圖：越靠左 = 越穩定，越靠上 = 分數越高；理想是走向左上方的 A 區。

A 區 · 績優且穩定

B 區 · 績優但不穩定

C 區 · 績弱且穩定

D 區 · 績弱但不穩定

你處於 B 區（績優但不穩定）：分數已在全年級前列，但波動偏大（本學期在 103—124 間起落）。波動主要來自中檔題的偶發失誤。穩定性，是你最確定的一塊提分空間——把 B 區拉到 A 區，往往只需消除「會做做錯」。

把波動變小的三件事

- 建立錯題歸因表：每道錯題只問一句——是「不會」還是「會做做錯」？後者就是穩定性的漏洞。
- 限時分段訓練：用 40 分鐘專練選擇 + 填空，把中檔題的「手感」固化為肌肉記憶。
- 考試節奏管理：先穩取中檔（1—15 題）再攻壓軸，避免在難題上擠壓了基礎題的檢查時間。

群體位置

07 你在不同群體中的位置

同一個分數，在不同對照群體裡的相對位置可能不同。下圖給出你在 **性別 / 年齡 / 地區** 三個群體中的位置（圓點越靠右 = 越靠前）。

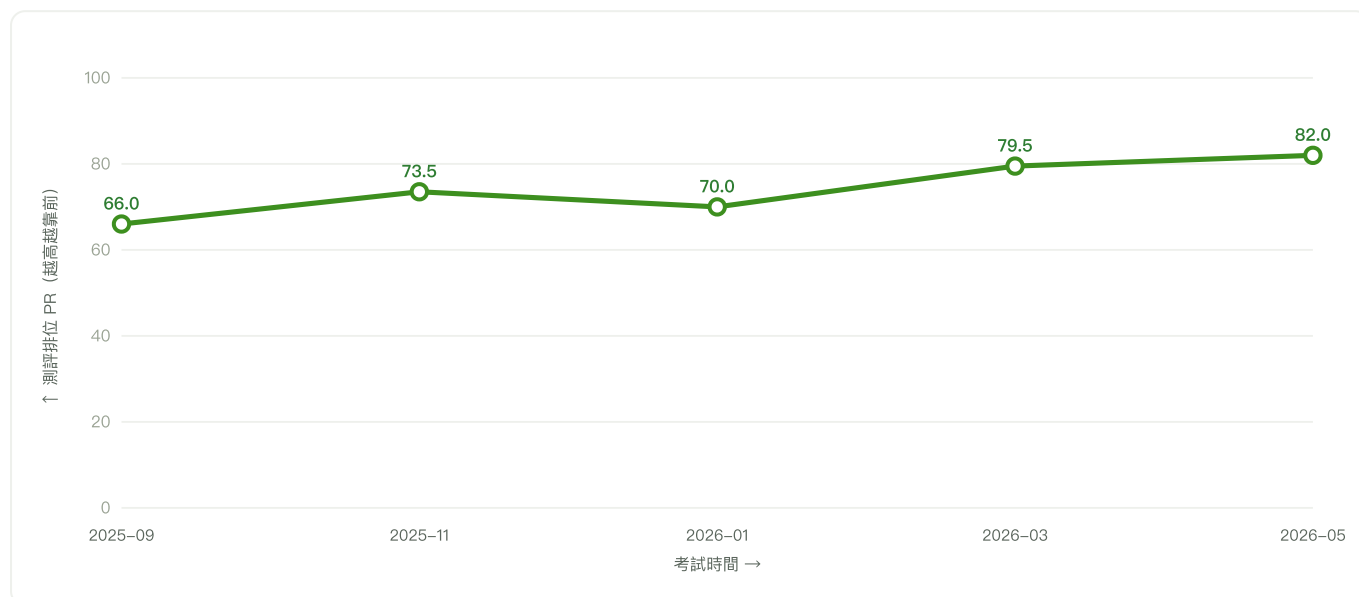


在性別與同齡群體中你已居前列；本地區競爭更激烈，是繼續向上的空間。

學業趨勢

08 測評排位變化趨勢

下圖展示你歷次考試的測評排位隨時間的變化。因每次考試人數不同、絕對名次不可比，這裡用百分位 PR (0–100，越高越靠前) 這一比例分值；線越靠上 = 排位越靠前，持續上行即穩步進步。



橫軸 = 考試時間，縱軸 = 測評排位 PR (比例分值，不受人數影響)；曲線越往上越靠前。 **如何看圖：**線向上 = 排位前進，向下 = 退步；留意拐點對應的那次考試，複盤當次得失。

你的測評排位整體穩步上升 (PR 66 → 82)；保持節奏，向更高百分位衝刺。

詳細數據

09 歷次考試明細

下表為你歷次考試明細。點擊任意一行，即可跳轉到該次考試的完整診斷報告。

日期	分數	測評排位	等級	查看報告
2025-09	101 / 150	66.0	3級·合格	查看 →
2025-11	108 / 150	73.5	3級·合格	查看 →
2026-01	104 / 150	70.0	3級·合格	查看 →
2026-03	112 / 150	79.5	4級·良好	查看 →
2026-05	118 / 150	82.0	4級·良好	查看 →

說明：「測評排位」為百分位 PR（比例分值，越高越靠前，不受人數影響）；「等級」與報告前面的 1-5 級一致。點擊任意一行跳轉到該次完整報告（樣板演示中連結指向示例報告）。

優先提升

10 優先提升清單

下列題目是你「能力夠、卻沒拿到分」的題（實際明顯低於預期），按提分性價比排序。逐一攻克它們，是本階段最高效的複習動作。

題號	難度	你的得分	考點	診斷與建議
Q5 單選	中	0/5	工業革命與資本主義	【工業革命與資本主義】掌握尚不牢固，本可拿到的分數沒拿到，建議優先回看。
Q11 多選	難	3/6	近代列強侵略與救亡	【近代列強侵略與救亡】掌握尚不牢固，本可拿到的分數沒拿到，建議優先回看。
Q14 填空	難	0/5	兩次世界大戰與冷戰	【兩次世界大戰與冷戰】掌握尚不牢固，本可拿到的分數沒拿到，建議優先回看。
Q17 解答	中	8/12	朝代與制度演變	【朝代與制度演變】掌握尚不牢固，本可拿到的分數沒拿到，建議優先回看。
Q10 多選	難	4/6	古代經濟與思想文化	【古代經濟與思想文化】掌握扎實、表現穩定。
Q16 解答	中	11/12	時空觀念與歷史解釋	在【時空觀念與歷史解釋】上超常發揮，較難的題也穩定拿分。

逐題考點

11 逐題考點與反饋 (Assessment Focus)

全卷逐題清單：難度、全體正確率 (Norm)、你的得分、考點與個人化反饋。這是你回看試卷時的「導航圖」。

題號	難度	Norm	你的得分	考點	反饋
Q1	易	90%	5/5	朝代與制度演變	【朝代與制度演變】掌握扎實、表現穩定。
Q2	易	86%	5/5	古代經濟與思想文化	【古代經濟與思想文化】掌握扎實、表現穩定。
Q3	中	66%	5/5	近代列強侵略與救亡	【近代列強侵略與救亡】掌握扎實、表現穩定。
Q4	中	71%	5/5	新中國建設與改革	【新中國建設與改革】掌握扎實、表現穩定。
Q5	中	58%	0/5	工業革命與資本主義	【工業革命與資本主義】掌握尚不牢固，本可拿到的分數沒拿到，建議優先回看。
Q6	中	62%	5/5	兩次世界大戰與冷戰	【兩次世界大戰與冷戰】掌握扎實、表現穩定。
Q7	難	38%	5/5	史料實證	在【史料實證】上超常發揮，較難的題也穩定拿分。
Q8	難	30%	5/5	時空觀念與歷史解釋	在【時空觀念與歷史解釋】上超常發揮，較難的題也穩定拿分。
Q9	中	55%	6/6	朝代與制度演變	在【朝代與制度演變】上超常發揮，較難的題也穩定拿分。
Q10	難	34%	4/6	古代經濟與思想文化	【古代經濟與思想文化】掌握扎實、表現穩定。
Q11	難	28%	3/6	近代列強侵略與救亡	【近代列強侵略與救亡】掌握尚不牢固，本可拿到的分數沒拿到，建議優先回看。
Q12	易	77%	5/5	新中國建設與改革	【新中國建設與改革】掌握扎實、表現穩定。
Q13	中	60%	5/5	工業革命與資本主義	【工業革命與資本主義】掌握扎實、表現穩定。
Q14	難	36%	0/5	兩次世界大戰與冷戰	【兩次世界大戰與冷戰】掌握尚不牢固，本可拿到的分數沒拿到，建議優先回看。
Q15	難	24%	5/5	史料實證	在【史料實證】上超常發揮，較難的題也穩定拿分。
Q16	中	58%	11/12	時空觀念與歷史解釋	在【時空觀念與歷史解釋】上超常發揮，較難的題也穩定拿分。
Q17	中	52%	8/12	朝代與制度演變	【朝代與制度演變】掌握尚不牢固，本可拿到的分數沒拿到，建議優先回看。
Q18	中	50%	11/14	古代經濟與思想文化	【古代經濟與思想文化】掌握扎實、表現穩定。
Q19	難	30%	14/17	近代列強侵略與救亡	在【近代列強侵略與救亡】上超常發揮，較難的題也穩定拿分。
Q20	難	15%	11/17	新中國建設與改革	在【新中國建設與改革】上超常發揮，較難的題也穩定拿分。

學習路徑

12 為你定制的提升路徑

把上面的診斷轉成可執行的行動。建議按「先穩後拔」推進：先用 4 週補穩定性與三角·數列，再用 4 週向建模與壓軸要分。

1 第 1—4 週 · 穩基礎、補短板

1. 建立錯題歸因表，專項清理「會做做錯」的中檔失分。
2. 針對最弱維度【世界歷史】每日定量訓練，糾正本次暴露的薄弱點。
3. 每週一次限時小卷，訓練考試節奏。

2 第 5—8 週 · 攻建模、拿壓軸

1. 綜合與應用專題：練「從情境到方法」的轉化。
2. 難題分步訓練，目標拿下高階題的前幾問。
3. 把高頻題型的標準流程固化為模板。

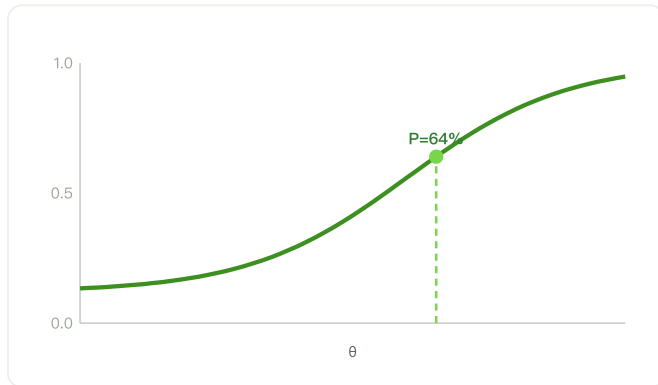
階段目標 在保持強項的前提下，把最弱維度【世界歷史】的得分率顯著提升、把成績波動收窄一半——總分有望穩步提升，並把象限位置從 B 區推入 A 區（績優且穩定）。

方法與讀圖指南

這份報告是怎麼算出來的

本報告採用國際主流的項目反應理論 IRT (2PL/3PL) 與認知診斷 DINA。下面三張圖分別解釋：單題難度與你的能力如何對應（項目特徵曲線）、這份卷子在你這個水平測得準不準（測驗信息函數）、以及你與全體考生、與各題難度的相對位置（能力—難度圖）。

① 項目特徵曲線 (ICC)



如何看圖：橫軸是能力 θ ，縱軸是答對該題的概率；曲線越靠右 = 題越難。豎線是你的 θ ，與曲線的交點 = 「以你的水平答對此題的概率」。

② 測驗信息函數 (TIF)



如何看圖：曲線越高 = 在該能力水平測得越精確（標準誤越小）。本卷信息峰值在中等偏上能力；你的 θ 落在高信息區，說明這份卷子對你測得相當準。

③ 能力—難度圖 (Wright Map)



如何看圖：同一把「logit 尺子」上：左側是全體考生的能力分佈，右側每個點是一道題的難度。落在你能力線以下的題，你有過半把握答對；線以上是挑戰區。

術語表

能力 θ (theta)

IRT 估計的潛在能力值；0 為該群體平均，正值高於平均。

能力量表分

由 θ 線性變換得到（均值 500、SD 100），便於跨卷比較。

SEM 測量標準誤

單次估計的不確定度；越小越精確，用於構造置信區間。

難度 b / 鑑別度 a

b = 答對概率 50% 所需的能力； a = 該題區分高低能力的靈敏度。

預期表現 Projected

以你當前 θ 推算的應答對概率；與實際對比可發現超常 / 失常。

DINA 認知診斷

由作答反推每項「認知能力」的掌握機率 (0—1)。



方法與聲明

■ 每一個結論，都來自你的真實作答

本報告的每一項指標——能力 θ 、實際 / 預期、維度等級、認知掌握、穩定性——都由你的逐題作答經統計模型計算得到，可複核、可解釋。它不是給你貼標籤，而是給你一張向上的地圖。

全卷 20 題 逐題作答	IRT 2PL/3PL 能力估計	DINA 認知診斷	約 2,000 萬 條 演算法驗證
------------------------	--------------------------------	---------------------	---------------------------------

說明：本報告為**樣板演示**，學生資料與數據為示例。指標定義與演算法見《YANZ 演算法數據說明書》。報告內容僅供學生與教師教學使用，請勿外傳。