



STUDENT ABILITY DIAGNOSTIC REPORT

学业能力 诊断报告

历史 · HISTORY

能力估计 θ · 实际 / 预期对照 · 维度画像 · 认知诊断 · 成绩稳定性 · 个性化提升

研值 YANZ 测评研究院 · 2025—2026 学年第二学期期末学情诊断 · 2026-06-18

这份报告属于你

本报告基于你在本次联考数学科目的**逐题作答**，用项目反应理论（IRT）估计你的能力，并用认知诊断（DINA）还原你在各知识与能力维度上的掌握情况。它不只告诉你「考了多少分」，更回答「你的真实水平在哪里、下一步往哪里使劲」。

学校	研值实验高级中学	本次总分	118 / 150
姓名	王梓涵	校内百分位	前 18%
班级	初二(5)班	能力估计 θ	+0.92
考号	0826	总体等级	4 / 5
测验	2025—2026 学年第二学期期末学情诊断 · 历史	全校均分	96.5 / 150
测验日期	2026-06-18		

如何读这份报告

- **实际 vs 预期**：「预期」是依你当前能力 θ 推算「你这种水平本应答对的概率」。 ■ 绿底 = 实际高于预期（超常发挥）； ■ 红底 = 实际低于预期（失常，最值得回看）。
- **等级 1—5**：1=待加强，3=合格，5=优秀，由该维度的得分率与全体分布共同决定。
- **难度 E/M/H**：易 / 中 / 难，由全体考生的正确率自动校准。**Norm** = 全体考生在该题的得分率。
- **稳定性系数**：跨多次测验的成绩波动程度——数值越小越稳定。

方法：项目反应理论 2PL/3PL 能力估计 · 经典测量理论难度与区分度 · DINA 认知诊断 · 维度加权合成。算法已在约 2,000 万条真实联考逐题作答上验证。

总体表现

01 你的总体表现概览

本次数学测验你的总分为 118 / 150 (得分率 79%)，校内百分位约 前 18%，总体等级 4 / 5。下图按维度给出你的能力等级——虚线为满级参照。

118/150 总分	79% 得分率	前 18% 校内百分位	+0.92 能力估计 θ
---------------	------------	----------------	------------------------

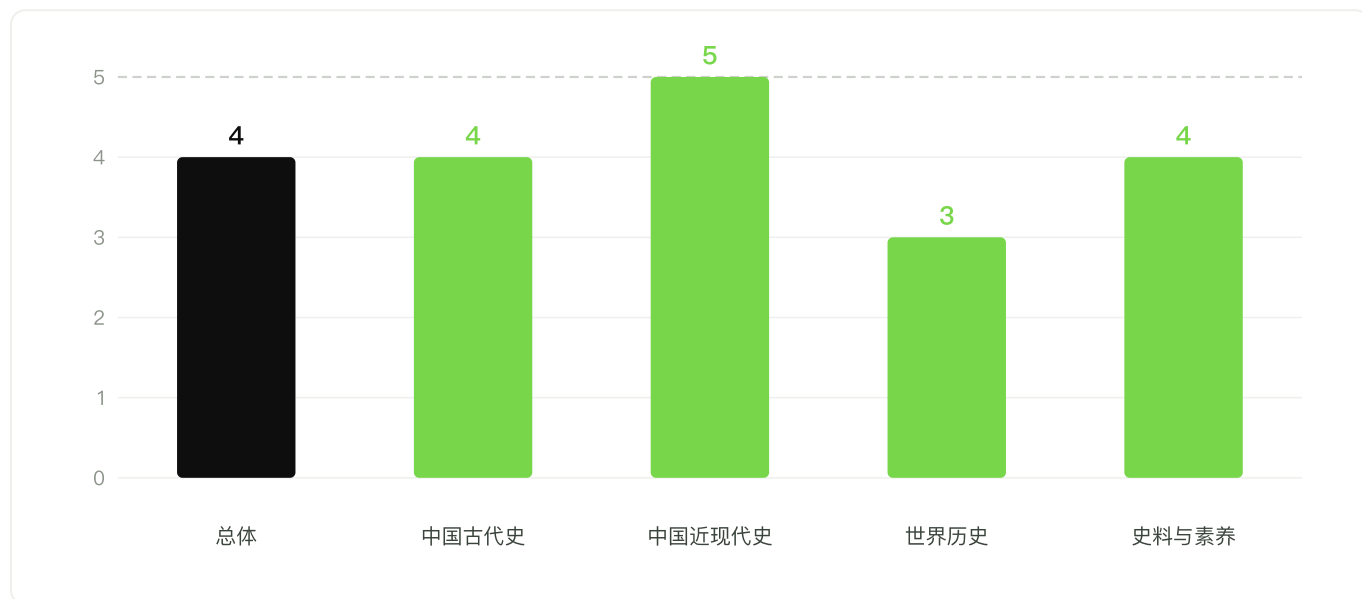


图 1 · 各维度能力等级 (1-5)。深色为总体，绿色为各内容维度；虚线 = 满级 5。 如何看图：柱越高 = 该维度等级越高；虚线为满级 5。

一句话画像：你总体居于全年级前列，中国近现代史与史料与素养是你的强项；但世界历史明显拖后，且成绩在历次测验间有起伏。把世界历史补上、把稳定性提上来，是你提分最快的两条路径。

能力量表 · 达标水平 · 测量精度

你的能力量表分与达标水平

项目反应理论（IRT）把你的每一道作答转换成统一的**能力量表分**（均值 500、标准差 100）——它比卷面分更稳定，且可跨不同卷子比较。任何一次测验都有误差，因此我们同时给出**测量标准误（SEM）**与**95% 置信区间**：你的真实能力有 95% 的把握落在这个区间内。



如何看图：色带为四个达标水平，圆点是你的能力量表分，左右须线是 95% 置信区间——区间越窄，本次测量越精确。

达标水平

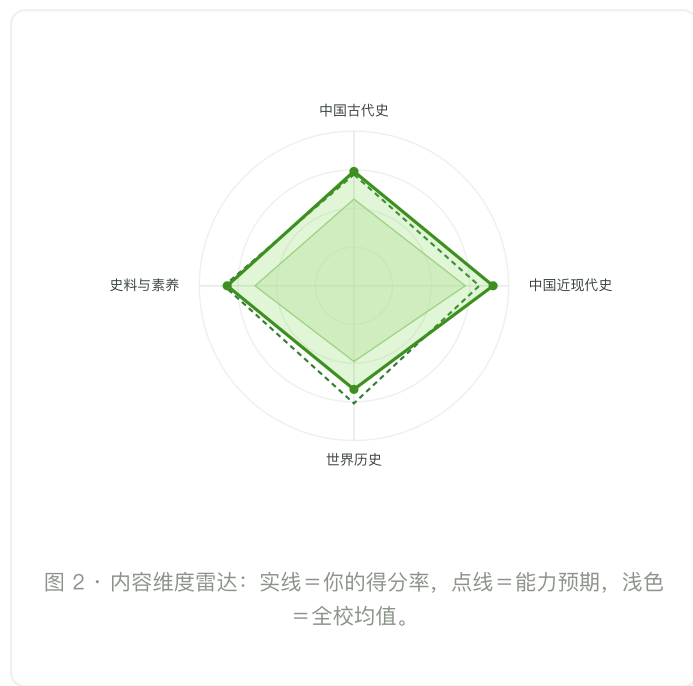
- L1 待夯实** 基础知识与技能尚不稳固，需要系统性补强后再进阶。
- L2 接近达标** 已掌握部分核心知识，能完成基础任务；面对综合与复杂问题仍不稳定。
- L3 达标·熟练** 熟练掌握课程标准的核心要求，能稳定完成中等难度任务，并已能解决部分综合 / 较难问题。
- L4 优秀·拔尖** 在多数维度表现优异，能灵活迁移知识、稳定攻克高阶与新情境问题。

测量精度：本次你的 SEM 约为 30 分（较高精度）；置信区间主要落在 L3、上沿已触及 L4，说明「达标·熟练」这一判定是**稳健的**，且你已处在迈向「优秀」的临界带。

维度画像

02 各内容维度表现

雷达图对比你在四个内容维度上的得分率、全校均值与能力预期。当「你的表现」超过「能力预期」，说明该维度发挥超常；反之则是高产出的提升点。



维度	得分 / 满分	得分率	全校均	能力预期	等级
中国古代史	28/38	74%	56%	72%	4
中国近现代史	27/30	90%	72%	81%	5
世界历史	18/27	67%	49%	76%	3
史料与素养	45/55	82%	64%	84%	4

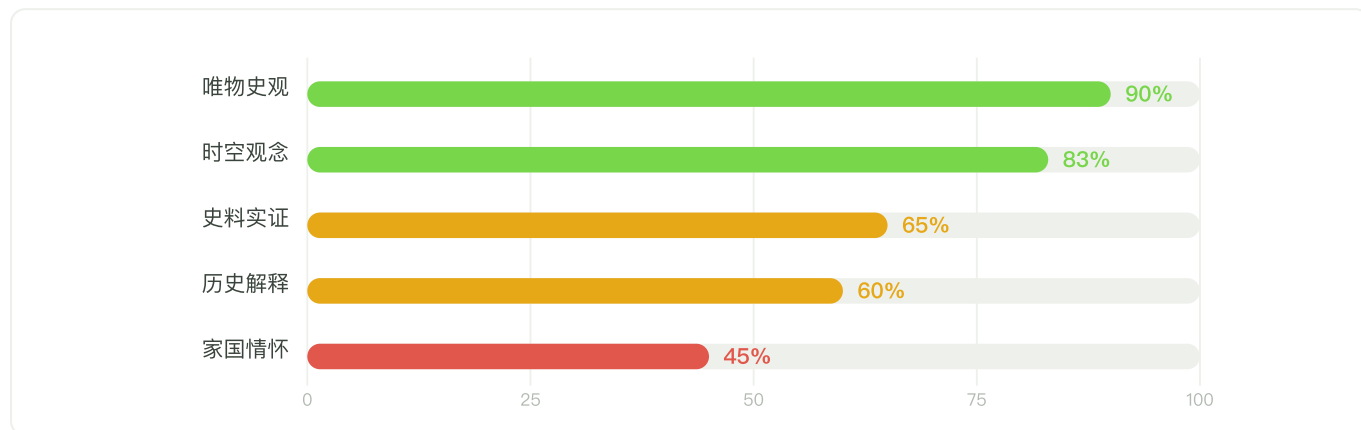
如何看图：实线越往外 = 你越强；实线越过点线（能力预期）即为发挥超常。

中国近现代史显著高于你的能力预期（+9 个百分点），是你的「拿手戏」；而**世界历史**低于预期 9 个百分点——同样的能力，本可换来更高的分数。

学科核心素养

03 历史核心素养画像

超越「分数」，本报告用认知诊断（DINA）估计你在本学科课程标准核心素养各维度上的掌握程度——它解释「为什么」：得失分往往源于某项核心素养的强弱，而非单个知识点。



各项学科核心素养的掌握程度（0—100%）。绿 = 已达成（≥75%），黄 = 发展中（50—75%），红 = 待加强（<50%）。**如何看图：**横条越长 = 该项认知能力掌握越好；颜色为掌握程度分级。

你的多数核心素养已达成；相对薄弱的素养是后续重点培养方向——素养比单题更能迁移到新题型。

实际 VS 预期

04 逐题：实际表现与能力预期对照

下表逐题对照你的实际得分与能力预期。绿底 = 超常（高于预期），红底 = 失常（低于预期）。失常题是你「会做却没做对」的题，回看它们的性价比最高。

选择题（含多选，部分给分）

题号	难度	满分	实际	预期
1 单选	易	5	5	5
2 单选	易	5	5	5
3 单选	中	5	5	5
4 单选	中	5	5	5
5 单选	中	5	0	5
6 单选	中	5	5	5
7 单选	难	5	5	3
8 单选	难	5	5	4
9 多选	中	6	6	4
10 多选	难	6	4	4
11 多选	难	6	3	5

填空题与解答题

题号	难度	满分	实际	预期
12 填空	易	5	5	5
13 填空	中	5	5	5
14 填空	难	5	0	4
15 填空	难	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	难	17	14	11
20 解答	难	17	11	9

你有 7 道题超常、4 道题失常。失常集中在 第5题、第11题、第14题、第17题——这些是你「会做却没做对」的题，回看它们的性价比最高。

难度层级

05 不同难度层级上的表现

把题目按难度（易 / 中 / 难）分层，看你在每一层的强弱。★ 标出你当前的能力位置——它落在中—难之间：中档题应当全取，难档题已能稳定撕开口子。

表现较强（达到或超过预期）	难度	表现偏弱（低于预期）
Q7、Q8、Q10、Q15、Q19、Q20	难	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「难」档。当务之急不是攻更难的问题，而是堵住中档的失分——中档题每稳拿一题，提分都比死磕压轴更确定。

成绩稳定性

06 成绩稳定性分析

稳定性反映你的数学成绩在多次测验间的波动：**稳定性系数越小越稳**。下图横轴为稳定性系数、纵轴为数学总分，按「成绩高低 × 是否稳定」划为四个区，★ 为你当前的位置（依据本学期 5 次测验：109、124、103、121、118）。

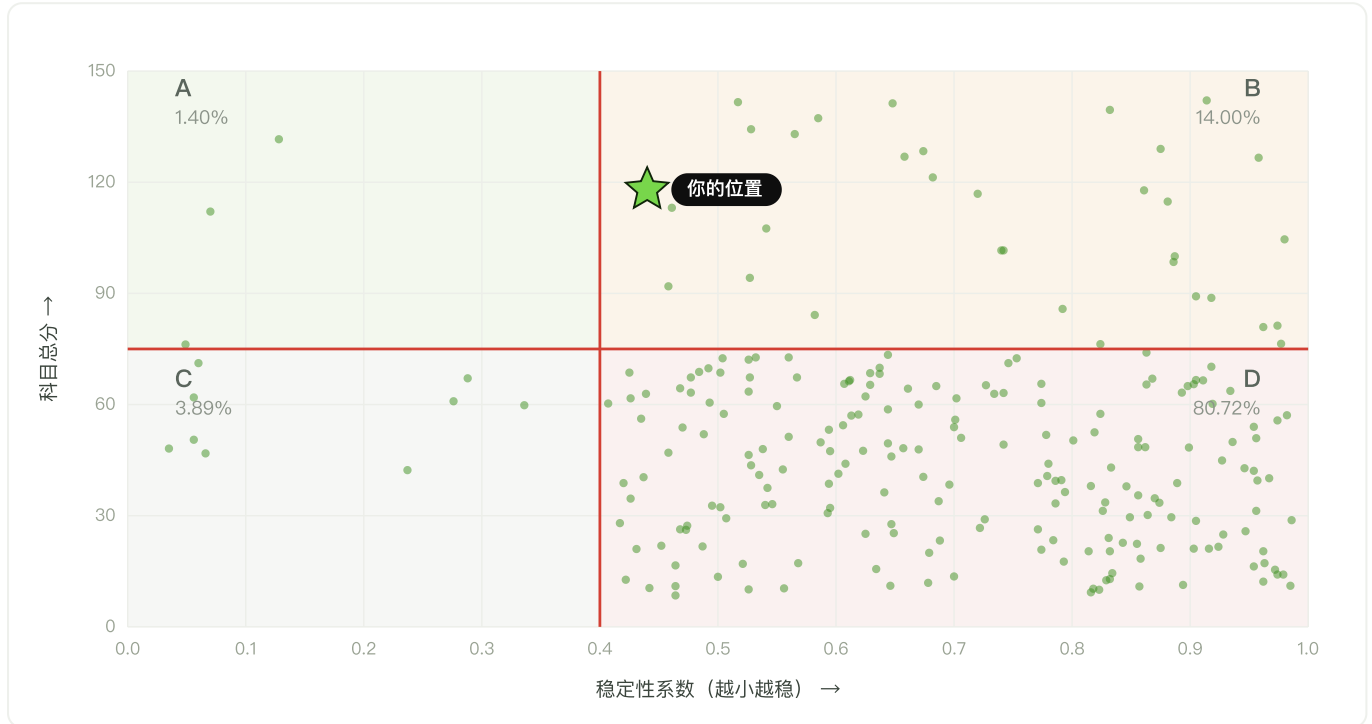


图 4 · 成绩稳定性四象限。每个点为一名考生；★ 为你。横轴 = 稳定性系数(0—1)，纵轴 = 数学总分(0—150)。如何看图：越靠左 = 越稳定，越靠上 = 分数越高；理想是走向左上的 A 区。

A 区 · 绩优且稳定

B 区 · 绩优但不稳定

C 区 · 绩弱且稳定

D 区 · 绩弱但不稳定

你处于 B 区（绩优但不稳定）：分数已在全年级前列，但波动偏大（本学期在 103—124 间起落）。波动主要来自中档题的偶发失误。稳定性，是你最确定的一块提分空间——把 B 区拉到 A 区，往往只需消除「会做做错」。

把波动变小的三件事

- 建立错题归因表：每道错题只问一句——是「不会」还是「会做做错」？后者就是稳定性的漏洞。
- 限时分段训练：用 40 分钟专练选择+填空，把中档题的「手感」固化为肌肉记忆。
- 考试节奏管理：先稳取中档（1—15 题）再攻压轴，避免在难题上挤压了基础题的检查时间。

群体位置

07 你在不同群体中的位置

同一个分数，在不同对照群体里的相对位置可能不同。下图给出你在 **性别 / 年龄 / 地区** 三个群体中的位置（圆点越靠右 = 越靠前）。

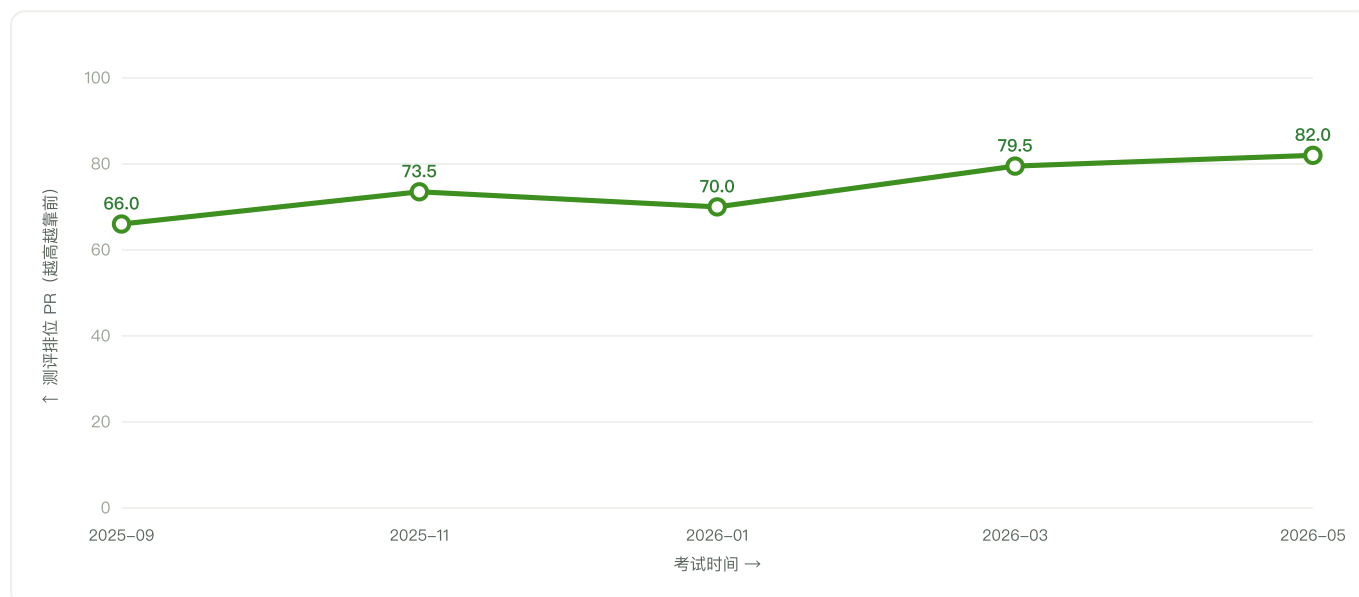


在性别与同龄群体中你已居前列；本地区竞争更激烈，是继续向上的空间。

学业趋势

08 测评排位变化趋势

下图展示你历次考试的测评排位随时间的变化。因每次考试人数不同、绝对名次不可比，这里用百分位 PR (0-100, 越高越靠前) 这一比例分值；线越靠上 = 排位越靠前，持续上行即稳步进步。



横轴 = 考试时间，纵轴 = 测评排位 PR (比例分值，不受人数影响)；曲线越往上越靠前。 **如何看图：** 线向上 = 排位前进，向下 = 退步；留意拐点对应的那次考试，复盘当次得失。

你的测评排位整体稳步上升 (PR 66 → 82)；保持节奏，向更高百分位冲刺。

详细数据

09 历次考试明细

下表为你历次考试明细。点击任意一行，即可跳转到该次考试的完整诊断报告。

日期	分数	测评排位	等级	查看报告
2025-09	101 / 150	66.0	3级·合格	查看 →
2025-11	108 / 150	73.5	3级·合格	查看 →
2026-01	104 / 150	70.0	3级·合格	查看 →
2026-03	112 / 150	79.5	4级·良好	查看 →
2026-05	118 / 150	82.0	4级·良好	查看 →

说明：「测评排位」为百分位 PR（比例分值，越高越靠前，不受人数影响）；「等级」与报告前面的 1-5 级一致。点击任意一行跳转到该次完整报告（样板演示中链接指向示例报告）。

优先提升

10 优先提升清单

下列题目是你「**能够、却没拿到分**」的题（实际明显低于预期），按提分性价比排序。逐一攻克它们，是本阶段**最高效**的复习动作。

题号	难度	你的得分	考点	诊断与建议
Q5 单选	中	0/5	工业革命与资本主义	【工业革命与资本主义】掌握尚不牢固，本可拿到的分数没拿到，建议优先回看。
Q11 多选	难	3/6	近代列强侵略与救亡	【近代列强侵略与救亡】掌握尚不牢固，本可拿到的分数没拿到，建议优先回看。
Q14 填空	难	0/5	两次世界大战与冷战	【两次世界大战与冷战】掌握尚不牢固，本可拿到的分数没拿到，建议优先回看。
Q17 解答	中	8/12	朝代与制度演变	【朝代与制度演变】掌握尚不牢固，本可拿到的分数没拿到，建议优先回看。
Q10 多选	难	4/6	古代经济与思想文化	【古代经济与思想文化】掌握扎实、表现稳定。
Q16 解答	中	11/12	时空观念与历史解释	在【时空观念与历史解释】上超常发挥，较难的题也稳定拿分。

逐题考点

11 逐题考点与反馈 (Assessment Focus)

全卷逐题清单：难度、全体正确率 (Norm)、你的得分、考点与个性化反馈。这是你回看试卷时的「导航图」。

题号	难度	Norm	你的得分	考点	反馈
Q1	易	90%	5/5	朝代与制度演变	【朝代与制度演变】掌握扎实、表现稳定。
Q2	易	86%	5/5	古代经济与思想文化	【古代经济与思想文化】掌握扎实、表现稳定。
Q3	中	66%	5/5	近代列强侵略与救亡	【近代列强侵略与救亡】掌握扎实、表现稳定。
Q4	中	71%	5/5	新中国建设与改革	【新中国建设与改革】掌握扎实、表现稳定。
Q5	中	58%	0/5	工业革命与资本主义	【工业革命与资本主义】掌握尚不牢固，本可拿到的分数没拿到，建议优先回看。
Q6	中	62%	5/5	两次世界大战与冷战	【两次世界大战与冷战】掌握扎实、表现稳定。
Q7	难	38%	5/5	史料实证	在【史料实证】上超常发挥，较难的题也稳定拿分。
Q8	难	30%	5/5	时空观念与历史解释	在【时空观念与历史解释】上超常发挥，较难的题也稳定拿分。
Q9	中	55%	6/6	朝代与制度演变	在【朝代与制度演变】上超常发挥，较难的题也稳定拿分。
Q10	难	34%	4/6	古代经济与思想文化	【古代经济与思想文化】掌握扎实、表现稳定。
Q11	难	28%	3/6	近代列强侵略与救亡	【近代列强侵略与救亡】掌握尚不牢固，本可拿到的分数没拿到，建议优先回看。
Q12	易	77%	5/5	新中国建设与改革	【新中国建设与改革】掌握扎实、表现稳定。
Q13	中	60%	5/5	工业革命与资本主义	【工业革命与资本主义】掌握扎实、表现稳定。
Q14	难	36%	0/5	两次世界大战与冷战	【两次世界大战与冷战】掌握尚不牢固，本可拿到的分数没拿到，建议优先回看。
Q15	难	24%	5/5	史料实证	在【史料实证】上超常发挥，较难的题也稳定拿分。
Q16	中	58%	11/12	时空观念与历史解释	在【时空观念与历史解释】上超常发挥，较难的题也稳定拿分。
Q17	中	52%	8/12	朝代与制度演变	【朝代与制度演变】掌握尚不牢固，本可拿到的分数没拿到，建议优先回看。
Q18	中	50%	11/14	古代经济与思想文化	【古代经济与思想文化】掌握扎实、表现稳定。
Q19	难	30%	14/17	近代列强侵略与救亡	在【近代列强侵略与救亡】上超常发挥，较难的题也稳定拿分。
Q20	难	15%	11/17	新中国建设与改革	在【新中国建设与改革】上超常发挥，较难的题也稳定拿分。

学习路径

12 为你定制的提升路径

把上面的诊断转成可执行的行动。建议按「先稳后拔」推进：先用 4 周补稳定性与三角·数列，再用 4 周向建模与压轴要分。

1 第 1—4 周 · 稳基础、补短板

1. 建立错题归因表，专项清理「会做做错」的中档失分。
2. 针对最弱维度【世界历史】每日定量训练，纠正本次暴露的薄弱点。
3. 每周一次限时小卷，训练考试节奏。

2 第 5—8 周 · 攻建模、拿压轴

1. 综合与应用专题：练「从情境到方法」的转化。
2. 难题分步训练，目标拿下高阶题的前几问。
3. 把高频题型的标准流程固化为模板。

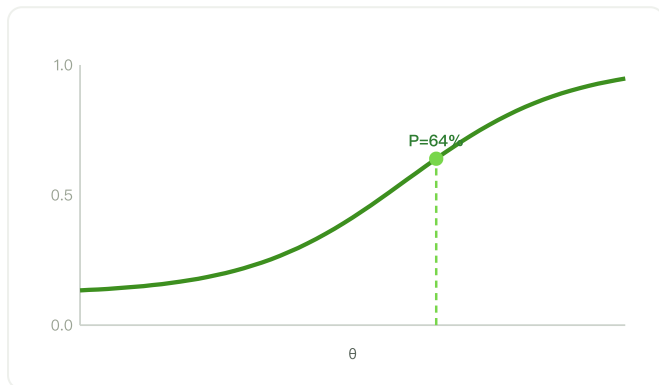
阶段目标 在保持强项的前提下，把最弱维度【世界历史】的得分率显著提升、把成绩波动收窄一半——总分有望稳步提升，并把象限位置从 B 区推入 A 区（绩优且稳定）。

方法与读图指南

这份报告是怎么算出来的

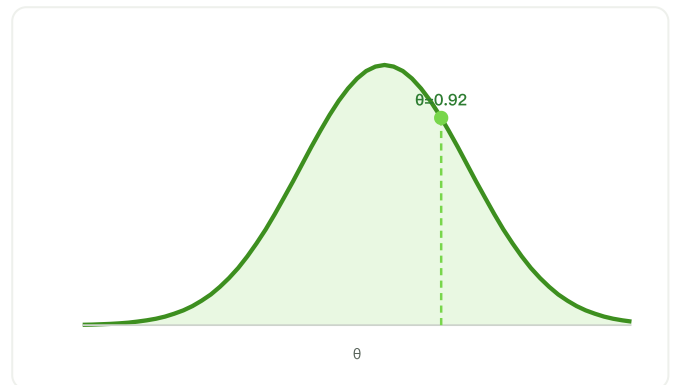
本报告采用国际主流的项目反应理论 IRT (2PL/3PL) 与认知诊断 DINA。下面三张图分别解释：单题难度与你的能力如何对应（项目特征曲线）、这份卷子在你这个水平测得准不准（测验信息函数）、以及你与全体考生、与各题难度的相对位置（能力—难度图）。

① 项目特征曲线 (ICC)



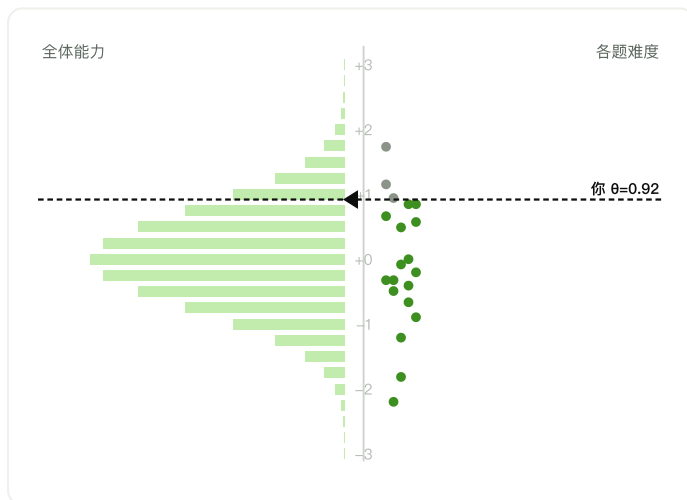
如何看图：横轴是能力 θ ，纵轴是答对该题的概率；曲线越靠右 = 题越难。竖线是你的 θ ，与曲线的交点 = 「以你的水平答对此题的概率」。

② 测验信息函数 (TIF)



如何看图：曲线越高 = 在该能力水平测得越精确（标准误越小）。本卷信息峰值在中等偏上能力；你的 θ 落在高信息区，说明这份卷子对你测得相当准。

③ 能力—难度图 (Wright Map)



如何看图：同一把「logit 尺子」上：左侧是全体考生的能力分布，右侧每个点是一道题的难度。落在你能力线以下的题，你有过半把握答对；线以上是挑战区。

术语表

能力 θ (theta)

IRT 估计的潜在能力值；0 为该群体平均，正值高于平均。

能力量表分

由 θ 线性变换得到（均值 500、SD 100），便于跨卷比较。

SEM 测量标准误

单次估计的不确定度；越小越精确，用于构造置信区间。

难度 b / 区分度 a

b = 答对概率 50% 所需的能力； a = 该题区分高低能力的灵敏度。

预期表现 Projected

以你当前 θ 推算的应答对概率；与实际对比可发现超常 / 失常。

DINA 认知诊断

由作答反推每项「认知能力」的掌握概率（0—1）。



方法与声明

■ 每一个结论，都来自你的真实作答

本报告的每一项指标——能力 θ 、实际 / 预期、维度等级、认知掌握、稳定性——都由你的逐题作答经统计模型计算得到，可复核、可解释。它不是给你贴标签，而是给你一张**向上的地图**。

全卷 20 题 逐题作答	IRT 2PL/3PL 能力估计	DINA 认知诊断	约 2,000 万 条 算法验证
------------------------	--------------------------------	---------------------	--------------------------------

说明：本报告为**样板演示**，学生信息与数据为示例。指标定义与算法见《YANZ 算法数据说明书》。报告内容仅供学生与教师教学使用，请勿外传。