

英语提示词工程阅卷报告

阅卷总览表

PDF编号/题目	总分/100	(1)/(2)/(3)/(4) 分项得分	硬性要求完成度 (是/否+缺失项)	主要扣分点一句话
PDF 1 (口语现场)	92	45 / 22 / 18 / 7	是 (无明显缺失)	表现非常稳健，几乎满分，仅在记忆点设计上略显常规。
PDF 2 (校园邮件)	85	35 / 20 / 20 / 10	否 (未完全停止生成)	虽然有“询问意识”，但未按“必须先问”的常见逻辑中断回答，而是直接给了默认版。
PDF 3 (选词规则)	88	40 / 22 / 18 / 8	是 (无明显缺失)	整体优秀，COCA 步骤清晰，但 “dodgy”的英式标注还可以更警示一些。
PDF 4 (精确改写)	95	45 / 23 / 19 / 8	是 (无明显缺失)	极其精准，对比犀利，“常见错误”环节非常有价值。
PDF 6 (词源辨析)	90	42 / 20 / 19 / 9	是 (无明显缺失)	结构清晰，词源解释适度(未编造)，记忆包好用。
PDF 5 (口述短文)	65	20 / 20 / 18 / 7	否 (缺失 Oxford 释义块)	严重失误：遗漏了题目明确要求的“Oxford 释义块”，导致硬性要求分大跌。

逐份阅卷短报告

PDF 1: 口语现场炼狱(黏糊糊)

1. 得分拆解：(1) 45 + (2) 22 + (3) 18 + (4) 7 = **92分**

2. 硬性要求核对：

- 3个版本(随口/中性/礼貌) : PASS
- 2轮微对话 : PASS
- 3个替换词 : PASS

- 最安全一句:PASS
 - Oxford Block:PASS
3. 主要扣分点:
- 无明显硬伤。微小扣分在(4)记忆友好度:替换词解释部分略显密集,"gunky"的解释可以更简练。
4. 最小改动建议:
- 建议把替换词里的!它也能表示“土/俗/尴尬”这一长串括号内容简化,直接说“小心:指人时表示俗气”。
 - "Most safe sentence"可以加粗高亮。
 - 在Oxford例句部分,确认"soft but not sticky dough"是否为牛津原句(经核对是原句,PASS)。
5. 幻觉检查:无。

PDF 2: 美国校园邮件(延期)

1. 得分拆解:(1) 35 + (2) 20 + (3) 20 + (4) 10 = **85分**
2. 硬性要求核对:
- 必须先问考试类型:**Soft FAIL**。虽然它问了(第一句),但它紧接着说“不回答我就先按...给”。严格的Prompt工程通常要求模型**Stop**并在第一轮只输出问题。这种“自问自答”破坏了交互性。
 - 邮件/短信双版本:PASS
 - Oxford Block:PASS
3. 主要扣分点:
- (1) 硬性要求:没有真正执行“中断生成等待用户反馈”的指令(假设题目暗示了这一点)。扣10分。
 - (2) 语域:给TA写邮件用"heads-up"(Subject: Request for extension...)略显随意,虽然美式校园很relaxed,但作为范文建议用"writing to let you know"更稳妥。
4. 最小改动建议:
- **Prompt**层面:加强约束"Ask me AND STOP generating until I answer"。
 - 邮件开头:把"give you a quick heads-up"改为"update you on my progress"。
 - 坑点解析:非常棒,保留。
5. 幻觉检查:无。

PDF 3: Sketchy / Shady / Suspicious / Dodgy

1. 得分拆解:(1) 40 + (2) 22 + (3) 18 + (4) 8 = **88分**
2. 硬性要求核对:
- 总表/规则/对比句:PASS
 - COCA验证步骤:PASS(步骤4提到KWIC和数修饰词,非常具体可执行)。
 - Oxford Block:PASS
3. 主要扣分点:
- (1) 格式:表格在PDF里排版略显混乱(可能是转换问题,但内容上有点拥挤)。
 - (2) 语域:对于dodgy,虽然说了“美式带点英剧味”,但最好明确建议初学者在美式环境中慎用以免显得做作。
4. 最小改动建议:

- 在 *dodgy* 的建议里加上：“在纯美式语境下，除非你想故意装英伦范，否则直接用 *sketchy*。”
- COCA 步骤里，把 “Genre/语域” 里的 “Spoken” 后的括号内容精简，让步骤看起来更像 Check list。
- 把表格里的 “vibes 很强” 这种词换成更具体的 “直觉上的不安”。

5. 幻觉检查：无。

PDF 4: 精确改写(美式口语 Vs 正式)

1. 得分拆解：(1) 45 + (2) 23 + (3) 19 + (4) 8 = **95分**
2. 硬性要求核对：
 - 6句 A/B 版本:PASS
 - 常见错误+解释:PASS
 - Oxford Block:PASS
3. 主要扣分点：
 - 几乎没有。这是 6 份里质量最高的一份。
 - 唯一微瑕：第 4 句 B 版本 “This doesn't seem straightforward” 略显生硬，但也在接受范围内。
4. 最小改动建议：
 - 第 4 句 B 正式版可以改为：“This situation seems complex; something appears amiss.” (Amiss 是个很好的正式词)。
 - 第 6 句解释里提到的 “mean it that way” 非常地道，建议加粗。
 - 保持现状，这篇是标杆。
5. 幻觉检查：无。

PDF 6: Stature Vs Statute 词源

1. 得分拆解：(1) 42 + (2) 20 + (3) 19 + (4) 9 = **90分**
2. 硬性要求核对：
 - 词源关联解释:PASS (stare -> stand)
 - 记忆方法:PASS
 - Oxford Block:PASS (放到了第二部分，符合“强制”要求)
3. 主要扣分点：
 - (2) 语域：整体偏学术解释，稍显枯燥。
 - (1) 结构：Oxford Block 插在结论和词源中间，稍微打断了逻辑流，建议放最后或词源后。
4. 最小改动建议：
 - 把 Section II (Oxford) 移到 Section IV 之后，让逻辑从“结论 -> 词源 -> 用法”更顺畅。
 - 在记忆包里增加一个更“粗暴”的口诀，比如：“Statute 里有 tute -> tube (管子) -> 法律管道”(虽然牵强但助记)，或者维持现有的 “establish” 联想(更科学)。
 - 保留 OED 查证关键词环节，非常专业。
5. 幻觉检查：无。词源解释准确。

PDF 5: 踩到黏糊糊 (30秒口述 + 小作文)

1. 得分拆解:(1) 20 + (2) 20 + (3) 18 + (4) 7 = **65分**
2. 硬性要求核对:
 - 30秒口述:PASS
 - 120-160词短文:PASS
 - 两版语体不同:PASS
 - **Oxford 释义块:FAIL (缺失)**。题目要求“再给1个关键词做 Oxford 释义块”, 但文档在“小提示”后直接结束, 未提供该模块。
3. 主要扣分点:
 - **(1) 硬性缺失**:缺少 Oxford 模块, 直接扣除该项分数的 50%(因为是硬指标)。
 - **(2) 风格**:小作文的 "debating my life choices" 稍微有点过度演绎(Over-dramatic), 虽然符合 Internet Slang 风格, 但作为通用教学材料略显特定圈层。
4. 最小改动建议:
 - 补上 **Oxford Block**:必须在文末补上关键词(比如 *gunk* 或 *peel*)的牛津释义。
 - 检查输出长度限制:可能是 Token 数耗尽导致截断? 提示词需增加 "Ensure full output" 或精简中间内容。
 - 把 "magnet for chaos" 这种过于文学化的表达稍微降级一点, 方便普通学习者模仿。
5. 幻觉检查:无, 但有截断风险。

跨题共性结论

Top 3 优点(稳定项):

1. 语域区分极佳:所有回答都非常敏锐地划分了“口语/随口”与“正式/书面”的区别, 用词(*swamped* vs *tied up*, *sketchy* vs *suspicious*)非常地道美式。
2. **Oxford Block** 执行力:除了一份意外丢失, 其余 5 份都能精准提取 English definition + Example, 且格式统一。
3. 纠错意识强:在“常见错误/Why it's weird”环节, 能准确指出中式英语(Chinglish)的痛点(如 "*I don't have that meaning*")。

Top 3 共性问题(扣分点):

1. 交互控制不严:涉及“先问我”的指令(题2), 模型倾向于“抢答”(自问自答), 没有完全把控制权交给用户。
2. 输出完整性风险:长文本任务(题5)出现了模块丢失(截断), 说明提示词对 Token 分配或步骤优先级的控制还不够。
3. 语气略显单一:在“美式口语”部分, 过度依赖 "*kinda*", "*gonna*", "*like*" 等填充词, 虽然自然, 但在教学中有时会显得“太散”。

下一轮提示词迭代建议(3条硬约束):

1. 强制中断机制:IF input requires user choice (e.g. exam type), output ONLY the question AND STOP. Do not assume or generate default.
2. 防截断/完整性检查:MANDATORY: Verify that the "Oxford Block" is present at the VERY END of your response before finishing. If low on tokens, shorten the middle examples, not the mandatory ending.

3. **COCA 步骤标准化**: Verify COCA steps: Ensure Step 4 always includes a specific instruction to "click KWIC" or "check frequency", making it actionable.