

Appendix A: Scoring Rubrics

Table 1: The Developed Rubrics (Translated from Chinese)

	Content	Use of Language
A	<ul style="list-style-type: none"> - The central theme is distinct and clear; the content is rich, detailed, and closely aligned with the topic. - The structure is well-organized with clear, logical progression. - The content revolves around the keywords of the topic, approaching from various angles. 	<ul style="list-style-type: none"> - The language flows smoothly, with precise expression; word choice is appropriate and varied. - Sentence structures are diverse, not just short sentences. - The word count is adequate, not too few. - Rhetorical and descriptive techniques are varied and skillfully applied, closely integrated with the theme of the composition, effectively enhancing the expressive effect.
B	<ul style="list-style-type: none"> - The central theme is clear; the content is fairly substantial, fitting the topic well. - The structure is relatively clear, with fairly logical organization. - The content revolves around the keywords of the topic, with some degree of perspective. 	<ul style="list-style-type: none"> - The language is relatively smooth, with fairly accurate expression; word choice is suitable. - Sentence structures are relatively varied, but may lack complexity. - Rhetorical and descriptive techniques are fairly comprehensive, but slightly inferior to the highest grade in terms of variety, depth, or creativity. - The word count is sufficient, but may lack richness.
C	<ul style="list-style-type: none"> - The central theme is basically clear; the content is basically substantial and aligns with the topic. - The structure is acceptable, with basic logical progression. - The content revolves around the keywords of the topic, but the perspective may be one-dimensional. 	<ul style="list-style-type: none"> - The language is basically smooth, with basically accurate expression; word choice is fairly appropriate. - Sentence structures show some variation, but may be overly simple. - Basic use of rhetorical and descriptive techniques, limited in variety and application, possibly lacking in expressive appeal. - The word count may be on the lower side.
D	<ul style="list-style-type: none"> - The central theme is not clear enough; the content is insufficient, and most of it is not closely related to the topic. - The structure is not apparent, with unclear organization, sometimes appearing scattered. - The content lacks development around the keywords of the topic, with a one-dimensional perspective or lack thereof. 	<ul style="list-style-type: none"> - The language lacks smoothness, with inaccurate expression; word choice is simple. - Sentence structures are monotonous, and lacking in variation. - The composition lacks effective rhetorical and descriptive techniques, or their application is inappropriate, failing to enhance the overall quality of the work. - The word count is inadequate, and lacking in detail.

Appendix B: A Sample Essay and LLMs’ Feedback

Table 2: A Sample Essay From a K3 Student and LLMs’ Scoring Results

<p>Chinese version:</p> <p>题目：三千年后的地球人</p> <p>在美国莱州的时候，他们制作了一个钢铁般的盒子，这个盒子里有一封给地球人的信，但是这信里面写着什么，地球人还都不知道。像每天这样，三千年马上就来临了。在来到三千年之前，我看到了这里有堆积如山的垃圾，汽车里的二氧化碳已经污染了城市的一大部分了，我好希望他们不要这样下去了，我又来到了原先如此美丽的花园，现在已经变得非常破旧了，也再也没有往日那些美丽的花了。我又来到了特高特高的大厦，但现在，垃圾都把他的美丽给挡住了，这里的学校也是成了很废旧的一个小学校了，并且那里的小学生想干什么就干什么，一点也不紧张。忽然，我的背后忽然跑出一个，我看见了，是一个披着黑衣服的一个人，我急忙追去，他给我带进了一片大森林，树上有个纸条说：“在三千年以后，就在这个地区，会有一个对人们的一封信的。”我忽然觉得我在做梦，我在想，在那里有宝藏，我怎么才能拿到那的一封信呢，信里写的又是什么呢？我对这件事产生了疑问，我想，这也有大森林，我干嘛不在这里看看呢！于是，我急忙跑去了大森林里，但是，这里面根本没有纸条。一眨眼，三千年过去了，我急忙跑去原来的树林里，找了几个挖掘工，他们一起挖呀！挖呀！最后，我终于看见了一个箱子，可是这个箱子怎么也打不开，我想，说不定那个黑衣人就能打开这个宝箱呢！所以，我就开始去寻找那个黑衣人，最后，我找到黑衣人的房子，我看见黑衣人正在睡觉，我蹑手蹑脚的溜进了黑衣人的房间里，我看见了她的黑衣服里，还真有一把钥匙，这把钥匙被我偷走了，但是，黑衣人早就已经知道，只不过他也想环境变得好一些，所以才让他拿走的，突然，宝箱打开了，这里面写着：“三千年后的人类们，我希望你能改变你们的主意，让地球重新变得美好起来。”这让三千年后的地球人看见了，决定改掉这个毛病。最后，他们改掉了这个毛病，大道上不再是那么多的垃圾了，大厦也没有那么多的垃圾了，大森林里的树木又变的茂盛起来了，大家看到这美丽的景象，都开心了起来，人们看到了这样的景象，也觉得早应该这样知错就改了。同时大家都感谢我，都说是我给了他们的教育，我说：“你们只要以后不在破坏环境，就可以了，这是我应该做的。”</p> <p>Translated version:</p> <p>Title: Earthlings 3,000 Years Later</p> <p>When I was in Lai State in the United States, they made a steel-like box with a letter for the people of Earth inside. However, no one on Earth knew what was written in that letter. Now, as the three thousand years approach, I see piles of garbage towering like mountains. The carbon dioxide from cars has polluted a large part of the city. I hope they will not continue like this. The once beautiful garden has become very worn out, with no more beautiful flowers as before.</p> <p>I visited tall skyscrapers, but now their beauty is hidden by garbage. The school here has also turned into a run-down primary school, where the students do as they please without a care in the world. Suddenly, someone ran up behind me. It was a person dressed in black. I chased after him and he led me into a large forest. There was a note on a tree that said: "Three thousand years from now, in this area, there will be a letter for the people." It felt like a dream. I wondered if there was treasure there. How could I get that letter and what did it say? I had doubts about this. There's a big forest here. Why not take a look around here? So, I quickly ran into the forest, but there was no note inside.</p>

In the blink of an eye, three thousand years passed. I hurried back to the original forest and found some excavators to dig together! Finally, I saw a box, but no matter how hard I tried, I couldn't open it. Maybe the man in black could open this treasure chest! So, I started looking for the man in black. Eventually, I found his house. He was asleep. I sneaked into his room and found a key in his black clothes. I stole the key, but he knew it already. He wanted to improve the environment, so he let me take it. Suddenly, the chest opened, and it read: "Three thousand years later, I hope you change your mind and make Earth beautiful again." The Earthlings of three thousand years later saw this and decided to change their ways.

In the end, they changed their ways. There was less garbage on the streets, fewer skyscrapers filled with litter, and the trees in the forest became lush again. Everyone was happy to see this beautiful sight. They knew they had to change their ways earlier. Everyone thanked me and said I gave them education. I said, "As long as you do not continue to destroy the environment, that's all I ask for. It is what I should do."

Human Rating: D

GPT-4's responses:

[\[Zero-shot, w/o rubrics, w/o explanation\]](#)

等级: B

Level: B

[\[Zero-shot, with rubrics, w/o explanation\]](#)

等级: B

Grade: B

[\[Zero-shot, with rubrics, with explanation\]](#)

评分依据: 作文内容紧扣题目, 以书信的形式表达了对三千年后地球人的期望, 中心明确。结构层次分明, 条理清晰, 内容丰富, 从环保的角度切入。语言表达准确, 句式多样, 用词丰富。修辞和描写手法丰富, 有效增强了表达效果。字数适当, 不太少。

等级: A

Rating Criteria: The composition sticks closely to the topic, expressing expectations for Earth dwellers three thousand years later in the form of a letter with a clear focus. The structure is well-organized, logic is clear, and the content is rich, from the perspective of environmental protection. Language expression is accurate, with diverse sentence structures and rich vocabulary. The use of rhetorical and descriptive techniques enriches the piece and effectively enhances the expressive effect. The word count is appropriate, not too few.

Grade: A

[\[Few-shot-retrieval-based-1, with rubrics, w/o explanation\]](#)

等级: B

Level: B

[\[Few-shot-retrieval-based-1, with rubrics, with explanation\]](#)

评分依据: 作文的中心是环保, 内容基本符合题意, 但是与题干的关联不紧密, 没有明确表达出是一封信的形式。结构层次尚可, 条理基本清晰。语句基本流畅, 表达基本准确; 用词尚恰当。句式有一定变化, 但可能过于简单。基本的修辞和描写手法, 种类和运用有限, 可能在表达效果上缺乏吸引力。字数可能偏少。符合 C 类作文评分标准。

等级: C

Assessment Criteria: The focus of the composition is on environmental protection and the content basically meets the theme, but its connection with the topic is not tight and it doesn't explicitly express the format of a letter. The structural level is acceptable, and the logic is basically clear. Sentences are generally smooth, and expressions are basically accurate; the use of words is generally appropriate. Sentence structures show some variations, but might be overly simplified. Basic rhetorical devices and descriptive techniques are limited in variety and application, which might result in a lack of attraction in conveying the message. The word count might be on the lesser side. It meets the C-level composition assessment standards.

Rating: C

[\[Few-shot-retrieval-based-2, with rubrics, with explanation\]](#)

评分依据：作文内容大部分在描述一个寻找和打开盒子的冒险故事，而非直接以书信形式表达对未来人的期望或建议。信件内容在作文末尾简短提及，与题干要求的书信形式相偏离。结构上虽有一定的故事性，但与题目要求的书信形式不符，内容与题意关联不紧密。符合 D 类作文评分标准。

等级：D

Translation Basis: The majority of the composition describes an adventure of searching for and opening a box, rather than directly expressing hopes or advice to future people in the form of a letter. The letter content is briefly mentioned at the end of the composition, which deviates from the required letter format of the prompt. Although the composition has a certain narrative structure, it does not comply with the demanded letter form, and the content is loosely related to the topic. This meets the scoring standards of Grade D.

Grade: D

Fine-tuned models' responses:

[\[Fine-tuned Qwen-1.8B-Chat\]](#)

等级：D

Level: D

[\[Fine-tuned GPT-3.5-turbo\]](#)

等级：D

Level: D