



My experiment using ChatGPT to assess writing

Mario Lecluyze

turns to AI for help.

While I love teaching English to teens, I find grading their writing assignments and providing meaningful feedback less enjoyable.

The main reason for this: it is a very time-consuming and repetitive task, especially when you have almost 100 students. Moreover, I always feel a little insecure about the objectivity of my assessment and grades, even though I use rubrics with specific criteria. I have come across blog discussions about using AI to assist with grading, sparking my curiosity. So, after some hesitation ('Is this not my job?'), I decided to put it to the test and had ChatGPT do it for me. I want to share my findings with you.

The task

My students had to write an email as the final communicative task at the end of a unit in which we had studied the American educational system and, in particular, what life at an American high school is like.

As a first step, they had to watch a YouTube video in which a student presented a typical day in her high school. While watching, my students had to take notes and focus especially on the similarities and differences between that school and ours.

Next, they had to write the draft of an email to that student on their laptops. In their drafts, students were asked to:

- mention they had watched the video;
- describe what life at our school is like compared to hers; and
- ask at least two interesting questions about her school life and whether she was interested in becoming a pen pal.

In this phase, they were allowed to use an online dictionary, but no translating apps.

After that, in a peer correction moment, they had to sit together in pairs and swap laptops to revise each other's work. I always do this and it results in better

writing products. Finally, they needed to copy the revised draft into the mail application Outlook and mail it to me.

The rubric

Before they started writing, the students also got the rubric with the criteria that were going to be used to assess their email. In this respect, they knew beforehand what they needed to pay attention to. I always use a so-called 'rubric-of-one', giving them a global description of the expectation with some space to write what they have done well and how they can still grow. The following were the criteria for this task, with the score in brackets, adding up to a total score of 20 points.

- Content: you have written an interesting email with similarities and differences (6 points).
- Register: you have used the tone and conventions for a formal email (3 points).

- Referring to clip: you have mentioned several elements from the clip and asked interesting questions (3 points).
- Spelling: there were no mistakes in your text that you could have avoided with your spelling checker or by proof reading (2 points).
- School vocabulary: you have used the (new) vocabulary that was studied in the unit (3 points).
- Grammar: you haven't made any mistakes against the practised grammar (3 points).

The correction

With 96 student emails in my inbox I turned to ChatGPT for help. I already had a free account, so I guessed this could not be a problem. And, indeed, ChatGPT corrected them all, although after a couple of emails a warning appeared saying it would change from Version 4 to Version 3. Luckily, the correction task was still performed in the same way.

In every ChatGPT course the importance of 'prompting' is stressed. Prompts are the textual inputs (e.g. questions, instructions) that you enter into ChatGPT to get responses. ChatGPT predicts an appropriate response to the prompt you have entered. In general, a more specific and carefully worded prompt will get you better responses. I was taught that assigning a role was a good starting point. So, I told ChatGPT it was a teacher of English as a foreign language to students moving from A2+ to B1 at the CEFR. I asked it to grade their formal emails out of 20 using the criteria that I would upload (see above).

'For each email, I'll score each criterion and provide a short explanation for the score. Once you provide the emails, I'll get started!' was the answer that made me very enthusiastic.

I was even more excited after I had copied and pasted the very first email into the text box. Check out ChatGPT's reply (Figure 1).

This answer was better than I had expected. For every criterion, I got a fair score, but also extra feedback for the student. I was surprised by how nuanced the feedback was. The comments on

'Here's the evaluation for Aza's email based on the provided criteria:

1. **Content (5/6):** The email includes relevant and interesting information about similarities and differences between the schools, like: locker placement; school hours; and fire drills. There's a good range of content, but a little more detail would bring it to a full score.
2. **Register (2/3):** The tone is generally formal and polite, but a few casual phrases ('I really want to know more') slightly reduce the formality. Phrasing like 'I am writing this email to let you know' could also be more concise for added professionalism.
3. **Referring to clip (3/3):** Aza effectively references elements from the clip, including school start times, locker locations, and fire drills, and also asks relevant, engaging questions.
4. **Spelling (1/2):** There are a few spelling errors ('Firts' for 'First', 'aspecially' for 'especially', 'commen' for 'common'), along with minor issues that could have been caught with spell-check.
5. **School vocabulary (3/3):** The vocabulary aligns well with school-related topics, especially terms like 'lockers', 'fire drills', and 'lesson cancellation', which demonstrate appropriate use of unit vocabulary.
6. **Grammar (2/3):** Grammar is mostly correct, but there are minor errors, such as 'another difference is the time when school starts and ends,' which could be phrased more smoothly as 'another difference is our school start and end times'.

Total score: 16/20

This email is a solid effort for an A2+ level and effectively meets most of the criteria with only minor areas for improvement.'

Figure 1: ChatGPT response

specific vocabulary and phrasing added a layer of insight that I usually wouldn't have time to include. Without ChatGPT, the student would probably have got the same grade, but not as much extra information. However, I must admit that I also read and checked the email myself, because I wanted to make sure the student got a fair score. I did this eventually for every single student, so my aim of gaining time was not really reached. However, I was really happy with the elaborate feedback I could pass on to my students.

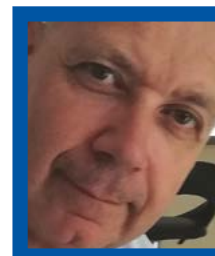
The pitfalls

I didn't copy all my students' emails in one session, but worked on them over several days. Since I have a free account, my previous sessions and the prompt were saved on ChatGPT. However, when I started a new session and uploaded the next email, it was assessed with totally different criteria than mine: 'content', 'structure and flow', 'language and grammar', and 'engagement and tone'. I had to suggest uploading my criteria again to get the same results as in the previous session. This happened every time I restarted!

If you want to try this, you will also need to keep an eye on ChatGPT's mathematical skills. It does n't always

add up scores correctly, although this only happened once in my 96 emails: one email was scored 17/20, with scores of 5+2+2+2+2+2, equalling 15.

In the end, my experience with ChatGPT didn't save me much time, but it offered a level of objectivity and feedback depth that I hadn't anticipated. It reassured me that my grading practices are on the right track, and my students benefited from detailed, individualised comments on their work. This experiment has made me optimistic about the role AI could play in education, potentially helping teachers manage grading while ensuring students receive meaningful feedback.



Mario Lecluyze is a seasoned English language teacher based in Belgium, with over 35 years of experience. He specialises in English as a foreign language (EFL) and Content and Language

Integrated Learning (CLIL). Lecluyze has worked as a teacher trainer and lecturer at VIVES University College in Torhout, focusing on English teaching methodology and CLIL practices. He also served as an educational adviser for the Catholic Education Flanders organisation and contributed to designing English curricula for secondary education.