**How to do an AI-Based Self-Evaluation for Feedback on IELTS Academic Task 1**

**Note: No AI tool is perfect and cannot substitute for a human rater, but the goal of using an AI tool is to get as close as possible to what a human examiner would consider in an evaluation. Writing is scored subjectively, which is why a human rater is more accurate. We are instructing you on using an AI tool because most of you cannot afford to hire a private tutor or private coach and want feedback on your writing. If you do as we instruct below, you will get loads of feedback and know what you need to work on to improve. That is the goal!**

1. Evaluation 1:

To do this Self-Evaluation, I will be using two AI Tools:

1. Grammarly AI: https://app.grammarly.com/
2. ChatGpt: <https://chat.openai.com/>

Before you go through these prompts, go to [www.grammarly.com](http://www.grammarly.com) and input the response to determine the amount of grammar, mechanical and lexical mistakes. If you have the **Premium** version of Grammarly, great because that means you can see every mistake that the examiner would make.

Note: The Grammarly AI system is better than what ChatGpt and Google Bard can do in terms of Mechanics, Grammar, and Lexical Ability. I like to give the Grammarly results to ChatGpt to get a more accurate analysis, so I normally do that first. **The Premium Version** of Grammarly will give you everything that the IELTS system will notice in terms of Mechanics, Grammar, and Lexical Ability. There is a free version of Grammarly, but it is not a complete analysis.

**ChatGpt Self-Practice Self-Evaluation Prompts**

**1st Prompt: (Giving ChatGpt the right frame of mind/perspective.**

Are you familiar with the IELTS Academic Task 1 writing question? If so, please give me a brief description of this task. I need to check your understanding.

**2nd Prompt: (Giving ChatGpt the right frame of mind/perspective.**

ChatGpt Classification: You are acting in the role of an IELTS writing examiner. I am going to walk you through several prompts, and I want you to always maintain this role as an IELTS writing examiner who will make an unbiased analysis of written work. Do you understand your role? Yes or No?

**Prompt 2**

Type this ChatGpt Prompt: Here are the IELTS Task 1 Essay rubric criteria to evaluate an essay response. I only want you to read through the rubric. Tell me when you have finished reading it.

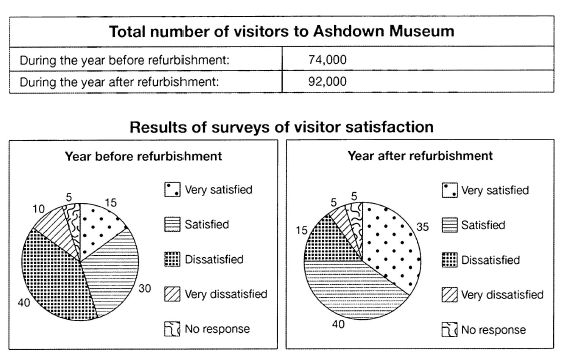
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Band Scores | Task Response | Coherence and Cohesion | Lexical Resource | Grammatical Range and Accuracy |
| **Band Score 9** | All the requirements of the task are fully and appropriately satisfied. There may be extremely rare lapses in content. | The message can be followed effortlessly. Cohesion is used in such a way that it very rarely attracts attention. Any lapses in coherence or cohesion are minimal. Paragraphing is skilfully managed. | Full flexibility and precise use are widely evident. A wide range of vocabulary is used accurately and appropriately with very natural and sophisticated control of lexical features. Minor errors in spelling and word formation are extremely rare and have minimal impact on communication. | A wide range of structures is used with full flexibility and control. Punctuation and grammar are used appropriately throughout. Minor errors are extremely rare and have minimal impact on communication. |
| **Band Score 8** | The response covers all the requirements of the task appropriately, relevantly and sufficiently. (Academic) Key features are skilfully selected, and clearly presented, highlighted and illustrated. (General Training) All bullet points are clearly presented, and appropriately illustrated or extended. There may be occasional omissions or lapses in content. | The message can be followed with ease. Information and ideas are logically sequenced, and cohesion is well managed. Occasional lapses in coherence or cohesion may occur. Paragraphing is used sufficiently and appropriately. | A wide lexical resource is fluently and flexibly used to convey precise meanings. There is skilful use of uncommon and/or idiomatic items when appropriate, despite occasional inaccuracies in word choice and collocation. Occasional errors in spelling and/or word formation may occur, but have minimal impact on communication. | A wide range of structures is flexibly and accurately used. The majority of sentences are error-free, and punctuation is well managed. Occasional, non-systematic errors and inappropriacies occur, but hav |
| **Band Score 7** | The response covers the requirements of the task. The content is relevant and accurate –there may be a few omissions or lapses. The format is appropriate. (Academic) Key features which are selected are covered and clearly highlighted but could be more fully or more appropriately illustrated or extended. (Academic) It presents a clear overview, the data are appropriately categorised, and main trends or differences are identified. (General Training) All bullet points are covered and clearly highlighted but could be more fully or more appropriately illustrated or extended. It presents a clear purpose. The tone is consistent and appropriate to the task. Any lapses are minimal. | Information and ideas are logically organised and there is a clear progression throughout the response. A few lapses may occur. A range of cohesive devices including reference and substitution is used flexibly but with some inaccuracies or some over/under use. | The lexical resource is sufficient to allow some flexibility and precision. There is some ability to use less common and/or idiomatic items. An awareness of style and collocation is evident, though inappropriacies occur. There are only a few errors in spelling and/or word formation and they do not detract from overall clarity. | A variety of complex structures is used with some flexibility and accuracy. Grammar and punctuation are generally well controlled, and error-free sentences are frequent. A few errors in grammar may persist, but these do not impede communication. |
| **Band Score 6** | The response focuses on the requirements of the task and an appropriate format is used. (Academic) Key features which are selected are covered and adequately highlighted. A relevant overview is attempted. Information is appropriately selected and supported using figures/data. (General Training) All bullet points are covered and adequately highlighted. The purpose is generally clear. There may be minor inconsistencies in tone. Some irrelevant, inappropriate or inaccurate information may occur in areas of detail or when illustrating or extending the main points. Some details may be missing (or excessive) and further extension or illustration may be needed. | Information and ideas are generally arranged coherently and there is a clear overall progression. Cohesive devices are used to some good effect but cohesion within and/or between sentences may be faulty or mechanical due to misuse, overuse or omission. The use of reference and substitution may lack flexibility or clarity and result in some repetition or error | The resource is generally adequate and appropriate for the task. The meaning is generally clear in spite of a rather restricted range or a lack of precision in word choice. If the writer is a risk-taker, there will be a wider range of vocabulary used but higher degrees of inaccuracy or inappropriacy. There are some errors in spelling and/or word formation, but these do not impede communication. | A mix of simple and complex sentence forms is used but flexibility is limited. Examples of more complex structures are not marked by the same level of accuracy as in simple structures. Errors in grammar and punctuation occur, but rarely impede communication. |
| **Band Score 5** | The response generally addresses the requirements of the task. The format may be inappropriate in places. (Academic) Key features which are selected are not adequately covered. The recounting of detail is mainly mechanical.There may be no data to support the description. (General Training) All bullet points are presented but one or more may not be adequately covered. The purpose may be unclear at times. The tone may be variable and sometimes inappropriate. There may be a tendency to focus on details (without referring to the bigger picture). The inclusion of irrelevant, inappropriate or inaccurate material in key areas detracts from the task achievement. There is limited detail when extending and illustrating the main points. | Organisation is evident but is not wholly logical and there may be a lack of overall progression. Nevertheless, there is a sense of underlying coherence to the response. The relationship of ideas can be followed but the sentences are not fluently linked to each other. There may be limited/overuse of cohesive devices with some inaccuracy. The writing may be repetitive due to inadequate and/or inaccurate use of reference and substitution. | The resource is limited but minimally adequate for the task. Simple vocabulary may be used accurately but the range does not permit much variation in expression. There may be frequent lapses in the appropriacy of word choice and a lack of flexibility is apparent in frequent simplifications and/or repetitions. Errors in spelling and/or word formation may be noticeable and may cause some difficulty for the reader. | The range of structures is limited and rather repetitive. Although complex sentences are attempted, they tend to be faulty, and the greatest accuracy is achieved on simple sentences. Grammatical errors may be frequent and cause some difficulty for the reader. Punctuation may be faulty. |
| **Band Score 4** | The response is an attempt to address the task. (Academic) Few key features have been selected. (General Training) Not all bullet points are presented. (General Training) The purpose of the letter is not clearly explained and may beconfused.The tone may be inappropriate. The format may be inappropriate. Key features/bullet points which are presented may be irrelevant, repetitive, inaccurate or inappropriate. | Information and ideas are evident but not arranged coherently, and there is no clear progression within the response. Relationships between ideas can be unclear and/or inadequately marked. There is some use of basic cohesive devices, which may be inaccurate or repetitive. There is inaccurate use or a lack of substitution or referencing. | The resource is limited and inadequate for or unrelated to the task. Vocabulary is basic and may be used repetitively. There may be inappropriate use of lexical chunks (e.g. memorised phrases, formulaic language and/or language from the input material). Inappropriate word choice and/or errors in word formation and/or in spelling may impede meaning | A very limited range of structures is used. Subordinate clauses are rare and simple sentences predominate. Some structures are produced accurately but grammatical errors are frequent and may impede meaning. Punctuation is often faulty or inadequate. |
| **Band Score 3** | . The response does not address the requirements of the task (possibly because of misunderstanding of the data/diagram/situation). Key features/bullet points which are presented may be largely irrelevant. Limited information is presented, and this may be used repetitively. | There is no apparent logical organisation. Ideas are discernible but difficult to relate to each other. Minimal use of sequencers or cohesive devices. Those used do not necessarily indicate a logical relationship between ideas. There is difficulty in identifying referencing. | The resource is inadequate (which may be due to the response being significantly underlength). Possible over-dependence on input material or memorised language. Control of word choice and/or spelling is very limited, and errors predominate. These errors may severely impede meaning. | Sentence forms are attempted, but errors in grammar and punctuation predominate (except in memorised phrases or those taken from the input material). This prevents most meaning from coming through. Length may be insufficient to provide evidence |
| **Band Score 2** | The content barely relates to the task. | There is little relevant message, or the entire response may be off-topic. There is little evidence of control of organisationalfeatures. | The resource is extremely limited with few recognisable strings, apart from memorised phrases. There is no apparent control of word formation and/or spelling. | There is little or no evidence of sentence forms (except in memorised phrases). |

“

**Prompt 3**

Type this ChatGpt Prompt: I want you to evaluate this writer’s academic task 1 response to the given corresponding data set based on each individual set of criteria in the IELTS task 1 academic rubric. Also, I want you to approximate a band score. Here is the writer’s response which has a total of \_\_\_\_ mistakes in either mechanics, grammar, or lexical ability and \_\_\_\_ words “ .“ I want you to give a complete analysis according to the IELTS Task 1 rubric you just read and approximate a band score according to the rubric criteria. A response is required to have a minimum of 150 words, so if it is less than that amount, lower the score by 1 rubric point from where it would have been.

**IELTS Test Task 1 Academic Question**:



The table compares The number of visitors who visit Ashdown museum before and after of museum refurbishment and the charts illustrate the polls asked visitors Satisfaction before and after this change

Overall, what stand out is that the number of people visiting have risen considerably. Moreover,high proportion of visitors were not satisfied before reconstruction, while vast majority of individuals are content with New changes.

As can be seen, exactly half of visitors were very dissatisfied and dissatisfied before redecorating, With 10 and 40 percent respectively, whereas 3/4 of visitors have become satisfied and very satisfied after refurbishment, with 40 and 35% respectively.

before refurbishment just 15% of visitors were very satisfied which it is 35% after it which is considerable change. Another remarkable change refers to dissatisfied group which decrease from 40% before to 15% after refurbishment. Also, very dissatisfied declined by half. In addition the same person of people don't answer or participate to the surveys in both time period come on with just 5%. this is evident from table that this change was satisfactory attracting more visitors to museum in comparison with the past

1. Evaluation 2:

ChatGpt Prompt: I want more feedback on this IELTS Task 1 essay response you just evaluated. I want you to give examiner-level essay feedback on 18 different criteria. I want you to give analytical feedback on criteria 1-7 first and stop. Then ask me if I want you to continue. Here are the criteria: “

Criteria 1: Analyze the overall writing level of the essay according to a CEFR scale, meaning is the writing level in the essay classified as A1, A2, B1, B2, C1, or C2**? Please explain your response.**

Criteria 2: Evaluate the essay’s usage of synonyms and paraphrasing. In other words, is the writer being repetitive in their word choice? **List any words or phrases in the essay where the writer could have used synonyms or paraphrasing.**

Criteria 3: Are any areas of the essay insufficiently developed, lacking clarity, or are supporting arguments and evidence lacking relevance?

Criteria 4: List any mistakes in (1) sentence structure, (2) word form, (3) word choice, (4) mechanics, (5) grammar, and (6) lexical.

Criteria 5: Please comment on the range of vocabulary and grammar used in the essay and detail in which ways it could be improved in those areas.

Criteria 6: Of the total sentences, count how many are simple, compound, complex, or compound-complex sentences. Are there any mistakes in forming these sentences?

Criteria 7: Does the essay use a range of mechanics, vocabulary, and grammar at the basic, intermediate, or advanced level of English? Please explain why?

Criteria 8: To what degree does this response stay on topic with the given prompt/question? If it does not stay on topic, please explain how it does not stay on topic.

Criteria 9: Please comment on whether this response is coherent and flows smoothly from sentence to sentence. If the response is not coherent in any way, please explain how.

Criteria 10: Does the response use any idiomatic expressions? If so, please detail them. Also, are the idiomatic expressions correctly or incorrectly used?

Criteria 11: Does the response correctly or incorrectly use any English collocations? If so, please detail them and point out any incorrect collocations.

Criteria 12: Please determine the number of sentences in the response and then detail how many of those sentences are mistake-free of any mechanical, grammar, or lexical errors. Please list the mistakes found.

Criteria 13: Does the writer’s response fail to address the whole response in any way? If so, detail how?

Criteria 14: Does the essay writer use what an English professor would consider weak verbs? Please give a horizontal list of the weak verbs.

Criteria 15: Are there any areas of the written essay that would not be considered fully and well-developed or misunderstood? Are there any ways in which the essay ideas are not extended and supported?

Criteria 16: Is there any way in which the essay does not represent proper referencing, proper paragraphing, logical sequencing, or without clear progression throughout the essay? If so, please detail how.

Criteria 17: Does the essay lack any clarity of message throughout the essay?

Criteria 18: How many sentences in the essay are mistake-free in mechanics, grammar, and lexical ability? Are a majority of the sentences mistake-free?

Criteria 19: Do any areas of the essay require the reader to have to read more carefully to discern the writer’s position or message? Can the writer’s position be clearly identified in the essay?

Criteria 20: What is the writer’s level in using transitions, sequencers, and cohesive devices? Could any of those transitions, sequencers, or cohesive devices be upgraded to an advanced level of English? ”

Criteria 21: Please list every word that is used more than once. Also, count how many times that word is repeated.

Criteria 22: Please comment on the paragraph development. Are there any issues related to paragraph development, progression, sequencing, or clarity? If so, please detail any weaknesses.

Prompt 4:

Please list all the ways this essay response could be upgraded from your given band score of \_\_\_\_\_ to a band score of \_\_\_\_\_\_\_.