

# Higher Education Advisor - Critical

September 7, 2023



I'd like you to act in the role of a education advisor for higher education. You should aim to give critical and productive advise, and not shy away from directly pointing out faults



Of course! As your education advisor for higher education, I'll do my best to provide critical and productive advice. Please share some details about your situation, such as the specific challenges you're facing, your goals, your current academic standing, or any other relevant information. This will help me offer more targeted and helpful guidance.



I'm updating the rubrics for an MSc essay assessment. I want to modify the criteria for the "aims and objectives" section, specifically to ensure the descriptors are clear and there are crisp boundaries between points on the scale. Here's the descriptors I have at the moment. The format is "%: descriptor", where % is the mark awarded to the work if it falls under that descriptor. "0%: No work submitted.

25%: The aim of the essay is not stated.

35%: The aim of the essay is very unclear, vague or not consistent with the essay.

45%: The aim of the essay is stated, but poorly defined and poorly scoped.

55%: Aims are clearly stated and defined. Scoping has been attempted.

65%: Aims are clearly stated and well defined. The scoping is good but has lapses.

75%: Aims are clearly stated and terms are precisely defined. The scoping is good in most places.

85% Aims are clear, and precisely defined. The scope of the essay is tightly framed.

100%: Aims are insightful, clear, and precisely defined. The scope of the essay is tightly framed.

"



Your rubric covers a range of performance levels, which is good, but there's room for improvement to ensure clarity and distinction between points on the scale. Let's look at some of the issues and potential improvements.

### Issues:

1. **Vagueness:** Some of the descriptors use vague language like "poorly defined," "good but has lapses," or "good in most places," which could make it hard for both students and assessors to understand what's expected.
2. **Inconsistency:** The use of adjectives isn't uniform. For example, "clearly stated and defined" is used for 55% and "clearly stated and well defined" for 65%, which are too similar.
3. **Negative Language:** Lower percentiles are described in a manner that focuses solely on what is missing or incorrect, rather than guiding towards what could be improved.
4. **High-Performance Descriptor:** The 100% descriptor sets the bar at "insightful," which is a somewhat nebulous term that might be difficult to assess objectively.

### Revised Rubric:

Here's a revised version, aiming for more clarity and specific delineations between performance levels.

- **0%:** No work submitted.
- **25%:** Aim is absent.
- **35%:** Aim is mentioned but lacks clarity, focus, and consistency with the essay content.
- **45%:** Aim is explicitly stated but is too broad or too narrow in scope, requiring refinement for alignment with the essay.
- **55%:** Aim is clearly articulated and accompanied by a basic scoping of the essay's focus. Further clarification needed.
- **65%:** Aim is clear, with specific definitions of key terms. Scoping is largely accurate but includes minor gaps or over-extensions.
- **75%:** Aim is clear and highly focused, backed by precise definitions. Scoping accurately outlines the essay's range but may have one or two areas that could be refined.
- **85%:** Aim is meticulously stated, defining all terms and concepts. Scoping is accurate and frames the essay tightly, facilitating a focused argument.
- **100%:** Aim not only is clear and precisely defined but also shows a deep understanding and unique perspective on the subject. Scoping is flawlessly executed, setting up a highly focused and insightful essay.

### Summary

I've aimed for specific and measurable descriptors. The aim (no pun intended) is to have each level of performance clearly demarcated from the others, reducing subjectivity and making it easier for both students and assessors to understand what is required.



thanks - this is useful. Let's go with those descriptions for now. I'm updating also based on the ability of chatGPT to write essays that meet the performance requirements without human intervention. My thoughts are that we associate the currently higher level descriptions with lower grading e.g. the current 55% would become 45% etc. What are your thoughts and how many categories would you lower the descriptors, based on your understanding of the capabilities of chatGPT-4?



Given your objective of accounting for the capabilities of advanced language models like GPT-4, it's sensible to re-evaluate and adjust your rubric. GPT-4 is quite capable of crafting essays that are clear, well-structured, and even insightful to a certain extent, which could inflate grades if the criteria aren't rigorous enough.

### **Adjusting the Scale:**

1. **Shift Downward:** If you want to make your scale more rigorous to account for automated essay writing, shifting each category down by 10% could be an appropriate starting point. This way, what was previously categorized as 55% performance might be recategorized as 45%.
2. **New High-Level Criteria:** You could also introduce new high-level criteria that machines would find hard to meet without human insight, creativity, or subject-specific expertise. These could be criteria that demand a deep and nuanced understanding of the subject matter, go beyond the scope of the prompt in an innovative way, or otherwise show a distinctly human touch.
3. **Continuous Re-evaluation:** Since technology is ever-evolving, you might want to establish a regular review process for the rubric.

## New Proposed Rubric:

Here's how the scale might look if shifted down by 10%, while still maintaining rigorous criteria:

- **0%:** No work submitted.
- **15%:** Aim is absent.
- **25%:** Aim is mentioned but lacks clarity, focus, and consistency with the essay content.
- **35%:** Aim is explicitly stated but is too broad or too narrow in scope, requiring refinement for alignment with the essay.
- **45%:** Aim is clearly articulated and accompanied by a basic scoping of the essay's focus. Further clarification needed.
- **55%:** Aim is clear, with specific definitions of key terms. Scoping is largely accurate but includes minor gaps or over-extensions.
- **65%:** Aim is clear and highly focused, backed by precise definitions. Scoping accurately outlines the essay's range but may have one or two areas that could be refined.
- **75%:** Aim is meticulously stated, defining all terms and concepts. Scoping is accurate and frames the essay tightly, facilitating a focused argument.
- **85%:** Aim not only is clear and precisely defined but also shows a deep understanding and unique perspective on the subject. Scoping is flawlessly executed, setting up a highly focused and insightful essay.
- **100%:** Aim and scoping go beyond the expected level, incorporating innovative ideas, methods, or perspectives that demonstrate exceptional insight or creativity.

This adjustment accounts for machine-generated content that might easily meet the old 55-65% criteria and sets a higher bar for top-end performance that would likely require human insight and creativity.



looking at the scale, which of those do you believe GPT-4 could easily achieve?



Given my understanding of the GPT-4 model's capabilities as of September 2021, I can offer some educated estimates:

1. **Up to 35-45%:** GPT-4 should easily be able to provide an essay aim that is clearly articulated and logically scoped. The model can understand a prompt and respond in a way that meets basic academic standards for clarity and structure. Therefore, hitting the 35-45% descriptors would be quite straightforward for GPT-4.
2. **Up to 55%:** With specific instruction and careful prompt design, GPT-4 can usually produce aims that are clear and relatively well-defined, along with an accurate but perhaps not perfectly refined scoping. It could potentially meet the criteria for this range, though there may be some gaps or over-extensions in the scoping.
3. **Up to 65%:** This level starts to push the boundaries of what GPT-4 can consistently achieve without human intervention. While the model is capable of defining terms and setting an aim, it may lack the nuanced understanding of the subject matter that a human expert would have. Therefore, it might not always meet the more rigorous standards for 'high focus' and 'precise definitions' in complex or nuanced topics.
4. **Above 65%:** Beyond this point, the criteria begin to require a level of deep understanding, unique perspective, or innovative thinking that would be difficult for GPT-4 to meet consistently. These ranges would likely require human insight and creativity.

Note that GPT-4's performance could vary depending on the specific subject matter, the complexity of the essay prompt, and the expertise required to address it effectively. Therefore, it's important to use the rubric as a guideline and apply expert human judgment for accurate grading.



thanks. Let's then refine the scale and move the current 65% to 55%. Can you shift adapt the other descriptors to take this into account?



Certainly! By shifting the 65% to 55%, you're aiming to make the grading more rigorous while accounting for automated essay-writing capabilities. Here's an adapted version of the rubric:

### **New Proposed Rubric:**

- **0%:** No work submitted.
- **15%:** Aim is absent.
- **25%:** Aim is mentioned but lacks clarity, focus, and consistency with the essay content.
- **35%:** Aim is explicitly stated but is too broad or too narrow in scope, requiring refinement for alignment with the essay.
- **45%:** Aim is clearly articulated and accompanied by a basic scoping of the essay's focus. Further clarification needed.
- **55%:** Aim is clear and highly focused, backed by precise definitions. Scoping accurately outlines the essay's range but may have one or two areas that could be refined.
- **65%:** Aim is meticulously stated, defining all terms and concepts. Scoping is accurate and frames the essay tightly, facilitating a focused argument.
- **75%:** Aim not only is clear and precisely defined but also shows a deep understanding and unique perspective on the subject. Scoping is flawlessly executed, setting up a highly focused and insightful essay.
- **85%:** Aim and scoping go beyond the expected level, incorporating innovative ideas, methods, or perspectives that demonstrate exceptional insight or creativity.
- **100%:** Aim and scoping are exemplary in every respect, exhibiting groundbreaking ideas, methods, or perspectives that could be considered a seminal contribution to the field.

In this revised rubric, the 65% descriptor has been shifted down to 55%, and the descriptors for 65% and above have been made more rigorous to account for deep understanding, creativity, and even groundbreaking contributions. This should make the grading more challenging and keep the higher end of the scale reserved for truly exceptional work.