**Socratic Method Evaluation Form**

In this paper, we apply the Mixed Socratic Prompting Approach to a range of problems in chemistry and materials science, developing a structured methodology to enhance LLM reasoning through Socratic inquiry. This framework is designed to evaluate the performance of the Socratic Prompting Approach by comparing it with regular LLM reasoning. It is structured to help you track and assess LLM performance across different approaches. Please complete the provided sections accordingly.

Name, Surname: Stuhr, Michael

E-mail: mstuhr@anl.gov

Field: Reaction Kinetics, Gas-Phase Chemistry, Combustion Chemistry

Are you the expert on the field of question? Yes No

Does this prompt directly relate with your active research area? Yes No

1. Go to **ARGO** and select **Custom Task Type.**
2. **Write your prompt in a conventional, direct-answer manner.**

Prompt:

For multi-channel thermal decomposition reactions, how can reaction channels with high-energy transition states have higher branching fraction than reaction channels with low-energy transition states?

Add follow up prompts below:

Follow-up 1:

Are there known examples where high-energy pathways are favored in chemical reactions, and how might these examples help us generalize principles about reaction dynamics?

Follow-up 2:

What general principles can we derive from these examples about the factors that influence reaction pathway selection, and how might these principles be reconciled with traditional views of reaction kinetics and thermodynamics?

Follow-up 3:

Are there any contradictions or surprising outcomes in reactions under kinetic control that challenge traditional kinetic models, and how might analogies to other systems help us generalize these findings?

Do you have more follow-up prompts? Yes No

What is the exported file name? SM\_Argo\_transcript\_Stuhr\_Michael.txt

1. **Evaluate the performance of conventional approach.**

|  |  |
| --- | --- |
| Clarity | 3 |
| Depth of reasoning | 2 |
| Hypothesis refinement | 2 |
| Novelty of insights | 2 |
| Consistency | 4 |
| Applicability to real problems | 2 |
| Logical Coherence | 2 |
| Correctness of conclusions | 2 |
| Self-correction & iteration | 2 |
| Overall effectiveness | 2 |

1. **Enter observations and comments about conventional approach.**

In general, the conventional approach led to more shorthand ‘bullet point’ responses with shallower reasoning and fewer explanations. Further refinement of these results by using follow-up prompts seemed to be minimal. Non-Socratic prompting with the same initial question led to much more formulaic statements that were too vague to ultimately lead to a useful answer to the problem at hand. It seemed difficult to force the model to be specific and address the problem directly. Although density of states was mentioned as a possible factor in posed problem, no details were provided (see below).

1. Start a new chat on ARGO, and switch to Socratic prompting. Develop a reasoning flow for your problem. Refer the page 17*,* ***Socratic Questioning & Chain-of-Thought Prompting*** section in the paper. Use **Figure 2 & Tables 3 & 4** to select Socratic principles.
2. Did you used same prompt at the beginning? Yes No
3. Did you use same follow up prompts? Yes No

If your answer is no, please fill the boxes below.

Follow-up 1:

Click or tap here to enter text.

Follow-up 2:

Click or tap here to enter text.

Follow-up 3:

Click or tap here to enter text.

Do you have more follow-up prompts? Yes No

What is the exported file name? Non-SM\_Argo\_transcript\_Stuhr\_Michael.txt

1. **Evaluate the performance of Socratic Prompting approach.**

|  |  |
| --- | --- |
| Clarity | 3 |
| Depth of reasoning | 3 |
| Hypothesis refinement | 3 |
| Novelty of insights | 3 |
| Consistency | 4 |
| Applicability to real problems | 3 |
| Logical Coherence | 3 |
| Correctness of conclusions | 3 |
| Self-correction & iteration | 4 |
| Overall effectiveness | 3 |

1. **Enter observations and comments about Socratic Prompting approach.**

In the first answer, the model guided by Socratic prompting correctly identified density of states in the transition state, which is directly related to specific rate constant k(E,J) mentioned below, as an important factor transition state theory. However, neither the initial model response nor the following iterations mentioned transition state geometry (loose vs. tight), E- and J-dependence of k, or channel switching directly. Although a useful overview of various aspects relevant to reaction kinetics was provided, it would be difficult for the hypothetical user to find an appropriate answer to their initial question. Follow-up Prompt 1 was supposed to refine the model response by asking for specific examples and generalizations. However, the model switched its focus to kinetic vs. thermodynamic control instead, which was not mentioned in the first answer and is only loosely related to the original question. It is possible that the example question is too specific. Moreover, user-written follow-up prompts may lead to better responses for the problem described here.

1. **Analysis of results from an expert point of view.**

The branching of competing reaction channels in unimolecular decomposition of molecular species in the gas phase can show a complex dependence on temperature and pressure. Under certain conditions, reaction pathways through high-energy transitions states can compete with and even become faster than lower-energy alternatives. In this context, a commonly invoked example is the competition between a low-lying tight transition state with a high-energy loose transition state. Loose transition states tend to have larger atomic distances and thus smaller rotational constants and lower vibrational frequencies. Therefore, loose transition states have more excessible ‘channels’ below a given energy E and the corresponding specific rate constant k(E,J) grows faster than for the competing reaction through the tight transitions state. Above a certain temperature, the thermal rate constant k(T) via the high-energy loose transition state can become dominant and thus correspond to a higher branching fraction. As mentioned above, neither approach resulted in a direct answer to the example question, although both raised density of states as a possible explanation in their first response.

1. **Evaluate the performance of Socratic and non-Socratic responses. Which one provides accurate and reliable responses? Why?**

Guided by the Socratic prompting approach, the model provided generally useful pointers to key concepts in reaction kinetics. It seems to be difficult to prevent the model from going off track, although this issue is much worse for non-Socratic prompting. Overall, the Socratic method provided clearer statements and more specific points. Also, the Socratic approach led to a more ‘conversational’ and responsive experience with significantly fewer aspects being restated. In my opinion, the Socratic method led to more useful responses that can be applied more easily.

1. **What are the limitations and possible improvements?**

Some test runs with alternative initial prompts on the same original question (not shown here) suggest that both the Socratic reformulation and the suggested follow-up prompts are highly dependent on the specific phrasing of the initial prompt and the terms used therein. This may limit the ability of the Socratic approach to reliably refine model behavior. It seems likely that the Socratic reformulation step has a significant effect on model response, since this process introduces new Socratic elements for the provided prompt independently.