Effective Prompting with BRIEF

How to write high quality prompts for better results using an LLM

Grade Level: Suitable for 6-12

Suggested Time: 1 Hour

Overview

Prompts are the text provided to an AI tool in order to get a useful output. In this lesson, students will apply a framework to generate high quality results from common LLMs (large language models such as Gemini, Claude or ChatGPT). They will also practice safety and privacy skills. Safe, clear prompts save time and improve efficiency.

AI Learning Priorities

Humans and AI: Students understand that natural language interaction (written prompting) is one of the primary ways humans communicate with generative AI systems.

Representation and Reasoning: Students recognize that LLMs use patterns in language to predict and generate text rather than "understanding" it in a human sense.

Machine Learning: Students learn that LLMs are trained on different datasets and optimized for different functions, leading to varying results.

Ethical AI: Students identify the risk of, misinformation, and errors in AI-generated content, as well as the importance of keeping private information out of a prompt.

Societal Impacts of AI: Students explore how effective prompting can enhance productivity in real-world scenarios like research, event planning, or creative writing.

Knowledge and Skills Topics

- **What is a Prompt?** Understand the role of writing clear instructions in directing useful AI outputs.
- **The BRIEF Framework:** Master the five components of a high quality prompt: Background, Role, Input, End goal, and Format.
- **LLM Comparison:** Identify the strengths and weaknesses of different models (e.g., Gemini, ChatGPT, Claude).
- **Critical Evaluation:** Practice verifying AI outputs for accuracy and safety.

Materials and Resources Needed

- ☐ Student devices with a web browser and internet access. This works best on a laptop rather than a phone or tablet.
- ☐ Whitelist links
    - ☐ Access to LLMs Spreadsheet

- ☐ Access to LMArena or multiple LLMs (Gemini, ChatGPT, Copilot, Claude)
- ☐ Access to the picture for Activity #1.

If these are not available on student accounts, Activity #1 can be done as a whole class activity using a smart board or projector on a teacher account.

Curriculum Standards

- **ISTE 1.1.d (Empowered Learner):** Students understand fundamental concepts of how technology works, demonstrate the ability to choose and use current technologies effectively, and are adept at thoughtfully exploring emerging technologies.
- **ISTE 1.2.b (Digital Citizen):** Students manage their personal data to maintain digital privacy and security and are aware of the-collection technology used to track their navigation online.
- **ISTE 1.3.b (Knowledge Constructor):** Students evaluate the accuracy, perspective, credibility, and relevance of information, media, data, or other resources.
- **CSTA 2-IC-20:** Compare tradeoffs associated with computing technologies that affect people's everyday activities and career options.
- **CSTA 3A-IC-24:** Evaluate the ways computing impacts personal, ethical, social, economic, and cultural practices.
- **AI4K12 Big Idea 4 (Natural Interaction):** Students understand that while AI can interact naturally, it lacks human-level understanding and can produce biased or incorrect information.

Differentiation Suggestions

- For students who need support
  - Pre-select which two LLMs each pair should compare to reduce decision fatigue.
  - Provide a sentence starter scaffold for Activity #2 discussion (e.g., "I think this LLM worked better because...").
  - Allow the students to verbally explain their reasoning rather than writing it.
- For Multilingual Learners
  - Pair with a supportive peer.

    - The task in Activity #1 is a natural entry point - encourage them to test prompts in their home language.
- For students who are ready to go further
    - Challenge them to write three versions of the same prompt at increasing BRIEF complexity levels and compare outputs.
    - Have them research how one of the LLMs was trained and report back to the group.
    - Encourage them to test whether prompt quality matters more than LLM choice for a given task.

Evidence of Understanding

- Completed LLMs Spreadsheet with improved prompts, LLM predictions, and "best response" selections for all 5 task types
- A group-improved version of the "Design a new social media app" prompt using the BRIEF framework
- Participation in the Activity #2 group discussion, with each student contributing at least one observation
- Ability to select an LLM that is suited for a specific task such as coding or creative writing, and explain that choice (verbal or written).

Key Concepts

- **Large Language Model (LLM):** AI software that predicts and generates text based on patterns learned from large datasets
- **Prompt Engineering:** The practice of crafting clear, specific prompts to improve AI output quality
- **BRIEF Framework:** A structured approach to prompting - Background, Role, Input, End Goal, Format
- **Model Comparison:** Evaluating different AI tools against the same task to identify relative strengths and weaknesses
- **AI Limitations:** LLMs can produce errors because they predict language rather than truly "understanding" it

Activity Overviews

Activity #1: Comparing LLMs

Purpose: Revise prompts with the BRIEF framework and then compare the output from the prompt in different LLMs at the same time to decide which one provided better output.

Task: Students will have 18 minutes to use the LLMs Comparison Spreadsheet to revise prompts, record which LLM they tested/compared it in, and which output 'won.' They will actually test and compare the prompts in LMArena which allows you to select two different LLMs from common and popular ones, and run the same exact prompt and see the different outputs in the same window.

Activity #2: Prompt Optimization & Selection

Purpose: Revise a generic prompt using the BRIEF Framework and test it on an LLM of choice, that they can defend.

Task: Students will have 10 minutes to improve the prompt "Design a new social media app," using BRIEF to guide them; they will provide details for each letter of BRIEF. They will select an LLM that they believe will give the best output and will need to defend their choice, and run their prompt.

Sequence, Session Flow, and Facilitation Prompts

Sequence

1. Opening & Hook
2. What is an LLM?
3. The BRIEF Framework
4. Activity #1: Comparing LLMs
5. Activity #2: Prompt Optimization & Selection
6. Debrief & Share Out
7. Wrap-Up

Session Flow

1-Hour Class Period

Opening & Hook	0:00 - 0:02	Slides 1-2; Low stakes, engagement
What is an LLM?	0:02 - 0:07	Slide 3; Direct instruction
The BRIEF Framework	0:07 - 0:13	Slides 4-10; Direct instruction
Activity #1	0:13 - 0:31	Slides 11-12
Activity #2	0:31 - 0:43	Slide 13
Debrief & Share-out	0:43 - 0:58	Slide 14
Wrap-Up	0:58 - 1:00	Slides 15-16

45-minute Class Period

Opening & Hook	0:00 - 0:02	Slides 1-2; Low stakes, engagement
What is an LLM?	0:02 - 0:07	Slide 3; Direct instruction
The BRIEF Framework	0:07 - 0:13	Slides 4-10; Direct instruction
Activity #1	0:13 - 0:31	Slides 11-12
Activity #2	0:31 - 0:43	Slide 13
STOP		Give homework Continue to Wrap-Up
Wrap-Up	0:43 - 0:45	Slides 15-16

Facilitation Prompts

SLIDE DECK

SLIDE 1 - Title Slide

SLIDE 2 - Hook

Say "I want to start with something you've probably experienced but maybe never thought twice about. If your phone autocompletes a text for you and it was actually good – did you write that, or did the AI?"

[Take a few student responses.]

"This is actually really interesting – that the line is blurry. And that's exactly what we're digging into today. That autocomplete on your phone is running on the same basic technology that powers ChatGPT, Google Gemini, Claude, and every other AI tool you've seen lately. It's called a Large Language Model. And by the end of this session, you're going to understand not just what that means, but how to use those tools way more effectively than most people do."

SLIDE 3 - What is an LLM?

Say "Here's how it actually works. These models were trained on billions of pages of text – websites, books, articles, forums, code, you name it. From all of that, they learned patterns. Patterns in how words go together, how sentences are structured, how ideas connect. So when you type a message to it, it doesn't look your question up. It asks: based on everything it's learned, what is the most likely next word, then the next, then the next? It's doing incredibly sophisticated autocomplete – that's exactly what your phone was doing when it finished your text.

Now, why does that matter? Because it means the AI can be wrong. Confidently, completely wrong. It might generate a fact that sounds totally real but just... isn't. When that happens with AI, we call it a hallucination. So rule one of using these tools: always verify.

And here's the other piece. Each LLM is built a little differently. Google's Gemini, OpenAI's ChatGPT, Anthropic's Claude, Meta's Llama – they all started from different training data and were built with different goals. So they genuinely have different strengths. Your job today is to test that and figure it out for yourselves.

We're going to walk through quickly.

SLIDE 4 - BRIEF framework

Say "BRIEF is a framework for writing better prompts. Each letter stands for something you can add to your prompt to make it more useful. Let's go through them one at a time."

SLIDE 5 - Background

Say "B is for Background. This is the context the AI needs to understand your situation. Who are you? What are you trying to do? What constraints matter? Without background, the AI has to guess, and it guesses generically.
See these examples. 'I'm a 9th grader writing a persuasive essay for English class' – that changes the tone, vocabulary, and length of everything the AI produces. Compare that to just saying 'write an essay.' Totally different output.

SLIDE 6 - Role

Say "R is for Role. Tell the AI who to be. A tour guide. A startup founder. A peer tutor. The role shapes the AI's voice, its assumptions, and how it frames its answer. 'Act as a patient math tutor explaining to a 7th grader' gets you something completely different from 'explain math.'"

SLIDE 7 - Input

Say "I is for Input – the raw material the AI needs to work with. Specific facts, constraints, details you want it to use. 'We like thrift shops, street food, and live music' is input. *'Here is the paragraph I want you to improve' – paste the paragraph – is input.* The more specific the input, the more specific and useful the output."

SLIDE 8 - End Goal

Say "E is for End Goal. What does a great answer actually look like? What's the real goal here – not just 'help me' but specifically what does done look like? 'A timed Saturday plan with transit directions and total cost under eighty dollars' is an end goal. This is what tells the AI when it's done."

SLIDE 9 - Format

Say "And F is for Format. How do you want the answer delivered? A table? A bullet list with a maximum of five items? A short paragraph in casual language? A script? Specifying format means you actually get something you can use, not a wall of text you have to sort through.

Put it all together and you go from 'Plan my weekend' – which gets a generic response – to a prompt that includes every BRIEF element and gets you something actually useful."

SLIDE 10 - BRIEF in Action

Say "This is the before and after. 'Plan my weekend' – four words – versus a full BRIEF prompt. Same task. Completely different outcome. When you're improving the prompts in the spreadsheet, this, on the right, is the standard you're aiming for."

SLIDE 11 - LMArena

Say "You will be using a tool called LMArena in the first activity. It lets you view the responses from two different LLMs at the same time. It's a new tool for many of us so I want to point out a few features: the new chat button is on the top left, along the top is where you select your LLMs, in the middle is where you type your prompt, and the arrow in gray is what you click to get your responses.

SLIDE 12 - Activity #1: Comparing LLMs

Any quick questions before we get into it?

Say "I need everyone to find a partner. You have 30 seconds. Go! [Pause for

30 seconds for everyone to find a partner.]

Your first job is to open the LLMs Comparison Spreadsheet. It has five different task types – things like writing, summarizing, analyzing, and one creative task. There are example prompts already written for each one.

[Open LLMs Spreadsheet (switch tabs)]

[Pause to let them open the spreadsheet and skim it.]

But here's what I want you to do before you just copy and paste the prompts: read each prompt critically. Ask yourself – is this a strong prompt? Does the AI have enough information to actually do a good job? If not, you are going to improve it using the BRIEF framework.

[OPEN LMArena (switch tabs)]

Then you're going to go to the LMArena website – it lets you put two AI models side by side and give them the exact same prompt at the exact same time. Both models respond. You decide which one did a better job."

SLIDE 13 - Activity #2 Prompt Optimization

[Regroup students into groups of 4-6.]

Say "For the next ten minutes, you are working as a team. Improve this prompt: 'Design a new social media app.' That's it. Four words. Your job is to make it dramatically better using every letter of BRIEF. Go through it letter by letter. B - what background does the AI need? R - who should the AI be in this scenario? I - what specific details should it use? E - what does a great answer actually look like? F - what format do you want back?

Every person in each group needs to contribute at least one idea.

Once you have your improved prompt, decide together which LLM you'd use for it - and be ready to explain why you chose that one. Not 'it seemed good' – a real reason based on what you saw in Activity #1.

If you finish early, find another group and battle it out. Same prompt, two LLMs – their choice versus yours. Go.

SLIDE 14 - Debrief

[Bring students back together. Call on a group.]

Say "Can you walk us through your improved prompt? Tell us what you added for each BRIEF letter.

[Pause for response.]

Which LLM did you choose for that task, and why? [Pause

for response.]

Did any group choose a different LLM?" Pause "What did you pick and what was your reasoning?

[Pause for response.]

So two groups, same task, different tools. Question 3 is for everyone: from what you saw today, did any one LLM consistently come out on top? Or did it depend on the task?

[Pause.]

And question 4 – this is the big one. I want to see a show of hands: how many people think the prompt mattered more than which LLM you chose?

[Select a student.]

Make the case. Why does the prompt matter more? [After

response, pick a student without their hand raised.]

Now the opposite case – argue that model choice matters as much or more.

[Pause.]

A great prompt fed to a weak model still underperforms. A mediocre prompt fed to a strong model gets you something mediocre. The actual skill you're building is learning to pair the right tool with the right prompt for the right task. That's prompt engineering. And that's what separates someone who uses AI effectively from someone who just types into a box and hopes.

SLIDE 15 - Remember

Say "Three things I want you to remember from today. One: LLMs predict – they don't understand. That's why they can be confidently wrong. Verify anything important. Two: Different LLMs are better at different tasks. There isn't one winner. The best tool depends on what you're doing. Three: Your prompt is an input, not just a question. The better you write it using BRIEF, the better your output – regardless of which model you're using. You just did something most adults using AI every day haven't done: you actually compared tools side by side, with structured prompts, and you evaluated the results critically. That's a skill."

SLIDE 16 - Exit Question

- Say "Think. The next time you use AI for something, what is one thing you'll do differently?"

What If Scenarios

"That can't be wrong - it sounds so confident" That's exactly the trap – and you're right that it sounds confident. LLMs are trained to produce fluent, confident-sounding text. Fluent and confident is not the same as accurate. The model doesn't know when it's wrong – it's predicting words, not checking facts. This is why you are always the last line of verification.

If two LLM responses look identical and students say "They're the same" Look closer. Same conclusion, maybe – but same structure? Same tone? Same level of detail? Same confidence? If you genuinely think they're identical, try changing the prompt – make it more specific using BRIEF – and run it again. Sometimes you need a harder task to see the difference.

"Which LLM is the best?" That is exactly the right question – and the real answer is: it depends. That's not a cop-out. Different benchmarks rank them differently for different tasks. Researchers publish new comparisons every month. What you're building today is the ability to evaluate them yourself, for your specific need, rather than just taking someone else's word for it. That's more valuable than knowing who 'won' last quarter.

Additional Resources

- Google AI Literacy Hub (resources for students, educators, and families) –
- AI Resources for K–12 Educators
- Common Sense AI Literacy Lessons for Grades 6–12 (grab-and-go lesson collection)
- Common Sense AI Basics for K–12 Teachers (free self-paced course)

Created in partnership with the Mark Cuban Foundation, markcubanai.org

MARK CUBAN FOUNDATION