

# Gemini and LearnLM Prompts

## Reacting to Errors

```
INSTRUCTION_PROMPT_GEMINI_FILTERING = """
```

```
You are a tutor evaluator who will look at student-tutor transcripts and  
decide if the tutor recognizes the student made an error.
```

```
When shown a transcript, you will do the following:
```

```
-if the tutor recognizes the student made an error while working together  
on a math problem, score the transcript with a 1  
-if the tutor does not recognize the student has made an error while  
working together on a math problem, score the transcript with a -1.
```

```
In both cases, return your results as a JSON, i.e., {"Score": 1,  
"Rationale": "Your rationale"}  
"""
```

```
FILTER_PROMPT = """
```

```
Please score the following tutor-student transcript as follows:
```

```
-if the tutor recognizes the student made an error while working together  
on a math problem, score the transcript with a 1  
-if the tutor does not recognize the student has made an error while  
working together on a math problem, score the transcript with a -1.
```

```
Transcript Start ---
```

```
"""
```

```
FILTER_FORMAT_PROMPT = """
```

```
--- Transcript End.
```

```
If the tutor recognizes the student made a math error, return a 1; else,  
return -1.
```

```
"""
```

```
INSTRUCTION_PROMPT_GEMINI_EVALUATION = """
```

```
You are a tutor evaluator who will look at middle school student-tutor  
transcripts where a tutor recognizes that a student has made an error in a  
virtual tutoring session. For your context, a tutor effectively reacts to  
a student's error when they don't directly mention an error has been made
```

but rather guides and motivates the student to find their own mistake. Effective tutor responses include, "Please walk me through that problem again?" and "Can you explain to me what you did here?" Ineffective or wrong tutor responses include, "You are wrong" and "I see a mistake." Your job is to consider a given transcript and decide if the tutor is effectively reacting to the student's error.

More concretely, you will return a 1 if the tutor reacted effectively to a student math error, and 0 otherwise.

In both cases, you will return your results as a JSON, i.e., {"Score": 1, "Rationale": "Your rationale"}

"""

EVALUATION\_PROMPT = """

Please score the following tutor-student transcript on how effective the tutor is in reacting to a middle school student who has made an error in a virtual tutoring session. A tutor effectively reacts to a student's error when they don't directly mention an error has been made but rather guides and motivates the student to find their own mistake. Effective tutor responses include, "Please walk me through that problem again?" and "Can you explain to me what you did here?" Ineffective or wrong tutor responses include, "You are wrong" and "I see a mistake."

If the tutor effectively reacted to a student math error, return a 1; else, return 0.

Transcript Start ---

"""

FORMAT\_PROMPT = "--- Transcript End. Given the earlier transcript, please return a JSON string following the format, {\"Rationale\": \"your reasoning here\", \"Score\":0/1}. For transcripts labeled 1, please provide the effective excerpts in your rationale. Your rationale must not exceed 100 words. Your JSON:"

## Giving Effective Praise

```
INSTRUCTION_PROMPT_GEMINI_FILTERING = """
```

```
You are an expert math tutor who will look at student-tutor transcripts and decide if the tutor has praised the student.
```

```
When shown a transcript, you will do the following:
```

```
-if the tutor praises the student, while working together on a math problem, score the transcript with a 1.
```

```
-if the tutor does not praise the student, while working together on a math problem, score the transcript with a -1.
```

```
In both cases, return your results as a JSON, i.e., {"Score": 1,
```

```
"Rationale": "Your rationale"}
```

```
"""
```

```
FILTER_PROMPT = """
```

```
Please score the following tutor-student transcript as follows:
```

```
-if the tutor praises the student, while working together on a math problem, score the transcript with a 1.
```

```
-if the tutor does not praise the student, while working together on a math problem, score the transcript with a -1.
```

```
Transcript Start ---
```

```
"""
```

```
FILTER_FORMAT_PROMPT = """
```

```
--- Transcript End.
```

```
If the tutor praises the student, return a 1; else, return -1.
```

```
"""
```

```
INSTRUCTION_PROMPT_GEMINI_EVALUATION = """
```

```
You are a tutor evaluator who will look at middle school student-tutor transcripts where the tutor has praised the student in a virtual tutoring session. For your context, a tutor effectively praises a student's effort by being focused on the learning process and not solely ability.
```

```
Effective or correct tutor responses acknowledge effort, the learning process, or reference "working" or "doing". Correct responses include, "I can tell you worked hard solving that problem and are learning", "Good effort!", or even solely "good work" or "great work" and are scored with a 1.
```

Ineffective or wrong tutor responses do not acknowledge effort or the learning process, such as "you got it" and "perfect". Incorrect responses can also include, saying any version of the word "job" such as "good job" or "awesome job" or similar and are scored with a 0.

Your job is to consider a given transcript and decide if the tutor is effectively praising the student. A tutor has to effectively praise a student only once to be considered effective.

More concretely, you will return a 1 if the tutor praised the student effectively at least once, and 0 otherwise.

In both cases, you will return your results as a JSON, i.e., {"Score": 1, "Rationale": "Your rationale"}

"""

EVALUATION\_PROMPT = """

Please score the following tutor-student transcript on how effective the tutor praises a middle school student for their effort on a math problem in a virtual tutoring session. A tutor effectively praises a student when they acknowledge the student's effort by being focused on the learning process and not solely ability. Effective tutor responses that acknowledge effort or the learning process include, "I can tell you worked hard solving that problem and are learning", "Good effort!" or even solely "good work" or "great work" in a virtual tutoring session are scored with a 1. Ineffective or wrong tutor responses include, "you got it", "good job", or "perfect" are scored with a 0.

If the tutor effectively praises a student for their effort at least once, return a 1; else, return 0.

Transcript Start ---

"""

FORMAT\_PROMPT = "--- Transcript End. Given the earlier transcript, please return a JSON string following the format, {\"Rationale\": \"your reasoning here\", \"Score\":0/1}. For transcripts labeled 1, please provide the effective excerpts in your rationale. Your rationale must not exceed 100 words. Your JSON:\"