

## (a): Meta-Learning Capabilities



### Qwen3FT(Zero – Shot):

A little girl in an Elmo costume holding a guitar.

### Qwen3FT(Few – Shot):

A young blond child in a red Elmo suit is playing on a red Elmo guitar.

### GroundTruth:

A child in a Elmo suit is playing the guitar.

*Success : Few – shotpromptsrecovereddetails('blond', 'redguitar').*

## (b): Contextual Interference



### Llama3.2FT(Zero – Shot):

Two young men with brown hair... sit at a table surrounded by cords...

### Llama3.2FT(Few – Shot):

Two men... one with a cigarette... one with a pencil in his mouth... playing a game.

### GroundTruth:

Two men looking at little white laptops and holding a guitar.

$\text{\it{Failure: Few-shot prompts caused hallucinations ('cigarette', 'pencil').}}$

## (c): Fine-Grained Limitation



### Llama3.2FT(Zero – Shot):

A man in a white shirt walks toward the camera as people walk away...

### Qwen3FT(Few – Shot):

An elderly man in a white shirt is walking in a public area while other people walk behind him.

### GroundTruth:

An old man walking with a folder in his hand.

$\text{\it{Failure: Both models missed the small object ('folder') mentioned in GT.}}$