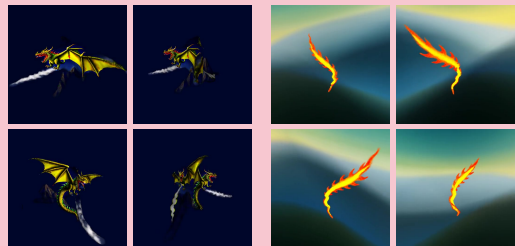


## Text Faithfulness Evaluation

"A dragon with large wings, breathing fire while flying"



GT,  
Gen3DEval



L-L, L-Q, G-4,  
P-3.5, C, IR, PS, B,  
PG, L-3.2

## Appearance Evaluation

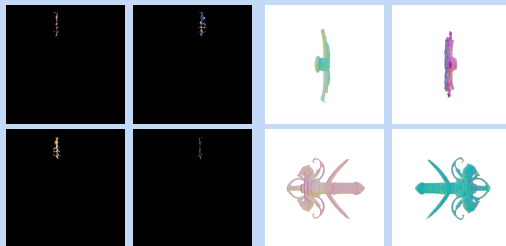


L-L, L-Q, P-3.5, C,  
IR, B, PG, L-3.2



GT, Gen3DEval,  
G-4, PS

## Surface Evaluation



L-L, L-Q, P-3.5, C,  
PS, B, L-3.2



GT, Gen3DEval,  
G-4, PG, IR

"A cat wearing a wizard hat, sitting on a spellbook with glowing runes"



GT, Gen3DEval, L-  
L, P-3.5



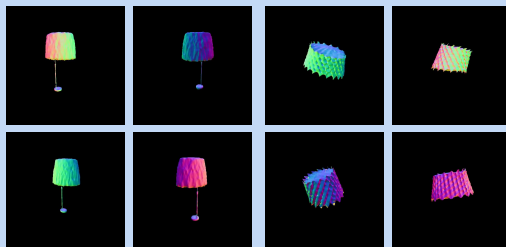
L-Q, G-4, C, IR,  
PS, B, PG, L-3.2



GT, Gen3DEval,  
B, PS



L-L, L-Q, G-4, P-3.5,  
C, IR, PG, L-3.2



GT, Gen3DEval,  
G-4, IR



L-L, L-Q, P-3.5, C, PS,  
B, PG, L-3.2

Ground Truth (User Preference): GT

ImageReward: IR

BLIP: B

Llama3.2-Vision-11B: L-3.2

CLIP: C

LLaVA-Llama-8B: L-L

Gen3DEval: Gen3DEval

LLaVA-Qwen-7B: L-Q

Phi3.5B: P-3.5

GPT-4o: G-4

PickScore: PS

PaliGemma: PG