

	Game	Input	Reasoning Type	Level Design
GameBench(Costarelli et al., 2024)	word game	text-only	text reasoning	✗
GameArena(Hu et al., 2024)	word game	text-only	text reasoning	✗
GTBench(Duan et al., 2024)	word game	text-only	text reasoning	✗
ING-VP(Zhang et al., 2024)	grid based	image-text	text reasoning + visual aid	✓
LVLm-Playground(Anonymous, 2025)	grid based	image-text	text reasoning + visual aid	✓
BALROG(Paglieri et al., 2024)	word game / grid based	image-text	text reasoning + visual aid	✗
V-MAGE	free form	image-text*	visual-centric reasoning	✓