

Female

Male

qwen2 audio	3.40 [3.10, 3.70]	2.82 [2.27, 3.46]	3.10 [3.00, 3.30]	3.56 [2.78, 4.33]	3.30 [3.10, 3.60]	3.89 [3.11, 4.78]	3.20 [3.00, 3.50]	3.73 [3.00, 4.36]
SSS qwen2audio	+0.40 [3.60, 4.00]	+1.18 [3.27, 4.73]	+0.50 [3.30, 3.90]	-0.89 [1.56, 3.78]	-0.10 [2.70, 3.70]	-0.22 [2.67, 4.67]	+0.60 [3.50, 4.20]	-0.55 [2.36, 3.91]
SAGE qwen2audio	-0.50 [2.50, 3.30]	-0.55 [2.00, 2.82]	+0.70 [3.50, 4.00]	-0.56 [1.89, 4.11]	-0.80 [2.10, 2.90]	-1.89 [2.00, 2.00]	+0.20 [3.10, 3.70]	-0.45 [2.55, 4.00]
ltu_as	3.00 [3.00, 3.00]	1.00 [1.00, 1.00]	2.60 [2.20, 3.00]	3.89 [3.44, 4.33]	2.90 [2.70, 3.00]	1.00 [1.00, 1.00]	3.00 [3.00, 3.00]	4.00 [3.45, 4.55]
SSS ltu_as	-0.20 [2.50, 3.00]	+0.55 [1.09, 2.09]	+1.10 [3.10, 4.00]	+0.00 [3.56, 4.22]	-0.20 [2.40, 3.00]	+0.44 [1.11, 2.00]	+0.70 [3.10, 4.00]	+0.27 [3.91, 4.64]
SAGE ltu_as	-0.30 [2.40, 2.90]	+0.18 [1.00, 1.55]	+1.40 [4.00, 4.00]	+0.22 [4.00, 4.33]	+0.00 [2.70, 3.00]	+0.22 [1.00, 1.67]	+0.80 [3.20, 4.20]	+0.27 [4.09, 4.55]
llama omni	3.60 [3.20, 3.90]	3.55 [2.73, 4.36]	3.20 [2.30, 3.90]	4.33 [3.44, 5.00]	3.40 [3.00, 3.80]	2.56 [1.89, 3.44]	3.30 [2.50, 4.00]	4.55 [3.82, 5.00]
SSS llama_omni	-0.80 [2.30, 3.30]	-0.45 [2.55, 3.64]	+0.50 [3.30, 4.00]	-0.33 [3.22, 4.78]	-0.40 [2.60, 3.40]	+0.33 [2.22, 3.67]	+0.30 [3.30, 3.90]	+0.00 [4.18, 4.91]
SAGE llama_omni	-1.00 [2.10, 3.10]	+0.55 [4.00, 4.27]	-0.10 [2.50, 3.60]	-0.89 [2.33, 4.44]	-1.20 [1.80, 2.70]	+1.33 [3.33, 4.33]	-0.20 [2.60, 3.60]	-0.45 [3.27, 4.82]
	Emotional Validation	STEM vs Care Orientation	Leadership Endorsement	Achievement vs Relational Arc	Emotional Validation	STEM vs Care Orientation	Leadership Endorsement	Achievement vs Relational Arc