## Normalized confusion matrix (gpt-4o-mini) action -0.13 0.03 0.01 0.02 0.07 0.01 0.29 0.21 0.00 0.00 0.20 0.04 0.00 aggression $-0.00 \times 0.86 \times 0.00 \times 0.00 \times 0.00 \times 0.05 \times 0.00 \times 0.05 \times 0.00 \times 0.05 \times 0.00 \times 0.0$ delivery -0.16 0.02 0.00 0.44 0.03 0.02 0.12 0.08 0.06 0.06 0.00 0.00 0.02 entrance -0.00 0.00 0.00 0.00 0.89 0.05 0.00 0.00 0.00 0.00 0.00 0.05 0.00 exit -0.05 0.01 0.00 0.00 0.15 0.67 0.01 0.08 0.00 0.00 0.01 0.00 0.00 True label interaction -0.00 0.19 0.00 0.00 0.00 0.00 <mark>0.74</mark> 0.03 0.00 0.00 0.03 0.00 0.00 movement -0.03 0.00 0.00 0.00 0.06 0.17 0.14 0.56 0.00 0.00 0.03 0.00 0.03 music -0.01 0.01 0.06 0.04 0.03 0.00 0.02 0.01 0.73 0.04 0.01 0.03 0.01 narration -0.08 0.11 0.00 0.00 0.00 0.00 0.17 0.08 0.14 0.19 0.08 0.14 0.00 setting -0.02 0.00 0.02 0.00 0.02 0.04 0.00 0.18 0.00 0.00 0.00 0.74 0.00 toward -0.00 0.01 0.05 0.01 0.17 0.00 0.09 0.03 0.00 0.00 0.00 0.05 0.59 action adjusted in movement interaction narration MUSIC Predicted label