



Figure R1: Two cases selected from the CSR-generated datasets

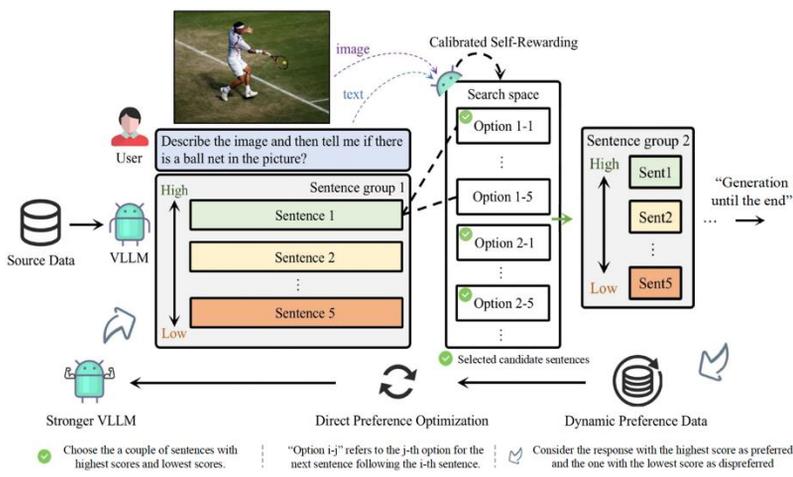


Figure R2: Polished Figure 2 in the main paper.

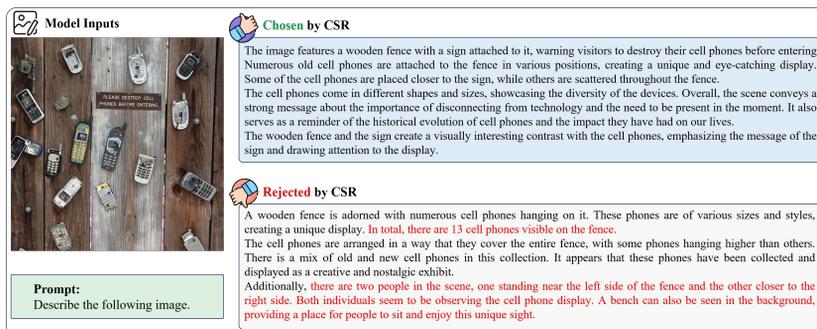


Figure R3: A case including both self-generated preferred and dispreferred responses

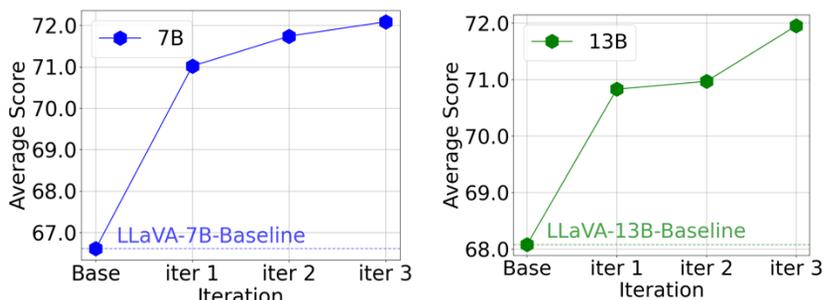


Figure R6: Updated Figure 3 in the main paper

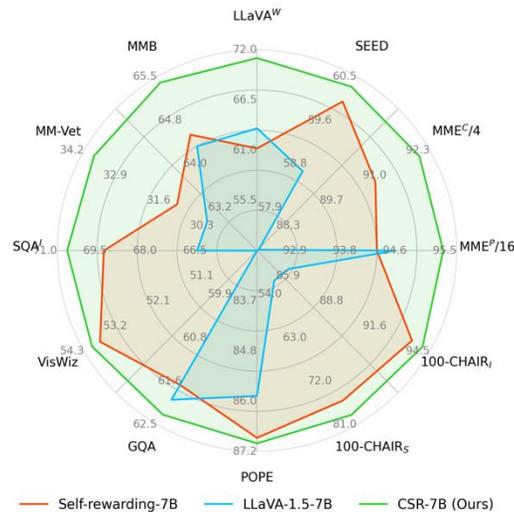


Figure R4: Polished Figure 1 in the main paper

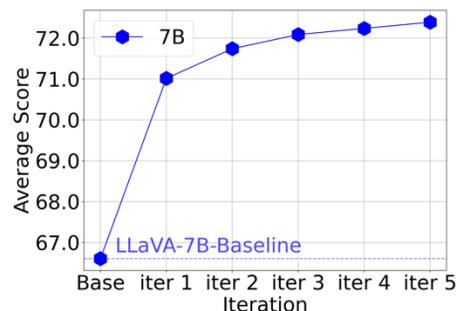


Figure R5: The average score of CSR across different iterations, with additional iterations included.

Method	7B	13B
Base	66.61	68.08
Only R_T	68.46	68.12
Only R_I	67.49	69.23
CSR (Ours)	72.39	71.95

Table A1: Updated Table 2 in the main paper