Average Human Judgements

Legend:
- 16plays_36LM
- Dictionary
- Video Baseline

Categories: semantic adequacy, dissimilarity, style, overall