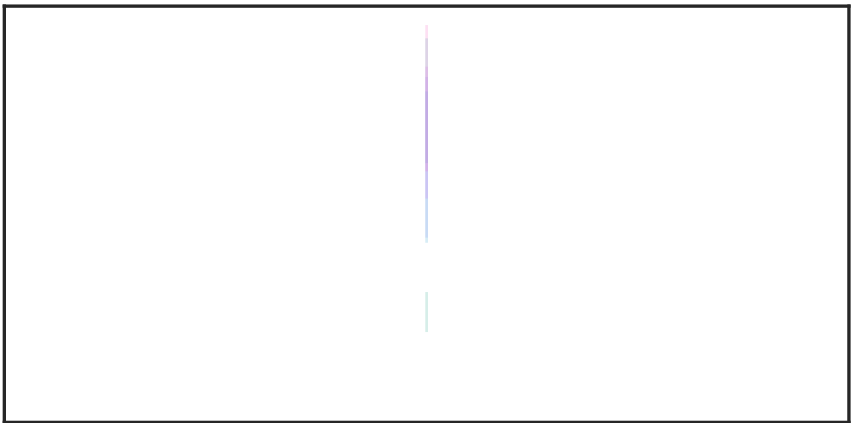
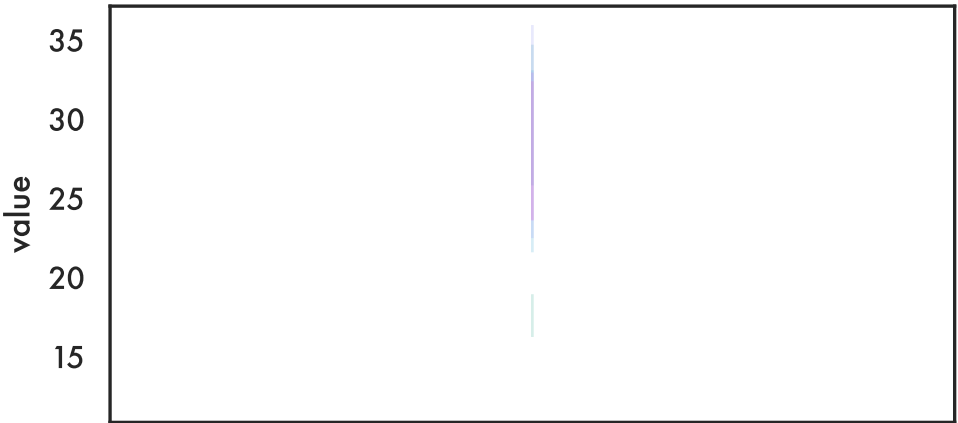


Overview of Turn Scores: referencegame

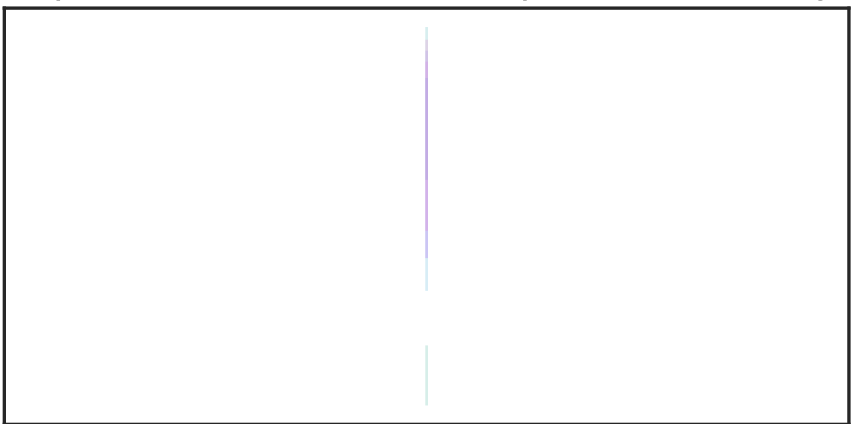
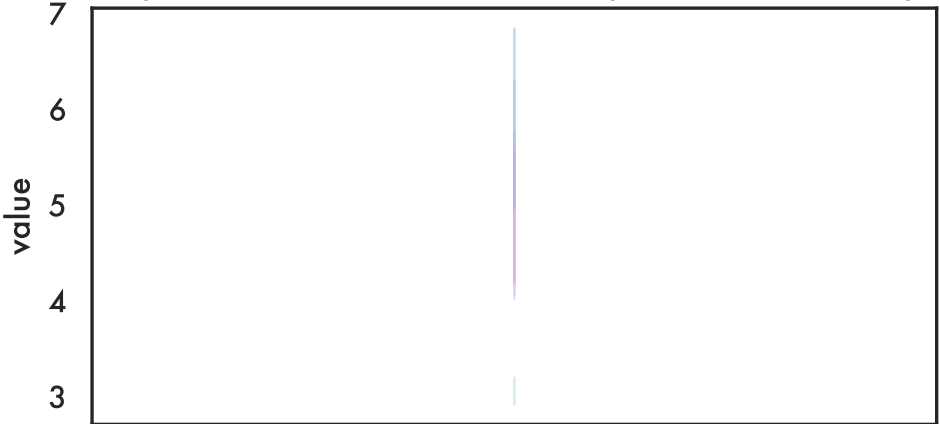
metric = Generated Expression Length | experiment = 0_hard_grids_edit_distance_2

metric = Generated Expression Length | experiment = 1_hard_grids_edit_distance_4



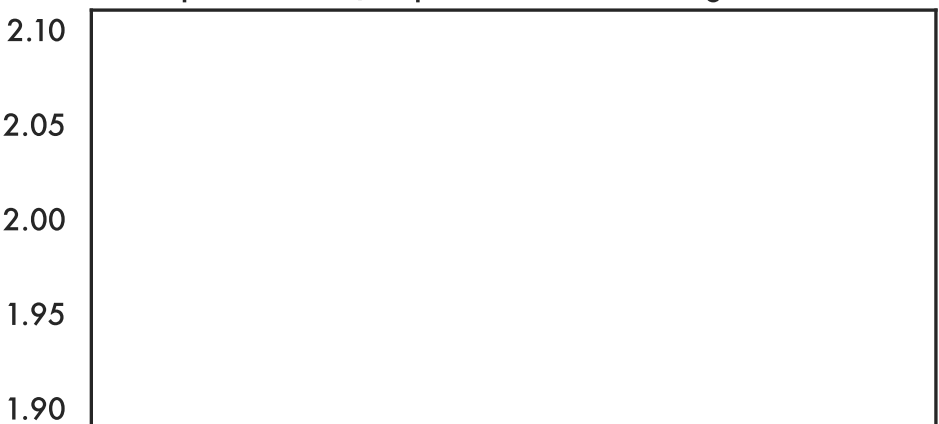
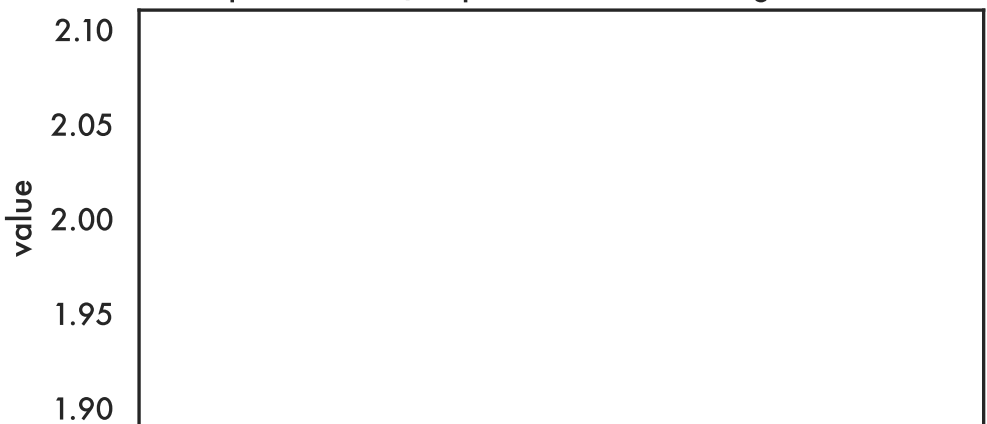
metric = Generated Expression Number of Tokens | experiment = 0_hard_grids_edit_distance_2

metric = Generated Expression Number of Tokens | experiment = 1_hard_grids_edit_distance_4



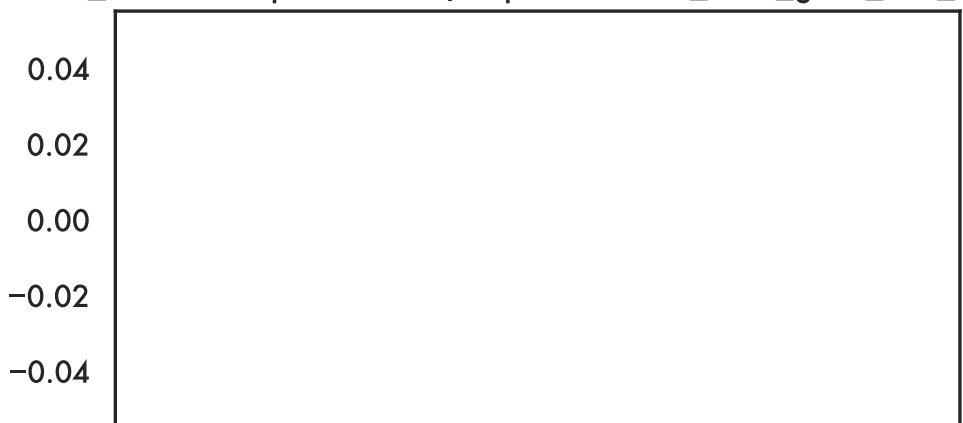
metric = Request Count | experiment = 0_hard_grids_edit_distance_2

metric = Request Count | experiment = 1_hard_grids_edit_distance_4



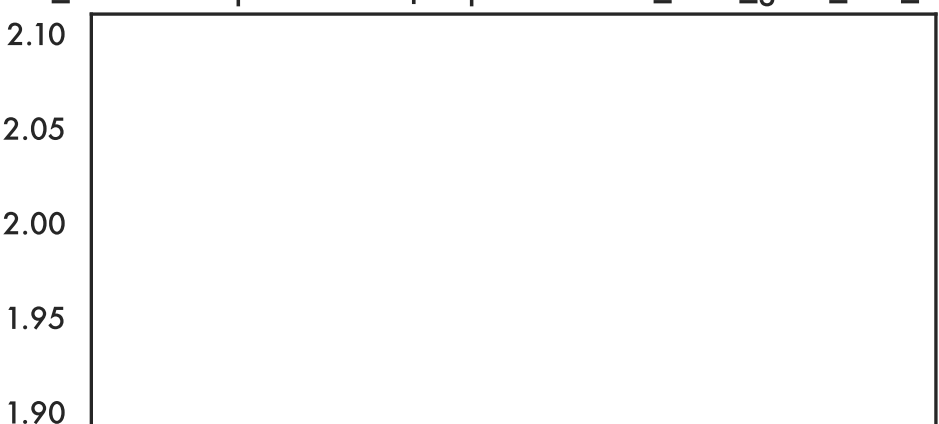
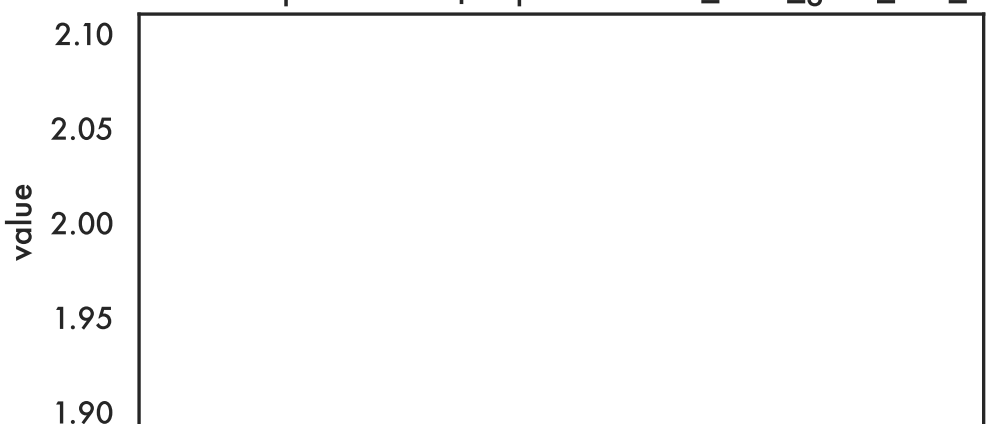
metric = Violated Request Count | experiment = 0_hard_grids_edit_distance_2

metric = Violated Request Count | experiment = 1_hard_grids_edit_distance_4



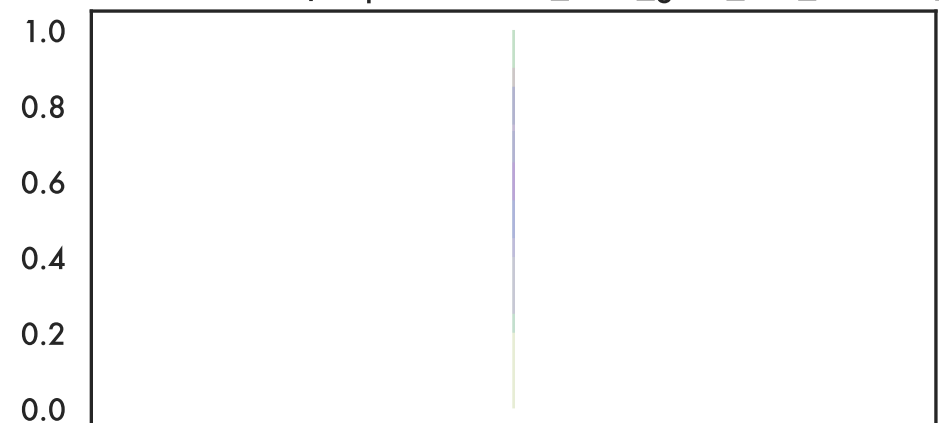
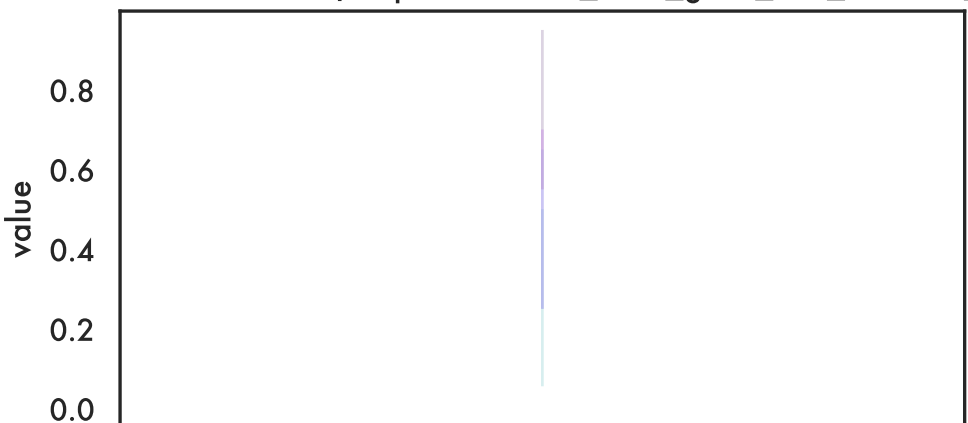
metric = Parsed Request Count | experiment = 0_hard_grids_edit_distance_2

metric = Parsed Request Count | experiment = 1_hard_grids_edit_distance_4



metric = Success | experiment = 0_hard_grids_edit_distance_2

metric = Success | experiment = 1_hard_grids_edit_distance_4



- model
- lm-lm
 - ko-ko
 - flc-flc
 - ost-ost
 - vcn-vcn
 - cl-cl
 - 3-3
 - 3.5-3.5
 - 3.5-4
 - 4-3.5
 - 4-4