n_screenshots vs is_correct 1.0 8.0 0.6 is_correct 0.4 0.2 0.0 0 10 20 30 40 70 50 60 n_screenshots