gemgem ll8bll8b	iolated Request Count experiment = 0_easy II70bII70b mismis mixmix openopenqwenqwen = Violated pattern experiment = 0_easy	gemge	em		is mixmix openopenqwenqwen	metric = Violated Request Count e gemgem	s mixmix openopenqwenqwen	metric = Violated Request Count exp	1.0	metric = Violated Request Count ex gemgem	nixmix openopenqwenqwen	metric = Violated Request Count gemgem		metric = Violated Req 1.0 0.8 0.6 0.4 0.2 gemgem 8b 8b 70b metric = Violated 1.0	rest Count experiment = 7_themselves 70b mismis mixmix openopenqwenqwe pattern experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen Violated token length experiment = 0_easy		em	0.8 0.6 0.4 0.2 0.0 gemgem 8b 8b 70b 70b mism metric = Violated token length		gemgem		gemgem	0.8 0.6 0.4 0.2 0.0 ixmix openopenqwenqwen riment = 1_medium 0.20	gemgem 8b 18b 70b 170b mismis m metric = Violated token length exp		gemgem		0.8 0.6 0.4 0.2 0.0 gemgem 8b 8b 70b metric = Violated tok	70b mismis mixmix openopengwenqwe en length experiment = 7_themselves
-0.02 -0.04 gemgem II8bII8b	o II70bII70b mismis mixmix openopenqwenqwen Parsed Request Count experiment = 0_easy	-0.02 -0.04 gemge	em 8b 8b 70b 70b mismis mixmix openopenqwenqweretric = Parsed Request Count experiment = 10_same_gender	0.10 0.05		gemgem II8bII8b II70bII70b mismis metric = Parsed Request Count ex	-0.02 -0.04 s mixmix openopenqwenqwen	gemgem 8b 8b 70b 70b mismis mi metric = Parsed Request Count expe	0.10 0.05 0.00 ixmix openopenqwenqwen	gemgem II8bII8b II70bII70b mismis m metric = Parsed Request Count exp	-0.02 -0.04 nixmix openopenqwenqwen	gemgem	s mixmix openopenqwenqwer	-0.02 -0.04	70b mismis mixmix openopenqwenqwe est Count experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen = Request Count experiment = 0_easy		em 8b 8b 70b 70b mismis mixmix openopenqwenqwer metric = Request Count experiment = 10_same_gender	gemgem 8b 8b 70b 70b mism metric = Request Count exp		gemgem		gemgem		gemgem II8bII8b II70bII70b mismis m metric = Request Count expering		gemgem			70b mismis mixmix openopenqwenqwe Count experiment = 7_themselves
10 8 9 6 4 2 0 gemgem 8b 8b metric = Re	o II70bII70b mismis mixmix openopenqwenqwen Request Success Ratio experiment = 0_easy		em	gemgem		gemgem II8bII8b II70bII70b mismis metric = Request Success Ratio ex		gemgem 8b 8b 70b 70b mismis mi metric = Request Success Ratio expe	10 8 6 4 2 0 ixmix openopenqwenqwen eriment = 1_medium 1.0	gemgem II8bII8b II70bII70b mismis m metric = Request Success Ratio exp		gemgem II8bII8b II70bII70b mismis metric = Request Success Ratio	mixmix openopenqwenqwer experiment = 4_reprompt	gemgem II8bII8b II70bI metric = Request Suc	70b mismis mixmix openopenqwenqwe cess Ratio experiment = 7_themselves
gemgem ll8bll8b	o II70bII70b mismis mixmix openopenqwenqwen mber of completed turns experiment = 0_easy	gemge	em 8b 8b 70b 70b mismis mixmix openopenqwenqwerc = Number of completed turns experiment = 10_same_gender	en gemgem ll8bll8b ll70bll70b mism	is mixmix openopenqwenqwen	gemgem II8bII8b II70bII70b mismis metric = Number of completed turns	s mixmix openopenqwenqwen	gemgem II8bII8b II70bII70b mismis mi metric = Number of completed turns ex	xmix openopenqwenqwen	gemgem II8bII8b II70bII70b mismis m metric = Number of completed turns 0	nixmix openopenqwenqwen	gemgem II8bII8b II70bII70b mismis metric = Number of completed turn	mixmix openopenqwenqwer	gemgem ll8bll8b ll70bl	70b mismis mixmix openopenqwenqwe pleted turns experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen dumber of Agreements experiment = 0_easy		em 8b 8b 70b 70b mismis mixmix openopenqwenqwertric = Number of Agreements experiment = 10_same_gender			gemgem 8b 8b 70b 70b mismis metric = Number of Agreements e		gemgem 8b 8b 70b 70b mismis mi metric = Number of Agreements exp		gemgem II8bII8b II70bII70b mismis m metric = Number of Agreements ex		gemgem			70b mismis mixmix openopenqwenqwe greements experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen Agreement Efficiency experiment = 0_easy		em 8b 8b 70b 70b mismis mixmix openopenqwenqweretric = Agreement Efficiency experiment = 10_same_gender			gemgem		gemgem		gemgem		gemgem			70b mismis mixmix openopenqwenqwe Efficiency experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen of Mismatched Agreements experiment = 0_easy		em 8b 8b 70b 70b mismis mixmix openopenqwenqwer Number of Mismatched Agreements experiment = 10_same_gend		ais mixmix openopenqwenqwen ments experiment = 12_less_turns 3.5	gemgem II8bII8b II70bII70b mismis metric = Number of Mismatched Agreeme		gemgem 8b 8b 70b 70b mismis mi metric = Number of Mismatched Agreements	20 ixmix openopenqwenqwen s experiment = 1_medium 4.0 3.5	gemgem II8bII8b II70bII70b mismis m metric = Number of Mismatched Agreemen		gemgem II8bII8b II70bII70b mismis metric = Number of Mismatched Agreen	mixmix openopenqwenqwernents experiment = 4_reprompt	60 40 20 gemgem	70b mismis mixmix openopenqwenqwe hed Agreements experiment = 7_themselve
3.0 2.5 9	o II70bII70b mismis mixmix openopenqwenqwen Number of Reprompts experiment = 0_easy	2.5 2.0 1.5 1.0 0.5 0.0 gemge	em 8b 8b 70b 70b mismis mixmix openopenqwenqweretric = Number of Reprompts experiment = 10_same_gender	2.5 2.0 1.5 1.0 0.5 gemgem 8b 8b 70b 70b mism metric = Number of Reprompts		gemgem II8bII8b II70bII70b mismis metric = Number of Reprompts ex	3 2 1 2 s mixmix openopenqwenqwen xperiment = 13_more_turns 0.04	gemgem	3.0 2.5 2.0 1.5 1.0 0.5 0.0 eriment = 1_medium	gemgem II8bII8b II70bII70b mismis m metric = Number of Reprompts exp	2.5 2.0 1.5 1.0 0.5 0.0 0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	gemgem	mixmix openopenqwenqwerexperiment = 4_reprompt	2.5 2.0 1.5 1.0 0.5 gemgem 8b 8b 70b metric = Number of F	70b mismis mixmix openopenqwenqwe eprompts experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen umber of Redundancy experiment = 0_easy	-0.02 -0.04 gemge	em 8b 8b 70b 70b mismis mixmix openopenqwenqwertric = Number of Redundancy experiment = 10_same_gender	-0.02 -0.04 en gemgem ll8bll8b ll70bll70b mism	-0.02 -0.04 his mixmix openopenqwenqwen	gemgem II8bII8b II70bII70b mismis metric = Number of Redundancy e	-0.02 -0.04 s mixmix openopenqwenqwen	gemgem 8b 8b 70b 70b mismis mi metric = Number of Redundancy exp	-0.02 -0.04 ixmix openopenqwenqwen	gemgem II8bII8b II70bII70b mismis m metric = Number of Redundancy ex	0.4 0.2 0.0 nixmix openopenqwenqwen	gemgem II8bII8b II70bII70b mismis metric = Number of Redundancy			70b mismis mixmix openopenqwenqweedundancy experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen ic = Friendzoned experiment = 0_easy		em 8b 8b 70b 70b mismis mixmix openopenqwenqwer metric = Friendzoned experiment = 10_same_gender	1.0 0.8 0.6 0.4 0.2 0.0 gemgem 8b 8b 70b 70b mism metric = Friendzoned expe		gemgem II8bII8b II70bII70b mismis metric = Friendzoned experir		gemgem 8b 8b 70b 70b mismis mi metric = Friendzoned experime		gemgem II8bII8b II70bII70b mismis m metric = Friendzoned experim		gemgem			70b mismis mixmix openopenqwenqwe oned experiment = 7_themselves
-0.02 -0.04 gemgem II8bII8b	o II70bII70b mismis mixmix openopenqwenqwen ic = Out of turns experiment = 0_easy	0.15 0.10 0.05 0.00 gemge	em 8b 8b 70b 70b mismis mixmix openopenqwenqwer metric = Out of turns experiment = 10_same_gender	0.02 0.00 -0.02 -0.04 gemgem 8b 8b 70b 70b mism metric = Out of turns expe	0.10 0.09 0.09 0.00 0.00 0.00	gemgem		gemgem		gemgem II8bII8b II70bII70b mismis m metric = Out of turns experim		gemgem			70b mismis mixmix openopenqwenqwe urns experiment = 7_themselves
0.15 0.00 gemgem II8bII8b metric 0.04	o II70bII70b mismis mixmix openopenqwenqwen c = Out of retries experiment = 0_easy		em 8b 8b 70b 70b mismis mixmix openopenqwenqwer metric = Out of retries experiment = 10_same_gender	0.6 0.4 0.2 gemgem 8b 8b 70b 70b mism metric = Out of retries exp	-0.02 -0.04 -0.04 -0.05 mixmix openopenqwenqwen	gemgem	0.10 0.05 0.00 s mixmix openopenqwenqwen	gemgem 8b 8b 70b 70b mismis mi metric = Out of retries experime	0.3 0.2 0.1 0.0 0.1 0.0 0.0 0.0 0.0	gemgem II8bII8b II70bII70b mismis m metric = Out of retries experim	-0.02 -0.04 nixmix openopenqwenqwen	gemgem	s mixmix openopenqwenqwer	0.10 0.05 0.00 gemgem 8b 8b 70b	70b mismis mixmix openopenqwenqwe etries experiment = 7_themselves
-0.02 -0.04 gemgem II8bII8b	o II70bII70b mismis mixmix openopenqwenqwen	0.02 0.00 -0.02 -0.04 gemge	em 8b 8b 70b 70b mismis mixmix openopemqwenqwermetric = Turn penalty experiment = 10_same_gender	0.00 -0.02 -0.02 -0.04 gemgem	-0.02 -0.04 nis mixmix openopenqwenqwen	gemgem	-0.02 -0.04 s mixmix openopenqwenqwen	gemgem 8b 8b 70b 70b mismis mi metric = Turn penalty experime	-0.02 -0.04 ixmix openopenqwenqwen	gemgem	0.4 0.2 0.0 nixmix openopenqwenqwen	gemgem	mixmix openopenqwenqwer riment = 4_reprompt	0.00	70b mismis mixmix openopenqwenqwenalty experiment = 7_themselves
	II70bII70b mismis mixmix openopenqwenqwen Agreement penalty experiment = 0_easy		em	0.00 -0.02 -0.04 gemgem 8b 8b 70b 70b mism metric = Agreement penalty 6		gemgem II8bII8b II70bII70b mismis metric = Agreement penalty exp		gemgem		gemgem II8bII8b II70bII70b mismis m metric = Agreement penalty expe		gemgem II8bII8b II70bII70b mismis metric = Agreement penalty e			70b mismis mixmix openopenqwenqwe penalty experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen Inefficiency penalty experiment = 0_easy		em 8b 8b 70b 70b mismis mixmix openopenqwenqwer netric = Inefficiency penalty experiment = 10_same_gender	15 10 gemgem		gemgem II8bII8b II70bII70b mismis metric = Inefficiency penalty exp		gemgem		gemgem		gemgem		15 10 gemgem II8bII8b II70bI metric = Inefficiency	70b mismis mixmix openopenqwenqwe penalty experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen c = Error penalty experiment = 0_easy		em 8b 8b 70b 70b mismis mixmix openopenqwenqwer metric = Error penalty experiment = 10_same_gender	8 6 4 2 gemgem		gemgem	s mixmix openopenqwenqwen iment = 13_more_turns 0.04	gemgem 8b 8b 70b 70b mismis mi metric = Error penalty experime	20 10 ixmix openopenqwenqwen ent = 1_medium 0.04	gemgem 8b 8b 70b 170b mismis m metric = Error penalty experim	1 nixmix openopenqwenqwen nent = 2_difficult 5	gemgem	mixmix openopenqwenqwereriment = 4_reprompt	2 1 gemgem	70b mismis mixmix openopenqwenqwe nalty experiment = 7_themselves
-0.02 -0.04 gemgem II8bII8b	o II70bII70b mismis mixmix openopenqwenqwen Location agreement experiment = 0_easy	-0.02 -0.04 gemge	em 8b 8b 70b 70b mismis mixmix openopemwenqweretric = Location agreement experiment = 10_same_gender	-0.02 -0.04	-0.02 -0.04 his mixmix openopenqwenqwen	gemgem	-0.02 -0.04 s mixmix openopenqwenqwen	gemgem 8b 8b 70b 70b mismis mi metric = Location agreement exper	-0.02 -0.04 ixmix openopenqwenqwen	<u>.</u>	1 nixmix openopenqwenqwen	gemgem		-0.02 -0.04 gemgem	70b mismis mixmix openopenqwenqwe reement experiment = 7_themselves
0.6 0.4 0.2 gemgem II8bII8b metric = 1.0 0.8	o II70bII70b mismis mixmix openopenqwenqwen = Action agreement experiment = 0_easy		em 8b 8b 70b 70b mismis mixmix openopenqwenqwermetric = Action agreement experiment = 10_same_gender	0.6 0.4 0.2 gemgem 8b 8b 70b 70b mism metric = Action agreement e	0.0 0.1 0.2 0.3 0.5 0.6 1.6 0.6 0.7 0.7 0.7 0.8 0.8	gemgem II8bII8b II70bII70b mismis metric = Action agreement expe	0.6 0.4 0.2 0.0 s mixmix openopenqwenqwen periment = 13_more_turns 1.0 0.8	gemgem	0.6 0.4 0.2 0.0 0.0 0.0 0.0 0.0 0.0	gemgem	0.6 0.4 0.2 0.0 nixmix openopenqwenqwen riment = 2_difficult 1.0 0.8	gemgem II8bII8b II70bII70b mismis metric = Action agreement ex	mixmix openopenqwenqwer eperiment = 4_reprompt	0.6 0.4 0.2 gemgem 8b 8b 70b metric = Action agr 1.0 0.8	70b mismis mixmix openopenqwenqwe eement experiment = 7_themselves
0.6 0.4 0.2 0.0 gemgem	o II70bII70b mismis mixmix openopenqwenqwen = Time agreement experiment = 0_easy	0.6 0.4 0.2 0.0 gemge r 1.0	em 8b 8b 70b 70b mismis mixmix openopenqwenqwermetric = Time agreement experiment = 10_same_gender	0.6 0.4 0.2 gemgem	0.4 0.5 is mixmix openopenqwenqwen periment = 12_less_turns 1.6 0.8	gemgem	0.6 0.4 0.2 s mixmix openopenqwenqwen eriment = 13_more_turns 1.0 0.8	gemgem	0.6 0.4 0.2 0.0 0.0 0.0 0.0 0.0 0.0	gemgem II8bII8b II70bII70b mismis m metric = Time agreement exper	0.6 0.4 0.2 0.0 0.0 0.0 0.0 0.0 0.0	gemgem	mixmix openopenqwenqwer periment = 4_reprompt	0.6 0.4 0.2 gemgem 8b 8b 70b metric = Time agree 0.8 0.7 0.6	70b mismis mixmix openopenqwenqwe ement experiment = 7_themselves
	o II70bII70b mismis mixmix openopenqwenqwen etric = Aborted experiment = 0_easy	0.6 0.4 0.2 0.0 gemge	em Bb Bb T0b T0b mismis mixmix openopenqwenqwer metric = Aborted experiment = 10_same_gender	0.6 0.4 0.2 gemgem 8b 8b 70b 70b mism metric = Aborted experi		gemgem II8bII8b II70bII70b mismis metric = Aborted experime		gemgem		gemgem II8bII8b II70bII70b mismis m metric = Aborted experimer		gemgem		0.5 0.4 0.3 0.2 0.1 gemgem 8b 8b 70b metric = Abort	70b mismis mixmix openopenqwenqwe ed experiment = 7_themselves
0.6 0.4 0.2 gemgem II8bII8b met 1.0 0.8	o II70bII70b mismis mixmix openopenqwenqwen etric = Success experiment = 0_easy	0.8 0.6 0.4 0.2 0.0 gemge 1.0 0.8	em II8bII8b II70bII70b mismis mixmix openopenqwenqwer metric = Success experiment = 10_same_gender	0.8 0.6 0.4 0.2 0.0 gemgem 8b 8b 70b 70b mism metric = Success experi	0.4 0.5 0.6 0.7 0.7 0.7 1.6 1.6	gemgem II8bII8b II70bII70b mismis metric = Success experiments	0.6 0.4 0.2 s mixmix openopenqwenqwen ent = 13_more_turns 1.0	gemgem II8bII8b II70bII70b mismis mi metric = Success experiment	0.6 0.4 0.2 0.0 0.0 1.0	gemgem 8b 8b 70b 170b mismis m metric = Success experimer	0.6 0.4 0.2 0.0 nixmix openopenqwenqwen nt = 2_difficult 1.0	gemgem II8bII8b II70bII70b mismis metric = Success experin	mixmix openopenqwenqwerment = 4_reprompt	0.6 0.4 0.2 0.0 gemgem II8bII8b II70bI metric = Success 1.0 0.8	70b mismis mixmix openopenqwenqwe ess experiment = 7_themselves
0.6 0.4 0.2 gemgem II8bII8b me	o II70bII70b mismis mixmix openopenqwenqwen netric = Lose experiment = 0_easy	0.6 0.4 0.2 0.0 gemge	em	0.8 0.6 0.4 0.2 gemgem 8b 8b 70b 70b mism metric = Lose experim	o.s o.s o.s o.s o.s o.s o.s o.s	gemgem	0.8 0.6 0.4 0.2 0.0 0.0 0.0 0.0 0.0 0.0	gemgem	0.8 0.6 0.4 0.2 0.7 0.8 0.8 0.8 0.9 0.9 0.9 0.9 0.9	gemgem II8bII8b II70bII70b mismis m metric = Lose experiment		gemgem		0.6 0.4 0.2 gemgem II8bII8b II70bI metric = Los	70b mismis mixmix openopengwenqwe e experiment = 7_themselves
-0.02 -0.04 gemgem II8bII8b	o II70bII70b mismis mixmix openopenqwenqwen ric = Main Score experiment = 0_easy	0.20 0.15 0.10 0.05 0.00 gemge	em II8bII8b II70bII70b mismis mixmix openopenqwenqwer metric = Main Score experiment = 10_same_gender	0.02 0.00 -0.02 -0.04 gemgem	0.10 0.09 0.00 0.00 0.00 0.00	gemgem	0.15 0.10 0.05 mixmix openopengwenqwen ment = 13_more_turns 100	gemgem 8b 8b 70b 70b mismis mi metric = Main Score experime	0.15 0.10 0.05 0.00 0.00 0.00 0.00 0.00 0.00	gemgem II8bII8b II70bII70b mismis m metric = Main Score experime	0.15 0.10 0.05 0.00 nixmix openopenqwenqwen ent = 2_difficult 100	gemgem	mixmix openopenqwenqwer riment = 4_reprompt	0.15 0.00 gemgem	70b mismis mixmix openopenqwenqwe core experiment = 7_themselves
60 ang 40 20 gemgem II8bII8b mei	o II70bII70b mismis mixmix openopenqwenqwen etric = Played experiment = 0_easy	80 60 40 20 gemge	em	60 40 20 gemgem	nis mixmix openopenqwenqwen ment = 12_less_turns	gemgem	mixmix openopenqwenqwen ent = 13_more_turns 1.0	gemgem	60 40 20 21 22 22 32 32 32 33 34 35 36 36 37 37 38 30 30 30 30 30 30 30 30 30 30 30 30 30	gemgem	anixmix openopenqwenqwen $t = 2_{difficult}$ 1.0	gemgem	mixmix openopenqwenqwennent = 4_reprompt	60 40 20 gemgem	70b mismis mixmix openopenqwenqwe ed experiment = 7_themselves
0.8 0.6 0.4 0.2	o II70bII70b mismis mixmix openopengwenqwen	0.8 0.6 0.4 0.2 gemge	em II8bII8b II70bII70b mismis mixmix openopenqwenqwer	0.8 0.6 0.4 0.2 gemgem 8b 8b 70b 70b mis-m mode	0.6 0.6 0.7 0.7 0.7 0.8 0.8 0.8 0.9 0.9 0.9 0.9 0.9	gemgem II8bII8b II70bII70b mismis model	0.8 0.6 0.4 0.2 mixmix openopenqwenqwen	gemgem II8bII8b II70bII70b mismis mi model	0.8 0.6 0.4 0.2 0.0 0.0 0.0	gemgem II8bII8b II70bII70b mismis model	0.8 0.6 0.4 0.2 0.0 0.0	gemgem II8bII8b II70bII70b mismis model	mixmix openopenqwenqwer	0.8 0.6 0.4 0.2 gemgem	