	S2	L1	S1	L0
Actual1	$\phi \cdot (L_1(s \mid w) + (1-\phi) \cdot V[L_1(s \mid w)]$		$L_0(s \mid w)$	$[\![w]\!](s)$
Actual2	$\phi_2 \cdot (L_1(s \mid w) + (1-\phi_2) \cdot V[L_1(s \mid w)]$		$\phi_1 \cdot (L_0(s \mid w) + (1-\phi_1) \cdot V[L_0(s \mid w)]$	same as above
Self1 (Cogsci 2017)	$L_1(s, \phi_1 \mid w)$	same as above		same as above
Self2	$\phi_3 \cdot (L_1(s, \phi_1 \mid w) + (1-\phi_3) \cdot V[L_1(s \mid w)]$	same as above		same as above
Self3*	$L_1(s \mid w)$	same as above		same as above
Self4*	$L_1(\phi_1 \mid w)$	same as above		same as above

1) Actual1:

- S1 only considers epistemic goal

- S2 has both epistemic and social goals (actual, not self-presentational)

2) Actual2:

- S1 considers both epistemic and social goal

- S2 has both epistemic and social goals (actual, not self-presentational)

3) Self1:

- S1 considers both epistemic and social goal

- S2 reasons about L1's inferred goals and state

4) Self2:

- S1 considers both epistemic and social goal

- S2 reasons about L1's inferred goals and state AND his true social goals

5) Self3:

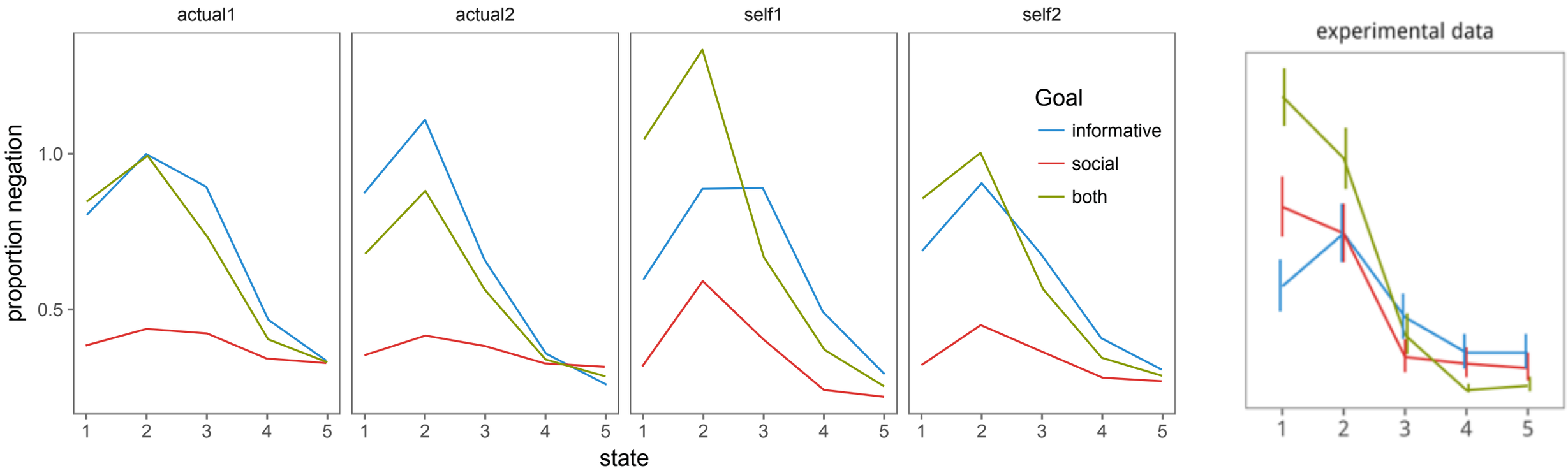
- S1 considers both epistemic and social goal

- S2 reasons about L1's state inference only (inferred goals don't matter)

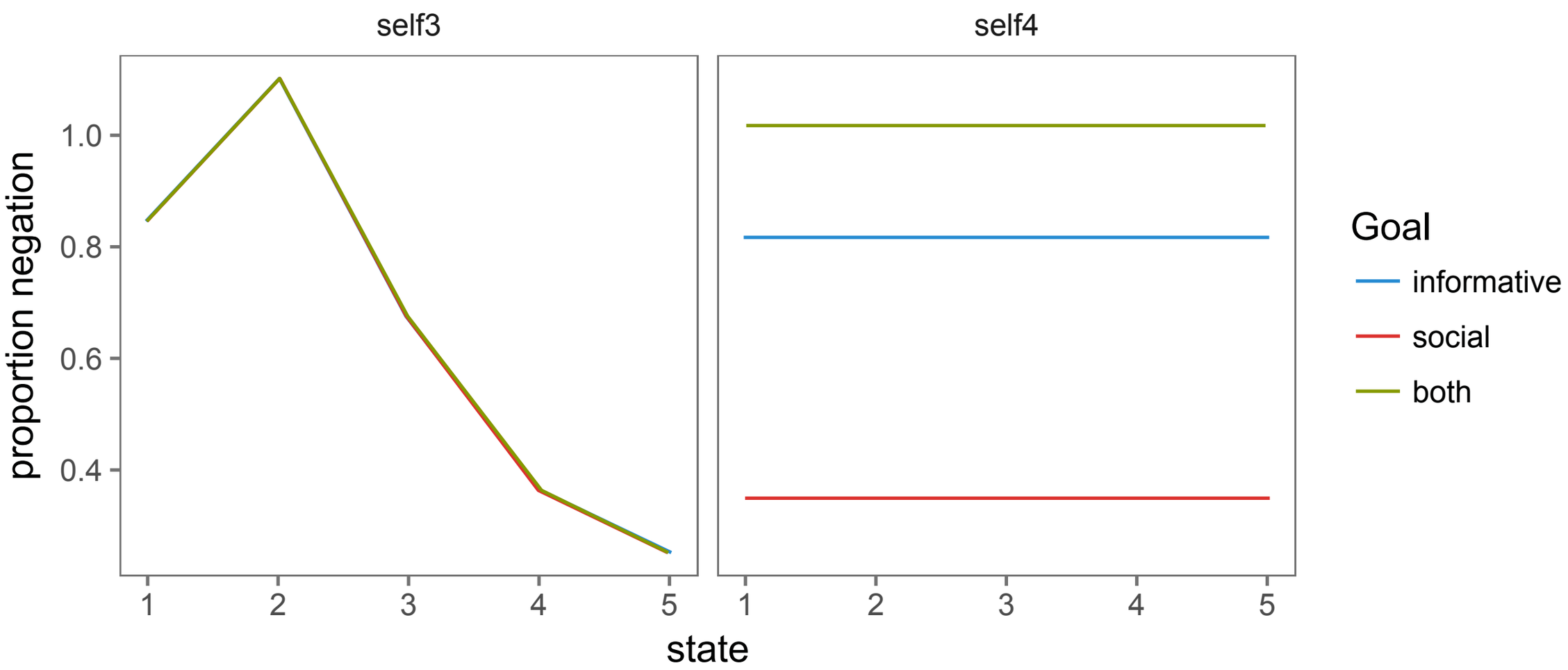
6) Self4:

- S1 considers both epistemic and social goal

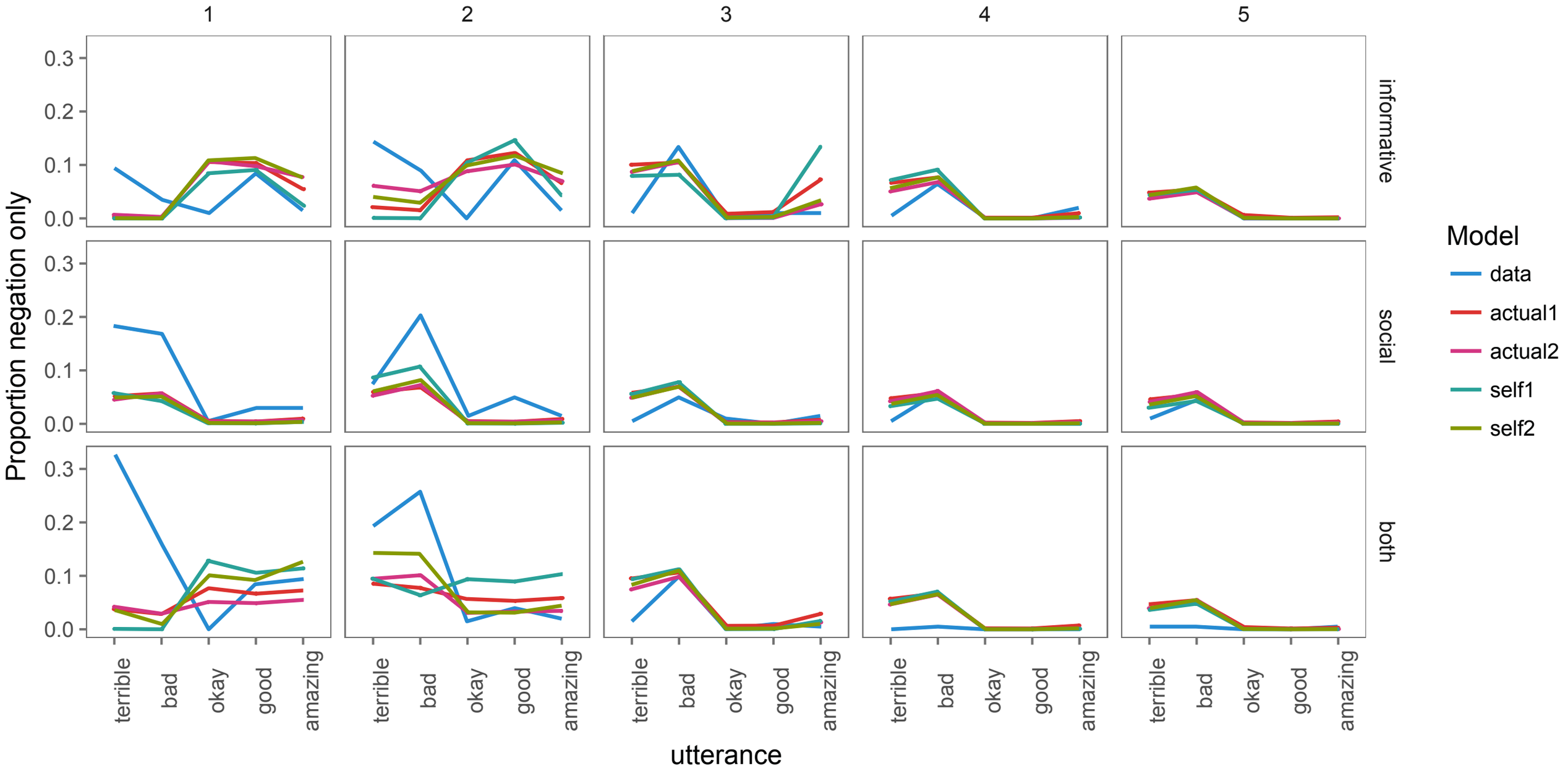
- S2 reasons about L1's goal inference only (true state doesn't matter)



* where goal is actual for actual1 and actual2 models (i.e. the speaker wanted to BE informative and/or social), but self-presentational for self1 and self2 models (i.e. the speaker wanted to LOOK informative and/or social)



self3: speaker wanted to convey state (but not goals)
self4: speaker wanted to convey goals (but not state)



Looking at expected proportion of utterances with negation only (e.g. “not terrible”). Just based on shape, self2 seems to capture it best, but still not very successfully…