Social Projection Meets Social Reality: A Probabilistic Implementation of the Inductive Reasoning Model

Samuel G. B. Johnson

APPENDIX

Here, I detail one full run of the simulation. The results are summarized in Tables A1 (ground truth parameters), A2 (observed and virtual samples used for inference), A3 (posterior means for each model), and A4 (resulting social judgment phenomena).

First, each of 8 traits is randomly simulated. The true mean level of each of the 8 traits for the ingroup (μ_I) and outgroup (μ_O) is drawn from a uniform distribution, and then perturbated in such a way as to produce an average correlation of .47 (SD = .31) between these variables (Table A1, Columns (a) and (b)). In this run of the simulation, these variables are correlated at r = .48. The desirability of each trait (δ) is then drawn from a uniform [0,1] distribution (Table A1, Column (c)).

	(a)	(b)	(c)	(d)	(e)	(f)
	Group-Level Trait Means		Desirability	Target-Level Trait Values		
Trait	μ_I	μ_O	δ	τ_I	τ_O	τ_S
(i)	.71	.40	.71	.69	.44	.86
(ii)	.16	.19	.25	.19	.60	.09
(iii)	.82	.26	.39	.94	.12	.88
(iv)	.91	.83	.09	.96	.83	.99
(v)	.47	.69	.96	.88	.45	.38
(vi)	.59	.64	.01	.50	.73	.72
(vii)	.69	.53	.57	.89	.66	.63
(viii)	.39	.32	.76	.55	.26	.32

Table A1. Ground truth parameters in one run of the simulation.

We then translate these group-level trait means μ_I and μ_O into beta distributions. Since this information is available to social perceivers, these parameters are recorded in Table A2, Columns (a)–(d). The simulation assumes that the heterogeneity of each group's distribution corresponds to a virtual sample size of $S_G = \alpha_I + \beta_I = \alpha_O + \beta_O = 5$. (Larger numbers indicate more homogeneous distributions, reflecting greater confidence in the true mean of the distribution.) Since $\mu_I = \alpha_I/(\alpha_I + \beta_I)$ and $\mu_O = \alpha_O/(\alpha_O + \beta_O)$, we can solve for α_I, β_I, α_O, and β_O. Thus, α_I and α_O are just the trait means multiplied by S_G, and β_I and β_O are the remainder of subtracting α_I and α_O from S_G.

Next, we set values of τ_I and τ_O for the ingroup and outgroup targets and τ_S for the self; τ_I and τ_S are both drawn from the Beta(α_I, β_I) distribution while τ_O is drawn from the Beta(α_O, β_O) distribution. Since these are ground truth parameters, they are recorded in Table A1, Columns (d)–(f).

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)	(o)	(p)	(q)	(r)
	Group Knowledge (Virtual Observations)				Individual Knowledge (Observations)				Self-Knowledge (Projection)						Noise (Hallucinations)			
Trait	α_I	β_I	α_O	β_O	a_I	b_I	a_O	b_O	a_S	b_S	$\pi_I a_S$	$\pi_I b_S$	$\pi_O a_S$	$\pi_O b_S$	a_{NI}	b_{NI}	a_{NO}	b_{NO}
(i)	3.56	1.44	2.02	2.98	4.00	1.00	1.00	4.00	10.00	0.00	5.00	0.00	0.00	0.00	0.93	8.07	7.60	1.40
(ii)	0.78	4.22	0.95	4.05	0.00	5.00	3.00	2.00	0.00	10.00	0.00	5.00	0.00	0.00	1.21	7.79	8.61	0.39
(iii)	4.11	0.89	1.29	3.71	5.00	0.00	1.00	4.00	7.00	3.00	3.50	1.50	0.00	0.00	5.56	3.44	1.61	7.39
(iv)	4.53	0.47	4.17	0.83	5.00	0.00	4.00	1.00	7.00	3.00	3.50	1.50	0.00	0.00	5.18	3.82	7.76	1.24
(v)	2.35	2.65	3.47	1.53	5.00	0.00	4.00	1.00	5.00	5.00	2.50	2.50	0.00	0.00	3.97	5.03	2.55	6.45
(vi)	2.96	2.04	3.21	1.79	1.00	4.00	4.00	1.00	3.00	7.00	1.50	3.50	0.00	0.00	0.15	8.85	1.57	7.43
(vii)	3.44	1.56	2.67	2.33	5.00	0.00	3.00	2.00	6.00	4.00	3.00	2.00	0.00	0.00	1.68	7.32	7.88	1.12
(viii)	1.94	3.06	1.60	3.40	3.00	2.00	1.00	4.00	3.00	7.00	1.50	3.50	0.00	0.00	0.61	8.39	8.48	0.52

Table A2. Sources of data used to induce trait estimates by social perceivers.

From here, we can generate observations of each target. For the ingroup and outgroup targets, we do this by simulating S_I or S_O draws from a Binomial(τ_I) or Binomial(τ_O) distribution, respectively. Both S_I and S_O are set to 5. The number of successes and failures for the ingroup (a_I and b_I) and outgroup (a_O and b_O) target are summarized in Table A2, Columns (e)–(h).

For observations of the self, we draw a biased sample of S_S observations for each trait—for positive traits ($\delta > 0.5$), self-observations are drawn from Binomial($\tau_S+\gamma(1–\tau_S)(2\delta–1)$) and for negative traits ($\delta < 0.5$) from Binomial ($\tau_S– \gamma(\tau_S)(1–2\delta)$) distribution, resulting in a_S successes and b_S failures. Intuitively, this sampling bias increases or decreases the probability of success proportional to how desirable a trait is. We set $S_S = 10$ and $\gamma = 0.56$ reflecting a moderate degree of wishful thinking. The number of successes and failures in this biased sample, a_S and b_S, are summarized in Columns (i)–(j). These are then projected onto the ingroup and outgroup targets as though they were observed directly of those targets, but scaled by projection constants, π_I and π_O, as shown in Columns (k)–(n). These are set to $\pi_I = .50$ and $\pi_O = .00$, reflecting moderate ingroup projection and no outgroup projection.

Finally, "hallucinatory" observations (noise) are drawn, with the number of successes a_{NI} and a_{NO} for ingroup and outgroup targets, respectively, drawn from a uniform [0,S_N] distribution (Columns (o)–(r)). The failure counts b_{NI} and b_{NO} are just $S_N – a_{NI}$ and $S_N – a_{NO}$, respectively. S_N was set to 9.

The mean of each model's posterior (Table 1 in the main text) can be calculated as the sum of all "successes" divided by the sum of all sample sizes, for whichever inductive sources the model uses. The numerator and denominator of each estimate can be calculated from the data in Table A2 using the expressions summarized in Table A3. Each model's numerical estimate is also listed in Table A3.

Model		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	
Source	**Group**	No	Yes	No	Yes	No	Yes	No	Yes	**Ground Truth**
	Individual	No	No	Yes	Yes	No	No	Yes	Yes	
	Self	No	No	No	No	Yes	Yes	Yes	Yes	
	Noise	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
						Estimated τ_I				True τ_I
Numerator		$1+a_{NI}$	α_I+a_{NI}	$1+a_I+a_{NI}$	$\alpha_I+a_I+a_{NI}$	$1+\pi_I a_S+a_{NI}$	$\alpha_I+\pi_I a_S+a_{NI}$	$1+a_I+\pi_I a_S+a_{NI}$	$\alpha_I+a_I+\pi_I a_S+a_{NI}$	
Denominator		$2+S_N$	S_G+S_N	$2+S_I+S_N$	$S_G+S_I+S_N$	$2+\pi_I S_S+S_N$	$S_G+\pi_I S_S+S_N$	$2+S_I+\pi_I S_S+S_N$	$S_G+S_I+\pi_I S_S+S_N$	
Trait	**(i)**	.18	.32	.37	.45	.43	.50	.52	.56	.69
	(ii)	.20	.14	.14	.10	.14	.10	.11	.08	.19
	(iii)	.60	.69	.72	.77	.63	.69	.72	.76	.94
	(iv)	.56	.69	.70	.77	.60	.70	.70	.76	.96
	(v)	.45	.45	.62	.60	.47	.46	.59	.58	.88
	(vi)	.10	.22	.13	.22	.17	.24	.17	.23	.50
	(vii)	.24	.37	.48	.53	.36	.43	.51	.55	.89
	(viii)	.15	.18	.29	.29	.19	.21	.29	.29	.55
						Estimated τ_O				True τ_O
Numerator		$1+a_{NO}$	α_O+a_{NO}	$1+a_O+a_{NI}$	$\alpha_O+a_O+a_{NO}$	$1+\pi_O a_S+a_{NO}$	$\alpha_O+\pi_O a_S+a_{NO}$	$1+a_O+\pi_O a_S+a_{NO}$	$\alpha_O+a_O+\pi_O a_S+a_{NO}$	
Denominator		$2+S_N$	S_G+S_N	$2+S_O+S_N$	$S_G+S_O+S_N$	$2+\pi_O S_S+S_N$	$S_G+\pi_O S_S+S_N$	$2+S_O+\pi_O S_S+S_N$	$S_G+S_O+\pi_I S_S+S_N$	
Trait	**(i)**	.78	.69	.60	.56	.78	.69	.60	.56	.44
	(ii)	.87	.68	.79	.66	.87	.68	.79	.66	.60
	(iii)	.24	.21	.23	.21	.24	.21	.23	.21	.12
	(iv)	.80	.85	.80	.84	.80	.85	.80	.84	.83
	(v)	.32	.43	.47	.53	.32	.43	.47	.53	.45
	(vi)	.23	.34	.41	.46	.23	.34	.41	.46	.73
	(vii)	.81	.75	.74	.71	.81	.75	.74	.71	.66
	(viii)	.86	.72	.65	.58	.86	.72	.65	.58	.26

Table A3. Estimates of τ_I and τ_O in each model from the main text. Numerical estimates can be derived by separately computing the numerator and denominator from the quantities in Table A2 and taking their ratio.

With these estimates in hand, we can check for intergroup accentuation, self-enhancement, ingroup favoritism, and differential accuracy for this particular model run. The outputs calculated across 50,000 simulation runs in Table 2 of the main text are calculated in Table A4 for this run only.

First, we see that self-judgments are strongly projected onto the ingroup target (high $r_{S,I}$) but not to the outgroup target (low $r_{S,O}$). Intergroup accentuation can be thought of as either a low value of $r_{I,O}$ or as a value of $r_{I,O}$ lower than the true correlation. Although the $r_{I,O}$ is negative in all models (prima facie evidence of intergroup accentuation), this inference was not an error in this particular simulation run because the traits of the ingroup and outgroup target really *were* negatively correlated ($r_{<I>,<O>} = -.10$). Self-enhancement is operationalized as higher $r_{S,D}$ than $r_{S,I}$, which is true in every model. Ingroup favoritism is operationalized as higher $r_{I,D}$ than $r_{O,D}$, which is true in about half of the models. Finally, differential accuracy is defined as higher $r_{I,<I>}$ than $r_{O,<O>}$, which is also true for every model. Overall, this simulation run is fairly representative of the model results as a whole.

Model	Correlations among Judgments							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
	$r_{S,I}$	$r_{S,O}$	$r_{I,O}$	$r_{S,D}$	$r_{I,D}$	$r_{O,D}$	$r_{I,<I>}$	$r_{O,<O>}$
1	.38	–.06	–.32	.25	–.04	.10	.73	.24
2	.61	.04	–.19	.25	–.11	.04	.86	.47
3	.63	–.20	–.24	.25	.21	–.07	.92	.59
4	.72	–.09	–.10	.25	.12	–.11	.95	.74
5	.77	–.06	–.26	.25	.10	.10	.88	.24
6	.82	.04	–.12	.25	.01	.04	.90	.47
7	.80	–.20	–.25	.25	.24	–.07	.94	.59
8	.83	–.09	–.10	.25	.16	–.11	.95	.74

Table A4. Extent of intergroup accentuation, self-enhancement, ingroup favoritism, and differential accuracy for each model, using only the example simulation run. (See Table 2 in main text for results of 50,000 runs.)