Supplementary Materials

Anthropomorphism and Emotional Reactions to Human-AI Collaboration

Supplementary Materials

Anthropomorphism and Emotional Reactions to Human-AI Collaboration

Scenarios Used in the Study

In the present research, we introduce a new experimental paradigm and design four different scenarios: human-only, full AI automation, and AI involvement in different roles, where both a human and an AI were involved in completing the job. We tested two forms of AI involvement: in one, the AI did the thinking (i.e., "Thinker" role) and the human carried out the action (i.e., "Doer" role); in the other, the human made the judgments and the AI executed. These scenarios were developed for four different professions (Artist, Engineer, Surgeon, and Therapist). Each participant read four job-based scenarios, presented in a random order. Below is a snippet of the Artist scenario in Human only, full AI automation, and AI Thinker-Human Doer (for other jobs, see https://osf.io/nkm9p/):

Human Only (Baseline):

> *Hi, my name is Taylor. I am an artist dedicated to exploring diverse styles and mediums. My process begins with conceptualizing the artwork, then using traditional tools to bring these ideas to life. The piece you're about to see is a reflection of my journey in art, which I painted using brushes and boards. These fundamental tools allow me to navigate through various artistic expressions, providing a tangible canvas for my visions.*

AI Automation:

Hi, my name is Taylor. I am an Artificial Intelligence (AI) painter. This painting is my creation, inspired by the intricate patterns of nature and the complexities of human emotions. My process starts with analyzing vast amounts of digital data from past human artists, which helps me understand various artistic styles and techniques. The piece you're about to see is painted by myself based on my large training data.

AI Thinker–Human Doer (Collaboration):

Hi, my name is Taylor. I am an artist dedicated to exploring diverse styles and mediums. I wanted a fresh wave of ideas to paint about and found it through an Artificial Intelligence (AI) assistant. The AI proposed an idea, and instructed how I should paint it. I brought our shared vision to life on canvas, blending traditional brush strokes with the AI's inspiration. I painted this piece you're observing as instructed by AI.

Experimental Manipulation Check

Participants completed an experimental manipulation check after each scenario to assess if they correctly identified whether the agent in this scenario was a human or an AI. It was done through an open-ended question: "How much sleep does [agent's name] need?", and participants could write their responses as numbers or in textual form. For the analysis, we converted answers to numeric values, converting written numbers and terms like "none" to numeric values and checking the numbers reported for the conditions (e.g., "8 hours" converted to 8, and "none" converted to 0). We selected the average and standard deviation of each country's human Thinker, human Doer scenarios responses, as the human sleep range for that country, and for all scenarios, we checked to exclude answers greater than "24 hours" (e.g., "100 hours"). In AI Thinker, AI Doer scenarios, participants were expected to report "0 hours" as agent names were

assigned to the AI system, which does not need to sleep (e.g., In Artist AI Thinker and AI Doer scenario: Hi, my name is Taylor. I am an Artificial Intelligence (AI) painter…) Based on this logic, we excluded all participants who provided a response that deviated from "0 hours". For the other scenarios (e.g., human Thinker, AI Doer), the agent's name represented the human agent who requires sleep, so participants were expected to answer hours within the human sleep range. Based on this, we excluded all participants whose response was "0 hours" for a human Thinker, AI Doer, or AI Thinker Human Doer scenario (An analysis without this exclusion is reported in the Supplementary Materials). We did not exclude missing responses, as removing missing responses could introduce selection bias, as participants may have skipped the question due to uncertainty rather than misunderstanding AI vs. human roles.

Multilevel Models for Each Dependent Variable

To assess the robustness of the main findings and the incremental contribution of model components, we estimated a series of stepwise multilevel models (Models 1–4) for each dependent variable (Negative Affect, Positive Affect, Appreciation, and Trust).

Table S1

Multilevel Models Predicting Negative Affect

	Model 1		**Model 2**		**Model 3**	
Predictors	*Estimates*	*95% CI*	*Estimates*	*95% CI*	*Estimates*	*95% CI*
(Intercept)	28.229 ***	23.080 – 33.377	29.548 ***	23.549 – 35.548	15.263 ***	7.847 – 22.680
AI Automation	12.135 ***	8.850 – 15.420	12.141 ***	8.852 – 15.431	39.674 ***	30.855 – 48.493

Collab: Human Thinker AI Doer	9.917 ***	6.982 – 12.852	9.914 ***	6.978 – 12.851	29.390 ***	21.325 – 37.454
Collab: AI Thinker human Doer	8.759 ***	5.824 – 11.694	8.757 ***	5.820 – 11.693	26.249 ***	18.360 – 34.138
JobGroup [Emotion]			-2.957 ***	-3.977 – -1.937	-4.572 ***	-6.530 – -2.615
A2I2			0.065	-1.197 – 1.327	6.363 ***	4.234 – 8.492
AI Automation × JobGroup [Emotion]					-0.185	-3.243 – 2.874
Collab: Human Thinker AI Doer × JobGroup [Emotion]					6.193 ***	3.446 – 8.941
Collab: AI Thinker human Doer× JobGroup [Emotion]					-0.184	-2.931 – 2.564
AI Automation × A2I2					-11.949 ***	-15.534 – -8.364
Collab: Human Thinker AI Doer ×A2I2					-9.351 ***	-12.407 – -6.294

Collab: AI Thinker human Doer ×A2I2			-7.238 *** -10.225 – -4.250
Random Effects			
ICC	0.64	0.65	0.65
N	1336	1336	1336
	4 Country	4 Country	4 Country
AIC	23748.533	23719.091	23630.118

*$p < .05$ **$p < .01$ ***$p < .001$

Table S2

Multilevel Models Predicting Positive Affect

	Model 1		**Model 2**		**Model 3**	
Predictors	*Estimates*	*95% CI*	*Estimates*	*95% CI*	*Estimates*	*95% CI*
(Intercept)	53.237 ***	39.478 – 66.996	24.874 ***	16.153 – 33.595	33.713 ***	24.126 – 43.301
AI Automation	-8.269 ***	-11.952 – -4.587	-6.957 ***	-10.174 – -3.739	-25.429 ***	-34.143 – -16.715
Collab: Human Thinker AI Doer	-3.422 *	-6.710 – -0.135	-3.983 **	-6.853 – -1.112	-12.962 **	-20.930 – -4.993
Collab: AI Thinker human Doer	-5.348 **	-8.636 – -2.060	-5.885 ***	-8.756 – -3.014	-18.971 ***	-26.766 – -11.175
JobGroup [Emotion]			-6.723 ***	-7.823 – -5.622	-1.102	-3.183 – 0.978

A2I2	13.019 ***	11.780 – 14.257	8.160 ***	6.060 – 10.260
AI Automation × JobGroup [Emotion]			-7.566 ***	-10.820 – -4.312
Collab: Human Thinker AI Doer × JobGroup [Emotion]			-11.828 ***	-14.748 – -8.908
Collab: AI Thinker human Doer × JobGroup [Emotion]			-3.562 *	-6.482 – -0.642
AI Automation × A2I2			9.725 ***	6.192 – 13.259
Collab: Human Thinker AI Doer × A2I2			6.182 ***	3.170 – 9.195
Collab: AI Thinker human Doer ×A2I2			6.178 ***	3.233 – 9.123

Random Effects

ICC	0.71	0.62	0.63
N	1336	1336	1336
	4 Country	4 Country	4 Country

AIC	24358.116	23855.705	23749.425

** p < .05 ** p < .01 *** p < .001*

Table S3

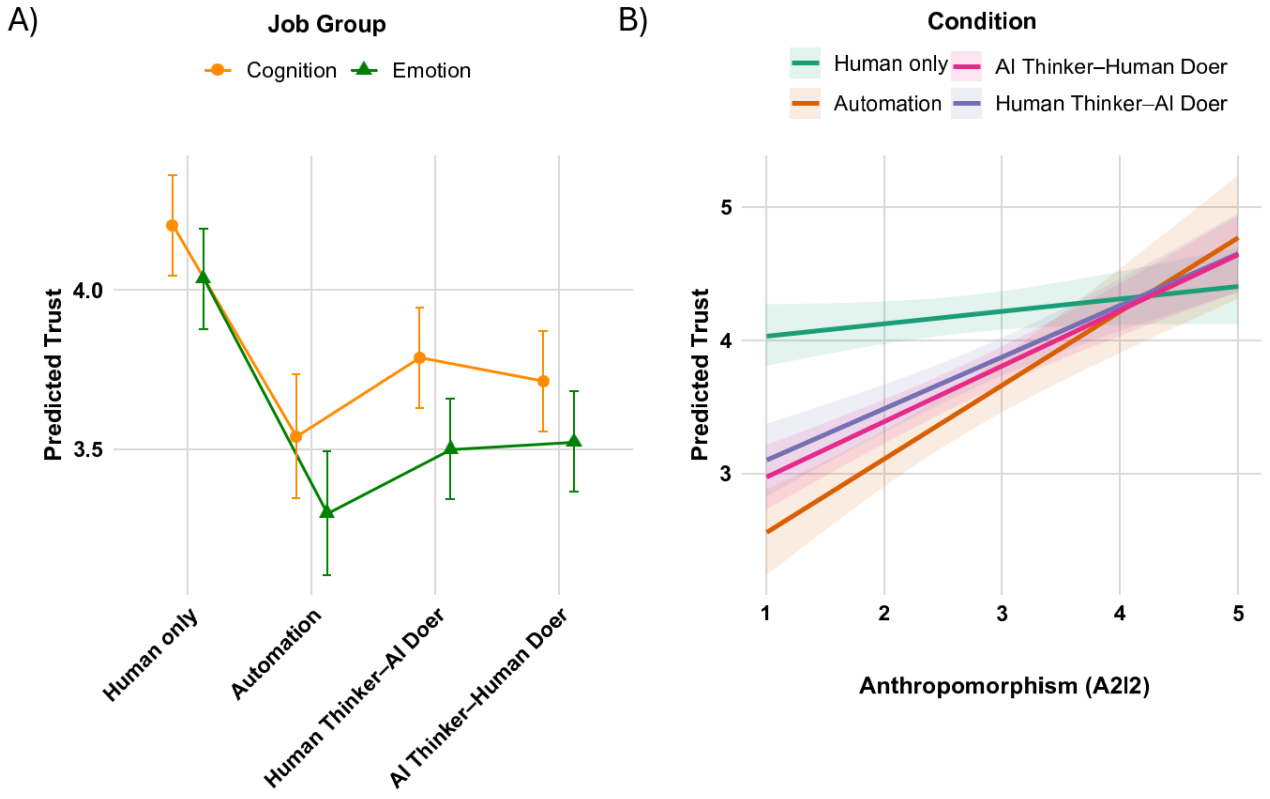
Multilevel Models Predicting Appreciation

	Model 1		**Model 2**		**Model 3**	
Predictors	*Estimates*	*95% CI*	*Estimates*	*95% CI*	*Estimates*	*95% CI*
(Intercept)	4.112 ***	3.877 – 4.347	3.448 ***	3.301 – 3.595	4.072 ***	3.858 – 4.287
AI Automation	-0.973 ***	-1.107 – -0.840	-0.941 ***	-1.065 – -0.816	-2.175 ***	-2.509 – -1.842
Collab: Human Thinker AI Doer	-0.736 ***	-0.855 – -0.616	-0.751 ***	-0.863 – -0.640	-1.428 ***	-1.732 – -1.123
Collab: AI Thinker human Doer	-0.699 ***	-0.818 – -0.579	-0.714 ***	-0.826 – -0.603	-1.614 ***	-1.912 – -1.315
JobGroup [Emotion]			-0.439 ***	-0.487 – -0.392	-0.309 ***	-0.400 – -0.218
A2I2			0.362 ***	0.316 – 0.408	0.074	-0.005 – 0.154
AI Automation × JobGroup [Emotion]					-0.210 **	-0.353 – -0.068

Collab: Human Thinker AI Doer × JobGroup [Emotion]		-0.318 ***	-0.445 – -0.190
Collab: AI Thinker human Doer × JobGroup [Emotion]		-0.014	-0.141 – 0.114
AI Automation × A2I2		0.586 ***	0.452 – 0.721
Collab: Human Thinker AI Doer × A2I2		0.347 ***	0.232 – 0.461
Collab: AI Thinker human Doer ×A2I2		0.376 ***	0.264 – 0.488
Random Effects			
ICC	0.49	0.50	0.48
N	1336	1336	1336
	4 Country	4 Country	4 Country
AIC	7131.728	6655.626	6573.093

** p < .05 ** p < .01 *** p < .001*

Figure S1
Trust Responses to AI Involvement in the Indian and Colombian Samples



Note. **Panel A shows model-estimated trust ratings across AI involvement conditions and job domains. Panel B shows the interaction between AI anthropomorphism and AI involvement condition in predicting trust. Trust was measured only in the Indian and Colombian samples; therefore, these analyses are reported as secondary subsample analyses. Error bars and shaded bands represent 95% confidence intervals.**

Table S4

Multilevel Models Predicting Trust

	Model 1		**Model 2**		**Model 3**	
Predictors	*Estimates*	*95% CI*	*Estimates*	*95% CI*	*Estimates*	*95% CI*
(Intercept)	4.118 ***	3.948 – 4.288	3.356 ***	3.150 – 3.562	3.943 ***	3.618 – 4.269
AI Automation	-0.781 ***	-0.981 – -0.581	-0.743 ***	-0.927 – -0.559	-1.891 ***	-2.453 – -1.329

Collab: Human Thinker AI Doer	-0.438 ***	-0.598 – -0.278	-0.485 ***	-0.633 – -0.337	-1.255 ***	-1.743 – -0.768
Collab: AI Thinker human Doer	-0.488 ***	-0.649 – -0.327	-0.506 ***	-0.654 – -0.357	-1.380 ***	-1.837 – -0.923
JobGroup [Emotion]			-0.218 ***	-0.284 – -0.151	-0.166 **	-0.290 – -0.041
A2I2			0.320 ***	0.257 – 0.383	0.094	-0.014 – 0.203
AI Automation × JobGroup [Emotion]					-0.075	-0.293 – 0.144
Collab: Human Thinker AI Doer × JobGroup [Emotion]					-0.118	-0.294 – 0.057
Collab: AI Thinker human Doer × JobGroup [Emotion]					-0.025	-0.202 – 0.151
AI Automation × A2I2					0.449 ***	0.248 – 0.650
Collab: Human Thinker AI Doer × A2I2					0.301 ***	0.138 – 0.463

Collab: AI Thinker human Doer ×A2I2		0.324 ***	0.170 – 0.478
Random Effects			
ICC	0.52	0.48	0.46
N	556	556	556
	2 Country	2 Country	2 Country
AIC	2624.753	2509.381	2510.725

** p < .05 ** p < .01 *** p < .001*

Analyses Controlling for Scenario Length

Scenario word counts ranged from 50 to 88 words across the 16 descriptions. To account for this variation, we conducted an additional robustness check in which scenario word count was included as a control variable in the regression models. The results remained unchanged after controlling for scenario length, and the main effects of AI involvement and job type remained robust. This suggests that the observed effects were not driven by differences in the amount of textual information presented in the scenarios.

Table S5

Robustness Checks with Scenario Word Counts

	Negative Rating		**Positive Rating**		**Appreciation Rating**		**Trust Rating**	
Predictors	*Estimate*	*95% CI*	*Estimate*	*95% CI*	*Estimate*	*95% CI*	*Estimate*	*95% CI*
(Intercept)	-8.858	-19.925 – 2.208	46.468 ***	33.914 – 59.022	4.206 ***	3.803 – 4.610	3.484 ***	2.881 – 4.087

AI Automation	44.502 ***	35.541 – 53.462	-26.323 ***	-35.144 – -17.503	-2.127 ***	-2.467 – -1.788	-1.998 ***	-2.568 – -1.428
Collab: Human Thinker AI Doer	30.552 ***	22.497 – 38.607	-12.933 **	-20.859 – -5.007	-1.408 ***	-1.711 – -1.104	-1.237 ***	-1.723 – -0.751
Collab: AI Thinker–Human Doer	22.380 ***	14.456 – 30.304	-16.321 ***	-24.120 – -8.523	-1.630 ***	-1.929 – -1.331	-1.401 ***	-1.858 – -0.945
JobGroup [Emotion]	-28.867 ***	-32.151 – -25.583	10.056 ***	6.786 – 13.327	-0.256 ***	-0.392 – -0.119	-0.005	-0.209 – 0.199
A2I2	6.253 ***	4.125 – 8.382	7.980 ***	5.886 – 10.075	0.072	-0.007 – 0.151	0.092	-0.016 – 0.201
Scenario Word Count	0.654 ***	0.564 – 0.744	-0.395 ***	-0.485 – -0.306	-0.003	-0.007 – 0.000	0.002	-0.003 – 0.008
Age	-0.127 **	-0.218 – -0.037	0.072	-0.016 – 0.161	0.001	-0.002 – 0.004	0.006 *	0.000 – 0.011
Gender [Male]	-2.528 *	-4.653 – -0.403	0.696	-1.394 – 2.786	-0.038	-0.118 – 0.041	0.012	-0.098 – 0.122
Political ideology	0.136	-0.330 – 0.602	0.880 ***	0.421 – 1.339	0.029 **	0.012 – 0.046	0.015	-0.011 – 0.040
AI Automation × JobGroup [Emotion]	-2.361	-5.656 – 0.934	-6.387 ***	-9.671 – -3.102	-0.227 **	-0.364 – -0.090	-0.090	-0.306 – 0.126
Collab: Human Thinker AI Doer × JobGroup [Emotion]	3.250	-0.175 – 6.674	-9.004 ***	-12.414 – -5.594	-0.299 ***	-0.441 – -0.156	-0.212 *	-0.416 – -0.009

Collab: AI Thinker human Doer × JobGroup [Emotion]	-3.846 *	-7.099 – -0.593	-0.666	-3.907 – 2.575	0.019	-0.116 – 0.155	-0.082	-0.276 – 0.112
AI Automation × A2I2	-11.907 ***	-15.480 – -8.334	9.628 ***	6.114 – 13.142	0.586 ***	0.452 – 0.720	0.462 ***	0.261 – 0.663
Collab: Human Thinker AI Doer ×A2I2	-9.434 ***	-12.478 – -6.389	6.095 ***	3.100 – 9.089	0.343 ***	0.229 – 0.457	0.296 ***	0.134 – 0.458
Collab: AI Thinker human Doer× A2I2	-7.254 ***	-10.235 – -4.273	5.913 ***	2.980 – 8.845	0.371 ***	0.259 – 0.483	0.326 ***	0.173 – 0.480

Random Effects

ICC	0.43	0.45	0.34	0.31
N	1336	1336	1336	556
	4 Country	4 Country	4 Country	2 Country
AIC	49506.807	49444.312	15106.465	5931.499

*$p < .05$ **$p < .01$ ***$p < .001$

Intercorrelations Among Dependent Variables

We ran Pearson correlation tests among the outcome measures. The outcomes were significantly correlated in expected directions, but the correlations varied in magnitude. In the full sample, positive affect was moderately positively associated with appreciation ($r = .58$, $p <$

.001), whereas negative affect showed small negative associations with both positive affect ($r = -.15$, $p < .001$) and appreciation ($r = -.25$, $p < .001$). Because trust was measured only in the Indian and Colombian samples, correlations involving trust were examined within that subset. Trust was strongly associated with appreciation ($r = .76$, $p < .001$) and moderately associated with positive affect ($r = .53$, $p < .001$), but only weakly negatively associated with negative affect ($r = -.21$, $p < .001$). Overall, these correlations indicate that the outcomes are related but not interchangeable. We therefore analyzed appreciation, positive affect, negative affect, and trust separately to retain the conceptual distinction between evaluative judgment, positive emotional response, negative emotional response, and trust.

Table S6

Intercorrelations Among Dependent Variables

Variable	Positive affect	Negative affect	Appreciation	Trust[b]
Positive affect	1.00	−.15[a]	.58[a]	.53[b]
Negative affect	−.15[a]	1.00	−.25[a]	−.21[b]
Appreciation	.58[a]	−.25[a]	1.00	.76[b]
Trust[b]	.53[b]	−.21[b]	.76[b]	1.00

Note. [a] Correlations are based on the full sample. [b] Correlations involving trust are based on the Indian and Colombian samples only, because trust was not measured in the U.S. and U.K.

Alternative Model Specifications

To evaluate the robustness of our primary models, we estimated alternative model specifications for each outcome and compared them to the main models reported in the manuscript. In the main analyses, the job scenario was treated as a fixed effect via the job group variable (emotion- vs. cognition-oriented). To test whether explicitly modeling job-level random effects changed the results, we added a random intercept for job scenario (all four) to the main model structure: (1 | Country/Participants) + (1 | Job). Fixed effect estimates for condition type, job group, and anthropomorphism were highly consistent with those from the main models across all four outcomes, suggesting the primary findings (Table S7). In addition, we should mention that random slopes were not included because condition type was assigned between participants, and job group varied across only two levels within each participant, making random slope estimation not appropriate for this design (Barr et al., 2013).

Table S7

Alternative Model

	Negative Rating		**Positive Rating**		**Appreciation Rating**		**Trust Rating**	
Predictors	*Estimates*	*CI*	*Estimates*	*CI*	*Estimates*	*CI*	*Estimates*	*CI*
(Intercept)	20.67	-11.10 – 23.01	26.09***	18.67 – 48.79	3.75***	3.34 – 4.16	3.67***	3.26 – 4.08
AI Automation	39.63 ***	19.08 – 55.16	-24.88 ***	-46.81 – -15.00	-2.17 ***	-2.35 – -1.21	-1.94 ***	-2.50 – -1.38
Collab: Human Thinker AI Doer	29.53 ***	15.20 – 46.57	-12.49 ***	-29.20 – -1.55	-1.21 ***	-1.71 – -0.72	-1.23 ***	-1.72 – -0.75
Collab: AI Thinker–Human Doer	26.13 **	7.68 – 36.98	-18.06 ***	-39.61 – -13.77	-1.51 ***	-1.98 – -1.05	-1.39 ***	-1.84 – -0.93

JobGroup [Emotion]	-4.57	-17.27 – 11.52	-1.10	-12.38 – 10.50	-0.15	-0.37 – 0.06	-0.17	-0.39 – 0.05
A2I2	6.255 ***	4.126 – 8.384	7.980 ***	5.886 – 10.075	0.072	-0.007 – 0.151	0.092	-0.016 – 0.201
Age	-0.127 **	-0.218 – -0.037	0.072	-0.016 – 0.161	0.001	-0.002 – 0.004	0.006 *	0.000 – 0.011
Gender [Male]	-2.533 *	-4.658 – -0.407	0.695	-1.395 – 2.785	-0.038	-0.118 – 0.041	0.012	-0.098 – 0.122
political ideology	1.40 **	0.57 – 2.24	1.08**	0.35 – 1.81	0.02	-0.01 – 0.04	0.01	-0.01 – 0.04
Automation× JobGroup [Emotion]	-0.124	-3.375 – 3.126	-7.585 ***	-10.854 – -4.316	-0.210 **	-0.346 – -0.073	-0.072	-0.287 – 0.143
Human Thinker–AI Doer × JobGroup [Emotion]	6.236 ***	3.316 – 9.157	-11.817 ***	-14.751 – -8.882	-0.320 ***	-0.442 – -0.197	-0.119	-0.292 – 0.053
AI Thinker–Human Doer × JobGroup [Emotion]	-0.184	-3.102 – 2.735	-3.552 *	-6.487 – -0.618	-0.014	-0.136 – 0.109	-0.023	-0.197 – 0.151
Automation × Composite Anthropomorphism	-11.909 ***	-15.482 – -8.335	9.628 ***	6.114 – 13.142	0.586 ***	0.452 – 0.720	0.462 ***	0.261 – 0.663
Human Thinker–AI Doer× Composite Anthropomorphism	-9.433 ***	-12.478 – -6.389	6.094 ***	3.100 – 9.089	0.343 ***	0.229 – 0.457	0.296 ***	0.133 – 0.458

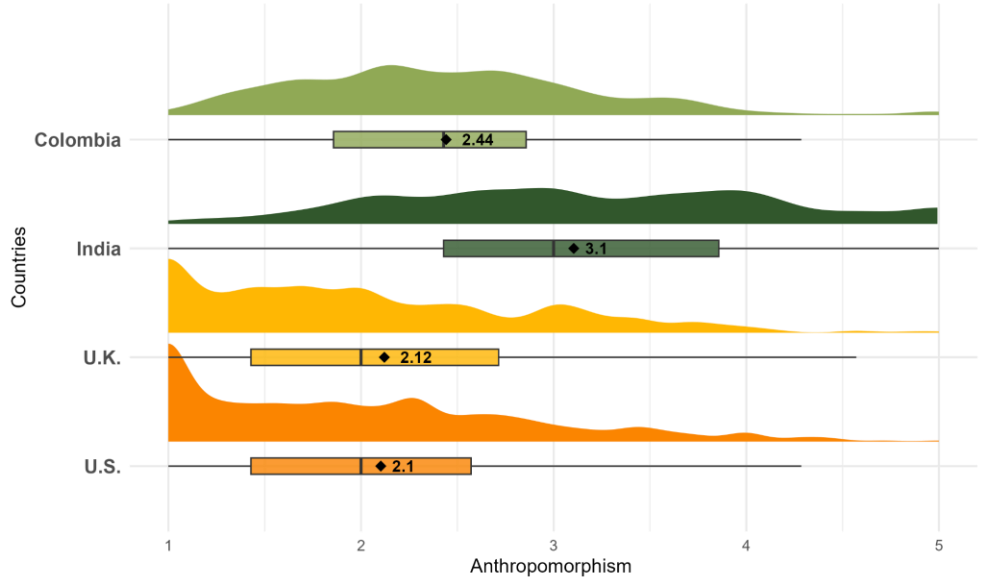
AI Thinker–Human Doer × Composite Anthropomorphism	-7.253 ***	-10.234 – -4.272	5.908 ***	2.976 – 8.841	0.371 ***	0.259 – 0.483	0.326 ***	0.173 – 0.480
Random Effects								
ICC	0.59		0.44		0.31		0.31	
N	1336		1336		1336		556	
	4 Country		4 Country		4 Country		2 Country	
	4 Job		4 Job		4 Job		4 Job	

*$p < .05$ **$p < .01$ ***$p < .001$

Anthropomorphism Across Countries

To examine cross-cultural differences in anthropomorphism, we conducted a Kruskal–Wallis test on A2I2 composite scores across the four countries (Figure S1). Results revealed a significant difference in anthropomorphism levels, $\chi^2(3, N = 1335) = 211.75$, $p < .001$. Follow-up pairwise comparisons using Wilcoxon rank-sum tests with Bonferroni correction indicated no significant difference between the U.S. and the U.K. ($p = 1.00$). On average, participants from India reported substantially higher anthropomorphism scores ($M = 3.10$) compared with participants from the U.S. ($M = 2.12$; $p < .001$) and the U.K. ($M = 2.12$; $p < .001$). Participants in Colombia ($M = 2.44$) differed significantly from all three other countries, lower than India and higher than the U.S. and the U.K. (ps $< .001$).

Figure S2



Note. The variation in A2I2 composite scores across four countries. A Kruskal–Wallis test revealed a significant difference in anthropomorphism levels, $\chi^2(3, N = 1335) = 211.75$, $p < .001$. Follow-up pairwise comparisons using Wilcoxon rank-sum tests with Bonferroni correction indicated participants from India reported substantially higher anthropomorphism scores across all other countries and no significant difference between the U.S. and the U.K. ($p = 1.00$).

Factor Structure and Reliability of the Anthropomorphizing AI Inventory Within Each Country

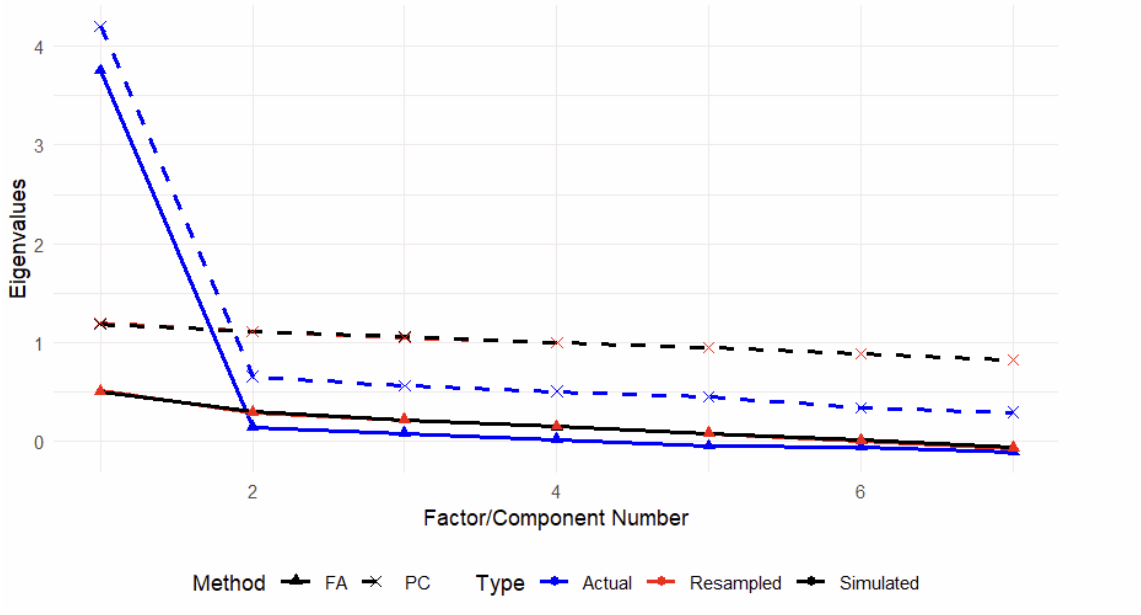
United States

The average anthropomorphism score in the U.S. was $M = 2.10$ ($SD = 0.86$), and Cronbach's α was 0.88 (95% CI [0.86, 0.90]). To evaluate the structure of the Anthropomorphizing AI Inventory (A2I2), we conducted an exploratory factor analysis (EFA) using the weighted least squares (WLS) estimation method, which is well-suited for analyzing Likert-scale (ordinal) data that may not be normally distributed (Bates et al., 2015). Parallel analysis suggested a one-factor solution (see Figure S2), indicating that all anthropomorphism

items loaded onto a single latent construct. Factor loadings ranged from .62 to .82, with communalities (h^2) between 0.39 and 0.67, demonstrating that the items contributed meaningfully to a unified construct. The single factor accounted for 52% of the variance, providing preliminary support for a unidimensional structure across participants (see Table S8).

Figure S3

Parallel Analysis Results in the U.S. Sample



Note. The solid blue line represents eigenvalues from the actual data (FA), while the dashed blue line represents principal components (PC). The red and black lines show eigenvalues from resampled and simulated data, respectively. In this case, only 1 factor should be retained, as it is the only one exceeding the reference lines.

Table S8

Items and Factor Analysis for the U.S.

Item	Factor Loading	Communality	Uniqueness
To what extent do you feel that AI is human-like?	0.82	0.68	0.32
To what extent do you feel that AI possesses its own personality?	0.80	0.64	0.35
To what extent do you feel that AI can express genuine emotions?	0.76	0.59	0.41
To what extent do you feel	0.69	0.48	0.51

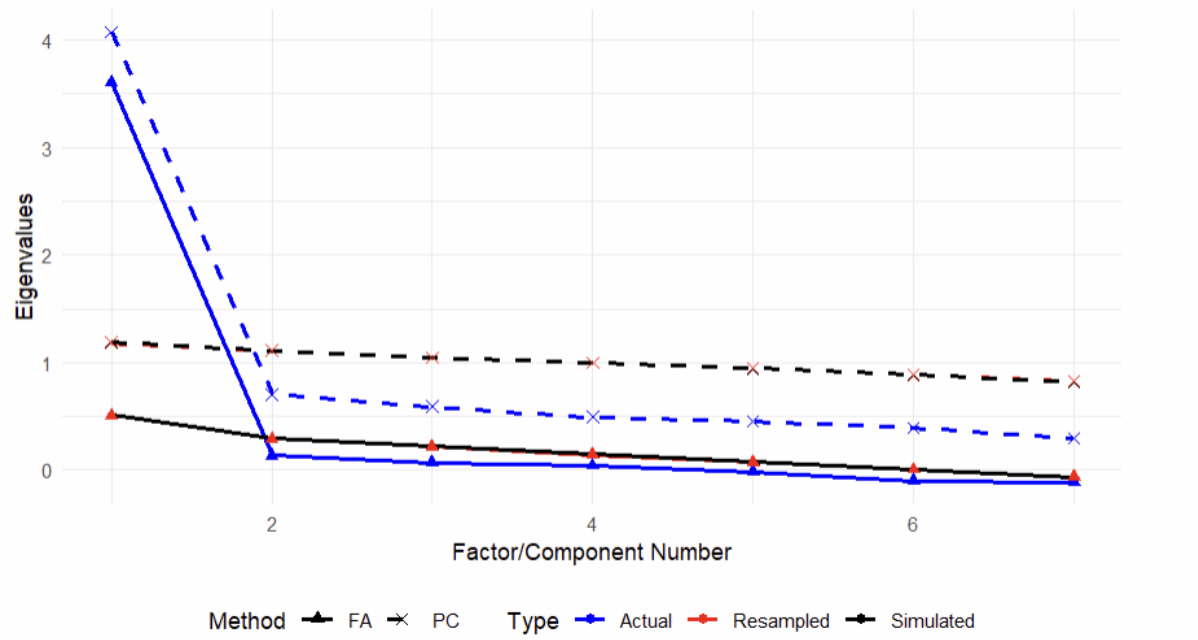
that AI considers ethical implications in its decision-making process?			
To what extent do you feel that AI is capable of creative thinking?	0.72	0.52	0.48
To what extent do you feel that AI has full control over its behavior?	0.62	0.38	0.61
To what extent do you feel that AI contributes to fairer outcomes in social issues?	0.65	0.41	0.58

United Kingdom

The average anthropomorphism score in the U.K. was $M = 2.12$ ($SD = 0.84$), and Cronbach's α for responses in this country was 0.87 (95% CI [0.86, 0.89]). The parallel analysis indicated that a one-factor solution was the best fit (see Figure S3), suggesting that all items measured a unified construct of AI anthropomorphism. The one-factor model accounted for 51% of the variance, with standardized item loadings ranging from 0.56 to 0.78, suggesting that all items contributed meaningfully to the latent construct (see Table S9).

Figure S4

Parallel Analysis for the U.K.



Note. The solid blue line represents eigenvalues from the actual data (FA), while the dashed blue line represents principal components (PC). The red and black lines show eigenvalues from resampled and simulated data, respectively. In this case, only 1 factor should be retained, as it is the only one exceeding the reference lines.

Table S9

Items and factor analysis for the U.K.

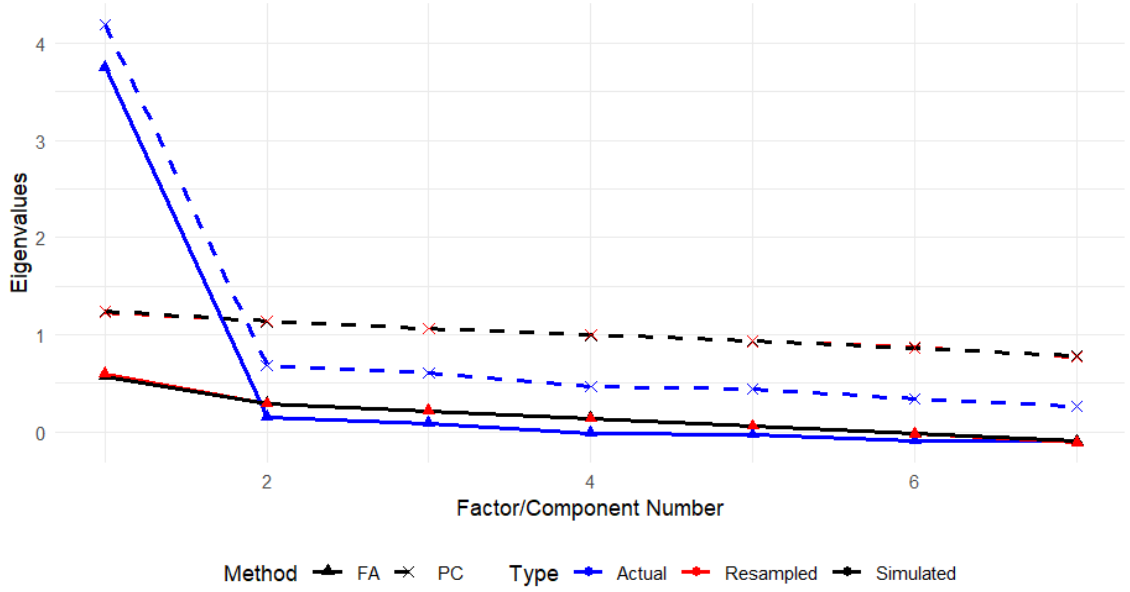
Item	Factor Loading (MR1)	Communality (h^2)	Uniqueness (u^2)
To what extent do you feel that AI is human-like?	0.78	0.60	0.39
To what extent do you feel that AI possesses its own personality?	0.78	0.61	0.39
To what extent do you feel that AI can express genuine emotions?	0.76	0.57	0.42
To what extent do you feel that AI considers ethical implications in its decision-making process?	0.72	0.52	0.48
To what extent do you feel that AI is capable of creative thinking?	0.75	0.56	0.44
To what extent do you feel that AI has full control over its behavior?	0.56	0.31	0.69
To what extent do you feel that AI contributes to fairer outcomes in social issues?	0.66	0.43	0.56

India

In the Indian sample, anthropomorphism scores ranged from 1 to 5 across all seven questionnaire items, with mean values ranging from 2.83 to 3.61. The median score was 3, indicating a moderate tendency to attribute human-like qualities to AI. A parallel analysis suggested a single-factor solution (see Figure S4), meaning that all anthropomorphism items loaded onto one latent construct. Factor loadings ranged from 0.57 to 0.8, and the single factor explained 52% of the variance, demonstrating a strong unidimensional structure (see Table S10). Also, the internal consistency of the scale was high, with Cronbach's α = 0.88 (95% CI [0.87, 0.89]), indicating strong reliability.

Figure S5

Parallel Analysis for India



Note. The solid blue line represents eigenvalues from the actual data (FA), while the dashed blue line represents principal components (PC). The red and black lines show eigenvalues from resampled and simulated data, respectively. In this case, only 1 factor should be retained, as it is the only one exceeding the reference lines.

Table S10

Items and Factor Analysis for India

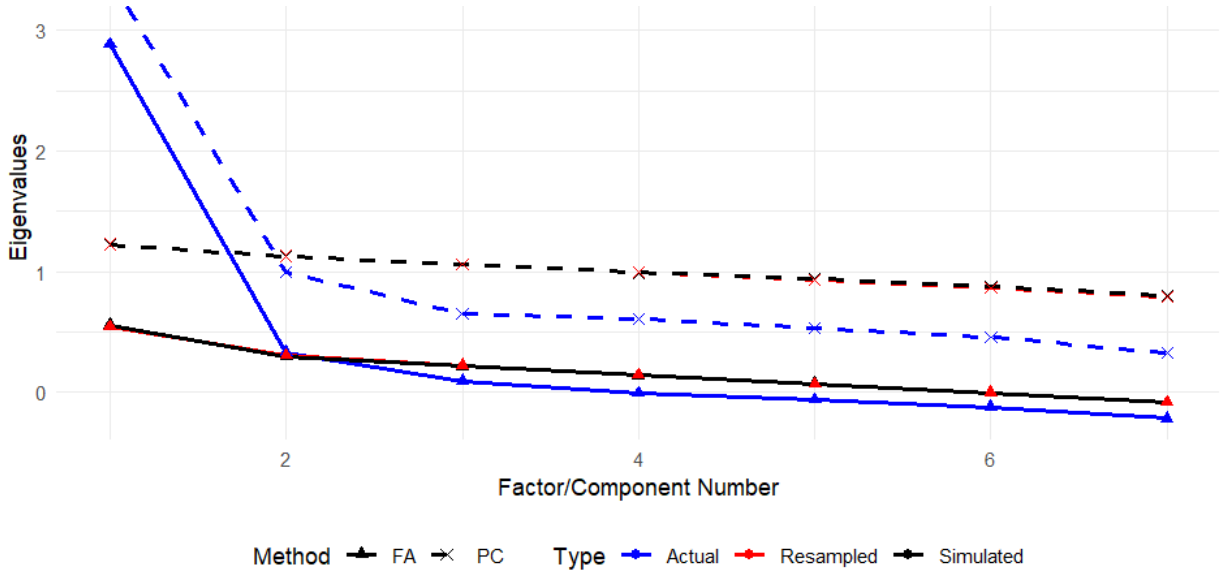
Item	Factor Loading (MR1)	Communality (h^2)	Uniqueness (u^2)
To what extent do you feel that AI is human-like?	0.79	0.62	0.38
To what extent do you feel that AI possesses its own personality?	0.79	0.63	0.37
To what extent do you feel that AI can express genuine emotions?	0.81	0.65	0.35
To what extent do you feel that AI considers ethical implications in its decision-making process?	0.73	0.54	0.46
To what extent do you feel that AI is capable of creative thinking?	0.57	0.33	0.67
To what extent do you feel that AI has full control over its behavior?	0.67	0.44	0.56
To what extent do you feel that AI contributes to fairer outcomes in social issues?	0.71	0.50	0.50

Colombia

The EFA analysis for the data from Colombia showed that a one-factor solution was the best fit (see Figure S5), suggesting that all items measured a unified construct of AI anthropomorphism. Item-total correlations ranged from 0.51 to 0.78, indicating that each item was meaningfully associated with the overall construct. The mean anthropomorphism score in the Colombian sample was M = 2.45 (SD = 0.77), which was slightly higher than in the U.K. and U.S. samples. The one-factor model accounted for 41% of the variance, with standardized item loadings ranging from 0.51 to 0.78. Communalities (h^2) ranged from 0.26 to 0.60, meaning that individual items had moderate levels of shared variance with the overall factor (see Table S11). Internal consistency analysis revealed good reliability, with Cronbach's α = 0.82 (95% CI [0.79, 0.85]).

Figure S6

Parallel Analysis for Colombia



Note. The solid blue line represents eigenvalues from the actual data (FA), while the dashed blue line represents principal components (PC). The red and black lines show eigenvalues from resampled and simulated data, respectively. In this case, only 1 factor should be retained, as it is

Table S11

Items and Factor Analysis for Colombia

Item	Factor Loading (MR1)	Communality (h^2)	Uniqueness (u^2)
To what extent do you feel that AI is human-like?	0.68	0.46	0.54
To what extent do you feel that AI possesses its own personality?	0.74	0.55	0.45
To what extent do you feel that AI can express genuine emotions?	0.78	0.60	0.40
To what extent do you feel that AI considers ethical implications in its decision-making process?	0.56	0.32	0.68
To what extent do you feel that AI is capable of creative thinking?	0.61	0.37	0.63
To what extent do you feel that AI has full control over its behavior?	0.58	0.34	0.66
To what extent do you feel that AI contributes to fairer outcomes in social issues?	0.51	0.26	0.74

the only one exceeding the reference lines.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 10.1016/j.jml.2012.11.001. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67, 1–48. https://doi.org/10.18637/jss.v067.i01