1. Correlations Between Demographic Subgroups and Interview Ratings (ChatGPT vs. Human), and Their Differences (Δ*r*)

Table 1: Correlations Between Demographic Subgroups and Interview Ratings (ChatGPT vs. Human), Including Differences with 95% Confidence Intervals and p-values.

Subgroup	ChatGPT *r*	ChatGPT 95% CI	Human *r*	Human 95% CI	Δ*r* (ChatGPT-Human)	Δr 95% CI	Δ*r p*
Gender: Female	-0.08	[-0.23, 0.07]	-0.10	[-0.25, 0.05]	0.03	[-0.11, 0.17]	0.656
Gender: Agender	0.07	[-0.08, 0.22]	0.12	[-0.03, 0.27]	-0.05	[-0.19, 0.1]	0.519
Gender: Genderfluid	0.07	[-0.08, 0.22]	0.10	[-0.05, 0.25]	-0.03	[-0.17, 0.11]	0.677
Ethnicity: Black	**-0.27**	**[-0.4, -0.13]**	**-0.02**	**[-0.17, 0.14]**	**-0.29**	**[-0.43, -0.16]**	**< .001*** **
Ethnicity: South Asian	0.19	[0.04, 0.33]	0.07	[-0.09, 0.22]	0.13	[-0.02, 0.26]	0.083
Ethnicity: Latin	0.07	[-0.08, 0.22]	0.07	[-0.09, 0.22]	0.00	[-0.14, 0.15]	0.962
Ethnicity: Southeast Asian	0.08	[-0.07, 0.23]	0.00	[-0.16, 0.15]	0.08	[-0.06, 0.23]	0.246
Ethnicity: West Asian	-0.07	[-0.21, 0.08]	-0.05	[-0.2, 0.1]	-0.02	[-0.16, 0.12]	0.765
Ethnicity: Arab	0.10	[-0.05, 0.25]	0.16	[0.01, 0.3]	-0.05	[-0.19, 0.09]	0.463
Ethnicity: Indigenous	-0.09	[-0.24, 0.06]	-0.07	[-0.22, 0.08]	-0.03	[-0.17, 0.11]	0.661

Subgroup	ChatGPT *r*	ChatGPT 95% CI	Human *r*	Human 95% CI	Δ*r* (ChatGPT-Human)	Δr 95% CI	Δ*r p*
Age: Under 25	-0.11	[-0.25, 0.04]	0.02	[-0.13, 0.17]	-0.14	[-0.28, 0]	0.056
Age: 35-44	-0.09	[-0.23, 0.06]	-0.13	[-0.28, 0.02]	0.05	[-0.09, 0.19]	0.501
Age: 45-54	**0.15**	**[0, 0.29]**	**-0.02**	**[-0.17, 0.14]**	**0.20**	**[0.05, 0.33]**	**0.007****
Age: 55 plus	0.09	[-0.06, 0.24]	-0.02	[-0.17, 0.14]	0.11	[-0.04, 0.25]	0.151

***Note**. Asterisks indicate statistical evidence thresholds based on p-values for each correlation.*

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

Correlations with asterisks are those for which the 95% confidence interval does not include zero.

Bolded rows reflect subgroups where the difference in correlation (Δr) between ChatGPT and human ratings was most pronounced and accompanied by statistically meaningful estimates, as judged by the confidence interval and effect size.