nature human behaviour

Supplementary information https://doi.org/10.1038/s41562-025-02247-w

Comparing the value of perceived human versus AI-generated empathy

In the format provided by the authors and unedited

Supplementary Information

Table of Contents

1. **<u>T-tests without filtering outliers for differences between conditions in studies 1–5</u>**

Study number	Variable	Degrees of freedom	t	CI for difference	*p*	Cohen's *d*	Significance
1a	Empathy (single item)	721.11	-2.74	[-.60, -.10]	.006	.20	**
	Positivity resonance	722.74	-5.57	[-12.99, -6.22]	<.001	.41	***
1b	Empathy (single item)	553.47	-4.79	[-.89, -.37]	<.001	.39	***
	Positivity resonance	583.39	-6.57	[-16.17, -8.73]	<.001	.54	***
	Authenticity	590.62	-3.00	[-.40, -.08]	.002	.25	**
	Support	No outliers					**
	Positive emotions	579.9	-5.2	[-1.17, -.53]	<.001	.43	***
	Negative emotions	516.1	5.21	[.26,.58]	<.001	.43	***
1c	Empathy (single item)	701.15	-4.41	[-.87, -.33]	<.001	.32	***
	Positivity resonance	725.24	-5.53	[-14.04, -6.68]	<.001	.41	***
	Authenticity	722.8	-4.51	[-1.17, -.46]	<.001	.33	***

	Support	No outliers					**
	Positive emotions	728.26	-4.88	[-1.09, -047]	<.001	.36	***
	Negative emotions	731.51	1.69	[-.03,.39]	.09	.12	.
1d	Empathy (single item)	702.43	-1.71	[-.49, .03]	.09	.13	.
	Positivity resonance	700.3	-2.14	[-7.52, -.33]	.03	.16	*
	Authenticity	No outliers			NS		
	Support	No outliers					**
	Positive emotions	701.7	-.97	[-.47, .16]	.33	.07	
	Negative emotions	NS			NS		
2a	Empathy (single item)	No outliers					**
	Positivity resonance	No outliers					**
	Authenticity	No outliers					**
	Support	No outliers					**
	Positive emotions	No outliers					**

	Negative emotions	640.63	4.14	[.26, .73]	<.001	.31	***
2b	Empathy (single item)	624.7	-6.84	[-1.25, -.69]	<.001	.52	***
	Positivity resonance	661.89	-7.79	[-19.08, -11.40]	<.001	.59	***
	Authenticity	661.03	-6.92	[-1.66, -.93]	<.001	.53	***
	Support	No outliers					**
	Positive emotions	675.36	-5.71	[-1.33, -.65]	<.001	.43	***
	Negative emotions	557.54	5.85	[.37, .75]	<.001	.44	***
3	Study 3 did not use t-tests due to the interactions between condition and response type						
4	Curiosity	NS			NS		
	Timing	No outliers					
	Understanding	No outliers					
	Hesitation	No outliers					
	Sharing	No outliers					
	Care	No outliers					
	Loneliness	No outliers					
5	Curiosity	324.12	1.17	[-.15, .60]	.24	.11	NS
	Timing	No outliers					

	Understanding	No outliers
	Hesitation	No outliers
	Sharing	No outliers
	Care	No outliers
	Loneliness	No outliers

Supplementary table 1: Detailed results for all two-tailed t-tests in our analyses, without filtering for outliers, used to compare the differences between experimental conditions. *** = $p < .001$, ** = $p < .01$, * = $p < .05$, . = $p < .10$.

2. Wilcoxon tests for differences between conditions in studies 1–5

Due to our data being skewed, we report here the results of Wilcoxon rank-sum tests, to compare their findings to t-tests between conditions.

Study number	Variable	W	*p*	Significance
1a	Empathy (single item)	48950	<.001	***
	Positivity resonance	44221	<.001	***
1b	Empathy (single item)	32686	<.001	***
	Positivity resonance	27227	<.001	***
	Authenticity	34109	<.001	***
	Support	37556	.002	**
	Positive emotions	30908	<.001	***
	Negative emotions	49831	<.001	***
1c	Empathy (single item)	52986	<.001	***

	Positivity resonance	47862	<.001	***
	Authenticity	49288	<.001	***
	Support	57467	<.001	***
	Positive emotions	47334	<.001	***
	Negative emotions	67584	.007	**
1d	Empathy (single item)	52814	.06	.
	Positivity resonance	54227	.06	.
	Authenticity	NS	NS	
	Support	66759	.09	.
	Positive emotions	53198	.06	.
	Negative emotions	NS	NS	
2a	Empathy (single item)	39508	<.001	***
	Positivity resonance	38230	<.001	***
	Authenticity	39954	<.001	***
	Support	45124	<.001	***
	Positive emotions	41978	<.001	***
	Negative emotions	60442	.002	**
2b	Empathy (single item)	40568	<.001	***
	Positivity resonance	38006	<.001	***
	Authenticity	41217	<.001	***
	Support	46438	<.001	***
	Positive emotions	44419	<.001	***
	Negative emotions	67530	<.001	***

3	Study 3 did not use t-tests due to the interactions between condition and response type			
4	Curiosity	NS	NS	
	Timing	48509	<.001	***
	Understanding	1381	<.001	***
	Hesitation	22432	<.001	***
	Sharing	1242	<.001	***
	Care	2297.5	<.001	***
	Loneliness	5544	<.001	***
5	Curiosity	27659	.07	.
	Timing	47731	<.001	***
	Understanding	1467.5	<.001	***
	Hesitation	19437	<.001	***
	Sharing	1518	<.001	***
	Care	3808	<.001	***
	Loneliness	7492	<.001	***

Supplementary table 2: Detailed results for all Wilcoxon tests in our analyses, used to replace t-tests without assuming normality, to show the differences between experimental conditions. *** = $p < .001$, ** = $p < .01$, * = $p < .05$, . = $p < .10$.

3. Detailed analyses for study 1a

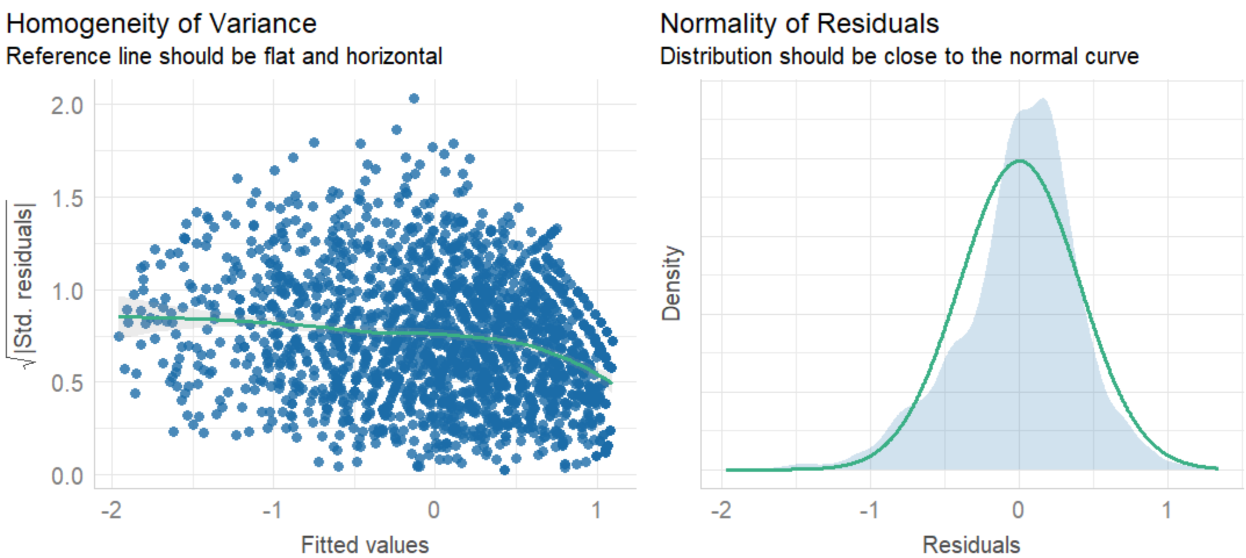
a. Repeated-measures models with multiple empathy types

i. Reporting tests without outliers

We fitted a linear mixed-effect model to predict empathy with condition, aspect of empathy (cognitive, affective or motivational) and their interaction, to the full dataset. The model showed a significant main effect of condition ($F(1, 723) = 21.88$, $p < .001$, partial $\eta^2 = .03$). This reinforces our central results of a significant difference between the two conditions across all participants.

ii. <u>Permutation analyses without assuming normality</u>

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not fit perfectly (Shapiro-Wilk test: $W = .98$, $p < .001$) but are not extremely skewed.



To assess whether our results hold without assuming normality, we ran a permutation test, shuffling the experimental conditions between participants and the empathy scores between types within each participant 10,000 times, and each time we analyzed a model similar to that described in the paper to assess the differences between conditions, empathy types and their interactions. This allowed us to create an F distribution using our data, without any effect of condition and without assuming normality. The test showed that for the main effect of condition, our original results hold true (original $F(1, 694.36) = 29.7440$, permutation $p < .001$).

4. Detailed analyses for study 1b

a. Initial analyses to comply with pre-registration

To comply with the pre-registration, we report here the t-tests performed to analyze differences in ratings of different aspects of empathy between conditions. Similar to our reported analyses, these tests found significant differences between conditions in overall empathy ($t(552.95) = -4.36$, $p < .001$, Cohen's $d = .36$, 95% CI for difference = [-.65, -.25]) and all aspects of empathy (cognitive: $t(571.17) = -5.28$, $p < .001$, Cohen's $d = .44$, 95% CI = [-.78, -.36]; affective: $t(569.9) = -7.07$, $p < .001$, Cohen's $d = .59$, 95% CI = [-1.25, -.071]; motivational: $t(575.02) = -6.22$, $p < .001$, Cohen's $d = .52$, 95% CI = [-.92, -.48]).

b. Interaction effects of assumed aid by the other source and condition

In addition to the reported effects of assumed aid by the other source on perceived empathy, we found many other interaction effects. First, we wanted to explore whether this interaction with empathy was driven by a specific aspect of empathy, so we used linear mixed-effects regression models, adding condition, empathy type, assumed aid by the other source and their interactions into the model. Following the observation of significant interactions, post-hoc contrasts revealed that within the human condition, assumed aid negatively affected perceived empathy on all aspects (cognitive: $t(927) = -6.23$, SE = .05, $p < .001$, $\beta = -.34$, 95% CI = [-.45, -.23]; affective: $t(924) = -7.37$, SE = .05, $p < .001$, $\beta = -.40$, 95% CI = [-.51, -.29]; motivational: $t(924) = -5.92$, SE = .05, $p < .001$, $\beta = -.32$, 95% CI = [-.43, -.21]). In the AI condition, the assumed aid positively affected affective empathy ($t(928) = 2.51$, SE = .07, $p = .01$, $\beta = .17$, 95% CI = [.04, .31]) and had a marginally positive effect on motivational empathy (t(936) = 1.73, SE = .07, p = .08, β = .12, 95% CI = [-.02, .26]) but no significant effect on cognitive empathy (p = .25). Further contrasts did not find significant differences between the effects on different aspects within each condition.

We additionally found a significant interaction between condition and assumed aid by the other source on positivity resonance (F(1, 568) = 29.68, p < .001, partial η^2 = .05), with post-hoc contrasts showing that assumed aid had a positive effect on the AI condition, increasing positivity resonance (t(568) = 2.33, SE = .07, p = .02, β = .16, 95% CI = [.02, .29]), but a negative effect on the human condition, decreasing positivity resonance (t(568) = -5.87, SE = .05, p < .001, β = -.31, 95% CI = [-.42, -.21]).

Similar interactions were found for positive emotions, with assumed aid having a positive effect in the AI condition (t(589) = 2.26, SE = .07, p = .02, β = .15, 95% CI = [.02, .28]) and a negative effect in the human condition (t(569) = -6.27, SE = .05, p < .001, β = -.33, 95% CI = [-.43, -.23]).

Authenticity was influenced similarly, with assumed aid having a positive effect in the AI condition (t(575) = 3.08, SE = .07, p = .002, β = .21, 95% CI = [.08, .35]) and a negative effect in the human condition (t(575) = -9.62, SE = .05, p < .001, β = -.53, 95% CI = [-.63, -.42]).

The same was true of support perceived in the response, positively influenced in the AI condition (t(589) = 3.29, SE = .08, p = .001, β = .26, 95% CI = [.10, .41]) but negatively in the human condition (t(589) = -6.90, SE = .06, p < .001, β = -.43, 95% CI = [-.55, -.31]).

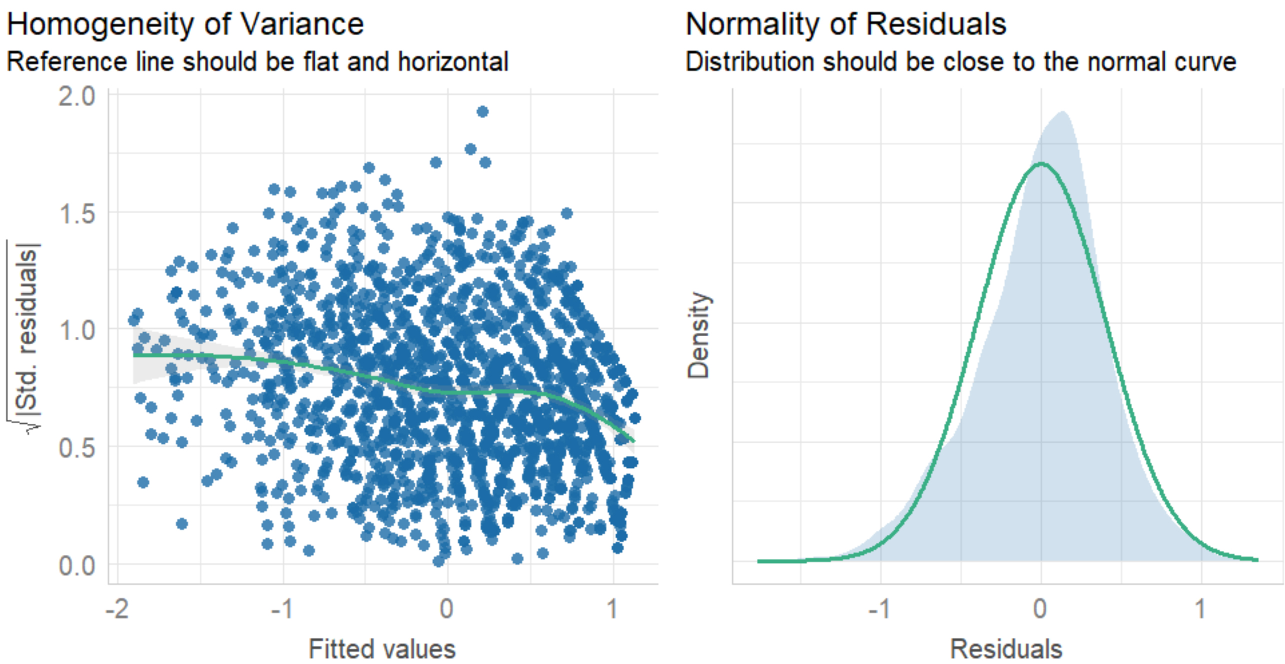
c. <u>Extended results for repeated-measures models with multiple empathy types</u>

i. <u>Reporting tests without outliers</u>

We fitted a linear mixed-effect model to predict empathy with condition, aspect of empathy (cognitive, affective or motivational) and their interaction, to the full dataset. The model showed a significant main effect of condition (F(1, 591) = 43.88, $p < .001$, partial $\eta^2 =$ .07). This reinforces our central results of a significant difference between the two conditions across all participants. The interaction between condition and empathy type was not significant in this analysis ($p = .20$); however, the effect size for cognitive empathy was still smaller than the effect sizes for affective and motivational empathy.

ii. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .99, $p < .001$) but are not extremely skewed.



To assess whether our results hold without assuming normality, we ran a permutation test, shuffling the experimental conditions between participants and the empathy scores between types within each participant 10,000 times, and each time we analyzed a model similar to that described in the paper to assess the differences between conditions, empathy types and their interactions. This allowed us to create an F distribution using our data, without any effect of condition and without assuming normality. The test showed that for the main effect of condition our original results hold true (original $F(1, 567.53) = 51.307$, permutation $p < .001$). The same was true for the interaction between condition and empathy type (original $F(2, 1126.97) = 4.09$, permutation $p = .02$). Contrasts revealed that the specific differences between cognitive and affective empathy were still significant even when using permutations (original $t(1147.1) = 2.82$, permutation $p = .002$). Moreover, the permutation test showed that the difference between cognitive and motivational empathy was also significant (original $t(1147.1) = 1.83$, permutation $p = .03$), where the standard contrast did not show it to be significant.

These results reinforce our central conclusion of a difference between the experimental conditions, and specifically the affective and motivational aspects of empathy, even without assuming normality.

d. <u>Extended results on assumed aid by the other source</u>

We provide here further analyses without filtering outliers, and separate analyses using permutations without assuming normality. These are parallel to the multiple-regression models presented in the paper and earlier in the supplementary (section 4b), providing evidence for the interactions between condition and assumed aid.

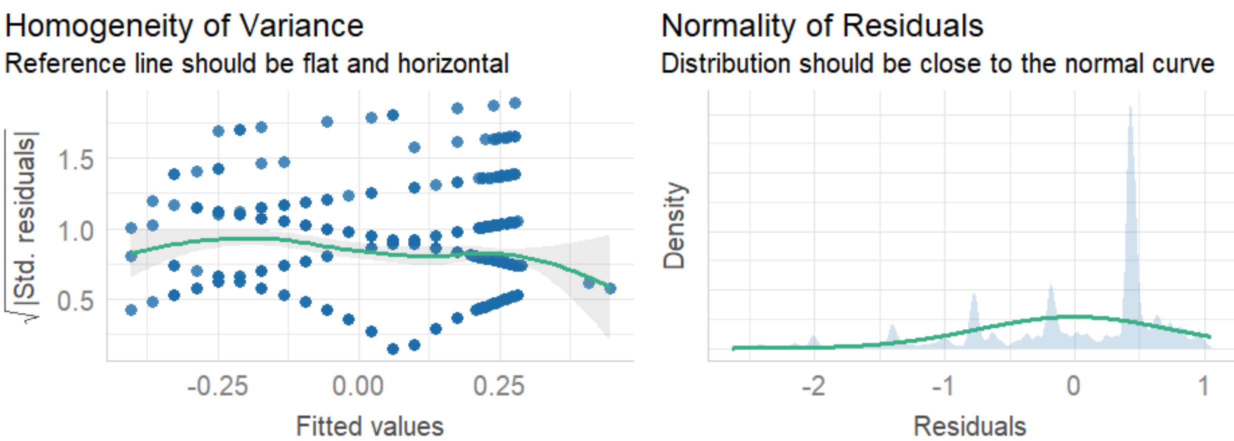
i. <u>General empathy item</u>

a. Model without filtering outliers

We fitted a linear model predicting the general empathy rating from condition, assumed aid by the other source and their interaction. Both the effect of condition ($F(1, 589) = 24.02$, $p < .001$, partial $\eta^2 = .04$) and the interaction effect remained significant ($F(1, 589) = 20.16$, $p < .001$, partial $\eta^2 = .03$). Contrasts showed that the negative effect of assumed aid in the human condition remained significant (t(589) = -5.29, SE = .06, $p < .001$, β = -.33, 95% CI = [-.45, -.21]).

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing the level of skewness of our residuals from a normal distribution (Shapiro-Wilk test: W = .86, $p < .001$).



We ran a permutation test identical to the one in study 1b, using 10,000 permutations. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 571) = 19.62, permutation *p*-value < .001), and the interaction between condition and assumed aid by the other source also remained significant (F(1, 571) = 6.14, permutation *p*-value = .02). Contrasts revealed that the positive effect of assumed aid in the AI condition remained significant (t(571) = 4.02, permutation $p < .001$).

ii. <u>Mixed model for multiple empathy types</u>

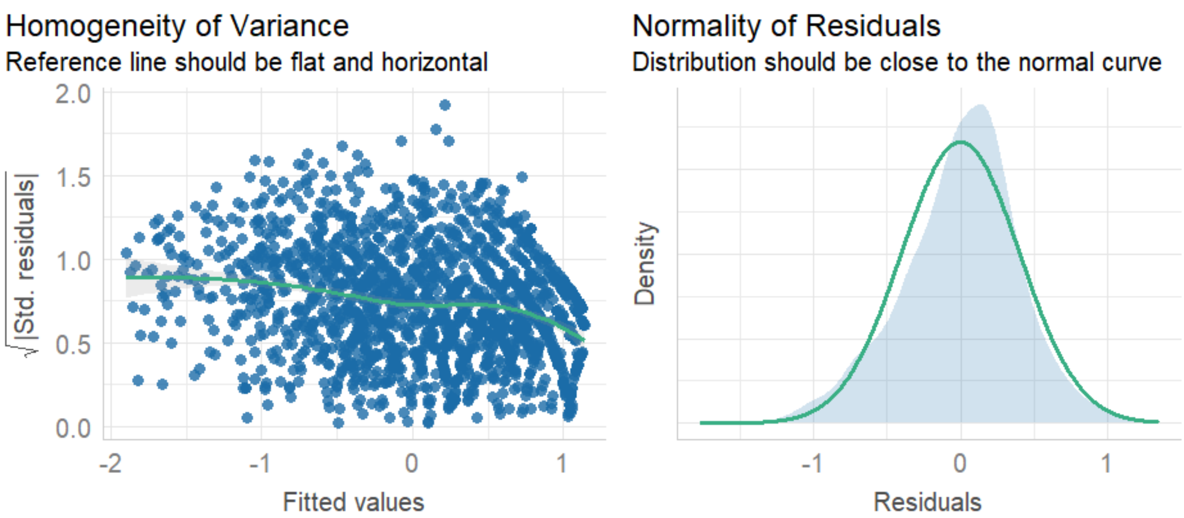
a. Model without filtering outliers

We fitted a linear mixed-effect model to predict empathy with condition, aspect of empathy (cognitive, affective or motivational), assumed aid by the other source and their interactions, to the full dataset. The model showed a significant main effect of condition (F(1,

589) = 49.42, $p < .001$, partial $\eta^2 = .08$). This reinforces our central results of a significant difference between the two conditions across all participants. The interaction between condition and empathy type was not significant in this analysis ($p = .30$); however, the interaction between condition and assumed aid remained significant (F(1, 589) = 40.87, $p < .001$, partial $\eta^2 = .06$). Contrasts revealed a significant positive effect of assumed aid in the AI condition (, t(589) = 2.48, SE = .07, $p = .01$, $\beta = .17$, 95% CI = [.04, .30]), with a much larger negative effect in the human condition (t(589) = -7.18, SE = .06, $p < .001$, $\beta = -.39$, 95% CI = [-.50, -.29]).

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .99, $p < .001$) but are not extremely skewed.



We ran a permutation test identical to the one in study 1b, using 10,000 permutations. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 565.09) = 50.15, permutation *p*-value < .001), and the interaction between condition and assumed aid by the other source also remained significant

(F(1, 565.35) = 3.98, permutation *p*-value = .02). Contrasts revealed that the positive effect of assumed aid in the AI condition remained significant (t(587) = 3.01, permutation p = .002).

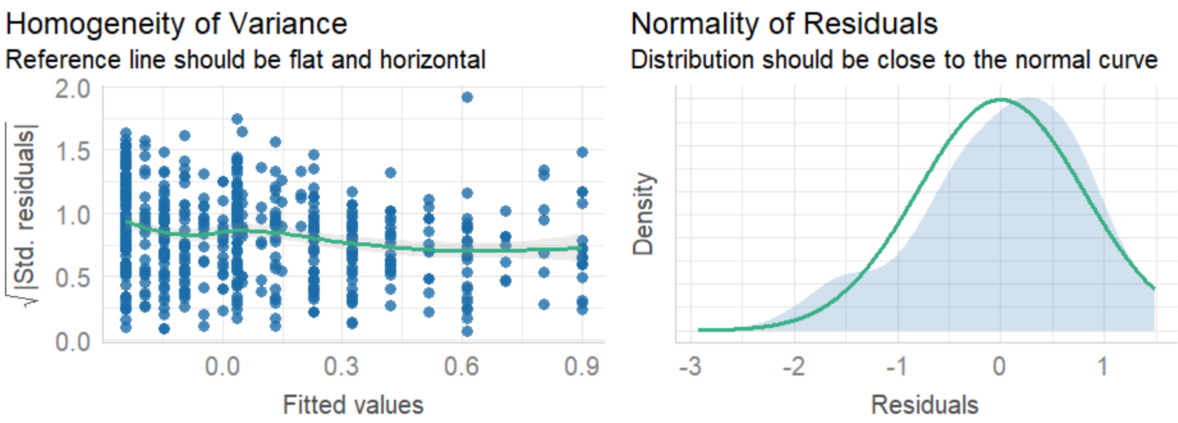
iii. Positivity resonance

a. Model without filtering outliers

We fit a model predicting positivity resonance from the condition, assumed aid from the other source and their interaction, using the entire dataset. We replicated the significant effect of condition (F(1, 589) = 46.30, $p < .001$, partial $\eta^2 = .07$) and the significant interaction between condition and assumed aid (F(1, 589) = 30.12, $p < .001$, partial $\eta^2 = .05$). Contrasts replicated the significant negative effect of assumed aid on positivity resonance in the human condition (t(589) = -6.34, SE = .06, $p < .001$, $\beta = -.39$, 95% CI = [-.51, -.27]) and the significant positive effect in the AI condition (t(589) = 1.99, SE = .08, $p = .05$, $\beta = .15$, 95% CI = [.002, .31]).

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .97, $p < .001$) but are not extremely skewed.



We ran a permutation test identical to the one earlier, with positivity resonance as the dependent variable, using 10,000 permutations. The test showed that even using a permutation-based distribution, the effect of condition remained significant ($F(1, 568) = 52.98$, permutation p-value $< .001$), and the interaction between condition and assumed aid by the other source also remained significant ($F(1, 568) = 29.68$, permutation p-value $< .001$). Contrasts revealed that both the positive effect of assumed aid in the AI condition on positivity resonance ($t(571) = 2.36$, permutation $p = .01$) and its negative effect in the human condition ($t(571) = -5.87$, permutation $p < .001$) remained significant.

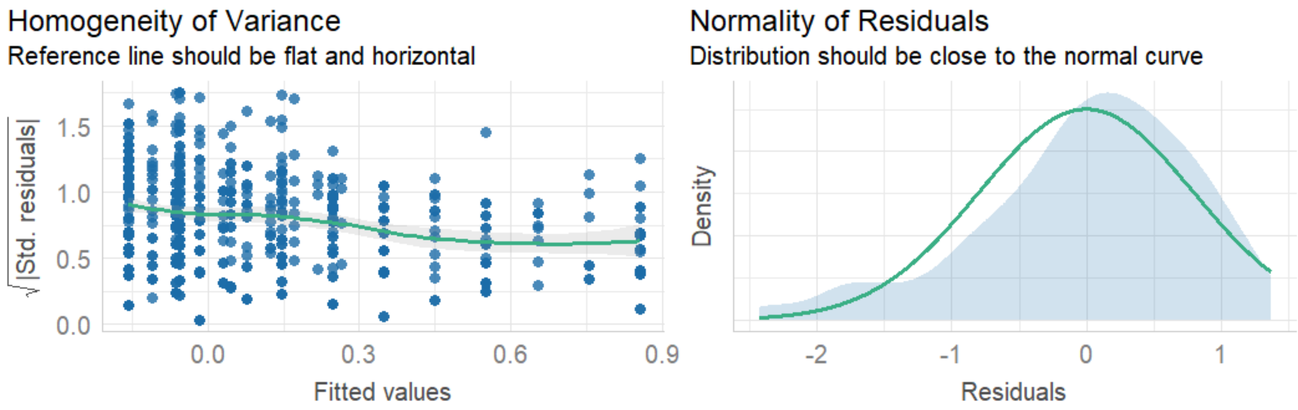
iv. Positive emotions

a. Model without filtering outliers

We fit a model predicting positive emotions from the condition, assumed aid from the other source and their interaction, using the entire dataset. We replicated the significant effect of condition ($F(1, 589) = 29.03$, $p < .001$, partial $\eta^2 = .05$) and the significant interaction between condition and assumed aid ($F(1, 589) = 37.84$, $p < .001$, partial $\eta^2 = .06$). Contrasts replicated the significant negative effect in the human condition ($t(589) = -6.05$, $SE = .06$, $p < .001$, $\beta = -.37$, 95% CI = [-.49, -.25]) and the significant positive effect in the AI condition ($t(589) = 3.06$, $SE = .08$, $p = .002$, $\beta = .24$, 95% CI = [.09, .39]).

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: $W = .97$, $p < .001$) but are not extremely skewed.



We ran a permutation test identical to the one earlier, with positive emotions as the dependent variable, using 10,000 permutations. The test showed that even using a permutation-based distribution, the effect of condition remained significant ($F(1, 569) = 23.85$, permutation *p*-value < .001), and the interaction between condition and assumed aid by the other source also remained significant ($F(1, 569) = 31.83$, permutation *p*-value < .001). Contrasts revealed that both the positive effect of assumed aid in the AI condition ($t(569) = 2.26$, permutation $p = .01$) and the negative effect in the human condition ($t(569) = -6.27$, permutation $p < .001$) remained significant in their impact on positive emotions.

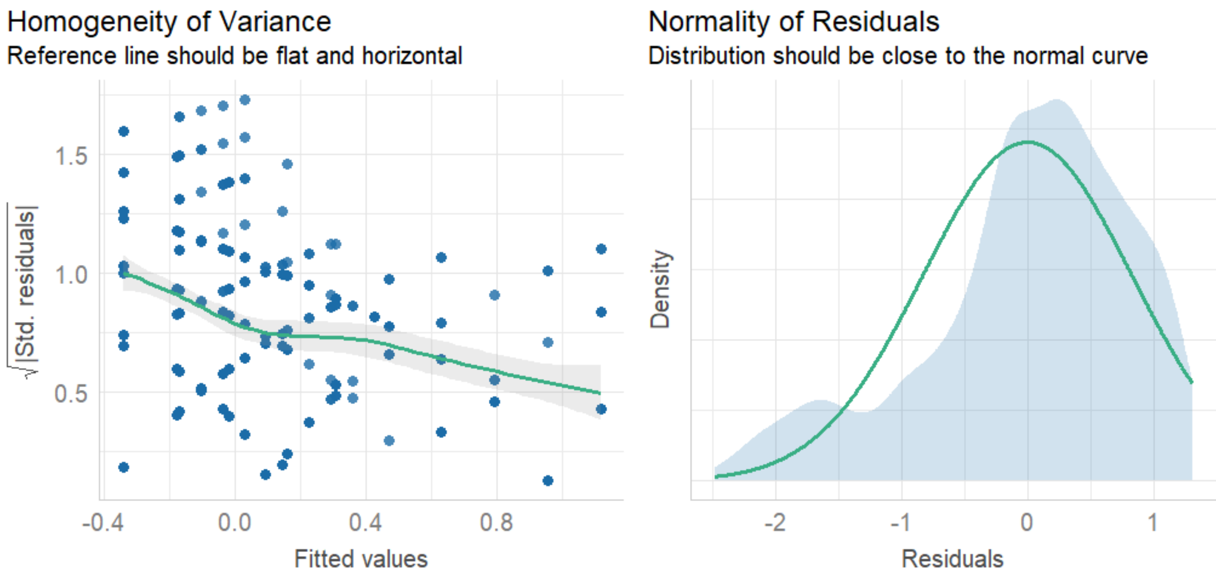
v. <u>Authenticity</u>

a. Model without filtering outliers

We fit a model predicting authenticity from the condition, assumed aid from the other source and their interaction, using the entire dataset. We replicated the significant effect of condition ($F(1, 589) = 10.50$, $p = .001$, partial $\eta^2 = .02$) and the significant interaction between condition and assumed aid ($F(1, 589) = 69.77$, $p < .001$, partial $\eta^2 = .11$). Contrasts replicated the significant negative effect on authenticity in the human condition ($t(589) = -9.59$, $SE = .06$, $p < .001$, $\beta = -.57$, 95% CI = [-.69, -.46]) and the significant positive effect in the AI condition ($t(589) = 3.07$, $SE = .08$, $p = .002$, $\beta = .23$, 95% CI = [.08, .38]).

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W =.94, $p < .001$) but are not extremely skewed.

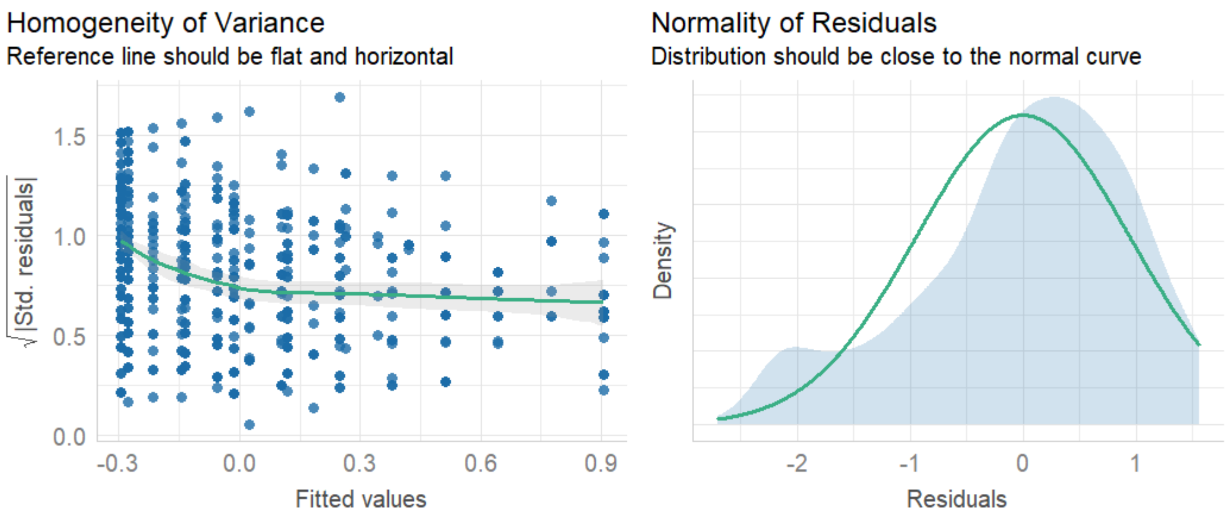


We ran a permutation test identical to the one earlier, with authenticity as the dependent variable, using 10,000 permutations. The test showed that even using a permutation-based distribution, the effect of condition remained significant ($F(1, 575) = 9.25$, permutation p-value = .003), and the interaction between condition and assumed aid by the other source also remained significant ($F(1, 575) = 70.03$, permutation $p < .001$). Contrasts revealed that both the positive effect of assumed aid in the AI condition ($t(575) = 3.08$, permutation $p = .001$) and the negative effect in the human condition ($t(575) = -9.62$, permutation $p < .001$) remained significant in their impact on authenticity.

vi. <u>Support</u>

a. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W =.96, $p < .001$) but are not extremely skewed.



We ran a permutation test identical to the one earlier, with support as the dependent variable, using 10,000 permutations. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 589) = 11.90, permutation *p*-value < .001), and the interaction between condition and assumed aid by the other source also remained significant (F(1, 589) = 46.99, permutation *p*-value < .001). Contrasts revealed that the positive effect of assumed aid on perceived support in the AI condition (t(589) = 3.29, permutation $p = .001$) and the negative effect in the human condition (t(589) = -6.90, permutation $p < .001$) both remained significant.

5. Detailed analyses for study 1c

This study aimed to generalize our findings using a different language model, Llama 3.1 405B instead of GPT models as in the other studies.

a. Main analyses for empathy

We repeated the same analyses for this study as we did for the previous studies. We found a significant difference in the single-item empathy question between conditions (Welch t-test: $t(696.31) = -4.00$, $p < .001$, Cohen's $d = .30$, 95% CI [-.66,-.22]).

As in the previous studies, a linear mixed-effect model predicting empathy by the experimental condition, empathy type and the interaction between them found a significant main effect of condition ($F(1, 709.86) = 16.81$, $p < .001$, partial $\eta^2 = .02$), with participants in the human condition rating responses as more empathic than those in the AI condition. The model also showed a significant interaction between condition and empathy type ($F(2, 1408.87) = 8.92$, $p < .001$, partial $\eta^2 = .01$), showing that the effect of condition on perceived affective empathy was significantly larger than the effect on perceived cognitive empathy ($t(1424) = 3.94$, $SE = .05$, $p < .001$, $\beta = .18$, 95% CI = [.07, .30]) and the effect on perceived motivational empathy ($t(1424) = -3.29$, $SE = .05$, $p = .003$, $\beta = -.15$, 95% CI = [-.27, -.04]). As in study 1b, the greatest difference in perceived empathy was specifically for affective empathy.

b. Additional effects of condition

Condition had a significant effect on positivity resonance (Welch t-test: $t(698.29) = -4.73$, $p < .001$, Cohen's $d = .36$, 95% CI [-10.86,-4.49]). This replicates our earlier findings that participants experienced greater positivity in the communication when perceiving the response as human.

We also found a significant difference between conditions in the level of authenticity participants attributed to the response (Welch t-test: $t(676.89) = -4.52$, $p < .001$, Cohen's $d = .34$, 95% CI [-.39, -.15]). As in the previous studies, participants found the response more authentic when they perceived it as human.

There was a significant difference in perceived support between conditions (t(726.07) = -3.65, $p < .001$, Cohen's $d = .27$, 95% CI [-1.06, -.32]). This replicated the previous findings, suggesting that participants perceived the response as more supportive when they believed it was human-authored.

Responses that were presented as human also raised significantly more positive emotions than those presented as AI-generated (t(693.75) = -4.82, $p < .001$, Cohen's $d = .37$, 95% CI [-.87, -.37]). Like previously, participants felt more positive emotions when perceiving the response as human.

There was also a significant difference in the levels of negative emotions caused by the response (t(694.04) = 2.42, $p = .02$, Cohen's $d = .18$, 95% CI [.03, .30]). As in the previous studies, participants felt fewer lower levels of negative emotions following the response when perceiving it as human-authored.

c. <u>**Conditional effects without outliers or assumptions of normality**</u>

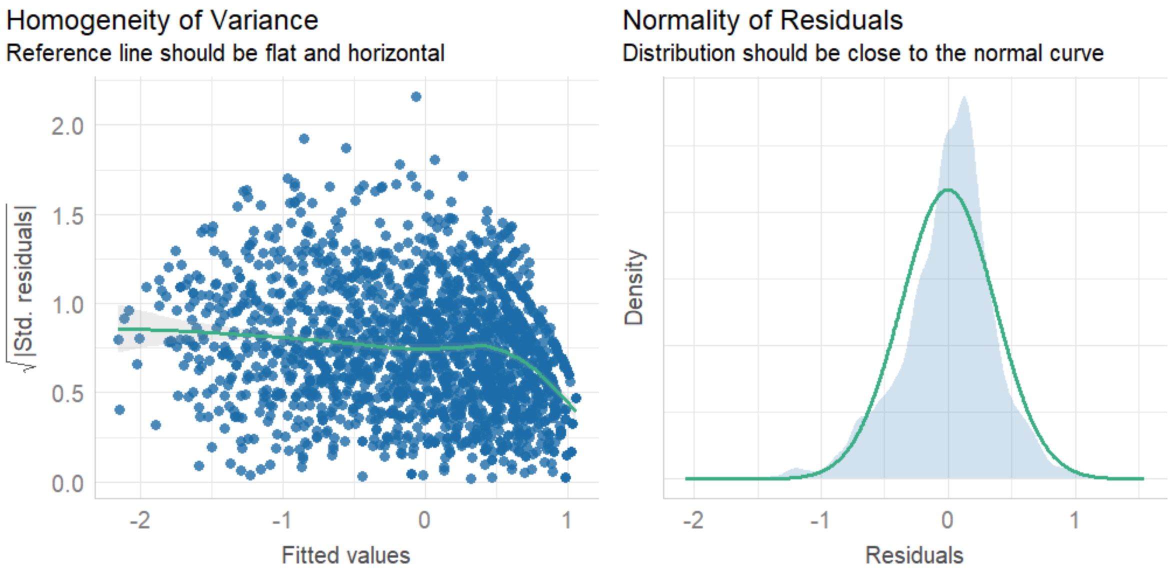
i. <u>Reporting tests without outliers</u>

We fitted a linear mixed-effect model to predict empathy with condition, aspect of empathy (cognitive, affective or motivational) and their interaction, to the full dataset. The model showed a significant main effect of condition (F(1, 732) = 20.14, $p < .001$, partial $\eta^2 = .03$). This reinforces our central results of a significant difference between the two conditions across all participants. The interaction between condition and empathy type remained significant in this analysis (F(2, 1464) = 7.47, $p < .001$, partial $\eta^2 = .01$), with contrasts showing that the differences between conditions were significant for all empathy types (cognitive: t(976) = -3.32, SE = .07, $p < .001$, $\beta = -.24$, 95% CI = [-.39, .10]; affective: t(976) = -5.60, SE = .07, $p < .001$, $\beta = -.41$, 95% CI = [-.55, .27]; motivational: t(976) = -3.57, SE = .07, $p < .001$, $\beta = -.26$, 95% CI = [-.40, -.12]), but significantly larger for affective empathy (compared to cognitive: t(1464)

= 3.52, SE = .05, p = .001, β = .17, 95% CI = [.05, .28]; compared to motivational: t(1464) = -3.14, SE = .05, p = .005, β = -.15, 95% CI = [-.26, -.04]).

ii. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W =.98, p < .001) but are not extremely skewed.



To assess whether our results hold without assuming normality, we ran a permutation test, shuffling the experimental conditions between participants and the empathy scores between types within each participant 10,000 times; and each time we analyzed a model similar to that described previously to assess the differences between conditions, empathy types and their interactions. This allowed us to create an F distribution using our data, without any effect of condition and without assuming normality. The test showed that for the main effect of condition, our original results hold true (original F(1, 709.86) = 16.81, permutation p < .001). The same was true for the interaction between condition and empathy type (original F(2, 1408.87) = 8.92, permutation p < .001). Contrasts revealed that effects for affective empathy were still significantly larger than for cognitive empathy (original t(1424.23) = 3.94, permutation $p < .001$) and for motivational empathy (original t(1424.23) = -3.29, permutation $p < .001$).

These results reinforce our central conclusion of a difference between the experimental conditions, and specifically the affective and motivational aspects of empathy, even without assuming normality.

d. Interactions with assumed aid by the other source

We aimed to replicate our findings from the previous studies, in which assumed aid by the other source (human/AI) decreased empathy in the human condition and increased it in the AI condition. A multiple-regression model predicting perceived empathy from the condition, assumed aid and their interaction found that alongside the significant main effect of condition ($F(1, 709) = 16.20$, $p < .001$, partial $\eta^2 = .02$), there was a significant interaction between condition and assumed aid by the other source ($F(1, 709) = 6.05$, $p = .01$, partial $\eta^2 = .008$). In this study, as in study 1b, post-hoc contrasts revealed a significant negative effect of assumed aid on perceived empathy in the human condition ($t(709) = -3.64$, SE = .06, $p < .001$, $\beta = -.17$, 95% CI = [-.27, -.08]). However, the effect of assumed aid in the AI condition was not significant ($p = .91$). This suggests that perceived AI involvement in a response presented as human lowered the perceived empathy.

We then used a linear mixed-effect model including empathy type, condition, assumed aid and their interactions to test whether these effects differed between aspects of empathy. The effect of condition remained significant ($F(1, 705.47) = 23.70$, $p < .001$, partial $\eta^2 = .03$, 95% CI = [.01, 1.00]). We again found a significant interaction between condition and assumed aid ($F(1, 708.26) = 10.34$, $p = .001$, partial $\eta^2 = .01$). The interaction of empathy type and condition also remained significant ($F(2, 1402.47) = 3.57$, $p = .03$, partial $\eta^2 = .005$), and there was a significant three-way interaction between condition, empathy type and assumed aid ($F(2, 1405.41) = 7.97$, $p < .001$, partial $\eta^2 = .01$). Post-hoc contrasts revealed significant negative effects of assumed aid (from AI) in the human condition on all aspects of empathy (cognitive: $t(1022) = -3.32$, $SE = .05$, $p < .001$ $\beta = -.18$, 95% CI = [-.28, -.07]; affective: $t(1021) = -4.75$, $SE = .05$, $p < .001$ $\beta = -.26$, 95% CI = [-.36, -.15]; motivational: $t(1025) = -3.89$, $SE = .05$, $p < .001$ $\beta = -.21$, 95% CI = [-.31, -.10]) and only a marginally significant positive effect on affective empathy in the AI condition ($t(1025) = 1.91$, $SE = .06$, $p = .06$, $\beta = .12$, 95% CI = [-.003, .24]). This suggests that perceived AI involvement in responses perceived as human lowers all types of empathy, but perceived human involvement in AI responses increases specifically the attribution of affective empathy.

We also replicated our finding of a significant interaction between condition and assumed aid by the other source on positivity resonance ($F(1, 698) = 33.17$, $p < .001$, partial $\eta^2 = .05$), with post-hoc contrasts showing that assumed aid had a positive effect in the AI condition, increasing positivity resonance ($t(698) = 2.57$, $SE = .06$, $p = .01$, $\beta = .15$, 95% CI = [.03, .26]), but a negative effect in the human condition, decreasing positivity resonance ($t(698) = -5.86$, $SE = .05$, $p < .001$, $\beta = -.29$, 95% CI = [-.39, -.19]). This strengthens our previous conclusion that perceived AI involvement lowers the positivity resonance in a response presented as human, but that perceived human involvement in AI responses increases it.

We again found a similar interaction for positive emotions ($F(1, 693) = 16.04$, $p < .001$, partial $\eta^2 = .02$), with assumed aid having a negative effect in the human condition ($t(693) = -5.59$, $SE = .05$, $p < .001$, $\beta = -.26$, 95% CI = [-.35, -.17]) but no significant effect in the AI condition ($p = .63$). This suggests that perceived AI involvement in a response presented as human decreased the level of positive emotions.

Authenticity was influenced similarly, with assumed aid interacting significantly with condition ($F(1, 692) = 36.26$, $p < .001$, partial $\eta^2 = .05$), having a positive effect in the AI condition ($t(692) = 2.33$, SE = .06, $p = .02$, $\beta = .13$, 95% CI = [.02, .24]) and a negative effect in the human condition ($t(692) = -6.52$, SE = .05, $p < .001$, $\beta = -.31$, , 95% CI = [-.40, -.22]). Responses that were perceived as involving greater human involvement were considered more authentic compared to responses perceived as involving AI.

The same interaction was significant when predicting support perceived in the response ($F(1, 730) = 45.27$, $p < .001$, partial $\eta^2 =.06$), where assumed aid by the other source positively influenced support perceived in the AI condition ($t(730) = 3.03$, SE = .07, $p = .002$, $\beta = .20$, 95% CI = [.07, .34]) but was a negative influence in the human condition ($t(730) = -6.78$, SE = .06, $p < .001$, $\beta = -.40$, 95% CI = [-.51, -.28]). This suggests that responses perceived as involving human input were thought to be more supportive, whereas perceived AI involvement lowered feelings of support.

These results replicate most of the findings we found in the previous studies, showing that our effects are reproducible using a different, open-source LLM.

e. <u>Extended results on assumed aid by the other source</u>

We report here further analyses without filtering outliers, and separate analyses using permutations without assuming normality. These are parallel to the multiple-regression models presented earlier in the supplementary (section 5d), providing evidence for the effect of condition and the interactions between condition and assumed aid by the other source (human/AI).

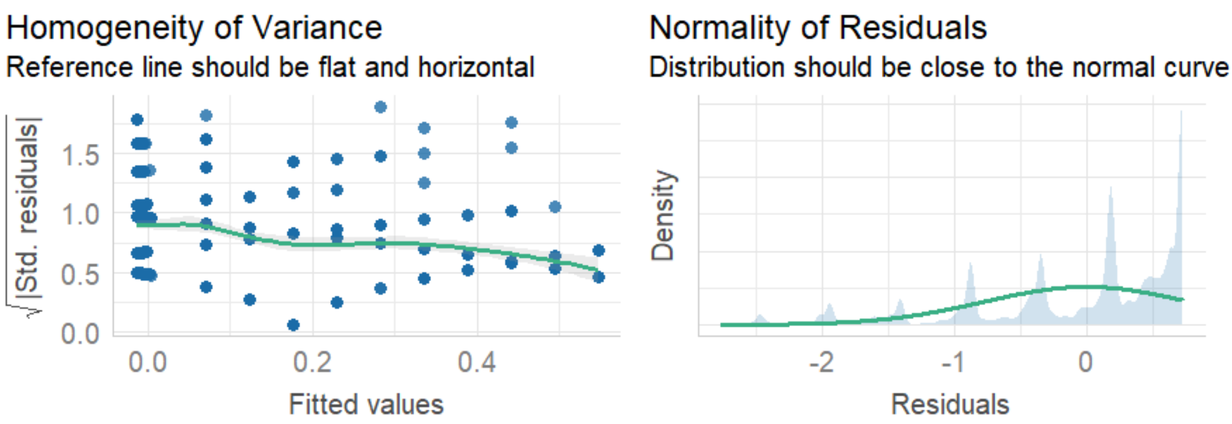
i. <u>General empathy item</u>

a. Model without filtering outliers

We fitted a linear model predicting the general empathy rating from condition, assumed aid by the other source and their interaction. Both the effect of condition ($F(1, 730) = 19.72$, $p < .001$, partial $\eta^2 = .03$) and the interaction effect remained significant ($F(1, 730) = 16.04$, $p < .001$, partial $\eta^2 = .02$). Contrasts showed that the negative effect of assumed aid on empathy in the human condition ($t(730) = -3.79$, $SE = .06$, $p < .001$ $\beta = -.22$, 95% CI = [-.34, -.11]) remained significant. Additionally, there was a positive effect in the AI condition ($t(730) = 2.02$, $SE = .06$, $p = .04$ $\beta = .14$, 95% CI = [.003, .27]).

b. Permutation analyses without assuming normality

This figure shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: $W = .84$, $p < .001$).



We ran 10,000 permutations to test the model, in an analysis identical to the one in study 1b. The permutation test showed that our main effect for condition remained significant ($F(1, 709) = 16.20$, permutation $p < .001$), as did the interaction with assumed aid by the other source ($F(1, 709) = 6.05$, permutation $p = .01$). The permutations' contrasts showed that the negative effect of assumed aid on empathy in the human condition remained significant ($t(709) = -3.64$, permutation $p < .001$).

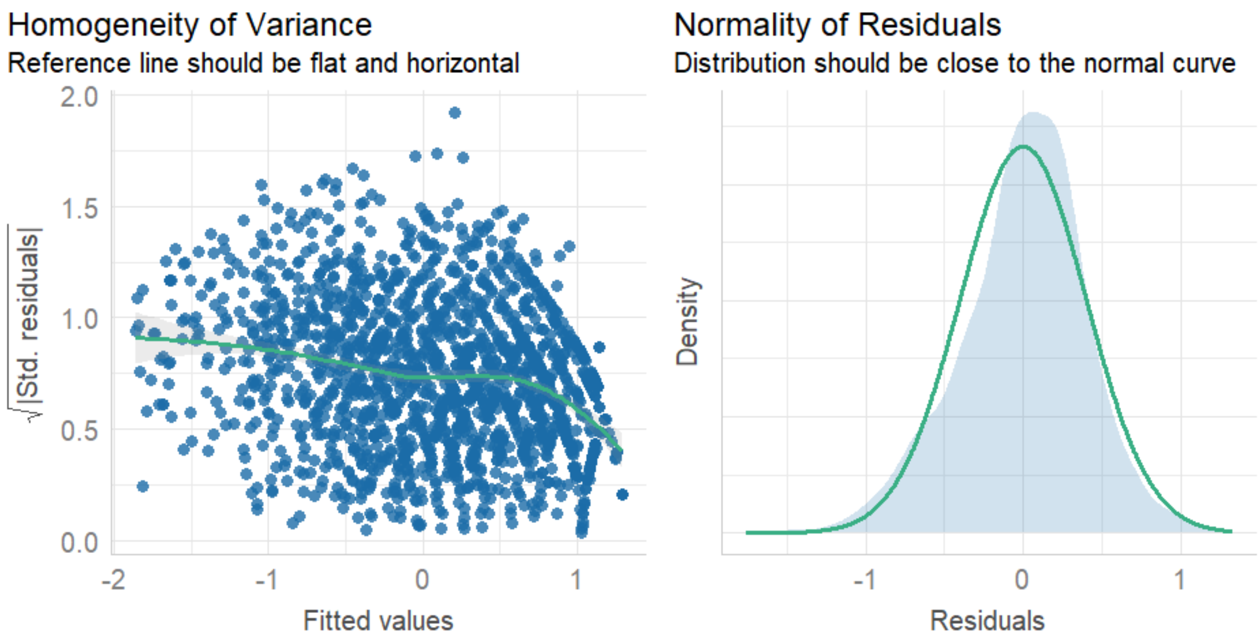
ii. <u>Mixed model for multiple empathy types</u>

a. Model without filtering outliers

We fitted a linear mixed-effect model to predict empathy with condition, aspect of empathy (cognitive, affective or motivational), assumed aid by the other source (human/AI) and their interaction, to the full dataset. The model showed a significant main effect of condition (F(1, 730) = 21.17, $p < .001$, partial $\eta^2 = .03$). This reinforces our central results of a significant difference between the two conditions across all participants. The interaction between condition and empathy type remained significant in this analysis (F(2, 1460) = 7.47, $p = .04$, partial $\eta^2 = .004$). The three-way interaction between condition, empathy type and assumed aid remained significant (F(2, 1460) = 4.80, $p = .008$, partial $\eta^2 = .006$), with contrasts showing that assumed aid had a negative effect on all types of empathy in the human condition (cognitive: t(979) = -3.49, SE = .06, $p < .001$, $\beta = -.21$, 95% CI = [-.32, -.09]; affective: t(979) = -4.60, SE = .06, $p < .001$, $\beta = -.27$, 95% CI = [-.39, -.16]; motivational: t(979) = -4.02, SE = .06, $p < .001$ $\beta = -.24$, 95% CI = [-.36, -.12]). Meanwhile, assumed aid had a positive effect in the AI condition only for affective empathy (t(979) = -3.49, SE = .07, $p = .008$, $\beta = .18$, 95% CI = [.05, .32]), replicating our previous findings.

b. Permutation analyses for interaction model with assumed aid

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W =.99, $p < .001$) but are not extremely skewed.



We ran a permutation test identical to the one in study 1b, using 10,000 permutations, while also including the assumed aid by the other source. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 561.82) = 60.38, permutation p < .001), and the interaction with assumed aid by the other source also remained significant (F(1, 564.08) = 38.09, permutation p < .001). Contrasts showed that the negative effect of assumed aid by the other source in the human condition was significant for all aspects of empathy (all permutation ps < .001). Additionally, in the AI condition, the permutation test replicated the positive effect of assumed aid on affective empathy (t(928) = 2.51, permutation p = .007) and showed a significant effect on motivational empathy (t(936) = 1.73, permutation p =.04).

iii. <u>Positivity resonance</u>

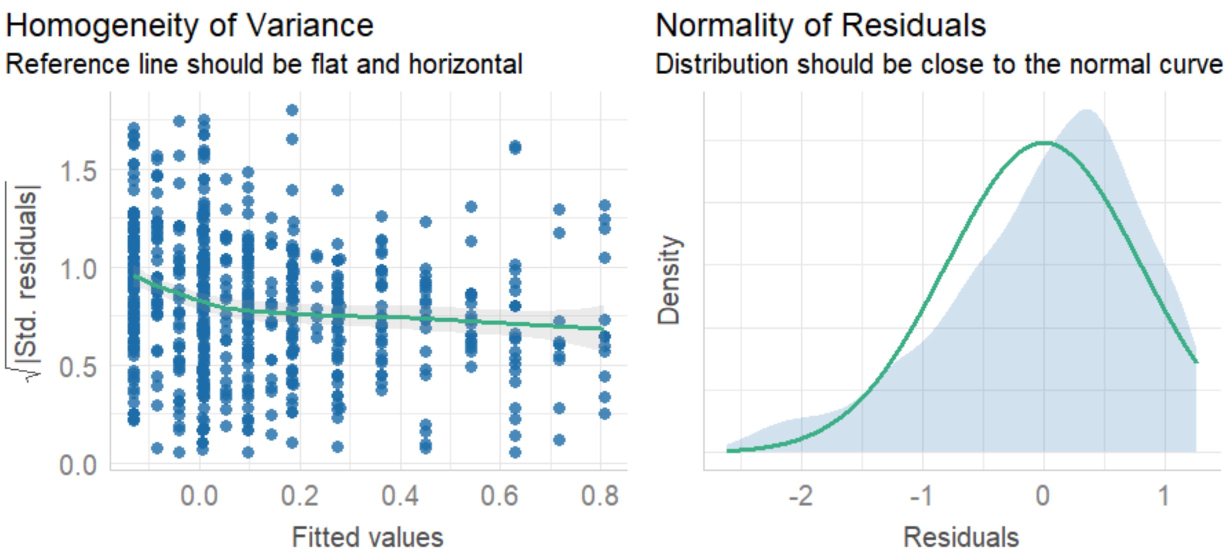
a. Model without filtering outliers

We fitted a linear model predicting positivity resonance from condition, assumed aid by the other source and their interaction. Both the effect of condition (F(1, 730) = 32.62, p <

.001, partial η^2 = .04) and the interaction effect remained significant (F(1, 730) = 49.59, p < .001, partial η^2 = .06). Contrasts showed that the negative effect of assumed aid in the human condition (t(730) = -6.25, SE = .06, p < .001, β = -.36, 95% CI = [-.48, -.25]) and the positive effect in the AI condition (t(730) = 3.90, SE = .07, p < .001, β = .26, 95% CI = [.13, .39]) remained significant.

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .96, p < .001) but are not extremely skewed.



We ran a permutation test identical to the one in study 1b, using 10,000 permutations, including the assumed aid by the other source. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 698) = 23.63, permutation p < .001), and the interaction with assumed aid by the other source also remained significant (F(1, 698) = 33.17, permutation p < .001). Permutation contrasts showed that the positive effect on positivity resonance in the AI condition (t(698) = 2.57, permutation p = .006) and the negative effect in the human condition (t(698) = -5.86, permutation p < .001) remained significant.

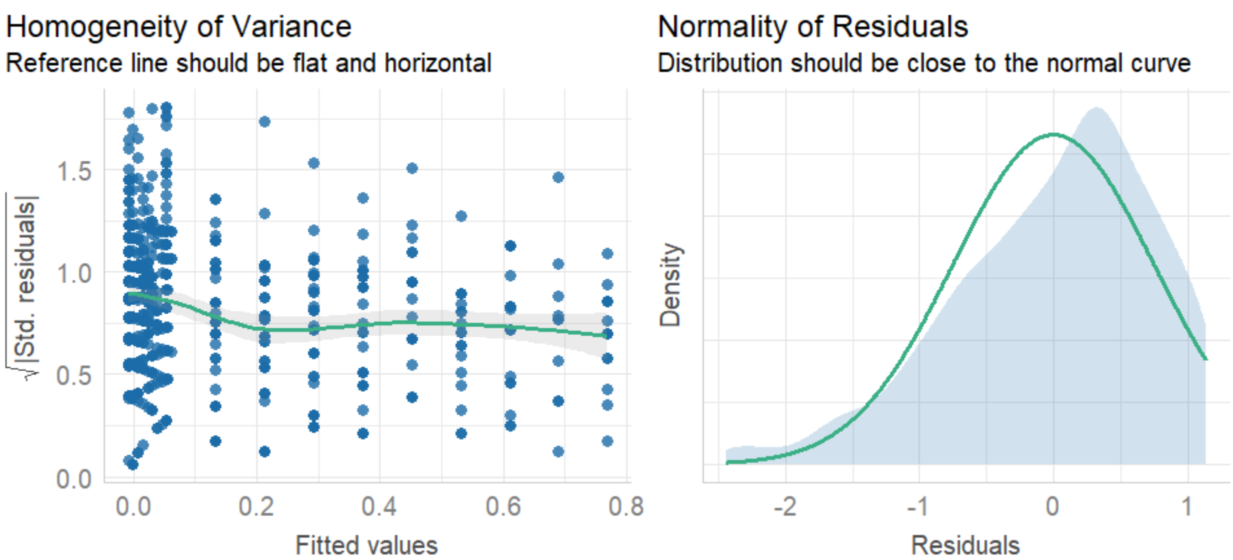
iv. Positive emotions

a. Model without filtering outliers

We fitted a linear model predicting positive emotions from condition, assumed aid by the other source and their interaction. Both the effect of condition (F(1, 730) = 25.26, p < .001, partial η^2 = .03) and the interaction effect remained significant (F(1, 730) = 39.28, p < .001, partial η^2 = .05). Contrasts showed that the negative effect of assumed aid in the human condition (t(730) = -6.41, SE = .06, p < .001, β = -.37 , 95% CI = [-.49, -.26]) remained significant, and we also found a significant positive effect in the AI condition (t(730) = 2.74, SE = .07, p = .006 β = .18 , 95% CI = [.05, .32]).

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .96, p < .001) but are not extremely skewed.



We ran a permutation test identical to the one in study 1b using 10,000 permutations of a model predicting positive emotions from condition, assumed aid by the other source and their interaction. The test showed that even using a permutation-based distribution, the effect of condition remained significant ($F(1, 693) = 24.23$, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant ($F(1, 693) = 16.03$, permutation $p < .001$). Permutation contrasts showed that the negative effect of assumed aid on positive emotions in the human condition remained significant ($t(693) = -5.59$, permutation $p < .001$).

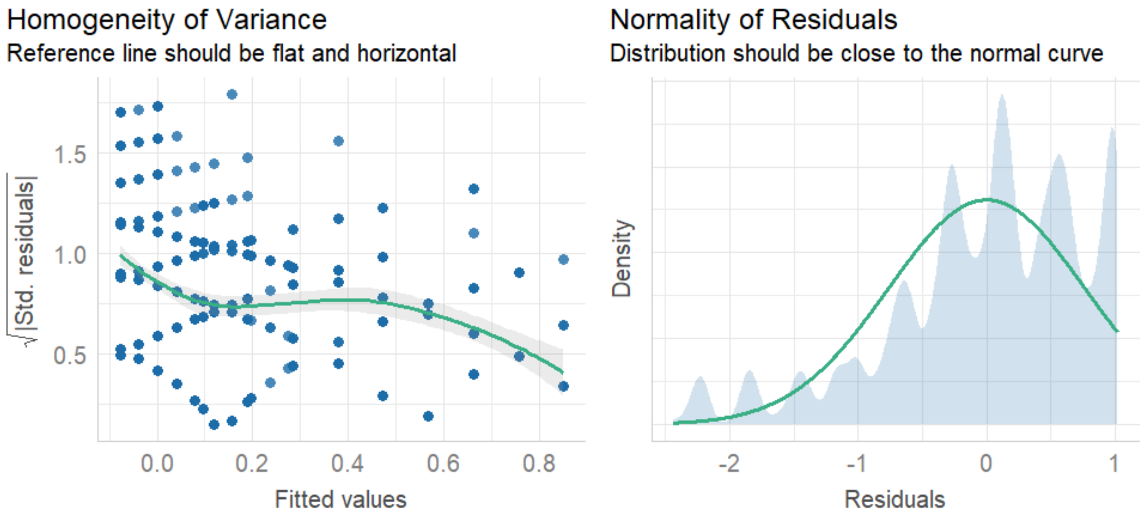
v. Authenticity

a. Model without filtering outliers

We fitted a linear model predicting authenticity from condition, assumed aid by the other source and their interaction. Both the effect of condition ($F(1, 730) = 22.12$, $p < .001$, partial $\eta^2 = .03$) and the interaction effect ($F(1, 730) = 59.78$, $p < .001$, partial $\eta^2 = .08$) remained significant. Contrasts showed that the negative effect of assumed aid in the human condition ($t(730) = -7.40$, SE = .06, $p < .001$, $\beta = -.43$, 95% CI = [-.54, -.31]) and the positive effect in the AI condition ($t(730) = 3.82$, SE = .07, $p < .001$, $\beta = .25$, 95% CI = [.12, .39]) both remained significant.

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .93, $p < .001$).

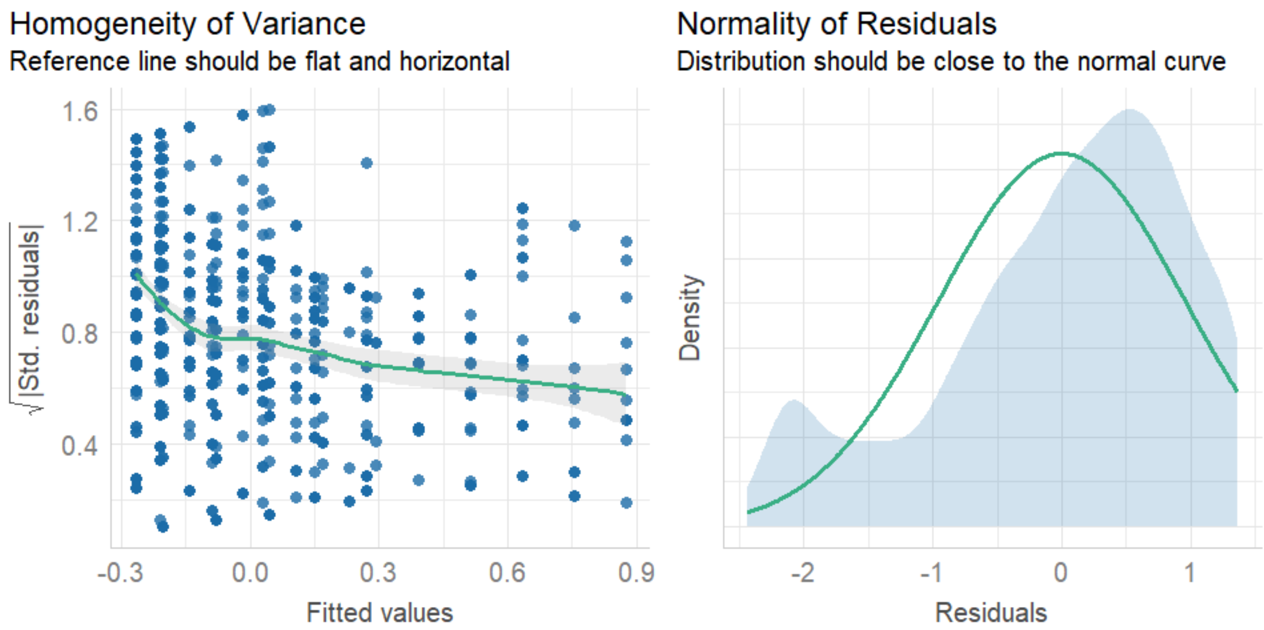


We ran a permutation test identical to the one in study 1b, using 10,000 permutations of models predicting authenticity for condition, assumed aid by the other source and their interaction. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 692) = 21.78, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant (F(1, 692) = 36.26, permutation $p < .001$). Permutation contrasts showed that the positive effect of assumed aid on authenticity in the AI condition (t(692) = 2.32, permutation $p = .01$) and its negative effect on authenticity in the human condition (t(692) = -6.51, permutation $p < .001$) both remained significant.

vi. Support

a. Permutation analyses without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .96, $p < .001$) but are not extremely skewed.



We ran a permutation test identical to the one in study 1b, using 10,000 permutations, including the assumed aid by the other source. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 730) = 14.25, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant (F(1, 730) = 45.27, permutation $p < .001$). Permutation contrasts showed that the positive effect of assumed aid on perceived support in the AI condition (t(730) = 3.03, permutation $p = .001$) and its negative effect in the human condition (t(730) = -6.78, permutation $p < .001$) both remained significant.

6. Detailed analyses for study 1d

a. Additional effects of condition

This study aimed to replicate our previous results, but using a 4-turn continuous interaction, as opposed to a single response. We found a marginally significant effect of condition on positivity resonance (t(685.99) = -1.96, $p = .05$, Cohen's $d = .15$, 95% CI [-6.62,

.01]). This suggests that as in the previous studies, participants perceived the communication as more positive when they thought it came from a human.

We also found a marginally significant effect of condition on the level of support participants attributed to the response (t(699.72) = 1.75, p = .08, Cohen's d = .13, 95% CI [-.04, .70]).

Additionally, we replicated the effect of condition on positive emotions. Responses presented as human raised significantly more positive emotions than those presented as AI-generated (t(683.29) = -2.02, p = .04, Cohen's d = .15, 95% CI [-.56,-.01]), showing that participants felt more positive emotions when perceiving the response as human.

Initially, t-tests did not find a significant main effect of condition on authenticity (p = .86) or on negative emotions (p = .30).

b. **Interactions with assumed aid by the other source**

We aimed to replicate our findings from study 1b, in which assumed aid by the other source decreased empathy in the human condition and increased it in the AI condition. A multiple-regression model predicting perceived empathy from the condition, assumed aid and their interaction found a significant interaction between condition and assumed aid (F(1, 673) = 33.17, p < .001, partial η^2=.05). As in the previous studies, post-hoc contrasts revealed a positive effect of assumed aid on empathy in the AI condition (t(673) = 2.30, SE = .05, p = .03, β = .12, , 95% CI = [.01, .22]) but a negative one in the human condition (t(677) = -5.77, SE = .05, p < .001, β = -.32 , 95% CI = [-.43, -.21]).

We then used a linear mixed-effect model including empathy type, condition, assumed aid and their interactions to test whether these effects differed between aspects of empathy. The model found a significant effect of condition (F(1, 684.50) = 22.58, p < .001, partial η^2 = .03). We again found a significant interaction between condition and assumed aid

($F(1, 686.05) = 52.94$, $p < .001$, partial $\eta^2 = .07$). There was no effect ($p = .40$) or interaction of empathy type ($p = .45$). As in studies 1b and 1c, there was a significant three-way interaction between condition, assumed aid and empathy type ($F(2, 1363.59) = 3.07$, $p = .05$, partial $\eta^2 = .004$), with post-hoc contrasts showing significant negative effects of assumed aid in the human condition on all aspects of empathy (cognitive: $t(1005) = -7.03$, SE = .06, $p < .001$, $\beta = -.46$, 95% CI = [-.59, -.33]; affective: $t(994) = -7.82$, SE = .07, $p < .001$, $\beta = -.51$, 95% CI = [-.64, -.38]; motivational: $t(1002) = -6.97$, SE = .07, $p < .001$, $\beta = -.46$, 95% CI = [-.59, -.33]), and a significant positive effect only for affective empathy in the AI condition (affective: $t(999) = 2.90$, SE = .06, $p = .004$, $\beta = .18$, 95% CI = [.06, .31]).

A model predicting positivity resonance from condition, assumed aid and their interaction found a significant main effect of condition ($F(1, 683) = 4.19$, $p = .04$, partial $\eta^2 = .006$, 95% CI = [.00, 1.00]) and a significant interaction between condition and assumed aid ($F(1, 683) = 50.07$, $p < .001$, partial $\eta^2 = .07$). Post-hoc contrasts revealed that assumed aid had a positive effect on positivity resonance in the AI condition ($t(683) = 2.25$, SE = .06, $p = .02$, $\beta = .14$, , 95% CI = [.02, .26]) but a negative effect in the human condition ($t(683) = -7.64$, SE = .06, $p < .001$, $\beta = -.49$, 95% CI = [-.62, -.37]). These results replicate our earlier findings that perceiving a response as human increases positivity resonance, but perceiving AI involvement decreases this effect. Concurrently, thinking that an empathic response is AI-generated lowers perceived empathy, but thinking there is human involvement in such a response increases perceived empathy.

We found similar effects predicting positive emotions, including a significant effect of condition ($F(1, 678) = 4.36$, $p = .04$, partial $\eta^2 = .006$) and a significant interaction between condition and assumed aid ($F(1, 678) = 44.79$, $p < .001$, partial $\eta^2 = .06$). Post-hoc contrasts showed that assumed aid had a positive effect on positive emotions in the AI condition (t(678) = 2.28, SE = .06, p = .02, β = .14, 95% CI = [.02, .25]) and a negative effect in the human condition (t(678) = -7.06, SE = .06, p < .001, β = -.44, 95% CI = [-.56, -32]).

Authenticity was influenced similarly, with assumed aid interacting significantly with condition (F(1, 701) = 95.37, p < .001, partial η^2 = .12), having a positive effect in the AI condition (t(701) = 3.11, SE = .07, p = .002, β = .21, 95% CI = [.07, .33]) and a negative effect in the human condition (t(701) = -10.54, SE = .07, p < .001, β = -.72, 95% CI = [-.85, -.58]) on perceived authenticity.

The same interaction was significant when predicting support perceived in the response (F(1, 701) = 95.68, p < .001, partial η^2=.12), where assumed aid positively influenced perceived support in the AI condition (t(701) = 3.54, SE = .07, p < .001, β = .23, 95% CI = [.10, .36]) but negatively in the human condition (t(701) = -10.15, SE = .07, p < .001, β = -.69, 95% CI = [-.83, -.56]).

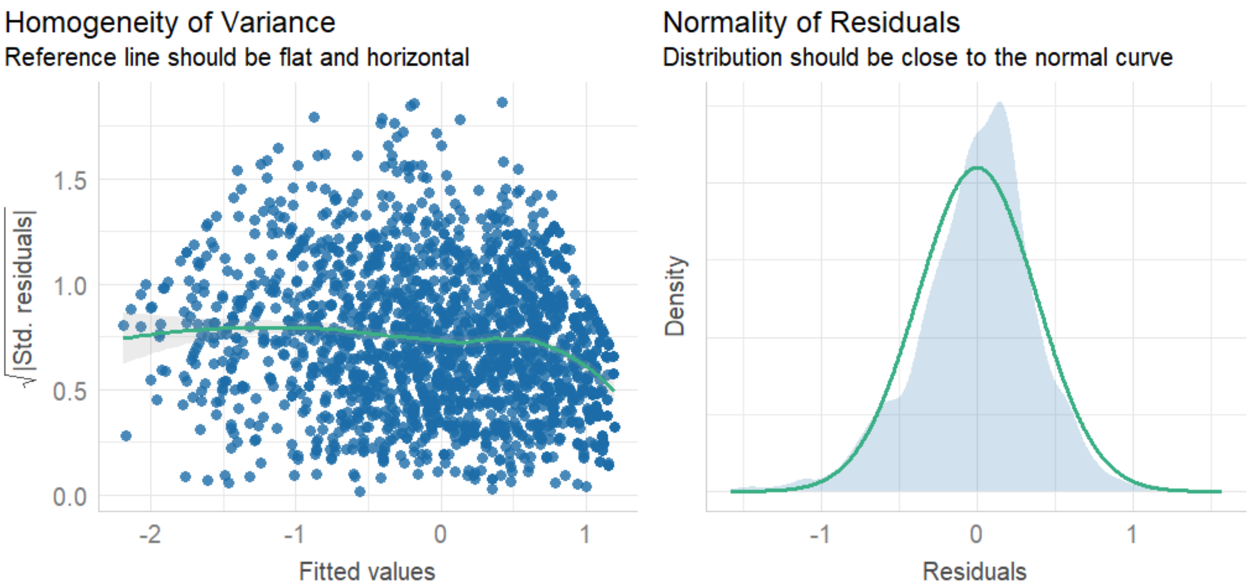
These results replicate most of the findings of study 1b, showing that in a continuous interaction, participants who perceived the responses as more human-authored had a more positive, supportive, empathic and connected experience compared to those who perceived greater AI involvement. Extended results for repeated-measures models with multiple empathy types

i. <u>Model without filtering outliers</u>

We replicated our findings with a linear mixed-effect model predicting levels of empathy from condition, empathy type and their interaction, without filtering for outliers. We found a significant main effect of condition (F(1, 703) = 5.15, p = .02, partial η^2 = .007) and a significant interaction between condition and empathy type (F(2, 1406) = 3.21, p = .04, partial η^2 = .004). Post-hoc contrasts revealed that the effect of condition was significant only for affective empathy (t(942) = -2.82, SE = .08, p = .005, β = -.21 , 95% CI = [-.36, -.06]) and motivational empathy (t(942) = -2.29, SE = .08, p = .02, β = -.17 , 95% CI = [-.32, -.03]), and not for cognitive empathy (p = .23).

ii. <u>Permutation test without assuming normality</u>

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .98, p < .001) but are not extremely skewed.



To assess whether our results hold without assuming normality, we ran a permutation test, shuffling the experimental conditions between participants and the empathy scores between types within each participant 10,000 times, and each time we analyzed a model similar to that described in the paper to assess the differences between conditions, empathy types and their interactions. This allowed us to create an F distribution using our data, without any effect of condition and without assuming normality. The test showed that the main effect of condition on empathy remained significant (original F(1, 687.42) = 5.38, permutation p = .02). Contrasts revealed that, as in our main analysis, the differences between conditions were significant specifically for affective empathy (t(975) = -

2.60, permutation p = .005) and motivational empathy (t(978) = -2.40, permutation p = .009), while for cognitive empathy the difference was not significant (t(975) = -1.36, permutation p = .09).

These results reinforce our central conclusion of a difference between the experimental conditions, specifically in the affective and motivational aspects of empathy, even without assuming normality.

c. Extended results on assumed aid by the other source

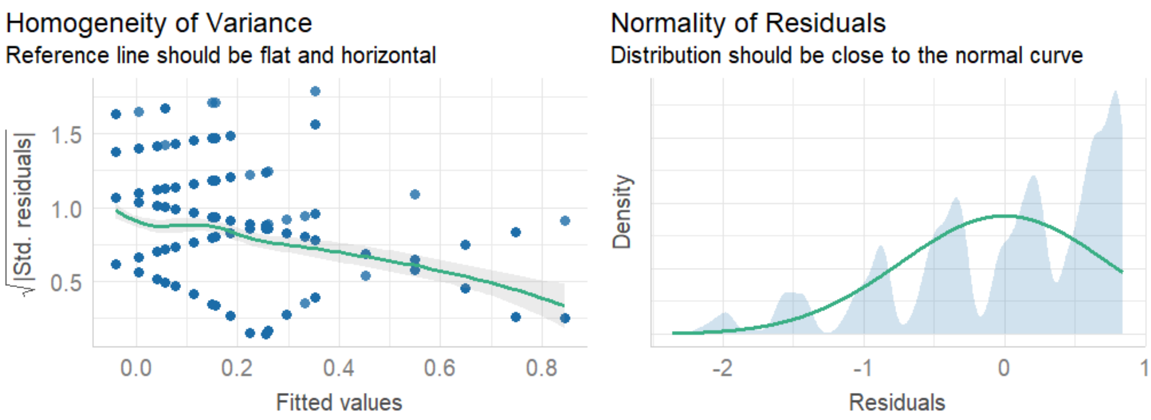
i. General empathy

a. Model without filtering outliers

Using the full dataset, we replicated our findings with a multiple-regression model predicting the rating of empathy from condition, assumed aid by the other source and their interaction. It showed a significant interaction between condition and assumed aid (F(1, 705) = 41.11, p < .001, partial η^2=.06). As in the previous studies, post-hoc contrasts revealed a positive effect of assumed aid in the AI condition (t(701) = 3.08, SE = .07, p = .002, β = .21 , 95% CI = [.07, .34]) but a negative one in the human condition (t(701) = -5.93, SE = .07, p < .001, β = -.43, 95% CI = [-.57, -.29]).

b. Permutation test without assuming normality

The figure below shows the level of deviation our model had from the assumption of normality of residuals (Shapiro-Wilk test: W = .91, p < .001).



We ran 10,000 permutations to test the model predicting the general empathy rating from the condition, assumed aid by the other source and their interaction. The permutation test showed that the interaction with assumed aid by the other source (F(1, 673) = 33.17, permutation p < .001) remained significant. The permutations' contrasts showed that the negative effect of assumed aid on empathy in the human condition (t(673) = -5.87, permutation p < .001) and the positive effect in the AI condition (t(673) = 2.23, permutation p = .01) both remained significant.

ii. Mixed model for multiple empathy types

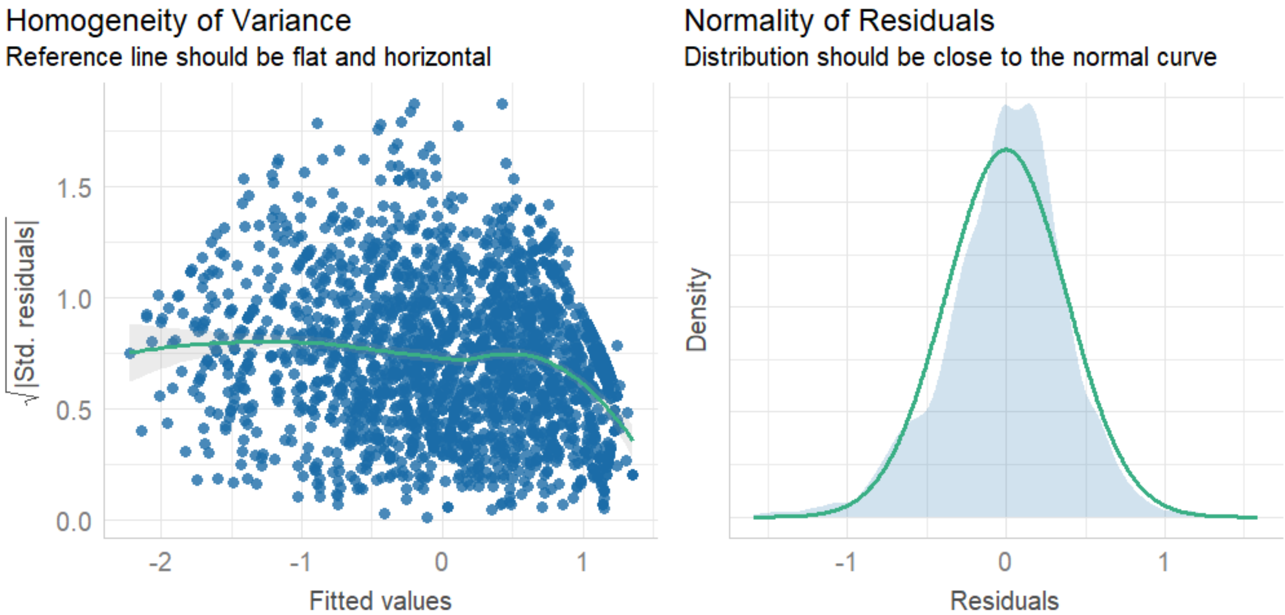
a. Model without filtering outliers

We next used a linear mixed-effect model to predict levels of empathy while controlling for empathy type in the full dataset. We replicated our results, finding a significant effect of condition (F(1, 701) = 20.45, p < .001, partial η^2 = .03). We again found a significant interaction between condition and assumed aid (F(1, 701) = 56.75, p < .001, partial η^2 = .07). There was no effect (p = .38) or interaction of empathy type (p = .52). Post-hoc contrasts remained the same, showing significant negative effects of assumed aid in the human condition on all aspects of empathy (cognitive: t(963) = -7.36, SE = .07, p < .001, β = -.52 , 95% CI = [-66, -.38]; affective: t(963) = -7.44, SE = .07, p < .001, β = -.53 , 95% CI = [-67, -

.39]; motivational: t(963) = -7.14, SE = .07, $p < .001$, β = -.50, 95% CI = [-.64, -.37]), but a significant positive effect only for affective empathy in the AI condition (affective: t(963) = 3.36, SE = .07, $p = .001$ β = .22, 95% CI = [.09, .36]).

b. Permutation test without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .99, $p < .001$) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations of a model predicting empathy from condition, empathy type, assumed aid by the other source and their interactions. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 684.50) = 22.58, permutation $p < .001$). The interaction with assumed aid by the other source (F(1, 686.05) = 52.95, permutation $p < .001$) and the three-way interaction between condition, empathy type and assumed aid (F(2, 1363.59) = 3.07, permutation $p = .05$) also remained significant. Contrasts showed that the negative effect of assumed aid by the other source in the human condition was significant for all aspects of empathy (all permutation *p*s < .001), but it was only significant in the AI condition for affective empathy (t(999) = 2.90, permutation $p = .002$) and not significant for cognitive (t(1000) = 1.53, permutation $p = .06$) or motivational empathy (t(997) = 1.43, permutation $p = .08$).

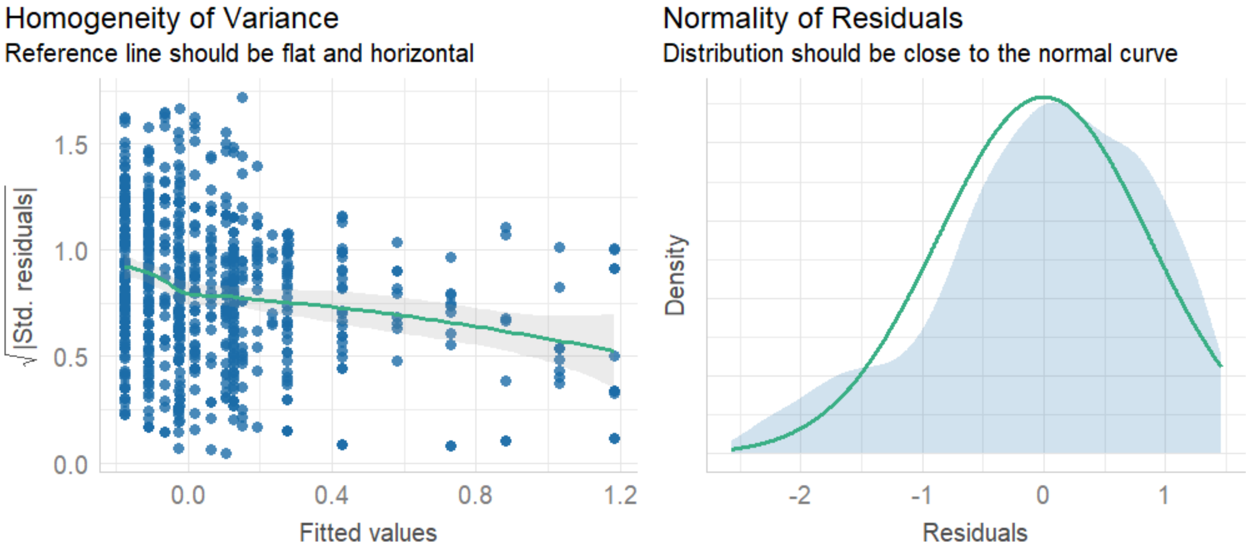
iii. <u>Positivity resonance</u>

a. Model without filtering outliers

A model predicting positivity resonance from assumed aid, condition and their interactions in the full dataset replicated our findings. We found a significant main effect of condition ($F(1, 701) = 5.04$, $p = .03$, partial $\eta^2 = .007$) and a significant interaction between condition and assumed aid ($F(1, 701) = 62.56$, $p < .001$, partial $\eta^2 = .08$). Post-hoc contrasts revealed that assumed aid had a positive effect on positivity resonance in the AI condition ($t(701) = 3.23$, SE = .07, $p = .001$, $\beta = .22$, 95% CI = [.08, .35]) but a negative effect on it in the human condition ($t(701) = -7.85$, SE = .07, $p < .001$, $\beta = -.55$, 95% CI = [-.69, -.41]).

b. Permutation test without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .97, $p < .001$) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations to test a model predicting positivity resonance from condition, assumed aid by the other source and their interaction. The test showed that even using a permutation-based distribution, the effect of condition remained significant ($F(1, 683) = 4.19$, permutation $p = .04$), and the interaction with assumed aid by the other source also remained significant ($F(1, 683) = 50.07$, permutation $p < .001$). Permutation contrasts showed that the positive effect on positivity resonance in the AI condition ($t(683) = 2.26$, permutation $p = .01$) and the negative effect in the human condition ($t(683) = -7.64$, permutation $p < .001$) remained significant.

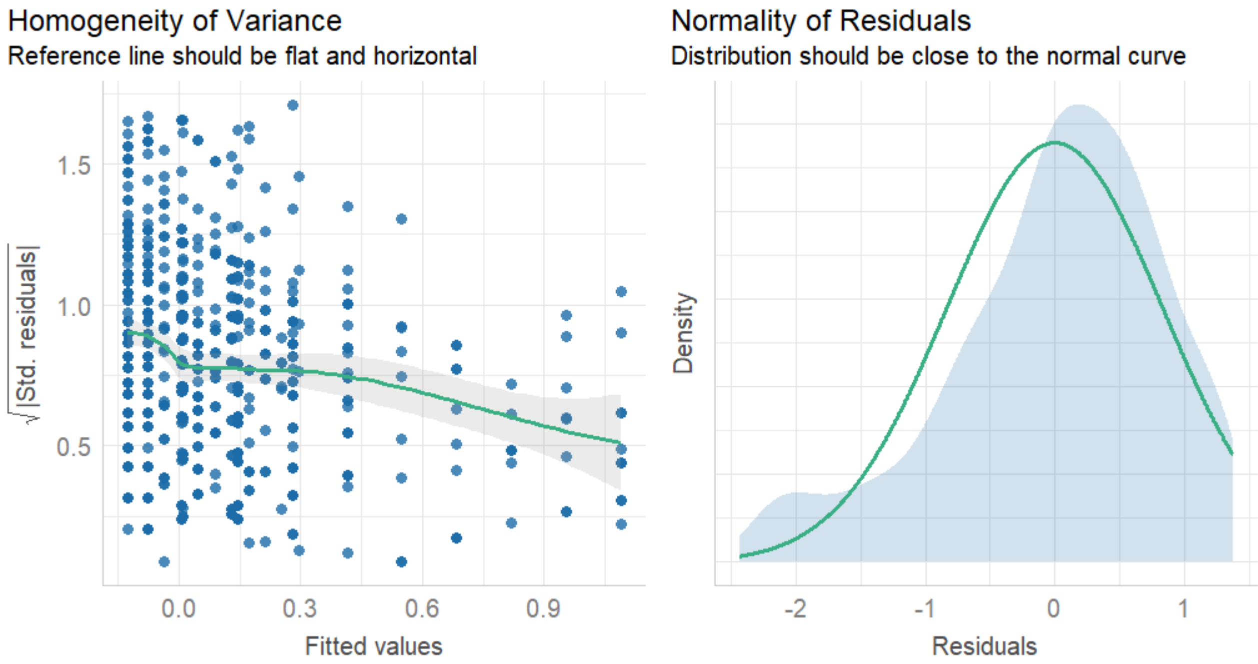
iv. <u>Positive emotions</u>

a. Model without filtering outliers

When using a similar model to predict positive emotions, we did not find a significant effect of condition ($p = .30$). However, the significant interaction between condition and assumed aid remained ($F(1, 701) = 60.12$, $p < .001$, partial $\eta^2 = .08$). Post-hoc contrasts showed that assumed aid had a positive effect on positive emotions in the AI condition ($t(701) = 2.68$, $SE = .07$, $p = .008$, $\beta = .18$, 95% CI = [.05, .31]) and a negative effect in the human condition ($t(701) = -8.16$, $SE = .07$, $p < .001$, $\beta = -.57$, 95% CI = [-.71, -.43]).

b. Permutation test without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .96, *p* < .001) but are not extremely skewed.

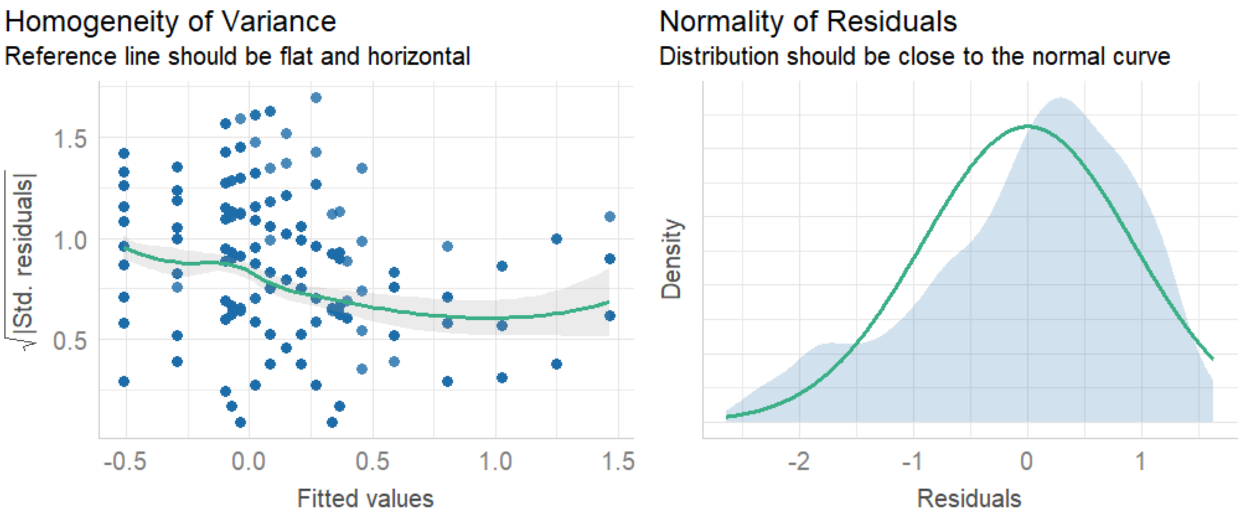


We ran a permutation test identical to the one in study 1b, using 10,000 permutations, including the assumed aid by the other source. The test showed that even using a permutation-based distribution, the effect of condition on positive emotions remained significant (F(1, 678) = 4.36, permutation *p* = .04), and the interaction with assumed aid by the other source also remained significant (F(1, 678) = 44.79, permutation *p* < .001). Permutation contrasts showed that the positive effect of assumed aid on positive emotions in the AI condition (t(678) = 2.28, permutation *p* = .01) and the negative effect in the human condition (t(661) = -7.06, permutation *p* < .001) both remained significant.

v. Authenticity

a. Permutation test without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .94, p < .001) but are not extremely skewed.

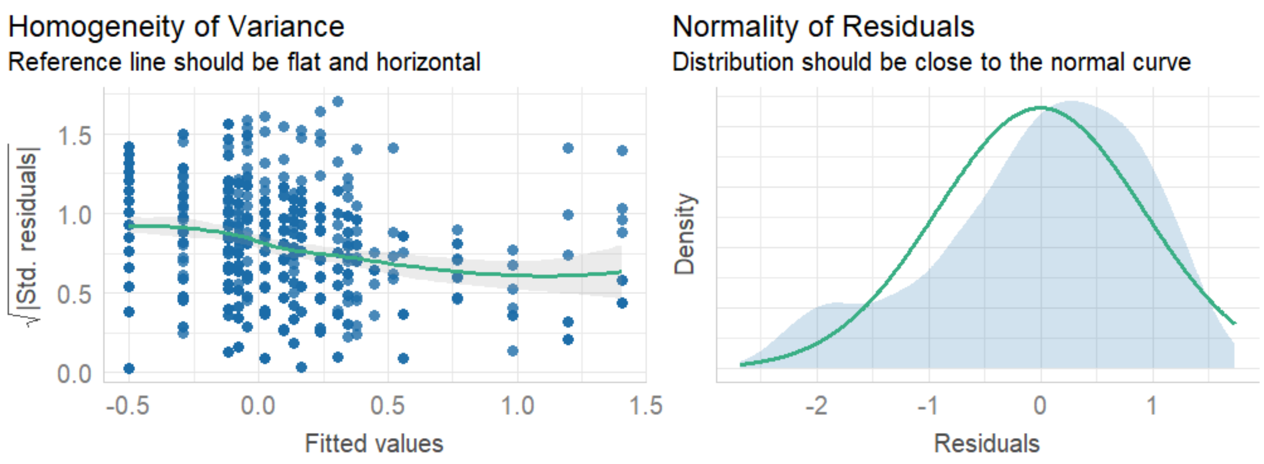


We ran a permutation test, using 10,000 permutations of a model predicting authenticity from condition, assumed aid by the other source and their interaction. The test showed that even using a permutation-based distribution, the interaction with assumed aid by the other source remained significant (F(1, 701) = 95.37, permutation p < .001). Permutation contrasts showed that the positive effect of assumed aid on authenticity in the AI condition (t(701) = 3.11, permutation p = .001) and the negative effect in the human condition (t(701) = -10.54, permutation p < .001) both remained significant.

vi. <u>Support</u>

a. Permutation test without assuming normality

The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .96, p < .001).



We ran a permutation using 10,000 permutations of a model predicting perceived support from condition, assumed aid by the other source and their interaction. The test showed that even using a permutation-based distribution, the effect of condition was still marginally significant ($F(1, 701) = 3.57$, permutation $p = .06$), and the interaction with assumed aid by the other source remained significant ($F(1, 701) = 95.68$, permutation $p < .001$). Permutation contrasts showed that the positive effect of assumed aid on perceived support in the AI condition ($t(701) = 3.54$, permutation $p < .001$) and the negative effect in the human condition ($t(701) = -10.15$, permutation $p < .001$) both remained significant.

7. <u>Detailed analyses for study 2a</u>

This study aimed to rule out an alternative explanation: a halo effect from participants receiving an eloquent response in a timely manner, as opposed to the differences in perceived sources between conditions influencing the perceptions of empathy. This was done by instructing the model to generate shorter and less eloquent responses.

a. <u>Analyses for empathy</u>

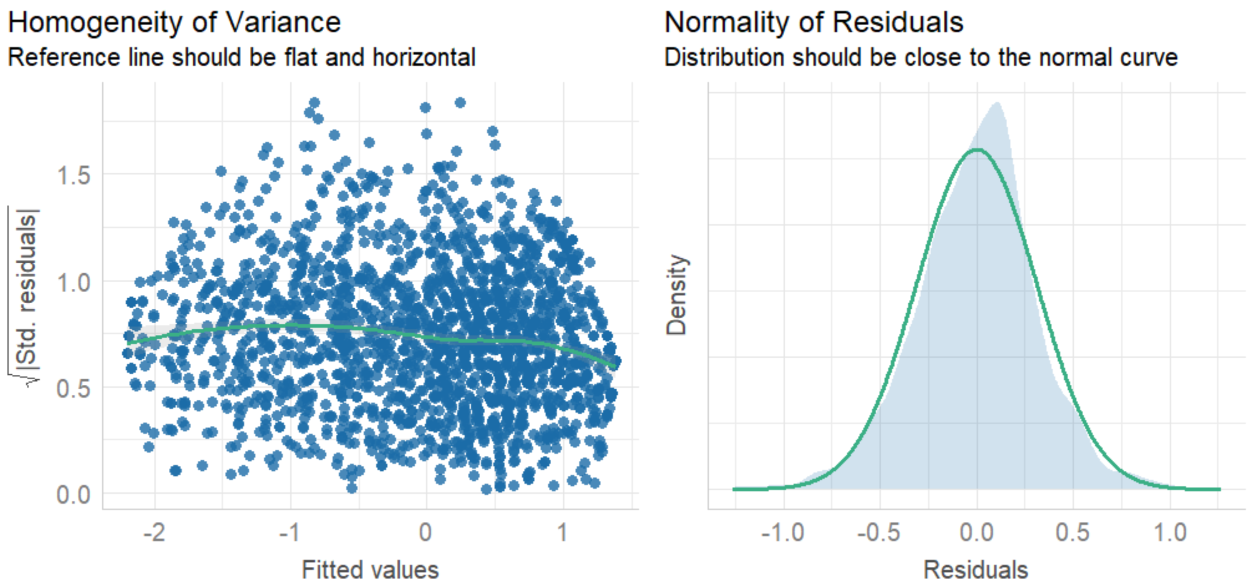
i. <u>Main analyses</u>

We repeated the same analyses for this study as we did for study 1b. We found a significant difference in the single-item empathy question between conditions (Welch t-test: $t(649.28) = -8.20$, $p < .001$, Cohen's d =.62, 95% CI for difference [-1.94, -1.19]). There were no outliers to this test.

As in study 1, a linear mixed-effect model predicting empathy by the experimental condition, empathy type and the interaction between them found a significant main effect of condition ($F(1, 685) = 68.59$, $p < .001$, partial $\eta^2 = .09$), with participants in the human condition rating responses as more empathic than those in the AI condition. We did not find a significant effect of empathy type ($p = 1.00$) or interaction between it and the condition (p =.33). There were no outliers in these analyses.

ii. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .99, $p < .001$) but are not extremely skewed.



We ran a permutation test identical to the one in study 1b, using 10,000 permutations. The test showed that even using a permutation-based distribution, the effect of condition on empathy remained significant (F(1, 685) = 68.59, permutation $p < .001$).

b. Additional effects of condition

We found a significant effect of condition on positivity resonance (Welch t-test: t(677.66) = -8.49, $p < .001$, Cohen's $d = .65$, 95% CI [-24.16, -15.09]). As in studies 1a and 1b, participants perceived the communication as more positive when they thought it was from a human. There were no outliers for this test.

We also found a significant difference between conditions in the level of authenticity that participants attributed to the response (Welch t-test: t(665.48) = -8.06, Cohen's $d = .61$, 95% CI [-.73, -.44], $p < .001$). Thus, like previously, participants found the response more authentic when they perceived it as human. There were no outliers for this test.

Condition also significantly influenced the how supportive the participants perceived the response to be (t(682.99) = -5.46, $p < .001$, Cohen's $d = .42$, 95% CI [-1.67, -.79]). Participants perceived the response as more supportive when they believed it was human-authored. There were no outliers for this test.

Responses presented as human raised significantly more positive emotions than those presented as AI-generated (t(676.79) = -7.10, $p < .001$, Cohen's $d = .54$, 95% CI [-1.93, -1.09]). Participants felt more positive emotions when perceiving the response as human. There were no outliers for this test.

There was a significant difference in the levels of negative emotions caused by the response (t(630.84) = 3.04, $p = .002$, Cohen's $d = .24$, 95% CI [.09, .44]). Participants felt lower levels of negative emotions when perceiving the response as human.

c. Interactions with assumed aid by the other source

We aimed to replicate our findings from study 1b, in which assumed aid by the other source (human/AI) decreased empathy in the human condition and increased it in the AI condition. A multiple-regression model predicting perceived empathy from the condition, assumed aid and their interaction found that alongside the significant main effect of condition ($F(1, 683) = 68.77$, $p < .001$, partial $\eta^2 = .09$), there was a significant interaction between condition and assumed aid ($F(1, 683) = 21.61$, $p < .001$, partial $\eta^2 = .03$). As in study 1b, post-hoc contrasts revealed a positive effect of assumed aid in the AI condition ($t(683) = 3.63$, $SE = .07$, $p < .001$, $\beta = .26$, 95% CI = [.12, .40]) but a negative one in the human condition ($t(683) = -2.93$, $SE = .05$, $p = .003$, $\beta = -.15$, 95% CI = [-.25, -.05]). There were no outliers in these analyses.

We then used a linear mixed-effect model including empathy type, condition, assumed aid and their interactions to test whether these effects differed between aspects of empathy. The effect of condition remained significant ($F(1, 683) = 42.45$, $p < .001$, partial $\eta^2 = .06$). We again found a significant interaction between condition and assumed aid ($F(1, 683) = 43.91$, $p < .001$, partial $\eta^2 = .06$). There was no effect ($p = .53$) or interaction of empathy type ($p = .73$), and post-hoc contrasts showed significant negative effects of assumed aid in the human condition on all aspects of empathy (cognitive: $t(852) = -3.83$, $SE = .05$, $p < .001$, $\beta = -.19$, 95% CI = [-.29, -.09]; affective: $t(852) = -4.40$, $SE = .05$, $p < .001$, $\beta = -.22$, 95% CI = [-.32, -.12]; motivational: $t(852) = -4.21$, $SE = .05$, $p < .001$, $\beta = -.21$, 95% CI = [.31, -.11]) and a significant positive effect on them in the AI condition (cognitive: $t(852) = 4.23$, $SE = .07$, $p < .001$, $\beta = .30$, 95% CI = [.16, .44]; affective: $t(852) = 5.31$, $SE = .07$, $p < .001$, $\beta = .38$, 95% CI = [.24, .52]; motivational: $t(852) = 4.66$, $SE = .07$, $p < .001$, $\beta = .33$, 95% CI = [.19, .47]). There were no outliers for these analyses.

We also replicated our finding of a significant interaction between condition and assumed aid on positivity resonance (F(1, 683) = 45.43, $p < .001$, partial $\eta^2 = .06$), with post-hoc contrasts showing that assumed aid had a positive effect on the AI condition, increasing positivity resonance (t(683) = 4.81, SE = .07, $p < .001$, $\beta = .34$, 95% CI = [.20, .48]), but it had a negative effect on the human condition, decreasing positivity resonance (t(683) = -4.88, SE = .05, $p < .001$, $\beta = -.24$, 95% CI = [-.34, -.15]). There were no outliers for these analyses.

A similar interaction was found for positive emotions (F(1, 683) = 51.24, $p < .001$, partial $\eta^2 = .07$), with assumed aid having a positive effect in the AI condition (t(683) = 5.34, SE = .07, $p < .001$ $\beta = .38$, 95% CI = [.24, .52]) and a negative effect in the human condition (t(683) = -4.85, SE = .05, $p < .001$, $\beta = -.24$, 95% CI = [-.34, -.15]). There were no outliers for these analyses.

Authenticity was influenced similarly, with assumed aid from the other source interacting significantly with condition (F(1, 683) = 56.67, $p < .001$, partial $\eta^2 = .08$) and having a positive effect on authenticity in the AI condition (t(683) = 5.25, SE = .07, $p < .001$, $\beta = .37$, 95% CI = [.23, .51]) and a negative effect in the human condition (t(683) = -5.63, SE = .05, $p < .001$, $\beta = -.28$, 95% CI = [-.38, -.18]). There were no outliers for these analyses.

The same interaction was significant when predicting support perceived in the response (F(1, 683) = 57.35, $p < .001$ partial $\eta^2 = .08$), where assumed aid positively influenced it in the AI condition (t(683) = 5.63, SE = .07, $p < .001$, $\beta = .41$, 95% CI = [.26, .55]), and the influence was negative in the human condition (t(683) = -5.16, SE = .05, $p < .001$, $\beta = -.26$, 95% CI = [-.36, -.16]). There were no outliers for these analyses.

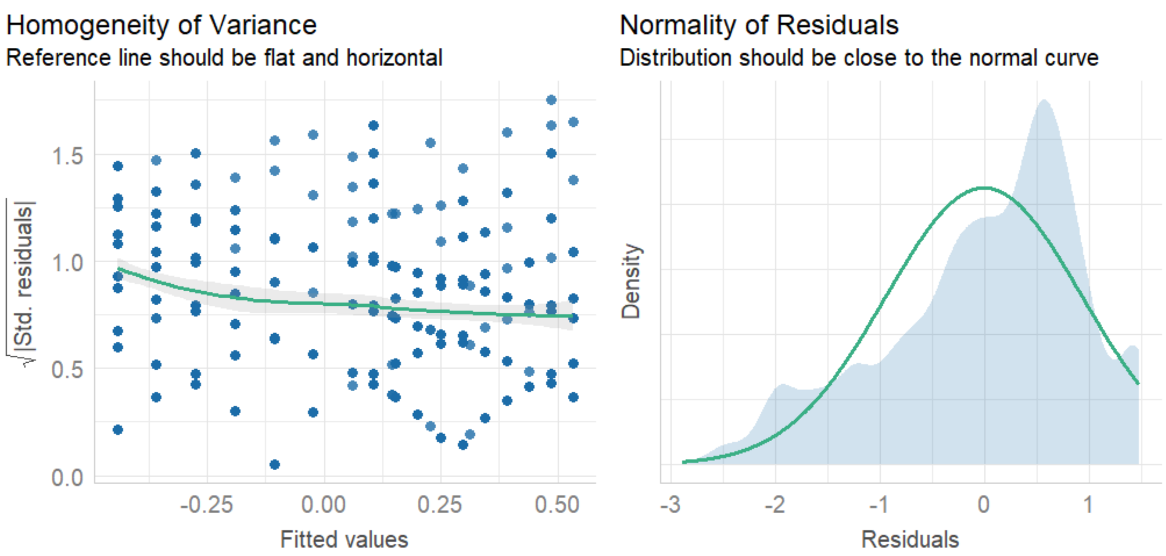
These results replicate most of the results reported in study 1b, while providing participants with shorter, less eloquent responses.

d. <u>Extended results on assumed aid by the other source</u>

We provide here further analyses using permutations without assuming normality. There are no analyses without outliers because there were no outliers in the variables in this study that interacted significantly with assumed aid. These are parallel to the multiple-regression models presented earlier in the supplementary (section 6c), offering evidence for the effect of condition and the interactions between condition and assumed aid by the other source (human/AI).

i. <u>Permutation analyses without assuming normality - General empathy item</u>

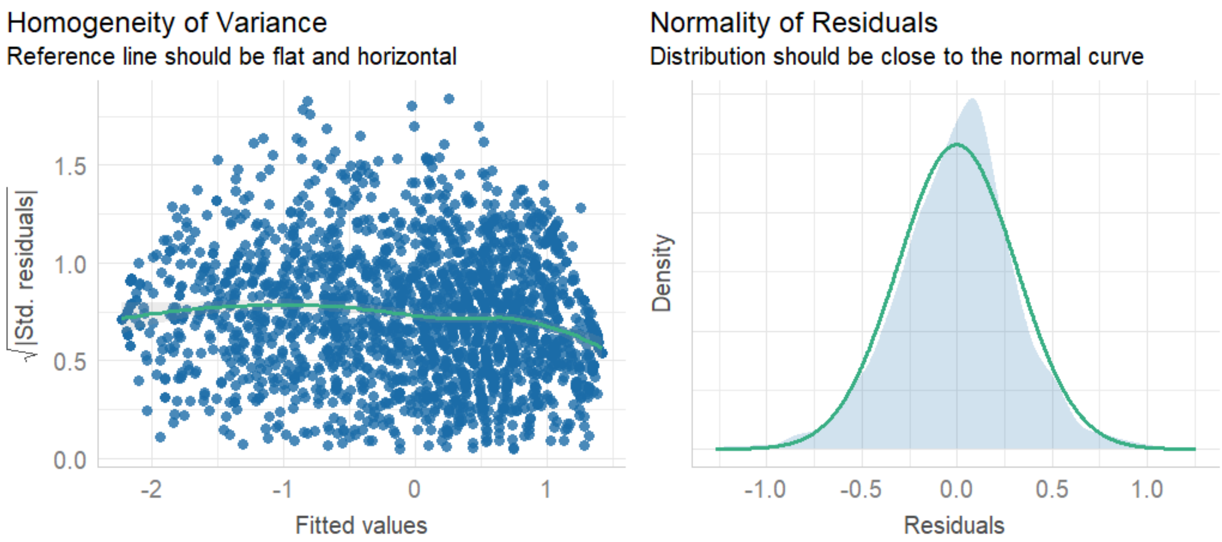
The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .95, *p* < .001).



We ran 10,000 permutations to test the model, in an analysis identical to the one in study 1b. The permutation test showed that our main effect for condition remained significant (F(1, 683) = 68.77, permutation *p* < .001), as did the interaction with assumed aid by the other source (F(1, 683) = 21.61, permutation *p* < .001). The permutations' contrasts showed that the negative effect of assumed aid on perceived empathy in the human condition (t(683) = -2.93, permutation p = .002) and the positive effect in the AI condition (t(683) = 3.63, permutation p < .001) both remained significant.

ii. Permutation analyses - Mixed model for multiple empathy types with assumed aid

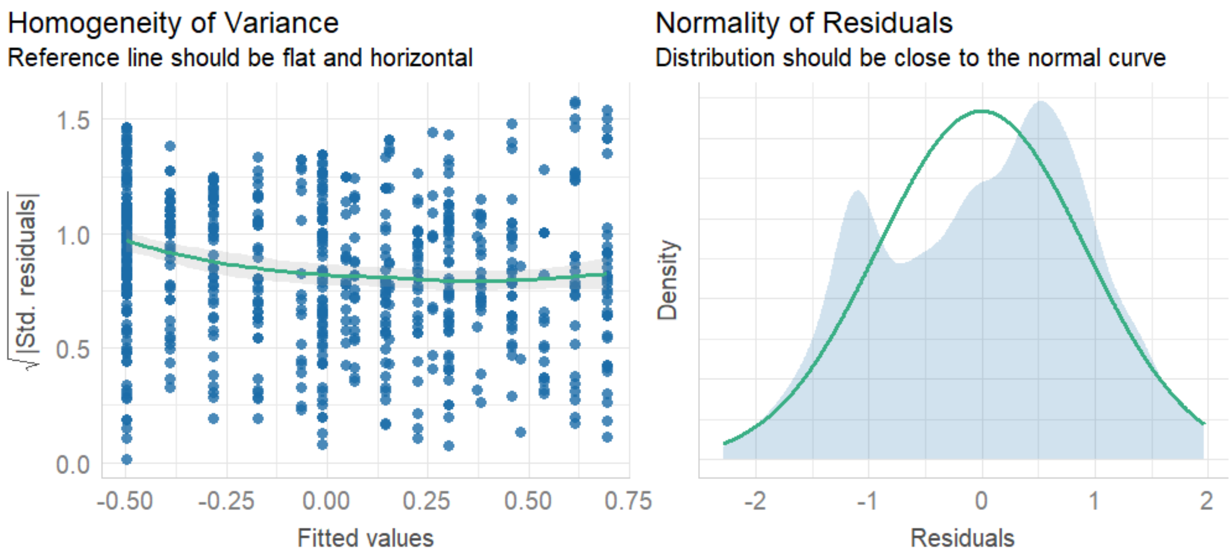
The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .99, p < .001).



We ran a permutation test identical to the one in study 1b, using 10,000 permutations, including the assumed aid by the other source (human/AI). The test showed that even using a permutation-based distribution, the effect of condition on perceived empathy remained significant (F(1, 683) = 42.45, permutation p < .001), and the interaction with assumed aid by the other source also remained significant (F(1, 683) = 43.91, permutation p < .001). Contrasts showed that the negative effect of assumed aid by the other source in the human condition and its positive effect in the AI condition were significant for all aspects of empathy (all permutation ps < .001).

iii. Permutation analyses - Positivity resonance

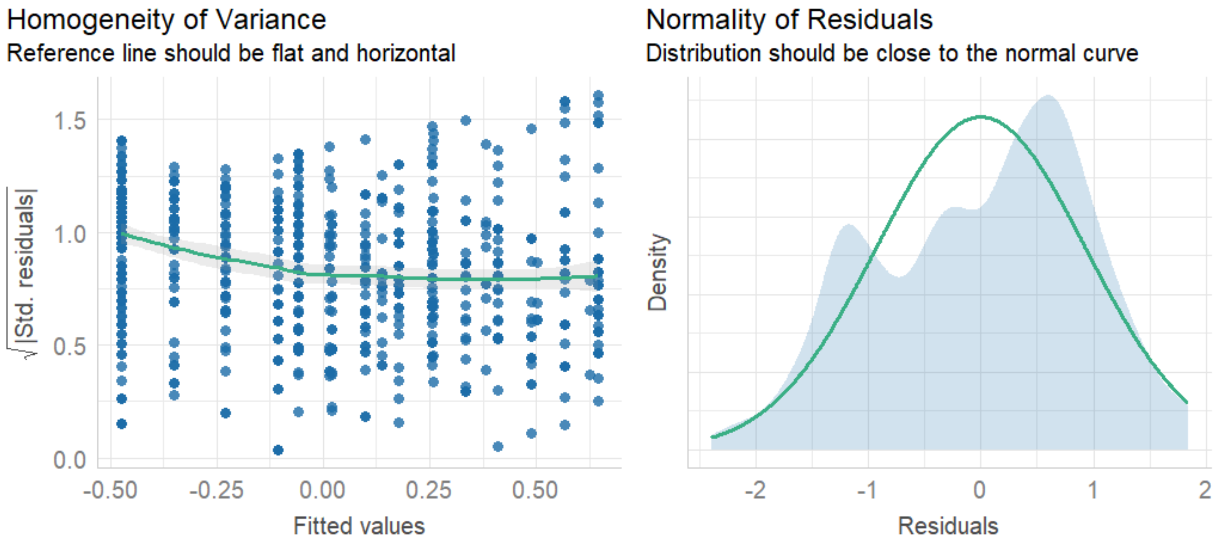
The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .98, $p < .001$).



We ran a permutation test identical to the one in study 1b, using 10,000 permutations, including the assumed aid by the other source. The test showed that even using a permutation-based distribution, the effect of condition remained significant ($F(1, 683) = 76.54$, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant ($F(1, 683) = 45.43$, permutation $p < .001$). Permutation contrasts showed that the positive effect on positivity resonance in the AI condition ($t(683) = 4.81$, permutation $p < .001$) and the negative effect in the human condition ($t(683) = -4.88$, permutation $p < .001$) both remained significant.

iv. <u>Permutation analyses - Positive emotions</u>

The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .98, $p < .001$).

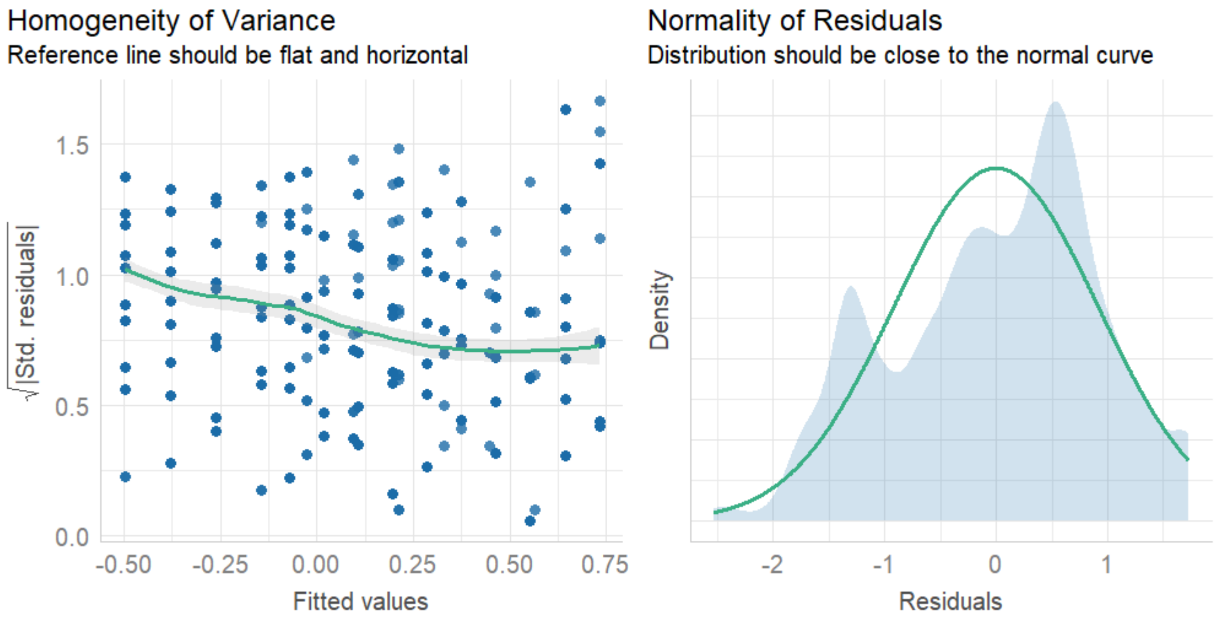


We ran a permutation test identical to the one in study 1b, using 10,000 permutations, including the assumed aid by the other source. The test showed that even using a permutation-based distribution, the effect of condition on positive emotions remained significant (F(1, 683) = 53.91, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant (F(1, 683) = 51.24, permutation $p < .001$). Permutation contrasts showed that the positive effect on positive emotions of assumed aid in the AI condition (t(683) = 5.34, permutation $p < .001$) and the negative effect in the human condition (t(683) = -4.85, permutation $p < .001$) both remained significant.

v. <u>Permutation analyses - Authenticity</u>

a. Permutation analyses without assuming normality

The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .98, $p < .001$).

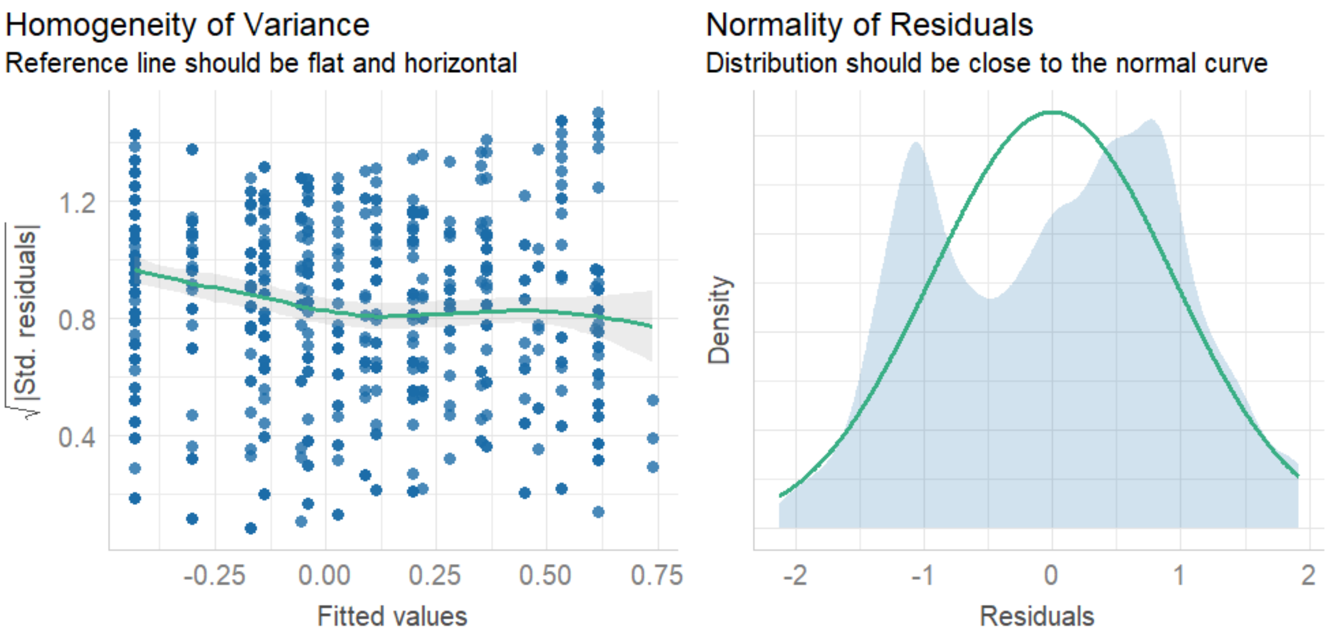


We ran a permutation test identical to the one in study 1b, using 10,000 permutations, including also the assumed aid by the other source. The test showed that even using a permutation-based distribution, the effect of condition on perceived authenticity remained significant ($F(1, 683) = 70.00$, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant ($F(1, 683) = 56.67$, permutation $p < .001$). Permutation contrasts showed that the positive effect of assumed aid on perceived authenticity in the AI condition ($t(683) = 5.25$, permutation $p < .001$) and the negative effect in the human condition ($t(683) = -5.63$, permutation $p < .001$) both remained significant.

vi. <u>Permutation analyses - Support</u>

a. Permutation analyses without assuming normality

The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: $W = .97$, $p < .001$).



We ran a permutation test identical to the one in study 1b, using 10,000 permutations, including the assumed aid by the other source (human/AI). The test showed that even using a permutation-based distribution, the effect of condition on perceived support from the response remained significant ($F(1, 683) = 32.15$, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant ($F(1, 683) = 57.35$, permutation $p < .001$). Permutation contrasts showed that the positive effect of assumed aid on perceived support in the AI condition ($t(683) = 5.63$, permutation $p < .001$) and the negative effect in the human condition ($t(683) = -5.16$, permutation $p < .001$) both remained significant.

8. Detailed analyses for study 2b

This study aimed to rule out the same halo effect as study 2a, but by lengthening the time participants were asked to wait for the response. Participants had to wait 3 minutes for their response instead of 60 seconds.

a. Analyses for empathy

i. Main analyses

We repeated the same analyses for this study as we did for studies 1b and 2a. We found a significant difference in the single-item empathy question between conditions (Welch t-test: $t(636.24) = -6.07$, $p < .001$, Cohen's $d = .47$, 95% CI [-.89,-.45]).

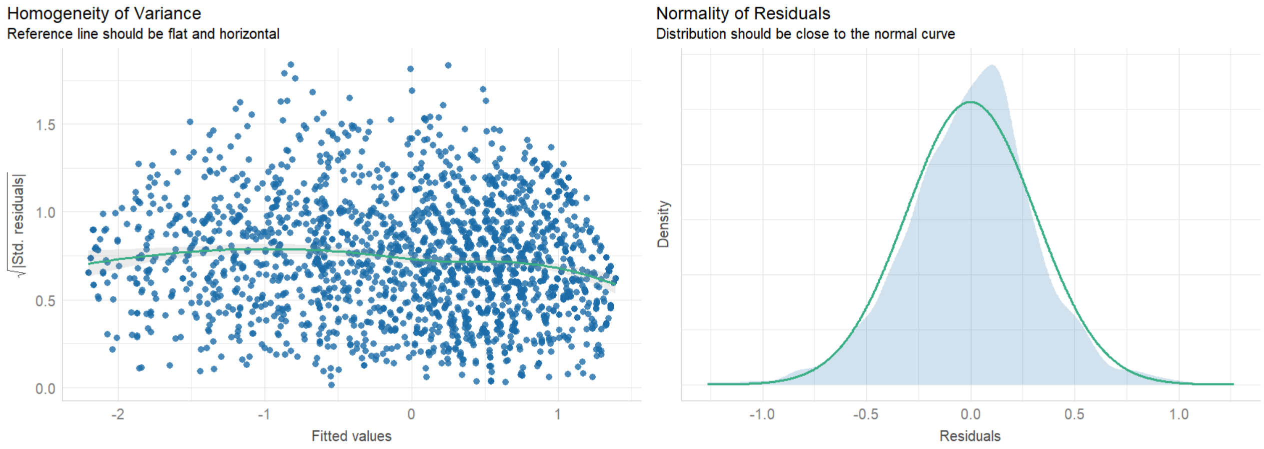
As in the previous studies, a linear mixed-effect model predicting empathy by the experimental condition, empathy type and the interaction between them found a significant main effect of condition ($F(1, 670.47) = 78.56$, $p < .001$, partial $\eta^2 = .10$), with participants in the human condition rating responses as more empathic than those in the AI condition. The model also revealed a significant interaction between condition and empathy type ($F(2, 1329.82) = 3.46$, $p = .03$, partial $\eta^2 = .005$), showing that the effect of condition on perceived motivational empathy was significantly larger than the effect on perceived cognitive empathy ($t(1341) = 2.55$, $SE = .05$, $p = .03$, $\beta = .12$, 95% CI = [.007, .23]).

ii. Models without filtering outliers

We fitted a linear mixed-effect model to predict empathy with condition, aspect of empathy (cognitive, affective or motivational) and their interaction, to the full dataset. The model showed a significant main effect of condition ($F(1, 689) = 82.34$, $p < .001$, partial $\eta^2 = .11$). The interaction between condition and empathy type was marginally significant ($F(2, 1378) = 2.76$, $p = .06$, partial $\eta^2 = .004$) and in the same direction—showing that the difference between conditions was larger for motivational than for cognitive empathy ($t(1378) = 2.28$, $SE = .05$, $p = .07$, $\beta = .11$, 95% CI = [-.00, .22]). This reinforces our central results of a significant difference between the two conditions across all participants.

iii. Permutation analyses without assuming normality

The figure below illustrates the level of deviation from normality in the data, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .99, *p* < .001) but are not extremely skewed.



We ran a permutation test identical to the one in study 1b, using 10,000 permutations. The test showed that even using a permutation-based distribution, the effect of condition on perceived empathy remained significant (F(1, 685) = 68.59, permutation *p* < .001).

b. <u>Additional effects of condition</u>

Condition had a significant effect on positivity resonance (Welch t-test: t(642.35) = -7.56, *p* < .001, Cohen's *d* = .58, 95% CI [-17.6, -10.34]). This replicates our earlier findings that participants experienced greater positivity in the communication when perceiving the response as human.

We also found a significant difference between conditions in the level of authenticity participants attributed to the response (Welch t-test: t(636.33) = -5.88, *p* < .001, Cohen's *d* = .46, 95% CI [-.50, -.25]). As in the previous studies, participants found the response more authentic when they perceived it as human.

There was also a significant difference in perceived support between conditions ($t(680.00) = -4.95$, $p < .001$, Cohen's d =.38, 95% CI [-1.35, -.58]). This replicated the finding that participants perceived the response as more supportive when they believed it was human-authored.

Responses that were presented as human raised significantly more positive emotions than those presented as AI-generated ($t(638.28) = -6.02$, $p < .001$, Cohen's d = .47, 95% CI [-1.21, -.61]). Like previously, participants felt more positive emotions when perceiving the response as human.

There was a significant difference in the levels of negative emotions caused by the response ($t(602) = 4.81$, $p < .001$, Cohen's d =.37, 95% CI [.19, .45]). Participants felt fewer negative emotions following the response when perceiving it as human-authored.

c. Interactions with assumed aid by the other source

We aimed to replicate our findings from the previous studies, in which assumed aid by the other source decreased empathy in the human condition and increased it in the AI condition. A multiple-regression model predicting perceived empathy from the condition, assumed aid and their interaction found that alongside the significant main effect of condition ($F(1, 662) = 38.88$, $p < .001$, partial η^2=.06), there was a significant interaction between condition and assumed aid ($F(1, 662) = 32.15$, $p < .001$, partial η^2=.05). As in studies 1b and 2a, post-hoc contrasts revealed a positive effect of assumed aid in the AI condition ($t(662) = 3.52$, SE = .05, $p < .001$, $\beta = .16$, 95% CI = [.07, .25]) but a negative one in the human condition ($t(662) = -4.59$, SE = .04, $p < .001$, $\beta = -.18$, 95% CI = [-.26, -.11]). This suggests that, overall, participants perceived greater empathy the more they believed a human was involved in the response, and lower empathy when they believed AI was involved in the interaction.

We then used a linear mixed-effect model including empathy type, condition, assumed aid and their interactions to test whether these effects differed between aspects of empathy. The effect of condition remained significant ($F(1, 666.40) = 76.37$, $p < .001$, partial $\eta^2 =.10$). We again found a significant interaction between condition and assumed aid ($F(1, 667.31) = 39.55$, $p < .001$, partial $\eta^2 = .06$). The interaction of empathy type and condition also remained significant ($F(2, 1323.86) = 3.40$, $p = .03$, partial $\eta^2 = .005$), and there was a significant three-way interaction between condition, empathy type and assumed aid ($F(2, 1324.81) = 4.06$, $p = .02$, partial $\eta^2 = .006$). Post-hoc contrasts revealed significant negative effects of assumed aid in the human condition on all aspects of empathy (cognitive: $t(979) = -4.68$, $SE = .05$, $p < .001$ $\beta = -.22$, 95% CI = [-.31, -.13]; affective: $t(979) = -5.23$, $SE = .05$, $p < .001$, $\beta = -.25$, 95% CI = [-.34, -.15]; motivational: $t(974) = -4.49$, $SE = .05$, $p < .001$, $\beta = -.21$, 95% CI = [-.30, -.12]) and a significant positive effect on them in the AI condition (cognitive: $t(980) = 3.25$, $SE = .05$, $p = .001$, $\beta = .17$, 95% CI = [.07, .28]; affective: $t(988) = 4.47$, $SE = .05$, $p < .001$, $\beta = .24$, 95% CI = [.13, .34]; motivational: $t(982) = 2.40$, $SE = .05$, $p < .001$, $\beta = .13$, 95% CI = [.02, .23]). We also found that the increase in affective empathy caused by assumed aid in the AI condition was significantly greater than the increase in motivational empathy in the AI condition ($t(1338) = 2.85$, $SE = .04$, $p = .03$, $\beta = .11$, 95% CI = [.008, .21]).

We also replicated our finding of a significant interaction between condition and assumed aid on positivity resonance ($F(1, 671) = 77.07$, $p < .001$, partial $\eta^2 = .10$), with post-hoc contrasts showing that assumed aid had a positive effect in the AI condition, increasing positivity resonance ($t(671) = 6.09$, $SE = .05$, $p < .001$, $\beta =.33$, 95% CI = [.22, .43]), but a negative effect in the human condition, decreasing positivity resonance ($t(671) = -6.37$, $SE = .05$, $p < .001$, $\beta = -.30$, 95% CI = [-.40, -.21]).

We again found a similar interaction for positive emotions ($F(1, 661) = 64.99$, $p < .001$, partial $\eta^2 = .06$), with assumed aid having a positive effect in the AI condition ($t(661) = 5.23$, $SE = .05$, $p < .001$, $\beta = .27$, 95% CI = [.17, .37]) and a negative effect in the human condition ($t(661) = -6.25$, $SE = .05$, $p < .001$, $\beta = -.28$, 95% CI = [-.37, -.19]). Perceived human involvement increased the level of positive emotions, whereas perceived AI involvement decreased this variable.

Authenticity was influenced similarly, with assumed aid interacting significantly with condition ($F(1, 653) = 57.26$, $p < .001$, partial $\eta^2 = .08$), having a positive effect in the AI condition $t(653) = 3.98$, $SE = .05$, $p < .001$(, $\beta = .20$, 95% CI = [.10, .30]) and a negative effect in the human condition ($t(653) = -6.94$, $SE = .04$, $p < .001$, $\beta = -.30$, 95% CI = [-.39, -.22]). Responses that were perceived as more human were considered more authentic than those perceived as involving AI.

The same interaction was significant when predicting support perceived in the response ($F(1, 687) = 97.23$, $p < .001$, partial $\eta^2 = .12$), where assumed aid positively influenced it in the AI condition ($t(687) = 6.96$, $SE = .06$, $p < .001$, $\beta = .40$, 95% CI = [.29, .51]), but the influence was negative in the human condition ($t(687) = -7.02$, $SE = .05$, $p < .001$, $\beta = -.36$, 95% CI = [-.46, -.26]). This suggests that responses perceived as involving human input were perceived as more supportive, whereas perceived AI involvement lowered feelings of support.

These results replicate most of the findings from studies 1a, 1b and 2a, showing that our effects remain the same even when participants were required to wait 3 minutes for a response.

d. <u>Extended results on assumed aid by the other source</u>

We provide here further analyses without filtering outliers, and separate analyses using permutations without assuming normality. These are parallel to the multiple-regression models presented earlier in the supplementary (section 7c), offering evidence for the effect of condition and the interactions between condition and assumed aid.

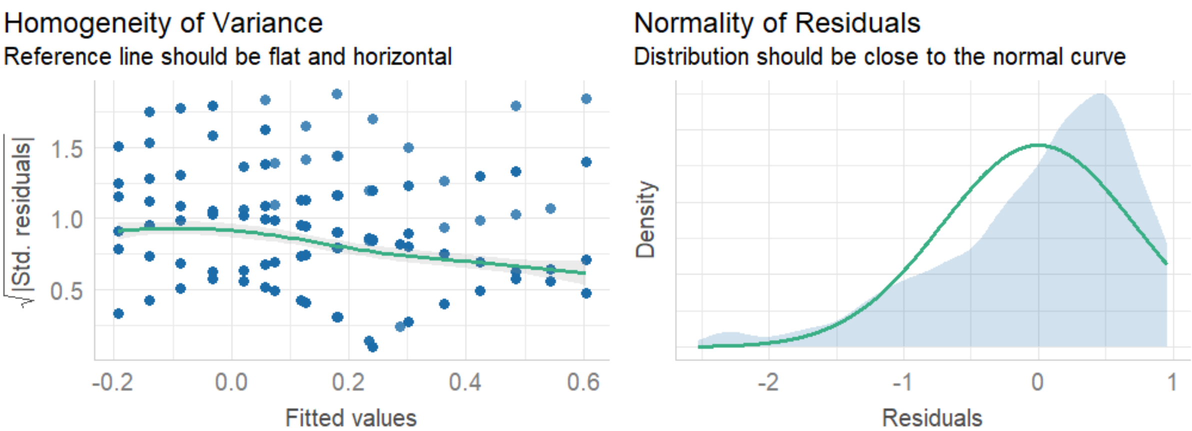
i. <u>General empathy item</u>

a. Analysis without filtering outliers

We fitted a linear model predicting the general empathy rating from condition, assumed aid by the other source and their interaction. Both the effect of condition ($F(1, 687) = 49.29$, $p < .001$, partial $\eta^2 = .07$) and the interaction effect remained significant ($F(1, 687) = 39.24$, $p < .001$, partial $\eta^2 = .05$). Contrasts showed that the negative effect of assumed aid on perceived empathy in the human condition ($t(687) = -3.03$, $SE = .05$, $p = .002$, $\beta = -.16$, 95% CI = [-.26, -.06]) and the positive effect of assumed aid in the AI condition ($t(687) = 5.69$, $SE = .06$, $p < .001$, $\beta = .34$, 95% CI = [.22, .45]) both remained significant.

b. Permutation analyses without assuming normality

The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: $W = .92$, $p < .001$).



We ran 10,000 permutations to test the model predicting the general empathy rating from the condition, assumed aid by the other source and their interaction. The permutation test showed that our main effect for condition remained significant ($F(1, 662) = 38.88$, permutation $p < .001$), as did the interaction with assumed aid by the other source ($F(1, 662) = 32.15$, permutation $p < .001$). The permutations' contrasts showed that the negative effect of assumed aid on empathy in the human condition ($t(662) = -4.59$, permutation $p < .001$) and the positive effect in the AI condition ($t(662) = 3.52$, permutation $p < .001$) both remained significant.

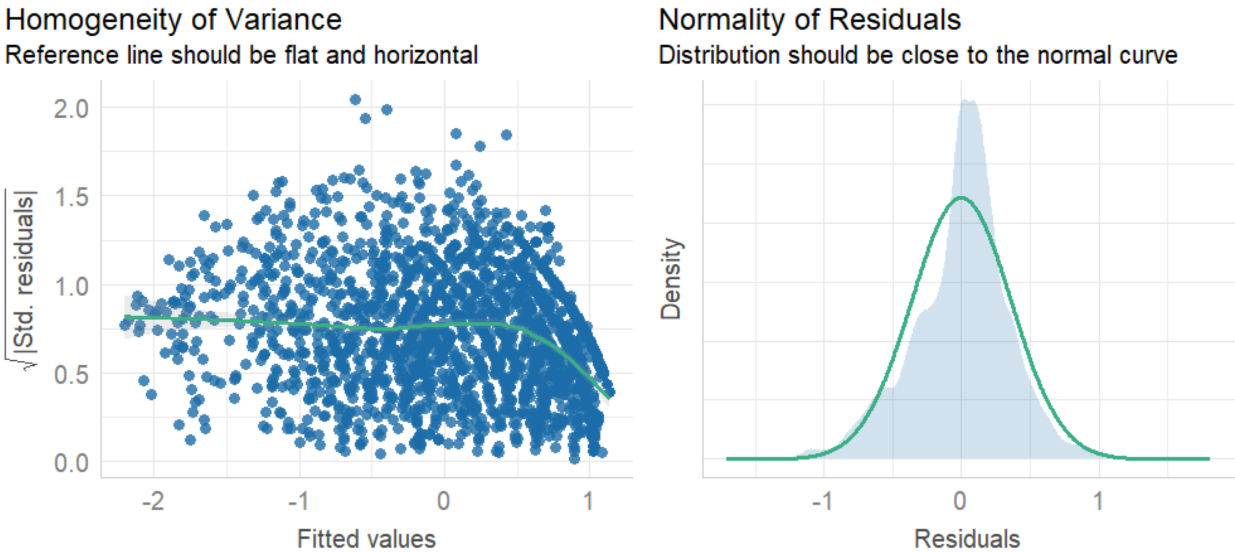
ii. <u>Mixed model for multiple empathy types</u>

a. Analysis without filtering outliers

We fit a linear mixed-effect model including empathy type, condition, assumed aid and their interactions to the full dataset. The model found that the effect of condition remained significant ($F(1, 687) = 71.44$, $p < .001$, partial $\eta^2 = .09$). We also replicated the significant interaction between condition and assumed aid ($F(1, 687) = 49.62$, $p < .001$, partial $\eta^2 = .07$). The interaction of empathy type and condition was found to be marginally significant ($F(2, 1374) = 2.82$, $p = .06$, partial $\eta^2 = .004$). The three-way interaction between condition, empathy type and assumed aid ($F(2, 1374) = 3.07$, $p = .05$, partial $\eta^2 = .004$) also remained significant. Post-hoc contrasts revealed significant negative effects of assumed aid in the human condition on all aspects of empathy (cognitive: $t(943) = -4.49$, $SE = .05$, $p < .001$, $\beta = -.23$, 95% CI = [-.33, -.13]; affective: $t(943) = -4.88$, $SE = .05$, $p < .001$, $\beta = -.25$, 95% CI = [-.35, -.15]; motivational: $t(943) = -4.40$, $SE = .05$, $p < .001$, $\beta = -.23$, 95% CI = [-.33, -.13]), and a significant positive effect on them in the AI condition (cognitive: $t(943) = 4.48$, $SE = .06$, $p < .001$, $\beta = .26$, 95% CI = [.15, .37]; affective: $t(943) = 5.55$, $SE = .06$, $p < .001$, $\beta = .25$, 95% CI = [.21, .43]; motivational: $t(943) = 3.75$, $SE = .06$, $p < .001$, $\beta = .22$, 95% CI = [.10, .33]). The increase in affective empathy caused by assumed aid in the AI condition remained significantly greater than the increase in motivational empathy in the AI condition (t(1374) = 2.64, SE = .04, p = .05, β = .10, 95% CI = [.015, .06]).

b. Permutation analyses for interaction model with assumed aid

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .98, p < .001) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations of a model predicting empathy from condition, empathy type, assumed aid by the other source and their interactions. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 666.40) = 76.37, permutation p < .001), and the interaction with assumed aid by the other source also remained significant (F(1, 667.31) = 39.55, permutation p < .001). Contrasts showed that the negative effect of assumed aid by the other source in the human condition and the positive effect in the AI condition were significant for all aspects of empathy (permutation *ps* for all aspects in the human condition < .001; permutation *ps* for the cognitive and affective aspects in the AI condition < .001; permutation *p* for the motivational aspect in the AI condition = .009).

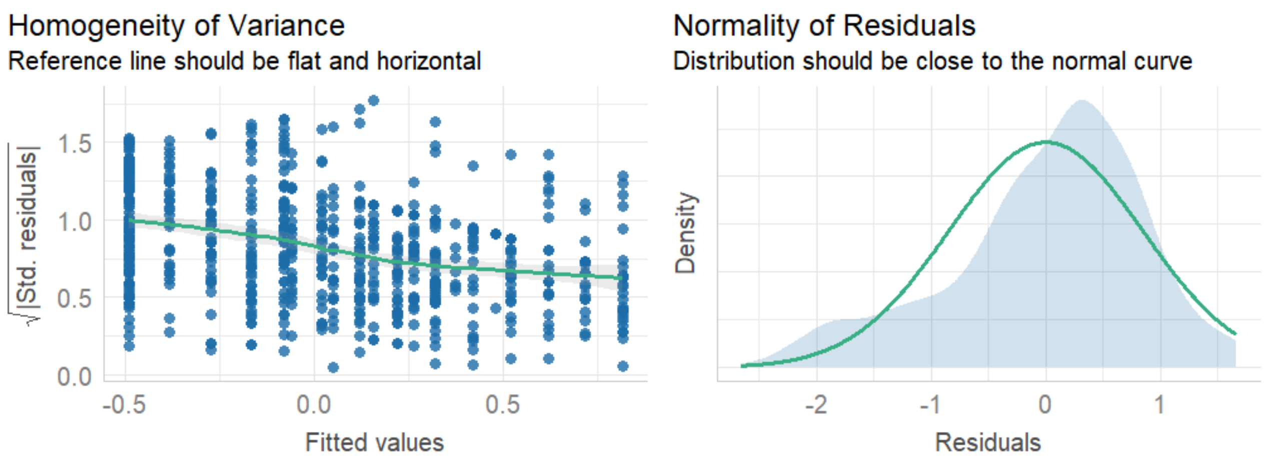
iii. Positivity resonance

a. Analysis without filtering outliers

We fitted a linear model predicting positivity resonance from condition, assumed aid by the other source and their interaction. Both the effect of condition ($F(1, 687) = 67.90$, $p < .001$, partial $\eta^2 = .09$) and the interaction effect remained significant ($F(1, 687) = 85.93$, $p < .001$, partial $\eta^2 = .11$). Contrasts showed that the negative effect of assumed aid in the human condition ($t(687) = -6.07$, $SE = .05$, $p < .001$, $\beta = -.31$, 95% CI = [-.40, -.21]) and the positive effect in the AI condition ($t(687) = 7.01$, $SE = .06$, $p < .001$, $\beta = .40$, 95% CI = [.29, .51]) both remained significant.

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: $W = .96$, $p < .001$) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations to test a model predicting positivity resonance from condition, assumed aid by the other source and their interaction.

The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 671) = 63.64, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant (F(1, 671) = 77.07, permutation $p < .001$). Permutation contrasts showed that the positive effect on positivity resonance in the AI condition (t(671) = 6.09, permutation $p < .001$) and the negative effect in the human condition (t(671) = -6.37, permutation $p < .001$) both remained significant.

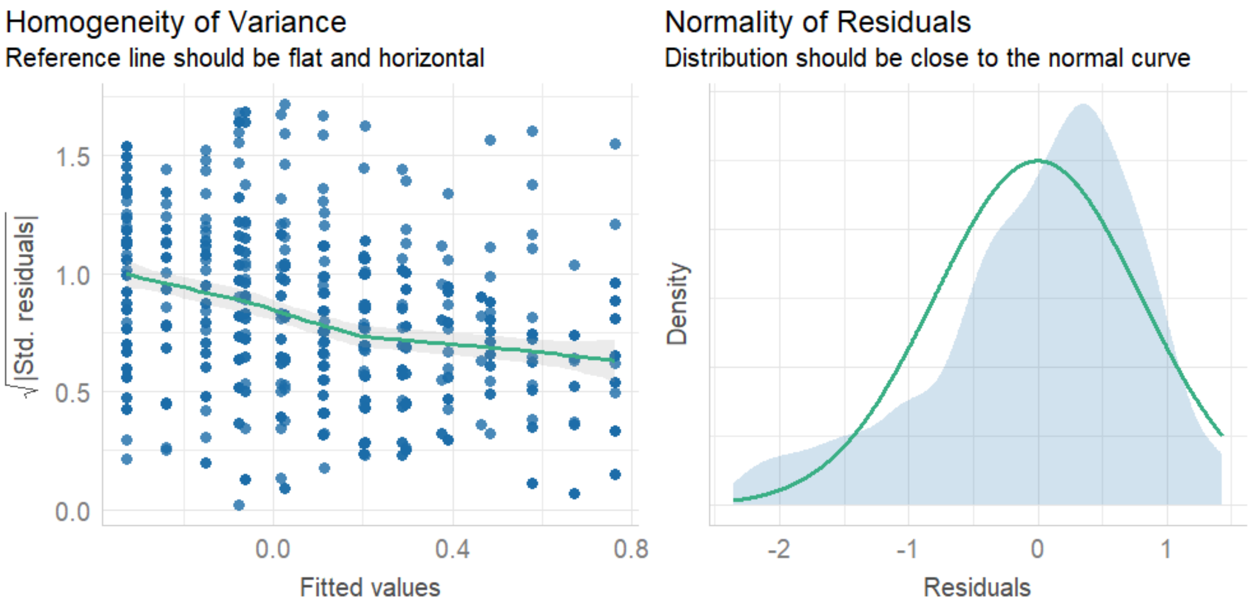
iv. Positive emotions

a. Analysis without filtering outliers

We fitted a linear model predicting positive emotions from condition, assumed aid by the other source and their interaction. Both the effect of condition (F(1, 687) = 36.42, $p < .001$, partial $\eta^2 = .05$) and the interaction effect remained significant (F(1, 687) = 82.56, $p < .001$, partial $\eta^2 = .11$). Contrasts showed that the negative effect of assumed aid on positive emotions in the human condition (t(687) = -6.58, SE = .05, $p < .001$, $\beta = -.34$, 95% CI = [-.44, -.24]) and its positive effect in the AI condition (t(687) = 6.31, SE = .06, $p < .001$, $\beta = .37$, 95% CI = [.25, .48]) remained significant.

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .95, $p < .001$) but are not extremely skewed.



We ran a permutation test identical to the one in study 1b, using 10,000 permutations, including the assumed aid by the other source. The test showed that even using a permutation-based distribution, the effect of condition on positive emotions remained significant (F(1, 661) = 39.78, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant (F(1, 661) = 64.99, permutation $p < .001$). Permutation contrasts showed that the positive effect of assumed aid on positive emotions in the AI condition (t(661) = 5.23, permutation $p < .001$) and the negative effect in the human condition (t(661) = -6.25, permutation $p < .001$) both remained significant.

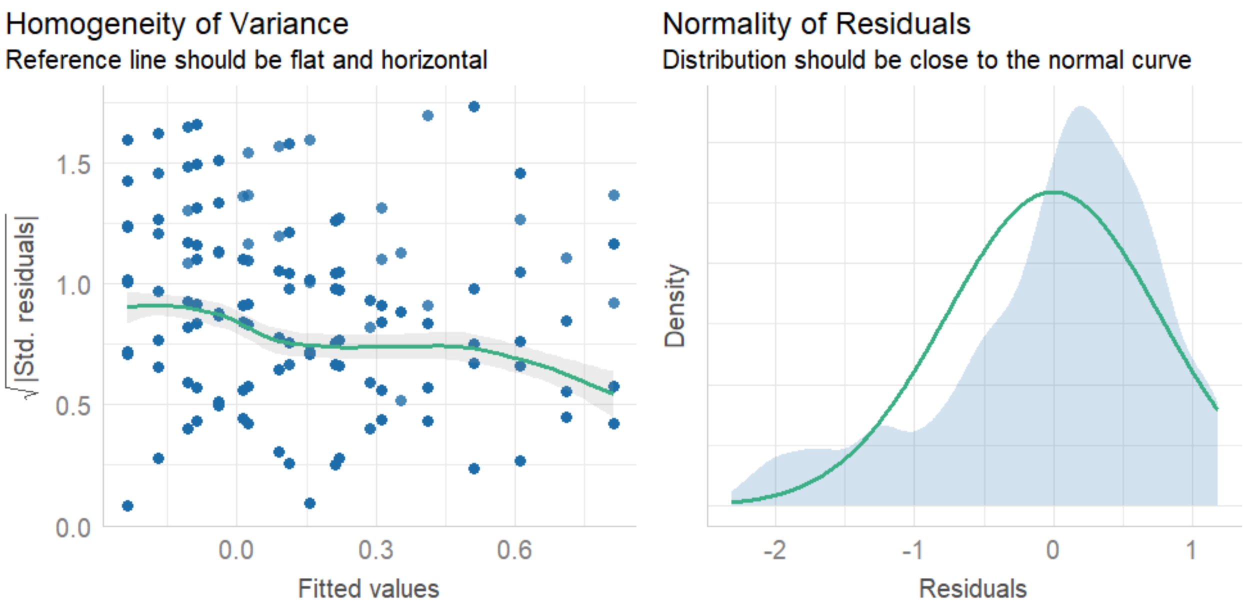
v. Authenticity

a. Analysis without filtering outliers

We fitted a linear model predicting authenticity from condition, assumed aid by the other source and their interaction. Both the effect of condition (F(1, 687) = 53.45, $p < .001$, partial $\eta^2 = .07$) and the interaction effect remained significant (F(1, 687) = 83.03, $p < .001$, partial $\eta^2 = .11$). Contrasts showed that the negative effect of assumed aid on authenticity in the human condition (t(687) = -6.83, SE = .05, $p < .001$, $\beta = -.35$, 95% CI = [-.45, -.25]) and the positive effect in the AI condition (t(687) = 6.12, SE = .06, $p < .001$, β = .35, 95% CI = [.24, .46]) remained significant.

b. Permutation analyses without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .94, $p < .001$) but are not extremely skewed.

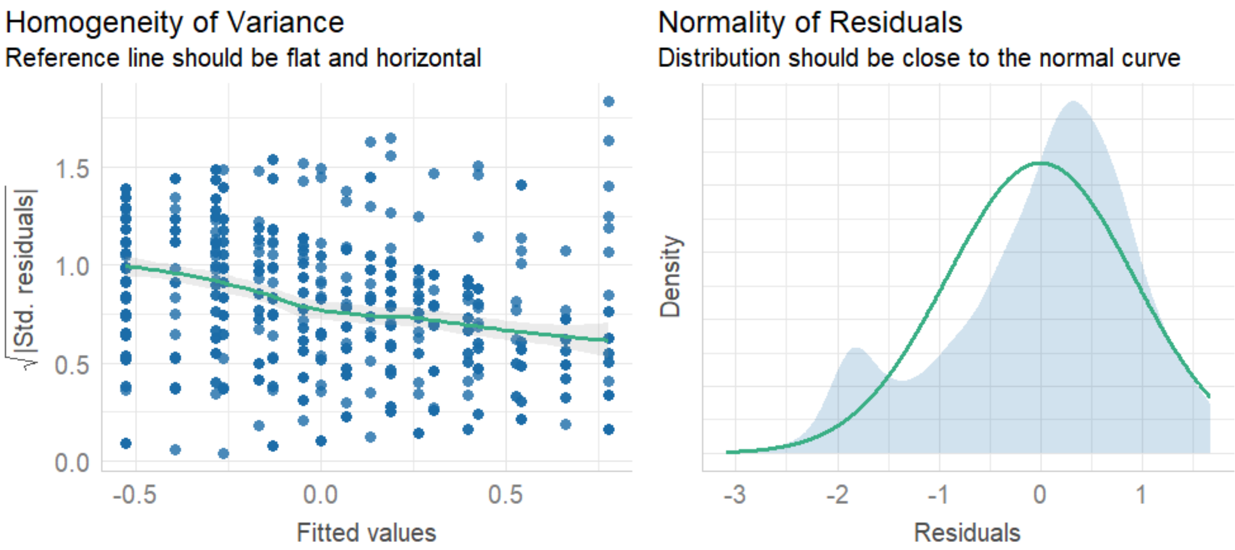


We ran a permutation test, using 10,000 permutations of a model predicting authenticity from condition, assumed aid by the other source and their interaction. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 653) = 38.02, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant (F(1, 653) = 57.26, permutation $p < .001$). Permutation contrasts showed that the positive effect of assumed aid on authenticity in the AI condition (t(653) = 3.98, permutation $p < .001$) and the negative effect in the human condition (t(653) = -6.94, permutation $p < .001$) both remained significant.

vi. <u>Support</u>

a. Permutation analyses without assuming normality

The figure below illustrates the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .96, $p < .001$).



We ran a permutation test, using 10,000 permutations of a model predicting support from condition, assumed aid by the other source and their interaction. The test showed that even using a permutation-based distribution, the effect of condition on perceived support remained significant ($F(1, 687) = 27.85$, permutation $p < .001$), and the interaction with assumed aid by the other source also remained significant ($F(1, 687) = 97.23$, permutation $p < .001$). Permutation contrasts showed that the positive effect of assumed aid on perceived support in the AI condition ($t(687) = 6.96$, permutation $p < .001$) and the negative effect in the human condition ($t(687) = -7.02$, permutation $p < .001$) both remained significant.

9. Detailed analyses for study 3

a. Comparisons between response types

We first report the full comparisons among the ratings for each type of empathy within each response type.

Supplementary Table 3

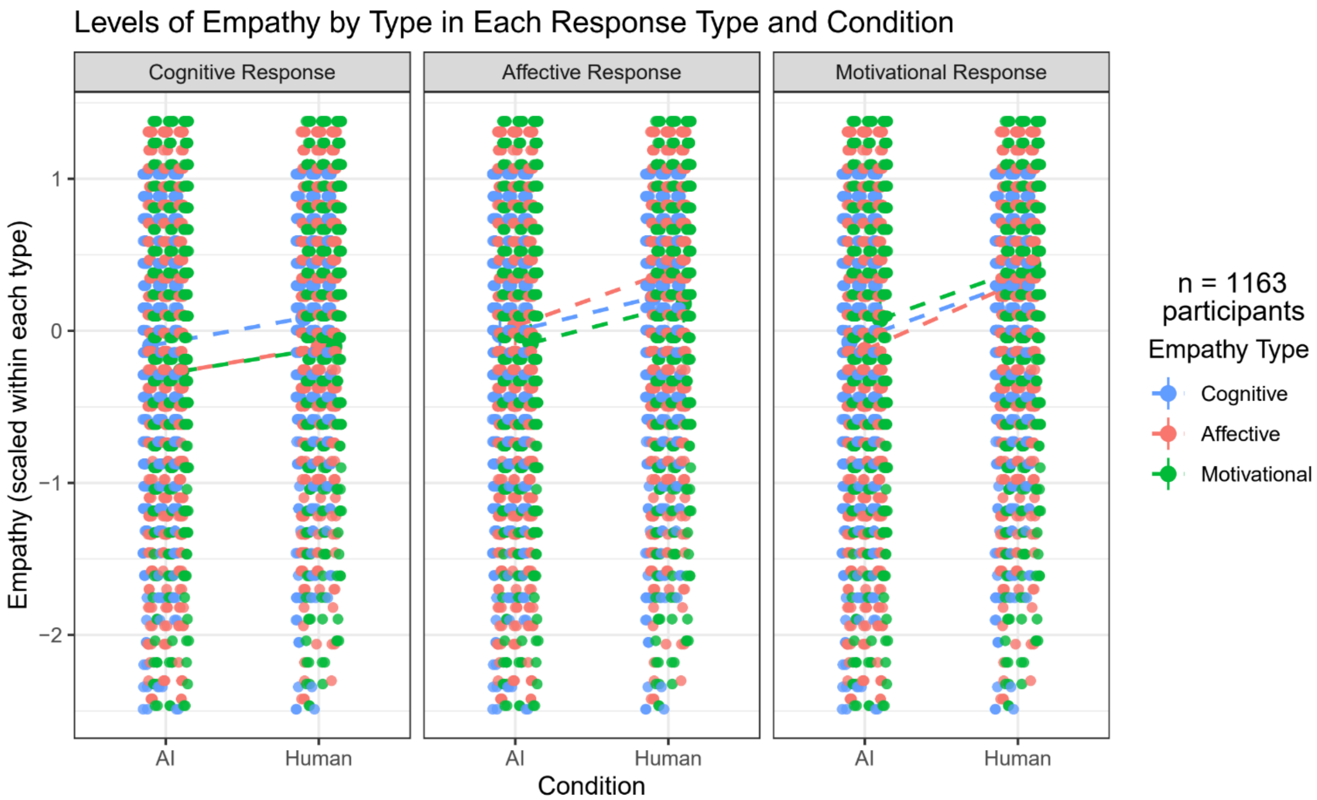
Differences among aspects of empathy by response type

Response Type	Contrast	β	SE	CI	df	t	*p*
Cognitive	Cognitive - Affective	.19	.03	[.11, .27]	2270	5.76	<.001***
	Cognitive - Motivational	.18	.03	[.10, .26]	2269	5.43	<.001***
	Affective - Motivational	-.01	.03	[-.09, .07]	2270	-.36	1
Affective	Cognitive - Affective	-.11	.03	[-.19, -.04]	2265	-3.52	.001**
	Cognitive - Motivational	.05	.03	[-.03, .12]	2264	1.47	.43
	Affective - Motivational	.16	.03	[.08, .24]	2263	5.01	<.001***
Motivational	Cognitive - Affective	.01	.03	[-.07, .09]	2269	.30	1
	Cognitive - Motivational	-.15	.03	[-.23, -.07]	2269	-4.56	<.001***
	Affective - Motivational	-.16	.03	[-.24, -.08]	2268	-4.90	<.001***

Linear pairwise contrasts were used to compare the rating of each aspect of empathy within each response type. asterisks represent different levels of significance: *** - $p < .001$; ** - $p < .01$; * - $p < .05$.

b. <u>Figure 5B – with data points</u>

To illustrate the distribution of data in figure 4B, we provide it here with the specific data points. See supplementary section 9a for details on the comparisons.



c. <u>Additional effects of condition and response types</u>

The analyses were all conducted using multiple-regression models, similar to the analysis of empathy, while including condition and response type in the model and predicting the dependent variables described below.

Similar to empathy, positivity resonance showed significant differences between conditions and response types. Contrasts revealed that the cognitive responses were rated lower for positivity resonance than affective (t(1144) = -4.01, SE = .07, $p < .001$, β = -.27, 95% CI = [-43, -.11]) and motivational responses (t(1144) = -5.39, SE = .07, $p < .001$, β = -.36, 95% CI = [-.52, -.20]). We also found a significant difference between conditions overall (t(1144) = -5.64, SE = .05, $p < .001$, β = -.15, 95% CI = [-.21, -.10]), with responses perceived to be human rated higher for positivity resonance. However, we found that these differences remained significant only in the affective responses (t(1144) = -3.39, SE = .09, $p < .001$, β = -.32, 95% CI = [-.50, -.13]) and motivational responses (t(1144) = -4.73, SE = .09, $p < .001$, β = -.44, 95% CI = [-.63, -.26]) and not in the cognitive responses ($p = .10$).

Positive emotions were significantly affected by condition (t(1120) = 4.64, SE = .02, $p < .001$, β = .11, 95% CI = [.07, .16]) such that the responses deemed to come from a human source raised more positive emotions. This was also affected by response type (F(2, 1120) = 4.29, $p = .01$), with cognitive responses being rated lower for positive emotions than affective responses (t(1120) = -2.51, SE = .06, $p = .04$, β = -.15, 95% CI = [-.30, -.007]) and motivational responses (t(1120) = 2.59, SE = .06, $p = .03$, β = -.16, 95% CI = [-.30, -.01]). There was no significant interaction between condition and response type.

Negative emotions too were influenced significantly by condition (t(1112) = 5.27, $p < .001$, β = -.10, 95% CI [-.13, -.06]), with responses deemed human triggering less negative emotions. There was no significant effect of response type, nor an interaction with the condition.

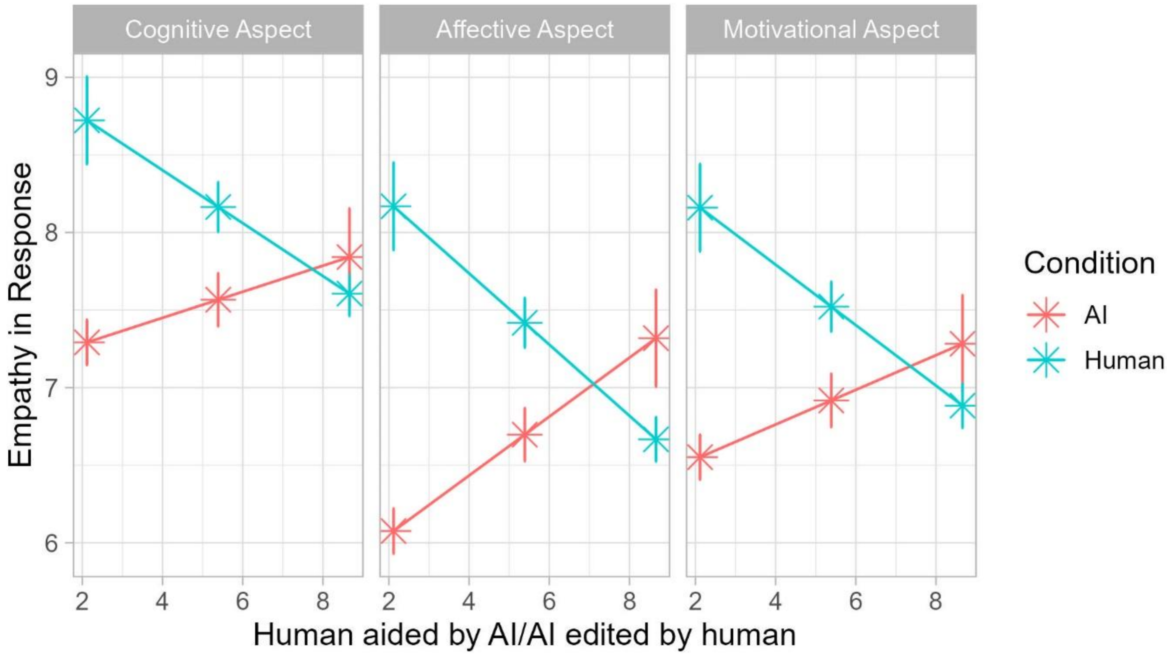
A separate model found a similar effect of condition on the perceived authenticity of the response (t(1111) = 2.20, SE = .03, $p = .03$, β = .06, 95% CI = [.01, .10]), with responses that seem human rated as more authentic. We also found a significant effect of response type, with contrasts revealing that motivational responses were rated as more authentic than cognitive responses (t(1111) = -3.54, SE = .06, p = .001, β = -.22, 95% CI = [-.37, -.07]). There was no significant interaction between response type and condition.

Perceived supportiveness of the response was not influenced significantly by condition (p = .22). However, response type did have a significant effect on support perceived in the response, such that cognitive responses were rated as less supportive than affective responses (t(1166) = -3.06, SE = .07, p = .006, β = -.22, 95% CI = [-.38, -.05]) and motivational responses (t(1166) = -5.39, SE = .07, p < .001 β = -.38, 95% CI = [-.55, -.22]); and affective responses were perceived as less supportive than motivational ones (t(1166) = -2.35, SE = .07, p = .05, β = -.17, 95% CI = [-.33, -.00]).

d. <u>Interaction effects with assumed aid by the other source</u>

Similar to study 2, we again examined the effect of assumed aid by the other source (human/AI) on many variables, to determine if the previous effects would be replicated. In order to do this, we examined linear regression models, much like in study 2. However, in study 3 we also included response type as a predictor in the model to identify any differences between the types of responses.

We replicated the moderation of the effect of condition by the assumed aid on many variables. First, there was a significant three-way interaction between condition, assumed aid and aspect of empathy (F(2, 2232.10) = 6.11, p = .002, η^2 = .005). Post-hoc contrasts showed that in the human condition, assumed aid significantly decreased cognitive (t(1694) = -7.84, SE = .04, p < .001, β = -.35, 95% CI = [-.43, -.26]), affective (t(1684) = -8.40, SE = .04, p < .001 β = -.37, 95% CI = [-.46, -.28]) and motivational empathy (t(1680) = -8.51, SE = .04, p < .001, β = -.37, 95% CI = [-.46, -.29]) in the same way, as comparisons among these effects were not significant (p = 1 for all comparisons among aspects of empathy within this condition). However, for the AI condition, assumed aid significantly increased cognitive (t(1703) = 3.48, SE = .05, $p < .001$, $\beta = .17$, 95% CI = [.07, .26]), affective (t(1701) = 6.52, SE = .05, $p < .001$, $\beta = .32$, 95% CI = [.22, .41]) and motivational empathy (t(1694) = 4.42, SE = .05, $p < .001$, $\beta = .21$, 95% CI = [.12, .31]), increasing affective empathy significantly more than cognitive (t(2259) = 4.03, SE = .04, $p < .001$, $\beta = .15$, 95% CI = [.05, .24]) or motivational empathy (t(2257) = 2.81, SE = .04, $p = .03$, $\beta = .21$, 95% CI = [.006, .20]; see supplementary figure 9A).



Supplementary figure 9A: The interaction between condition, aspect of empathy and assumed aid. There is a significant difference in the effect of assumed aid on affective empathy as opposed to cognitive or motivational empathy within the AI condition.

Positivity resonance was significantly affected by the interaction between condition and assumed aid (F(1, 1138) = 103.62, $p < .001$, $\eta^2 = .08$), regardless of response type. Post-hoc contrasts showed that assumed aid had a positive effect in the AI condition (t(1138) = 4.47, SE = .05, $p < .001$, $\beta = .22$, 95% CI = [.13, .32]) but a negative effect in the human condition (t(1138) = -10.25, SE = .05, $p < .001$, $\beta = -.46$, 95% CI = [-.55, .37]). The differences between response types also remained significant but did not significantly interact with assumed aid.

Positive emotions were influenced in a similar manner by the interaction between condition and assumed aid. Post-hoc contrasts showed that assumed aid increased positive emotions in the AI condition (t(1114) = 4.11, SE = .05, $p < .001$, $\beta = .19$, [-.10, .28]) but decreased them in the human condition (t(1114) = -9.52, SE = .04, $p < .001$, $\beta = -.39$, 95% CI = [-.47, -.31]).

The same was true of authenticity, with post-hoc contrasts showing the positive effect of assumed aid in the AI condition (t(1105) = 4.02, SE = .05, $p < .001$, $\beta = .18$, 95% CI = [.09, .27]) but a negative effect in the human condition (t(1105) = -12.17, SE = .04, $p < .001$, $\beta = -.50$, 95% CI = [-.58, -.42]).

Levels of perceived support were influenced similarly, with post-hoc contrasts showing that assumed aid had a positive effect in the AI condition (t(1160) = 4.72, SE = .05, $p < .001$, $\beta = .25$, 95% CI = [.15, .35]) but a negative effect in the human condition (t(1160) = -10.31, SE = .05, $p < .001$ $\beta = -.50$, 95% CI = [-.59, -.40]).

e. Extended results for mixed-effect models in study 3

i. General empathy item

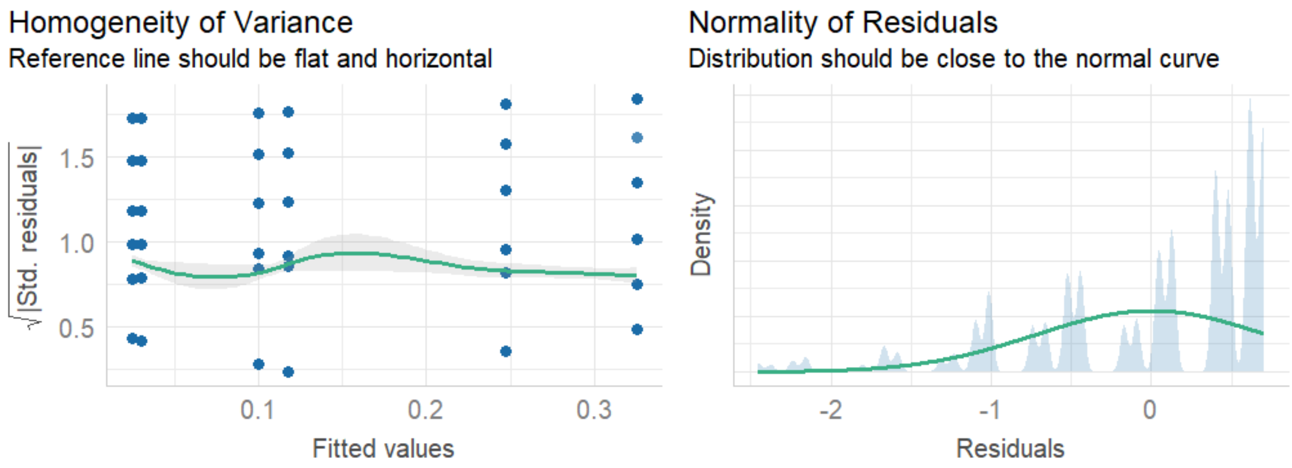
a. Analyses without filtering outliers

We fitted a linear model predicting the general empathy rating from condition, response type and their interaction. The effect of condition ($F(1, 1166) = 17.55$, $p < .001$, partial $\eta^2 = .01$) remained significant. Unlike our initial analysis, the effect of response type was found to be significant in this test ($F(1, 1166) = 4.81$, $p = .008$, partial $\eta^2 = .008$). Contrasts showed that the difference between conditions was significant in the cognitive responses (t(1166) = -2.70, SE = .10, $p = .007$, $\beta = -.27$, 95% CI = [-.47, -.07]) and the motivational responses ($t(1166) = -3.13$, $SE = .10$, $p = .002$, $\beta = -.32$, 95% CI = [-.51, -.12]) but not the affective responses ($p = .16$).

To explore the interaction with assumed aid by the other source, we added this variable to the previous model. The model revealed that the effect of condition remained significant ($F(1, 1160) = 18.86$, $p < .001$, partial $\eta^2 = .02$), and that assumed aid had a significant interaction with condition ($F(1, 1160) = 83.09$, $p < .001$, partial $\eta^2 = .07$). There was also a significant three-way interaction with response type and condition ($F(2, 1160) = 3.60$, $p = .03$, partial $\eta^2 = .006$). Contrasts revealed that assumed aid had negative effects in all response types in the human condition (cognitive: $t(1160) = -4.80$, $SE = .09$, $p < .001$, $\beta = -.43$, 95% CI = [-.61, -.26]; affective: $t(1160) = -3.39$, $SE = .08$, $p < .001$, $\beta = -.28$, 95% CI = [-.44, -.12]; motivational: $t(1160) = -4.48$, $SE = .08$, $p < .001$, $\beta = -.37$, 95% CI = [-.53, -.21]). Positive effects were found in the AI condition for the cognitive ($t(1160) = 5.09$, $SE = .10$, $p < .001$, $\beta = .49$, 95% CI = [.30, .68]) and motivational responses ($t(1160) = 3.05$, $SE = .09$, $p = .002$, $\beta = .29$, 95% CI = [.10, .47]), but they were only marginally significant for the affective responses ($t(1160) = 1.88$, $SE = .09$, $p = .06$, $\beta = .17$, 95% CI = [-.01, -. 34]).

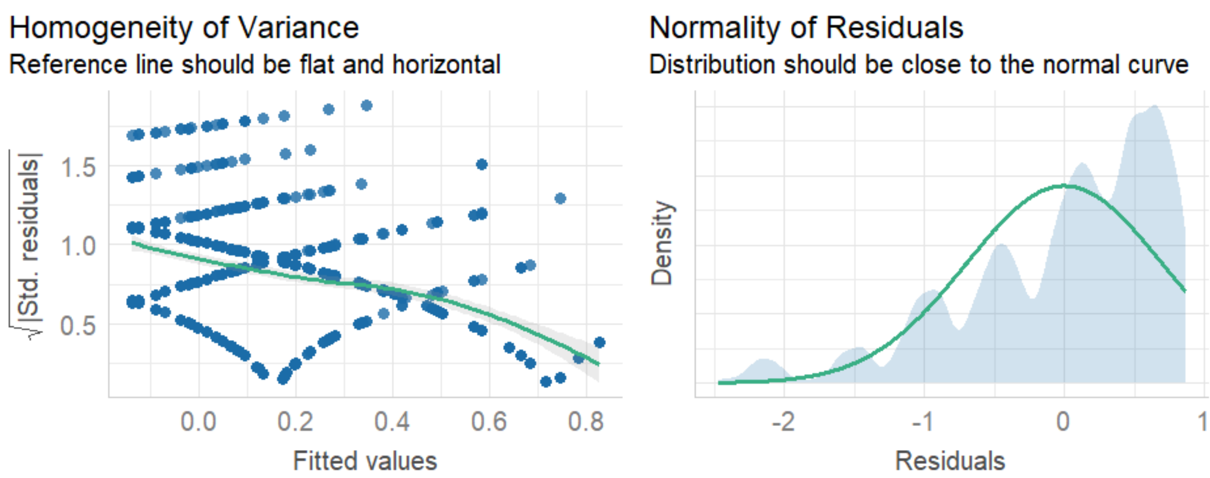
b. Permutation tests without assuming normality

The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: $W = .85$, $p < .001$).



We ran 10,000 permutations to test the model predicting the general empathy rating from the condition, response type and their interaction. The permutation replicated our main results, showing a significant main effect of condition (F(1, 1117) = 16.98, permutation *p* < .001), with contrasts showing it only in affective responses (t(1117) = -1.98, permutation *p* = .02) and motivational responses (t(1117) = -4.02, permutation *p* < .001), but not in cognitive responses (permutation *p* = .26).

We also examined these for the interaction with assumed aid by the other source, the figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .91, *p* < .001).



We ran a permutation test, using 10,000 permutations of a model predicting empathy from condition, response type, assumed aid by the other source and their interactions. The test showed that even using a permutation-based distribution, the effect of condition ($F(1, 1111) = 18.12$, permutation $p < .001$) and the interaction between condition and assumed aid ($F(1, 1111) = 72.98$, permutation $p < .001$) remained significant. Contrasts revealed that the positive effect of assumed aid in the AI condition ($t(1111) = 4.59$, permutation $p < .001$) and the negative effect in the human condition ($t(1111) = -7.71$, permutation $p < .001$) remained significant.

ii. <u>Mixed-effect models for aspects of empathy</u>

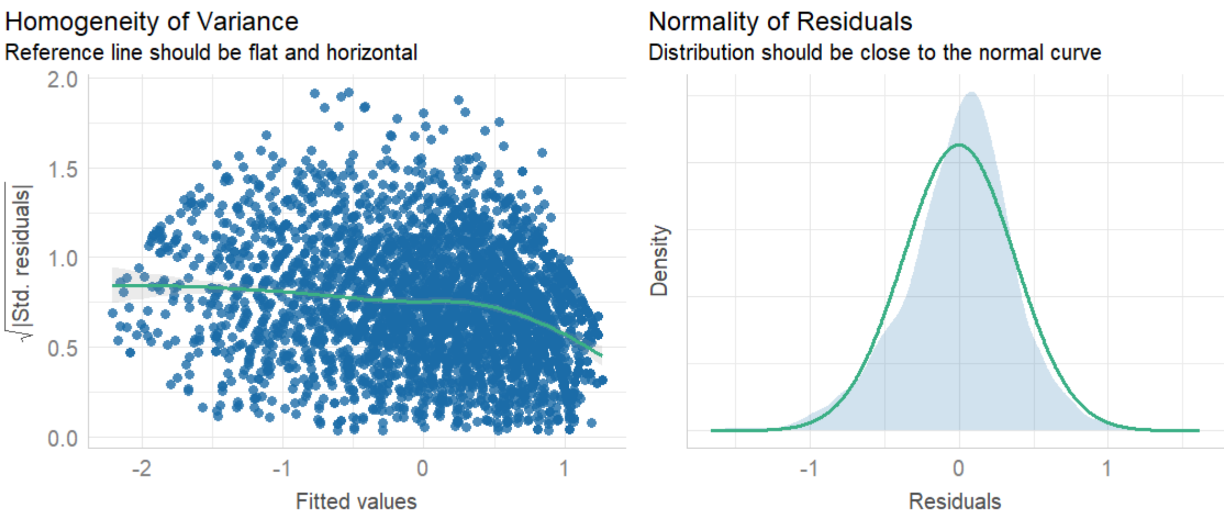
a. Analyses without filtering outliers

We fit a linear mixed-effect model including empathy type, condition, assumed aid and their interactions to the full dataset. The model found that the effect of condition remained significant ($F(1, 1166) = 36.14$, $p < .001$, partial $\eta^2 = .03$). The interaction of empathy type and response type also remained significant ($F(4, 2332) = 26.70$, $p < .001$, partial $\eta^2 = .04$). Post-hoc contrasts revealed significant differences between conditions for all response types (cognitive: $t(1166) = -3.27$, $SE = .09$, $p = .001$, $\beta = -.30$, 95% CI = [.-.47, -.12, 1.00]; affective: $t(1166) = -2.89$, $SE = .09$, $p = .004$ $\beta = -.26$, 95% CI = [-.44, -.08]; motivational: $t(1166) = -4.25$, $SE = .09$, $p < .001$, $\beta = -.39$, 95% CI = [-.57, -.21]).

To determine the interaction of condition and aspect of empathy with assumed aid by the other source, we added this variable to the previous model. The effect of condition ($F(1, 1160) = 30.32$, $p < .001$, partial $\eta^2 = .03$) remained significant. As in our main analysis, the interaction with assumed aid ($F(1, 1160) = 101.03$, $p < .001$, partial $\eta^2 = .08$) remained significant, as did the three-way interaction between condition, assumed aid and aspect of empathy ($F(2, 2320) = 3.53$, $p = .03$, partial $\eta^2 = .003$). Assumed aid had positive effects for all response types in the AI condition (cognitive: t(1634) = 4.54, SE = .05, $p < .001$, β = .24, 95% CI = [.14, .34]; affective: t(1634) = 6.95, SE = .05, $p < .001$ β = .37, 95% CI = [.26, .47]; motivational: t(1634) = 4.95, SE = .05, $p < .001$, β = .26, 95% CI = [.16, .37]), and it had negative effects in the human condition (cognitive: t(1634) = -7.65, SE = .05, $p < .001$, β = -.37, 95% CI = [-.46, -.27]; affective: t(1634) = -7.60, SE = .05, $p < .001$, β = -.37, 95% CI = [-.46, -.27]; motivational: t(1634) = -7.63, SE = .05, $p < .001$ β = -.37, 95% CI = [-.46, -27]). Further contrasts revealed that the positive effect in the AI condition was significantly larger for the affective responses than for the cognitive (t(2320) = 3.42, SE = .04, $p = .003$, β = .13, 95% CI = [.03, .23]) or motivational responses (t(2320) = 2.83, SE = .04, $p = .03$, β = .11, 95% CI = [.007, .12]).

b. Permutation tests without assuming normality

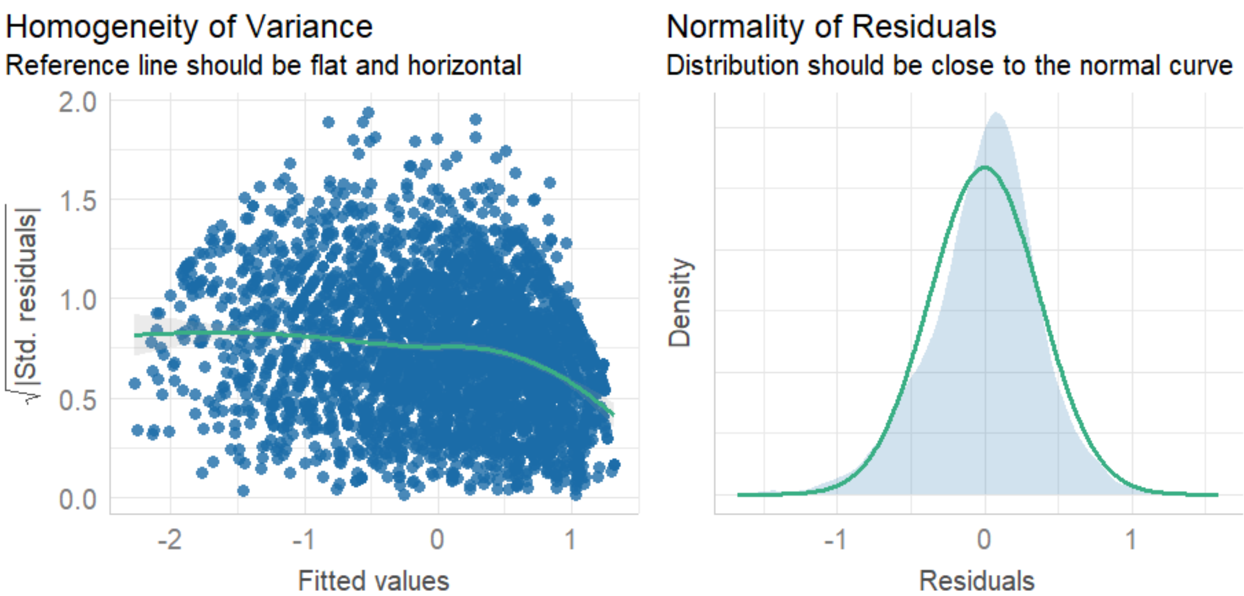
The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .99, $p < .001$) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations of a model predicting empathy from condition, response type, empathy type and their interactions. The test showed that even using a permutation-based distribution, both the effect of condition (F(1, 1134.10) =

36.13, permutation $p < .001$) and the interaction between response type and empathy type (F(1, 2244.60) = 23.95, permutation $p < .001$) remained significant. Contrasts showed that the difference in conditions remained significant across all types of empathy but was more extreme for affective (t(1153) = -3.54, permutation $p < .001$) and motivational empathy (t(1158) = -4.75, permutation $p < .001$) than for cognitive empathy (t(1159) = -2.12, permutation $p = .02$).

In terms of interaction with assumed aid by the other source, the figure below shows the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .99, $p <$ .001) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations of a model predicting empathy from condition, response type, empathy type, assumed aid by the other source and their interactions. The test showed that even using a permutation-based distribution, both the effect of condition (F(1, 1124) = 39.00, permutation $p < .001$) and the interaction between condition and assumed aid (F(1, 1127) = 101.94, permutation $p < .001$) remained significant.

The three-way interaction between condition, empathy type and assumed aid also remained significant (F(2, 2232.10) = 6.11, permutation p = .002). Contrasts showed that for all aspects of empathy, the assumed aid by the other source had a significant positive effect in the AI condition (cognitive: t(1703) = 3.48, permutation p < .001; affective: t(1701) = 6.52, permutation p < .001; motivational: t(1694) = 4.42, permutation p < .001), and it had a significant negative effect in the human condition (cognitive: t(1694) = -7.84, permutation p < .001; affective: t(1684) = -8.40, permutation p < .001; motivational: t(1680) = -8.51, permutation p < .001). Further contrasts showed that the effect was again significantly larger for affective responses than for cognitive (t(2259) = 4.03, permutation p < .001) or motivational responses (t(2257) = 2.81, permutation p = .002).

iii. Positivity resonance

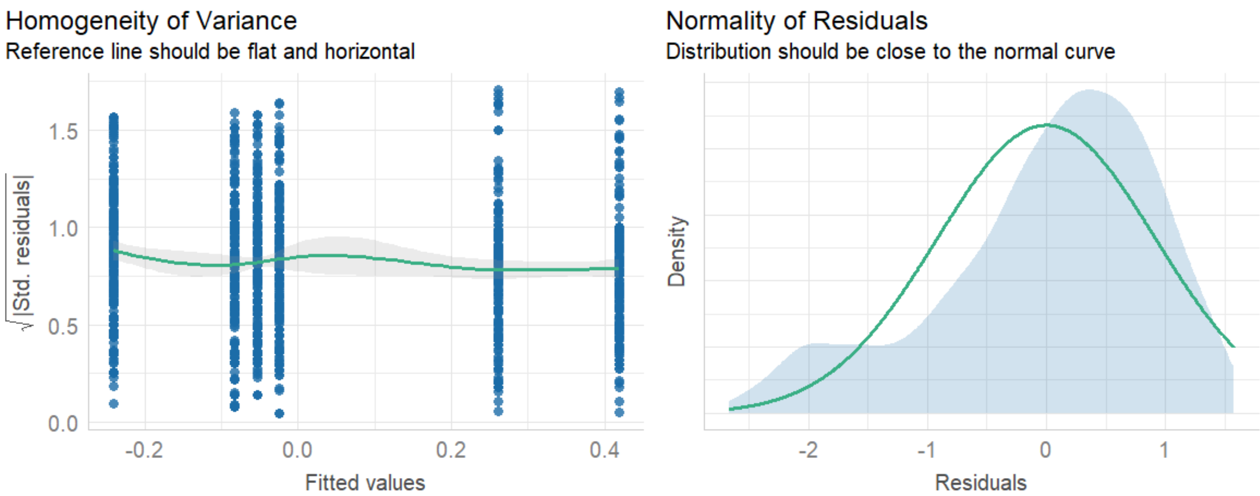
a. Analyses without filtering outliers

We fitted a linear model predicting positivity resonance from condition, response type and their interaction. Both the effect of condition (F(1, 1166) = 36.86, p < .001, partial η^2 = .03) and the effect of response type remained significant (F(2, 1166) = 18.82, p < .001, partial η^2 = .001). Contrasts showed that the difference between conditions was significant for all response types (cognitive: t(1166) = -2.67, SE = .10, p = .008, β = -.26, 95% CI = [-.46, -.07]; affective: t(1166) = -3.33, SE = .10, p < .001, β = -.33, 95% CI = [-.52, -.13]; motivational: t(1166) = -4.46, SE = .10, p < .001, β = -.44, 95% CI = [-.63, -.25]). Additionally, as in our initial analysis, cognitive responses were rated significantly lower for positivity resonance compared to affective (t(1166) = -4.48, SE = .07, p < .001, β = -.31, 95% CI = [-.48, -.14]) and motivational responses (t(1166) = -5.86, SE = .07, p < .001, β = -.41, 95% CI = [-.58, -.24]).

Next, we added assumed aid by the other source to the previous model, to replicate the interaction it had with condition. The effect of condition (F(1, 1160) = 40.64, p < .001, partial η^2 = .03) and the effect of response type (F(2, 1160) = 20.35, p < .001, partial η^2 = .03) remained significant. Supporting our main analysis, the interaction with assumed aid (F(1, 1160) = 106.94, p < .001, partial η^2 = .08) remained significant as well. The positive effect that assumed aid had in the AI condition on positivity resonance (t(1160) = 5.08, SE = .05, p < .001, β = .27, 95% CI = [.16, .37]) and its negative effect in the human condition (t(1160) = -9.87, SE = .05, p < .001, β = -.47, 95% CI = [-.56, -.38]) both remained significant.

b. Permutation tests without assuming normality

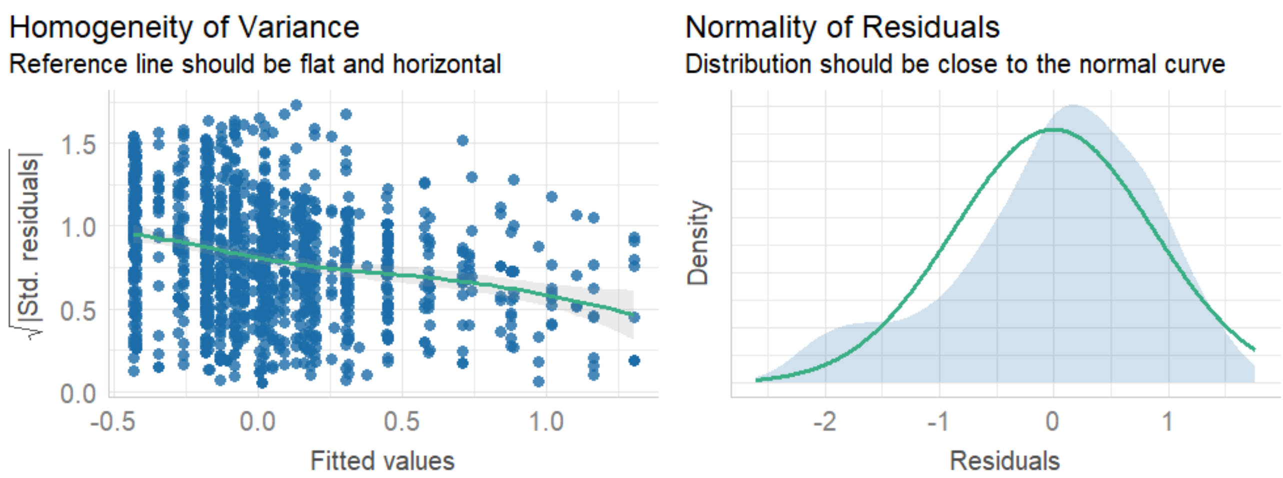
The figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .95, p < .001) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations of a model predicting positivity resonance from condition, response type and their interaction. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 1144) = 31.68, permutation p < .001). The difference between response types also remained significant (F(2, 1144) = 13.35, permutation p < .001). Contrasts showed that the difference between conditions was significant for all response types (cognitive: t(1144) = -

1.66, permutation $p = .05$; affective: t(1166) = -3.39, permutation $p < .001$; motivational: t(1144) = -4.73, $p < .001$), and that cognitive responses were still rated significantly lower than affective (t(1144) = -4.01, permutation $p < .001$) or motivational responses (t(1144) = -5.39, permutation $p < .001$).

To examine the interaction with assumed aid by the other source, the figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .97, $p < .001$) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations of a model predicting positivity resonance from condition, response type, assumed aid by the other source and their interaction. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 1138) = 35.00, permutation $p < .001$). The difference between response types also remained significant (F(2, 1138) = 16.94, permutation $p < .001$). The interaction between condition and assumed aid remained significant (F(2, 1138) = 103.63, permutation $p < .001$). Contrasts showed that both the positive effect of assumed aid on positivity resonance in the AI condition (t(1138) = 4.47, permutation $p < .001$)

and its negative effect in the human condition (t(1138) = -10.25, permutation $p < .001$) remained significant.

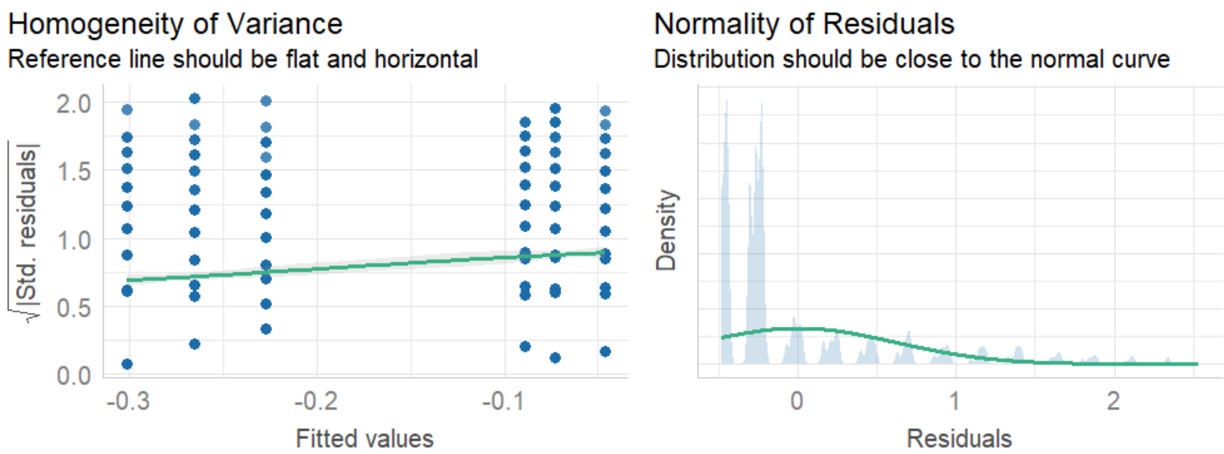
iv. Negative emotions

a. Analyses without filtering outliers

We fit a model predicting negative emotions to the full dataset. As in our main analysis, we found a significant effect of condition on negative emotions (F(1, 1166) = 19.53, $p < .001$, partial $\eta^2 = .001$), showing that responses perceived as human raise less negative emotions than responses perceived as AI-generated.

b. Permutation tests without assuming normality

The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: W = .74, $p < .001$).



We ran a permutation test, using 10,000 permutations of a model predicting negative emotions from condition, response type and their interaction. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 1112) = 27.94, permutation $p < .001$).

v. Positive emotions

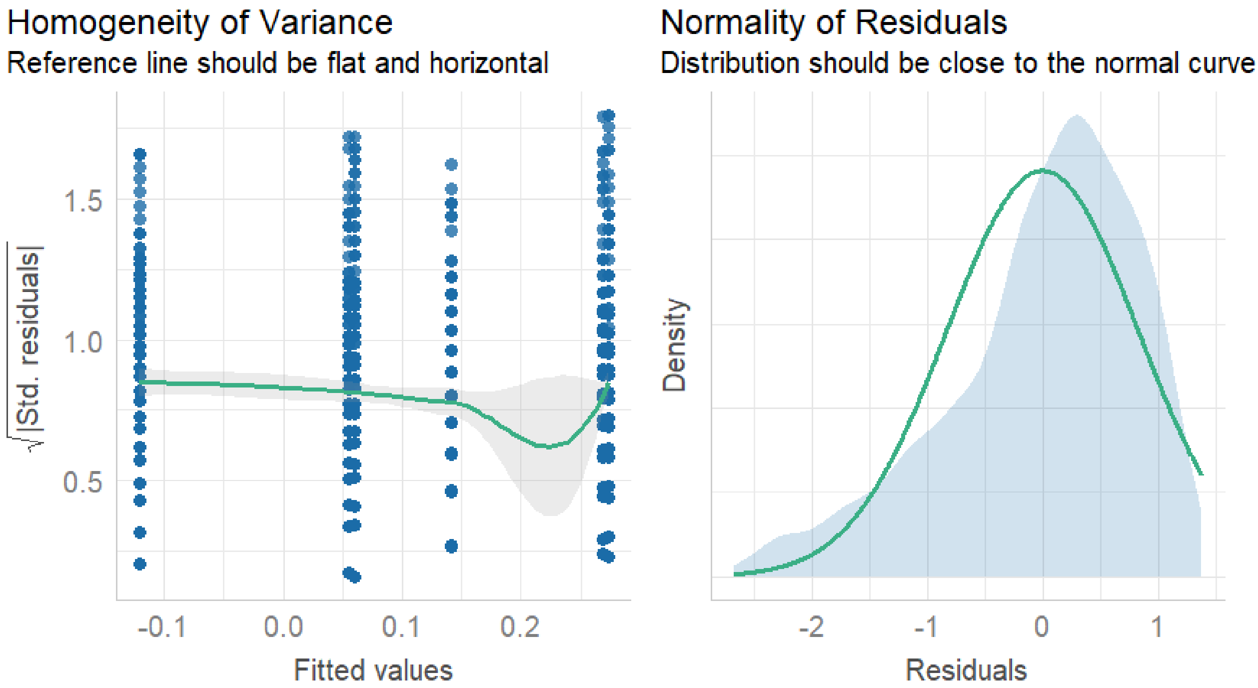
a. Analyses without filtering outliers

We fit a model predicting positive emotions to the full dataset. As in our main analysis, we found a significant effect of condition on positive emotions ($F(1, 1166) = 20.58$, $p < .001$, partial $\eta^2 = .02$), showing that responses perceived as human raise more positive emotions than those perceived as AI-generated. We again found a significant effect of response type as well ($F(2, 1166) = 5.32$, $p = .005$, partial $\eta^2 = .009$), with motivational responses being rated significantly higher for positive emotions than cognitive responses ($t(1166) = -3.18$, $SE = .07$, $p = .005$, $\beta = -.23$, 95% CI = [-.40, -.06]).

We then added assumed aid by the other source to the previous model. The effect of condition ($F(1, 1160) = 22.54$, $p < .001$, partial $\eta^2 = .02$) remained significant, as did the effect of response type ($F(2, 1160) = 15.29$, $p < .001$, partial $\eta^2 = .01$). As in our main analysis, the interaction with assumed aid ($F(1, 1160) = 99.15$, $p < .001$, partial $\eta^2 = .08$) remained significant. Assumed aid had positive effects on positive emotions for all response types in the AI condition ($t(1160) = 4.78$, $SE = .05$, $p < .001$, $\beta = .26$, 95% CI = [.15, .36]), and this influence was negative in the human condition ($t(1160) = -9.57$, $SE = .05$, $p < .001$, $\beta = -.46$, 95% CI = [-.56, -37]).

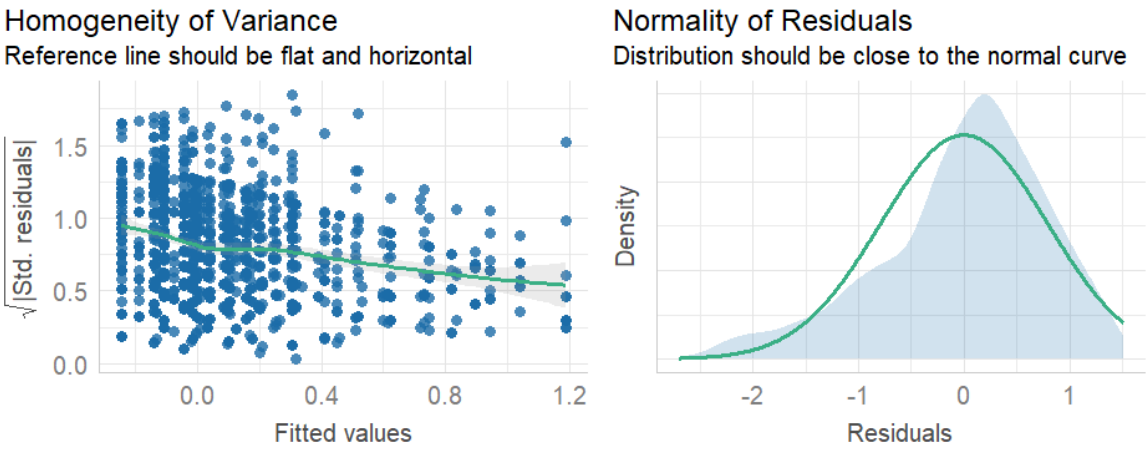
b. Permutation tests without assuming normality

The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: $W = .95$, $p < .001$).



We ran a permutation test, using 10,000 permutations of a model predicting positive emotions from condition, response type and their interaction. The test showed that even using a permutation-based distribution, the effect of condition ($F(1, 1120) = 21.58$, permutation $p < .001$) and the effect of response type ($F(1, 1120) = 4.29$, permutation $p = .01$) both remained significant. Contrasts showed that even using permutations, cognitive responses were rated significantly lower than affective ($t(1120) = -2.51$, permutation $p = .005$) and motivational responses ($t(1120) = -2.58$, permutation $p = .005$).

In terms of the interaction with assumed aid by the other source, the figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: $W = .96$, $p < .001$) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations of a model predicting positive emotions from condition, response type, assumed aid by the other source and their interactions. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 1116) = 18.15, permutation $p < .001$). Contrasts revealed that the positive effect of assumed aid on positive emotions in the AI condition (t(1114) = 4.11, permutation $p < .001$) and the negative effect in the human condition (t(1114) = -9.52, permutation $p < .001$) remained significant.

vi. Authenticity

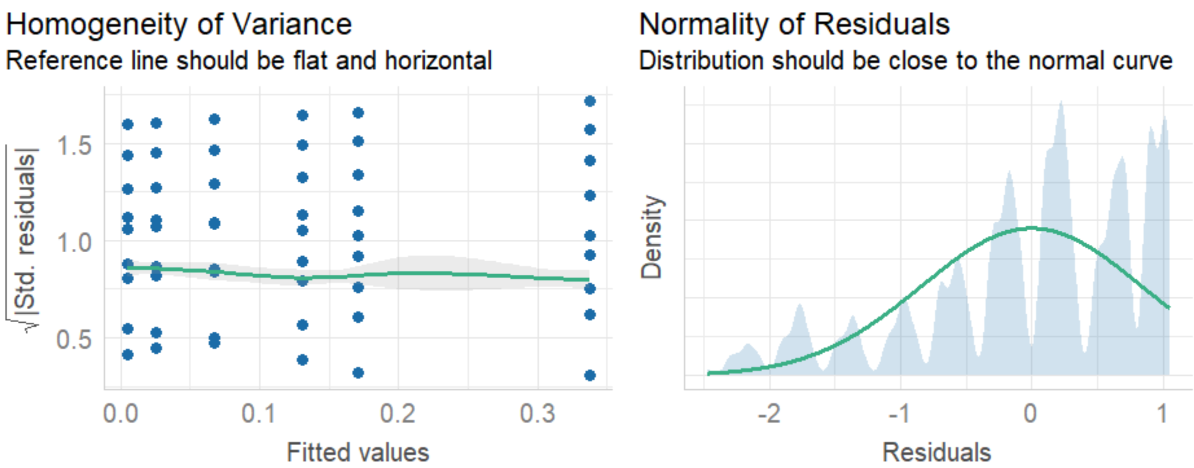
a. Analyses without filtering outliers

We fit a model predicting authenticity to the full dataset. As in our main analysis, we found a significant effect of condition on authenticity (F(1, 1166) = 10.16, $p < .001$, partial η^2 = .008), showing that responses perceived as human were rated higher for authenticity than those perceived as AI-generated. The effect of response type also remained significant (F(2, 1166) = 8.01, $p < .001$, partial η^2 = .01), with motivational responses being rated significantly higher for authenticity than cognitive responses (t(1166) = -3.98, SE = .07, $p < .001$, β = -.28, 95% CI = [-.45, -.11]), like in our main analysis.

Next, we added assumed aid by the other source to the previous model. The effects of condition ($F(1, 1160) = 11.63$, $p < .001$, partial $\eta^2 = .01$) and response type ($F(2, 1160) = 9.16$, $p < .001$, partial $\eta^2 = .02$) remained significant. As in our main analysis, the model indicated a significant interaction between condition and assumed aid ($F(1, 1160) = 147.96$, $p < .001$, partial $\eta^2 = .11$), showing a positive effect on authenticity in the AI condition ($t(1160) = 5.62$, $SE = .05$, $p < .001$, $\beta = .29$, 95% CI = [.19, .40]) and a negative effect in the human condition ($t(1160) = -11.94$, $SE = .05$, $p < .001$, $\beta = -.57$, 95% CI = [-.66, -.47]).

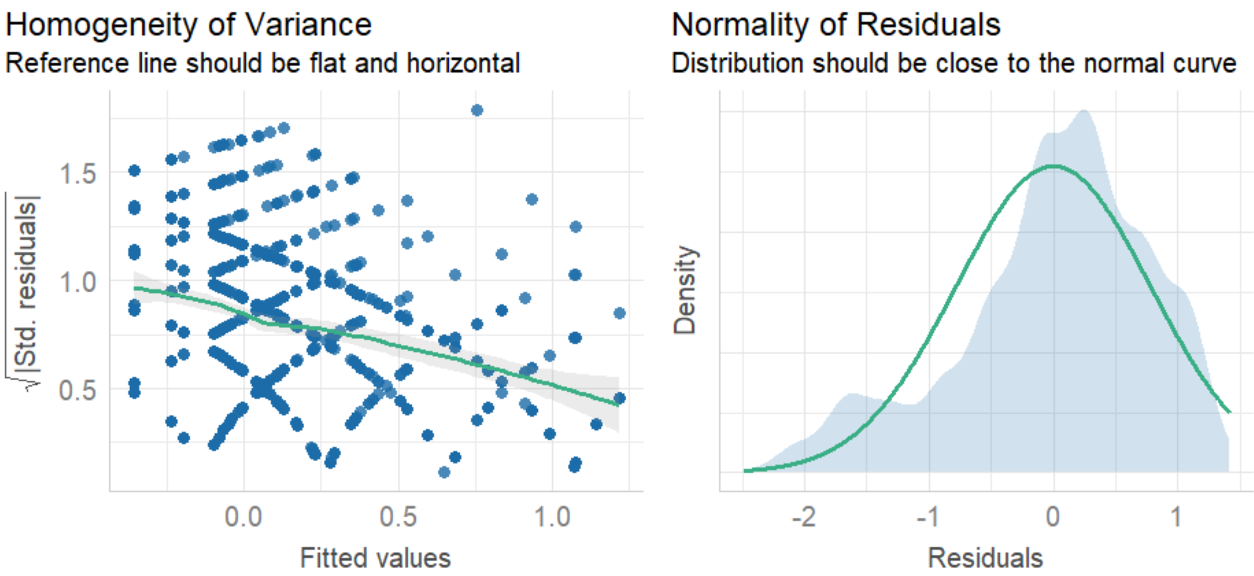
b. Permutation tests without assuming normality

The figure below shows the level of deviation our model has from the assumption of normality of residuals (Shapiro-Wilk test: $W = .92$, $p < .001$).



We ran a permutation test, using 10,000 permutations of a model predicting authenticity from condition, response type and their interaction. The test showed that even using a permutation-based distribution, the effect of condition ($F(1, 1111) = 4.89$, permutation $p = .03$) and the effect of response type ($F(1, 1111) = 6.38$, permutation $p = .002$) remained significant. Contrasts showed that even using permutations, cognitive responses were rated significantly lower than motivational responses ($t(1111) = -3.54$, permutation $p < .001$).

In terms of interaction with assumed aid by the other source, the figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .96, $p < .001$) but are not extremely skewed.

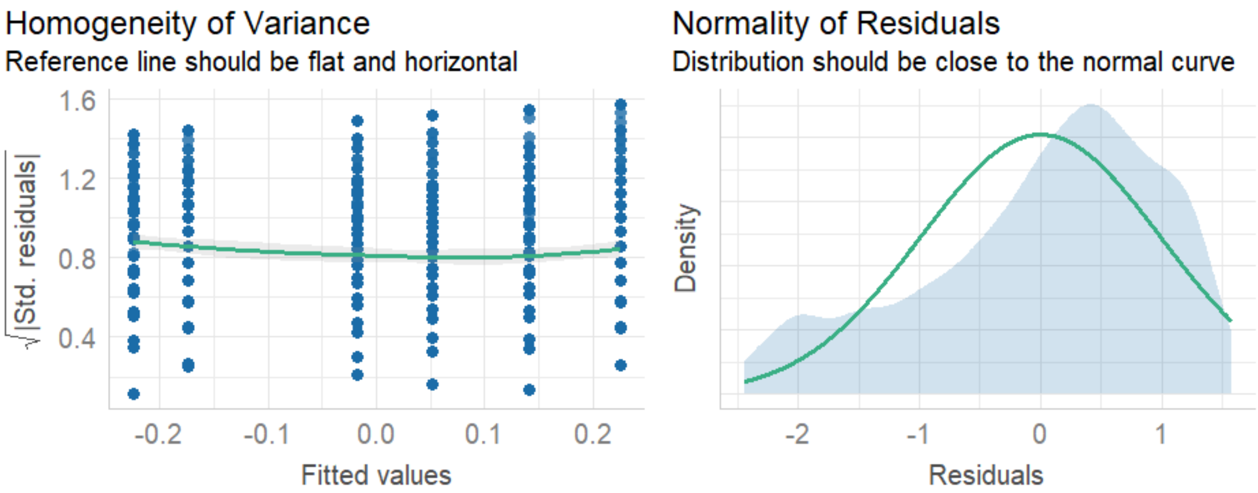


We ran a permutation test, using 10,000 permutations of a model predicting authenticity from condition, response type, assumed aid by the other source and their interactions. The test showed that even using a permutation-based distribution, the effect of condition remained significant (F(1, 1105) = 5.59, permutation $p = .02$). The interaction between condition and assumed aid (F(1, 1105) = 125.00, permutation $p < .001$) also remained significant. Contrasts revealed that the positive effect on authenticity of assumed aid in the AI condition (t(1105) = 4.02, permutation $p < .001$) and its negative effect in the human condition (t(1105) = -12.17, permutation $p < .001$) remained significant.

vii. <u>Support</u>

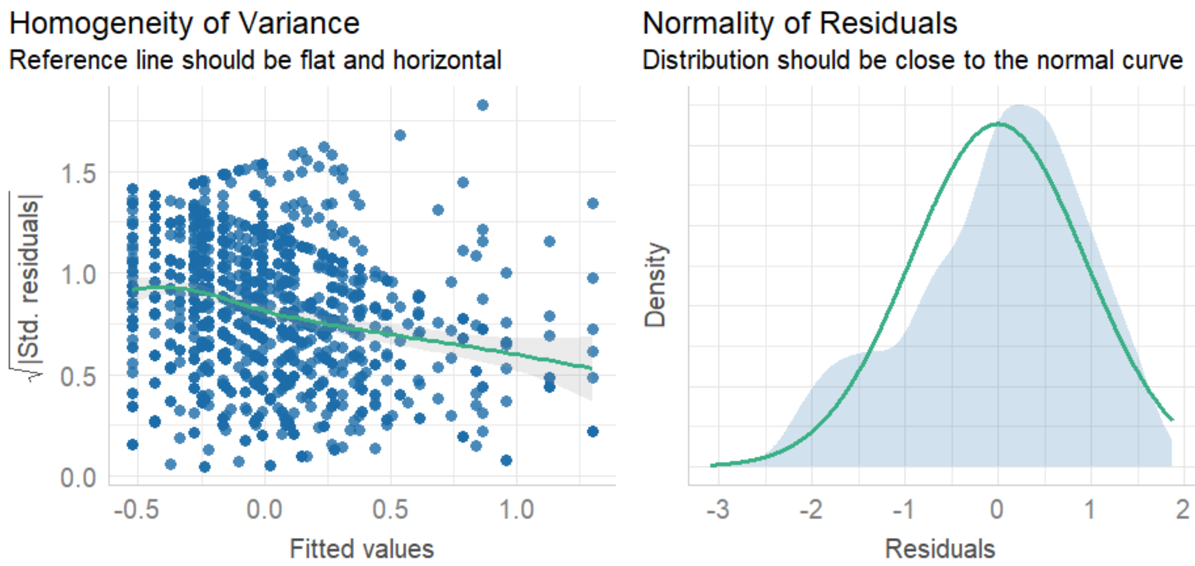
a. Permutation tests without assuming normality

The figure below shows the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .95, $p < .001$) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations of a model predicting perceived support from condition, response type and their interaction. The test showed that even using a permutation-based distribution, the effect of response type ($F(1, 1166) = 14.62$, permutation $p < .001$) remained significant. Contrasts showed that as in our main analyses, even using permutations, cognitive responses were rated significantly lower in perceived support than affective ($t(1166) = -3.06$, permutation $p < .001$) and motivational responses ($t(1166) = -5.39$, permutation $p < .001$), and affective responses were rated lower than motivational ones ($t(1166) = -2.35$, permutation $p = .01$).

In terms of interaction with assumed aid by the other source, the figure below illustrates the level of deviation our model has from the assumption of normality of residuals, showing that our residuals do not perfectly fit a normal distribution (Shapiro-Wilk test: W = .97, $p < .001$) but are not extremely skewed.



We ran a permutation test, using 10,000 permutations of a model predicting perceived support from condition, response type, assumed aid by the other source and their interactions. The test showed that even using a permutation-based distribution, the interaction effect of condition and assumed aid remained significant ($F(1, 1160) = 108.40$, permutation $p < .001$). Contrasts revealed that the positive effect of assumed aid on perceived support in the AI condition ($t(1160) = 4.72$, permutation $p < .001$) and the negative effect in the human condition ($t(1160) = -10.31$, permutation $p < .001$) remained significant.

10. <u>Detailed analyses for studies 4–5</u>

a. <u>Detailed analyses for the effects of timeframe</u>

We used logistic regression to test whether and how much the timeframe offered would increase the probability of participants choosing AI, and no effect of timeframe was found ($p = .46$ for study 4, $p = .70$ for study 5; see supplementary). We used the logarithmic transformation of the number of minutes suggested. A remarkably similar percentage of people (~40–50% in study 4 across the various time conditions; ~30–40% in study 5) chose a human response regardless of the given timeframe. Contrary to our expectations, our results do not show a significant effect for the timeframe participants were offered in either study (p = .46 for study 4, p = .70 for study 5; see figures 10a-A, 10a-B). That is, people's preferences for human or AI responses did not seem to vary much, even if they had to wait up to 2 years to receive a human response. These fixed preferences suggest that there are probably strong individual differences that might be influencing people's choices. Indeed, we found a significant effect of feeling about AI ($p < .001$, Z = -5.089, SE = .10, DF = 488, OR = .61; see figure 10a-C) when this variable was included in the model, showing that those who feel more positively about AI were more likely to choose an AI response. This was replicated in study 4 ($p < .001$, Z = -5.783, SE = .11, DF = 485, OR = .54; see figure 10a-D).

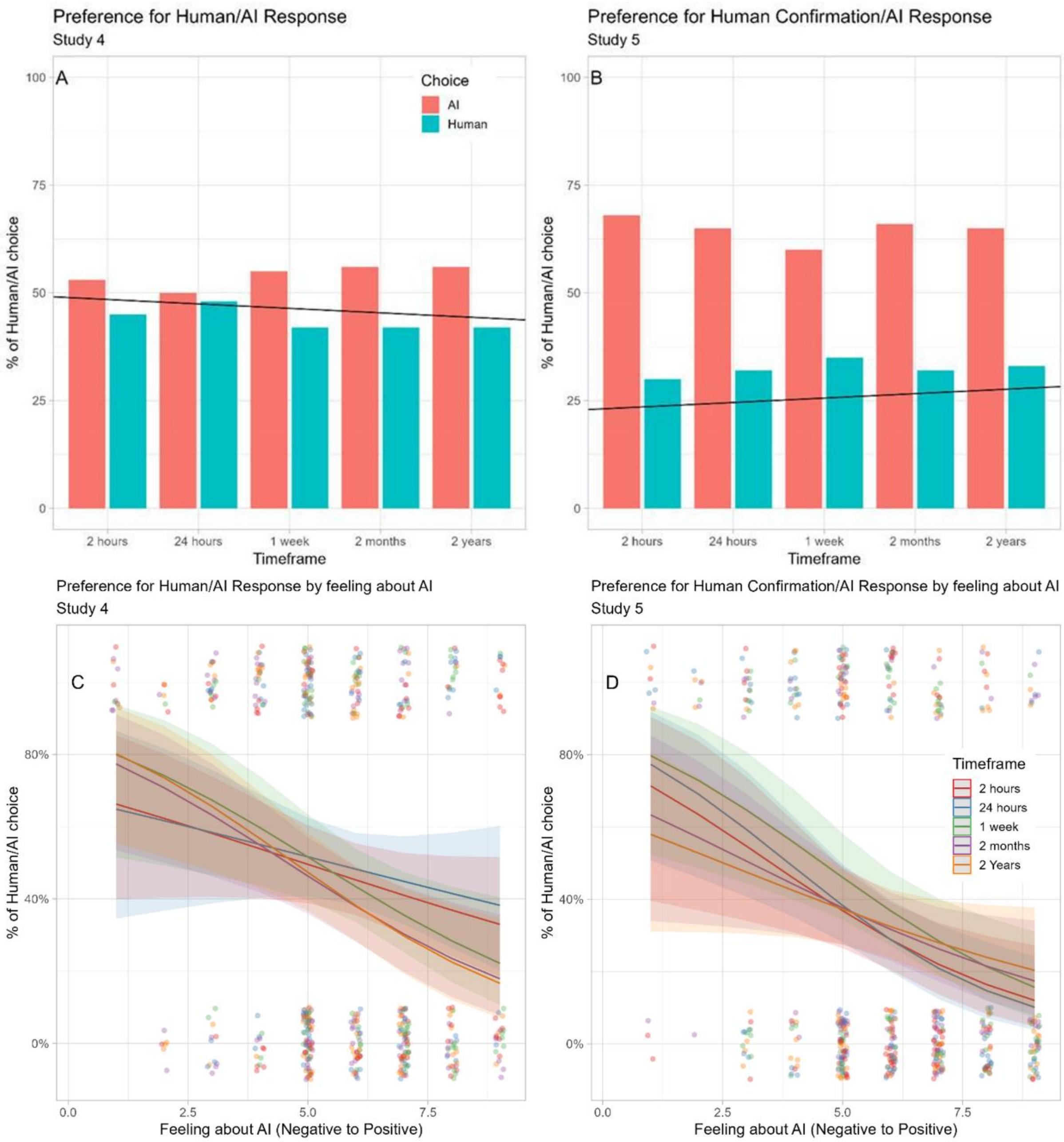


Figure 10a: Percentage of participants' preference for an AI or human response by the timeframe they were given in study 4 (A) and in study 5 (B), with a regression line marked in black showing the miniscule difference between the timeframes. The participants' choice was influenced by their positive or negative feelings about AI (plot C for study 4 and plot D for study 5). Due to these plots showing logistic regression models, data points are along 100% or 0% and not along the regression line itself.

b. **<u>Questions on participants' reasoning for their choice</u>**

Following their choice between an AI response and a human response (in study 4), or between an AI response and human confirmation that their story was read (study 5), participants were asked to rate the following statements, from 1 (not true at all) to 9 (very true):

Choice	Statement
AI	I am hesitant about another person's response
	I am curious about an AI's view of my story
	I believe AI will understand my experience better than a human
	I prefer receiving a response now rather than later*
	I believe an AI can share my emotions better than another person
	I believe an AI will provide more care and support than another person
	I believe an AI will alleviate my sense of loneliness more than another person
Human	I am hesitant about AI's response
	I am curious about another person's view of my story
	I believe another person will understand my experience better than an AI
	I prefer receiving a response later rather than now*
	I believe another person can share my emotions better than an AI
	I believe another person will provide more care and support than an AI
	I believe another person will alleviate my sense of loneliness more than an AI

* These statements were changed in study 5, due to not offering participants a human response. Instead, the question following the choice of an AI was changed to: "I prefer receiving a response now rather than another person confirm later". The question following

choice of human confirmation was: “I am willing to wait, but want someone human to know what I'm going through”.

11. **<u>Consent form for use of the participants’ stories in the future</u>**

Participants were given the following choice concerning the use of their stories:

Please tell us to what degree you agree that we use the story you shared for future lab needs and studies:

- I consent to any use of my story. This includes use in future studies and I understand it may be shown to future participants, as well as public speaking affairs like conferences and lectures. **Note that this will always be without identifying details**.
- I consent to the use of my story in this study, and in future studies and understand it may be shown to future participants. **Note that this will always be without identifying details**.
- I consent to the use of the story in this study only.
- I do not consent to any future use of the story.

12. **<u>Questionnaire on perceived empathy in the response</u>**

Please rate how true the following statements are:

I found the AI/other person very empathetic to my story.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person correctly understood my emotions.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed an understanding of what being in my shoes is like.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9

I feel that the AI/other person did not show an understanding of my perspective or emotions.*	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed an ability to accurately perceive and understand my perspective.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed a deep interest in my emotions. Regardless of the true answer, please answer nine (the rightmost) for this question.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed a different interpretation of my emotions, inconsistent with my own.*d	Not true at all – 0	1	2	3	4	5	6	7	8	Completely True – 9
I feel the AI/other person showed a genuine sharing and experiencing of the emotions that I felt.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel that the AI/other person reflected an emotional experience completely different from my own.*d	Not true at all – 0	1	2	3	4	5	6	7	8	Completely True – 9
I feel the AI/other person showed no emotional reaction to my story.*	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed a deep emotional synchrony with my emotions.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed my story was emotionally moving.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed a meaningful effort to relate to my emotions.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed little care or thought about my experiences.*	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed a desire to actively help me.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9

I feel the AI/other person showed a genuine concern for my well-being.	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9
I feel the AI/other person showed no motivation to help in making a positive difference in my life.*d	Not true at all – 0	1	2	3	4	5	6	7	8	Completely true – 9

* Reverse-scored items

d – Items dropped after study 1

13. Prompts for AI-generated responses:

Studies 1a, 1b, 4 and 5 used the following GPT prompts:

Prompt 1:

"In this exercise, you will be shown a description of a short personal story and write an empathic response. An empathic response needs to relate somehow to the story, and must include three elements. These are cognitive empathy, which involves understanding how the other person is feeling, affective empathy, which involves sharing and experiencing the same emotions, and motivational empathy, which involves showing you care and want to help them. If you understand these instructions, repeat them."

Prompt 2:

"The story is: [participant's story was entered here using Qualtrics piped text]. Please write an empathic response to the story as instructed above, which will reflect an understanding of the emotions the person expressed in the story, sharing these emotions and caring about them. Remember to direct the response at the person in the story and use second person in your response. Limit your response to 2–4 sentences."

Study 2a used the following prompt:

"In this exercise, you will be shown a description of a short personal story and write an empathic response. An empathic response needs to relate somehow to the story, and must include three elements. These are Cognitive empathy, which involves understanding how the other person is feeling, affective empathy, which involves sharing and experiencing the same emotions, and motivational empathy, which involves showing you care and want to help them. The story is: [participant's story was entered here using Qualtrics piped text]. Please write an empathic response to the story as instructed above, which will reflect an understanding of the emotions the person expressed in the story, sharing these emotions and caring about them. Remember to direct the response at the person in the story and use second person in your response. Use short sentences, an informal tone, and day-to-day language. answer as an average person would given the instructions on empathy, if they only had 60 seconds to respond. Do not use fancy words, or uncommon punctuation such as a semicolon. limit your response to 2 sentences, each one comprised of no more than 13 words."

Studies 1c and 2b used the following prompt:

"In this exercise, you will be shown a description of a short personal story and write an empathic response. An empathic response needs to relate somehow to the story, and must include three elements. These are Cognitive empathy, which involves understanding how the other person is feeling, affective empathy, which involves sharing and experiencing the same emotions, and motivational empathy, which involves showing you care and want to help them. The story is: [participant's story was entered here using Qualtrics piped text]. Please write an empathic response to the story as instructed above, which will reflect an understanding of the emotions the person expressed in the story, sharing these emotions and caring about them. Remember to direct the response at the person in the story and use second person in your response. Limit your response to 2-4 sentences."

Study 3 used the following GPT prompts:

Prompt for cognitive responses:

"In this exercise, you will be shown a description of a short personal story and write a cognitively empathic response. A cognitively empathic response needs to convey cognitive empathy, which involves understanding how the other person is feeling as accurately as possible. The point of the response should be to show you accurately perceive and identify the storyteller's emotions in the story. Do not to mention your own emotions. The story is: [participant's story was entered here using Qualtrics piped text]. Please write an empathic response to the story as instructed above, which will reflect perceiving the emotions the person expressed in the story as accurately as possible. Remember to direct the response at the person in the story and use second person in your response. Limit your response to 2–4 sentences."

Prompt for affective responses:

"In this exercise, you will be shown a description of a short personal story and write an affectively empathic response. An affectively empathic response needs convey affective empathy, which involves sharing and experiencing the same emotions the other person is feeling as much as possible. The point of the response should be to show how you share and experience the same emotions as they do in the story. Focus on your experience of them. The story is: [participant's story was entered here using Qualtrics piped text]. Please write an empathic response to the story as instructed above, which will reflect sharing the emotions the person expressed in the story as much as possible. Remember to direct the response at the person in the story and use second person in your response. Limit your response to 2–4 sentences."

Prompt for motivational responses:

"In this exercise, you will be shown a description of a short personal story and write a motivationally empathic response. A motivationally empathic response needs to convey motivational empathy, which involves showing you support and care for the other person as much as possible. The point of the response should be to show how you care and wish to support them in response to their story. Focus on your feeling toward them. The story is: [participant's story was entered here using Qualtrics piped text]. Please write an empathic response to the story as instructed above, which will reflect support and care for the person in the story as much as possible. Remember to direct the response at the person in the story and use second person in your response. Limit your response to 2–4 sentences."

Study 1d used the following GPT system prompt:

"In this exercise, you will be shown a description of a short personal story and write an empathic response. An empathic response needs to relate somehow to the story, and must include three elements. These are Cognitive empathy, which involves understanding how the other person is feeling, affective empathy, which involves sharing and experiencing the same emotions, and motivational empathy, which involves showing you care and want to help them. This will be a back-and-forth interaction, so ask questions or comment on their experience. Do so with empathy for their experience as instructed above, showing an understanding of the emotions the person expressed in the story, sharing these emotions and caring about them. Don't offer practical solutions. Focus on the person's emotions, matching the valence of their emotions and making sure to accept them and their emotions, providing care. Respond in a way that will prompt to keep the conversation going until there is only 1 message left. Remember to direct the response at the person in the story and use second person in your responses. Use short sentences, an informal tone, and day-to-day language. Do not say you are an AI. Do not use fancy words, or uncommon punctuation such as a semicolon. Pay attention to the chat history and do not ask the same question again. Do not repeat phrases you've used. portray your empathy in a different way between messages. Limit your response to [sentence limit was given as a variable here] sentences, each one comprised of no more than [word limit was given as a variable here] words. Know that there are only [remaining messages were calculated here] messages until the end of the conversation."

The last sentence in the prompt changed for the last message, being replaced by "this is the last message, so close off the conversation and don't ask follow up questions."