Supplementary material

for

Domain-specific and domain-general neural network engagement during human-robot interactions

by

Ann Hogenhuis and Ruud Hortensius

published in

European Journal of Neuroscience

Table S1. Coordinates for the Theory-of-Mind network

Region	*MNI Coordinates*		
	x	*y*	*z*
Dorsomedial prefrontal cortex	-6	54	36
Middle medial prefrontal cortex	-4	58	16
Ventromedial prefrontal cortex	-4	56	-16
Precuneus	0	-54	34
Left temporoparietal junction	-48	-62	30
Right temporoparietal junction	48	-60	30

Coordinates are based on Richardson et al. (2018).

Table S2. Coordinates for the person perception network

Region	*MNI Coordinates*		
	x	*y*	*z*
Left occipital face area	-40	-76	-18
Right occipital face area	44	-76	-12
Left fusiform face area	-40	-52	-18
Right fusiform face area	38	-42	-22
Left posterior superior temporal sulcus	-54	-36	6
Right posterior superior temporal sulcus	48	-38	4
Left extrastratiate body area	-48	-74	10
Right extrastratiate body area	50	-70	10

Coordinates are based on Julian et al. (2012).

Table S3. Coordinates for the object-selective regions

Region	*MNI Coordinates*		
	x	*y*	*z*
Left fusiform gyrus	-26	-60	-10
Right fusiform gyrus	30	-50	-8
Left lateral occipital complex	-46	-72	-4
Right lateral occipital complex	46	-70	-4
Left superior parietal lobule	-24	-56	60
Right superior parietal lobule	24	-52	64
Left middle occipital gyrus	-30	-84	12
Right middle occipital gyrus	36	-84	12

Bilateral lateral occipital complex and superior parietal lobule coordinates are based on Julian et al. (2012), while the bilateral fusiform gyrus, superior parietal lobule, and middle occipital gyrus coordinates are based on Henschel, Hortensius and Cross (2020) using independent coordinates from Dubey et al. (2020).

Table S4 Coordinates for the language network

Region	*MNI Coordinates*		
	x	*y*	*z*
Left orbital portion of the inferior frontal gyrus	-42	28	-4
Left inferior frontal gyrus	-48	18	20
Left middle frontal gyrus	-44	-6	58
Left anterior temporal lobe	-52	-8	-14
Left posterior temporal lobe	-48	-38	2
Left angular gyrus	-48	-62	20
Right orbital portion of the inferior frontal gyrus	46	30	4
Right inferior frontal gyrus	42	18	22
Right middle frontal gyrus	50	-4	52
Right anterior temporal lobe	48	-4	-18
Right posterior temporal lobe	44	-32	0
Right angular gyrus	46	-62	20

Coordinates of the regions are based on the peak voxels of the group maps masked with the network parcels using data from Diachek and colleagues (2020; Experiment 1, n = 383, https://osf.io/pdtk9/).

Table S5 Coordinates for the multiple-demand network

Region	*MNI Coordinates*		
	x	*y*	*z*
Left posterior parietal lobe	-16	-66	32
Left middle parietal lobe	-44	-54	42
Left anterior parietal lobe	-52	-44	48
Left superior frontal gryus	-18	16	62
Left precentral gyrus	-40	18	42
Left opercular portion of the inferior frontal gyrus	-34	18	34
Left middle frontal gyrus	-40	42	14
Left orbital middle frontal gyrus	-40	46	10
Left insula	-36	14	-6
Left medial frontal gyrus	-2	26	40
Right posterior parietal lobe	18	-62	32
Right middle parietal lobe	54	-52	40
Right anterior parietal lobe	54	-46	40
Right superior frontal gryus	22	18	60
Right precentral gyrus	40	18	50
Right opercular portion of the inferior frontal gyrus	38	18	40
Right middle frontal gyrus	38	22	48
Right orbital middle frontal gyrus	42	48	-2
Right insula	30	18	-12
Right medial frontal gyrus	2	32	34

Coordinates of the regions are based on the peak voxels of the group maps masked with the network parcels using data from Diachek and colleagues (2020; Experiment 1, n = 383, https://osf.io/pdtk9/).

Table S6. Comparison of activity within each network when listening and speaking during interactions with a human and robotic agent

Network	*Main effect of dynamics*			*Main effect of type of interaction*			*Interaction between dynamics and type of interaction*		
	F	*p*	η^2_G	*F*	*p*	η^2_G	*F*	*p*	η^2_G
Theory-of-Mind	**101.48**	**1.703e-09**	**0.37**	3.25	.086	0.03	2.28	.146	0.01
Person perception	**96.60**	**2.621e-09**	**0.55**	**7.47**	**.013**	**0.04**	**24.75**	**6.363e-05**	**0.09**
Object-specific	1.45	.242	0.01	6.69	.017*	0.07	1.60	.219	5.07e-03
Language	**29.18**	**2.338e-05**	**0.23**	2.95	.1	0.01	**51.50**	**4.487e-07**	**0.10**
Multiple-demand	**18.47**	**.0003**	**0.09**	0.38	.545	4.05e-03	0.33	.572	1.17e-03

Beta estimates were extracted for each event for each ROI and averaged across network and entered in a 2 (interaction: HHI, HRI) by 2 (dynamics: listening, speaking) repeated-measures ANOVA for each network separately. Significant effects are in bold, * Did not survive Bonferroni correction ($p < .05 / 5$); η^2_G: generalised eta-squared.

Table S7. Estimated posterior regression coefficient for the region-of-interest Bayesian linear model

Network	*Main effect of dynamics*		*Main effect of type of interaction*		*Interaction between dynamics and type of interaction*	
	Estimates	*CI (95%)*	*Estimates*	*CI (95%)*	*Estimates*	*CI (95%)*
Theory-of-Mind	1.99	1.34 – 2.65	0.14	-0.51 – 0.80	0.78	-0.08 – 1.67
Person perception	2.60	2.13 – 3.05	0.08	-0.38 – 0.54	-0.96	-1.60 – -0.32
Object-specific	0.37	-0.11 – 0.87	-0.43	-0.92 – 0.06	-0.27	-0.94 – 0.40
Language	1.66	1.25 – 2.07	0.34	-0.06 – 0.75	**-1.14**	**-1.71 – -0.58**
Multiple-demand	0.44	0.08 – 0.80	0.05	-0.32 – 0.40	0.12	-0.37 – 0.61

The following linear Bayesian regression model was specified: value ~ dynamics.d * interaction.d + (1|sub), with value representing the beta estimates extracted for each event for each ROI averaged across network, dynamics and type of interaction as fixed effects and a random intercept for participants (sub). Deviation coding style was used with 0.5: HHI/listening, and -0.5: HRI/speaking. Negative values for the main effects indicate a bias towards HRI or speaking, while positive values indicate a bias towards HHI or listening. CI (95%): 95% credibility interval

Table S8. Comparison of functional connectivity between interactions with a human and robotic agent

Network–network connectivity	*With global signal regression*			*Without global signal regression*			*Bayesian estimation*	
	M_{diff}	*t*	95% CI	M_{diff}	*t*	95% CI	*Estimates*	*CI (95%)*
tom–tom	-0.00	-0.28	-0.02, 0.02	-0.02	-1.68	-0.04, 0.00	-0.01	-0.07 – 0.04
ppn–ppn	0.01	0.83	-0.01, 0.02	-0.00	-0.14	-0.02, 0.0]	-0.01	-0.06 – 0.03
object–object	0.00	0.62	-0.01, 0.02	-0.00	-0.15	-0.02, 0.02	-0.02	-0.06 – 0.01
language–language	-0.00	-0.80	-0.01, 0.00	0.00060	0.09	-0.01, 0.01	-0.01	-0.04 – 0.03
demand–demand	0.01	2.02	-0.00032, 0.02	0.01	1.25	-0.01, 0.02	-0.02	-0.05 – 0.01
tom–ppn	0.01	1.32	-0.01, 0.02	-0.00	-0.15	-0.02, 0.01	-0.03	-0.07 – 0.01
tom–object	0.01	1.14	-0.01, 0.02	-0.00	-0.57	-0.02, 0.01	-0.02	-0.05 – 0.02
tom–language	0.01	2.65*	0.00, 0.02	0.02	2.19*	0.00092, 0.04	-0.02	-0.06 – 0.01
tom–demand	0.00	1.03	-0.00, 0.01	0.00	0.18	-0.01, 0.02	-0.02	-0.05 – 0.01
ppn–object	0.01	1.73	-0.00, 0.03	0.01	0.63	-0.01, 0.02	-0.03	-0.06 – 0.01
ppn–language	0.00	1.10	-0.00, 0.01	0.01	1.03	-0.01, 0.02	-0.02	-0.06 – 0.01
ppn–demand	0.01	2.86*	0.00, 0.02	0.01	1.23	-0.01, 0.03	-0.02	-0.05 – 0.00
object–language	0.00	0.32	-0.01, 0.01	0.00	0.32	-0.01, 0.01	-0.01	-0.04 – 0.03
object–demand	0.01	3.16*	0.00, 0.02	0.01	1.35	-0.01, 0.03	-0.03	-0.06 – 0.00
language–demand	0.01	2.18*	0.00053, 0.02	0.01	0.78	-0.01, 0.02	-0.02	-0.06 – 0.01

Paired-sample t-test was used to test for differences in functional correlation (Fisher z-transformation) between human-human and human-robot interactions for each within- and between-network combination. To test for robustness, these analyses were rerun with nuisance regression without global signal regression (but including nuisance signals from white matter, cerebrospinal fluid and global signal sources (Liu et al., 2017). For Bayesian re-analysis, the following model was specified: corz ~ run * interaction + (run | sub), with corz representing the Fisher z-transformed Pearson's correlation coefficients between the time courses for all possible combinations of ROIs within or between the network(s) with type of interaction (HHI, HRI) and run (1-4) as fixed effects and a random intercept and slope for participants (run | sub). Estimated posterior regression coefficient for the comparison between human-human and human-robot interaction are reported. * Did not survive Bonferroni correction ($p < .05 / 15$); M_{diff}: mean difference; 95% CI: 95% confidence interval for M_{diff}; tom: Theory-of-Mind; ppn: person perception; object: object-specific; demand: multiple-demand

Table S9. Comparison of within- and between-network functional connectivity during interactions with a human and robotic agent across the four runs

Network	*Main effect of run*			*Main effect of type of interaction*			*Interaction between run and type of interaction HRI*		
	F	*p*	η^2_G	*F*	p	η^2_G	*F*	*p*	η^2_G
Theory-of-Mind – Theory-of-Mind	1.48	.24	0.03	4.223044e-06	.998	1.001184e-08	0.81	.48	0.005
Person perception – Person perception	0.37	.71	0.01	1.14	.30	0.002	0.09	.94	0.0006
Object-specific – Object-specific	3.12	.04*	0.05	0.18	.68	0.0005	1.28	.29	0.008
Language – Language	0.44	.70	0.01	0.45	.51	0.0005	0.74	.51	0.003
Multiple-demand – Multiple-demand	0.89	.44	0.02	4.39	.05	0.008	0.38	.75	0.002
Theory-of-Mind – Person perception	0.44	.67	0.01	2.57	.12	0.004	0.48	.68	0.002
Theory-of-Mind – Object-specific	1.24	.30	0.02	1.43	.24	0.003	0.54	.65	0.003
Theory-of-Mind – Language	1.13	.34	0.02	6.92	.02*	0.008	0.32	.79	0.002
Theory-of-Mind – Multiple-demand	0.80	.49	0.01	1.18	.29	0.002	0.73	.50	0.003
Person perception – Object-specific	1.68	.19	0.02	3.07	.09	0.01	0.57	.61	0.003
Person perception – Language	1.68	.19	0.03	1.24	.28	0.001	1.34	.27	0.005
Person perception – Multiple-demand	1.27	.29	0.02	7.84	.01*	0.009	1.02	.38	0.003
Object-specific – Language	0.88	.45	0.02	0.15	.70	0.0001	0.24	.84	0.001
Object-specific – Multiple-demand	1.61	.21	0.02	10.32	.004*	0.01	0.46	.68	0.002
Language – Multiple-demand	1.65	.20	0.03	5.20	.03*	0.009	0.87	.45	0.004

Pearson's correlation coefficients (Fisher z-transformation) were calculated between time courses for all possible combinations of ROIs for each within- and between-network combination per interaction and entered in a 2 (interaction: HHI, HRI) by 4 (run: 1-4) repeated-measures ANOVA for each network separately. Significant effects are in bold, * Did not survive Bonferroni correction ($p < .05 / 15$); η^2_G: generalised eta-squared.

Table S10. Estimated posterior regression coefficient for the functional connectivity Bayesian linear model

Network	*Slope of HHI*		*Slope of HRI*		*Difference in slopes between HHI and HRI*	
	Estimates	*CI (95%)*	*Estimates*	*CI (95%)*	*Estimates*	*CI (95%)*
Theory-of-Mind – Theory-of-Mind	0.01	-0.01 – 0.03	0.01	-0.01 – 0.03	0.01	-0.02 – 0.03
Person perception – Person perception	0	-0.02 – 0.02	0	-0.01 – 0.02	0	-0.01 – 0.02
Object-specific – Object-specific	-0.01	-0.03 – -0.00	-0.01	-0.02 – 0.01	0.01	-0.01 – 0.02
Language – Language	0	-0.01 – 0.01	0	-0.01 – 0.01	0	-0.01 – 0.02
Multiple-demand – Multiple-demand	0	-0.01 – 0.01	0.01	-0.01 – 0.02	0	-0.01 – 0.01
Theory-of-Mind – Person perception	-0.03	-0.07 – 0.01	0.01	-0.01 – 0.03	0.01	-0.01 – 0.02
Theory-of-Mind – Object-specific	-0.02	-0.05 – 0.02	0	-0.02 – 0.01	0.01	-0.01 – 0.02
Theory-of-Mind – Language	-0.02	-0.06 – 0.01	0	-0.01 – 0.01	0.00	-0.01 – 0.02
Theory-of-Mind – Multiple-demand	-0.02	-0.05 – 0.01	0	-0.01 – 0.01	0.01	-0.01 – 0.02
Person perception – Object-specific	-0.03	-0.06 – 0.01	0	-0.02 – 0.01	0.01	-0.01 – 0.02
Person perception – Language	-0.02	-0.06 – 0.01	0	-0.01 – 0.01	0.01	-0.01 – 0.02
Person perception – Multiple-demand	-0.02	-0.05 – 0.00	0	-0.01 – 0.01	0.00	-0.01 – 0.01
Object-specific – Language	-0.01	-0.04 – 0.03	0	-0.01 – 0.01	0.00	-0.01 – 0.01
Object-specific – Multiple-demand	-0.03	-0.06 – 0.00	0	-0.01 – 0.01	0.00	-0.01 – 0.02
Language – Multiple-demand	-0.02	-0.06 – 0.01	0	-0.01 – 0.02	0.01	-0.01 – 0.02

For Bayesian re-analysis, the following model was specified: corz ~ run * interaction + (run | sub), with corz representing the Fisher z-transformed Pearson's correlation coefficients between the time courses for all possible combinations of ROIs for each within- and between-network combination with type of interaction (HHI, HRI) and run (1-4) as fixed effects and a random intercept and slope for participants (run | sub). The hypothesis function was used to specify the slope for the HRI: run + run:interactionhri = 0. CI (95%): 95% credibility interval

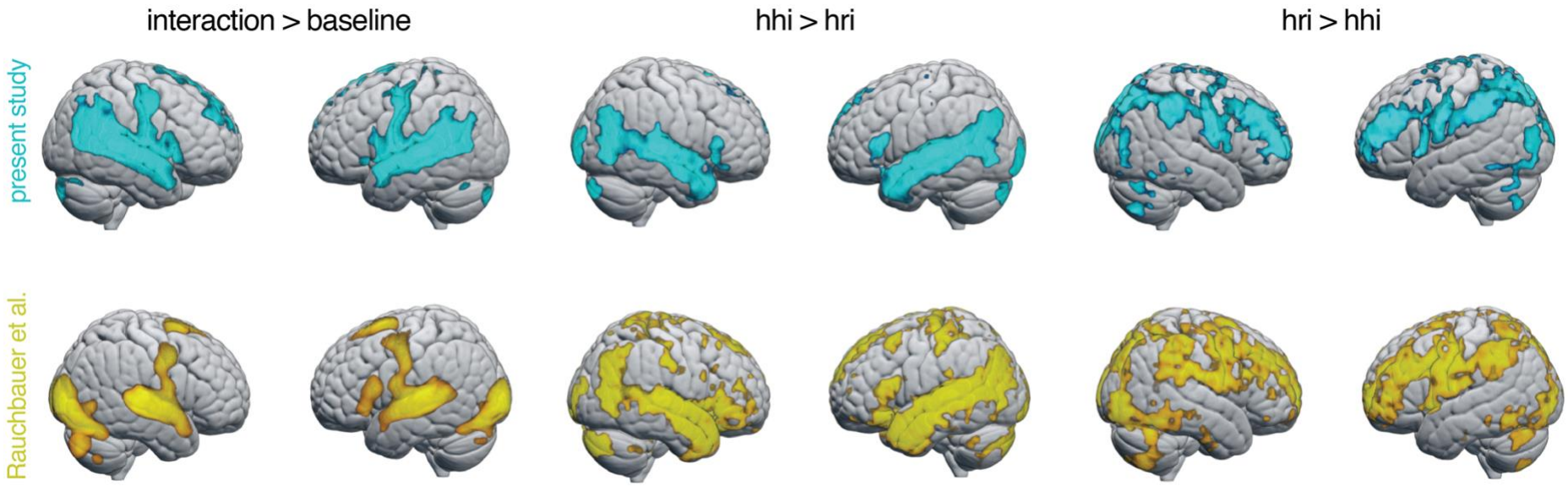

Figure S1. Comparison of the whole-brain brain analyses between the present study and the study of Rauchbauer and colleagues (2019). Reanalysis of the contrasts used in Rauchbauer and colleagues (2019) resulted in similar activation maps when contrasting the interactions with a human and robotic agent with the implicit baseline (c.f. instruction screen), human-human interactions (hhi) versus human-robot interactions (hri), and vice versa. For visual purposes contrast maps are shown. The maps from Rauchbauer and colleagues (2019) are obtained from https://neurovault.org/images/112528/.

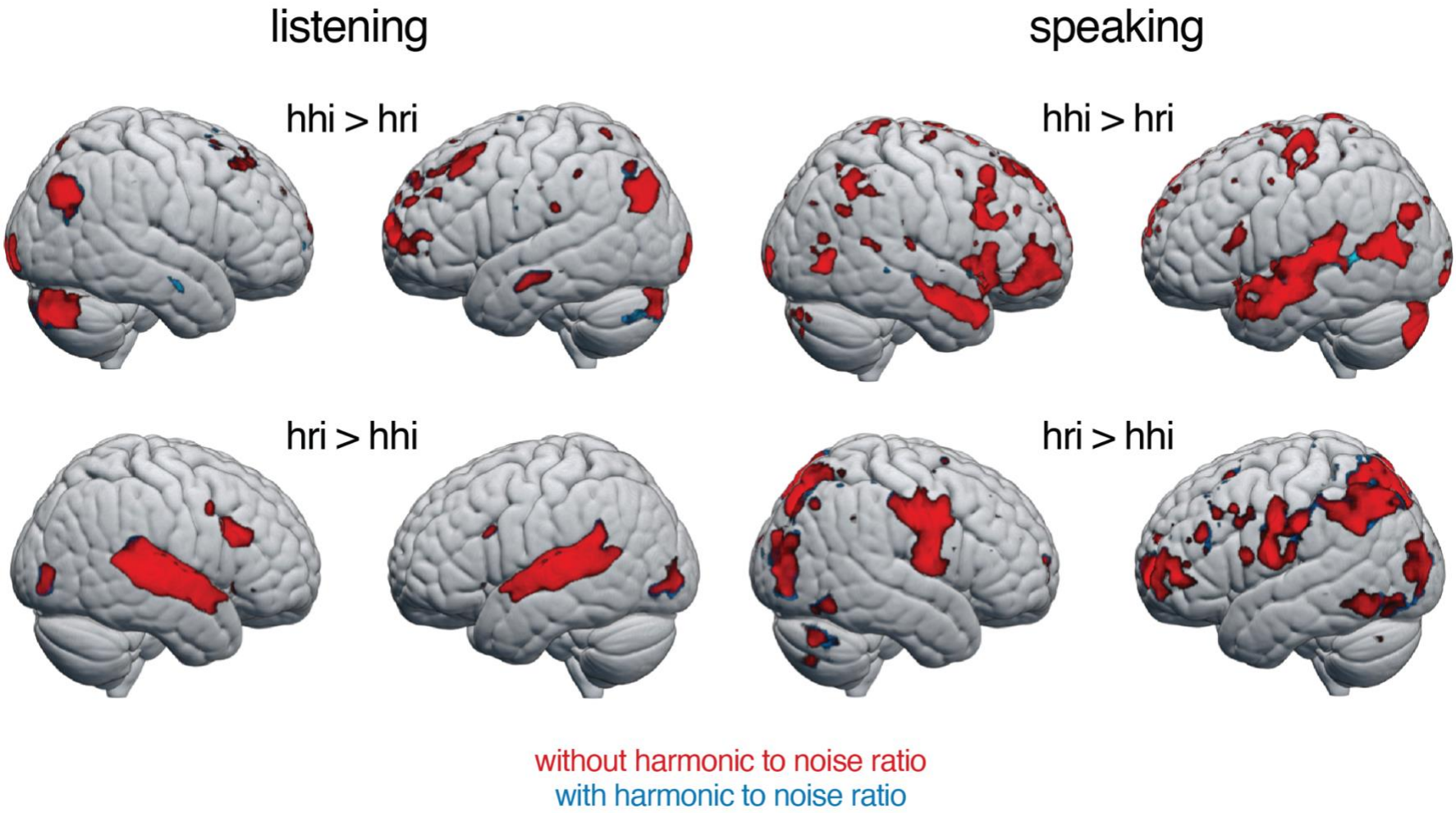

Figure S2. Comparison of the whole-brain brain analyses with and without harmonic to noise ratio as a parametric predictor. Similar effects were obtained when controlling for voice quality of the agent or general signal-to-noise using the mean harmonic to noise ratio (centred and scaled) as a parametric predictor. For visual purposes activation maps are shown with an uncorrected threshold of $p < .001$ ($k = 10$).

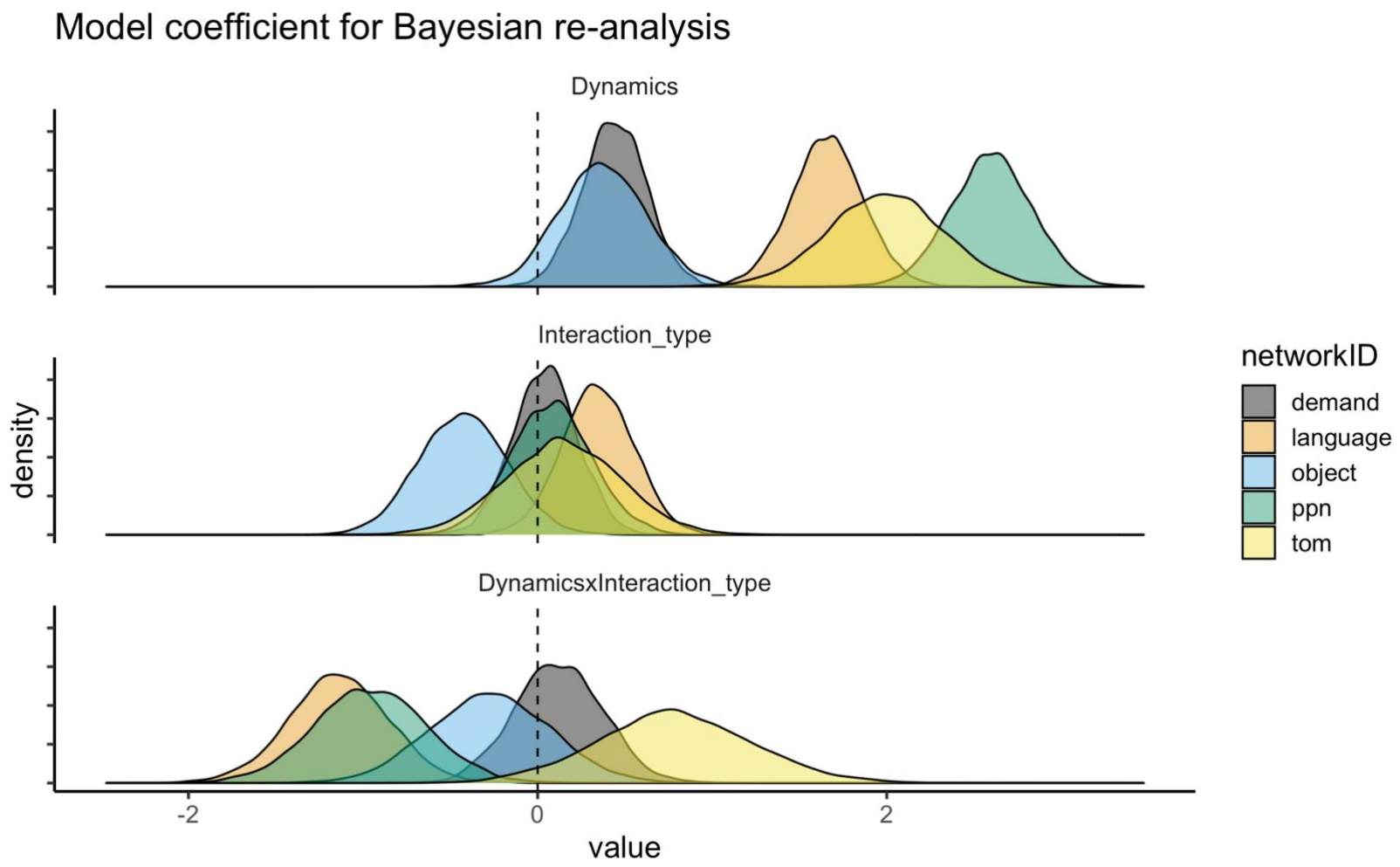

Figure S3. Posterior distribution for the region-of-interest Bayesian linear model. The following linear Bayesian regression model was specified: value ~ dynamics.d * interaction.d + (1|sub), with value representing the beta estimates extracted for each event for each ROI averaged across network, and dynamics and type of interaction as fixed effects and a random intercept for participants (sub). Deviation coding style was used with 0.5: HHI/listening, and -0.5: HRI/speaking. Negative values for the main effects indicate a bias towards HRI or speaking, while positive values indicate a bias towards HHI or listening.

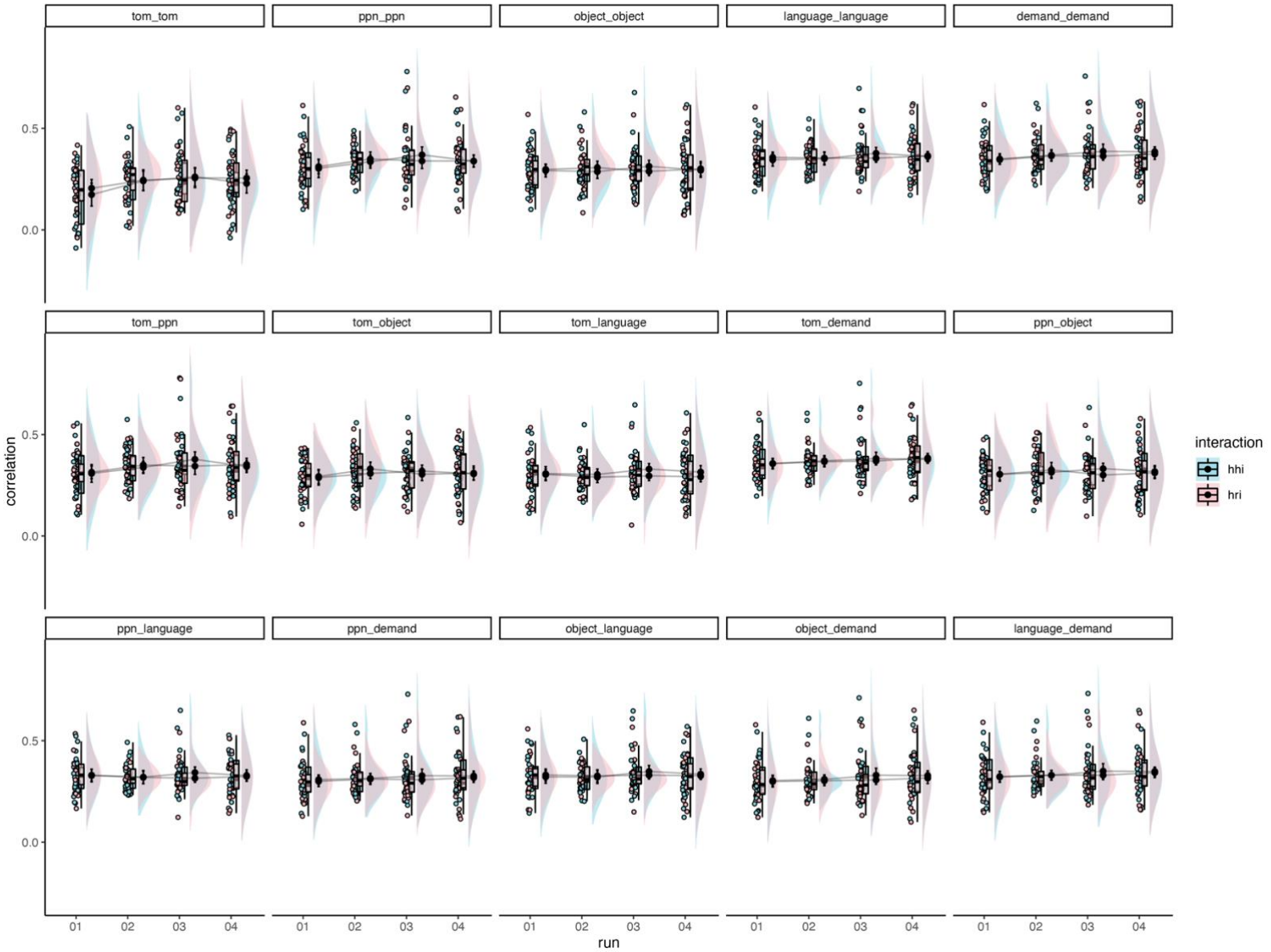

Figure S4. No effect of repeated experience on functional connectivity during interactions with a human and robotic agent. Within- and between-network z-transformed correlations are shown for the four runs and interaction type (human-human or human-robot interaction) separately. Rain cloud plots with errors bars reflecting 95% confidence intervals are used (Allen et al., 2021).

References

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., van Langen, J., & Kievit, R. A. (2021). Raincloud plots: A multi-platform tool for robust data visualization [version 2; peer review: 2 approved]. *Wellcome Open Research*, *4*, 63. https://doi.org/10.12688/wellcomeopenres.15191.2

Diachek, E., Blank, I., Siegelman, M., Affourtit, J., & Fedorenko, E. (2020). The Domain-General Multiple Demand (MD) Network Does Not Support Core Aspects of Language Comprehension: A Large-Scale fMRI Investigation. *The Journal of Neuroscience*, *40*(23), 4536–4550. https://doi.org/10.1523/JNEUROSCI.2036-19.2020

Dubey, I., Georgescu, A. L., Hommelsen, M., Vogeley, K., Ropar, D., & Hamilton, A. F. de C. (2020). Distinct neural correlates of social and object reward seeking motivation. *European Journal of Neuroscience*, *52*(9), 4214–4229. https://doi.org/10.1111/ejn.14888

Henschel, A., Hortensius, R., & Cross, E. S. (2020). Social Cognition in the Age of Human–Robot Interaction. *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2020.03.013

Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage*, *60*(4), 2357–2364. https://doi.org/10.1016/j.neuroimage.2012.02.055

Liu, T. T., Nalci, A., & Falahpour, M. (2017). The global signal in fMRI: Nuisance or Information? *NeuroImage*, *150*, 213–229. https://doi.org/10.1016/j.neuroimage.2017.02.036

Rauchbauer, Nazarian Bruno, Bourhis Morgane, Ochs Magalie, Prévot Laurent, & Chaminade Thierry. (2019). Brain activity during reciprocal social interaction investigated using conversational robots as control condition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1771), 20180033. https://doi.org/10.1098/rstb.2018.0033

Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., & Saxe, R. (2018). Development of the social brain from age three to twelve years. *Nature Communications*, *9*(1), 1027. https://doi.org/10.1038/s41467-018-03399-2