Supplementary Materials

The State and Expectations of Interrater Reliability in Nonhuman Personality Research

Alexander J. Pritchard

Neuroscience and Behavior Unit, California National Primate Research Center, University of California Davis, Davis, CA, USA

Department of Population Health & Reproduction, School of Veterinary Medicine, University of California Davis, Davis, CA, USA

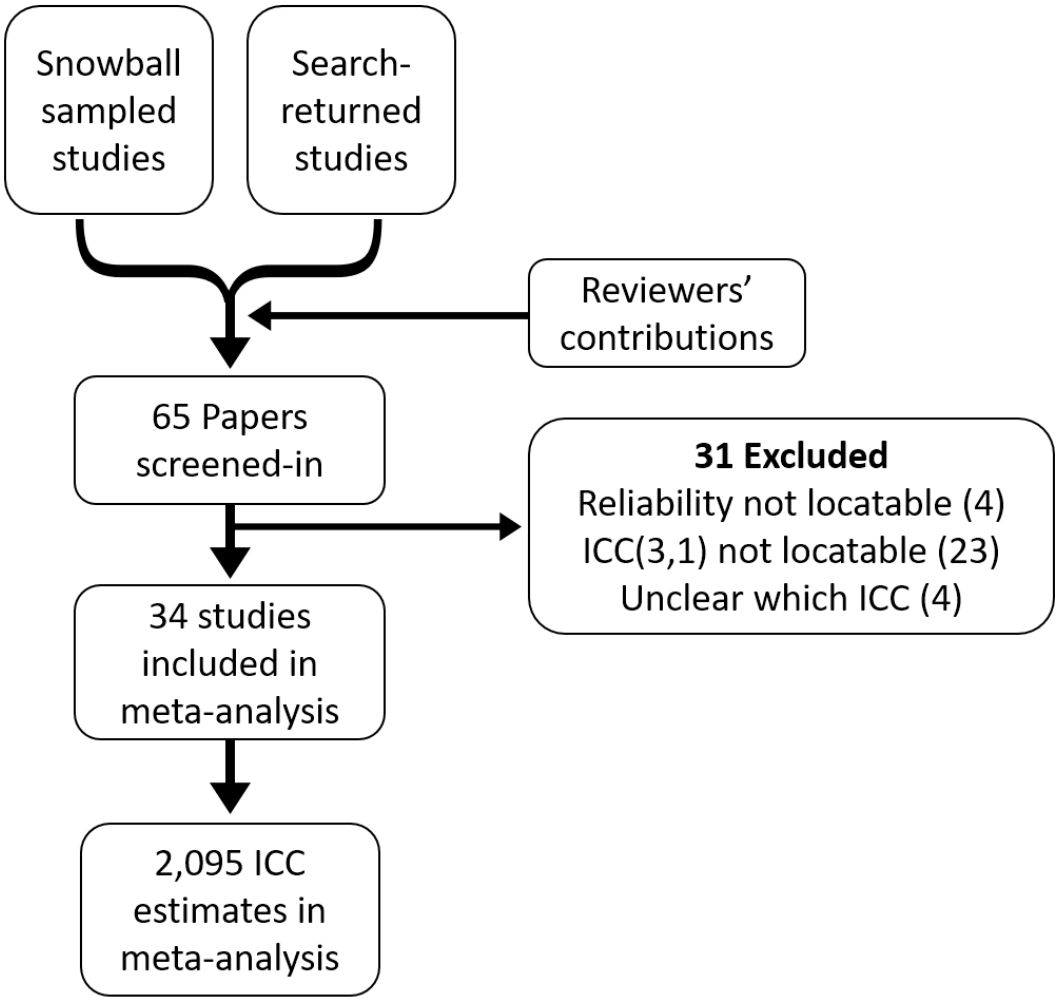
Author Note

Alexander J. Pritchard https://orcid.org/0000-0003-4922-3347

There are no original data published in this work. My simulation and meta-analytic code, as well as the ICC data from external sources and references, are available at: https://doi.org/10.17605/OSF.IO/J8RFX (Pritchard, 2025)

Correspondence concerning this article should be addressed to Alexander J. Pritchard, University of California, Davis, One Shields Avenue, Davis, CA, 95616.

Email (preferred): ajpritchard@ucdavis.edu

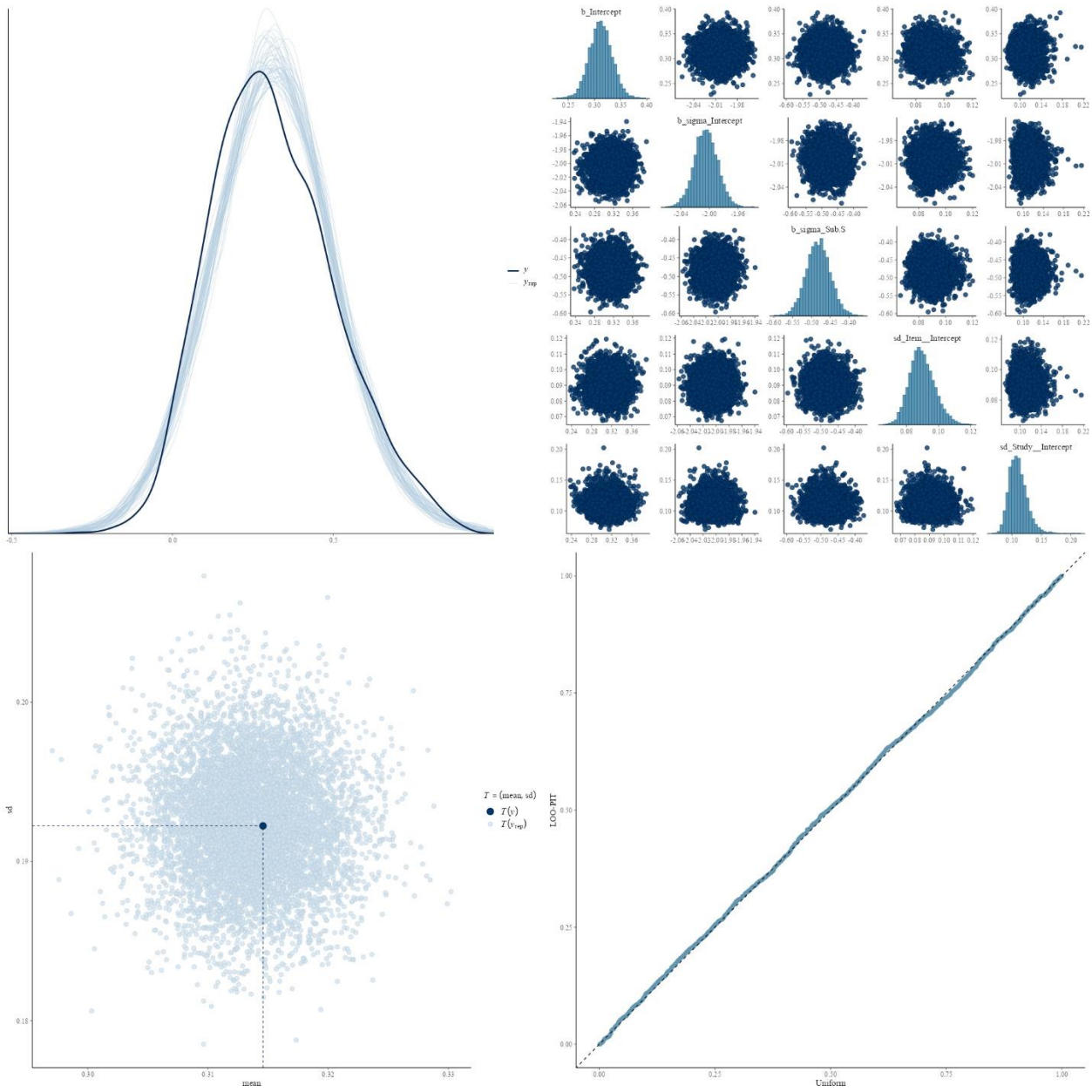


Supplementary Figure 1. Flowchart of article inclusion.

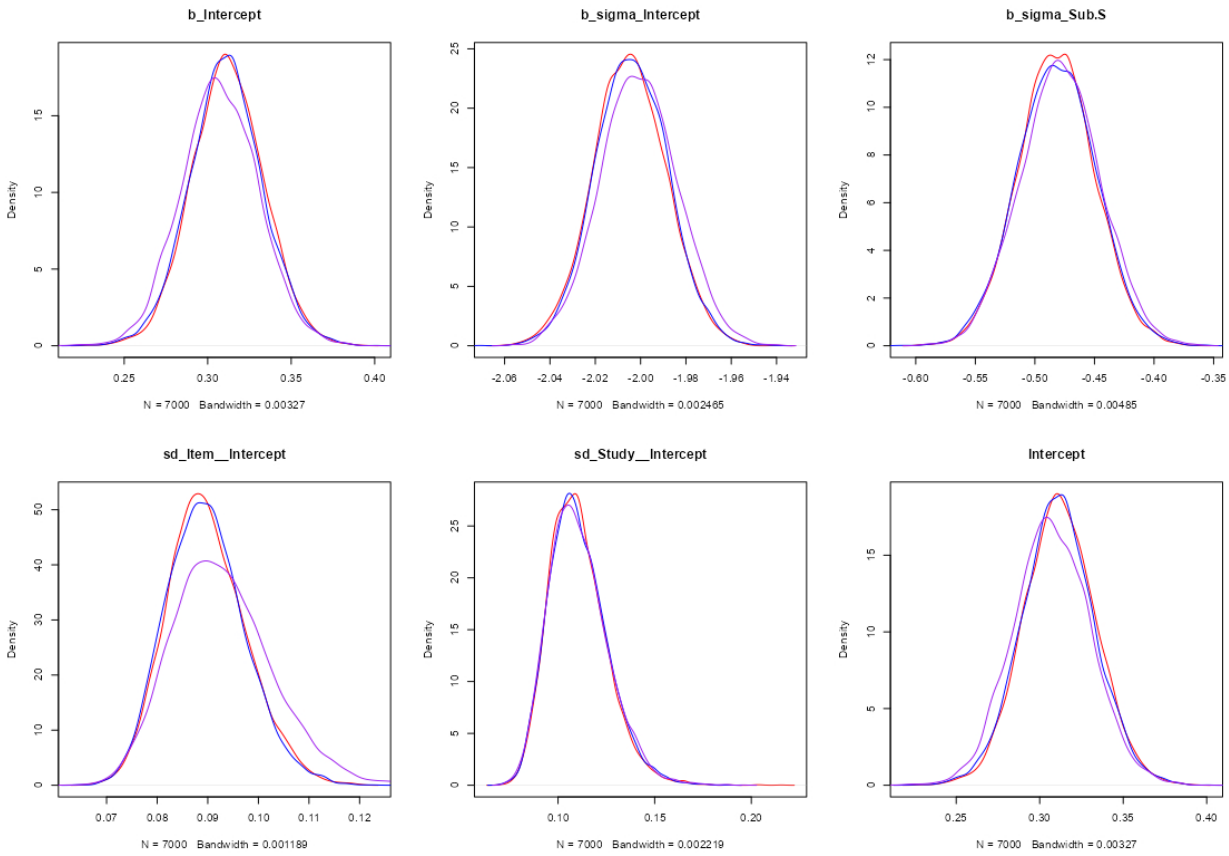
Supplementary Table 1. Studies included in the final model. Mean values are reported for rater and subjects counts.

Study	Year	Taxon	Raters	Subjects	Notes
Adams et al.	2015	*Macaca assamensis*	5.6	60	
		Macaca fuscata	1.4	74	
		Macaca mulatta	13.26	117.33	
		Macaca nigra	2	53	
		Macaca sylvanus	2.4	74	
		Macaca tonkeana	4	46	
Arlet et al.	2024	*Macaca radiata*	3	52	Visual inspection of ICC estimates with more raters but fewer subjects suggests that three raters might be representative
Baker et al.	2015	*Macaca nigra*	2	64	Rater number not reported but was more than one; I assumed two raters, which imbues the highest uncertainty
		Macaca sylvanus	2	62	
		Saimiri sciureus	2	69	
Eckardt et al.	2015	*Gorilla beringei beringei*	4.8	116	
Fernández-Bolaños et al.	2020	*Sapajus xanthosternos*	2.81	26	
Freeman et al.	2013	*Pan troglodytes*	12.36	99	Estimated the number of raters as: 12.36 72 (mean subjects per rater) * 17 (raters) / 99 (subjects)
Garai et al.	2016	*Pan paniscus*	5	16	
Hopper et al.	2018	*Macaca fuscata*	2	63	
Iwanicki & Lehmann	2015	*Callithrix jacchus*	4	63	
Konečná et al.	2008	*Semnopithecus entellus*	3.96	27	
Konečná et al.	2012	*Macaca sylvanus*	2	23	Only used season one
Koski et al.	2017	*Callithrix jacchus*	2.84	100	
Morton et al.	2013	*Sapajus apella*	3.35	121	
Murray & Diaz	2025	*Pan troglodytes*	5	28	
Parry-Howells et al.	2023	*Macaca nigra*	3.06	13	
Pohlner	2023	*Gorilla gorilla gorilla*	2.8	167	*In situ* population
			3.54	184	*Ex situ* population

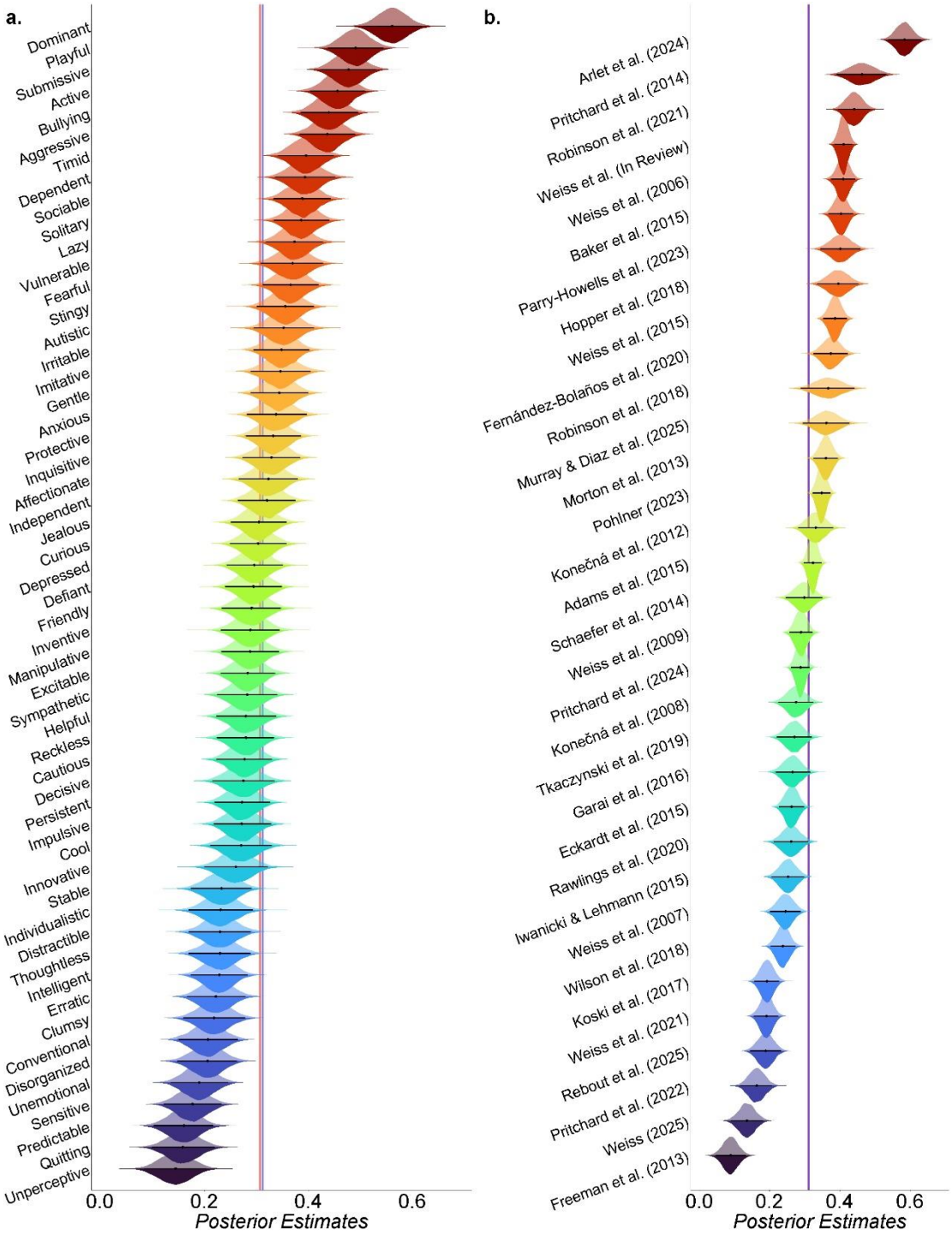
Study	Year	Taxon	Raters	Subjects	Notes
Pritchard et al.	2014	*Macaca thibetana*	6	12	
Pritchard & Palombit	2022	*Papio anubis*	4	44	Discarded second replicate used for intrarater reliability
Pritchard et al.	2024	*Macaca fascicularis*	2.42	101	
		Macaca mulatta	2.66	150	
		Macaca radiata	4.1	77	
Rawlings et al.	2020	*Pan troglodytes*	8	50	
Rebout et al.	2025	*Plecturocebus cupreus*	3	54	
Robinson et al.	2018	*Macaca mulatta*	2.23	85	
Robinson et al.	2021	*Macaca mulatta*	2	44	
Schaefer & Steklis	2014	*Gorilla gorilla gorilla*	9	8	
Tkaczynski et al.	2019	*Macaca sylvanus*	8	27	
Úbeda & Llorente	2015	*Pan troglodytes*	28	11	
Weiss et al.	2006	*Pongo* spp.	3	152	
Weiss et al.	2007	*Pan troglodytes*	2	75	
Weiss et al.	2009	*Pan troglodytes*	3.2	146	
Weiss et al.	2015	*Pan paniscus*	2.69	81	Wave one
			2.88	40	Wave two with a 5 year ±0.4*sd* separation from wave one
Weiss et al.	2021	*Callithrix jacchus*	2.63	128	
Weiss	2025	*Pan troglodytes*	7	40	
Weiss et al.	*In review*	Hylobatidae	2.49	247	
Wilson et al.	2018	*Saimiri boliviensis*	2.7	31	
		Saimiri sciureus	1.91	29	



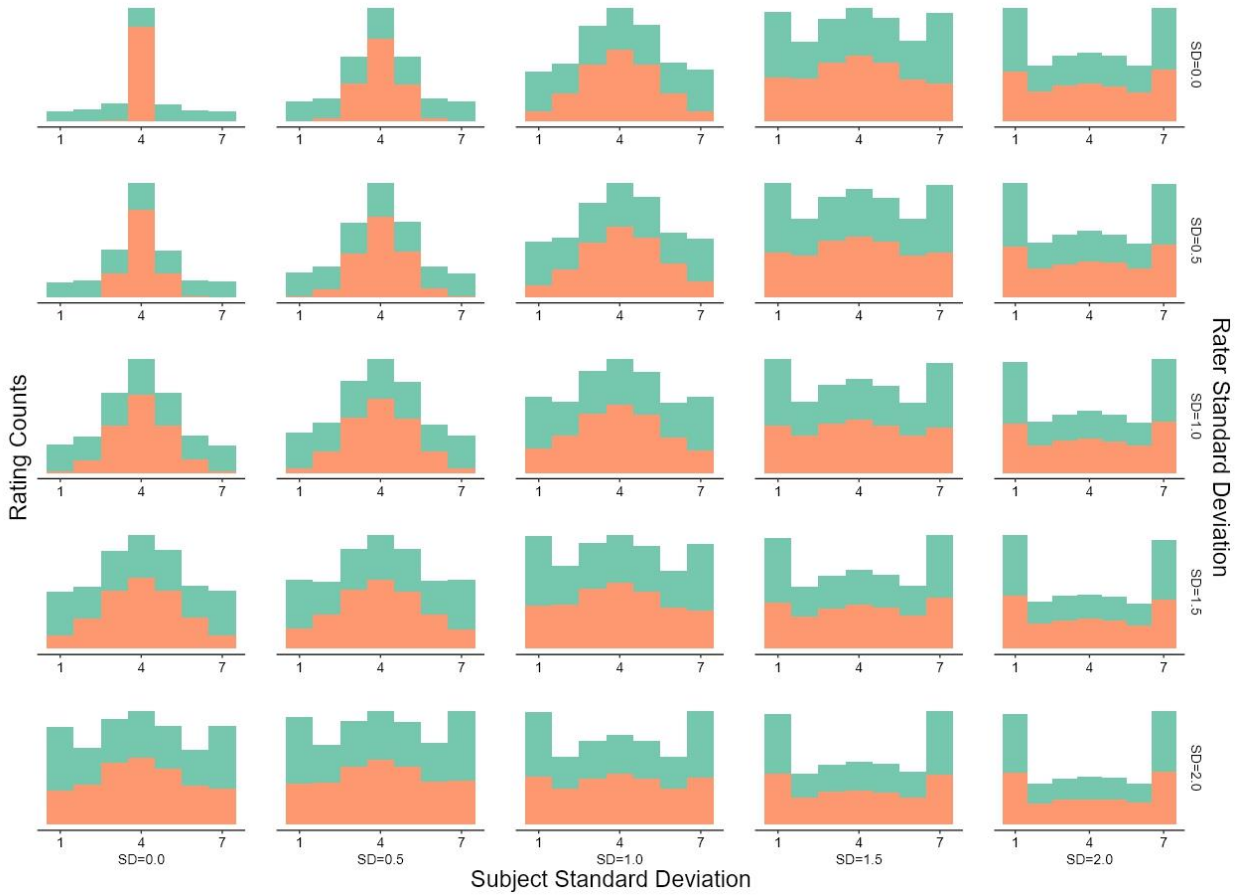
Supplementary Figure 2. Posterior predictive check and pairs plots for the final models.



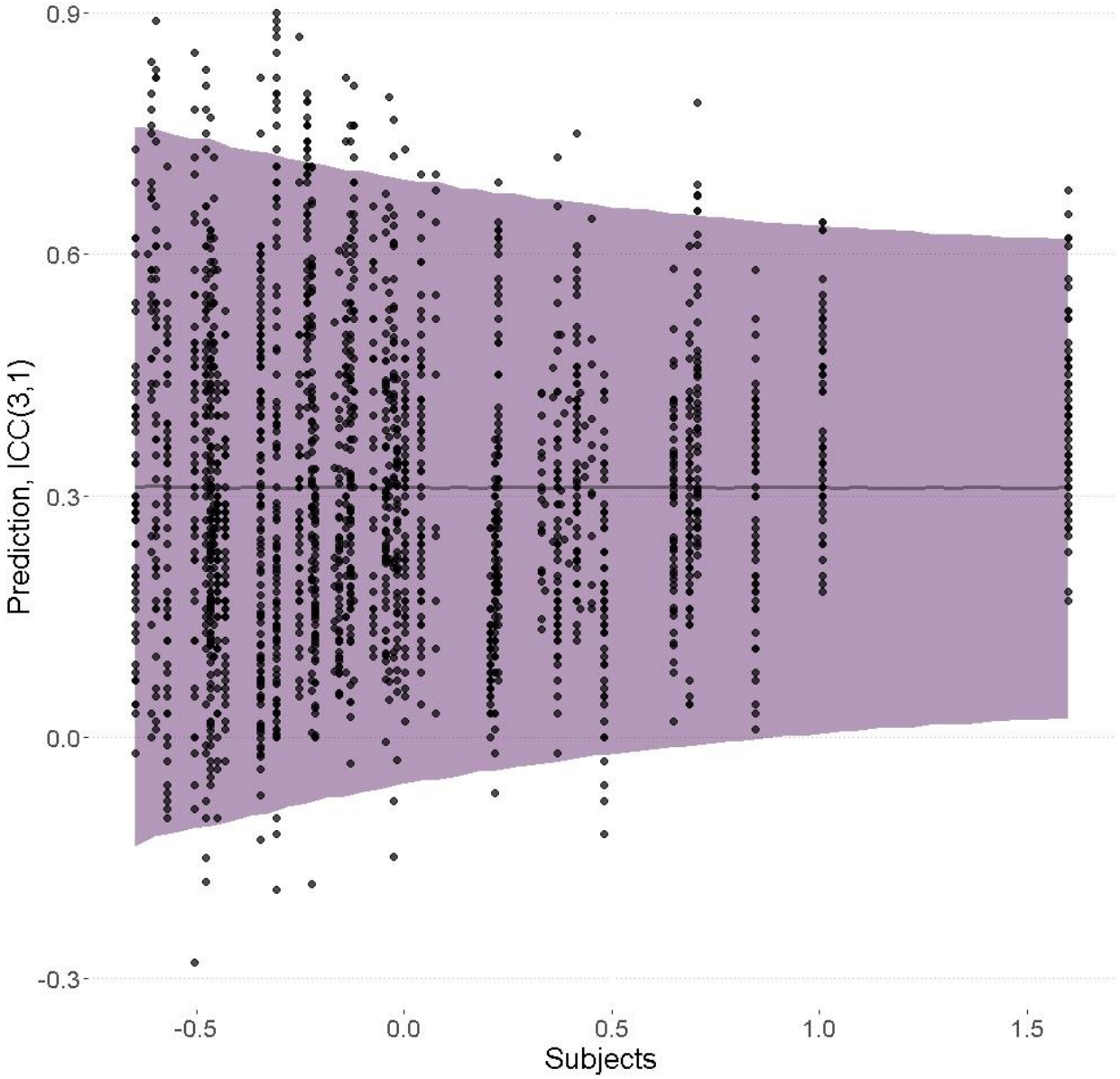
Supplementary Figure 3. Posterior parameter estimates for the final model (red) versus models constrained to studies with ≥ 4 reported ICC(3,1) estimates (blue) and items with ≥ 15 occurrences (purple). Note the high convergence around estimates means.



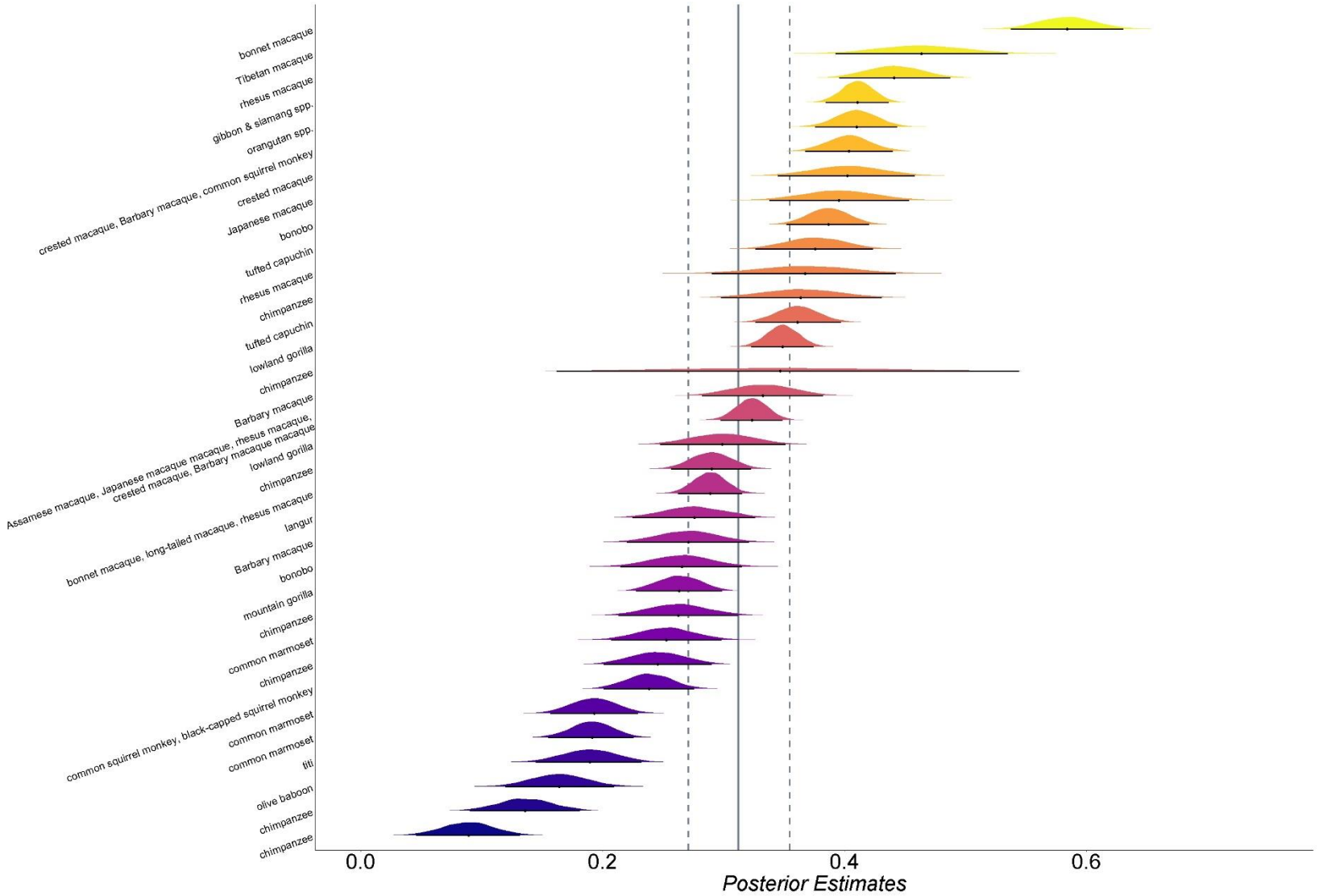
Supplementary Figure 4. Item (**a.**) and study (**b.**) estimates for the effect sizes of ICC(3,1) (x-axis). Fill color is used to emphasize items or studies, labelled on the y-axes. Eye densities above each slab show the posterior draws from the final model for a given study or item, with interval point and slab emphasizing the average and 95% interval, respectively. Eye densities below each slab show alternative model structures with fewer items (**a.**) or studies (**b.**). The central flanking vertical lines show the model average estimate (y-intercept) of the final model (blue) and the alternative models (red), note that the nearly complete overlap results in a purple appearance.



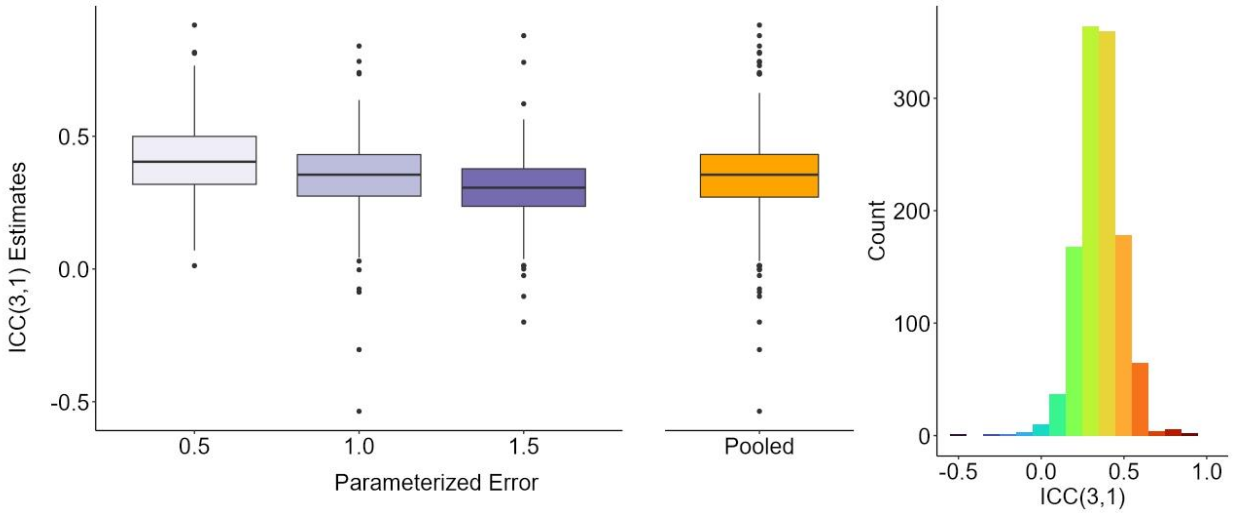
Supplementary Figure 5. Simulated distributions of ratings with different parameter estimates for subject (columns) and raters (rows). Each plot has two distributions with minimal error (0.2 – in teal) and maximal error (2.5 – in orange).



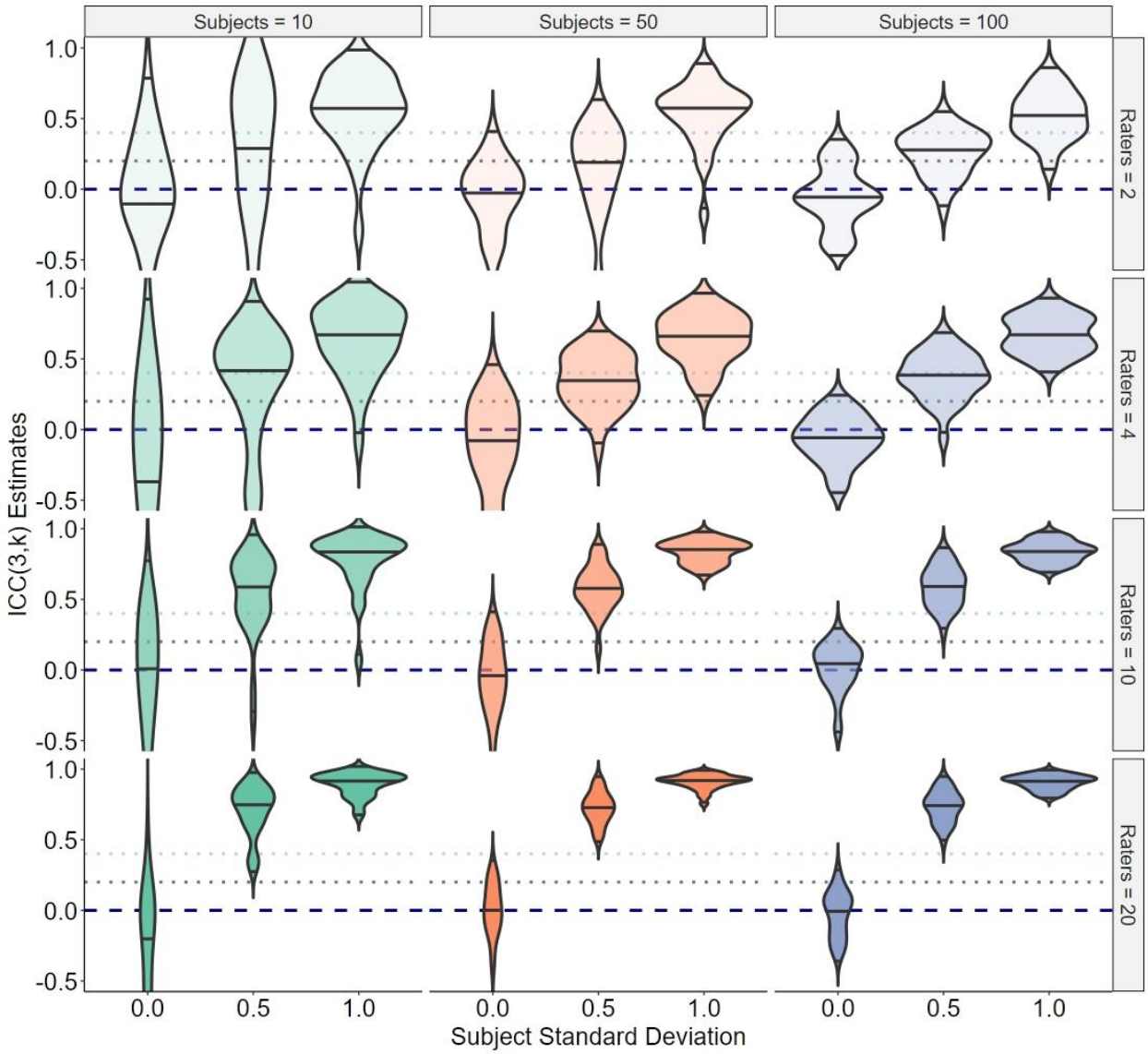
Supplementary Figure 6. The final model estimates based on the number of subjects, fit to σ. The x-axis is scaled by two standard deviations.



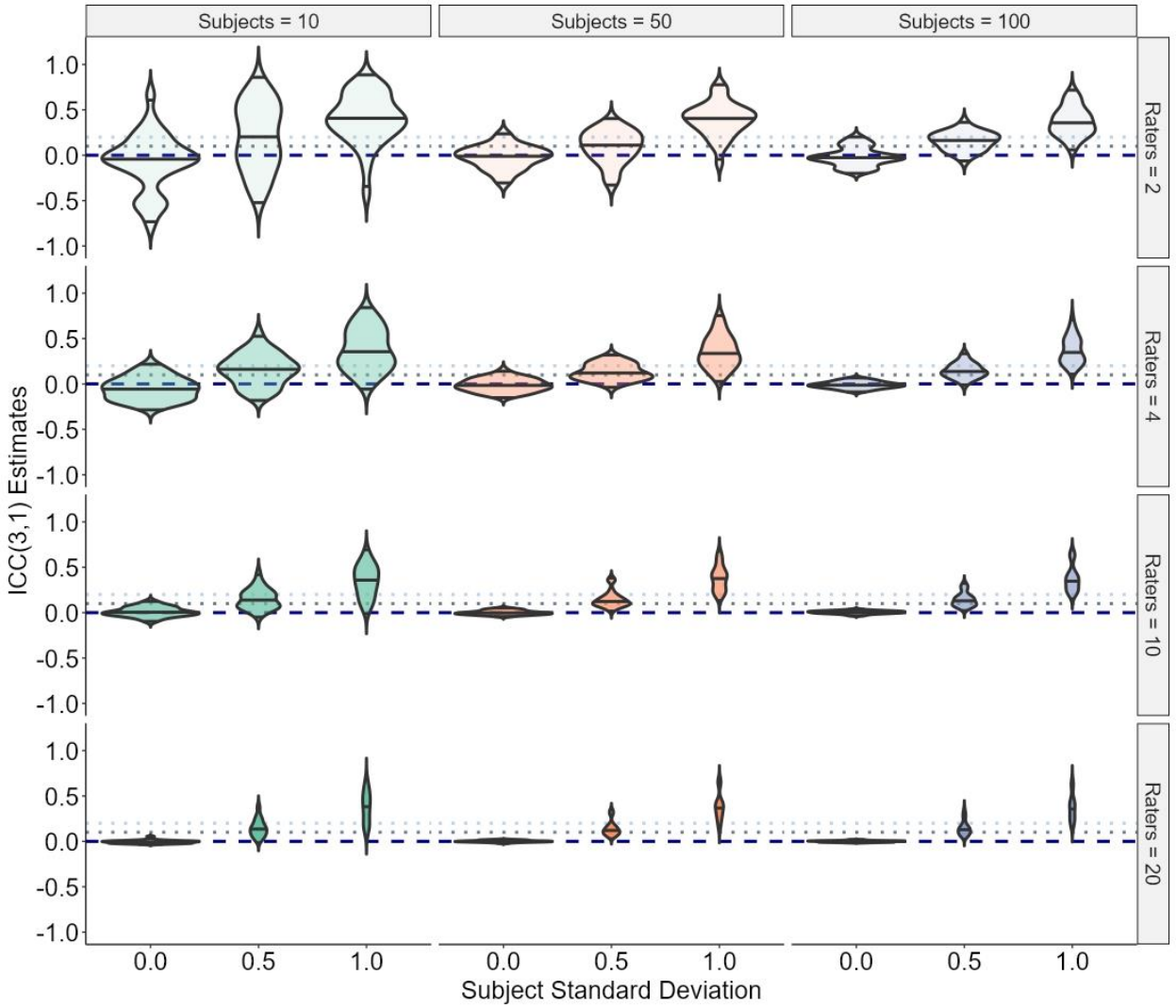
Supplementary Figure 7. Study estimates for the effect sizes of ICC(3,1) (x-axis), as in **Figure 2b**. Y-axis labels emphasize taxa within each study. Densities show posterior draws for each study, with interval point and slab emphasizing the average and 95% interval. Central and flanking vertical lines show the model average estimate (y-intercept) and its 95% credible intervals.



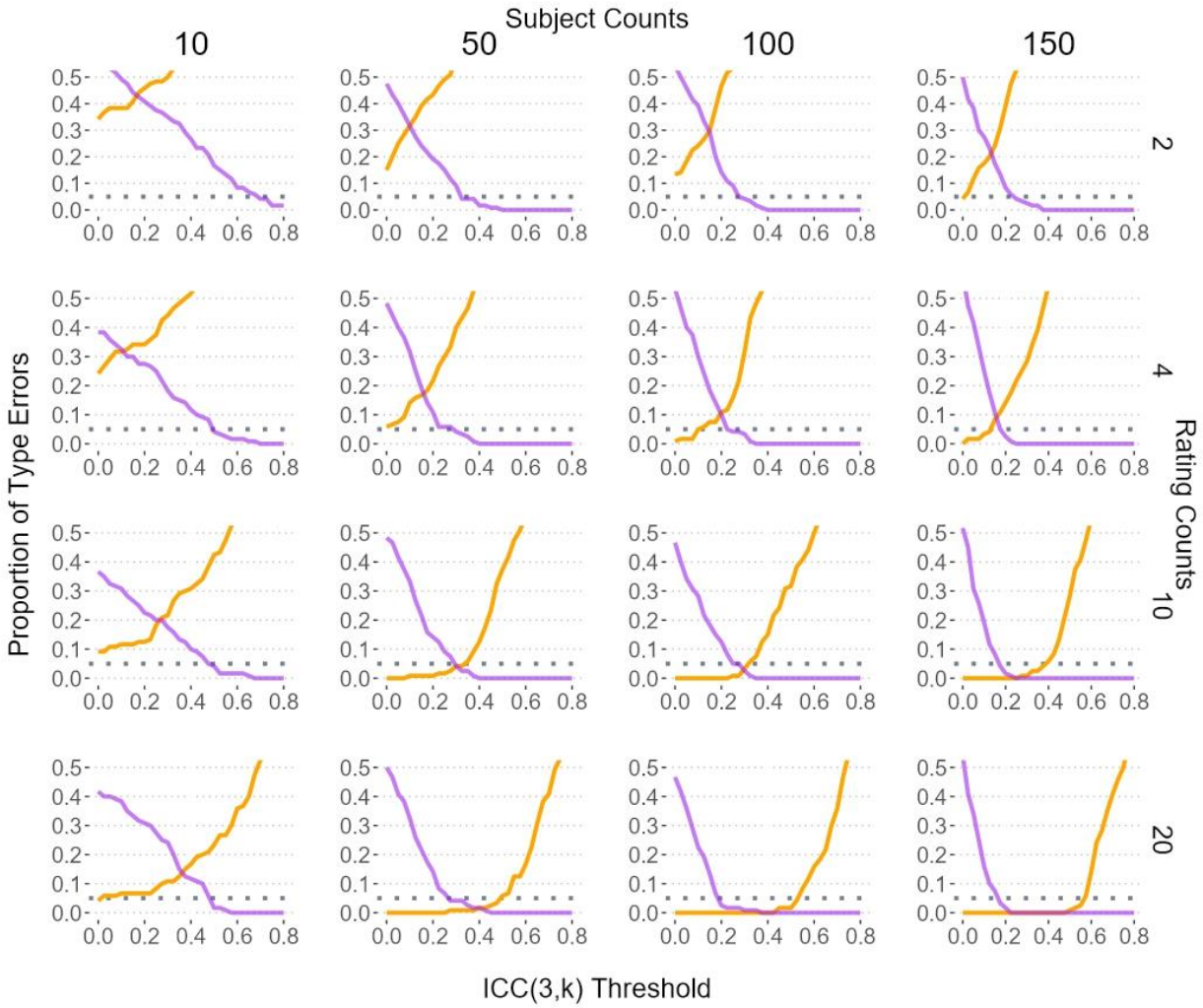
Supplementary Figure 8. Simulated ICC estimates with unreasonably high rater or subject variance (1.5 or 2.0) presented as a boxplot and histogram. Boxplots show the simulated ICC(3,1) estimates with several parameterized values for unattributed error (purple fill; x-axis) or for the pooled values (orange fill). The histogram shows the count of these ICC estimates by bins of 0.1.



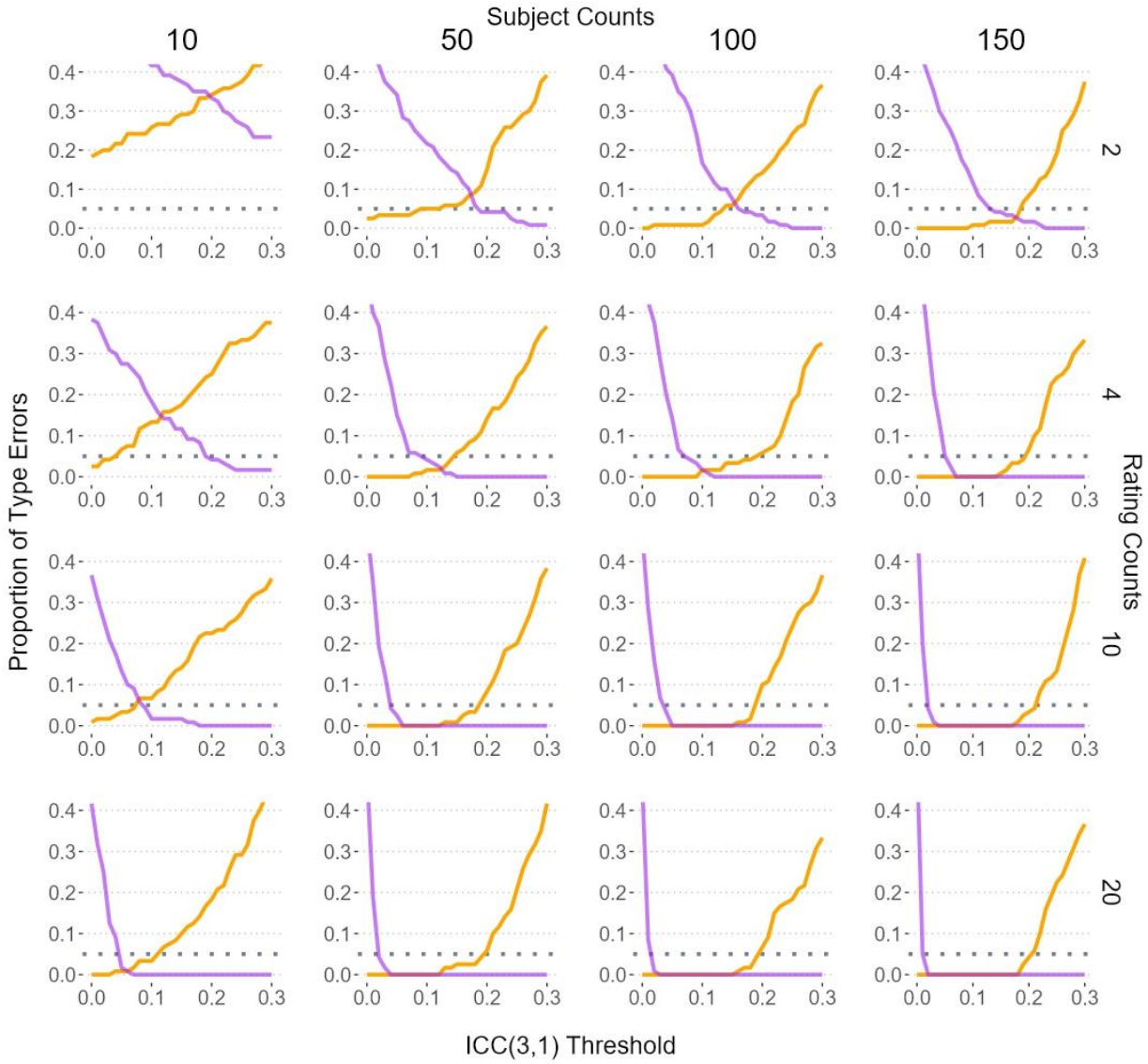
Supplementary Figure 9. Violin plots of ICC(3,k) (y-axes) for different numbers of subjects (columns; color fill) and raters (rows; hue fill). Subject variance parameters are on the x-axes. Dashed lines emphasize an ICC(3,k) of 0.00, with dotted lines at 0.20 and 0.40. Violins have quartile lines set at 2.5%, 50%, and 97.5%. These plots are obtained by extracting a subset of the simulation results for rater and subject variance parameters to 0.5 or 1.0, unattributed error to 0.5, 1.0, or 1.5, subject counts to 10, 50, and 100 (150 or 200 subjects performed similarly to 100), and rater counts to 2, 4, 10, and 20.



Supplementary Figure 10. Violin plots of ICC(3,1) (y-axes) for different numbers of subjects (columns; color fill) and raters (rows; hue fill). Subject variance parameters are on the x-axes. Dashed lines emphasize an ICC(3,1) of 0.00, with dotted lines at 0.10 and 0.20. Violins have quartile lines set at 2.5%, 50%, and 97.5%. These plots are obtained by extracting a subset of the simulation results for rater and subject variance parameters to 0.5 or 1.0, unattributed error to 0.5, 1.0, or 1.5, subject counts to 10, 50, and 100 (150 or 200 subjects performed similarly to 100), and rater counts to 2, 4, 10, and 20.



Supplementary Figure 11. Assuming a true subject variance of 0.0 (purple) or 0.5 (orange), proportion of errors (y-axis) with differing ICC(3,k) reliability cut-offs (x-axis). Purple lines, with a negative association, show erroneous acceptance of items that have no subject variance. Orange lines, with a positive or null association, show erroneous rejection of items that have known subject variance. Subject and rater counts are shown across columns and rows, respectively.



Supplementary Figure 12. Assuming a true subject variance of 0.0 (purple) or 1.0 (orange), proportion of errors (y-axis) with differing ICC(3,1) reliability cut-offs (x-axis). Purple lines, with a negative association, show erroneous acceptance of items that have no subject variance. Orange lines, with a positive or null association, show erroneous rejection of items that have known subject variance. Subject and rater counts are shown across column and row panels, respectively.

References

Adams, M. J., Majolo, B., Ostner, J., Schülke, O., De Marco, A., Ethoikos, F., di Piano dell'Abatino, P. F., Thierry, B., Engelhardt, A., Center, G. P., & others. (2015). Personality structure and social style in macaques. *Journal of Personality and Social Psychology*, *109*(2), 338–353. https://doi.org/10.1037/pspp0000041

Arlet, M. E., Sakson-Obada, O., Zakrzewska, M., & Singh, M. (2024). Personality traits and their correlation with behavior in the bonnet macaque (*Macaca radiata*) in southern India. *Scientific Reports*, *14*(1), 31071. https://doi.org/10.1038/s41598-024-82283-0

Baker, K. R., Lea, S. E. G., & Melfi, V. A. (2015). Comparative personality assessment of three captive primate species: *Macaca nigra*, *Macaca sylvanus*, and *Saimiri sciureus*. *International Journal of Primatology*, *36*(3), 625–646. https://doi.org/10.1007/s10764-015-9843-3

Eckardt, W., Steklis, H. D., Steklis, N. G., Fletcher, A. W., Stoinski, T. S., & Weiss, A. (2015). Personality dimensions and their behavioral correlates in wild Virunga mountain gorillas (*Gorilla beringei beringei*). *Journal of Comparative Psychology*, *129*(1), 26–41. https://doi.org/10.1037/a0038370

Fernández-Bolaños, M., Delval, I., de Oliveira, R. S., & Izar, P. (2020). Assessing the personality structure of wild capuchin monkeys (*Sapajus xanthosternos*) using trait rating and behavioral coding. *Journal of Comparative Psychology*, *134*(3), 349-360. https://doi.org/10.1037/com0000219

Freeman, H. D., Brosnan, S. F., Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Gosling, S. (2013). Developing a comprehensive and comparative questionnaire for measuring personality in chimpanzees using a simultaneous top-down/bottom-up design: Chimpanzee personality. *American Journal of Primatology*, *75*(10), 1042–1053. https://doi.org/10.1002/ajp.22168

Garai, C., Weiss, A., Arnaud, C., & Furuichi, T. (2016). Personality in wild bonobos (*Pan paniscus*). *American Journal of Primatology*, *78*(11), 1178–1189. https://doi.org/10.1002/ajp.22573

Hopper, L. M., Cronin, K. A., & Ross, S. R. (2018). A multi-institutional assessment of a short-form personality questionnaire for use with macaques. *Zoo Biology*, *37*(5), 281–289. https://doi.org/10.1002/zoo.21439

Iwanicki, S., & Lehmann, J. (2015). Behavioral and trait rating assessments of personality in common marmosets (*Callithrix jacchus*). *Journal of Comparative Psychology*, *129*(3), 205–217. https://doi.org/10.1037/a0039318

Konečná, M., Lhota, S., Weiss, A., Urbánek, T., Adamová, T., & Pluháček, J. (2008). Personality in free-ranging Hanuman langur (*Semnopithecus entellus*) males: Subjective ratings and recorded behavior. *Journal of Comparative Psychology*, *122*(4), 379–389. https://doi.org/10.1037/a0012625

Konečná, M., Weiss, A., Lhota, S., & Wallner, B. (2012). Personality in Barbary macaques (*Macaca sylvanus*): Temporal stability and social rank. *Journal of Research in Personality*, *46*(5), 581–590. https://doi.org/10.1016/j.jrp.2012.06.004

Koski, S. E., Buchanan-Smith, H. M., Ash, H., Burkart, J. M., Bugnyar, T., & Weiss, A. (2017). Common marmoset (*Callithrix jacchus*) personality. *Journal of Comparative Psychology*, *131*(4), 326–336. https://doi.org/10.1037/com0000089

Morton, F. B., Lee, P. C., Buchanan-Smith, H. M., Brosnan, S. F., Thierry, B., Paukner, A., de Waal, F., Widness, J., Essler, J. L., & Weiss, A. (2013). Personality structure in brown capuchin monkeys (*Sapajus apella*): Comparisons with chimpanzees (*Pan troglodytes*), orangutans (*Pongo spp.*), and rhesus macaques (*Macaca mulatta*). *Journal of Comparative Psychology*, *127*(3), 282–298. https://doi.org/10.1037/a0031723

Murray, L. E., & Diaz, S. (2025). Long-term stability of chimpanzee personality: Comparing trait ratings and behaviour codings over a quarter of a century. *Animal Behaviour*, 227, 123241. https://doi.org/10.1016/j.anbehav.2025.123241

Parry-Howells, N., Baker, K., & Farmer, H. L. (2023). The link between personality, subjective well-being, and welfare in zoo-housed Sulawesi crested macaques (*Macaca nigra*). *International Journal of Primatology*, *44*(5), 984–1006. https://doi.org/10.1007/s10764-023-00366-7

Pohlner, Y. (2023). *Personality and wellbeing in gorillas: Captive-wild comparisons and implications for zoos*. [Doctoral dissertation: University of Stirling] STORRE: Stirling Online Research Repository http://dspace.stir.ac.uk/handle/1893/36050

Pritchard, A. J. (2025, December 3). Data for: The state and expectations of interrater reliability in nonhuman personality research. https://doi.org/10.17605/OSF.IO/J8RFX

Pritchard, A. J., Bliss-Moreau, E., Balasubramaniam, K. N., Capitanio, J. P., Marty, P. R., Kaburu, S. S., Arlet, M. E., Beisner, B. A., & McCowan, B. (2024). Personality trait structures across three species of Macaca, using survey ratings of responses to conspecifics and humans. *PloS One*, *19*(9), e0309946. https://doi.org/10.1371/journal.pone.0309946

Pritchard, A. J., & Palombit, R. A. (2022). Survey-rated personality traits and experimentally measured coping style and stress reactivity, in wild baboons. *American Journal of Primatology*, *84*(11), e23429. https://doi.org/10.1002/ajp.23429

Pritchard, A. J., Sheeran, L. K., Gabriel, K. I., Li, J.-H., & Wagner, R. S. (2014). Behaviors that predict personality components in adult free-ranging Tibetan macaques *Macaca thibetana*. *Current Zoology*, *60*(3), 362–372. https://doi.org/10.1093/czoolo/60.3.362

Rawlings, B., Flynn, E., Freeman, H., Reamer, L., Schapiro, S. J., Lambeth, S., & Kendal, R. L. (2020). Sex differences in longitudinal personality stability in chimpanzees. *Evolutionary Human Sciences*, *2*, e46. https://doi.org/10.1017/ehs.2020.45

Rebout, N., Baxter, A., Bales, K. L., & Zablocki-Thomas, P. (2025). Exploring personality structure and similarity in coppery titi monkeys (*Plecturocebus cupreus*). *Journal of Comparative Psychology*, *139*, 301-314. https://doi.org/10.1037/com0000407

Robinson, L. M., Coleman, K., Capitanio, J. P., Gottlieb, D. H., Handel, I. G., Adams, M. J., Leach, M. C., Waran, N. K., & Weiss, A. (2018). Rhesus macaque personality, dominance, behavior, and health. *American Journal of Primatology*, *80*(2), e22739. https://doi.org/10.1002/ajp.22739

Robinson, L. M., Waran, N. K., Handel, I., & Leach, M. C. (2021). Happiness, welfare, and personality in rhesus macaques (*Macaca mulatta*). *Applied Animal Behaviour Science*, *236*, 105268. https://doi.org/10.1016/j.applanim.2021.105268

Schaefer, S., & Steklis, H. (2014). Personality and subjective well-being in captive male western lowland gorillas living in bachelor groups. *American Journal of Primatology*, *76*. https://doi.org/10.1002/ajp.22275

Tkaczynski, P. J., Ross, C., MacLarnon, A., Mouna, M., Majolo, B., & Lehmann, J. (2019). Measuring personality in the field: An in situ comparison of personality quantification methods in wild Barbary macaques (*Macaca sylvanus*). *Journal of Comparative Psychology*, *133*(3), 313–325. https://doi.org/10.1037/com0000163

Úbeda, Y., & Llorente, M. (2015). Personality in Sanctuary-Housed Chimpanzees: A Comparative Approach of Psychobiological and Penta-Factorial Human Models. *Evolutionary Psychology*, *13*(1), 182–196. https://doi.org/10.1177/147470491501300111

Weiss, A. (2025). Personality structure in wild Eastern chimpanzees (*Pan troglodytes schweinfurthii*). *Journal of Comparative Psychology*. https://doi.org/10.1037/com0000414

Weiss, A., Inoue-Murayama, M., Hong, K.-W., Inoue, E., Udono, T., Ochiai, T., Matsuzawa, T., Hirata, S., & King, J. E. (2009). Assessing chimpanzee personality and subjective well-being in Japan. *American Journal of Primatology*, *71*(4), 283–292. https://doi.org/10.1002/ajp.20649

Weiss, A., King, J. E., & Hopkins, W. D. (2007). A cross-setting study of chimpanzee (*Pan troglodytes*) personality structure and development: zoological parks and Yerkes National Primate Research Center. *American Journal of Primatology, 69*(11), 1264-1277. https://doi.org/10.1002/ajp.20428

Weiss, A., King, J. E., & Perkins, L. (2006). Personality and subjective well-being in orangutans (*Pongo pygmaeus* and *Pongo abelii*). *Journal of Personality and Social Psychology*, *90*(3), 501–511. https://doi.org/10.1037/0022-3514.90.3.501

Weiss, A., Masilkova, M., Gartner, M. C., Morton, F. B., Rowden, L. J., Cheyne, S. M., Cunningham, C., & Inoue-Murayama, M. (in review). Personality, happiness, and health in gibbons.

Weiss, A., Staes, N., Pereboom, J. J., Inoue-Murayama, M., Stevens, J. M., & Eens, M. (2015). Personality in bonobos. *Psychological Science*, *26*(9), 1430–1439. https://doi.org/10.1177/0956797615589933

Weiss, A., Yokoyama, C., Hayashi, T., & Inoue-Murayama, M. (2021). Personality, subjective well-being, and the serotonin 1a receptor gene in common marmosets (*Callithrix jacchus*). *Plos One*, *16*(8), e0238663. https://doi.org/10.1371/journal.pone.0238663

Wilson, V. A., Inoue-Murayama, M., & Weiss, A. (2018). A comparison of personality in the common and Bolivian squirrel monkey (*Saimiri sciureus* and *Saimiri boliviensis*). *Journal of Comparative Psychology*, *132*(1), 24-39. https://doi.org/10.1037/com0000093