**Figure ESM.1**

*Inter-rater agreement between the coders CG and CH (Species: mountain gorillas) when measuring the variables ‘MAU duration’ and ‘Gesture action duration’ (see worked example 1); the ‘Outcome time’, MAU.Outcome latency and GA.Outcome latency (see worked example 2)*

|  |
| --- |
| Graphical user interface  Description automatically generated |
| *Note.* Interobserver agreement in gesture time measurements A. Minimum action unit (MAU) duration, B. Gesture action (GA) duration C. Outcome time D. MAU.Outcome latency and E. GA.Outcome latency (n = 40 samples from the mountain gorilla dataset used in worked example 2 (n = 251, see manuscript for more details on the data)). The time measurements MAU duration and GA duration were analysed in worked example 1 (comparison of MAU and Gesture action duration across the three gesture types ‘Reach’, ‘Beckon’ and ‘Raise’). The ‘Outcome time’, MAU.Outcome and GA.Outcome latency variables were used in worked example 2 (inter-species comparison of response latencies using the gesture end points MAU and GA). ICC2 estimates and their 95% CIs were calculated using the *psych* package (v2.2.9 Revelle, 2022) based on an absolute-agreement and 2-way models and revealed high agreement (MAU.duration ICC = .97, 95% CI from .95 to .99, F(39, 39) = 73, p < 0.01; Gesture action duration ICC = 0.99, 95% CI from .99 to .99, F(39, 39) = 7796, p < 0.01). There seems to be no consistent bias in coding. |

**Figure ESM.2**

*Inter-rater agreement between the coders AS and CG (Species: EastAfrican chimpanzees) when measuring the variables ‘MAU end time’, ‘Gesture action end time’, Outcome time’, ‘MAU.Outcome latency’ and ‘GA.Outcome latency’ (see worked example 2)*

|  |
| --- |
| Chart, scatter chart  Description automatically generated |
| *Note.* Interobserver agreement in gesture time measurements A. Minimum action unit (MAU) end time, B. Gesture action (GA) end time, C. MAU.Outcome latency, D. GA.Outcome latency, E. Outcome time used in the inter-species comparison of response latencies in worked example 2. ICC2 estimates and their 95% CIs were calculated using the *psych* package (v2.2.9 Revelle, 2022) based on an absolute-agreement and 2-way models and suggest fair agreement (MAU.Outcome latency ICC value = 0.92 (95% IC .79 to .97, F(24, 24) = 29, p < 0.01); GA.Outcome latency ICC = 0.95 (95% CI .88 to .98, F(24, 24) = 50, p < 0.01) between the coders AS and CG. However, graphs C, D show that the deviation from the mean of the respective measurements (MAU.Outcome latency; GA.Outcome latency) is consistently slightly more negative for coder CG as compared to the coder AS. This leads to latencies for CG that are consistently slightly larger those for AS and this distinction seems to be driven by the Outcome time variable (see graph E, in the other timings used to calculate the latencies (A) MAU end time and (B) GA end time there is no obvious pattern observed). |

**Figure ESM.3**

*Inter-rater agreement between the coders GB and CG (Species: EastAfrican chimpanzees) when measuring the variables ‘MAU end time’, ‘Gesture action end time’, Outcome time’, ‘MAU.Outcome latency’ and ‘GA.Outcome latency’ (see worked example 2)*

|  |
| --- |
| Chart, scatter chart  Description automatically generated |
| *Note.* Interobserver agreement in gesture time measurements A. Minimum action unit (MAU) end time, B. Gesture action (GA) end time, C. MAU.Outcome latency, D. GA.Outcome latency, E. Outcome time used in the inter-species comparison of response latencies in worked example 2. ICC2 estimates and their 95% CIs were calculated using the *psych* package (v2.2.9 Revelle, 2022) based on an absolute-agreement and 2-way models and suggest fair agreement (MAU.Outcome latency ICC = .95, 95% CI from .85 to .98, F(19, 19) = 45, p < 0.01; GA.Outcome latency ICC = .98, 95% CI from .95 to .99, F(19, 19) = 88, p < 0.01) between the coders GB and CG. There seems to be no consistent bias in coding.  **Figure ESM.4**  *Replication of worked example 2.2 “Mountain gorilla and chimpanzee recipient latencies to (behaviourally) respond to a signaller’s gestures (successful communications, all goals except ‘play’)” accounting for coding deviation between AS and CG*   |  | | --- | | Chart, box and whisker chart  Description automatically generated | | *Note.* To see whether the potential coding bias between coders AS and CG contributes to the longer recipient response latencies observed in the results of worked example 2.2 we checked whether adding the mean coding difference between AS and CG (MAU.Outcome = 0.48, GA.Outcome = 0.38) from the IRR comparison to the original AS values changed the outcome (see graph B, D). Comparing the graphs (A and B; C and D), it seems unlikely that the species difference observed is due to coder bias alone. | |