

**Supplementary table 1.** Test category records for temperament collected by the seven evaluators.

Evaluator	Test category				
	Gentleness	Patience	Aggressiveness	Sensitivity	Friendliness
A	123	123	123	123	83
B	35	35	35	35	35
C	205	205	205	205	164
D	35	35	35	35	35
E	109	109	109	109	108
F	131	131	131	131	90
G	21	21	21	21	21

**Supplementary table 2.** Significance of the fixed effects for the five test categories.

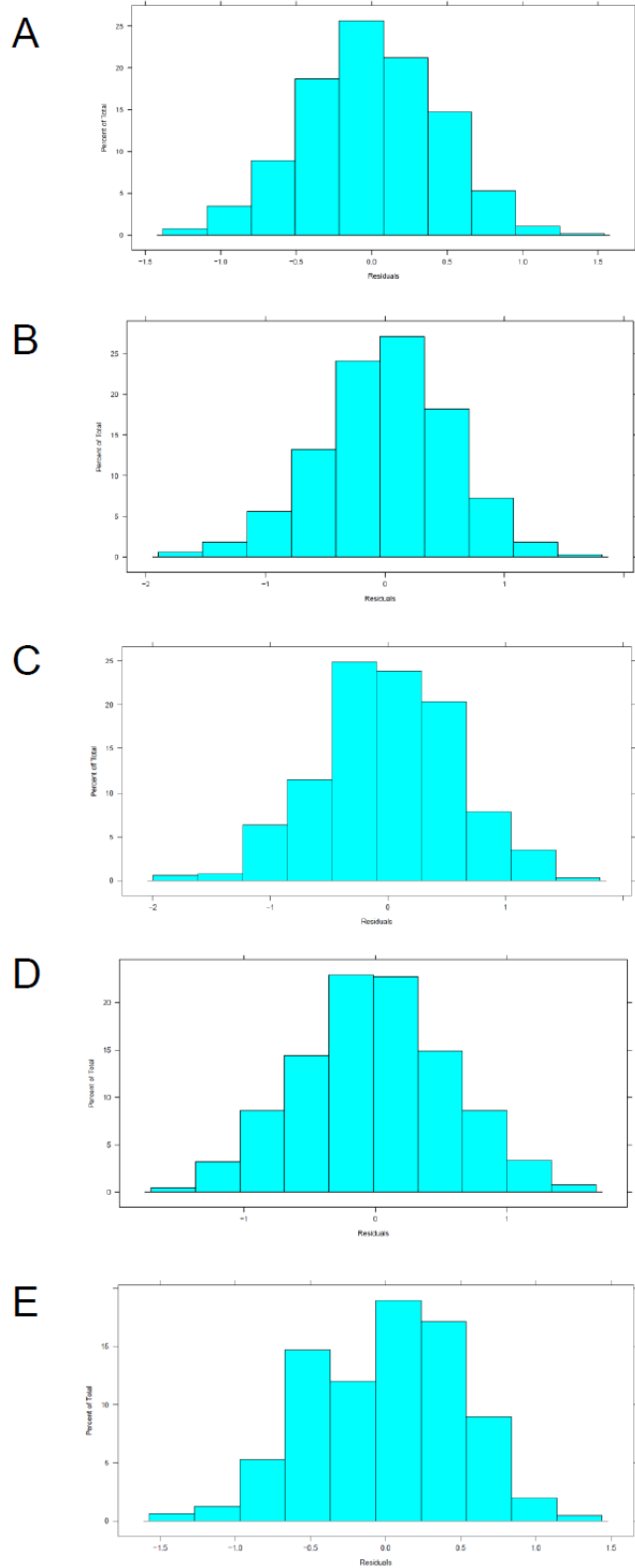
Term	Test category				
	Gentleness	Patience	Aggressiveness	Sensitivity	Friendliness
sex	N.S.	N.S.	N.S.	N.S.	N.S.
year	***	****	***	***	***
man	****	****	****	****	****
year × man	****	****	****	****	****
sex × year	N.S.	N.S.	N.S.	N.S.	N.S.
sex × man	N.S.	N.S.	N.S.	*	N.S.
sex × year × man	N.S.	N.S.	N.S.	N.S.	N.S.

N.S. no significance, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$

**Supplementary table 3.** Predicted value ( $\pm$  SE) of the year  $\times$  evaluator interaction

Year	Evaluator	Predicted value $\pm$ SE				
		Gentleness	Patience	Aggressiveness	Sensitivity	Friendliness
2010	A	3.83 $\pm$ 0.22	3.37 $\pm$ 0.26	3.26 $\pm$ 0.25	3.16 $\pm$ 0.28	- <sup>2</sup>
	B	N.A. <sup>1</sup>	N.A.	N.A.	N.A.	-
	C	3.95 $\pm$ 0.22	3.3 $\pm$ 0.26	3.73 $\pm$ 0.25	2.89 $\pm$ 0.28	-
	D	N.A.	N.A.	N.A.	N.A.	-
	E	N.A.	N.A.	N.A.	N.A.	-
	F	4.03 $\pm$ 0.22	4.22 $\pm$ 0.26	4.68 $\pm$ 0.25	3.86 $\pm$ 0.28	-
	G	N.A.	N.A.	N.A.	N.A.	-
2011	A	N.A.	N.A.	N.A.	N.A.	N.A.
	B	4.23 $\pm$ 0.23	3.5 $\pm$ 0.27	4.27 $\pm$ 0.26	30.9 $\pm$ 0.30	4.01 $\pm$ 0.20
	C	3.57 $\pm$ 0.23	3.5 $\pm$ 0.27	3.58 $\pm$ 0.26	2.89 $\pm$ 0.30	3.55 $\pm$ 0.20
	D	4.06 $\pm$ 0.23	3.58 $\pm$ 0.27	4.04 $\pm$ 0.26	3.61 $\pm$ 0.30	4.03 $\pm$ 0.20
	E	N.A.	N.A.	N.A.	N.A.	N.A.
	F	N.A.	N.A.	N.A.	N.A.	N.A.
	G	N.A.	N.A.	N.A.	N.A.	N.A.
2012	A	3.04 $\pm$ 0.22	2.61 $\pm$ 0.27	3.11 $\pm$ 0.26	2.32 $\pm$ 0.29	2.99 $\pm$ 0.19
	B	N.A.	N.A.	N.A.	N.A.	N.A.
	C	3.74 $\pm$ 0.22	3.49 $\pm$ 0.27	3.68 $\pm$ 0.26	3.01 $\pm$ 0.29	3.66 $\pm$ 0.19
	D	N.A.	N.A.	N.A.	N.A.	N.A.
	E	4.07 $\pm$ 0.22	3.13 $\pm$ 0.27	3.7 $\pm$ 0.26	3.29 $\pm$ 0.29	3.76 $\pm$ 0.19
	F	N.A.	N.A.	N.A.	N.A.	N.A.
	G	N.A.	N.A.	N.A.	N.A.	N.A.
2013	A	3.36 $\pm$ 0.22	2.9 $\pm$ 0.26	3.5 $\pm$ 0.25	2.93 $\pm$ 0.29	3.43 $\pm$ 0.18
	B	N.A.	N.A.	N.A.	N.A.	N.A.
	C	3.76 $\pm$ 0.22	3.25 $\pm$ 0.26	3.54 $\pm$ 0.25	2.92 $\pm$ 0.29	3.48 $\pm$ 0.18
	D	N.A.	N.A.	N.A.	N.A.	N.A.
	E	3.63 $\pm$ 0.22	2.63 $\pm$ 0.26	2.85 $\pm$ 0.25	2.32 $\pm$ 0.29	3.17 $\pm$ 0.18
	F	3.67 $\pm$ 0.22	3.05 $\pm$ 0.26	3.32 $\pm$ 0.25	2.81 $\pm$ 0.29	3.32 $\pm$ 0.18
	G	N.A.	N.A.	N.A.	N.A.	N.A.
2014	A	N.A.	N.A.	N.A.	N.A.	N.A.
	B	N.A.	N.A.	N.A.	N.A.	N.A.
	C	3.89 $\pm$ 0.25	3.73 $\pm$ 0.29	4.04 $\pm$ 0.29	3.27 $\pm$ 0.32	3.66 $\pm$ 0.24
	D	N.A.	N.A.	N.A.	N.A.	N.A.
	E	3.73 $\pm$ 0.25	2.89 $\pm$ 0.29	3.36 $\pm$ 0.29	2.99 $\pm$ 0.32	3.58 $\pm$ 0.24
	F	3.65 $\pm$ 0.25	3.73 $\pm$ 0.29	3.64 $\pm$ 0.29	3.11 $\pm$ 0.32	3.66 $\pm$ 0.24
	G	N.A.	N.A.	N.A.	N.A.	N.A.
2015	A	N.A.	N.A.	N.A.	N.A.	N.A.
	B	N.A.	N.A.	N.A.	N.A.	N.A.
	C	4.64 $\pm$ 0.28	4.39 $\pm$ 0.32	4.51 $\pm$ 0.31	4.01 $\pm$ 0.34	4.47 $\pm$ 0.25
	D	N.A.	N.A.	N.A.	N.A.	N.A.
	E	N.A.	N.A.	N.A.	N.A.	N.A.
	F	4.45 $\pm$ 0.28	4.24 $\pm$ 0.32	3.7 $\pm$ 0.31	3.49 $\pm$ 0.34	4.47 $\pm$ 0.25
	G	4.21 $\pm$ 0.28	4.05 $\pm$ 0.32	4.22 $\pm$ 0.31	3.58 $\pm$ 0.34	4.04 $\pm$ 0.25

<sup>1</sup>The corresponding evaluator did not participate the test.<sup>2</sup>The friendliness data were omitted in 2010 due to the incorrectness.



**Supplementary figure 1.** The residual distributions of the five temperamental traits. (A) Gentleness (B) Patience (C) Aggressiveness (D) Sensitivity (E) Friendliness. In each panel, the x-axis represents residuals and the y-axis represents percent of total.