Group and kin recognition via olfactory cues in chimpanzees (*Pan troglodytes*)

Stefanie Henkel and Joanna M. Setchell

Electronic Supplementary Material 2

Details of Methods and Results

Material and Methods

Behavioural Bioassay

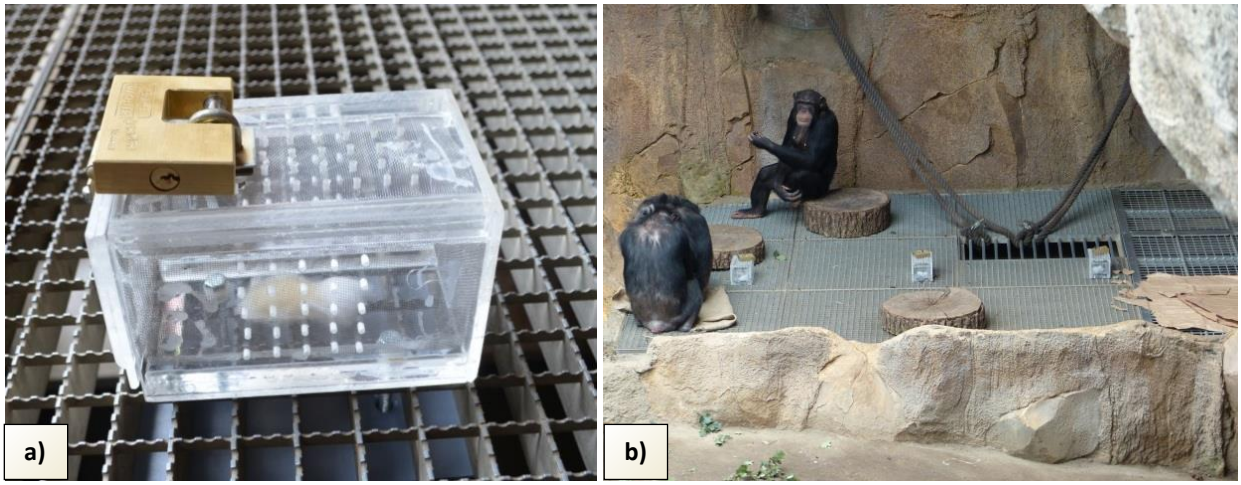

Fig. S1 a) Plexiglass box used to present urine samples. b) Experimental setup in the inside enclosure of chimpanzee group B at Leipzig Zoo

Table S1 Overview of individuals used as test subjects and/or odour donors

Individual	Test subject	Odour donor	Sex	Age [y]	Group
Alex	yes	yes	male	15	B
Alexandra	yes	yes	female	17	B
Annett	yes	no	female	16	B
Bangolo	yes	no	male	7	A
Corrie	yes	no	female	39	A
Daza	yes	no	female	30	B
Dorien	yes	yes	female	35	A
Fraukje	yes	yes	female	40	A
Frederike	yes	yes	female	42	B
Jahaga	yes	yes	female	23	B
Jeudi	yes	yes	female	50	B
Kara	yes	yes	female	11	A
Kisha	yes	no	female	12	A
Kofi	yes	yes	male	11	A
Lobo	yes	yes	male	12	A
Lome	yes	yes	male	15	A
Riet	yes	yes	female	38	A
Robert	no	yes	male	40	A
Sandra	yes	yes	female	23	A
Swela	yes	no	female	21	A
Tai	yes	no	female	14	A
Ulla	yes	yes	female	39	A

Table S2 Definitions of response behaviours (modified from [1])**.** Only touching and manipulating are mutually exclusive

Behaviour	Definition
Sniffing	Subject moves its nose towards the box to ≤ 3 cm
Nose within 20 cm	Subject directs its nose towards the box at an angle of ≤ 45° and within ≤ 20 cm while the animal pays active attention towards the box
Licking/Biting	Subject touches the box with its tongue or bites the box. Whereas sniffing and nose within 20 cm relate to perception of volatile chemosignals, licking mainly relates to perception of non-volatile signals [2,3]
Presence within 50 cm	≥ 50% of the subject's body is within 50 cm of the box. If the subject sits exactly in between two boxes, the head must be at ≤ 90° to a box to be within 50 cm. Includes when the subject interacted with the box and then engages in other behaviour while within 50 cm of the box. If the subject was not interested in the box before, other behaviours (like resting, eating) within 50 cm of the box are not included. Does not include when a subject walks by a box or lies next to a box without paying any attention to it.
Touching	Subject uses its hand or foot to touch the box (excluding the padlock)
Manipulating	Subject uses objects or sticks to manipulate the box (excluding the padlock)

Inter-observer reliability

Table S3 Results for the inter-observer reliability using Spearman's rank correlations to compare the total durations per behaviour, individual, box location and session

Response variable and box location	N	r_s	p
Licking left	12	0.391	0.209
Licking middle	15	0.815	<0.001
Licking right	16	0.566	0.022
Manipulating left	5	1.000	0.017
Manipulating middle	6	0.829	0.058
Manipulating right	7	0.883	0.015
Nose within 20 cm left	26	0.878	<0.001
Nose within 20 cm middle	21	0.885	<0.001
Nose within 20 cm right	31	0.803	<0.001
Present within 50 cm left	30	0.573	0.001
Present within 50 cm middle	23	0.961	<0.001
Present within 50 cm right	36	0.743	<0.001
Sniffing left	24	0.478	0.018
Sniffing middle	21	0.696	<0.001
Sniffing right	28	0.661	<0.001
Touching left	18	0.364	0.138
Touching middle	12	0.909	<0.001
Touching right	17	0.766	<0.001

Statistical analysis

We conducted all analyses using Linear Mixed Models with Gaussian error structure (LMMs, [4]) with significance set at $p<0.05$ and trends set at $0.05\leq p<0.1$. We fitted all models in R (version 3.4.0, [5]) using function lmer of R-package lme4 (version 1.1-13, [6]).

Influence of odour intensity

First, we tested whether intensity scores varied with the time odour samples were exposed to ambient air and whether in- and outgroup odours differed in intensity scores. We fitted a LMM on a larger dataset comprising more raters and odour samples than used for the ingroup-outgroup bioassay. We included odour condition (ingroup/outgroup) and ambient air time as fixed effects and rater ID, odour ID and session ID as random effects. To control for a potential effect of storage time, we included the time the samples were frozen as a further fixed effect. We calculated ambient air time as rating time post-session minus rating time pre-session and added 30 min because the samples were exposed to ambient air 30 min prior to rating time pre-session. Samples were always completely thawed when we rated them before the test session. To keep type I error rates at the nominal level of 5%, we included random slopes of time frozen within rater ID, odour ID and session ID as well as odour condition and ambient air time within session ID, but not the correlation parameters between random intercepts and random slopes terms [7,8].

Ambient air time and time frozen were both approximately symmetrically distributed and z-transformed (to a mean of zero and a standard deviation of one). We checked whether the assumptions of normally distributed and homogeneous residuals were fulfilled by visually inspecting a qqplot of the residuals and the residuals plotted against fitted values. Both indicated no obvious deviations from these assumptions. We checked for model stability by excluding levels of rater ID, odour ID and session ID one at a time from the data and comparing the model estimates derived for these subsets of the data with those derived for the full data set. This indicated no influential levels of any random effects. We derived Variance Inflation Factors (VIF, [9]) using the function vif of the R-package car (version 2.1-5, [10]) applied to a standard linear model excluding the random effects. These did not indicate a collinearity problem ($VIF_{max}=1.047$, $VIF_{mean}=1.033$).

We established the significance of the full model compared to the null model (comprising only time frozen and the random effects) using a likelihood ratio test (LRT, R function anova with argument test set to "Chisq", [11,12]). To allow a likelihood ratio test we fitted the models using Maximum Likelihood (rather than Restricted Maximum Likelihood, [13]). We based P-values for the individual effects on LRTs comparing the full with respective reduced models ([8], R function drop1). The sample size for this model was a total of 128 observations made by 9 raters on 20 odours in 30 sessions.

Overall, the full model was highly significant when compared to the null model (LRT: $\chi2=31.357$, df=2, P=<0.001). Intensity scores significantly decreased with ambient air time (Estimate=-1.574, SE=0.190, $\chi2=31.029$, df=1, $p<0.001$, fig. S2) and time frozen (Estimate=-0.893, SE=0.227, $\chi2=5.871$, df=1, $p=0.015$, fig. S3), but they did not differ between ingroup and outgroup odours (Estimate=0.092, SE=0.356, $\chi2=0.057$, df=1, $p=0.812$, table S4). We used intensity scores pre-session as a control predictor in subsequent analyses.

Table S4 Results of Linear Mixed Model with Gaussian error structure testing the effect of exposure to ambient air and odour stimuli on intensity scores. Degrees of freedom were 1 throughout. Significant effects are marked in bold. Ref = reference level

Predictor variable	Estimate	SE	CL_{lower}	CL_{upper}	χ^2	p
Intercept	4.554	0.465	3.646	5.487	(1)	(1)
Odour condition (ref = outgroup)	0.092	0.356	-0.582	0.832	0.057	0.812
Ambient air time	-1.574	0.190	-1.93	-1.217	31.029	**<0.001**
Time frozen	-0.893	0.227	-1.398	-0.365	5.871	**0.015**

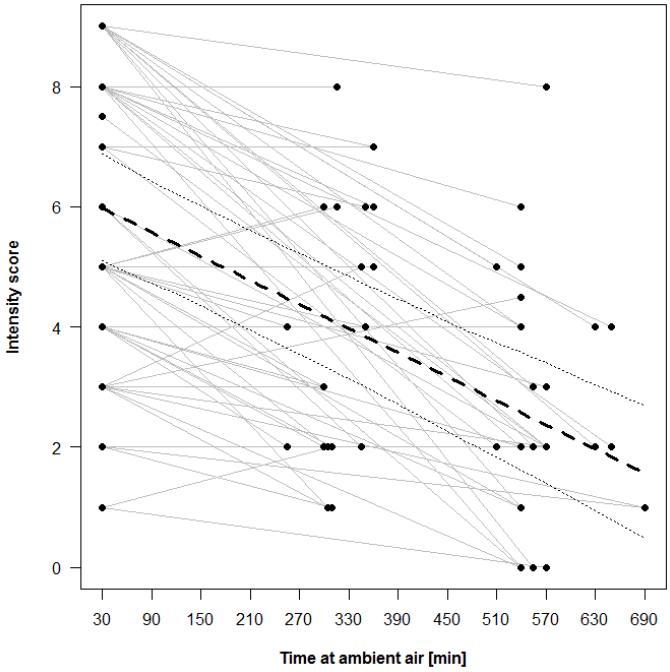

Fig. S2 Effect of the time samples were exposed to ambient air on intensity scores. The dashed line depicts the model (fitted based on odour condition manually dummy coded and then centered to a mean of zero), thin dotted lines the 95% confidence intervals of the model, and grey lines connect individual samples

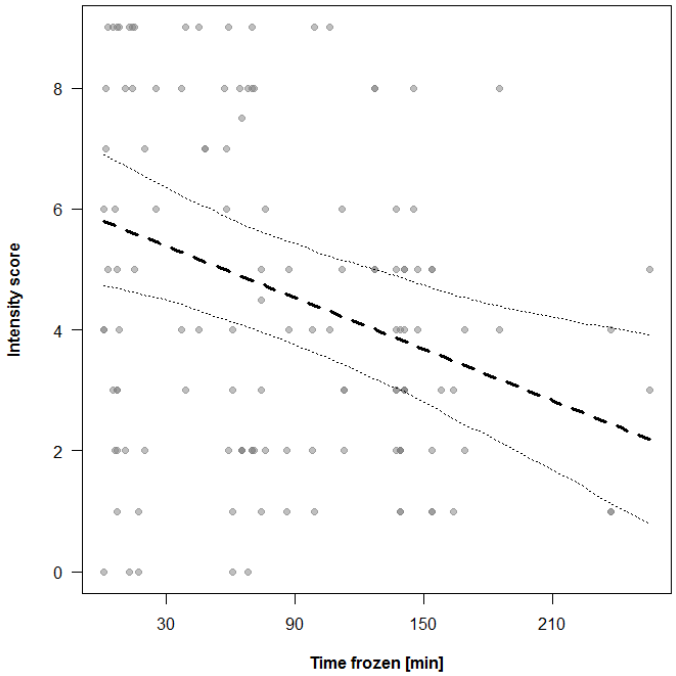

Fig. S3 Effect of the time samples were frozen on intensity scores. The dashed line depicts the model (fitted based on odour condition manually dummy coded and then centered to a mean of zero), the thin dotted lines the 95% confidence intervals of the model

Ingroup vs. outgroup and olfactory sensitivity

To test whether chimpanzees show an olfactory sensitivity to urine odours compared to the control and a differential behavioural response to ingroup and outgroup odours, we fitted two separate LMMs with Gaussian error structure and identity link for each of the response variables. We calculated the total duration of response behaviours per subject, odour stimulus and session. To standardise time between groups and to ensure odour intensity was high, we only used the first 4 hours per session. We excluded cases where the subject was also the odour donor to avoid a confounding influence of olfactory self-recognition.

The first set of models tested for a difference in response behaviours between control and odour stimuli in general. For these models, we included odour stimulus (odour/control), sex, group and age (in days) of the subject, session number and box location (left/middle/right) as fixed effects and subject ID and session ID as random effects.

The second set of models tested for a difference in response behaviours to ingroup and outgroup odours. For these models, we included odour stimulus (ingroup/outgroup), sex, group and age of the subject, sex and group of the odour donor, session number, box location (left/middle/right) and intensity score as fixed effects and subject ID, odour ID and session ID as random effects. To account for a potential differential response of males and females towards male and female odours, we also included the three-way interaction between odour stimulus, sex of the subject and sex of the odour

donor. Since there was only one individual in 2 combinations of factor levels for licking and manipulating, we did not include the 3-way interaction for these response variables but fitted only the two-way interaction between odour and subject sex. Odour stimulus was the test predictor and all other variables served as control predictors.

To keep type I error rates at a nominal level of 5%, we included a maximal random slopes structure but not the correlation parameters between random intercepts and random slopes terms ([8,14]; see table S5 and S6S4 for detailed model parameters and sample sizes).

Before running the model, we z-transformed the predictor variables session number, subject age (for all models) and intensity score (for the ingroup vs. outgroup models) to a mean of zero and a standard deviation of one to obtain easily interpretable estimates [15]. None of the response variables showed any obvious deviations from the assumptions of normally distributed and homogeneous residuals, based on visual inspection of a qqplot and plotting residuals against fitted values. We checked for model stability by excluding levels of random effects one at a time from the data and comparing the model estimates derived for these subsets of the data with those derived for the full data set. These did not indicate any influential cases.

We derived Variance Inflation Factors [VIF, 9] using function vif in R-package car [10] applied to a standard linear model excluding the random effects (for all models) and the interaction (for the ingroup vs. outgroup models). These did not indicate that collinearity was a problem (for max and mean VIFs see table S5 and S6).

We established the significance of the full model compared to the null model (comprising the control predictors, random effects and random slopes) using a likelihood ratio test (LRT, R function anova with argument test set to "Chisq"; [11,12]). We fitted models using Maximum Likelihood (rather than Restricted Maximum Likelihood [13]) to allow a likelihood ratio test. We based P-values for individual effects on likelihood ratio tests comparing the full model with respective reduced models [8]; R function drop1 with argument ‘test’ set to “Chisq”). We derived confidence intervals using the function bootMer of package lme4, using 1,000 parametric bootstraps and bootstrapping over the random effects, too.

Table S5 Model parameters for Linear Mixed Models with Gaussian error structure testing the effect of odour vs. control stimuli on response variables

Response variable	Fixed effects	Random effects	Random slopes	N_{RE}	N_{tot}	VIF_{max}	VIF_{mean}
Sniffing	Odour (TP) Subject sex (CP) Subject group (CP) Session (CP) Box location (CP) Subject age (CP)	Subject ID	Odour Session Box location Subject age	21	143	1.298	1.128
		Session ID	Odour Subject sex Box location Subject age	12			
Nose within 20cm	Odour (TP) Subject sex (CP) Subject group (CP) Session (CP) Box location (CP) Subject age (CP)	Subject ID	Odour Session Box location Subject age	21	157	1.287	1.117
		Session ID	Odour Subject sex Box location Subject age	12			
Licking	Odour (TP) Subject sex (CP) Subject group (CP) Session (CP) Box location (CP) Subject age (CP)	Subject ID	Odour Session Subject age	16	95	1.371	1.168
		Session ID	Odour Box location Subject age	12			
Presence within 50cm	Odour (TP) Subject sex (CP) Subject group (CP) Session (CP) Box location (CP) Subject age (CP)	Subject ID	Odour Session Box location Subject age	21	173	1.275	1.104
		Session ID	Odour Subject sex Box location Subject age	12			
Manipulating	Odour (TP) Subject sex (CP) Subject group (CP) Session (CP) Box location (CP) Subject age (CP)	Subject ID	Session Box location Subject age	18	116	1.276	1.104
		Session ID	Odour Subject sex Box location Subject age	12			

TP = test predictor; CP = control predictor; N_{RE} = number of levels per random effect; N_{tot} = total sample size; VIF_{max} = maximum variation inflation factor; VIF_{mean} = mean variation inflation factor

Table S6 Model parameters for Linear Mixed Models with Gaussian error structure testing the effect of ingroup vs. outgroup odours on response variables. Asterisks represent the interaction between fixed effects including all respective lower terms

Response variable	Fixed effects	Random effects	Random slopes	N_{RE}	N_{tot}	VIF_{max}	VIF_{mean}
Sniffing	Odour*Subject sex*Odour sex Subject group (CP) Odour group (CP) Session (CP) Box location (CP) Subject age (CP) Intensity score (CP)	Subject ID	Session Subject age Intensity score	21	96	1.789	1.351
		Odour ID	Subject sex Subject age	15			
		Session ID	Odour Odour group Box location Subject age	12			
Nose within 20cm	Odour*Subject sex*Odour sex Subject group (CP) Odour group (CP) Session (CP) Box location (CP) Subject age (CP) Intensity score (CP)	Subject ID	Session Subject age Intensity score	21	104	1.732	1.326
		Odour ID	Subject sex Subject age	15			
		Session ID	Odour Odour group Box location Subject age	12			
Licking	Odour*Subject sex (TP*CP) Odour*Odour sex (TP*CP) Subject group (CP) Odour group (CP) Session (CP) Box location (CP) Subject age (CP) Intensity score (CP)	Subject ID		16	63	1.837	1.460
		Odour ID	Subject age	13			
		Session ID	Odour Odour group Box location Subject age	12			
Presence within 50 cm	Odour*Subject sex*Odour sex Subject group (CP) Odour group (CP) Session (CP) Box location (CP) Subject age (CP) Intensity score (CP)	Subject ID	Session Subject age Intensity score	21	114	1.608	1.294
		Odour ID	Subject sex Subject age	15			
		Session ID	Odour Subject sex Odour group Box location Subject age Odour:Subject	12			
Manipulating	Odour*Subject sex (TP*CP) Odour*Odour sex (TP*CP)	Subject ID	Subject age Intensity score	18	83	1.626	1.297

	Subject group (CP)	Odour ID	Subject age	14
	Odour group (CP)	Session ID	Odour	12
	Session (CP)		Odour group	
	Box location (CP)		Box location	
	Subject age (CP)		Subject age	
	Intensity score (CP)			

TP = test predictor; CP = control predictor; N_{RE} = number of levels per random effect; N_{tot} = total sample size; VIF_{max} = maximum variation inflation factor; VIF_{mean} = mean variation inflation factor

Relatedness

To test whether the degree of relatedness influences behavioural responses, we fitted a LMM with Gaussian error structure and identity link for each response variable for ingroup odours only (taking only familiar individuals into account) because there were no related subject-odour donor dyads for outgroup odours. Relatedness coefficients ranged 0-0.5 (table S7). We included the relatedness coefficients and all predictor variables that were significant for the ingroup-outgroup models in the models. We included subject ID, odour ID and session ID as random effects and used a maximal random slopes structure (for detailed model parameters and sample sizes see table S8). All other procedures were the same as for the previous models. We detected no violations of the assumptions of normally distributed and homogeneous residuals and there was no indication of influential levels for any random effect or collinearity issues (for max and mean VIFs see table S8).

Table S7 Relatedness of familiar subject and odour donor dyads (group members) used for Linear Mixed Models testing the effect of relatedness on response behaviours. r = Relatedness coefficients

Subject	Odour donor	r	Relatedness
Alex	Frederike	0	unrelated
Alexandra	Alex	0	unrelated
Alexandra	Frederike	0	unrelated
Alexandra	Jeudi	0	unrelated
Bangolo	Dorien	0.500	son
Bangolo	Kara	0.125	great-nephew
Bangolo	Lobo	0.250	nephew
Bangolo	Lome	0.250	nephew
Bangolo	Robert	0.250	grandson
Bangolo	Sandra	0.125	great-nephew
Corrie	Kara	0	unrelated
Daza	Alex	0	unrelated
Daza	Frederike	0	unrelated
Daza	Jeudi	0	unrelated

Fraukje	Dorien	0	unrelated
Fraukje	Robert	0	unrelated
Frederike	Alex	0	unrelated
Frederike	Jeudi	0	unrelated
Jahaga	Alex	0	unrelated
Jahaga	Frederike	0	unrelated
Kara	Dorien	0	unrelated
Kisha	Dorien	0	unrelated
Kisha	Kara	0	unrelated
Kisha	Lobo	0	unrelated
Kofi	Dorien	0	unrelated
Kofi	Lobo	0.250	half-sibling
Lobo	Dorien	0	unrelated
Lobo	Kara	0.250	half-sibling
Lobo	Robert	0.500	son
Lome	Kara	0.250	half-sibling
Riet	Dorien	0	unrelated
Riet	Kara	0	unrelated
Riet	Lome	0	unrelated
Riet	Sandra	0.500	mother
Sandra	Lobo	0.250	half-sibling
Sandra	Robert	0.500	daughter
Swela	Dorien	0	unrelated
Swela	Kara	0	unrelated
Swela	Sandra	0	unrelated
Tai	Dorien	0	unrelated
Tai	Kara	0.250	half-sibling
Tai	Robert	0.500	daughter
Tai	Sandra	0.500	full-sibling

Table S8 Model parameters for Linear Mixed Models with Gaussian error structure testing the effect of relatedness coefficients (r) on response variables for ingroup odours. Asterisks represent the interaction between fixed effects including all respective lower terms

Response variable	Fixed effects	Random effects	Random slopes	N_{RE}	N_{tot}	VIF_{max}	VIF_{mean}
Sniffing	r (TP)	Subject ID		17	46	1.287	1.287
	Subject group (CP)	Odour ID		9			
		Session ID		12			
Nose within 20cm	r (TP)	Subject ID		17	46	1.505	1.361
	Subject group (CP)	Odour ID	Subject age	9			
	Session (CP)	Session ID	Subject age	12			
	Subject age (CP)						
Licking	r (TP)	Subject ID		13	31	1.426	1.426
	Subject age (CP)	Odour ID	Subject age	11			
		Session ID	Subject age	8			
Presence within 50cm	r (TP)	Subject ID		17	53	1.292	1.292
	Subject group (CP)	Odour ID		12			
		Session ID		9			
Manipulating	r*Subject sex (TP*CP)	Subject ID		18	83	1.029	1.029
		Odour ID		14			
		Session ID		12			

TP = test predictor; CP = control predictor; N_{RE} = number of levels per random effect; N_{tot} = total sample size; VIF_{max} = maximum variation inflation factor; VIF_{mean} = mean variation inflation factor

Results

Table S9 Overview of results for full-null model comparisons using likelihood ratio tests. Significances and trends are marked in bold

Model	Response variable	χ^2	df	p
Control vs. odour	Sniffing	4.660	1	**0.031**
	Nose within 20 cm	1.913	1	0.167
	Licking	1.572	1	0.210
	Presence within 50 cm	3.961	1	**0.047**
	Manipulating	1.985	1	0.159
Ingroup vs. outgroup	Sniffing	9.399	4	**0.052**
	Nose within 20 cm	5.225	4	0.265
	Licking	0.655	3	0.884
	Presence within 50 cm	5.620	4	0.231
	Manipulating	5.156	3	0.161
Relatedness	Sniffing	5.152	1	**0.023**
	Nose within 20 cm	1.461	1	0.227
	Licking	0.762	1	0.383
	Presence within 50 cm	8.719	1	**0.003**
	Manipulating	1.124	2	0.570

Table S10 Results of Linear Mixed Models with Gaussian error structure testing the effect of odour vs. control stimuli on response variables. Degrees of freedom (df) = 1 except for box location, where df = 2. Significances and trends are marked in bold. Ref = reference level

Response variable	Predictor variable	Estimate	SE	CL_{lower}	CL_{upper}	χ^2	p
Sniffing	Intercept	0.076	0.289	-0.501	0.667	(1)	(1)
	Odour (ref = control)	0.482	0.213	0.048	0.924	4.660	**0.031**
	Subject sex (ref = female)	0.119	0.409	-0.710	0.986	0.084	0.772
	Subject group (ref = Chimp A)	0.941	0.287	0.290	1.449	5.488	**0.019**
	Session[(2)]	-0.553	0.177	-0.924	-0.164	6.919	**0.009**
	Box location (middle)[(3)]	0.198	0.227	-0.261	0.648	1.262	0.532
	Box location (right)[(3)]	0.298	0.290	-0.333	0.899		
	Subject age[(2)]	-0.367	0.312	-1.018	0.280	1.320	0.251
Nose within 20 cm	Intercept	1.074	0.386	0.295	1.865	(1)	(1)
	Odour (ref = control)	0.362	0.243	-0.158	0.851	1.913	0.167
	Subject sex (ref = female)	-0.195	0.554	-1.337	0.961	0.122	0.726
	Subject group (ref = Chimp A)	1.397	0.446	0.380	2.382	6.539	**0.011**
	Session[(2)]	-0.579	0.111	-0.797	-0.325	12.108	**0.001**
	Box location (middle)[(3)]	-0.012	0.266	-0.552	0.575	0.913	0.634
	Box location (right)[(3)]	0.304	0.365	-0.435	1.073		
	Subject age[(2)]	-0.539	0.26	-1.132	0.066	3.168	**0.075**
Licking	Intercept	0.468	0.501	-0.546	1.548	(1)	(1)
	Odour (ref = control)	0.382	0.293	-0.223	1.004	1.572	0.210
	Subject sex (ref = female)	-0.901	0.615	-2.123	0.272	1.931	0.165
	Subject group (ref = Chimp A)	0.953	0.611	-0.307	2.149	2.215	0.137
	Session[(2)]	-0.451	0.169	-0.787	-0.122	2.205	0.138
	Box location (middle)[(3)]	0.242	0.309	-0.399	0.857	4.491	0.106
	Box location (right)[(3)]	0.825	0.360	0.102	1.556		
	Subject age[(2)]	-0.785	0.299	-1.388	-0.179	2.448	0.118
Presence within 50 cm	Intercept	2.752	0.342	2.046	3.444	(1)	(1)
	Odour (ref = control)	0.507	0.234	0.009	1.004	3.961	**0.047**
	Subject sex (ref = female)	0.251	0.433	-0.666	1.174	0.327	0.567
	Subject group (ref = Chimp A)	0.838	0.350	0.097	1.570	4.636	**0.031**
	Session[(2)]	-0.396	0.157	-0.708	0.039	3.292	**0.070**
	Box location (middle)[(3)]	-0.244	0.287	-0.860	0.336	2.189	0.335
	Box location (right)[(3)]	0.356	0.360	-0.408	1.120		
	Subject age[(2)]	-0.451	0.205	-0.915	-0.026	4.251	**0.039**
Manipulating	Intercept	2.119	0.412	1.271	2.943	(1)	(1)
	Odour (ref = control)	0.393	0.274	-0.159	0.935	1.985	0.159
	Subject sex (ref = female)	0.725	0.530	-0.419	1.742	1.686	0.194
	Subject group (ref = Chimp A)	0.714	0.453	-0.220	1.725	2.351	0.125
	Session[(2)]	-0.329	0.133	-0.608	-0.015	4.251	**0.039**
	Box location (middle)[(3)]	0.016	0.423	-1.088	0.857	0.767	0.681
	Box location (right)[(3)]	0.316	0.368	-0.444	1.020		
	Subject age[(2)]	-0.477	0.267	-1.119	0.042	3.084	**0.079**

(1): not shown because lacks a reasonable interpretation
(2): z-transformed to mean = 0 and sd = 1; mean and sd of the original variables are presented in table S5
(3): box location was dummy coded with left being the reference level

Table S11 Results of Linear Mixed Models with Gaussian error structure testing the effect of ingroup vs. outgroup odours on response variables. Degrees of freedom (df) = 1 except for box location, where df = 2. Significances and trends are marked in bold. Ref = reference level. Colons represent the interaction between fixed effects

Response variable	Predictor variable	Estimate	SE	CL_{lower}	CL_{upper}	χ^2	p
Sniffing	Intercept	-0.147	0.386	-0.918	0.661	(1)	(1)
	Odour (ref = ingroup)	0.564	0.251	0.055	1.065	4.675	**0.031**
	Subject sex (ref = female)	0.586	0.369	-0.230	1.341	2.101	0.147
	Odour sex (ref = female)	-0.389	0.359	-1.139	0.340	1.092	0.296
	Subject group (ref = Chimp A)	1.622	0.279	0.983	2.192	11.752	**0.001**
	Odour group (ref = Chimp A)	0.032	0.244	-0.474	0.528	0.016	0.899
	Session(2)	-0.295	0.269	-0.836	0.286	1.080	0.299
	Box location (middle)(3)	0.569	0.311	-0.058	1.202	3.582	0.167
	Box location (right)(3)	0.666	0.367	-0.091	1.459		
	Subject age(2)	-0.400	0.254	-0.969	0.122	2.358	0.125
	Intensity score(2)	-0.210	0.198	-0.610	0.196	1.063	0.303
Nose within 20 cm	Intercept	1.708	0.522	0.665	2.810	(1)	(1)
	Odour (ref = ingroup)	-0.784	0.475	-1.806	0.130	(1)	(1)
	Subject sex (ref = female)	-0.639	0.753	-2.128	0.892	(1)	(1)
	Odour sex (ref = female)	-0.962	0.560	-2.160	0.142	(1)	(1)
	Subject group (ref = Chimp A)	1.680	0.434	0.794	2.620	9.913	**0.002**
	Odour group (ref = Chimp A)	0.752	0.316	0.137	1.355	5.197	**0.023**
	Session(2)	-0.624	0.201	-1.070	-0.223	8.566	**0.003**
	Box location (middle)(3)	-0.256	0.355	-0.966	0.453	2.493	0.288
	Box location (right)(3)	0.276	0.428	-0.634	1.139		
	Subject age(2)	-0.499	0.269	-1.023	0.044	2.911	**0.088**
	Intensity score(2)	0.033	0.227	-0.420	0.491	0.020	0.888
	Odour : Subject sex	-0.411	0.777	-2.026	1.205		
	Odour : Odour sex	0.917	0.730	-0.470	2.600		
	Subject sex : Odour sex	2.474	0.985	0.530	4.520		
	Odour : Subject sex : Odour sex	-0.820	1.246	-3.357	1.810	0.423	0.515
Licking	Intercept	1.496	0.683	0.109	2.861	(1)	(1)
	Odour (ref = ingroup)	0.151	0.543	-1.003	1.220	(1)	(1)
	Subject sex (ref = female)	-0.949	0.746	-2.436	0.519	(1)	(1)
	Odour sex (ref = female)	-0.717	0.696	-2.214	0.668	(1)	(1)
	Subject group (ref = Chimp A)	0.538	0.658	-0.752	2.009	0.615	0.433
	Odour group (ref = Chimp A)	0.364	0.526	-0.769	1.462	0.464	0.496

	Session[2]	-0.317	0.247	-0.823	0.196	1.583	0.208
	Box location (middle)[3]	0.066	0.425	-0.887	1.038	0.325	0.850
	Box location (right)[3]	0.348	0.614	-0.840	1.658		
	Subject age[2]	-0.679	0.298	-1.307	-0.095	4.414	**0.036**
	Intensity score[2]	0.105	0.242	-0.395	0.566	0.186	0.667
	Odour : Subject sex	-0.222	0.696	-1.611	1.221	0.101	0.751
	Odour : Odour sex	-0.560	0.851	-2.310	1.092	0.419	0.517
Presence within 50 cm	Intercept	3.206	0.443	2.312	4.275	[1]	[1]
	Odour (ref = ingroup)	-0.658	0.479	-1.650	0.326	[1]	[1]
	Subject sex (ref = female)	-0.631	0.726	-2.448	1.019	[1]	[1]
	Odour sex (ref = female)	-0.543	0.542	-1.663	0.612	[1]	[1]
	Subject group (ref = Chimp A)	1.214	0.380	0.212	2.045	5.278	**0.022**
	Odour group (ref = Chimp A)	0.895	0.293	0.085	1.500	4.714	**0.030**
	Session[2]	-0.387	0.242	-0.963	0.162	2.021	0.155
	Box location (middle)[3]	-0.368	0.350	-1.187	0.351	4.147	0.126
	Box location (right)[3]	0.351	0.433	-0.601	1.272		
	Subject age[2]	-0.406	0.234	-0.933	0.122	2.233	0.135
	Intensity score[2]	-0.141	0.207	-0.583	0.368	0.409	0.523
	Odour : Subject sex	0.709	0.911	-1.665	2.629		
	Odour : Odour sex	0.521	0.760	-1.194	2.035		
	Subject sex : Odour sex	3.182	1.018	1.011	5.991		
	Odour : Subject sex : Odour sex	-2.718	1.413	-6.192	0.487	2.975	**0.085**
Manipulating	Intercept	2.535	0.538	1.493	3.683	[1]	[1]
	Odour (ref = ingroup)	0.015	0.544	-1.004	1.118	[1]	[1]
	Subject sex (ref = female)	-0.280	0.722	-1.817	1.091	[1]	[1]
	Odour sex (ref = female)	-0.207	0.614	-1.431	1.063	[1]	[1]
	Subject group (ref = Chimp A)	0.766	0.538	-0.345	1.915	1.676	0.195
	Odour group (ref = Chimp A)	0.557	0.346	-0.163	1.262	2.476	0.116
	Session[2]	-0.265	0.249	-0.776	0.257	1.010	0.315
	Box location (middle)[3]	0.143	0.421	-0.694	0.968	0.565	0.754
	Box location (right)[3]	-0.241	0.524	-1.410	0.820		
	Subject age[2]	-0.414	0.297	-0.971	0.146	1.607	0.205
	Intensity score[2]	-0.212	0.246	-0.717	0.301	0.654	0.419
	Odour : Subject sex	1.605	0.735	0.079	3.074	3.981	**0.046**
	Odour : Odour sex	0.032	0.886	-1.767	1.645	0.001	0.972

For sniffing, we fitted a reduced model without interactions since the three-way interaction odour : subject sex : odour sex and the two-way interactions odour : subject sex, odour : odour sex and subject sex : odour sex (tested by fitting a reduced model without the three-way interaction) were not significant. Results for the interactions were: odour : subject sex : odour sex: Estimate=-0.640, SE=0.974, $\chi2$=-0.657, df=1, p=0.515; odour : subject sex: Estimate=-0.920, SE= 0.504, $\chi2$=3.194, df=1, p= 0.074; odour : odour sex: Estimate=0.580, SE= 0.572, $\chi2$=- 1.004, df=1, p= 0.316; subject sex : odour sex: Estimate=0.599, SE= 0.717, $\chi2$=0.624, df=1, p= 0.429

[1]: not shown because lacks a reasonable interpretation

[2]: z-transformed to mean=0 and sd=1; mean and sd of the original variables are presented in table S5

[3]: box location was dummy coded with left being the reference level

Table S12 Results of Linear Mixed Models with Gaussian error structure testing the effect of relatedness coefficients (r) on response variables for ingroup odours. Degrees of freedom = 1 throughout. Significant effects are marked in bold. Ref = reference level. Colons represent the interaction between fixed effects

Response variable	Predictor variable	Estimate	SE	CL_{lower}	CL_{upper}	χ^2	p
Sniffing	Intercept	0.424	0.488	0.297	1.749	(1)	(1)
	r[(2)]	0.544	0.226	0.053	0.986	5.152	**0.023**
	Subject group (ref = Chimp A)	1.286	0.799	-0.212	3.027	2.152	0.142
Nose within 20 cm	Intercept	1.115	0.365	0.359	1.847	(1)	(1)
	r[(2)]	1.839	1.506	-1.181	4.871	1.461	0.227
	Subject group (ref = Chimp A)	2.736	0.545	1.601	3.863	15.048	**<0.001**
	Session[(2)]	-0.847	0.189	-1.227	-0.449	14.210	**<0.001**
	Subject age[(2)]	-0.623	0.256	-1.146	-0.086	4.984	**0.026**
Licking	Intercept	1.487	0.470	0.583	2.390	(1)	(1)
	r[(2)]	0.316	0.351	-0.440	1.060	0.762	0.383
	Subject age[(2)]	-0.135	0.399	-0.963	0.696	0.109	0.741
Presence within 50 cm	Intercept	3.182	0.391	2.416	3.934	(1)	(1)
	r[(2)]	0.791	0.246	0.284	1.241	8.719	**0.003**
	Subject group (ref = Chimp A)	1.832	0.610	0.651	3.012	6.931	**0.008**
Manipulating	Intercept	2.982	0.205	2.577	3.370	(1)	(1)
	r[(2)]	-0.040	0.206	-0.471	0.361	(1)	(1)
	Subject sex (ref = female)	0.841	0.410	0.006	1.700	(1)	(1)
	r : Subject sex	0.429	0.414	-0.405	1.271	1.040	0.308

[(1)]: not shown because lacks a reasonable interpretation
[(2)]: z-transformed to mean=0 and sd=1; mean and sd of the original variables are presented in table S5

Table S13 Mean and SD of the original predictor variables that were z-transformed in the Linear Mixed Models

Model	Response variable	Predictor variable	Mean	SD
Control vs. odour	Sniffing	Session	2.755	1.692
		Subject age	21.746	12.404
	Nose within 20 cm	Session	2.822	1.681
		Subject age	21.800	12.240
	Licking	Session	2.642	1.663
		Subject age	20.679	12.445
	Presence within 50 cm	Session	2.890	1.717
		Subject age	21.713	12.228
	Manipulating	Session	2.784	1.724
		Subject age	19.967	11.416
Ingroup vs. outgroup	Sniffing	Session	2.667	1.620
		Subject age	21.960	12.044
		Intensity score	6.151	2.700
	Nose within 20 cm	Session	2.750	1.624
		Subject age	21.990	12.024
		Intensity score	6.313	2.692
	Licking	Session	2.683	1.664
		Subject age	20.393	11.811
		Intensity score	6.230	2.738
	Presence within 50 cm	Session	2.851	1.679
		Subject age	21.797	12.059
		Intensity score	6.298	2.692
	Manipulating	Session	2.795	1.695
		Subject age	20.299	11.693
		Intensity score	6.187	2.757
Relatedness	Sniffing	r	0.084	0.158
	Nose within 20 cm	r	0.084	0.158
		Session	2.826	1.582
		Subject age	21.833	11.468
	Licking	r	0.060	0.124
		Subject age	19.294	10.407
	Presence within 50 cm	r	0.099	0.170
	Manipulating	r	0.068	0.145

References

1. Henkel S, Lambides AR, Berger A, Thomsen R, Widdig A. 2015 Rhesus macaques (*Macaca mulatta*) recognize group membership via olfactory cues alone. *Behavioral Ecology and Sociobiology* **69**, 2019–2034. (doi:10.1007/s00265-015-2013-y)

2. Keverne EB. 1999 The vomeronasal organ. *Science* **286**, 716–720.

3. Drea CM, Boulet M, Delbarco-Trillo J, Greene LK, Sacha CR, Goodwin TE, Dubay GR. 2013 The "Secret" in Secretions: Methodological Considerations in Deciphering Primate Olfactory Communication: Methods in Deciphering Olfactory Communication. *American Journal of Primatology* **75**, 621–642. (doi:10.1002/ajp.22143)

4. Baayen H. 2008 *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. 1st edn. Cambridge: Cambridge University Press.

5. R Core Team. 2017 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

6. Bates D, Mächler M, Bolker B, Walker S. 2015 Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**. (doi:10.18637/jss.v067.i01)

7. Schielzeth H, Forstmeier W. 2009 Conclusions beyond support: overconfident estimates in mixed models. *Behavioral Ecology* **20**, 416–420. (doi:10.1093/beheco/arn145)

8. Barr DJ, Levy R, Scheepers C, Tily HJ. 2013 Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* **68**, 255–278. (doi:10.1016/j.jml.2012.11.001)

9. Field A. 2009 *Discovering Statistics Using SPSS*. 3 Pap/Cdr. Sage Publications Ltd.

10. Fox J, Weisberg S. 2011 *An {R} Companion to Applied Regression*. Second Edition. Thousand Oaks CA: Sage.

11. Dobson AJ. 2002 *An introduction to generalized linear models*. 2nd edn. Boca Raton: Chapman & Hall/CRC.

12. Forstmeier W, Schielzeth H. 2011 Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behav Ecol Sociobiol* **65**, 47–55. (doi:10.1007/s00265-010-1038-5)

13. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White J-SS. 2009 Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution* **24**, 127–135. (doi:10.1016/j.tree.2008.10.008)

14. Schielzeth H, Forstmeier W. 2009 Conclusions beyond support: overconfident estimates in mixed models. *Behavioral Ecology* **20**, 416–420. (doi:10.1093/beheco/arn145)

15. Schielzeth H. 2010 Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* **1**, 103–113. (doi:10.1111/j.2041-210X.2010.00012.x)