# Computerized assessment of dominance hierarchy in baboons (Papio papio)

Julie Gullstrand, Nicolas Claidière, Joël Fagot

1

**Computerized assessment of dominance hierarchy in baboons (*Papio papio*)**

Julie Gullstrand, Nicolas Claidière, and Joël Fagot

4

Laboratory of Cognitive Psychology, CNRS and Aix-Marseille University

6

13

## Corresponding author

Joël Fagot, Laboratoire de Psychologie Cognitive, CNRS UMR 7290 / Aix-Marseille Université, 3 Place Victor Hugo, Bât. 9, Case D, 13331 Marseille cedex 1. E-mail: joel.fagot@univ-amu.fr

18

19

20

21  **Abstract**

22      Dominance hierarchies are an important aspect of Primate social life, and there is an

23  increasing need to develop new systems to collect social information automatically. The main

24  goal of this research was to explore the possibility to infer the dominance hierarchy of a group

25  of Guinea baboons (*Papio papio*) from the analysis of their spontaneous interactions with freely

26  accessible automated learning devices for monkeys (ALDM, Fagot & Bonté, 2010).

27  Experiment 1 compared the dominance hierarchy obtained from conventional observations of

28  agonistic behaviours to the one inferred from the analysis of automatically recorded supplanting

29  behaviours within the ALDM workstations. The comparison, applied to three different datasets,

30  shows that the dominance hierarchies obtained with the two methods are highly congruent (all

31  rs >0.7). Experiment 2 investigated the experimental potential of inferring dominance hierarchy

32  from ALDM testing. ALDM data previously published in Goujon & Fagot (2013) were re-

33  analysed for that purpose. Results indicate that supplanting events within the workstations lead

34  to a transient improvement of cognitive performance for the baboon supplanting its partners

35  and that this improvement depends on the difference in rank between the two baboons. This

36  study therefore opens new perspectives for cognitive studies conducted in a social context.

37

38      *Keywords: monkey, supplanting behaviours, dominance hierarchy, automatized testing,*

39  *social cognition*

40

41

42          Agonistic behaviours are widely present in the animal kingdom. These behaviours

43   usually occur between group members in a context of competition for limited resources (e.g.,

44   King, 1973) and are relevant to understand a variety of phenomena such as adaptive fitness

45   (e.g., Neat, Taylor & Huntingford, 1998) or dominance style (e.g., Flack, De Waal & Waal,

46   2004). The structure of dominance hierarchies and their mode of computation has been amply

47   discussed in the literature (Bang, Deshpande, Sumana & Gadagkar 2010; Balasubramaniam et

48   al., 2013; Hemelrijk, Wantia & Gygax, 2005; De Vries, 1998; De Vries, Stevens, & Vervaecke,

49   2006). However, this domain of research has witnessed very little evolution regarding the

50   methodology used for data recording. Agonistic behaviours can differ depending on the

51   dominance style of the considered species (e.g., Flack et al., 2004) and their outcome and

52   frequency can both potentially be used to infer group dominance hierarchies. According to

53   David (1987), in linear dominance hierarchies the most dominant individuals emit more

54   agonistic behaviours than they receive, increasing their proportion of contests won. In such

55   situations, dominance hierarchies can therefore be inferred from the difference in frequencies

56   of agonistic behaviours emitted and received by each individual within the group (Dugatkin,

57   1997). In animals with a less linear social organization, the analysis of dominance hierarchies

58   is often more difficult to establish due to a smaller number of agonistic behaviours emitted and

59   received, especially in species with more egalitarian styles (Watts, 1994). Fischer et al. (2017)

60   for instance reported in the male-male tolerant Guinea baboons (*Papio papio*) that they were

61   not able to establish a significant linear rank hierarchy because of the low number of agonistic

62   interactions, and despite the fact that agonistic interactions were mostly unidirectional (see also

63   Patzelt, 2013).

64          Another difficulty encountered in previous studies is that agonistic behaviours require

65   serious efforts to be recorded. In most studies on dominance hierarchies, data are collected by

66 effortful and time-consuming observations, either directly (Barton, Byrne & Whiten,1996) or

67 based on video recordings (Foris, Zebunke, Langbein, & Melzer, 2019). With the development

68 of modern technologies, new methods have emerged for automatic identifications of animals

69 (e.g., GPS tracking in baboons, Strandburg-Peshkin et al., 2015; bar code tracking in birds,

70 Alarcón-Nieto et al., 2018; deep learning in birds, : Ferreira et al., 2019), but the use of

71 automated identification methods remains limited in social studies on dominance hierarchy. In

72 rare noticeable exceptions, authors inferred dominance hierarchies from the combination of

73 video tracking with (1) RFID data (in mice: Weissbrod et al., 2013), (2) identity data obtained

74 from automated reading of ear-tags (in cows: Foris, Thompson, Keyserlingk, Melzer & Weary,

75 2019), or (3) automated barcode or GPS reading (in pigeons: Nagy et al., 2013), but these

76 methods remain time consuming because of video processing. In another study with olive

77 baboons (*Papio anubis*) the author used supplanting behaviours to assess dominance hierarchy

78 to overcome the lack of the other agonistic behaviours (Packer, 1979).

79       The main goal of the current research was to explore a new method to infer dominance

80 hierarchies automatically. Fagot and Bonté (2010) have implemented novel test systems to

81 study cognition in semi-free baboons. In this study, a colony of Guinea baboons has free access

82 to a battery of ten ALDM (for Automated Learning Devices for Monkeys) test systems. The

83 ALDM test systems are automatized operant conditioning systems equipped with touchscreen

84 on which the cognitive tasks are presented, as well as a RFID tag reader for the identification

85 of the subject. With this technology, each test system reads the identifying RFID microchip

86 implanted in each arm of the baboon when it enters a workstation. This allows the self-paced

87 presentation of the cognitive tasks to the members of the colony, without any social isolation

88 or physical capture. Claidière, Gullstrand, Latouche and Fagot (2017) reported that the spatial

89 use of these ten ALDM test systems provides information on the social structure of the baboon

90 colony: the social network inferred from the co-presence of individuals within the spatially

91      adjacent workstation correlated strongly and reliably with the social network inferred from the

92      affiliative behaviours (e.g. grooming or greeting behaviours) observed inside the enclosure

93      where the group of baboons lived. In this study, we assessed if the use of ALDM systems can

94      also provide information on dominance hierarchy, therefore complementing our previous

95      studies on affiliative network (Claidière et al., 2017; Gelardi, Fagot, Barrat & Claidière, 2019).

96      Following Packer (1979), we considered in our research that supplanting behaviours are

97      interesting behavioural markers of the dominance hierarchy. A supplanting behaviour can be

98      defined as a sequence of behaviour in which one animal A approaches another individual B

99      who then leaves. Supplanting behaviours have the advantage over many other agonistic

100     behaviours to have a clear winner (individual A).  According to Johnson (1989), supplanting

101     behaviours lead to the establishment and maintenance of dominance hierarchies in olive

102     baboons (*Papio anubis*), and these behaviours are observed in this species regardless of food

103     quality and rarity, between individuals closer in ranks, avoiding greater risks in escalation of

104     more aggressive displays. To the best of our knowledge, our study is the first study in which

105     dominance hierarchy was inferred from supplanting behaviours in Guinea baboons.

106     Given the proposed relation between supplanting behaviours and dominance hierarchy,

107     our study investigated the dominance hierarchy in Guinea baboons considering supplanting

108     behaviours within the ALDM test systems. In Experiment 1, we first assessed dominance

109     hierarchy from conventional observations of agonistic behaviours expressed within the

110     enclosure where the baboons lived. Secondly, we extracted hierarchy from ALDM dataset,

111     considering supplanting behaviours within the testing workstations and compared the

112     hierarchies obtained with the two methods. Using three different datasets, we found that the

113     dominance hierarchy converged with the two methods. Thirdly, experiment 2 aimed at

114     documenting the experimental potentials of inferring dominance hierarchy from automatically

115     recorded ALDM supplanting behaviours. Supplanting behaviours were extracted from a

116    previous data set already published in Goujon and Fagot (2013) and used as a factor for

117    reanalysing the cognitive performance of the baboons in this study. We found that supplanting

118    events lead to a transient improvement of cognitive performance for the baboon supplanting its

119    partner, confirming the potential usefulness of inferring dominance hierarchies to study

120    cognitive processes.

121

## Experiment 1: Assessing dominance hierarchy from ALDM testing

123    Experiment 1 aimed at comparing dominance hierarchies inferred from the ALDM

124    supplanting behaviours with those obtained from more naturalistic observations of the social

125    behaviours spontaneously expressed in the baboon's enclosure.

**Methods**

**Participants and living conditions**

128    All participants belonged to the same group of Guinea baboons (*Papio papio*) from the

129    CNRS primate Centre, Rousset-sur-Arc, France. The group consisted of males and females

130    from different families living with their offspring in a large enclosure. The entire group was

131    housed inside a 25 x 30 $m^2$ outdoor enclosure connected to an indoor area and to two

132    experimental trailers providing access to computerized (ALDM) test systems (see below). The

133    group had *ad libitum* access to water, and feeding was provided daily at 5 pm.

134    Individuals from this group were studied during three study periods, between 2013 and

135    2015. The size of the group varied among these periods, from 21 to 22 individuals, due to births

136    or deaths from natural causes. The subjects retained for our research where those who were old

137    enough to participate in computerized testing. Supplementary table 1 reports the name, sex,

138    year of birth and participation of the individuals in each test period.

139 **Ethics statements**

140     This research was carried out in accordance with European Union and French ethical

141 standards and received approval from the French Ministère de l'Education Nationale et de la

142 Recherche (approval no. APAFIS-2717-2015111708173794-V3).

143 **Behavioural observations in the enclosure**

144     The baboons' spontaneous behaviour in their outdoor enclosure were recorded during

145 three observation periods, which were approximately evenly interspersed within a 3-year time

146 window (from 2013 to 2015, see supplementary table 1). These three periods will be referred

147 to below as the main-, control-1 and control-2 study periods. The method used for behavioural

148 recording during each period is described below.

149     ***Observational methods: main study period.*** Data were recorded during the main study

150 period from the 21$^{st}$ of September to the 21$^{st}$ of October 2015. Behavioural observations in that

151 period used a focal sampling method (Altmann, 1974). The coding scheme contained a long list

152 of affiliative and agonistic behaviours but the current paper will only consider the following six

153 behaviours which are all known to occur in agonistic contexts: "aggression", "chase", "threat",

154 "supplanting behaviour", "avoidance" and "yakking", see supplementary table 2 for

155 definitions). Every day of observation was composed of three sessions during which two

156 experimenters observed the baboons. In every other observation session one experimenter was

157 assigned the role of observing the baboons (in order to identify the behaviour performed, its

158 producer and receiver) while the other simultaneously monitored timing and entered the data in

159 the Behavioural Observation Research Interactive Software (BORIS) program (Friard &

160 Gamba, 2006). The observer had to observe each baboon during five minutes in a predefined

161 randomized order varying from sessions to sessions, for a total of 120.8 hours of observation.

162 We recorded a total of 934 agonistic events during this observation period.

163     Inter-observer reliability was assessed using two hours of video recordings of the

164     baboons' behaviours. These videos contained sequences of five minutes of observation per

165     subject and were recorded at the beginning (one hour) and end (one hour) of the observation

166     periods. Each experimenter coded independently the behaviours observed in the video using

167     the same protocol as defined above. The concordance between the two observers was coded

168     using Kappa-Cohen's method (Cohen, 1960; Landis & Koch, 1977). Inter-observer agreement

169     was high regarding the category of the behaviour ($K_{session1}=0.68$; $K_{session2}=0.85$), and the identity

170     of the baboons ($K_{session1}=0.88$; $K_{session2}=0.85$).

171     ***Observational method: control periods 1 and 2.***  The database from control period-1

172     (from the 2$^{nd}$ to the 29$^{th}$ of July 2014) contained a total of 300 hours of behavioural observation

173     which were recorded in the context of Claidière et al. (2017). The data published in Claidière

174     et al. (2017) only concerned affiliative behaviours, and therefore used a scanning procedure

175     (Altman, 1974) adapted for that purpose. Information on the agonistic behaviours collected in

176     this study has remained thus far unpublished. The dataset from control period-2 (from the 2$^{nd}$

177     to the 16$^{th}$ of May 2013) used a total of 180 hours of observation (unpublished data) previously

178     collected for the same purpose and with the same method as in period-1. In control periods 1

179     and 2, the observers used the same coding scheme as in the main study period, the only

180     difference being therefore that the data were collected with a scan instead of the focal sampling

181     method. The datasets for control periods 1 and 2 contained a total of 600 and 351 agonistic

182     events respectively, with known agonistic dyads. The difference in the number of agonistic

183     events in periods 1 and 2 is explained by the difference in duration of these periods.

184     **ALDM testing**

185     In our laboratory, the baboons have a permanent access to ten Automated Learning

186     Device for Monkeys (ALDM) systems, and this was also true during the three study periods of

187     this research. The ALDM test systems are described in detail in Fagot & Paleressompoulle

188    (2009) and Fagot & Bonté (2010). In short, each ALDM workstation consisted of a 70 cm × 70

189    cm × 80 cm test chamber equipped with (1) a RFID microchip scanner aimed at reading

190    subjects' identity from RFID microchips implanted in their arms; (2) a 19 inches' computer

191    touch screen displaying the task, and (3) a food dispenser delivering a reward inside the test

192    chamber when the subject responded correctly to the task. In practice, the baboons could freely

193    enter and leave each experimental chamber via opened doors installed in the wire mesh of their

194    enclosure.

195        When a baboon voluntarily entered an ALDM test system, it was identified via the

196    forearms' microchip and this identification triggered the initiation of the test program. All

197    ALDM trials were recorded with detailed information regarding the baboon and its

198    performance, including the date and time of the trial, as well as the workstation used. This

199    information was used to detect supplanting behaviours within the ALDM test systems, and to

200    infer the dominance hierarchy from these behaviours (see below). In practice, the baboons

201    received several different computerised cognitive tasks (e.g., Wisconsin task) of varying

202    durations during the three study periods. However, for Experiment 1 we will consider that the

203    type of cognitive test is unimportant for our current scientific purpose, because our primarily

204    goal is to determine if we can use ALDM computerised data to determine the group hierarchy.

205        **Assessing dominance hierarchy**

206        To assess the dominance hierarchy from behaviours observed in the enclosure, we built

207    three "winner/loser" files from the original data, one per study period. Whenever an agonistic

208    behaviour was observed, the individuals displaying aggression, chase, threat, or supplanting

209    behaviour was considered "winner" in each dyad, while the baboons receiving these behaviours

210    were considered "losers". The coding was reversed for the avoidance and yakking behaviours,

211    which correspond to social submission.

212   To assess dominance from the data collected in the ALDM test systems, we reasoned

213 that supplanting behaviours in the ALDM test systems should be indicated by a very short

214 interval between the last trial of the supplanted subject ("loser") and the first trial of the

215 supplanting subject ("winner"). Based on this reasoning, we analysed the ALDM data in two

216 steps. Firstly, we defined a "run" as a succession of trials by the same individual within the

217 same workstation, and with inter-trial interval of less than 30 seconds (i.e. not exceeding twice

218 the average delay between two consecutive trials). A run therefore represents a succession of

219 trials corresponding to a single foraging bout. Secondly, to identify supplanting behaviours

220 within the workstations (i.e. cases in which the foraging bout was disrupted), we computed the

221 time interval between the last trial of one baboon within a given workstation and the first trial

222 of the next individual in the same workstation. We considered that this behavioural sequence

223 corresponded to a supplanting behaviour when this interval between two individuals was

224 smaller than 30 seconds (i.e. the foraging bout had been disrupted). Using this procedure, we

225 obtained an ALDM winner/loser files for each observation period, which had the same structure

226 as the winner/loser file from the spontaneous behaviour in the enclosure. This procedure

227 allowed the identification of 7844 cases of supplanting behaviours (for 22132 runs considered)

228 during the main study period (in 2015), and 7700 (22769 runs) and 7479 (17634 runs) during

229 the control periods 1 and 2 respectively.

230   The Elo-rating method (Elo, 1978) was used to calculate dominance hierarchy from the

231 sets of the winner/loser files described above. Interested readers are referred to Elo (1978) for

232 a detailed description of this method. In short, each individual started with the same score with

233 this method, and at each new contest the score of the winner increased by a certain amount

234 while it decreased for the loser by the same amount. To calculate the dominance hierarchy, we

235 used RandomElo function from the Elo-rating package (Neumann et al., 2011) that ran on the

236 statistical environment R (R version 4.0.2 ; R Core Team, 2015), and followed Neumann &

237    Kulik (2014)'s procedure, with starting value of 1000 and a constant $k$ optimized using

238    maximum log-likelihood (provided in supplemental table 4). For each inferred dominance

239    hierarchy, through the 3 sampling periods and both types of data sets, we assessed triangle

240    transitivity and linearity index using Elo-rating package viewable in supplemental table 4.

241    Finally, we ensured the repeatability of the hierarchies by computing repeatability scores using

242    AniDom package (Sánchez-Tójar, Schroeder & Farine, 2018), using the same constant $k$

243    optimized and 1000 randomisations.

244        **Statistical analyses**

245        ***Correlational analyses.*** A first way to validate our approach is to directly compare the

246    correlation between the dominance hierarchies obtained through behavioural observations and

247    through computerised testing. We used two different correlation tests on Elo-scores: Spearman

248    and Kendall correlations. The non-parametric Spearman correlation test compares the relative

249    order of dominance hierarchies (for example if A is above B in one hierarchy and B above A

250    in the other) and is widely used in the literature (Balasubramaniam et al., 2013; Hunter, Broom,

251    Edwards & Sibly, 1988; Williamson, Lee & Curley, 2006). The Kendall correlation test is less

252    frequently used (Bang et al., 2010; Hemelrijk et al., 2005) but compares the strict value of the

253    ranks of individuals in the hierarchy (for example if A is +3 ranks above B in one hierarchy and

254    -1 rank in another) and therefore seems more precise. For the purpose of exhaustivity, we report

255    below the results of both tests.

256        ***Modelling.*** A high correlation between the two hierarchies would show that both

257    hierarchies are similar but would not inform us on the cause of this similarity. For instance, it

258    is possible that dominant individuals are more frequently involved in cognitive testing than

259    subordinate individuals, and that a higher number of comings and goings in the workstations

260    increases the rank of individuals in the hierarchy calculated from ALDM testing. Exact

261    permutation tests (Fisher, 1954) were used to assess the contribution of such variables to our

262    measurement of dominance hierarchies. In an exact permutation test, we perform a large

263    number of permutations of the variables that may not account for the observed correlation while

264    maintaining the variable of interest constant, every time calculating the correlation between the

265    hierarchy obtained from the randomly permutated computerised data and the one obtained from

266    observed data. This generates a distribution of correlation coefficients obtained under the null

267    hypothesis that the factors that have been permutated do not account for the correlation. If the

268    observed correlation falls outside the 95% confidence interval of the distribution of simulated

269    coefficients, the null hypothesis can be rejected, and we can conclude that the factors that have

270    been permutated are important in explaining the correlation.

271            Two main parameters can influence the position of the individuals in the dominance

272    hierarchy, they include the frequency of agonistic interactions (i.e. supplanting behaviours)

273    each baboon has in the workstation (which could be linked to proportion of contest in traditional

274    hierarchy assessment; David, 1987) and the identity of the winner-loser dyads involved in each

275    agonistic interaction. We therefore used two random models, one testing each variable. This

276    modelling approach was applied to the analysis of the data collected during the main study

277    period only, because this dataset was the largest.

278            ***Random model 1.*** In the first model, the identity of individuals between two recordings

279    was randomly permutated before supplanting behaviours computation (we performed 50 000

280    permutations for 22132 runs for each of the 500 simulations). Each monkey maintained in

281    random model 1 the same frequency of participation in cognitive tasks, but the permutation of

282    the subjects' identity in the dataset altered the proportion of supplanting behaviours performed

283    by each participant as well as the identity of the winner-loser dyads. If the dominance hierarchy

284    obtained from the ALDM is determined mostly by the baboons' frequency of participation in

285    experiments, then we expect the correlation obtained from the simulations to be similar to that

286    obtained with the untransformed ALDM data.

287     ***Random model 2.*** The aim of the second model was to examine if the hierarchy can

288     depend on the frequency of contests won, irrespective of the identity of the dyad involved in

289     the supplanting behaviour. We performed 500 simulations in which we performed 50 000

290     permutations of the identity of the winners of two different supplanting behaviours (making

291     sure the same individual did not appear both as winner and loser). Each individual therefore

292     maintained the same number of contests won but with different losers. If the hierarchy is mostly

293     determined by the number of contests won and not by the winner-loser identity, then we expect

294     the correlation between the hierarchy obtained from the simulations to be similar to that

295     obtained with untransformed data.

296   **Results**

297   **Reliability and repeatability of the hierarchies**

298     The hierarchies obtained for the three periods and two types of data are highly repeatable

299     and reliable (Repeatability index were very high: $r_{Obs} = 0.92$ and $r_{ALDM} = 0.96$ for the main study

300     period; $r_{Obs} = 0.88$ and $r_{ALDM} = 0.97$ for the Control period-1 and $r_{Obs} = 0.94$ and $r_{ALDM} = 0.95$

301     for the Control period-2). Summary table in the appendix present the sets of Elo-scores and

302     ranks for each baboon (supplemental table 3), as well as the linearity and transitivity indices of

303     the triangles for each hierarchy (supplemental table 4).

304   **Main study period**

305     The correlation between the hierarchies obtained with observational and computerised

306     data is shown on the left-hand panel of Figure 1. It is high and different from chance (Kendall

307     correlation: $\tau = 0.75$, CI = [0.59; 0.89], N = 21, $p < 0.001$; Spearman correlation: $r_s = 0.91$, CI

308     = [0.76; 0.97], N = 21, $p < 0.001$). Individual Elo-scores and ranks are presented in

309    supplementary table 3. The absolute difference in rank between the ALDM and observational

310    dominance hierarchies remained relatively low, with an average equal to 2.04 ranks (max = 6).

311                                  Insert Figure 1 about here

312

313    **Random model 1**

314         For this model, we found that the correlation between the simulated hierarchies obtained

315    by permutations and  the observation's hierarchy was not  significantly different from 0 (see

316    Figure 2, mean Kendall correlation:  $\tau_{sim}$ = 0.15, 95% CI = [-0.16; 0.43]; mean Spearman

317    correlation: $r_{s\,sim}$ = 0.21, 95% CI = [-0.19; 0.60]), and  significantly smaller than the correlation

318    obtained with untransformed ALDM data ($\tau$ = 0.75, CI = [0.59; 0.89], $p < 0.001$; $r_s$ = 0.91, CI

319    = [0.76; 0.97], $p < 0.001$). This result shows that the mere variations in the frequency of the

320    baboons' participation to cognitive testing were not at the origin of the strong correlation that

321    we observed.

322                                  Insert Figure 2 about here

323    **Random model 2**

324         For  random  model  2,  the  distribution  of  simulated  correlation  coefficients  was

325    significantly different from 0 (see Figure 2, mean Kendall correlation: $\tau_{sim}$= 0.64, 95% CI =

326    [0.60; 0.68]; mean Spearman correlation: $r_{s\,sim}$ = 0.83, 95% CI = [0.81; 0.86]). Additionally, the

327    distribution  of  simulated  correlated  coefficients  was  also  significantly  smaller  than  the

328    correlation obtained with untransformed ALDM data ($\tau$ = 0.75, CI = [0.59; 0.89], $p < 0.001$; $r_s$

329    = 0.91, CI = [0.76; 0.97], $p < 0.001$), showing that ALDM dominance hierarchies reflect the

330    social assortment of the individuals in each supplanting dyad.

**Correlational analyses: all three study periods**

Figure 1 illustrates the correlations obtained between the observational and ALDM dominance hierarchies in the three study periods. As indicated above, that correlation was high and significant for the main study period (Figure 1, left panel). This was confirmed by the analysis of the two other study periods as shown in the middle and right panel of Figure 1 (control period-1: Kendall correlation: $\tau_{c1} = 0.76$, CI = [0.57; 0.91], N = 22, $p < 0.001$ and Spearman correlation: $r_{sc1} = 0.9$, CI = [0.71; 0.97], N = 22, $p < 0.001$; control period-2: Kendall correlation: $\tau_{c2} = 0.76$, CI = [0.57; 0.91], N = 21, $p < 0.001$ and Spearman correlation: $r_{sc2} = 0.9$, CI = [0.70; 0.98], N = 21, $p < 0.001$). On average, the absolute difference in ranking remained relatively low in both control periods, it was equal to 2.0 ranks (max = 6) and 2.1 ranks (max = 8) for the control periods 1 and 2, respectively (Individual Elo-scores and ranks are presented in supplementary table3).

**Discussion of Experiment 1**

Experiment 1 employed two methods to assess dominance hierarchy, conventional observations on the one hand and the analysis of supplanting behaviours during ALDM testing on the other hand. These two methods presented very similar rankings in all three datasets covering a period of 3 years, considering both Kendall and Spearman correlations and similarly high repeatability indices. Simulations further demonstrate that the hierarchies obtained with the ALDM method reflect the social dominance hierarchy, independently of variations in participation to cognitive testing. Both the frequency of supplanting behaviours and the identity of the individuals involved in the supplantation are crucial factors determining the dominance hierarchy. We believe that such results, at the group and individual levels, demonstrate that dominance hierarchy can be inferred reliably from the analysis of supplanting behaviours within the ALDM workstations

355

356                    **Experiment 2: Effect of supplanting behaviours on cognitive performance**

357          The main goal of Experiment 2 is to illustrate the possible use of ALDM hierarchies

358    obtained from supplanting behaviours in the context of cognitive studies. Goujon & Fagot,

359    (2013) have previously published data on the cognitive performance of our group of baboons

360    in the ALDM test system. This research used the so called "contextual cueing" task (Chun &

361    Jiang, 1998), in which the baboons had to detect a target in two types of visual displays, first

362    on non-predictive backgrounds providing no information on target location, and second on

363    predictive backgrounds which have been associated by training to specific target location on

364    the computer screen. Results indicated a strong effect of the background on the speed of target

365    detection, with faster response times (RTs) in the predictive compared to the non-predictive

366    condition.  In Experiment 2, we reanalysed Goujon and Fagot's (2013) data in order to assess

367    if the ALDM supplanting behaviours recorded during that research had subtle (but previously

368    undetected) effects on the baboon's cognitive performance.

369    **Methods**

370    **Participants**

371          The study used the same baboon colony as in Experiment 1. However, because this

372    study was conducted in 2012, and therefore one year before the control period 2 of Experiment

373    1, this colony had a slightly different composition. The data set therefore contained information

374    on a total 20 baboons (5 males and 15 females) whose name and biographical information are

375    indicated in the supplementary section.

376 **Experimental setup, computerized task and database**

377       Experiment 2 used the same ALDM workstations as in Experiment 1. The experimental

378 task is already presented in detail in Goujon and Fagot (2013) and will therefore only be briefly

379 described here. In short, the trials started with the display of a fixation stimulus on the screen.

380 Touching this stimulus triggered the display of 8 items, including one target (the capital letter

381 "T") and 7 distractors (capital "L" with different orientations). To be rewarded, the baboons

382 had to detect the target on the screen and to touch it. In the predictive condition (P), six patterns

383 of distractors were each associated with a fixed target position. In the non-predictive condition

384 (NP), six different configurations of distractors were presented independently of the target

385 position. The P and NP trials were presented for one month (July 3$^{rd}$ to August 3$^{rd}$ 2012) within

386 a balanced design. Due to a few technical problems, a small number of ALDM test boxes were

387 temporally stopped during testing. Because a too small number of boxes may have affected the

388 monkeys' turnover and the detection of supplanting behaviours, we only retained the data

389 corresponding to the days when a minimum of 8 boxes were accessible. This criterion lead to a

390 very small rate of rejection (only 3 days), while keeping high the likelihood of detecting

391 supplanting behaviours. The database analysed here contains a total of 411534 trials (mean =

392 20576, median = 21 134, SD = 6355 per baboon).

393 **Data analysis**

394       Investigating the effect of supplanting behaviours on loser of each dyad is difficult,

395 because the subjects either stop temporarily after a supplanting event, or move to a different

396 workstation, with a high variability in time and behaviour between the last trial after the

397 supplantation and the first trial of the next working bout. The analysis therefore focussed on the

398 winner who performed trials immediately after the supplanting event. We considered two types

399 of trials for that subject: those made after the baboon had supplanted another baboon (named

400 "Supplant" trial), and control trials in which it arrived in an empty workstation (baseline trial).

401 Data analysis therefore proceeded in two steps (1) identification of supplanting and baseline

402 trials, and (2) analysis of the baboons' RT depending on whether the trials were recorded or not

403 after a supplanting event. We used the same procedure as in Experiment 1 to identify runs and

404 instances of supplanting behaviours. A total of 20556 runs were identified with 9310 runs

405 starting after a supplanting event. The dominance hierarchy of the group was computed based

406 on these ALDM supplanting behaviours using the same procedure as in Experiment 1.

407 Because of a very low rate of errors in cognitive test (range: 96%-99%) statistical

408 analyses focused on RTs. Social challenges can induce extremely rapid biological effects in

409 animals (e.g., neuropeptides release, Kelly & Wilson, 2019), with potential consequences on

410 behaviour and cognition (see the discussion of Experiment 2 for additional information). We

411 therefore conducted a preliminary analysis of the data and found that RT differences across

412 conditions were mostly found on the very first trials following supplanting behaviour. On that

413 basis, our next analyses only considered the first 6 trials of each run. The RT of each trial was

414 considered for statistical analyses only if (1) the subject gave a correct response in the trial, and

415 (2) the RT did not exceed three standard deviations from the individual mean RT. This standard

416 filtering procedure avoided consideration of outlier trials resulting from transient inattention

417 (e.g., Lacreuse, Gullstrand & Fagot, 2016). The final dataset contained a total of 99 619 trials.

418 The average number of trials per individuals were as follows for each type of trial: Baseline-

419 NP: mean = 1351, median = 1201, SD = 730; Baseline-P: mean = 1368, median = 1213, SD =

420 743; Supplant-NP: mean = 1129, median = 890, SD = 674; Supplant-P: mean=1133, median =

421 893, SD = 667.

422 Linear mixed-effect models (LMM) were computed to analyse the results, using the

423 lme4 package in R (Bates & Machler, 2010). Goujon & Fagot (2013) have reported that the

424 RTs decreased in this task with the number of days of testing. We therefore included as random

425    factors the "Name" of the subjects with a random intercept and slope depending on the number

426    of "Days" of the experiment. This random factor controlled for repeated measurements and the

427    evolution of RT with learning. We also included a simple random intercept for the "Age" of the

428    subjects, because Goujon & Fagot (2013) reported inter-individual differences between age

429    groups. Based on the above considerations, we report the results of the three way interaction

430    between the following fixed factors: (1) the type of trial (baseline vs supplant trial, referred to

431    as the "Behavioural condition" in this analysis), (2) the type of contextual background (P or

432    NP, referred to as the "Experimental condition") and (3) the position of the trial in the run (from

433    1 to 6, referred to as "Ntrial" variable).

434        Finally, a last model investigated if the effect on RTs was modulated by the difference

435    of hierarchical position of both protagonists of the supplanting behaviour. In this model, we

436    only considered as dependent variable the RT of the very first trial performed by the winner

437    after a supplanting behaviour in the non-predictive condition, because the effect of

438    supplantation is stronger on the very first trial and then progressively fades. The explanatory

439    variable was the difference of rank between the winner and the loser of the ALDM supplanting

440    behaviour. With 20 individuals in the group, difference in ranks could vary between -19 (if the

441    most subordinated individual supplanted the most dominant) and +19 (if the most dominant

442    individual supplanted the most subordinated).

443    **Results**

444        The results of the three-way interaction model are shown in Table 1. Our results confirm

445    the main effect of experimental condition (P vs NP) previously described, with faster RTs in

446    the predicted condition (on average by 54 ms; RT mean = 541 ms, SD = 214 ms). In addition,

447    there was a small effect of the position of the trial in a run with RTs decreasing on average by

448    1.36 ms with the number of trials performed. Finally, RTs were shorter in trials following a

449  supplanting behaviour compared to trials performed without supplanting behaviours, on

450  average by 20 ms on the first trial. We found no evidence of interactions between these different

451  variables.

452                         Insert Table 1 about here

453  Regarding the effect of the difference of hierarchical position (between winner and loser of the

454  supplanting behaviour), we found significant RT differences depending on difference in ranks

455  between the winner and the loser: The most challenging supplanting behaviours, corresponding

456  to the supplanting of a higher-ranking individual by a lower-ranking individual gave rise to the

457  shortest RTs during the ALDM trials immediately following that behaviour, with an estimated

458  difference in RTs of about 1.75 ms (SE = 0.56 ms, t value = 3.12, $p < 0.001$) for each difference

459  in ranks. For instance, when Ewine (10th rank) supplants Angele (4th rank) the estimated RT

460  is 26 ms lower than if she supplants Dream (19th rank), because there is a difference of 15 ranks

461  between Angele and Dream.

462  **Discussion of Experiment 2**

463          Experiment 2 investigated the effect of supplanting behaviours on cognitive

464  performance. We analysed the first six trials of each run performed in the ALDM system,

465  distinguishing those performed after a supplanting behaviour from those without. The results

466  showed a significant effect on baboons' RTs: RTs were shorter in trials following supplanting

467  behaviour, compared to baseline trials. This difference was significant in both predictive and

468  non-predictive conditions. We also found that RTs decreased when the winner supplanted an

469  individual of a higher social rank suggesting that the response speed can be affected by social

470  challenges.

471          Since the goal of the current paper is mostly methodological, we will not discuss at

472  length the behavioural and brain mechanisms by which the supplanting behaviours can affect

473   RTs but will only briefly suggest three hypotheses that may account for our findings. A first

474   potential mechanism is a release of steroid hormones which are known to occur in animals after

475   social challenges (e.g., for a review see Kelly and Wilson, 2019). Androgens and especially

476   testosterone are known to affect cognitive performance (e.g., Janowsky, 2006) and to enhance

477   spatial abilities. Alternatively, the release of neuropeptides such as oxytocin could explain our

478   results, knowing their contribution to prosocial and aggressive behaviours. The action of

479   oxytocin has been reported on various types of social behaviours, including social approach

480   (Thompson and Walton, 2004) and aggression (Bosch, 2013). Note that oxytocin releases have

481   rapid effects on the behavioural response (e.g., Knobloch et al., 2012). This may account for

482   the transient effects of supplanting behaviour in our task. Finally, faster RTs after a supplanting

483   behaviour can be accounted for by a winner effect (Chase, Bartolomeo & Dugatkin, 1994).

484   Reeve, Olson & Cole (1985) argued that winning enhances intrinsic motivation in the task

485   relative to loosing. According to this view, the baboon winning the contest would have an

486   enhanced motivation to perform well in the computerized task, even more so if subordinates

487   were supplanting dominants, gaining a time advantage on the task.

488         Independently of the validity of these different hypotheses, which remains to be tested,

489   Experiment 2 has revealed that the dominance hierarchy can have subtle effects on cognitive

490   performance that would probably remain undetectable with more traditional experimental

491   approaches. We believe that the analysis of such effects can help answering questions regarding

492   the interaction between social and non-social cognition.

493                                   **General discussion**

494         The main goal of Experiment 1 was to demonstrate that the baboons' supplanting

495   behaviours within the ALDM test system can be used reliably to determine the dominance

496   hierarchy of the group. We showed that the dominance hierarchy inferred from untransformed

497   ALDM data differed from those obtained after a random permutation of the individuals

498   performing the trial, suggesting that the ALDM dominance hierarchies are independent of the

499   frequency of baboon's participation to ALDM testing. Secondly, the dominance hierarchies

500   inferred from untransformed ALDM data also differed from those obtained after random

501   permutation between winners, showing that the dominance hierarchy crucially depends on the

502   identity of the individuals interacting. Thirdly, and most importantly, strong correlations were

503   found between the dominance hierarchies obtained from conventional observations and those

504   obtained from ALDM testing for three independent study periods spanning a total of three years

505   (all correlations' r > 0.7). Earlier studies with sows (Hunter et al., 1988), birds (Evans, Devost,

506   Jones, & Morand-Ferron, 2018) and cows (Foris, Thompson et al., 2019) have already shown

507   the interest of computer-controlled identification systems to infer dominance hierarchies.

508          Dominance hierarchies are often difficult to measure in non-human primates, such as in

509   the male-male tolerant Guinea baboons (Fischer et al., 2017; Patzel, 2013). It might be the case

510   that agonistic behaviour are more frequent in captivity (such as in our study) than in the wild,

511   and this could explain why we found a linear dominance hierarchy in our group, while Fischer

512   et al. (2017) did not. Nevertheless, the present study on the same baboon species as in Fischer

513   et al. (2017) highlights the usefulness and reliability of supplanting behaviours in the

514   assessment of the dominance hierarchy. Its findings match perfectly with the previous studies

515   on baboons 'supplanting behaviours (Packer, 1979; Johnson, 1989) that reported that baboons

516   were using supplanting behaviours to reinforce existing social relationships. The consistency

517   of our results, across test periods and methods, opens new perspectives for future studies on

518   dominance hierarchy, and demonstrates that such hierarchies can be inferred from the analysis

519   of supplanting behaviours during ALDM testing.

520          The current study showing that dominance hierarchies can be calculated from

521   automatized ALDM testing also converges with earlier studies demonstrating that *ad-libitum*

522 computerized cognitive testing allows the extraction of social information. Previous studies

523 from our research group have already demonstrated that ALDM cognitive testing permits

524 studies on the static social network (Claidière et al., 2017), on the dynamic of social networks

525 and their instability (Gelardi et al., 2019), as well as on audience (Huguet, Barbet, Belletier,

526 Monteil & Fagot, 2014 ) and social comparison effects (Dumas, Fagot, Davranche & Claidière,

527 2017). It was also previously found that the cognitive performance measured during ALDM

528 cognitive testing is modulated by socially affiliative and agonistic events in the enclosure

529 (Marzouki, Gullstrand, Goujon & Fagot, 2014). Studies from other laboratories have further

530 shown that the ordering of the interactions with such automatized test systems depends on social

531 factors in Rhesus monkeys (*Macaca mulatta*): monkeys of different ranks tended not to work

532 at the same times to avoid social conflicts (Gaze, Lutz, Meyer, Hassetts & Hampton, 2019).

533 Social effects of freely accessible test systems were also reported in crested macaques (*Macaca*

534 *nigra*) who increased their association with others during the testing days (Whitehouse,

535 Michelatta, Powell, Border & Waller, 2013). From all these studies, we conclude that the

536 spontaneous interactions within automatized test systems presented to groups of individuals

537 should be interpreted as social events, and not only as purely cognitive events. Automatized

538 test systems therefore offer a unique opportunity to conduct experiments on either physical or

539 social cognition in non-human primates, as well as on the interactions between the two. The

540 transient effect of the supplanting behavior on the RTs recorded in the cognitive task of

541 Experiment 2 is a new illustration of the possibility offered by ALDM testing to work at the

542 interface between social and non-social cognition.

543 Modern analyses of behavior tend to be more and more sophisticated and call for the

544 development of new systems to collect social information automatically. The current study

545 shows that the provision of automatized operant conditioning test systems to monkeys

546 maintained in social groups can fulfil that need. One of the most obvious advantage of this

547     method is that extremely large datasets can be collected efficiently to infer dominance

548     hierarchy, in sharp contrast to the more effortful traditional observational methods. Of course,

549     this approach also has some limitations. One of them is that it requires sophisticated

550     technologies that can hardly be accessible to every laboratory working in comparative

551     cognition.  The last few years have witnessed the development of  freely accessible test systems

552     comparable to the ALDM: in monkeys (e.g., Gazes, Brown, Basile & Hampton, 2012; Fizet et al.,

553     2017; Tulip, Zimmermann, Farningham & Jackson, 2017; Butler & Kennerley, 2018), as well

554     as in other groups such as birds (Huber, Heise, Zenan & Palmers, 2015; Morand-Ferron,

555     Hamblin, Cole,  Aplin & Quinn, 2015). We are therefore convinced that such technology will

556     develop even more in the future, opening unprecedented opportunities to grasp the full

557     complexity of animal's social and non-social cognition.

558

559     **Open Practices Statements**:

560     The data analysed in this study are available here: https://osf.io/w39fr/

561

**References**

562

563    Alarcón-Nieto, G., Graving, J. M., Klarevas-Irby, J. A., Maldonado-Chaparro, A. A., Mueller,

564        I., & Farine, D. R. (2018). An automated barcode tracking system for behavioural

565        studies in birds. *Methods in Ecology and Evolution, 9*, 1536–1547.

566    Altmann, J. (1974). Observational study of behaviour: Sampling methods. *Behaviour*, *49*, 227-

567        266.

568    Bang, A., Deshpande, S., Sumana, A., & Gadagkar, R. (2010). Choosing an appropriate index

569        to construct dominance hierarchies in animal societies: A comparison of three indices.

570        *Animal Behaviour*, *79*, 631-636.

571    Bates, D. & Machler, M. (2010): Lme4: Linear mixed-effects models using S4 classes. R

572        package      version      0.999375-39.      Available      at:      http://cran.r-

573        project.org/web/packages/lme4/

574    Balasubramaniam, K. N., Berman, C. M., De Marco, A., Dittmar, K., Majolo, B., Ogawa, H.,

575        Thierry, B., & De Vries, H. (2013). Consistency of dominance rank order: A comparison

576        of David's scores with I&SI and Bayesian methods in macaques. *American journal of*

577        *primatology*, *75*, 959-971.

578    Barton, R. A., Byrne, R. W., & Whiten, A. (1996). Ecology, feeding competition and social

579        structure in baboons. *Behavioral Ecology and Sociobiology*, *38*, 321-329.

580    Bosch, O. J. (2013). Maternal aggression in rodents: Brain oxytocin and vasopressin mediate

581        pup defence. *Philosophical Transactions of the Royal Society B: Biological Sciences,*

582        *368*, 20130085.

583    Chase, I. D., Bartolomeo, C., & Dugatkin, L. A. (1994). Aggressive interactions and inter-

584        contest interval: How long do winners keep winning? *Animal Behaviour*, *48*, 393-400.

585    Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual

586        context guides spatial attention. *Cognitive psychology, 36*, 28-71.

587 Claidière, N., Gullstrand, J., Latouche, A., & Fagot, J. (2017). Using Automated Learning
588      Devices for Monkeys (ALDM) to study social networks. *Behaviour research methods,*
589      *49*, 24-34.

590 Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and*
591      *Psychological Measurement, 20,* 37-46.

592 David, H. A. (1987). Ranking from unbalanced paired-comparison data. *Biometrika*, *74*, 432-
593      436.

594 Dugatkin, L. A. (1997). Winner and loser effects and the structure of dominance hierarchies.
595      *Behavioral Ecology*, *8*, 583-587.

596 Dumas, F., Fagot, J., Davranche, K., & Claidière, N. (2017). Other better versus self better in
597      baboons: an evolutionary approach of social comparison. Proceedings of the Royal
598      Society B: Biological Sciences, 284(1855), 20170248.

599 Elo, A. E. (1978). *The rating of chessplayers, past and present*. New York, Arco Pub.

600 Evans, J. C., Devost, I., Jones, T. B., & Morand-Ferron, J. (2018). Inferring dominance
601      interactions from automatically recorded temporal data. *Ethology, 124,* 188-195.

602 Fagot, J., Paleressompoulle, D., (2009). Automatic testing of cognitive performance in baboons
603      maintained in social groups. Behavior Research Methods*, 41*, 396–404.

604 Fagot, J., Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of a
605      battery of computerized test systems by a troop of semi-free ranging baboons. *Behavior*
606      *Research Methods*, *42*, 507–516.

607 Ferreira, A. C., Silva, L. R., Renna, F., Brandl, H. B., Renoult, J. P., Farine, D. R., Covas, R.,
608      & Doutrelant, C. (2020). Deep learning-based methods for individual recognition in
609      small birds. *Methods in Ecology and Evolution, 11*, 1072-1085.

610 Fischer, J., Kopp, G. H., Dal Pesco, F., Goffe, A., Hammerschmidt, K., Kalbitzer, U.,
611      Klapproth, M., Maciej, P., Ndao, I., Patzelt, A., & Zinner, D. (2017). Charting the

612       neglected West: The social system of Guinea baboons. *American Journal of Physical*

613       *Anthropology, 162*, 15-31.

614 Fizet, J., Rimele, A., Pebayle, T., Cassel, J. C., Kelche, C., & Meunier, H. (2017). An

615       autonomous, automated and mobile device to concurrently assess several cognitive

616       functions in group-living non-human primates. *Neurobiology of learning and memory,*

617       *145*, 45-58.

618 Flack, J. C., De Waal, F. B., & Waal, D. (2004). Dominance style, social power, and conflict

619       management: a conceptual framework. In: *Macaque Societies: A Model for the Study of*

620       *Social Organization*. (Ed. by B. Thierry, M. Singh & W. Kaumanns, pp. 157–182). New

621       York: Cambridge University Press.

622 Foris, B., Thompson, A. J., Keyserlingk, M. A. G. von, Melzer, N., & Weary, D. M. (2019).

623       Automatic detection of feeding- and drinking-related agonistic behavior and dominance

624       in dairy cows. *Journal of Dairy Science, 102*, 9176-9186.

625 Foris, B., Zebunke, M., Langbein, J., & Melzer, N. (2019). Comprehensive analysis of

626       affiliative and agonistic social networks in lactating dairy cattle groups. Applied Animal

627       Behaviour Science, 210, 60–67.

628 Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software

629       for video/audio coding and live observations. *Methods in Ecology and Evolution*.

630 Gazes, R. P., Brown, E. K., Basile, B. M., & Hampton, R. R. (2013). Automated cognitive

631       testing of monkeys in social groups yields results comparable to individual laboratory-

632       based testing. *Animal cognition*, *16*, 445-458.

633 Gelardi, V., Fagot, J., Barrat, A., & Claidière, N. (2019). Detecting social (in) stability in

634       primates from their temporal co-presence network. Animal Behaviour, 157, 239-254.

635 Goujon, A., & Fagot, J. (2013). Learning of spatial statistics in nonhuman primates: contextual

636       cueing in baboons (Papio papio). *Behavioural Brain Research*, *247*, 101-109.

637 Hemelrijk, C. K., & Gygax, L. (2004). Dominance style, differences between the sexes and

638   individuals: An agent-based model. Interaction Studies, 5, 131-146.

639 Hemelrijk, C. K., Wantia, J., & Gygax, L. (2005). The construction of dominance order:

640   comparing performance of five methods using an individual-based model. Behaviour,

641   142, 1037-1058.

642 Huber, L., Heise, N., Zeman, C., & Palmers, C. (2015). The ALDB box: Automatic testing of

643   cognitive performance in groups of aviary-housed pigeons. Behavior research methods,

644   47, 162-171.

645 Huguet, P., Barbet, I., Belletier, C., Monteil, J. M., & Fagot, J. (2014). Cognitive control under

646   social influence in baboons. *Journal of Experimental Psychology: General, 143*, 2067.

647 Hunter, E. J., Broom, D. M., Edwards, S. A., & Sibly, R. M. (1988). Social hierarchy and feeder

648   access in a group of 20 sows using a computer-controlled feeder. *Animal Science*, *47*,

649   139-148.

650 Janowsky, J. S. (2006). Thinking with your gonads: testosterone and cognition. *Trends in*

651   *cognitive sciences*, *10*(2), 77-82.

652 Johnson, J. A. (1989). Supplanting by olive baboons: dominance rank difference and resource

653   value. *Behavioral Ecology and Sociobiology*, *24*, 277-283.

654 Kelly, A. M., & Wilson, L. C. (2019). Aggression: Perspectives from social and systems

655   neuroscience. *Hormones and behavior*. DOI: 10.1016/j.yhbeh.2019.04.010

656 King, J. A. (1973). The ecology of aggressive behavior. *Annual review of ecology and*

657   *systematics*, 117-138.

658 Knobloch, H. S., Charlet, A., Hoffmann, L. C., Eliava, M., Khrulev, S., Cetin, A. H., Osten, P.,

659   Schwarz, M. K., Seeburg, P. H., Stoop, R., & Grinevich, V. (2012). Evoked axonal

660   oxytocin release in the central amygdala attenuates fear response. *Neuron, 73*, 553-566.

661    Lacreuse, A., Gullstrand, J., & Fagot, J. (2016). Sex differences in inhibitory control in socially-

662         housed baboons (Papio papio). *Behavioural brain research, 312*, 231-237.

663    Landau, H. G. (1951). On dominance relations and the structure of animal societies: I. Effect

664         of inherent characteristics. *The bulletin of mathematical biophysics*, *13*, 1-19.

665    Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical

666         Data. *Biometrics, 33*, 159-174.

667    Marzouki, Y., Gullstrand, J., Goujon, A., & Fagot, J. (2014). Baboons' response speed is biased

668         by their moods. *PloS one*, *9*, e102562.

669    Morand-Ferron, J., Hamblin, S., Cole, E. F., Aplin, L. M., & Quinn, J. L. (2015). Taking the

670         operant paradigm into the field: associative learning in wild great tits. *PloS one, 10*,

671         e0133821.

672    Nagy, M., Vásárhelyi, G., Pettit, B., Roberts-Mariani, I., Vicsek, T., & Biro, D. (2013). Context-

673         dependent hierarchies in pigeons. *Proceedings of the National Academy of Sciences,*

674         *110*, 13049–13054.

675    Neat, F. C., Taylor, A. C., & Huntingford, F. A. (1998). Proximate costs of fighting in male

676         cichlid fish: the role of injuries and energy metabolism. *Animal Behaviour*, *55*, 875-882.

677    Neumann, Christof & Duboscq, Julie & Dubuc, Constance & Ginting, Andri & Maulana, Ardi

678         & Agil, Muhammad & Widdig, Anja & Engelhardt, Antje. (2011). Assessing

679         dominance hierarchies: Validation and advantages of progressive evaluation with Elo-

680         rating. *Animal Behaviour, 82,* 911-921..

681    Neumann, C., & Kulik, L. (2014). Elo-rating-a brief tutorial. R package version 0.41; R version

682         3.1.1, available at: http://CRAN.R-project.org/package=Elo-Rating.

683    Packer, C. (1979). Male dominance and reproductive activity in Papio anubis. *Animal*

684         *behaviour, 27*, 37-45.

685 Patzelt, A. (2013). The social system of Guinea baboons (Papio papio) with a focus on male-
686       male relationships (Doctoral dissertation, Niedersächsische Staats-und
687       Universitätsbibliothek Göttingen).

688 R Core Team (2015). R: A language and environment for statistical computing. R Foundation
689       for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

690 Reeve, J., Olson, B. C., & Cole, S. G. (1985). Motivation and performance: Two consequences
691       of winning and losing in competition. *Motivation and Emotion*, *9*, 291-298.

692 Sánchez-Tójar, A., Schroeder, J., & Farine, D. R. (2018). A practical guide for inferring reliable
693       dominance hierarchies and estimating their uncertainty. *Journal of Animal Ecology, 87*,
694       594-608.

695 Strandburg-Peshkin, A., Farine, D. R., Couzin, I. D., & Crofoot, M. C. (2015). Shared decision-
696       making drives collective movement in wild baboons. *Science, 348*, 1358-1361.

697 Thompson, R. R., & Walton, J. C. (2004). Peptide effects on social behavior: effects of
698       vasotocin and isotocin on social approach behavior in male goldfish (Carassius auratus).
699       *Behavioral neuroscience, 118*, 620.

700 Tulip, J., Zimmermann, J. B., Farningham, D., & Jackson, A. (2017). An automated system for
701       positive reinforcement training of group-housed macaque monkeys at breeding and
702       research facilities. *Journal of neuroscience methods, 285,* 6-18.

703 Vries, H. D. (1998). Finding a dominance order most consistent with a linear hierarchy: a new
704       procedure and review. *Animal Behaviour*, *55*, 827-843.

705 Vries, H., Stevens, J. M., & Vervaecke, H. (2006). Measuring and testing the steepness of
706       dominance hierarchies. *Animal Behaviour*, *71*, 585-592.

707 Watts, D. P. (1994). Agonistic relationships between female mountain gorillas (Gorilla gorilla
708       beringei). *Behavioral Ecology and Sociobiology, 34*, 347-358.

709 Weissbrod, A., Shapiro, A., Vasserman, G., Edry, L., Dayan, M., Yitzhaky, A., & Kimchi, T.

710    (2013).Automated long-term tracking and social behavioural phenotyping of animal

711    colonies within a semi-natural environment. *Nature communications, 4*, 2018.

712 Whitehouse, J., Micheletta, J., Powell, L. E., Bordier, C., & Waller, B. M. (2013). The impact

713    of cognitive testing on the welfare of group housed primates. *PloS one, 8*, e78308.

714 Williamson, C. M., Lee, W., & Curley, J. P. (2016). Temporal dynamics of social hierarchy

715    formation and maintenance in male mice. *Animal Behaviour*, *115*, 259-272.

716
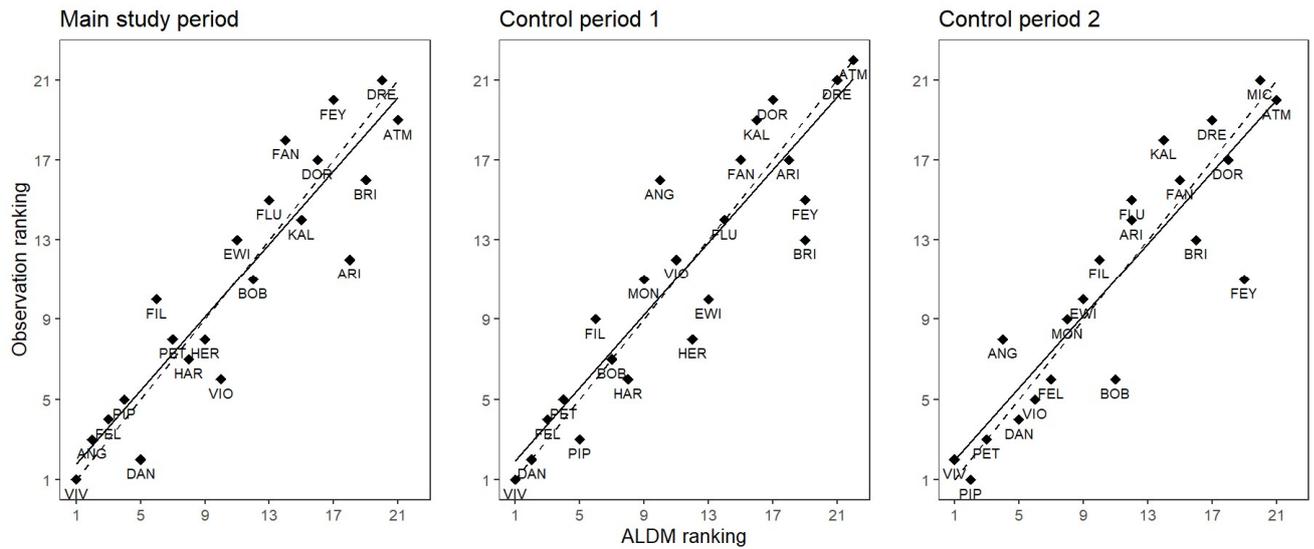
Figure 1: Correlation between the dominance ranks of each individual (labelled by the 3 first letters of their names) obtained from behavioural observations (y-axis) and from computerised ALDM testing (x-axis) for the three study periods. Bold lines represent the linear regression of the data and the dotted line a perfect positive (x=y) correlation. (Main study period: Kendall correlation: $\tau$ = 0.752, CI = [0.59; 0.89] and Spearman correlation: $r_s$ = 0.91, CI = [0.76; 0.97], Control period-1: Kendall correlation: $\tau_{c1}$ = 0.7665, CI = [0.57; 0.91] and Spearman correlation: $r_{sc1}$ = 0.9, CI = [0.71; 0.97], Control period-2: Kendall correlation: $\tau_{c2}$ = 0.760, CI = [0.57; 0.91] and Spearman correlation: $r_{sc2}$ = 0.9, CI = [0.70; 0.98]
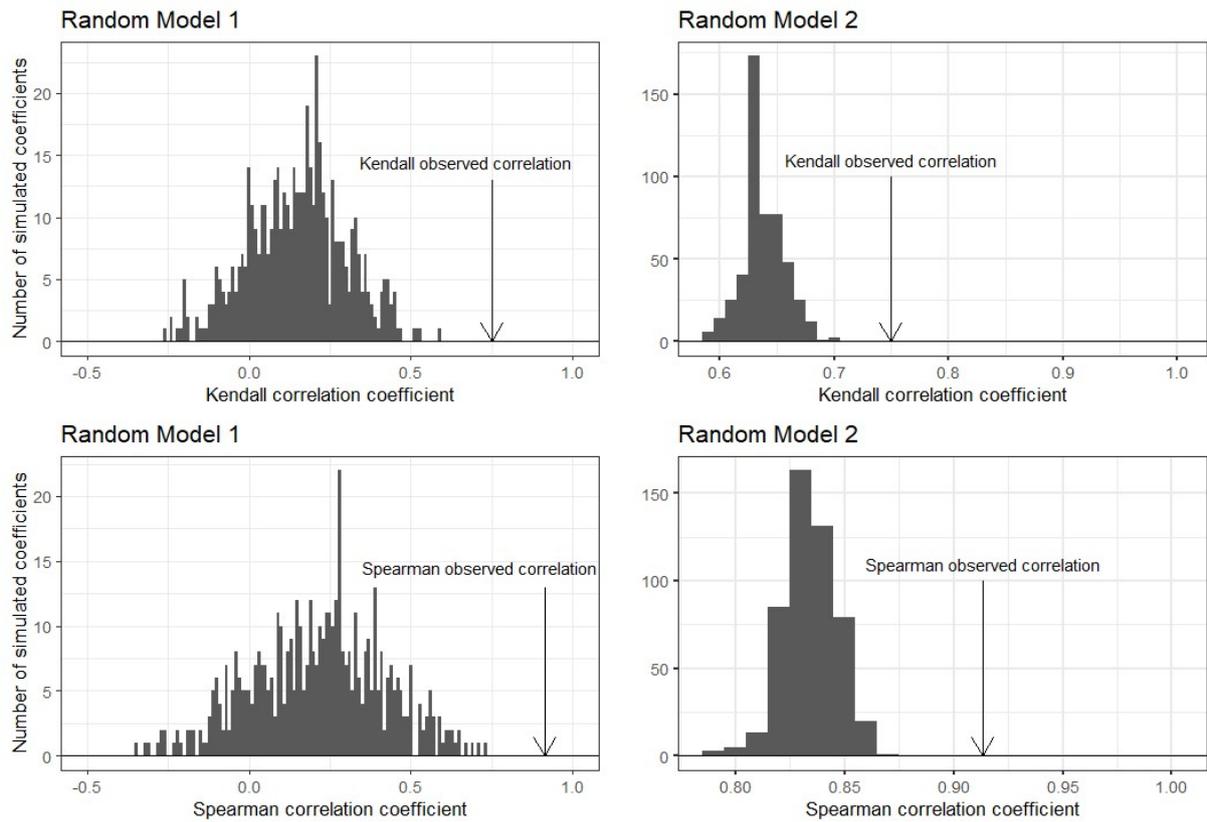
Figure 2. Distribution of the simulated Kendall and Spearman correlation coefficients between dominance hierarchies generated from the random models 1 (left figures) and 2 (right figures) and the dominance hierarchy from behavioural observations. The arrows indicate the observed correlation between the dominance hierarchy obtained from untransformed ALDM data and the one established from behavioural observations (Kendall correlation: $\tau$ = 0.752, CI = [0.59; 0.89] and Spearman correlation: $r_s$ = 0.91, CI = [0.76; 0.97].

738 Table 1: Results of the linear mixed effects model described in the text.

| Random effects: | | | | |
|---|---|---|---|---|
| Groups: | Name | Variance | SD | Corr |
| Name | (Intercept) | 620.4 | 24.91 | |
| | Day | 1523.5 | 39.03 | -0.93 |
| Age | (Intercept) | 10395.8 | 101.96 | |
| Residual | | 31242.1 | 176.75 | |

| Fixed effets: | | | | | |
|---|---|---|---|---|---|
| | Estimate | SE | df | t-value | p-value |
| Intercept | 562.00 | 23.61 | 1.92E+01 | 23.80 | <0.001 |
| Experimental condition: Predictive | -54.22 | 3.32 | 9.96E+04 | -16.31 | <0.001 |
| Ntrial | -1.36 | 0.63 | 9.96E+04 | -2.14 | 0.03 |
| Behavioural condition: Supplant | -19.73 | 3.52 | 9.96E+04 | -5.60 | <0.001 |
| Beh. cond. by Exp. cond. | 8.45 | 4.93 | 9.96E+04 | 1.71 | 0.09 |
| Beh. cond. by Ntrial | 1.70 | 0.94 | 9.96E+04 | 1.82 | 0.07 |
| Exp. Cond. by Ntrial | 0.29 | 0.89 | 9.96E+04 | 0.32 | 0.75 |
| Beh. Cond. by Exp. Cond. by Ntrial | -0.07 | 1.32 | 9.96E+04 | -0.05 | 0.96 |

739 Note. Number of observations: 99619, number of inidividuals:20

740

741     Table 2: Results of the linear mixed effects model described in the text.

**Random effects:**

| Groups: | Name | Variance | SD | Corr |
|---|---|---|---|---|
| Name | (Intercept) | 3402.9 | 58.334 | |
| | Day | 15.5 | 3.937 | 0.17 |
| Age | (Intercept) | 4966.9 | 70.476 | |
| Residual | | 30685.2 | 175.172 | |

742

**Fixed effects:**

| | Estimate | SE | df | t-value | p-value |
|---|---|---|---|---|---|
| (Intercept) | 633.1636 | 22.5610 | 18.5753 | 28.064 | <0,001 |
| DiffRank | 1.7580 | 0.5639 | 4277.7434 | 3.118 | 0.00183 |

743

744     Note. Number of observations: 4411, nmber of individuals: 20

745

746

747

748 Supplementary table 1: Identity and age of the participants for each study.

749

| Name | Year of birth | Sex | Experiment 1 | | | Experiment 2 |
| | | | Main study period | Control 1 | Control 2 | Fagot & Goujon (2013)* |
| | | | (N=21) | (N=22) | (N=21) | (N=20) |
|---|---|---|---|---|---|---|
| KALI | 1995 | F | + | + | + | + |
| BRIGITTE | 1996 | F | + | + | + | + |
| MICHELLE | 1996 | F | | | + | + |
| MONA | 1997 | F | | + | + | + |
| ATMOSPHERE | 1998 | F | + | + | + | + |
| PETOULETTE | 1999 | F | + | + | + | + |
| PIPO | 1999 | M | + | + | + | + |
| URANIE | 2003 | F | | | | + |
| VIVIEN | 2004 | M | + | + | + | + |
| VIOLETTE | 2004 | F | + | + | + | + |
| ANGELE | 2005 | F | + | + | + | + |
| ARIELLE | 2005 | F | + | + | + | + |
| BOBO | 2006 | M | + | + | + | + |
| DAN | 2008 | M | + | + | + | + |
| DREAM | 2008 | F | + | + | + | + |
| DORA | 2008 | F | + | + | + | + |
| EWINE | 2009 | F | + | + | + | + |
| FANA | 2010 | F | + | + | + | + |
| FELIPE | 2010 | M | + | + | + | + |
| FEYA | 2010 | F | + | + | + | + |
| FLUTE | 2010 | F | + | + | + | |
| FILO | 2010 | M | + | + | + | |
| HERMINE | 2012 | F | + | + | | |
| HARLEM | 2012 | M | + | + | | |

750

751 Note. (+) indicates if the subject is present in the considered study period (*) data in Goujon &

752 Fagot (2013) were collected in 2012.

753

754 Supplementary table 2: Ethogram with the name and definitions of the agonistic behaviours

755 coded during the observations.

| Behaviour | Description/example |
| --- | --- |
| Aggression | Hitting, biting, pinching, scratching, slapping another individual |
| Chase | Pursuing another individual (screaming, …) during a conflict before or without final aggression |
| Threat | Threatening another baboon raising its eyebrows with or without screaming and slapping on the ground |
| Supplanting behaviour | One individual is getting the place of another one, by pushing it |
| Avoid | Avoiding another individual in his approach (recordable at 2m max between both individuals) |
| Yak | Yakking behaviour (without avoidance) use to happen in a fear/surprise context or anticipating a conflict |

756

757

758

759

760

761

762

763    Supplementary table 3: Elo-scores and respective ranks of each individual in the three periods,
764    and for both ALDM and Observation data.

765

| ID | Main study period | | Control period-1 | | Control period-2 | |
|---|---|---|---|---|---|---|
| | Observation | ALDM | Observation | ALDM | Observation | ALDM |
| ANGELE | 1279 (3) | 1232 (2) | 911 (16) | 1058 (10) | 1101 (8) | 1104 (4) |
| ARIELLE | 1009 (12) | 824 (18) | 907 (17) | 810 (18) | 931 (14) | 961 (12) |
| ATMOSPHERE | 630 (19) | 336 (21) | 744 (22) | 514 (22) | 720 (20) | 550 (21) |
| BOBO | 1034 (11) | 959 (12) | 1015 (7) | 1087 (7) | 1129 (6) | 1014 (11) |
| BRIGITTE | 847 (16) | 716 (19) | 950 (13) | 778 (19) | 938 (13) | 895 (16) |
| DAN | 1294 (2) | 1174 (5) | 1156 (2) | 1200 (2) | 1162 (4) | 1079 (5) |
| DORA | 790 (17) | 845 (16) | 819 (20) | 822 (17) | 782 (17) | 870 (18) |
| DREAM | 509 (21) | 557 (20) | 772 (21) | 729 (21) | 749 (19) | 874 (17) |
| EWINE | 998 (13) | 968 (11) | 983 (10) | 995 (13) | 1025 (10) | 1024 (9) |
| FANA | 786 (18) | 944 (14) | 907 (17) | 950 (15) | 874 (16) | 902 (15) |
| FELIPE | 1155 (4) | 1226 (3) | 1061 (4) | 1165 (3) | 1129 (6) | 1033 (7) |
| FEYA | 611 (20) | 842 (17) | 937 (15) | 778 (19) | 983 (11) | 764 (19) |
| FILO | 1068 (10) | 1159 (6) | 1005 (9) | 1101 (6) | 971 (12) | 1017 (10) |
| FLUTE | 853 (15) | 957 (13) | 943 (14) | 966 (14) | 896 (15) | 961 (12) |
| HARLEM | 1073 (7) | 1127 (8) | 1055 (6) | 1075 (8) | NA | NA |
| HERMINE | 1070 (8) | 1118 (9) | 1011 (8) | 1007 (12) | NA | NA |
| KALI | 890 (14) | 938 (15) | 896 (19) | 888 (16) | 762 (18) | 951 (14) |
| MICHELLE | NA | NA | NA | NA | 695 (21) | 723 (20) |
| MONA | NA | NA | 969 (11) | 1060 (9) | 1097 (9) | 1028 (8) |
| PETOULETTE | 1070 (8) | 1137 (7) | 1060 (5) | 1163 (4) | 1186 (3) | 1163 (3) |
| PIPO | 1154 (5) | 1179 (4) | 1141 (3) | 1137 (5) | 1297 (1) | 1196 (2) |
| VIOLETTE | 1120 (6) | 1033 (10) | 960 (12) | 1029 (11) | 1139 (5) | 1050 (6) |
| VIVIEN | 1461 (1) | 1412 (1) | 1213 (1) | 1346 (1) | 1275 (2) | 1206 (1) |

766

767    Note. NA = Subject non-available for testing during the considered period, for instance due to
768    sickness. Ranks are indicated in parentheses.

769

770 Supplementary table 4: Hierarchical indexes through the three periods for both ALDM and
771 observation data.

772

| Index | | Main study period | | Control period-1 | | Control period-2 | | Exp2 |
|---|---|---|---|---|---|---|---|---|
| | | Observation | ALDM | Observation | ALDM | Observation | ALDM | ALDM |
| | N | 21 | 21 | 22 | 22 | 21 | 21 | 20 |
| Optimal K | optK | 101 | 18 | 34 | 12 | 116 | 10 | 10 |
| Triangle transitivity | Pt | 0.96 | 0.96 | 0.85 | 0.95 | 0.97 | 0.94 | 0.86 |
| | ttri | 0.85 | 0.83 | 0.39 | 0.80 | 0.88 | 0.77 | 0.45 |
| Linearity | h | 0.57 | 0.67 | 0.16 | 0.66 | 0.24 | 0.69 | 0.47 |
| | h' | 0.59 | 0.68 | 0.21 | 0.67 | 0.30 | 0.69 | 0.47 |
| Repeatability | r | 0.92 | 0.96 | 0.88 | 0.97 | 0.94 | 0.95 | 0.91 |

773

774 Note. This table shows OptK: the optimal K used in Elorating procedure of dominance

775 hierarchy assessment, Triangle transitivity (Pt: the proportion of triangles that are transitive,

776 and ttri: this proportion scaled so that it is 0 for the random expectation and 1 for maximum

777 transitivity) and Linearity indices (de Vries linearity index h (Appleby, 1983; de Vries, 1995)

778 and Landau corrected index h'(de Vries, 1998)) from Elorating package and r: the repeatability

779 index (Shizuka & McDonald, 2012) from AniDom package for 1000 randomisations.

780

781 Supplementary references:

782 Appleby, M. C. (1983). The probability of linearity in hierarchies. *Animal Behaviour, 31*, 600-

783 608.

784 De Vries, H. (1995). An improved test of linearity in dominance hierarchies containing

785 unknown or tied relationships. *Animal Behaviour, 50*, 1375-1389.

786 Shizuka, D., & McDonald, D. B. (2012). A social network perspective on measurements of

787 dominance hierarchies. *Animal Behaviour, 83*, 925-934.

788