

Supporting Information

Janmaat et al. 10.1073/pnas.1407524111

Statistical Analyses

Data and models were analyzed using Mann–Whitney U tests, a Spearman rank correlation test, and (Generalized) Linear Mixed Models (GLMM and LMM) (1, 2) in R (v3.0.3) (3) using the functions `wilcox.exact` (R package `exactRankTests`) (4), `cor.test`, `lmer` (R package `lme4`) (5), and `glmmadmb` (R package `glmmADMB`) (6). For the departure time and distance models with Gaussian error structure and identity link, we checked whether the assumptions of normally distributed and homogeneous residuals were fulfilled by visually inspecting QQ plots and the residuals plotted against fitted values. To obtain normally distributed and homogeneous residuals for the distance model, we \log_e -transformed the response distance from nest to breakfast site. After this procedure, the plots did not indicate any obvious deviations from these assumptions any longer. Because the difference in late evening and early morning travel direction was bound between 0° and 180° , we scaled the observed angles to lie between 0 and 1 and modeled them using a generalized linear mixed regression model (GLMM) with β -error structure and logit link (6). To check the overall significance of the combined set of predictor variables, we used likelihood ratio tests (7) to compare each full model with a respective null model containing only the random intercepts, random slopes, and control predictors (Tables S1–S4). We only considered the effect of individual predictors if the initial full models reached significance (8). To create stable models, we transformed the predictors in such a way that they had a roughly symmetric distribution, before running the models. To accomplish this, we \log_e -transformed distance from last feeding to nest site, distance from last feeding to breakfast site, the number of adult males and fruit size, and took the fourth root of the relative energy balance at departure time and feeding duration at breakfast site. All predictor variables were then z -transformed to obtain comparable estimates. We checked for collinearity by inspecting variance inflation factors (VIF) derived from multiple regression models with the random effects excluded using the function `vif` of the R package `car` (9). This process did not indicate collinearity problems in any of the models with the highest VIF value being 1.22. In addition, we plotted the factor fruit genus against fruit size and found no indication for any correlation and correlation coefficients between pairs of covariates were low (all $r < 0.24$). We assessed each model's validity by comparing the estimates derived from a model based on all data with those obtained from a model with data points, individuals, species, and tree identity dropped one by one, respectively, which indicated that all models were stable and that no influential cases existed. Because of a recent discussion about the validity of P values of fixed effects in the framework of LMM and GLMMs, we provide P values for the individual effects derived from likelihood ratio tests comparing the full with the respective null models (10). All tests were two-tailed.

Detailed Model Descriptions, Additional Results, and Interpretations

Departure Time Model. In addition to the results of the departure-time model described in the main text, we also found that the females departed later when it had been raining during the previous night (Table 1 and Fig. S1). Low body temperatures, because of the evaporation of wet fur, are unlikely to explain later departure times because no significant effect of night temperature was found (Table 1). The effect could be explained by females decreasing predation risk because leopards are proposed to hunt

more successfully in rainy periods when the forest floor is wet and prey can be approached more quietly (11). All five females had small offspring that were vulnerable to predation risk (12, 13). In fact, the male offspring of one of the target females (Lou) was attacked by a leopard and died of the consequences 1 y before our study.

Finally, feeding duration at the breakfast site and the number of adult males at the nest site had a significant interactive effect on departure time (Table 1 and Fig. S4). The direction of the interactive effect described in Fig. S4 does not allow a simple explanation. When the estimated breakfast food amount (i.e., feeding duration) was low, departure time decreased with an increasing number of males at the nest site. This finding suggests a negative effect of intragroup competition and a positive effect of protection from predators on departure time. [Males were observed to aggressively mob leopards, which are known to patrol fruiting trees (14), and male drumming, which occurs more frequently in the early morning, has been observed to deter leopards (11, 12, 15).] However, when the estimated food amount (i.e., feeding duration) was high and the number of males was low, the departure times were even earlier. To explain this complex result, we propose that the effect of feeding duration on departure time is socially confounded (i.e., influenced by behaviors of other group members). Because feeding duration was weakly correlated with the number of adult males ($r_{\text{Spearman exact}} = -0.15$, $n = 179$, $P = 0.0447$), but not with the total number of independent individuals in the party ($r_{\text{Spearman exact}} = -0.12$, $n = 179$, $P = 0.1134$), we propose that feeding duration was not only shortened by the amount of food in the breakfast tree, yet also by a large number of males, either through contest competition or different foraging and ranging interests of males and females. This explanation should be considered in future studies that try to estimate fruit amounts using proxies, such as feeding duration, when the amount of edible fruits are difficult to assess by visual cues (e.g., using color).

It is unlikely that feeding duration simply represented hunger levels as relative energy balance, which was based on the energy expenditure and intake of the day and night before departure, was controlled for and did not have a significant effect (Table 1).

The data in the departure-time model included repeated observations of individuals, necessitating a random-effects structure (2) to avoid pseudoreplication. In addition, to individual random effects, fruit species and tree individual were modeled with group level variation. Random slopes (of all predictors) were initially not included because the model became too complex and did not converge when doing so. However, when we excluded data from subsequent visits to a previously visited breakfast site and only analyzed the first visits, we were able to include all of the random slopes (because the model without the random effect breakfast site was less complex) and we found effects similar to the original model (Table S1).

Are earlier departure times the result of earlier waking? For part of the observation period, we also collected the time of first urination and used this as a proxy for wake-up time as we were unable to see when the target females had opened their eyes when they were high up in the nest (mean height \pm SD: 18.6 ± 5.8 m, range: 2–35 m). It is possible that departure time reflected wake-up time as departure time and first urine times were correlated ($r_{\text{Spearman}} = 0.81$, $n = 87$, $P < 0.0001$). However, the variation in time duration between the estimated wake-up and departure times is remarkable, varying between 0.03 and 53.61 min (median = 6.22 min, mean \pm SD = 10.57 ± 11.55 ; $n = 87$, excluding

the three observations of urination right after departure, $N_{\text{females}} = 3$). Whether this variation was influenced by an inhibition of hunger or fear remains elusive.

Positioning Models.

Proximity model. To investigate if chimpanzees were able to minimize predation risk at dawn by building their nest closer to popular breakfast sites compared with other sites, we designed a proximity model. Because variation in observed distances from nest to subsequent breakfast sites was substantial (mean \pm SD: 100.6 ± 144.9 m, range = 2–918 m), we investigated whether the chimpanzees had the possibility (with regard to the availability of sufficient suitable nest sites and preferences from other party members) to build their nest closer to the breakfast site when it contained figs or small fruit but also when the breakfast site was the same as the last site at which they fed before making their nest (dinner site). Hence, we included fruit genus type, size, and the categorical predictor “dinner is breakfast site” (yes or no) as main predictors in the model. Because large distances between dinner and subsequent breakfast sites could potentially restrict the females from making their nest at close distance to the breakfast site, we also included the distance from the last feeding site to the nest (Fig. 3A) and the duration between the end of last feeding and sunset time as control predictors in the model. To investigate the causal direction of the data, we also included energy balance and the number of adult males detected at the nest location to control for the possibility that distances between the nest and breakfast site were simply shorter, because (i) the females needed to feed and get energy more quickly and thus fed closer at dawn or (ii) a larger number of adult males had enabled them to reach a suitable nest at closer distance to the breakfast site through prolonged travel at dusk.

We found that the females made their evening nest closer to the breakfast site when they were going to have breakfast at the same site as where they had their last dinner (Fig. S2 and Table S3). This finding is consistent with the possibility that the females knew that they would eat at the same place the next day. However, a more parsimonious explanation is that they simply made their nest close to the last dinner site as was reported in other studies on great apes (16, 17), and breakfasted at the same close site because it provided salient cues. The effect of breakfast fruit genus and size was not significant as reported in the main article. The random effect of species was, however, significant ($\chi^2_1 = 19.05$, $P < 0.0001$). The effect of the control predictors relative energy balance, the number of adult males at the nest site and the duration between the end of last feeding and sunset were all not significant (Table S3).

Interestingly, when the breakfast was different from the last dinner site, the distance that the females traveled from the last dinner to the nest site was larger when the distance from the nest till the breakfast site (and thus the overall distance between last dinner and breakfast site) was larger (Fig. S2). This finding seems unlikely to be a result of the number of males present in the party, where a larger number could have triggered long travel distances in general (18) as no effect of the number of males were found on the distance traveled from nest to breakfast. The nonrandom pattern of distance distribution are therefore consistent with the possibility that the females were considering the location of next day's breakfast when leaving the last feeding site the day before, and traveled further when the overall distance to reach breakfast was larger. However, it is difficult to exclude the alternative explanation that the chimpanzees have a habit of searching for and feeding on trees that belong to the same species within fruiting seasons (use of fruiting synchrony) (19). If some of these species are simply more clustered or common than others, it will make it more likely that both late evening and early morning distances are small, and visa versa.

Direction model. To further investigate if chimpanzees planned to minimize predation risk at dawn by building their nest en route to

sought-after breakfast sites, we designed a direction model to describe the change in direction between the line connecting prior dinner with the night nest, and the line connecting the night nest to the breakfast site (Fig. 3A). The direction model only included first visits to breakfast sites and species for the following theoretical and practical reasons. First, by excluding repeated observations we could exclude the random effects and respective slopes for tree individual and tree species, which improved model stability. Second, by only considering first visits it is more difficult to explain results by learned multiple time–place associations (e.g., associations between nest positioning en route and high rewards at highly sought-after fruit versus nest positioning off route and low rewards), especially when we consider that feeding durations at more sought-after fruit like ripe figs are highly ephemeral (mean revisit frequency and interval is 1.9 and 4.5 d). In addition, we excluded all mornings for which the breakfast site was the same as the last dinner site to calculate relevant directions. To avoid pseudoreplication, we included a random intercept term for chimpanzee individual and the random slopes for all predictors. The results of this model were discussed in the main text.

Duration of Ripe Fruit Presence Calculations. To calculate the duration of ripe fruit presence for each plant species, we used long-term phenology data from individual trees growing in the research area collected at monthly intervals. Methods for tree selection and data collection are described in ref. 20, but notably the data (i) are of species representing the most common fruit food species for the Taï chimpanzees, and (ii) extend from February 1997 until November 2011, with many individuals being checked for multiple years. Because data collection was sometimes interrupted (e.g., because of observer illness or civil unrest), and because of the potential for observation error, simply computing the mean duration of time that fruit was observed from all distinct times of fruit presence was not robust (and would lead to downward bias and inflated variability in estimates). To avoid the problem of bias, we used a hidden Markov model (HMM) (21) to compute expected time durations of ripe fruit presence for individual trees, and excluded individuals for which less than 60 mo of data were available (total number of phenology trees = 1,477 from 82 species). HMMs allow for missing data, whereas the Markov property provides that the value of the latent state at time t is probabilistically related only to the previous latent state value at time $t - 1$ through the transition probability matrix. For our HMMs, the number of possible states is two, corresponding to ripe fruit either being present or absent. We set the latent fruit production states as the observed values with probability 1. The observation probability distribution for each observed state was given by the binomial distribution. Each estimated state transition matrix of the HMM models give the individual expected duration of ripe fruit presence as the sojourn times defined by the inverse of the elements of its diagonal entries by which individual expected duration of presence times were computed. We fitted the HMM models using the `hmmfit` function from the `mhmm` package (22) in the R v2.12.2 programming environment (3). The parameter optimization routine was initiated at multiple starting locations for each model to ensure convergence to the parameter values at the global maximum likelihood.

Relative Energy Balance Calculations. To control for the potential effect of energy balance at the moment of nest departure, we calculated the “relative energy balance” using the methods described in ref. 23. We first initiated calculations of each relative energy balance at zero at the moment when the target female left her nest the previous morning. We then added energy intake and subtracted energy expenditure until the moment the female departed her nest again on the subsequent day—the day of departure. As a slight improvement of the methodology of ref. 23, we calculated energy expenditure using velocity values that were

measured by following the respective traveling females with a Garmin 60Csx GPS. In addition, we used the climbing times of the respective female on the respective day instead of an estimation based on tree height. Furthermore, we used individual tree and female-specific fruit-intake rates. For most food items fed on, we used the average energy values calculated by A. N'Guessan et al. (23). For food items for which such information was missing, we used the energy values known for food items of similar type (e.g., fruit or young leaves) and of similar size and family. We realize our calculations resulted in rough estimations of energy balance, for example the oxygen intake per meter traveling, needed to calculate energy expenditure was based on only one captive chimpanzee walking on a treadmill (23). However, because we used the same estimates for each day and only included days that were preceded by complete follow days (days for which we did not lose the females and had complete knowledge of their food intake and traversed travel distances), these data allow accurate estimations of their relative energy balance at the moment of nest departure.

Nest Preference Calculations. To determine whether the target females preferred to build their nest in particular tree species, we used data from the two females for which sufficient data on nest tree species were collected (Isha: $n = 103$, Julia: $n = 54$). We ran two separate χ^2 tests for given probabilities, whereby P values were determined by simulation (based on 2,000 replicates) in

R (v3.0.3.), using the function “chisq.test” with the argument “simulate.p.value” set to TRUE. We found that both females made their nest more frequently in particular tree species than expected from the tree density distribution of the respective species (24) (Isha: $P = 0.0005$; Julia: $P = 0.0005$). In addition, we found that some tree species were preferred whereas others were avoided (Fig. S6). To attain preference scores per species, we calculated for each species a selection index defined as the proportion of trees of the respective species that were used to build a nest out of the total number of trees nested in, divided by the proportion of trees of the respective species that were encountered in a tree density transect of the total number of trees encountered in this transect (24, 25). A selection index of 1 indicates that the chimpanzee selected trees of the respective species purely based on the encounter rate of trees of that same species. To achieve symmetrical scales in the figure for positive and negative preference scores, we took the \log_e of the selection index (and added the value 1 to the observed frequencies as some species were never nested in). On the log scale, a value of 0 means that the proportion chosen is the expected proportion based on the tree species density, a value <0 means “avoidance” and a value >0 means a “preference” (Fig. S6). To quantify which preferences and avoidances were significant, we used a binomial test for each species’ preference scores.

- Bolker BM, et al. (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol* 24(3):127–135.
- Baayen RH (2008) *Analyzing Linguistic Data* (Cambridge Univ Press, Cambridge, UK).
- R Development Core Team (2014) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna).
- Hothorn T, Hornik K (2014) *ExactRankTests: Exact Distributions for Rank and Permutation Tests*. R Package Version 0.8-27. Available at <http://CRAN.R-project.org/package=exactRankTests>.
- Bates D, Maechler M, Bolker B (2014) *Lme4: Linear Mixed-effects Models using Eigen and Eigenfaces*. R Package Version 1.1-5. Available at <http://CRAN.R-project.org/package=lme4>.
- Fournier DA, et al. (2012) AD Model Builder: Using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim Methods Softw* 27(2):233–249.
- Dobson AJ (2002) *An Introduction to Generalized Linear Models* (Chapman & Hall, Boca Raton, FL).
- Schielzeth H, Forstmeier W (2009) Conclusions beyond support: Overconfident estimates in mixed models. *Behav Ecol* 20(2):416–420.
- Fox J, Weisberg S (2011) *An R Companion to Applied Regression*. Second Edition. (Sage, Thousand Oaks, CA).
- Barr DJ, Levy R, Scheepers C, Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang* 68(3):255–278.
- Jenny D, Zuberbühler K (2005) Hunting behavior in West African forest leopards. *Afr J Ecol* 43(3):197–200.
- Boesch C (1991) The effects of leopard predation on grouping patterns in forest chimpanzees. *Behavior* 117:220–242.
- Klailova M, et al. (2012) Non-human predator interactions with wild great apes in Africa and the use of camera traps to study their dynamics. *Folia Primatol (Basel)* 83(3-6):312–328.
- Hart JA, Katembo M, Punga K (1996) Diet, prey selection and ecological relations of leopard and golden cat in the Ituri Forest, Zaire. *Afr J Ecol* 34:364–379.
- Campbell G (2013) Applying the IUCN SSC A.P.E.S. (Ape Populations, Environments, and Surveys) database to conservation-oriented research. PhD thesis (University of Leipzig, Leipzig, Germany).
- MacKinnon J (1974) The behavior and ecology of wild orangutans (*Pongo pygmaeus*). *Anim Behav* 22(1):3–74.
- Goodall JM (1962) Nest building behavior in the free ranging chimpanzee. *Ann N Y Acad Sci* 102(2):455–467.
- Bates LA, Byrne RW (2009) Sex differences in the movement patterns of free-ranging chimpanzees (*Pan troglodytes schweinfurthii*): Foraging and border checking. *Behav Ecol Sociobiol* 64(2):247–255.
- Janmaat KRL, Ban SD, Boesch C (2013b) Tai chimpanzees use botanical skills to discover fruit: What we can learn from their mistakes. *Anim Cogn* 16(6):851–860.
- Anderson DP, et al. (2005) Factors influencing tree phenology in Tai National Park, Côte d'Ivoire. *Biotropica* 37(4):631–640.
- Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. *Proc IEEE* 77(2):257–286.
- O'Connell J, Hojsgaard S (2011) Hidden semi Markov models for multiple observation sequences: The mhsmm package for R. *J Stat Softw* 39:1–22.
- N'Guessan AK, Ortmann S, Boesch C (2009) Daily energy balance and protein gain among *Pan troglodytes verus* in the Tai National Park, Côte d'Ivoire. *Int J Primatol* 30(3):481–496.
- Goné Bi ZB (2007) Régime alimentaire des chimpanzés, distribution spatiale et phénologie des plantes dont les fruits sont consommés par les chimpanzés du Parc National de Tai, en Côte d'Ivoire. PhD thesis, (University of Cocody, Abidjan, Ivory Coast).
- Manly BFJ, McDonald LC, Thomas DL (1993) *Resource Selection by Animals: Statistical Design and Analysis for Field Studies* (Chapman & Hall, London).

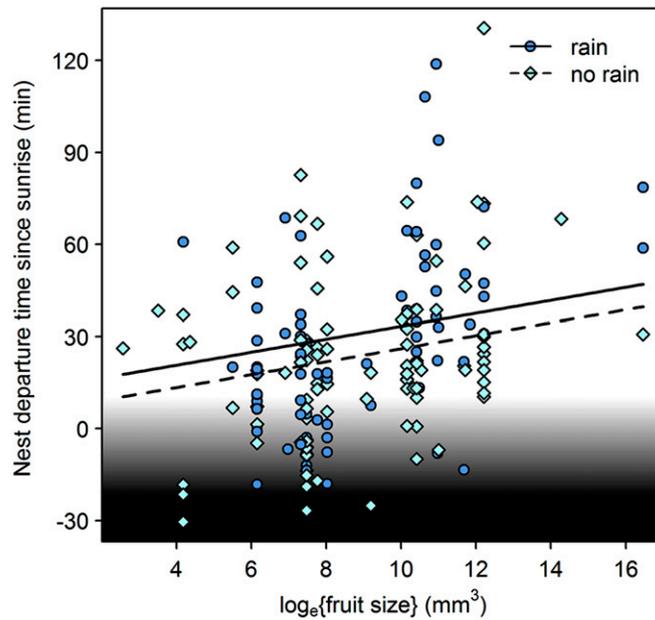


Fig. S1. The times at which female chimpanzees departed from their sleeping nest to feed on breakfast fruit of different size. Zero on the y axis represents sunrise and the background shading reflects time since the beginning of the average dawn start time (black) to 10 min after sunrise (white). The datapoints ($n = 179$) show observed nest departure times during rainy (dark blue circles) or dry (light blue diamonds) nights and lines illustrate expected departure times as predicted by the model (conditional on all other predictors being on their average and based on fruit genus manually dummy coded and centered). Both \log_e -transformed species-specific fruit sizes and whether or not it rained at night had a significant impact on departure times (Table 1).

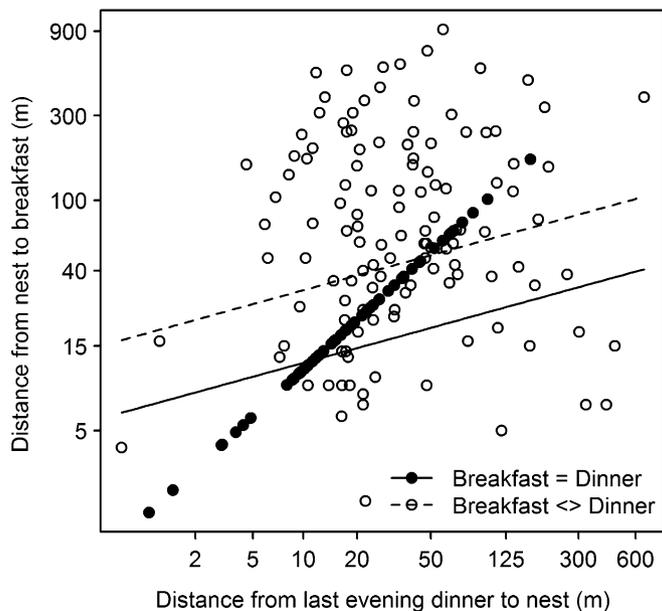


Fig. S2. The distances between the nest and the breakfast site for sites that were the same and different from the last dinner site. The y axis represents the distance from the sleeping nest to the breakfast site. The datapoints ($n = 180$) show observed distances when the last feeding site (dinner) was the same as the breakfast site (filled circles) or not (open circles) and lines illustrate expected distances as predicted by the model when dinner was breakfast (full) or not (dotted; conditional on all other predictors being on their average and based on fruit genus manually dummy coded and centered). Both \log_e -transformed distance from the last feeding to the nest site and whether or not last feeding was the same as breakfast influenced early morning travel distance to breakfast (Table S3).

Table S2. Summary of the LMM explaining the starting times of breakfasting

Predictors	Estimate	SE	P value
(Intercept)	1406.39	324.32	
Target predictors			
Breakfast fruit size (mm ³)	323.62	150.96	0.0384
Breakfast genus type (no fig)	887.29	365.01	
Distance nest to breakfast site	209.42	223.04	
Breakfast genus type × Distance nest to breakfast site	567.72	264.72	0.0393
Control predictors			
No. of adult males at nest	-169.04	124.93	
Feeding duration at breakfast site	-410.5	112.74	
Feeding duration at breakfast site × No. of adult males at nest	172.42	118.07	0.1533
Relative energy balance	-48.46	115.72	0.6768
Night temperature	-107.92	115.3	0.3651
Rain at nest (yes)	310.85	230.8	0.1793

Full versus control model comparison: $\chi^2_4 = 37.33$, $N_{\text{females}} = 5$, $N_{\text{days}} = 179$, $P < 0.0001$.

Table S3. Summary of the LMM explaining the distance between sleeping nest and early morning breakfast site

Predictors	Estimate	SE	P value
(Intercept)	4.25	0.26	
Target predictors			
Breakfast species (no fig)	-0.23	0.31	0.4654
Fruit volume (mm ³)	0.11	0.12	0.3586
Last dinner is breakfast (yes)	-0.94	0.17	0.0005
Control predictors			
Distance last feeding to nest	0.30	0.08	0.0002
Relative energy balance	0.09	0.08	0.2652
No. Adult males at nest site	-0.01	0.09	0.9574
Duration from last feeding till sunset	0.07	0.08	0.3452

Full versus control model comparison: $\chi^2_3 = 36.6$, $N_{\text{females}} = 5$, $N_{\text{days}} = 180$, $P < 0.0001$.

Table S4. Summary of the GLMM explaining the difference between late evening and early morning travel direction

Predictors	Estimate	SE	Z value	P value
Intercept	-0.332	0.271	-1.23	
Target predictors				
Breakfast genus type (no fig)	0.842	0.337	2.5	0.012
Breakfast fruit size (mm ³)	0.169	0.178	0.95	0.343
Control predictor				
Travel distance from last dinner to breakfast	-0.212	0.149	-1.43	0.154

Full versus control model comparison: $\chi^2_2 = 7.67$, $N_{\text{females}} = 5$, $N_{\text{days}} = 43$, $P = 0.0216$.