Supplementary Methods and Analyses:
Distinct developmental trajectories for risky and impulsive decision-making in chimpanzees
Alexandra G. Rosati, Melissa Emery Thompson, Rebeca Atencia, & Joshua W. Buckholtz

Supplemental Methods

Subjects

We tested 40 semi-free ranging chimpanzees (21 males, 19 females) living in Tchimpounga Chimpanzee Sanctuary in Republic of Congo. All care practices comply with the Pan-African Sanctuary Alliance (PASA) guidelines. Apes in PASA sanctuaries are typically born in the wild and enter the sanctuary after being confiscated at an early age (~2-3 years old) from the bushmeat and pet trade. All apes were socially housed, and the majority free-ranged in large tracts of tropical forest during the day (5-40 hectares across groups). In the evening, apes enter indoor dormitories (12 m^2-160 m^2), and were tested the day in these familiar dormitory buildings. Following testing, most apes were released back to their social groups outside. Apes had *ad libitum* access to water and were never food-deprived for testing. In addition to the food in the forest, they were fed a variety of fruits, vegetables, and species-appropriate foods. All tests were voluntary: if subjects stopped participating, the test was stopped. Previous work indicates that sanctuary apes in these contexts are physically and psychologically healthy relative to other captive populations (Cole et al., 2020; Rosati et al., 2013; Wobber & Hare, 2011).

Methodological overview

Apes completed a decision-making battery in 2012 comprising a total of eleven tasks administered across seven days of testing, with all individuals completing the tasks in the same order across days (see Cantwell, Buckholtz, Atencia, & Rosati, 2022; Rosati, DiNicola, & Buckholtz, 2018 for additional details). The order of the tasks across days was designed to ensure that each day of testing took approximately 30 minutes per subject, such that shorter tasks were grouped together on the same day. The current paper focuses on a subset of two tasks examining value-based decision, each completed on a different day of the battery (inter-temporal choice task on day two, and the risky choice task on day three). The remaining tasks addressed aspects of cognition, such as cooperation, that were unrelated to the current paper's questions and hypotheses. Interleaved with an individual's period of cognitive testing during the battery, chimpanzees also completed voluntary saliva sampling period to index hormones (cortisol and testosterone), as described below.

Cognitive task methods

The risky task used methods reported in prior work (Rosati & Hare, 2011, 2012, 2013; Rosati & Hare, 2016). Figure S1 provides a detailed diagram and photographs of the setup. The inter-temporal choice task also used methods reported in prior work (Rosati, et al., 2018; Rosati & Hare, 2013; Warneken & Rosati, 2015). Figure S2 provides a detailed diagram and photograph of the setup.

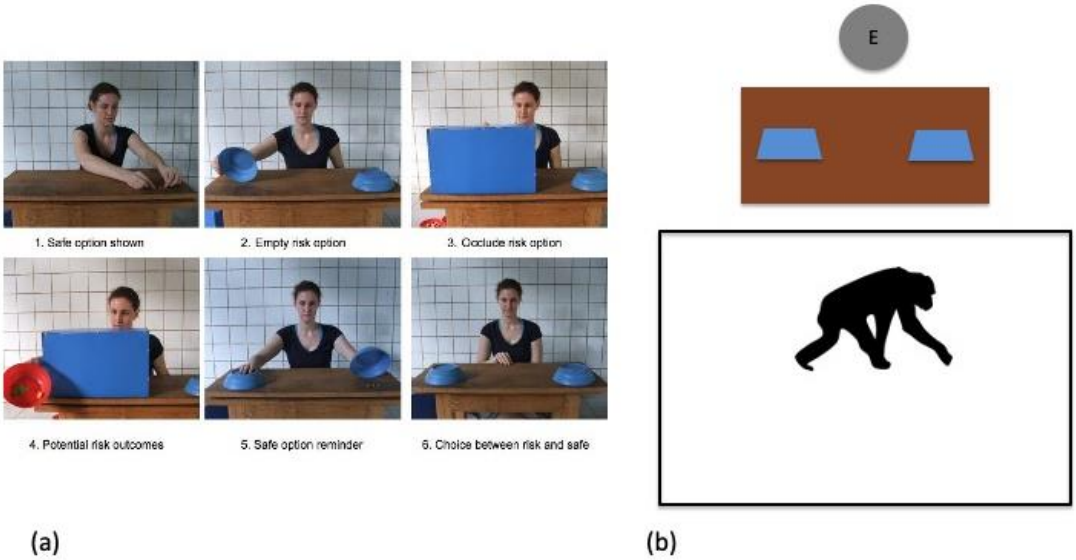

Figure S1: Setup for risky choice task. (a) Photo of setup (from chimpanzee's point of view): E sat across from the chimpanzee at a table and (1) visibly baited the safe option visible with peanuts, (2) showed the empty risk option; (3) occluded the risk option; (4) showed the two potential risk outcome banana and cucumber) in the risk outcome bowl, and then placed just one under the risky option; (5) reminded the chimpanzee of the safe option; and then (6) the chimpanzee could choose. (b). Birds-eye diagram of setup. (Figure adapted from supplemental figure in Rosati & Hare, 2013).

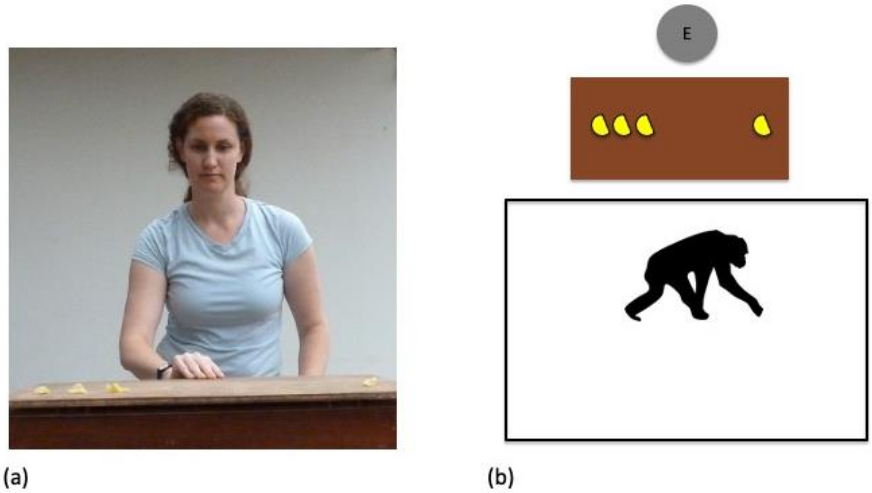

Figure S2: Setup for inter-temporal choice task. (a) Photo of setup (from chimpanzee's point of view): E sat across from the chimpanzee at a table, placing the smaller, immediate reward and larger, delayed reward on opposite sides of the table. (b). Birds-eye diagram of setup. (Figure adapted from supplementary figure in Rosati, DiNicola, & Buckholtz, "Chimpanzee cooperation Is fast and independent from self-control," Psychological Science (vol. 29, issue 11) pp. 1832–1845. Copyright © 2018 (Association for Psychological Science). DOI: 10.1177/0956797618800042)

Hormone sampling procedure and assays

Samples for endocrine analysis of cortisol and testosterone were collected concurrently during the six-week period of behavioral testing during which chimpanzees completed the task battery. We collected at least 4 (maximum of 8) saliva samples from each individual on different days; as some assay reads were excluded during the assay process so the total number per individual ranged from 3-8. All saliva samples were collected between 9:30 and 11 AM to account for circadian changes in hormone levels. Samples were always collected pre-task (e.g., before an individual participated in any cognitive testing that day).

The procedures for collecting and storing the saliva samples follow previous work in this population (Wobber & Hare, 2011; Wobber, Hare, Lipson, Wrangham, & Ellison, 2013; Wobber et al., 2010). First, the collector first put on gloves or thoroughly washed and disinfected their hands. They then poured ground Sweet Tarts candy onto a cotton round, and placed the cotton round inside the subject's lip so that it could suck on the cotton to absorb the saliva. The experimenter held on to the cotton throughout the collection procedure to avoid potential contamination. Once sufficient saliva was absorbed, the cotton round was placed into a syringe and squeezed to express the saliva into a test tube. The collection period for any particular sample did not span longer than 10 min. Immediately after collection, fifty microliters of 0.1% sodium azide solution was added to the saliva samples to prevent contamination, so the samples could be stored at room temperature on site. Prior work has validated aspects of this method (e.g., collection using cotton pad) for use with primates (Higham, Vitales, Mas Riveras, Ayala, & Maestripieri, 2010), as well as with this specific method for the chimpanzee population (Wobber & Hare, 2011; Wobber, et al., 2013; Wobber, et al., 2010).

Once the samples were returned to the US, they were frozen and later analyzed at the Hominoid Reproductive Ecology Laboratory at the University of New Mexico. Cortisol was assayed using the ALPCO cortisol radioimmunoassay (38-CORHU-R96, Salem, NH), following a secondary protocol for human saliva. Samples were run 1:1, but we diluted assay standards and controls to a range of 0.85-45 ng/ml using an assay buffer that mimics the properties of saliva (Salimetrics, Carlsbad, CA). For testosterone, we used the MP Biomedicals ImmunoChem double antibody testosterone radioimmunoassay kit (0718910-CF, Solon, OH) with modifications for saliva. Samples (1:4) and standards (1:20, 0-500 pg/ml) were diluted prior to assay in phosphate-buffered saline and introduced into the assay in 200ul duplicate aliquots. An extra 200ul of buffer was added to all tubes to promote parallelism. After addition of radiolabeled testosterone (50ul) and antibody (200ul), tubes were incubated in a water bath overnight at 37C. To precipitate bound fractions, 100ul of second antibody was added, and tubes were incubated at 37C for an additional hour before centrifuging at 4C for 2 hours and decanting. These methods were sufficient to achieve parallelism for both assays, and all samples were detectable on the curves. Additionally, we were able to rerun a small number of chimpanzee saliva samples that had been previously assayed with the discontinued DSL cortisol radioimmunoassay (Wobber et al. 2010), and the results were comparable ($r = 0.601$, $N = 13$, $p = 0.02$).

Intra-assay coefficients of variation (CVs) were 9.1% (low control) and 8.9% (high) for cortisol and 10.5% (low) and 12.9% (high) for testosterone. If sample volume allowed, samples with duplicate CVs higher than 15% were reanalyzed. We omitted results with CVs >25%. Intraassay CVs averaged 7.3% for cortisol and 8.2% for testosterone. Sample volumes did not

always allow both assays, in which case typically cortisol was assayed from that sample; see specific sample sizes in the results.

Supplemental Results

Risky choice task

We first confirmed chimpanzees exhibited appropriate patterns in the food preference pretest. Overall, chimpanzees chose the preferred food (good risk outcome - banana) on 100% of preference trials when it was contrasted with the non-preferred food (bad risk outcome - cucumber); they chose the preferred food on 89.6% of trials when contrasted with the intermediate food (safe outcome – peanuts); and finally, they chose the intermediate food on 92.9% of trials when contrasted with the non-preferred option. That is to say, the good risk outcome was preferred compared to the safe outcome, which was preferred over the bad risk outcome. In the main analyses of risky preferences reported in the main text, we then accounted for any potential differences in food preferences across individuals including their relative *preference score* as a covariate (as in prior work; Rosati & Hare, 2012, 2013; Rosati & Hare, 2016). We calculated this score from each individual's choices in the pretest, by averaging across all trials where they chose between the safe option food type and one of the risk outcome food types. This resulted in an index ranging from zero to one such that an individual who exhibited completely ordinal preferences (e.g., preferred the good risk outcome over the safe option, and preferred the safe option over the bad risk outcome) would have a score of 0.5, an individual who preferred the safe option when relatively more would have a lower score, and an individual preferred the bad outcome would have a higher score.

We then used mixed models to analyze trial-by-trial performance for risk choices, affective responses, and switching. These results are reported in the main text; parameter estimates from the models are shown here in Tables S1, S2, and S3. All reported estimates are unstandardized.

Predictor	*Estimate*	*SE*	*z-value*	*p-value*
Sex (reference: female)	0.805	0.509	1.582	0.114
Trial number	0.000	0.015	0.013	0.989
Food preference score	3.071	1.781	1.724	0.085
Good prev. outcome (reference: bad)	0.607	0.275	2.211	**0.027**
Safe prev. outcome (reference: bad)	0.109	0.247	0.443	0.658
Age (in years)	-0.098	0.040	-2.439	**0.015**

Table S1: Predictors of risk preferences. The base model included *sex, trial number*, and *food preference score*; *prior outcome, age,* and their interaction were added in subsequent models to test their importance. The best fit model included *age*. Reference levels for predictors indicated in the table.

Predictor	*Estimate*	*SE*	*t-value*	*p-value*
Sex (reference: female)	0.084	0.068	1.234	0.225
Trial number (within trial type)	-0.002	0.002	-0.776	0.438
Trial type (reference: exposure)	0.058	0.041	1.408	0.159
Good outcome (reference: bad)	-0.222	0.031	-7.061	< **0.0001**
Safe outcome (reference: bad)	-0.175	0.029	-6.093	< **0.0001**
Age (in years)	0.003	0.005	0.571	0.571

Table S2: Predictors of emotional responses to risky choice outcomes. The base model included *sex, trial number*, and *trial type* (test or exposure); *outcome, age,* and their interaction were added in subsequent models to test their importance. The best fit model did not include age.

Predictor	*Estimate*	*SE*	*t-value*	*p-value*
Sex (reference: female)	-0.215	0.761	-0.283	0.78
Trial number	-0.016	0.040	-0.391	0.70
Binary outcome (reference: bad)	-2.806	0.545	-5.145	< **0.0001**
Age (in years)	-0.0262	0.059	-0.441	0.66

Table S3: Predictors of switching responses. The base model included *sex* and *trial number*; *binary outcome, age,* and their interaction were added in subsequent models to test their importance. The best fit model did not include age.

Intertemporal choice task

We used mixed models to analyze trial-by-trial performance for choices and affective responses. These results are reported in the main text; parameter estimates from the models are shown here in Tables S4 and S5. Figure S3 shows a box plot of affective responding by age cohort (but note that the primary analyses included age as a continuous predictor, not categorical age cohorts).

Predictor	*Estimate*	*SE*	*z-value*	*p-value*
Sex (reference: female)	0.039	0.246	0.159	0.87
Trial number	-0.031	0.023	-1.348	0.18
Number pretest	0.699	0.600	1.165	0.24
Age (in years)	-0.017	0.020	-0.874	0.38

Table S4: Predictors of intertemporal preferences. The base model included *sex, trial number*, and *number pretest* performance; *age* was added to test its importance. The best fit model did not include *age*. Reference levels for predictors indicated in the table.

Predictor	*Estimate*	*SE*	*t-value*	*p-value*
Sex (reference: female)	0.277	0.122	2.270	**0.029**
Trial number (within trial type)	-0.020	0.006	-3.315	< **0.001**
Trial type (reference: exposure)	0.488	0.078	6.239	< **0.0001**
Choice (reference: immediate option)	1.052	0.111	9.502	< **0.0001**
Age (in years)	-0.006	0.010	-0.567	0.57
Age X Choice	-0.027	0.007	-3.993	< **0.0001**

Table S5: Predictors of emotional responses to intertemporal choice outcomes. The base model included *sex, trial number*, and *trial type* (test or exposure); *choice outcome, age,* and their interaction were added in subsequent models to test their importance. The best fit model included the age*choice outcome interaction.

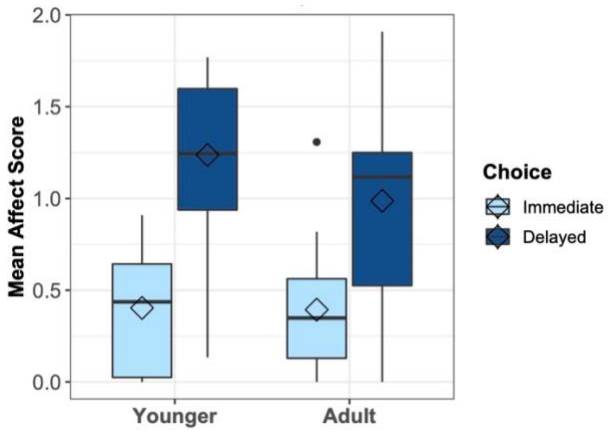

Figure S3: Affective responding in the inter-temporal choice task. Chimpanzees showed more negative affective responses to delayed compared to immediate rewards, and responses to delayed rewards declined with age. Boxplot hinges indicate the lower and upper quartile, the horizontal line represents the median, diamonds indicate the mean, and whiskers indicate the minimum and maximum range of the analyzed data. Outliers are plotted as individual points. Note that the primary analyses reported above included age as a continuous predictor, not categorical age cohorts.

Developmental trajectories for cortisol and testosterone

We then examined developmental changes in the chimpanzees' hormone levels. We analyzed a total of 169 cortisol samples comprising 95 from the 22 adults and 74 from the 18 younger chimpanzees (range: 3-8 across individuals); the mean coefficient of variation in the analyzed cortisol samples was 7.32 ng/ml within appropriate values. We analyzed a total of 160 testosterone samples, comprising 86 from the adults and 74 from the younger chimpanzees; the mean coefficient in the analyzed testosterone samples was 8.23 pg/ml, also within the appropriate range. Figure S4 depicts the relationships between individuals' log-transformed values and age.

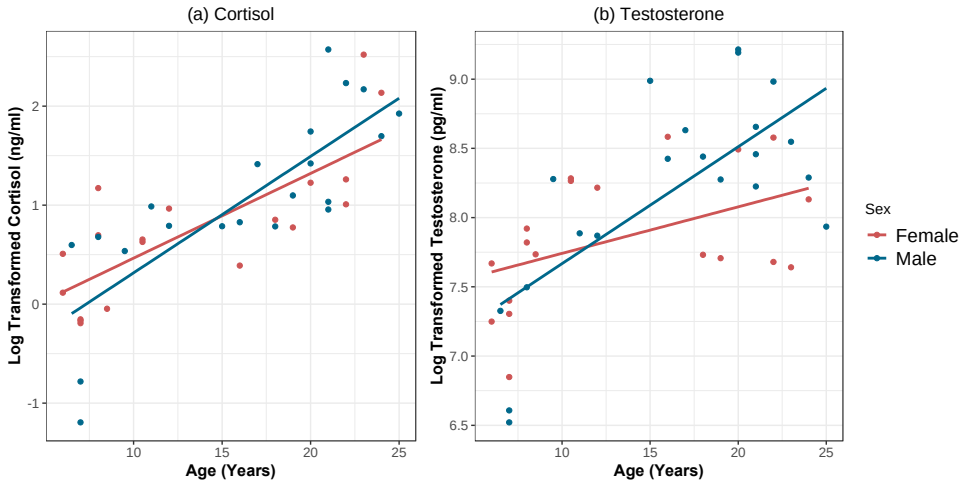

Figure S4: Physiological changes by age. (a) Male and female chimpanzees show increasing cortisol levels over the sampled age range. (b) Especially male chimpanzees showed increasing testosterone levels over this age range. Scatter plot depicts means of log-transformed average hormone values for each individual.

We used mixed models to analyze hormones values; these results are reported in the main text and parameter estimates from the models are shown here in Tables S6 and S7.

Predictor	*Estimate*	*SE*	*t-value*	*p-value*
Sex (reference: female)	-0.318	0.447	-0.711	0.48
Age	0.078	0.019	4.038	< **0.0001**
Age X Sex	0.028	0.028	1.005	0.31

Table S6: Predictors of cortisol levels. The base model included *sex; age* and the *age X sex* interaction were added in subsequent models to test their importance. The best fit model included age.

Predictor	*Estimate*	*SE*	*t-value*	*p-value*
Sex (reference: female)	-0.821	0.503	-1.631	0.10
Age	0.026	0.022	1.193	0.23
Age X Sex	0.062	0.031	1.988	**0.047**

Table S7: Predictors of testosterone levels. The base model included *sex; age* and the *age X sex* interaction were added in subsequent models to test their importance. The best fit model included the *age X sex* interaction.

Correlations across measures

To investigate chimpanzees' performance holistically across the cognitive, emotional, and hormone measures, we first conducted a series of bivariate Pearson correlations. For these correlations, we used the following variables: (1) individual's *age* (in years); (2) individual's mean *risky choice* in the risk task; (3) individual's *risk affect score* (a difference score indexing mean affective responses to bad outcomes – affect responses to good outcomes, for that individual); (4) individual's mean *temporal choices* for the delayed reward in the intertemporal choice task; (5) individual's mean *temporal affect score* (a difference score indexing mean affective responses to delayed rewards – affect responses to immediate rewards, for that individual); (6) mean log-transformed *testosterone* values for that individual; and (8) mean log-transformed *cortisol* values for that individual. Hormone values were log-transformed to address data skew, following standard practices for these kinds of data. All bivariate correlations values are reported in Table S8.

	Age	Risky choice	Risk affect score	Temporal choice	Temporal affect score	Cortisol
Risky choice	**r_p = -0.364** **p = 0.02**					
Risk affect score	r_p = 0.197 p = 0.22	r_p = 0.172 p = 0.29				
Temporal choice	r_p = -0.174 p = 0.28	r_p = 0.157 p = 0.33	r_p = -0.240 p = 0.14			
Temporal affect score	**r_p = -0.458** **p = 0.003**	**r_p = 0.343** **p = 0.03**	r_p = 0.154 p = 0.24	r_p = 0.118 p = 0.47		
Cortisol	**r_p = 0.790** **p < 0.001**	**r_p = -0.304** **p = 0.057**	r_p = 0.160 p = 0.32	r_p = -0.043 p = 0.79	**r_p = -0.347** **p = 0.028**	
Testosterone	**r_p = 0.618** **p < 0.001**	r_p = -0.226 p = 0.16	**r_p = 0.375** **p = 0.017**	r_p = -0.173 p = 0.28	r_p = -0.191 p = 0.24	**r_p = 0.642** **p < 0.001**

Table S8: Correlations between measures. Correlations for all pairwise comparisons across cognitive, affective, and physiological measures.

Principal component analysis

We next conducted a principal component analysis of these measures, designed to cluster performance variables into sets of measures that share significant variance so we could look at what measures clustered as well as compare these summary values across ages and sexes. We first assessed the adequacy of our correlation matrix by implementing a Bartlett's test for sphericity (Budaev, 2010). The test was significant (χ^2_6 = 44.61, df = 15, $p < 0.001$), indicating that the correlations between measures were sufficient for principal component analysis. A principal component analysis with these variables yielded two principal components with an adjusted eigenvalue >1 (PC1= 1.58 and PC2 = 1.14). Parallel analysis [26] confirmed retention of these two components. The first principal component explained 35.7% of the variance, while the second component explained 23.6% of the variance; together both components explained ~68% of the variance (see Table S9).

Measure	**PC1**	**PC2**
Risky choice	-0.373	0.443
Risk affect score	0.204	0.699
Temporal choice	-0.238	-0.215
Temporal affect score	-0.373	0.450
Cortisol	0.562	-0.008
Testosterone	0.555	0.258
Variance explained	35.7%	23.6%

Table S9. Principal components and eigenvectors. Hormone metrics and the risk affect score loaded positive on the first component, whereas risky choice, delayed choice, and discounting affect score loaded negatively. Risky choice, the affect scores, and testosterone all had a positive contribution to the second component, whereas discounting choices had a negative contribution. N = 40 individuals.

References

Budaev, S. V. (2010). Using principal components and factor analysis in animal behaviour research: caveats and guidelines. *Ethology, 116*, 472-480.

Cantwell, A., Buckholtz, J. W., Atencia, R., & Rosati, A. G. (2022). The origins of cognitive flexibility in chimpanzees. *Developmental Science, 25*, e13266.

Cole, M. F., Cantwell, A., Rukundo, J., Ajarova, L., Fernandez-Navarro, S., Atencia, R., & Rosati, A. G. (2020). Healthy cardiovascular biomarkers across the lifespan in wild-born chimpanzees (Pan troglodytes). *Philosophical Transactions of the Royal Society B, 375*, 20190609.

Higham, J. P., Vitales, A. B., Mas Riveras, A., Ayala, J. E., & Maestripieri, D. (2010). Measuring salivary analytes from free-ranging monkeys. *Physiology & Behavior, 101*, 601-607.

Rosati, A. G., DiNicola, L., & Buckholtz, J. W. (2018). Chimpanzee cooperation is fast and independent from self-control. *Psychological Science, 29*, 1832-1845.

Rosati, A. G., & Hare, B. (2011). Chimpanzees and bonobos distinguish between risk and ambiguity. *Biology Letters, 7*, 15-18.

Rosati, A. G., & Hare, B. (2012). Decision-making across social contexts: Competition increases preferences for risk in chimpanzees and bonobos. *Animal Behaviour, 84*, 869-879.

Rosati, A. G., & Hare, B. (2013). Chimpanzees and bonobos exhibit emotional respones to decision outcomes. *PLoS One, 8*, e63058.

Rosati, A. G., & Hare, B. (2016). Reward type modulates human risk preferences. *Evolution and Human Behavior, 37*, 159-168.

Rosati, A. G., Herrmann, E., Kaminski, J., Krupenye, C., Melis, A. P., Schroepfer, K., . . . Hare, B. (2013). Assessing the psychological health of captive and wild apes: A response to Ferdowsian et al. (2011). *Journal of Comparative Psychology, 127*, 329-336.

Warneken, F., & Rosati, A. G. (2015). Cognitive capacities for cooking in chimpanzees. *Proceeding of the Royal Society of London B, 282*, 20150229.

Wobber, V., & Hare, B. (2011). Psychological health of orphan bonobos and chimpanzees in African sanctuaries. *Plos One, 6*, e17147.

Wobber, V., Hare, B., Lipson, S., Wrangham, R., & Ellison, P. (2013). Different ontogenetic patterns of testosterone production reflect divergent male reproductive strategies in chimpanzees and bonobos. *Physiology & Behavior, 116-117*, 44-53.

Wobber, V., Hare, B., Maboto, J., Lipson, S., Wrangham, R., & Ellison, P. T. (2010). Differential changes in steroid hormones before competition in bonobos and chimpanzees. *Proceedings of the National Academy of Sciences, 107*, 12457-12462.