



Comparative Cognitive Development and Endocrinology in Pan and Homo

Citation

Wobber, Victoria Elizabeth. 2012. Comparative Cognitive Development and Endocrinology in Pan and Homo. Doctoral dissertation, Harvard University.

Link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:9453706>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#)

© 2012 Victoria Elizabeth Wobber
All rights reserved.

Comparative Cognitive Development and Endocrinology in *Pan* and *Homo*

Abstract

Key insights into the evolutionary origins of human social behavior can be gained via study of our closest living relatives, bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*). Despite being equally related to humans, these two species differ importantly in aspects of their morphology, physiology, behavior, and cognition. Morphological comparisons reveal numerous traits in bonobos that can be viewed as paedomorphic, or juvenile, relative to chimpanzees. Meanwhile, comparisons of endocrinology in the two species suggest that aspects of steroid physiology have changed significantly in bonobos in line with their reductions in male mating competition. Based on this evidence, I tested the hypothesis that behavioral and cognitive differences between bonobos and chimpanzees derive from changes in their 1) developmental trajectories of behavioral and cognitive traits and 2) neuroendocrine influences on behavior and cognition.

I tested this hypothesis by studying semi free-ranging populations of bonobos and chimpanzees. First, I found that bonobos retained juvenile levels of food sharing and social inhibition into adulthood, leading them to differ from chimpanzees in these traits as adults. Second, I found that bonobos showed muted elevations in their levels of testosterone from infancy to adulthood in comparison to chimpanzees, suggesting that numerous aspects of development differ between these two species. Third, I found that male bonobos and chimpanzees differ in their immediate neuroendocrine shifts surrounding competition,

implicating changes in proximate mechanisms influencing social behavior between the two species. Fourth, I found that patterns of cognitive development in these two apes differed significantly from those of human children.

These results provide substantial support for my hypothesis that phenotypic differences between bonobos and chimpanzees evolved via shifts in bonobo development and neuroendocrine physiology. More broadly, they illustrate how behavioral and cognitive evolution can occur through changes in ontogenetic trajectories and neuroendocrine mechanisms. These findings thus show the merits of integrating ultimate and proximate levels of analysis in studies of the evolution of human behavior and cognition.

TABLE OF CONTENTS

Abstract	iii
Acknowledgements	vi
Chapter 1. Introduction	1
Chapter 2. The value of comparative developmental studies of behavior and cognition to anthropology	9
Chapter 3. Bonobos exhibit delayed development of social behavior and cognition relative to chimpanzees	35
Chapter 4. Different ontogenetic patterns of testosterone production reflect divergent male reproductive strategies in chimpanzees and bonobos	51
Chapter 5. Differential changes in steroid hormones prior to competition in bonobos and chimpanzees	91
Chapter 6. Differences in the early cognitive development of children and great apes	115
Chapter 7. Discussion	174
Appendix 1. Distinguishing behavior and cognition	181
Appendix 2. Insights into human evolution from taxa distantly related to humans	184
Appendix 3. Supplemental methods, analyses, tables and figures for Chapter 3	190
Appendix 4. Supplemental tables and figures for Chapter 4	212
Appendix 5. Supplemental methods, analyses, tables and figures for Chapter 5	221
Appendix 6. Supplemental methods, analyses, tables and figures for Chapter 6	237

Acknowledgements

There are a great number of people to whom I owe thanks for their help throughout my graduate career. First and foremost I should thank my advisors, Richard Wrangham and Brian Hare, for their extensive support and guidance. Richard has been a wonderful source of advice on all levels for 6 (or in fact 10) years, whether commenting on a draft or chatting about ideas for the future. Brian was instrumental in helping me to begin my thesis research, and serves as an invaluable source of consistent, contagious enthusiasm. I would also like to thank the other members of my dissertation committee, Peter Ellison, Felix Warneken, and Daniel Lieberman, both for comments in the final stages of dissertation-writing but also for guidance throughout my graduate work. Michael Tomasello has also been extremely supportive of my research in its early and more recent stages.

I also owe thanks to numerous members of the Department of Human Evolutionary Biology, and the former Biological Anthropology wing. Nothing in the department would be possible without Meg Lynch, and I would like to thank her for answering even my most trivial questions. I would also like to thank Susan Lipson for her help in carrying out the endocrine analyses described here, and Melissa Emery Thompson for helping me to develop the saliva collection procedure. I am greatly indebted to the members of the Behavioral Ecology Group, past and present, for their feedback throughout the years – including Charlie Nunn, Katie Hinde, Zarin Machanda, Katie McAuliffe, Rachel Carmody, Alex Georgiev, Ian Gilby, Luke Glowacki, and Meg Crofoot, among others. Neil Roach and Amanda Lobell have also been great resources throughout graduate school, whether to discuss ideas or get the latest gossip.

I would like to thank fellow researchers whom I spent time with in the field, including Esther Herrmann, Alex Rosati, Evan MacLean, Kara Schroepfer, Jingzhi “Hippo” Tan and

Vanessa Woods, for discussing research methods, debating which chimps are the cutest, and putting up with my desires for food. Also thanks to Korrina Duffy for help in coordinating this research more recently, Zanna Clay for help in data collection, the research assistants at the Max Planck Institute for Evolutionary Anthropology who were fundamental in making the child research possible, Angela Loose, Sarah Girlich, Marie Schaefer, Isabelle Lehn, and Melanie Schreiner, and to Rob Tennyson, Tatjana Welikanowa, Maria Baumeister, Michael Caldwell, and Leah Kaiser for their help in reliability coding. Thanks also to Esther, Alicia Melis, Juliane Kaminski, and Jeff Stevens for help in the early stages of my research career.

None of the work at the sanctuaries would have been possible without the support of the staff there, including Rebeca Atencia, Debby Cox, Fernando Turmo, Lisa Pharoah, David Greer, Keith Brown, the Jane Goodall Institute, Claudine Andre, Valery Dhanani, Pierrot Mbonzo, Crispin Mahamba, Fanny Mehl, Suzy Kwetuenda, and Les Amis des Bonobos du Congo. Suzy deserves special thanks for her help in data collection and coordinating bonobo research. I would also like to thank Jean Maboto for helping me to develop the saliva collection method and for being consistently supportive of our research at Tchimpounga. In addition, I owe thanks to all of the chimpanzee and bonobo caretakers, with their knowledge of ape behavior essential to our being able to successfully carry out research. Finally, here I should also thank the chimpanzees and bonobos themselves, including Elikia, for his special gifts, and Luozi, for making it patently clear that bonobo psychology is peculiar.

I would like to thank my funding sources for making this research possible financially. I am grateful to the Harvard Anthropology Department for research funds to collect preliminary data, and to the National Science Foundation, Leakey Foundation, and Wenner Gren Foundation for funding collection of my dissertation data. Finally, support from the Harvard

Mind/Brain/Behavior Initiative and the Cora DuBois Foundation was instrumental in the final stages of analyzing and writing up this research.

Thanks to my parents, for being supportive of a non-traditional career choice and becoming more and more comfortable with the idea of me traveling to Congo.

And thanks to Brian Addison. I am lucky to have found someone whom I can look to for support on any level, even if it's just looking for an elbow drop or two.

Chapter 1: **Introduction**

How does human behavior differ from that of other species? What aspects of our early development and our physiology influence the way that we behave? In this thesis, I examine the evolutionary origins and proximate mechanisms of human social behavior through comparisons of humans and our closest living relatives, chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*). Bonobos and chimpanzees are equally related to humans, with the last common ancestor (LCA) of humans and these two species existing approximately 6 to 7 million years ago (Brunet et al., 2002; Glazko & Nei, 2003; Ruvolo, 1997). Chimpanzees and bonobos are also closely genetically related to one another, having diverged as recently as 850,000 years ago (Won & Hey, 2005). Despite their genetic similarity, these two species are notably distinct from one another in aspects of their morphology, physiology, behavior, and cognition. For example, bonobos are physically smaller and more gracile, have lesser sexual dimorphism in levels of testosterone, exhibit female as opposed to male dominance, and are more impatient relative to chimpanzees (Coolidge, 1933; Parish & de Waal, 2000; Rosati, Stevens, Hare, & Hauser, 2007; Sannen, Heistermann, van Elsacker, Moehle, & Eens, 2003; White & Wood, 2007).

Comparisons between bonobos and chimpanzees can therefore inform our knowledge of human origins in two main ways. First, given that there is some debate over whether the last common ancestor of humans and the genus *Pan* was more similar to modern-day chimpanzees (Wrangham & Pilbeam, 2001), or modern-day bonobos (Zihlman, Cronin, Cramer, & Sarich, 1978), studying both species is necessary in order to clarify the traits that were inherited in our hominid lineage. Second, any variation between the two *Pan* species can implicate specific trait changes and selection pressures that were part of recent ape evolution, providing examples of

evolutionary mechanisms that may have been present in human evolution as well. For instance, it is possible that humans and bonobos have undergone convergent selection pressures, if both were diverging from relatively chimpanzee-like LCAs (Hare, Wobber, & Wrangham, in press; Wrangham, Wilson, & Muller, 2006). Thus we can gain important insight into human evolution by understanding differences between bonobos and chimpanzees.

In this thesis, I test the specific hypothesis that phenotypic differences between bonobos and chimpanzees have arisen via delays in bonobo development and changes in bonobo neuroendocrine physiology relative to chimpanzees. This hypothesis is based on the presumption that the LCA of bonobos and chimpanzees was chimpanzee-like. Though there is some debate over this assertion (De Waal & Lanting, 1997; Wrangham & Pilbeam, 2001; Zihlman, et al., 1978), as mentioned above, morphological and behavioral comparisons of bonobos and chimpanzees to their nearest non-human living relative, gorillas (*Gorilla gorilla*), implicate a chimpanzee-like LCA of their clade (Wrangham & Pilbeam, 2001). Here therefore I propose that differences between extant bonobos and chimpanzees can be taken to represent changes that have occurred along the bonobo lineage in their recent evolutionary history.

The first component of my hypothesis, that bonobos exhibit delays in aspects of their development relative to chimpanzees, derives partly from skeletal comparisons of the two species. Studies of morphological development in bonobos have revealed that in specific aspects of their cranium they demonstrate paedomorphosis, or retention of juvenile characteristics into adulthood, relative to chimpanzees (Durrleman, Pennec, Trouve, Ayache, & Braga, 2012; Lieberman, Carlo, Ponce de Leon, & Zollikofer, 2007; Shea, 1984; Zihlman & Cramer, 1978). Preliminary evidence suggests that elements of bonobo behavior can also be viewed as paedomorphic relative to chimpanzees (Kuroda, 1989). Indications of paedomorphosis in both

the cranium and behavior raise the possibility that changes in the developmental timing of skeletal growth are the consequence of a unified selection pressure that has influenced the ontogenetic patterns of diverse phenotypic traits including physiology, behavior, and cognition. Specifically, selection for reduced aggression in bonobos may have caused them to retain numerous juvenile traits into adulthood, analogous to the suite of juvenilized traits tied to selection for reduced aggression in domesticated animals (Hare, et al., in press). To investigate whether bonobos show delays in multiple aspects of their development relative to chimpanzees, I performed comparative analyses of behavioral, cognitive, and endocrine development in both species, discussed here in Chapters 3, 4, and 6. In Chapter 2, I outline why this type of comparative developmental study is essential to our understanding of human evolution.

The second component of my hypothesis, that bonobos differ significantly from chimpanzees in aspects of their neuroendocrine physiology, derives from studies of social behavior and endocrinology in the two species – with my work presenting one of the first opportunities to link these two lines of inquiry. Bonobos differ from chimpanzees in numerous aspects of their social behavior, with reductions in aggressive behavior among bonobo males (Furuichi, 2011; Kano, 1992; Muller, 2002; Muller & Wrangham, 2009), increases in non-conceptive sexual behavior among bonobos of both sexes and all ages (de Waal, 1987; Hohmann, 2001; Hohmann & Fruth, 2000; Woods & Hare, 2011), and increases in tolerant behaviors such as food sharing and cooperation (Fruth & Hohmann, 2002; Hare, Melis, Woods, Hastings, & Wrangham, 2007). Differences have also been found between bonobos and chimpanzees in aspects of their steroid physiology. Bonobos appear to exhibit a lesser sex difference in androgen levels than do chimpanzees (Sannen, et al., 2003), with bonobo males also showing a weaker relationship between testosterone and dominance rank (Marshall &

Hohmann, 2005; Muller & Wrangham, 2004) and lesser elevations in testosterone when females are reproductively receptive than male chimpanzees (Muller & Wrangham, 2004; Surbeck, Deschner, Schubert, Weltring, & Hohmann, in press). Given that testosterone importantly influences energetic allocation toward male reproductive effort, including male reproductive behavior (Ellison, 2003; Wingfield, Hegner, Dufty, & Ball, 1990), there may be important links between the changes in steroid physiology and the changes in social behavior found among bonobos relative to chimpanzees. I investigated this possibility through comparisons of steroid responsiveness during competitive interactions in bonobos and chimpanzees, described in Chapter 5.

Finally, in Chapter 6 I use these two model species to investigate human origins through direct comparisons of cognitive development between young bonobos, chimpanzees, and humans. Much debate has centered on whether shifts in cognitive development during human evolution facilitated our species' cognitive complexity as adults, with some hypotheses proposing that a prolongation of cognitive development in human evolution was responsible for the overall extension of the human juvenile period relative to that of other apes (Bjorklund, 1997; Bogin & Smith, 1996; Kaplan, Hill, Lancaster, & Hurtado, 2000). However, little comparative data on behavioral and cognitive development in non-human apes has been available to fully test these hypotheses. I therefore performed a study of cognitive ontogeny in infant and juvenile humans, chimpanzees, and bonobos using both cross-sectional and longitudinal data. This comparison allowed me to examine how the rate and pattern of human cognitive development compares that of other ape species.

The results of this thesis support the hypothesis that there are substantive phenotypic differences between bonobos and chimpanzees that can provide insight into the evolutionary

mechanisms that have shaped human origins. I found several lines of support for the first component of my hypothesis, that bonobos show delays in their development relative to chimpanzees, with bonobos exhibiting delays in aspects of their social behavior (Chapter 3, published in *Current Biology* in 2010), social cognition (also Chapter 3), and endocrine maturation (Chapter 4). I also found support for the second component of my hypothesis, that shifts in bonobo neuroendocrine physiology correlate with differences in their social behavior relative to chimpanzees, finding differences between the two species in the rapid steroid shifts surrounding competition (Chapter 5, published in *Proceedings of the National Academy of Sciences* in 2010). Finally, I found that differences in the early ontogeny of cognitive skills exist not only between bonobos and chimpanzees (as shown in Chapter 3), but also between *Pan* and humans, with particular accelerations in human socio-cognitive development during infancy and juvenility (Chapter 6). On the whole, these findings are significant to human evolutionary biology in suggesting that important shifts in the pattern and pace of development, as well as the relationship between physiology and behavior, may underlie the differences between humans and our closest relatives. This work presents exciting directions for future inquiry, in using comparative models of development and neuroendocrinology to better understand aspects of human evolution.

Chapter 1 Literature Cited.

- Bjorklund, D. F. (1997). The role of immaturity in human development. *Psychological Bulletin*, 122(2), 153-169.
- Bogin, B., & Smith, B. (1996). Evolution of the Human Life Cycle. *American Journal of Human Biology*, 8, 703-716.
- Brunet, M., Guy, F., Pilbeam, D., Mackaye, H., Likius, A., Ahounta, D., . . . Zollikofer, C. (2002). A new hominid from the Upper Miocene of Chad, Central Africa. *Nature*, 418, 145-151.
- Coolidge, H. (1933). Pan paniscus, pigmy chimpanzee from south of the Congo river. *American Journal of Physical Anthropology*, 18(1), 1-59.
- de Waal, F. (1987). Tension regulation and nonreproductive functions of sex in captive bonobos (*Pan paniscus*). *National Geographic Research Reports*, 3, 318-335.
- De Waal, F., & Lanting, F. (1997). *Bonobo: the forgotten ape*. Berkeley: University of California Press.
- Durrleman, S., Pennec, X., Trouve, A., Ayache, N., & Braga, J. (2012). Comparison of the endocranial ontogenies between chimpanzees and bonobos via temporal regression and spatiotemporal registration. *Journal of Human Evolution*, 62(1), 74-88. doi: 10.1016/j.jhevol.2011.10.004
- Ellison, P. (2003). Energetics and reproductive effort. *American Journal of Human Biology*, 15(3), 342-351.
- Fruth, B., & Hohmann, G. (2002). How bonobos handle hunts and harvests: why share food? In C. Boesch, G. Hohmann & L. Marchant (Eds.), *Behavioural diversity in chimpanzees and bonobos* (pp. 231-243). Cambridge: Cambridge University Press.
- Furuichi, T. (2011). Female Contributions to the Peaceful Nature of Bonobo Society. *Evolutionary Anthropology*, 20(4), 131-142. doi: 10.1002/evan.20308
- Glazko, G., & Nei, M. (2003). Estimation of divergence times for major lineages of primate species. *Molecular Biology and Evolution*, 20(3), 424-434.
- Hare, B., Melis, A. P., Woods, V., Hastings, S., & Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology*, 17(7), 619-623.
- Hare, B., Wobber, V., & Wrangham, R. (in press). The self-domestication hypothesis: bonobo psychology evolved due to selection against aggression. *Animal Behaviour*.
- Hohmann, G. (2001). Association and Social Interactions Between Strangers and Residents in Bonobos (*Pan paniscus*). *Primates*, 42(1), 91-99.

- Hohmann, G., & Fruth, B. (2000). Use and function of genital contacts among female bonobos. *Animal Behaviour*, 60, 107-120.
- Kano, T. (1992). *The Last Ape: Pygmy Chimpanzee Behavior and Ecology*. Stanford, California: Stanford University Press.
- Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology*, 9(4), 156-185.
- Kuroda, S. (1989). Developmental retardation and behavioral characteristics of pygmy chimpanzees. In P. Heltne & L. Marquardt (Eds.), *Understanding chimpanzees* (pp. 184-193). Cambridge, Massachusetts: Harvard University Press.
- Lieberman, D. E., Carlo, J., Ponce de Leon, M., & Zollikofer, C. (2007). A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. *Journal of Human Evolution*, 52, 647-662.
- Marshall, A., & Hohmann, G. (2005). Urinary Testosterone Levels of Wild Male Bonobos (*Pan paniscus*) in the Lomako Forest, Democratic Republic of Congo. *American Journal of Primatology*, 65(1), 87-92.
- Muller, M. (2002). Agonistic relations among Kanyawara chimpanzees. In C. Boesch, G. Hohmann & L. Marchant (Eds.), *Behavioural diversity in chimpanzees and bonobos* (pp. 112-123). Cambridge: Cambridge University Press.
- Muller, M., & Wrangham, R. (2004). Dominance, aggression and testosterone in wild chimpanzees: a test of the 'challenge hypothesis'. *Animal Behaviour*, 67, 113-123.
- Muller, M., & Wrangham, R. (2009). *Sexual Coercion in Primates and Humans: An Evolutionary Perspective on Male Aggression Against Females*. Cambridge, MA: Harvard University Press.
- Parish, A., & de Waal, F. (2000). The Other "Closest Living Relative" How Bonobos (*Pan paniscus*) Challenge Traditional Assumptions about Females, Dominance, Intra- and Intersexual Interactions, and Hominid Evolution. *Annals of the New York Academy of Sciences*, 907(1), 97-113.
- Rosati, A., Stevens, J., Hare, B., & Hauser, M. (2007). The evolutionary origins of human patience: Temporal preferences in chimpanzees, bonobos, and human adults. *Current Biology*, 17(19), 1663-1668.
- Ruvolo, M. (1997). Molecular phylogeny of the hominoids: inferences from multiple independent DNA sequence data sets. *Molecular Biology and Evolution*, 14(3), 248-265.
- Sannen, A., Heistermann, M., van Elsacker, L., Moehle, U., & Eens, M. (2003). Urinary testosterone metabolite levels in bonobos: a comparison with chimpanzees in relation to social system. *Behaviour*, 140, 683-696.

- Shea, B. (1984). An allometric perspective on the morphological and evolutionary relationships between pygmy (*Pan paniscus*) and common (*Pan troglodytes*) chimpanzees. In R. Susman (Ed.), *The pygmy chimpanzee: evolutionary biology and behavior* (pp. 89-130). New York: Plenum Press.
- Surbeck, M., Deschner, T., Schubert, G., Weltring, A., & Hohmann, G. (in press). Mate competition, testosterone and intersexual relationships in bonobos, *Pan paniscus*. *Animal Behaviour*.
- White, F., & Wood, K. (2007). Female Feeding Priority in Bonobos, *Pan paniscus*, and the Question of Female Dominance. *American Journal of Primatology*, 69(8), 837-850.
- Wingfield, J., Hegner, R., Dufty, R., & Ball, G. (1990). The “challenge hypothesis”: theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. *American Naturalist*, 136, 829-846.
- Won, Y., & Hey, J. (2005). Divergence Population Genetics of Chimpanzees. *Molecular Biology and Evolution*, 22(2), 297-307.
- Woods, V., & Hare, B. (2011). Bonobo but not chimpanzee infants use socio-sexual contact with peers. *Primates*, 52(2), 111-116.
- Wrangham, R., & Pilbeam, D. (2001). African apes as time machines. In B. Galdikas, N. Briggs, L. Sheeran, G. Shapiro & J. Goodall (Eds.), *All apes great and small* (Vol. 1, pp. 5-18). New York: Kluwer Academic/Plenum Publishers.
- Wrangham, R., Wilson, M., & Muller, M. (2006). Comparative rates of violence in chimpanzees and humans. *Primates*, 47(1), 14-26.
- Zihlman, A., & Cramer, D. (1978). Skeletal differences between pygmy (*Pan paniscus*) and common chimpanzees (*Pan troglodytes*). *Folia Primatologica*, 29, 86-94.
- Zihlman, A., Cronin, J., Cramer, D., & Sarich, V. (1978). Pygmy chimpanzee as a possible prototype for the common ancestor of humans, chimpanzees, and gorillas. *Nature*, 275(5682), 744-746.

Chapter 2:

The value of comparative developmental studies of behavior and cognition to anthropology

INTRODUCTION

Human life history patterns are central in distinguishing our species from its closest relatives. The paradox in human life history research has been to account for aspects of our life course that are “slow” relative to other species, such as our extended lifespan and prolonged juvenile period, while at the same time explaining the aspects that are “fast,” such as our short inter-birth intervals and overall rapid reproductive rate (Kramer, 2005; Robson & Wood, 2008). Researchers have intensely debated the selection pressures that led to these observed features of human life history, citing shifts in energetics (Kramer & Ellison, 2010), morphology (Lieberman, 2012), behavior (Bogin & Smith, 1996; Hawkes, O'Connell, Blurton Jones, Alvarez, & Charnov, 1998; Kaplan, Hill, Lancaster, & Hurtado, 2000), cognition (Locke & Bogin, 2006), or some combination of these factors (Wrangham & Carmody, 2010) as central to the changes observed in our lineage.

In order to discriminate among these hypotheses, it is necessary to have adequate comparative data with which to juxtapose the human life history pattern to that of extant and fossil primates. Comparative data now exists for numerous primate taxa on rates of growth, aspects of dental development, and broad maturational indices such as the age at weaning and the age at which females first give birth (Dean et al., 2001; L. R. Godfrey, Samonds, Jungers, & Sutherland, 2001; Knott, 2001; Leigh, 1992; B. Smith, Crummett, & Brandt, 1994). However, comparative data on rates of behavioral and cognitive development is much more scarce. This likely owes to the logistical difficulties inherent in studying juvenile animals – because they are smaller, juveniles are often difficult to observe, and because they remain close to their mothers

they may not act independently or interact with other juveniles very often. Nonetheless, research on behavioral and cognitive development in non-human primates has proliferated in recent years, stemming both from heightened interest in the juvenile period and greater feasibility of studying juveniles at long-term field sites (Fairbanks & Pereira, 2002).

What I will argue here is that this type of comparative data on behavioral and cognitive development is critical in evaluating hypotheses of human life history evolution, reviewing a series of recent experiments supporting this conclusion. Though parameters such as somatic growth or age at weaning can provide a broad index of developmental stage, they represent only a proxy for concurrent behavioral and cognitive development (see Appendix 1 for a description of the distinctions between behavior and cognition). Thus particularly for hypotheses positing that changes in behavioral and cognitive development were central to human life history evolution, it is essential to have comparative data on these parameters from both humans and non-human primates.

Two major hypotheses have argued that changes in human behavioral and cognitive development in particular were instrumental in shaping human life history patterns. These models propose that with the greater complexities of adult life in humans, an extended period of immaturity was necessary in order to accommodate greater skill learning throughout development. The *embodied capital hypothesis* (Kaplan, et al., 2000; Lancaster & Kaplan, 2010) focuses on the demands of the human foraging niche, while the “*adolescence as apprenticeship*” hypothesis (Bogin, 2010; Bogin & Smith, 1996) focuses on the demands of adult social norms. Both models rely on two key assumptions: first, that aspects of adult behavior and cognition are somehow more complex in humans than among other hominoids, and second, that an extension (or addition) of early life stages is the primary means by which a species could accommodate any

such increase in behavioral or cognitive complexity. I illustrate these two assumptions in Figure 2.1, which I have created to model the embodied capital hypothesis. An analogous model could be created for the adolescence-apprenticeship hypothesis by moving the focal point to the adolescent/adult transition. Note that for the purposes of this figure, I use chimpanzees as the comparative taxon to provide the best available ancestral model for human patterns of behavioral and cognitive ontogeny, as done in the embodied-capital and adolescence-apprenticeship models (Bogin, 2010; Kaplan, et al., 2000) – a debatable but logistically reasonable presumption (Wrangham & Pilbeam, 2001) (see Appendix 2 for a description of advances that can be made by studies of more distant taxa). I also restrict my dependent variable to a mere index of behavioral or cognitive “complexity,” recognizing the limitations of this approach but doing so for simplicity of modeling developmental trajectories.

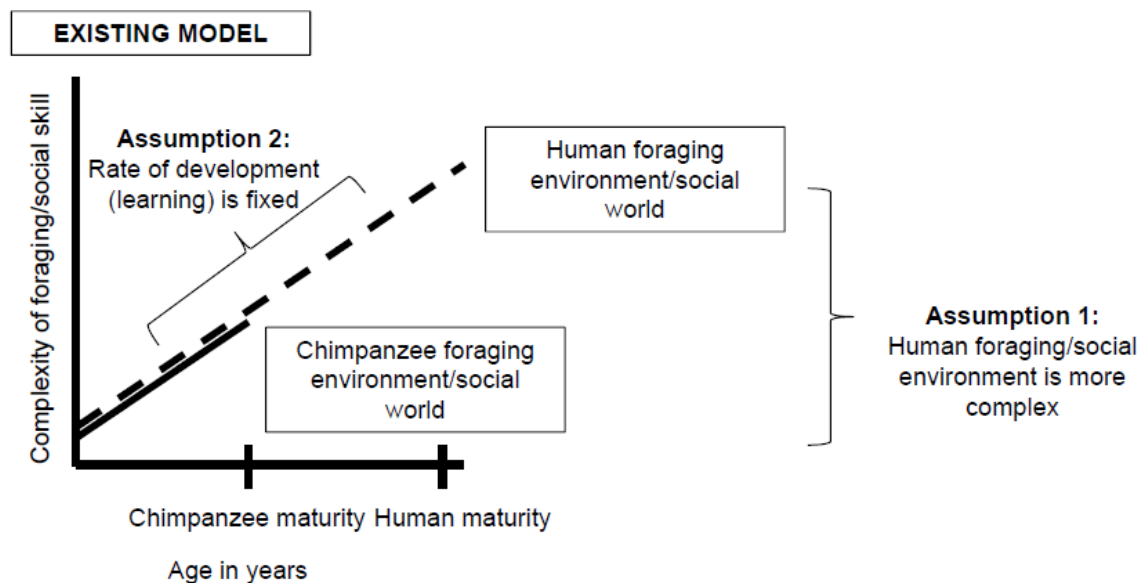


Figure 2.1. Model of the embodied capital hypothesis, based on the arguments of Kaplan and colleagues (2000). Age is depicted on the x-axis, with indications of the relative ages at chimpanzee and human sexual maturity. The y-axis denotes behavioral or cognitive “complexity” of tasks performed by adults. The relative end-stages of development for humans and chimpanzees are shown by their respective boxes. The human pattern of development is indicated by a dotted line, while the chimpanzee pattern is indicated by a solid line. The two assumptions of the model (greater complexity in humans, fixed rate of development) are indicated.

In this paper, I propose that the two assumptions made by these models are problematic without comparative data on behavioral and cognitive development. First, the assumption that human behaviors are more difficult to acquire than those of other species is challenging to address without comparable data on behavioral development collected from humans and non-human animals. Second, the assumption that a prolongation of juvenility is necessary to facilitate increases in behavioral complexity fails to incorporate the possibility for changes in the rate of development. In fact, the embodied capital and adolescence-apprenticeship models do not specify whether the rate of development in humans might have changed in addition to the prolonged duration of development. Thus here I simply intend to debate the point that a prolongation in juvenility would be the only means to shift ultimate behavioral and cognitive complexity, arguing that changes in the rate of development are an important additional mechanism to consider in studying behavioral and cognitive evolution. Before discussing this argument, I begin below by reviewing general life history parameters in human development. On the whole, I contend that studies of comparative behavioral and cognitive development are essential in discriminating among models of human life history evolution, in understanding the major changes between our lineage and those of other living primates and clarifying the selection pressures in our evolutionary past that may have driven these phenotypic changes.

“SLOW” ASPECTS OF HUMAN LIFE HISTORY

Numerous lines of evidence indicate that humans grow slowly in both somatic and reproductive traits relative to other primates. Humans have a later age of reproductive maturity, disperse from their natal groups at a later age, and typically have a later age of first birth than other apes (Anderson, Dallal, & Must, 2003; Coe, Connolly, Kraemer, & Levine, 1979; Knott,

2001; Sugiyama, 2004; Walker et al., 2006). Meanwhile, our early dental development is also slowed relative to other extant and fossil hominoids, with absolutely later emergence of the first molar (M1) and a proportionally later age of this landmark relative to weaning (T. Smith et al., 2010). (where in other taxa weaning and M1 emergence typically appear to coincide (B. Smith, 1992)). Finally, humans have an extended period of low growth rates between weaning and reproductive maturity, and ultimately reach adult body size at a later age than do chimpanzees (Bogin & Smith, 1996; Leigh, 2001; Walker, Hill, Burger, & Hurtado, 2006).

Intriguingly, recent investigations of gene expression in the brain have also supported the notion of slowed development in humans. Somel and colleagues examined patterns of expression across multiple genes active in the brains of humans, chimpanzees, and macaques and found evidence for “transcriptional neoteny” among genes expressed in the prefrontal cortex: the majority of genes showed a lesser magnitude of increase in their expression across development among humans relative to these other two primate species (Somel et al., 2009). In addition, several genes expressed in the prefrontal cortex were found to show a later age of peak expression in humans relative to chimpanzees and macaques, with similar patterns not present among genes expressed in the cerebellum (Liu et al., in press).

The finding that numerous aspects of human development are slowed raises the possibility that there is equivalent slowing in patterns of behavioral and cognitive development in humans compared to other apes. However, without comparative developmental data, it is open to question whether trajectories of behavioral and cognitive ontogeny scale precisely with aspects of somatic or reproductive ontogeny. Though some data is available in this area, in the next section I review why it is essential to have data collected comparably across both human and non-human species.

IS HUMAN BEHAVIOR OR COGNITION MORE COMPLEX?

The first major assumption of the embodied capital and adolescence-apprenticeship models is that there is a substantive difference in adult behavioral complexity between humans and chimpanzees – namely, that aspects of the human foraging and social environment are inherently more difficult to master during ontogeny than comparable traits in our closest living relatives (illustrated in Figure 2.1 as “Assumption 1”). However, this assumption is problematic without comparable data collected on both human and chimpanzee development, in showing that chimpanzees reach adult levels of foraging efficiency at an early age or that aspects of human social interactions are demonstrably more difficult to negotiate than the fission-fusion dynamics of chimpanzee social life. This type of data might in fact demonstrate that the skills needed for humans to forage and engage in their social world are comparably complex to the demands faced by chimpanzees. I have modeled this possibility in Figure 2.2, calling it the “what is complex?” model.

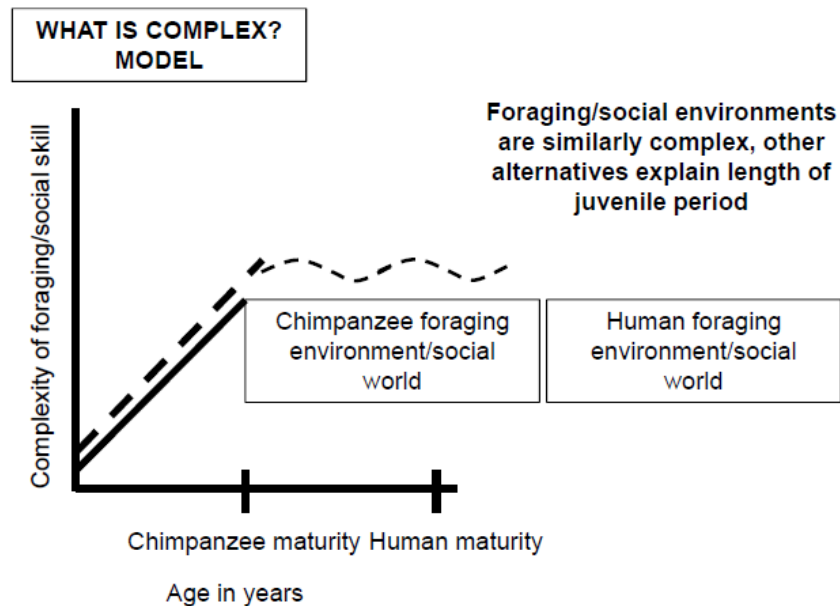


Figure 2.2. Depiction of the “what is complex?” model. As in Figure 2.1, age is depicted on the x-axis, with relative ages for chimpanzee and human maturity indicated. Again, the y-axis denotes behavioral or cognitive “complexity.” Human development is indicated by the dotted line and chimpanzee development by the solid line. In this model, I illustrate the possibility that the complexity of adult foraging/social life is similar for chimpanzees and humans, making it possible for humans acquire the necessary skills for adulthood in the same amount of developmental time as chimpanzees. This would leave open to question the function of the prolonged period of immaturity in humans, as depicted here by the curved line.

Evidence for the notion that human foraging is more complex than that of chimpanzees comes from the emphasis on hunting and extractive foraging in human societies, with Kaplan and colleagues pointing out that these components are much less substantial in the diet of chimpanzees (Kaplan, et al., 2000). Indeed, as they discuss, it is often difficult for human behavioral ecologists to match the skills of their study populations in food acquisition and processing, providing support for the notion that these tasks are difficult. Moreover, men in subsistence populations have been found to reach peak efficiency in hunting only relatively late in life, independent from gains in strength acquired with age, supporting the notion that hunting is difficult to master (Gurven, Kaplan, & Gutierrez, 2006). However, we know that human subsistence populations employ numerous different types of foraging techniques in addition to

hunting that might be less skill-based (e.g. gathering and processing plants or tubers) (Laden & Wrangham, 2005; Marlowe, 2001). Moreover, experimental investigations of these latter skills suggest that even inexperienced individuals may be able to match the capacities of seasoned foragers (Jones & Marlowe, 2002). The crucial point of the embodied capital hypothesis must then be that the extension of human immaturity facilitated greater skill acquisition in hunting, requiring that this extension be driven by males (since it is predominately males who engage in hunting behaviors) and that hunting be a substantial component of the human diet throughout our evolutionary history (with some debate on this topic (Laden & Wrangham, 2005; O'Connell, Hawkes, & Blurton Jones, 2002; Speth, 2011)). In turn, this would indicate that the compelling area for comparative data is whether acquisition of hunting proficiency is sufficiently more skill-demanding than acquisition of foraging proficiency in chimpanzees.

Unfortunately, developmental data with this level of resolution does not yet exist for chimpanzees, with no studies documenting the age at which chimpanzee individuals reach adult foraging capacity. For their comparisons of human and chimpanzee development, Kaplan and colleagues use a proxy for chimpanzee foraging rates by calculating caloric requirements from data on chimpanzee growth (Kaplan et al 2000, page 161). However, we know from studies of human foraging that there is significant individual and society-level variation in rates of production throughout development (Kaplan, et al., 2000; Kramer, 2005). Systematic data on chimpanzee feeding rates throughout ontogeny, taking into account inter-individual and inter-community variation, is thus essential to evaluate the degree to which human patterns are distinct. Such studies of foraging efficiency across development in chimpanzees will be complicated by the fact that age is often conflated with dominance status, and so lower foraging efficiency among young adults may be difficult to attribute to skill-based versus dominance-

based deficiencies. In this case, experimental investigations may prove the most valuable in determining capacities for foraging skill throughout chimpanzee development (see Appendix 1 for a discussion of how experimental methods complement observational methods). Specifically, comparisons between humans and chimpanzees in their acquisition of similar skills would reveal how difficult specific tasks are to master for each species. Importantly then, conclusions cannot be drawn regarding increases in the difficulty of the human foraging environment without comparable data from an “ancestral” form with which to document those relative increases. At present, few comparisons can be drawn regarding the ontogeny of these capacities in chimpanzees or other non-human primate taxa.

The complexities of the human and chimpanzee social environments may also be more similar than initially assumed. Bogin (2010) focuses on the difficulty for human adolescents in acquiring the socio-sexual practices of adults, arguing that maintaining a physical appearance of sexual maturity together with reduced fecundity is a valuable tool for adolescents to acquire skills of adult “sexual politics.” However, a period of adolescent subfecundity and lower mating success is also present among great apes (Knott, 2001), with chimpanzee females reaching sexual maturity at approximately 9 to 10 years in the wild, but typically not giving birth until many years later, between the ages of 13 and 15 years (Nishida, 2012; Pusey, 1990; Wallis, 1997). Though it is unclear to what degree this period of adolescent subfecundity is influenced by hormonal versus behavioral factors (Muller, Emery Thompson, & Wrangham, 2006; Nishida et al., 2003; Pusey, 1990), these findings indicate that the phenomenon of adolescent subfecundity is not unique to human females, calling into question whether it functions to accommodate social dynamics specific to our species. Among male chimpanzees meanwhile, there is also an important period of transition during adolescence, with younger males receiving

fewer opportunities to copulate with desirable females and encountering significant levels of aggression from older, more dominant males as they begin to travel more frequently with other males instead of with their mothers (Nishida, 2012; Pusey, 1990). These findings indicate that there are numerous complexities in the adolescent chimpanzee social world, again with additional data required to document this period in more detail in chimpanzees and other apes. The larger point here is that comparable data collected on human and non-human primate development will best enable us to best contextualize the unique features of human developmental patterns.

COMPARATIVE STUDIES OF BEHAVIORAL AND COGNITIVE DEVELOPMENT

In modeling “ancestral” patterns of behavioral development to juxtapose with those of humans, it will be valuable to collect data not only from chimpanzees but also more broadly across non-human primates. In particular, comparative developmental studies, conducted with identical methods across multiple species, may prove most useful in illuminating whether changes in adult foraging or social complexity precipitate changes in behavioral and cognitive development. Comparisons between closely-related species pairs where there are specific predictions regarding the nature of the developmental change will provide insight into the means by which behavior, cognition, and life history on the whole evolve across the primate lineage (MacLean et al., in press). Broader comparative research can also help to address the singularity problem in human evolution by illuminating the factors that influence evolutionary change more generally in order to understand the origins of the human lineage in particular.

One example where this type of comparative inquiry has already been conducted is in comparisons of chimpanzees and humans’ “other” closest living relative, the bonobo. Bonobos

and chimpanzees are equally genetically related to humans, and themselves diverged as recently as 850,000 years ago (Ruvolo, 1997; Satta, Klein, & Takahata, 2000; Won & Hey, 2005).

Despite their genetic similarity, bonobos and chimpanzees are notably distinct from one another in terms of their morphology, physiology, behavior, and cognition (Heilbronner, Rosati, Stevens, Hare, & Hauser, 2008; Rosati, Stevens, Hare, & Hauser, 2007; Shea, 1983; Wobber et al., 2010).

Behavioral and cognitive differences between the two species are proposed to derive from increases in the predictability of the bonobo foraging environment relative to that of chimpanzees, with these differences in habitat facilitating the heightened gregariousness observed among female bonobos relative to female chimpanzees and the significant reductions in aggression among male bonobos relative to male chimpanzees (Furuichi, 2009, 2011; Hare, Wobber, & Wrangham, in press; Wrangham, 1986). Bonobos have also been found to show lower levels of inhibitory control and less proficiency in areas such as tool use relative to chimpanzees (Herrmann, Hare, Call, & Tomasello, 2010; Rosati, et al., 2007), suggesting substantive changes in cognitive capacities between the two species in addition to their behavioral differences. Bonobos and chimpanzees therefore provide a compelling case in which to examine whether species differences in behavior are underlain by shifts in behavioral development. In addition, their known differences in aspects of morphological development (Durrleman, Pennec, Trouve, Ayache, & Braga, 2012; Lieberman, Carlo, Ponce de Leon, & Zollikofer, 2007; Shea, 1984) provide a preliminary indication that ontogenetic patterns have shifted in their recent evolutionary history.

A series of studies investigating behavioral and cognitive development among bonobos and chimpanzees provides evidence that their adult differences in behavior and cognition arise via shifts in developmental trajectory. First, adult bonobos' greater levels of inter-individual

tolerance relative to adult chimpanzees were found to derive from a retention of juvenile tolerance levels into adulthood (Wobber, Wrangham, & Hare, 2010) (Chapter 3). Second, lower levels of social inhibitory control among adult bonobos were underlain by slower acquisition of these capacities throughout bonobo development relative to chimpanzees (Wobber, Wrangham, et al., 2010) (Chapter 3). Finally, developmental differences between bonobos and chimpanzees were also found in a number of physical cognition tasks pertinent to foraging, with less skillful performance by adult bonobos in these areas relative to chimpanzees owing to slower rates of acquisition of these capacities throughout bonobo ontogeny (Rosati & Hare, submitted; Wobber, Herrmann, Hare, Wrangham, & Tomasello, in preparation). Taken together, these results support the hypothesis that changes in the dynamics of the adult environment can importantly influence the trajectories of behavioral and cognitive development across species. In the case of bonobos and chimpanzees in particular, differences in behavior and cognition among adults were underlain by changes in bonobo development relative to that of chimpanzees. Importantly, the notion of changes in the bonobo lineage presumes a chimpanzee-like last common ancestor of the two species (Wrangham & Pilbeam, 2001).

These comparisons suggest that behavioral differences between species can arise via shifts in behavioral development, and that such differences can emerge without any changes in the absolute length of the juvenile period. Bonobos and chimpanzees are similar in their timing of somatic and reproductive maturation, reaching menarche and first giving birth at comparable ages (Hashimoto, 1997; Knott, 2001; Kuroda, 1989; Nishida, et al., 2003; Wallis, 1997). This suggests that even between two species with similar general life history patterns, changes occurred in the rate of behavioral and cognitive development independent from overall trajectories of somatic or reproductive maturation. These findings thus urge caution in the use of

somatic or reproductive parameters to infer concurrent behavioral or cognitive development. Moreover, for the case of human life history, these results indicate that behavioral or cognitive change between species can be achieved through shifts in the rate of development, independent from changes in the duration of juvenility.

EVOLUTION OF DEVELOPMENTAL TRAJECTORIES

If we presume that, after comparable developmental data is collected, human foraging and sociality is ultimately more complex than that of chimpanzees, the second problematic assumption of the embodied-capital and adolescent-apprentice models still remains: namely, that greater skills in humans demand an extension of immaturity. Implicitly, this assumption presumes that the rate of skill learning must be fixed across taxa, or at least between humans and chimpanzees (see Figure 2.1) – though again, the rate of development is not explicitly discussed in either model. Here it is valuable to consider evidence from the study of ontogeny in skeletal traits suggesting that changes in development can occur through means other than an extension of the developmental endpoint.

Research in embryological and skeletal development has extensively investigated the evolution of ontogenetic timing, or heterochrony. Studies in this area have quantified the means by which traits change in their size and shape over development, determining changes in developmental trajectories in a descendant species through comparison with patterns in an ancestral species. Researchers in this area have classified three main ways in which developmental trajectories can vary across species, through changes in: 1) the start or end points of development, 2) the rate of development, and/or 3) the relationship between size and shape of a trait throughout development, with precise definitions created for each of these combinations

(Alberch, Gould, Oster, & Wake, 1979; Gould, 1977; Klingenberg, 1998). In the case of human evolution, evidence has shown that there are certain aspects of our skeletal anatomy that represent paedomorphosis, or a retention of ancestrally juvenile characteristics into adulthood, in comparison to extant apes (Lieberman, 2012; Penin, Berge, & Baylac, 2002; Shea, 1989). In contrast, other features of our anatomy display peramorphosis, developing beyond the adult ancestral form, with this owing both to the prolongation of the human developmental period but also to changes in the rate of ontogeny and size-shape relationships in humans (de Leon & Zollikofer, 2001; L. Godfrey & Sutherland, 1996; Vinicius, 2005).

What the heterochrony framework suggests for behavioral development is that there are multiple means by which to alter ontogenetic trajectories beyond prolonging development. In fact, research utilizing this framework has supported the notion that behavior can evolve through shifts in the rate of behavioral development. In a population of experimentally-selected mice bred for low levels of aggression, Gariepy and colleagues (Gariepy, Bauer, & Cairns, 2001) found that the rate of acquisition of both aggressive and defensive signals was slowed among the descendant, or selected line, leading to lower levels of these signals expressed among selected adults. Thus in this case selection acted on adult forms of a given behavior (aggression level), and changes occurred in that behavior through shifts in its rate of development – individuals continued to change in their signal frequency for the same duration of developmental time, but did so with changes of differing magnitude (Gariepy, et al., 2001).

In the case of human life history evolution then, a key insight from heterochrony research is that increased complexity in adult behavior or cognition could be achieved not only by an extension of the juvenile period, in changing the offset of development, but also through changes in the rate of development. If the rate of behavioral and cognitive development were accelerated

in human ontogeny, humans could achieve greater skill learning in the same amount of ontogenetic time (with some domains potentially enriched to a greater degree than others). I model this possibility in Figure 2.3, terming it the “social learning” model. As I will discuss below, the addition of extensive mechanisms for social learning in human development provides a means by which this model scenario could represent a realistic possibility in human evolution.

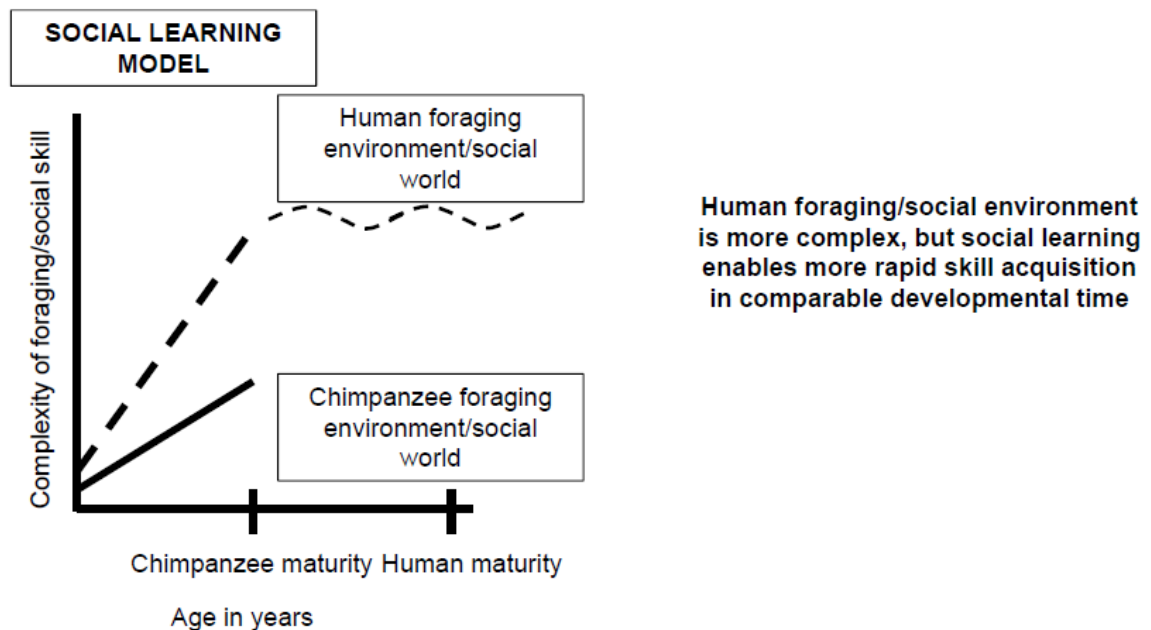


Figure 2.3. Depiction of the “social learning” model. As in Figure 2.1, age is depicted on the x-axis, and the y-axis denotes behavioral or cognitive “complexity.” Human development is indicated by the dotted line and chimpanzee development by the solid line. In this model, the human foraging/social environment is presumed to be more complex than that of chimpanzees. Under this presumption, I model the possibility that a more rapid rate of skill acquisition in humans would enable humans to master the skills necessary for adulthood in the same amount of developmental time as chimpanzees. As in Figure 2.2, this leaves the open to question the function of the prolonged human juvenile period, again depicted by the curved line.

Support for the notion that human cognitive ontogeny is accelerated in comparison to that of our closest living relatives comes from a comparison of early cognitive development between young humans, chimpanzees, and bonobos. This comparison presented identical cognitive tasks, spanning broad-reaching aspects of social and physical cognition, to same-age individuals of the three species. The results demonstrated that human children showed accelerated development in

their acquisition of cognitive skill across domains, with particularly early emergence of socio-cognitive skills in comparison to same-age *Pan* (Wobber, Herrmann, Hare, Wrangham, & Tomasello, submitted) (Chapter 6). Capacities to comprehend others' goals and intentions developed particularly early on in human ontogeny relative to that of *Pan*, suggesting the intriguing possibility that these early-emerging abilities to learn from others accelerate children's rate of skill acquisition on the whole throughout infancy and juvenility. This finding is consistent with a large body of literature comparing human children to chimpanzee adults demonstrating that children exhibit extensive capacities and motivations for social learning that are absent in chimpanzees (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007; Horner & Whiten, 2005; Tomasello, Carpenter, Call, Behne, & Moll, 2005). These findings suggest that initial capacity formation proceeds more rapidly in human than in chimpanzee development, prompting additional inquiry to determine whether the rate of skill mastery in complex tasks is also more rapid in human ontogeny.

The finding that human children show accelerated cognitive development in comparison to same-age *Pan* individuals supports the argument that behavioral and cognitive differences between species can emerge through shifts in the rate of behavioral and cognitive development, together with the comparisons of bonobos and chimpanzees discussed above. These findings support the proposal of the embodied-capital and adolescence-apprenticeship models that there have been substantive changes in patterns of behavioral and cognitive development in humans relative to other apes. However, these findings contradict the existing models in suggesting that such changes occurred in the rate of development, without necessitating a prolongation of development per se. Thus while humans are able to use our extended immaturity for ample behavioral and cognitive maturation, this maturation may not have been responsible for the

prolongation of these early life stages. This provokes the question of what selective pressure was responsible for the extension of the human juvenile period, with humans' slow growth in juvenility potentially owing to factors such as averting ecological risk (C. H. Janson & Van Schaik, 2002) or contributing to pooled energy budgets involved in cooperative breeding (Kramer & Ellison, 2010), to name a few. Further data is needed in order to fully evaluate the relative contribution of behavioral and cognitive ontogeny in shaping life history parameters across taxa.

DIRECTIONS FOR FUTURE WORK

Indeed, this is a crucial direction for future work in identifying the means by which patterns of behavioral and cognitive development evolve. This will help us to determine 1) whether trajectories of behavioral or cognitive development scale with trajectories of reproductive or somatic maturation, and in what cases they are dissociated, 2) how developmental trajectories of multiple traits interact, for example with energetic demands potentially constraining cognitive ontogeny, and 3) whether aspects of complexity in the adult environment affect the amount of time spent in juvenility across taxa, or whether rates of development in juvenility vary substantively to match these changes in complexity. Such inquiry will help us to better link multiple fields of research and to place studies of non-human primate behavioral and cognitive development into a broader theoretical framework of understanding the evolution of life history patterns. In the case of human evolution, comparative data from non-human apes will be essential to understand the degree to which we are unique in our life history patterns and to determine the selection pressures instrumental to human origins.

Chapter 2 Literature Cited.

- Alberch, P., Gould, S., Oster, G., & Wake, D. (1979). Size and shape in ontogeny and phylogeny. *Paleobiology*, 5(3), 296-317.
- Anderson, S. E., Dallal, G. E., & Must, A. (2003). Relative weight and race influence average age at menarche: Results from two nationally representative surveys of US girls studied 25 years apart. *Pediatrics*, 111(4), 844-850.
- Bogin, B. (2010). Evolution of Human Growth. In M. Muehlenbein (Ed.), *Human Evolutionary Biology* (pp. 379-395). Cambridge, UK: Cambridge University Press.
- Bogin, B., & Smith, B. (1996). Evolution of the Human Life Cycle. *American Journal of Human Biology*, 8, 703-716.
- Brannon, E. (2006). The representation of numerical magnitude. *Current Opinion in Neurobiology*, 16(2), 222-229.
- Cheney, D. L., & Seyfarth, R. M. (1988). Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, 36, 477-486. doi: 10.1016/s0003-3472(88)80018-6
- Coe, C. L., Connolly, A. C., Kraemer, H. C., & Levine, S. (1979). Reproductive development and behavior of captive female chimpanzees. *Primates*, 20(4), 571-582. doi: 10.1007/bf02373439
- Dally, J. M., Emery, N. J., & Clayton, N. S. (2006). Food-caching western scrub-jays keep track of who was watching when. *Science*, 312(5780), 1662-1665.
- de Leon, M. S. P., & Zollikofer, C. P. E. (2001). Neanderthal cranial ontogeny and its implications for late hominid diversity. *Nature*, 412(6846), 534-538.
- De Waal, F., & Lanting, F. (1997). *Bonobo: the forgotten ape*. Berkeley: University of California Press.
- Dean, C., Leakey, M. G., Reid, D., Schrenk, F., Schwartz, G. T., Stringer, C., & Walker, A. (2001). Growth processes in teeth distinguish modern humans from *Homo erectus* and earlier hominins. *Nature*, 414(6864), 628-631. doi: 10.1038/414628a
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567-631. doi: 10.1146/annurev.neuro.22.1.567
- Durrleman, S., Pennec, X., Troune, A., Ayache, N., & Braga, J. (2012). Comparison of the endocranial ontogenies between chimpanzees and bonobos via temporal regression and spatiotemporal registration. *Journal of Human Evolution*, 62(1), 74-88. doi: 10.1016/j.jhevol.2011.10.004

- Emery, N. (2006). Cognitive ornithology: the evolution of avian intelligence. *Philosophical Transactions of the Royal Society B*, 361, 23-43.
- Emery, N., & Clayton, N. (2001). Effects of experience and social context on prospective caching strategies in scrub jays. *Nature*, 414, 443-446.
- Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science*, 306(5703), 1903-1907. doi: 10.1126/science.1098410
- Fairbanks, L., & Pereira, M. (2002). *Juvenile Primates: Life History, Development, and Behavior*. Chicago: University of Chicago Press.
- Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, 15(5), 447-452.
- Furuichi, T. (2009). Factors underlying party size differences between chimpanzees and bonobos: a review and hypotheses for future study. *Primates*, 50(3), 197-209.
- Furuichi, T. (2011). Female Contributions to the Peaceful Nature of Bonobo Society. *Evolutionary Anthropology*, 20(4), 131-142. doi: 10.1002/evan.20308
- Gariepy, J., Bauer, D., & Cairns, R. (2001). Selective breeding for differential aggression in mice provides evidence for heterochrony in social behaviours. *Animal Behaviour*, 61, 933-947.
- Godfrey, L., & Sutherland, M. (1996). Paradox of peramorphic paedomorphosis: heterochrony and human evolution. *American Journal of Physical Anthropology*, 99(1), 17-42.
- Godfrey, L. R., Samonds, K. E., Jungers, W. L., & Sutherland, M. R. (2001). Teeth, brains, and primate life histories. *American Journal of Physical Anthropology*, 114(3), 192-214. doi: 10.1002/1096-8644(200103)114:3<192::aid-ajpa1020>3.0.co;2-q
- Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: Harvard University Press.
- Gould, S. (1977). *Ontogeny and Phylogeny*. Cambridge: Harvard University Press.
- Gruber, T., Muller, M. N., Strimling, P., Wrangham, R., & Zuberbuhler, K. (2009). Wild Chimpanzees Rely on Cultural Knowledge to Solve an Experimental Honey Acquisition Task. *Current Biology*, 19(21), 1806-1810. doi: 10.1016/j.cub.2009.08.060
- Gurven, M., Kaplan, H., & Gutierrez, M. (2006). How long does it take to become a proficient hunter? Implications for the evolution of extended development and long life span. *Journal of Human Evolution*, 51(5), 454-470. doi: 10.1016/j.jhevol.2006.05.003
- Hare, B. (2009). What is the effect of affect on bonobo and chimpanzee problem solving? In A. Berthoz & Y. Christen (Eds.), *The Neurobiology of the Umwelt: how living beings perceive the world* (pp. 89-102): Springer Press.

- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, 298, 1634-1636.
- Hare, B., & Kwetuenda, S. (2010). Bonobos voluntarily share their own food with others. *Current Biology*, 20(5), R230-R231.
- Hare, B., Wobber, V., & Wrangham, R. (in press). The self-domestication hypothesis: bonobo psychology evolved due to selection against aggression. *Animal Behaviour*.
- Hashimoto, C. (1997). Context and Development of Sexual Behavior of Wild Bonobos (*Pan paniscus*) at Wamba, Zaire. *International Journal of Primatology*, 18(1), 1-21.
- Hawkes, K., O'Connell, J., Blurton Jones, N., Alvarez, H., & Charnov, E. (1998). Grandmothering, menopause, and the evolution of human life histories. *Proceedings of the National Academy of Sciences*, 95, 1336-1339.
- Heilbrunner, S., Rosati, A., Stevens, J., Hare, B., & Hauser, M. (2008). A fruit in the hand or two in the bush? Divergent risk preferences in chimpanzees and bonobos. *Biology Letters*, 4(3), 246-249.
- Herrmann, E., Call, J., Hernandez-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis. *Science*, 317, 1360-1366.
- Herrmann, E., Hare, B., Call, J., & Tomasello, M. (2010). Differences in the Cognitive Skills of Bonobos and Chimpanzees. *PLOS One*, 5(8), e12438.
- Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, 8(3), 164-181.
- Janson, C., & Byrne, R. (2007). What wild primates know about resources: opening up the black box. *Animal Cognition*, 10(4), 357-367.
- Janson, C. H., & Van Schaik, C. (2002). Ecological risk aversion in juvenile primates: slow and steady wins the race. In M. Pereira & L. Fairbanks (Eds.), *Juvenile Primates: Life History, Development, and Behavior*. Chicago: University of Chicago Press.
- Jones, N. B., & Marlowe, F. W. (2002). Selection for delayed maturity - Does it take 20 years to learn to hunt and gather? *Human Nature-an Interdisciplinary Biosocial Perspective*, 13(2), 199-238.
- Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: evidence for "fast mapping". *Science*, 304, 1682-1683.
- Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology*, 9(4), 156-185.

- Klingenberg, C. (1998). Heterochrony and allometry: the analysis of evolutionary change in ontogeny. *Biological Review*, 73, 79-123.
- Knott, C. (2001). Female reproductive ecology of the apes. In P. Ellison (Ed.), *Reproductive ecology and human evolution* (pp. 429-464). New York: Aldine de Gruyter.
- Kramer, K. L. (2005). Children's help and the pace of reproduction: Cooperative breeding in humans. *Evolutionary Anthropology*, 14(6), 224-237. doi: 10.1002/evan.20082
- Kramer, K. L., & Ellison, P. T. (2010). Pooled Energy Budgets: Resituating Human Energy Allocation Trade-offs. *Evolutionary Anthropology*, 19(4), 136-147. doi: 10.1002/evan.20265
- Kuroda, S. (1989). Developmental retardation and behavioral characteristics of pygmy chimpanzees. In P. Heltne & L. Marquardt (Eds.), *Understanding chimpanzees* (pp. 184-193). Cambridge, Massachusetts: Harvard University Press.
- Laden, G., & Wrangham, R. (2005). The rise of the hominids as an adaptive shift in fallback foods: Plant underground storage organs (USOs) and australopith origins *Journal of Human Evolution*, 49(4), 482-498.
- Lancaster, J., & Kaplan, H. (2010). Embodied capital and extra-somatic wealth in human evolution and human history. In M. Muehlenbein (Ed.), *Human Evolutionary Biology* (pp. 439-455). Cambridge: Cambridge University Press.
- Lefebvre, L., Nicolakakis, N., & Boire, D. (2002). Tools and brains in birds. *Behaviour*, 139, 939-973. doi: 10.1163/156853902320387918
- Lefebvre, L., Reader, S., & Sol, D. (2004). Brains, Innovations and Evolution in Birds and Primates. *Brain Behavior and Evolution*, 63, 233-246.
- Leigh, S. R. (1992). Patterns of variation in the ontogeny of primate body size dimorphism. *Journal of Human Evolution*, 23(1), 27-50.
- Leigh, S. R. (2001). Evolution of human growth. *Evolutionary Anthropology*, 10(6), 223-236.
- Lieberman, D. E. (2012). *The Evolution of the Human Head*. Cambridge, MA: Harvard University Press.
- Lieberman, D. E., Carlo, J., Ponce de Leon, M., & Zollikofer, C. (2007). A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. *Journal of Human Evolution*, 52, 647-662.
- Liu, X., Somel, M., Tang, L., Yan, Z., Jiang, X., Guo, S., . . . Khaitovich, P. (in press). Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques. *Genome Research*.

- Locke, J. L., & Bogin, B. (2006). Language and life history: A new perspective on the development and evolution of human language. *Behavioral and Brain Sciences*, 29(3), 259-280. doi: 10.1017/s0140525x0600906x
- Lwanga, J. S., Struhsaker, T. T., Struhsaker, P. J., Butynski, T. M., & Mitani, J. C. (2011). Primate Population Dynamics Over 32.9 Years at Ngogo, Kibale National Park, Uganda. *American Journal of Primatology*, 73(10), 997-1011. doi: 10.1002/ajp.20965
- MacLean, E., Matthews, L., Hare, B., Nunn, C., Anderson, R., Aureli, F., . . . Wobber, V. (in press). How Does Cognition Evolve?: Phylogenetic Comparative Psychology. *Animal Cognition*.
- Marlowe, F. W. (2001). Male contribution to diet and female reproductive success among foragers. *Current Anthropology*, 42(5), 755-760. doi: 10.1086/323820
- Martin, P., & Bateson, P. (1986). *Measuring behavior*. Cambridge: Cambridge University Press.
- Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science*, 311(5765), 1297-1300.
- Mikkelsen, T. S., Hillier, L. W., Eichler, E. E., Zody, M. C., Jaffe, D. B., Yang, S. P., . . . Waterston, R. H. (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, 437(7055), 69-87.
- Muller, M., Emery Thompson, M., & Wrangham, R. (2006). Male Chimpanzees Prefer Mating with Old Females. *Current Biology*, 16(22), 2234-2238.
- Nishida, T. (2012). *Chimpanzees of the Lakeshore: Natural History and Culture at Mahale*. Cambridge: Cambridge University Press.
- Nishida, T., Corp, N., Hamai, M., Hasegawa, T., Hiraiwa-Hasegawa, M., Hosaka, K., . . . Zamma, K. (2003). Demography, Female Life History, and Reproductive Profiles Among the Chimpanzees of Mahale. *American Journal of Primatology*, 59(3), 99-121.
- O'Connell, J. F., Hawkes, K., & Blurton Jones, N. (2002). Meat-eating, grandmothering, and the evolution of early human diets. In P. Ungar & M. Teaforde (Eds.), *Human Diet: Its Origin and Evolution* (pp. 49-60). Westport, CT: Bergin and Garvey.
- Parish, A., & de Waal, F. (2000). The Other "Closest Living Relative" How Bonobos (*Pan paniscus*) Challenge Traditional Assumptions about Females, Dominance, Intra- and Intersexual Interactions, and Hominid Evolution. *Annals of the New York Academy of Sciences*, 907(1), 97-113.
- Penin, X., Berge, C., & Baylac, M. (2002). Ontogenetic study of the skull in modern humans and the common chimpanzees: Neotenic hypothesis reconsidered with a tridimensional procrustes analysis. *American Journal of Physical Anthropology*, 118(1), 50-62.

- Pusey, A. (1990). Behavioural Changes At Adolescence in Chimpanzees. *Behaviour*, 115(3-4), 203-246.
- Rawlins, R., & Kessler, M. (1987). *The Cayo Santiago Macaques*. New York: SUNY University Press.
- Robson, S., & Wood, B. (2008). Hominin life history: reconstruction and evolution. *Journal of Anatomy*, 212, 394-425.
- Rosati, A., & Hare, B. (submitted). Chimpanzees and bonobos exhibit divergent spatial memory development.
- Rosati, A., Stevens, J., Hare, B., & Hauser, M. (2007). The evolutionary origins of human patience: Temporal preferences in chimpanzees, bonobos, and human adults. *Current Biology*, 17(19), 1663-1668.
- Ruvolo, M. (1997). Molecular phylogeny of the hominoids: inferences from multiple independent DNA sequence data sets. *Molecular Biology and Evolution*, 14(3), 248-265.
- Salwiczek, L. H., Watanabe, A., & Clayton, N. S. (2010). Ten years of research into avian models of episodic-like memory and its implications for developmental and comparative cognition. *Behavioural Brain Research*, 215(2), 221-234. doi: 10.1016/j.bbr.2010.06.011
- Sapolsky, R. M., & Share, L. J. (2004). A pacific culture among wild baboons: Its emergence and transmission. *Plos Biology*, 2(4), 534-541. doi: e10610.1371/journal.pbio.0020106
- Satta, Y., Klein, J., & Takahata, N. (2000). DNA Archives and Our Nearest Relative: The Trichotomy Problem Revisited. *Molecular Phylogenetics and Evolution*, 14(2), 259-275.
- Savage-Rumbaugh, E. S., Shanker, S., & Taylor, T. (2001). *Apes, Language, and the Human Mind*. Oxford: Oxford University Press.
- Seed, A., Clayton, N., & Emery, N. (2008). Cooperative problem solving in rooks (*Corvus frugilegus*) *Proceedings of the Royal Society of London: B*, 275(1641), 1421-1429.
- Shea, B. (1983). Paedomorphosis and neoteny in the pygmy chimpanzee. *Science*, 222, 521-522.
- Shea, B. (1984). An allometric perspective on the morphological and evolutionary relationships between pygmy (*Pan paniscus*) and common (*Pan troglodytes*) chimpanzees. In R. Susman (Ed.), *The pygmy chimpanzee: evolutionary biology and behavior* (pp. 89-130). New York: Plenum Press.
- Shea, B. (1989). Heterochrony in human evolution: the case for neoteny reconsidered. *Yearbook of Physical Anthropology*, 32, 69-104.
- Smith, B. (1992). Life History and the Evolution of Human Maturation. *Evolutionary Anthropology*, 1(4), 134-142.

- Smith, B., Crummett, T., & Brandt, K. (1994). Ages of eruption of primate teeth: a compendium for aging individuals and comparing life histories. *Yearbook of Physical Anthropology*, 37, 177-231.
- Smith, T., Tafforeau, P., Reid, D., Pouech, J., Lazzari, V., Zermeno, J., . . . Hublin, J. J. (2010). Dental evidence for ontogenetic differences between modern humans and Neanderthals. *Proceedings of the National Academy of Sciences*, 107, 20923-20928.
- Somel, M., Franz, H., Yan, Z., Lorenc, A., Guo, S., Giger, T., . . . Khaitovich, P. (2009). Transcriptional neoteny in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 106(14), 5743-5748. doi: 10.1073/pnas.0900544106
- Speth, J. (2011). *The paleoanthropology and archaeology of big game hunting*. New York: Springer.
- Sugiyama, Y. (2004). Demographic Parameters and Life History of Chimpanzees at Bossou, Guinea. *American Journal of Physical Anthropology*, 124(2), 154-165.
- Surbeck, M., Mundry, R., & Hohmann, G. (2011). Mothers matter! Maternal support, dominance status and mating success in male bonobos (*Pan paniscus*). *Proceedings of the Royal Society B-Biological Sciences*, 278(1705), 590-598. doi: 10.1098/rspb.2010.1572
- Tomasello, M., & Call, J. (1997). *Primate Cognition*. New York: Oxford University Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675-735.
- Topal, J., Gergely, G., Erdohegyi, A., Csibra, G., & Miklosi, A. (2009). Differential Sensitivity to Human Communication in Dogs, Wolves, and Human Infants. *Science*, 325(5945), 1269-1272.
- Vinicius, L. (2005). Human encephalization and developmental timing. *Journal of Human Evolution*, 49(6), 762-776. doi: 10.1016/j.jhevol.2005.08.001
- Walker, R., Gurven, M., Hill, K., Migliano, H., Chagnon, N., De Souza, R., . . . Yamauchi, T. (2006). Growth rates and life histories in twenty-two small-scale societies. *American Journal of Human Biology*, 18(3), 295-311. doi: 10.1002/ajhb.20510
- Walker, R., Hill, K., Burger, O., & Hurtado, A. (2006). Life in the Slow Lane Revisited: Ontogenetic Separation Between Chimpanzees and Humans. *American Journal of Physical Anthropology*, 129, 577-583.
- Wallis, J. (1997). A survey of reproductive parameters in the free-ranging chimpanzees of Gombe National Park. *Journal of Reproduction and Fertility*, 109(2), 297-307.
- Warneken, F., & Tomasello, M. (2006). Altruistic Helping in Human Infants and Young Chimpanzees. *Science*, 311, 1301-1303.

- White, F., & Wood, K. (2007). Female Feeding Priority in Bonobos, *Pan paniscus*, and the Question of Female Dominance. *American Journal of Primatology*, 69(8), 837-850.
- Whiten, A., Horner, V., & Marshall-Pescini, S. (2003). Cultural Panthropology. *Evolutionary Anthropology*, 12(2), 92-105.
- Wobber, V., & Hare, B. (2009). Testing the social dog hypothesis: are dogs also more skilled than chimpanzees in non-communicative social tasks? *Behavioural Processes*, 81, 423-428.
- Wobber, V., & Hare, B. (2011). Psychological health of orphan bonobos and chimpanzees in African sanctuaries. *PLOS One*, 6(6), e17147.
- Wobber, V., Hare, B., Koler-Matznick, J., Wrangham, R., & Tomasello, M. (2009). Evidence for two waves of selection on the social skills of dogs. *Interaction Studies*, 10(2), 206-224.
- Wobber, V., Hare, B., Maboto, J., Lipson, S., Wrangham, R., & Ellison, P. (2010). Differential changes in steroid hormones before competition in bonobos and chimpanzees. *Proceedings of the National Academy of Sciences*, 107(28), 12457-12462.
- Wobber, V., Herrmann, E., Hare, B., Wrangham, R., & Tomasello, M. (in preparation). Distinctions between social and physical cognitive evolution illuminated by comparative developmental inquiry in chimpanzees and bonobos.
- Wobber, V., Herrmann, E., Hare, B., Wrangham, R., & Tomasello, M. (submitted). Differences in the early cognitive development of children and great apes.
- Wobber, V., Wrangham, R., & Hare, B. (2010). Bonobos exhibit delayed development of social behavior and cognition relative to chimpanzees. *Current Biology*, 20(3), 226-230.
- Won, Y., & Hey, J. (2005). Divergence Population Genetics of Chimpanzees. *Molecular Biology and Evolution*, 22(2), 297-307.
- Wrangham, R. (1986). Ecology and social relationships in two species of chimpanzee. In D. Rubenstein & R. Wrangham (Eds.), *Ecological aspects of social evolution* (pp. 352-378). Princeton, NJ: Princeton University Press.
- Wrangham, R., & Carmody, R. (2010). Human Adaptation to the Control of Fire. *Evolutionary Anthropology*, 19(5), 187-199. doi: 10.1002/evan.20275
- Wrangham, R., & Pilbeam, D. (2001). African apes as time machines. In B. Galdikas, N. Briggs, L. Sheeran, G. Shapiro & J. Goodall (Eds.), *All apes great and small* (Vol. 1, pp. 5-18). New York: Kluwer Academic/Plenum Publishers.
- Wright, P. C. (1999). Lemur traits and Madagascar ecology: Coping with an island environment. In C. Ruff (Ed.), *Yearbook of Physical Anthropology 1999*, Vol 42 (Vol. 42, pp. 31-72). New York: Wiley-Liss, Inc.

Zihlman, A., Cronin, J., Cramer, D., & Sarich, V. (1978). Pygmy chimpanzee as a possible prototype for the common ancestor of humans, chimpanzees, and gorillas. *Nature*, 275(5682), 744-746.

Chapter 3:
**Bonobos exhibit delayed development of social behavior
and cognition relative to chimpanzees.**

INTRODUCTION

Phenotypic changes between species can occur when evolution shapes development. Here, we tested whether differences in the social behavior and cognition of bonobos and chimpanzees derive from shifts in their ontogeny, looking at behaviors pertaining to feeding competition in particular. We found that as chimpanzees ($n = 30$) reached adulthood they became increasingly intolerant of sharing food, whereas as adults, bonobos ($n = 24$) maintained high, juvenile levels of food-related tolerance. We also investigated the ontogeny of inhibition during feeding competition. In two different tests, we found that bonobos ($n = 30$) exhibited developmental delays relative to chimpanzees ($n = 29$) in the acquisition of social inhibition, with these differences resulting in less skill among adult bonobos. The results suggest that these social and cognitive differences between two closely related species result from evolutionary changes in brain development.

Bonobos and chimpanzees differ extensively in their morphology, physiology, behavior, and cognition, despite the two species having diverged relatively recently (2.5 to 0.85 mya) (Heilbrunner, Rosati, Stevens, Hare, & Hauser, 2008; Parish & de Waal, 2000; Sannen, Heistermann, van Elsacker, Moehle, & Eens, 2003; Won & Hey, 2005). Their differences are thought to arise partly from shifts in developmental pathways. Relative to chimpanzees, bonobos have been shown to exhibit paedomorphism (retention of ancestrally juvenile traits into adulthood) in aspects of their cranial morphology (Durrleman, Pennec, Trouve, Ayache, & Braga, 2012; Lieberman, Carlo, Ponce de Leon, & Zollikofer, 2007). Bonobos also appear to retain juvenile levels of play and non-conceptive sexual behavior into adulthood, characteristics

that facilitate high inter-individual tolerance among adults when sharing food or cooperating to solve social problems (de Waal, 1987; Fruth & Hohmann, 2002; Hare, Melis, Woods, Hastings, & Wrangham, 2007; Kano, 1992; Kuroda, 1989; Palagi, 2006). However, there has been no direct test of the hypothesis that certain aspects of behavior or cognition in adult bonobos represent developmentally delayed forms of the traits found in chimpanzees. We tested this hypothesis by comparing the skills of semi free-ranging infant, juvenile and adult bonobos and chimpanzees in three tasks related to feeding competition, given the prediction that this area in particular differs between the two species.

RESULTS

Experiment 1: Inter-individual tolerance

In our first experiment, we examined inter-individual tolerance in competition for food. To assess whether bonobos' high levels of tolerance are in part a result of developmental delay, we administered a dyadic food sharing task similar to that used previously ((Hare, et al., 2007), distinctions in methodology from this prior study are described in Appendix 3) to 15 pairs of chimpanzees and 12 pairs of bonobos of varying age (mean dyad age in years ($\pm$ SEM): bonobos = 9.0 ($\pm$ 1.1), chimpanzees = 9.3 ($\pm$ 0.8), independent samples t-test, $p = \text{NS}$).

Subjects were paired with similarly aged partners. Equal numbers of male-male, male-female, and female-female dyads were tested (details in Appendix 3, Table A3.2). Each dyad received 9 trials of a food sharing task. There were 3 trial types, varying the food configuration in terms of the degree to which food could be monopolized. For each trial two measures of tolerant feeding behavior were coded: 1) *sharing* – both subjects obtained food; and 2) *co-feeding* – subjects fed from the same food source simultaneously. Play and sexual behavior were

also coded in each trial (see Appendix 3 for supplemental experimental procedures and supplemental analysis).

Chimpanzees showed a significant negative relationship between average dyad age and both measures of tolerance, *sharing* and *co-feeding* (linear regression, *sharing*: $r^2 = 0.31$, $p = 0.03$; *co-feed*: $r^2 = 0.46$, $p = 0.006$; Figure 3.1). In contrast, in bonobos there was no correlation between dyad age and *sharing* or *co-feeding* (*sharing*: $r^2 = 0.01$, $p = \text{NS}$; *co-feed*: $r^2 = 0.15$, $p = \text{NS}$) (Figure 3.1).

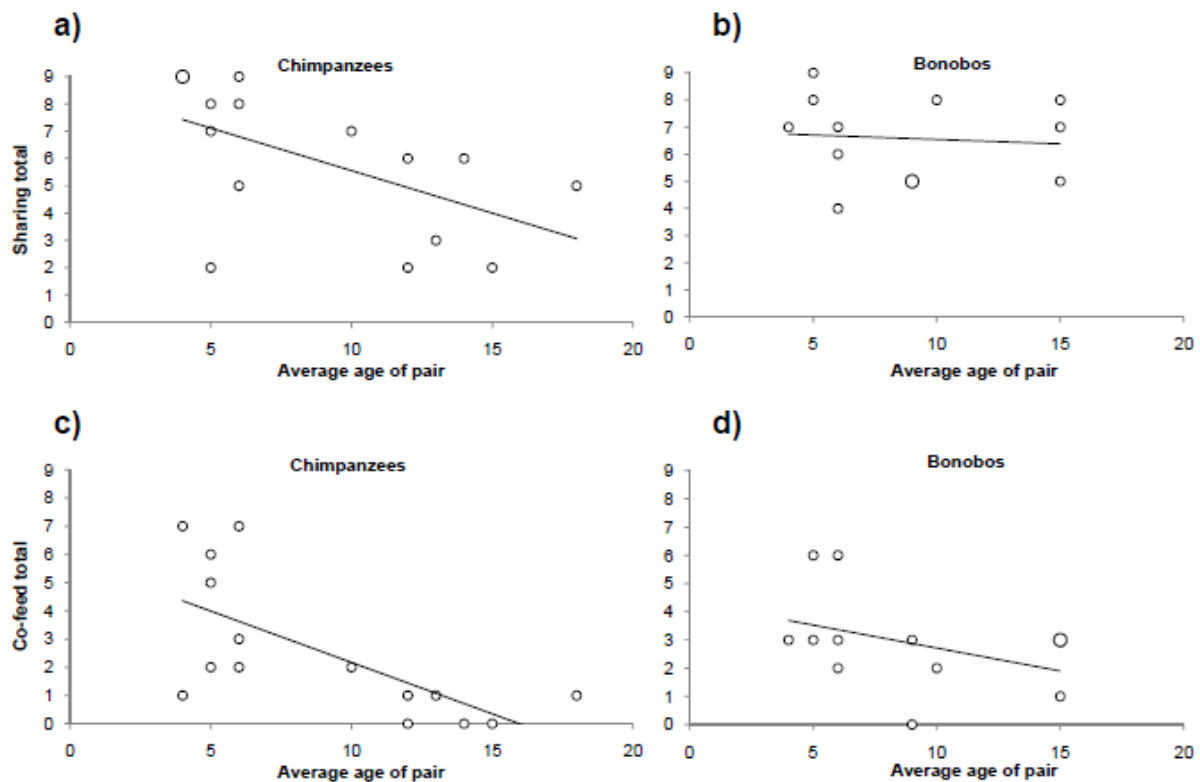


Figure 3.1 Feeding behavior according to species and age, experiment 1. a) Chimpanzees' average age of pair (dyad age) in relation to the number of trials (out of 9 total) where individuals shared food, b) bonobos' dyad age in relation to this measure, c) chimpanzees' dyad age in relation to the number of trials where they co-fed, and d) bonobos' dyad age in relation to this measure. Small circles represent one dyad while large circles represent multiple dyads with the same behavioral score.

To further probe the relationship between age and sharing we classified subjects as adults or juveniles. We defined adults as those possessing a 3rd molar at the time of testing (Smith, Crummett, & Brandt, 1994). We performed a 2x2 ANOVA of sharing with species and age category as factors, and found a significant effect of age category ($F(1,26) = 4.13$, $p = 0.05$). Post-hoc tests revealed that juvenile chimpanzees shared significantly more than adult chimpanzees (Tukey's HSD $p < 0.05$), while there was no difference in sharing between age categories of bonobos (Tukey's HSD $p > 0.05$) (Table 3.1). There was no significant difference in sharing between juvenile chimpanzees and juvenile bonobos, nor between adult chimpanzees and adult bonobos (Tukey's HSD $p > 0.05$).

Table 3.1 Performance across species and age groups in the tolerance test, experiment 1. The number of trials (out of 9 total) where individuals shared or co-fed during the food sharing task. Age groups are divided into juvenile and adult, as described in the manuscript. Means for each variable are listed with standard error in parentheses.

	Sharing	Co-feeding
Chimpanzee juveniles	7.12 (0.88)	4.12 (0.85)
Chimpanzee adults	4.43 (0.78)	0.71 (0.29)
Chimpanzee mean	5.87 (0.68)	2.53 (0.65)
Bonobo juveniles	6.83 (0.70)	3.83 (0.70)
Bonobo adults	6.33 (0.62)	2.00 (0.52)
Bonobo mean	6.58 (0.45)	2.92 (0.50)

We performed a similar ANOVA for co-feeding, and again found a significant effect of age category ($F(1,26) = 15.67$, $p = 0.001$). Post-hoc tests showed that juvenile chimpanzees co-fed significantly more than adult chimpanzees (Tukey's HSD $p < 0.01$), while there was no significant difference between age categories in bonobos (Tukey's HSD $p > 0.05$) (Table 3.1).

There was no difference between species in juvenile levels of co-feeding (Tukey's HSD $p > 0.05$), but adult bonobos co-fed significantly more than adult chimpanzees (Tukey's HSD $p < 0.05$).

Thus, both the *sharing* and *co-feeding* measures demonstrated that while chimpanzees became less tolerant as they reached adulthood, bonobos retained juvenile levels of sharing as adults. As a result bonobos were more tolerant than chimpanzees as adults (cf. (Hare, et al., 2007)). We also found that compared to chimpanzees, bonobos exhibited higher levels of play and sexual behavior, possibly facilitating their higher feeding tolerance (Appendix 3, Supplemental Analyses). Given these results, we conducted two experiments to test whether the more relaxed feeding style of bonobos is related to changes in the ontogeny of their inhibitory abilities in situations simulating feeding competition.

Experiment 2: Social Response Inhibition

In Experiment 2 we evaluated the ability of 20 infant and juvenile bonobos and 20 infant and juvenile chimpanzees to inhibit a social response (mean subject age in years ($\pm$ SEM): chimpanzees, 4.5 (± 0.3); bonobos, 4.3 (± 0.3), independent samples t-test, $p = \text{NS}$). In this task, a subject could beg for food from three human experimenters who stood shoulder-to-shoulder in front of him or her. Subjects were shown that only the outer two experimenters held a food reward. Subjects were successful if they chose these two experimenters (by touching their hands) without choosing the middle experimenter's (empty) hand, with 12 trials performed. This problem resembles what young apes can experience during competition over meat or attractive plant foods where individuals must inhibit the desire to beg from or feed near certain intolerant group members. We classify it as a social problem because subjects could use the identity or location of the experimenters as cues to the food location.

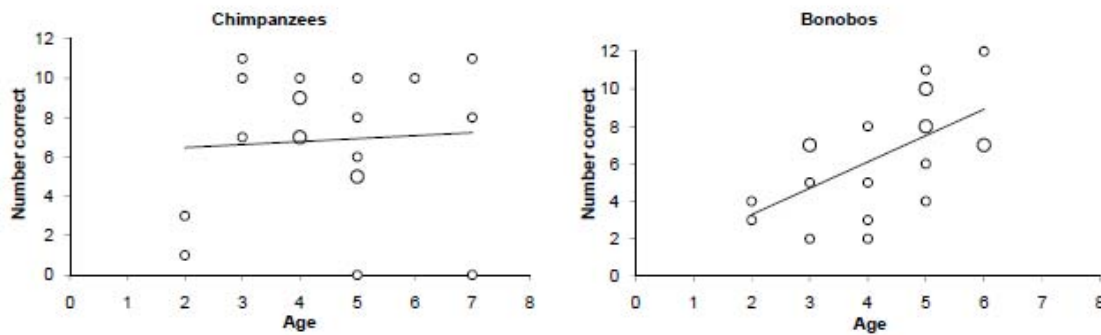


Figure 3.2 Social inhibition according to species and age, experiment 2. The relationship is shown between each subject's age and its overall number of correct choices in the 12 social response inhibition test trials. The small circles represent the performance of a single subject while the large circles represent multiple individuals.

Bonobos exhibited a significant positive relationship between age and performance on the test (linear regression, $r^2 = 0.35$, $p = 0.006$; Figure 3.2), while the performance of chimpanzees did not correlate with age ($r^2 = 0.06$, $p = \text{NS}$; Figure 3.2). We also performed a 2x2 ANOVA with species and age category as factors, classifying subjects as either pre-weaning (2-4 years, $N=10$ per species) or post-weaning (5-7 years, $N=10$ per species), based on the weaning age of 4-4.5 years observed in wild chimpanzees and bonobos (Goodall, 1986; Kuroda, 1989). There was no main effect of species or age category on test performance, but there was a significant species x age category interaction ($F(1,36) = 6.31$, $p = 0.02$). Post-hoc comparisons revealed that post-weaning individuals of the two species performed at similar levels (Tukey's HSD $p > 0.05$) (Table 3.2). However, pre-weaning bonobos performed less skillfully than post-weaning bonobos (Tukey's HSD $p < 0.01$), and pre-weaning chimpanzees (Tukey's HSD $p < 0.05$). In contrast, pre-weaning chimpanzees performed as well as post-weaning chimpanzees (Tukey's HSD $p > 0.05$) (Table 3.2).

Table 3.2. Performance across species and age groups in the social response inhibition task, experiment 2. There were 4 introduction trials and 12 test trials performed. Age groups are divided into pre- and post-weaning, as described in the manuscript. Means for each variable are listed with standard error in parentheses.

	Introduction	Test
Pre-weaning chimpanzees	2.80 (0.47)	7.40 (1.01)
Post-weaning chimpanzees	3.20 (0.29)	6.30 (1.24)
Chimpanzee mean	3.00 (0.27)	6.85 (0.79)
Pre-weaning bonobos	3.20 (0.20)	4.60 (0.69)
Post-weaning bonobos	3.30 (0.26)	8.30 (0.78)
Bonobo mean	3.25 (0.16)	6.45 (0.66)

Thus, our findings demonstrate a species difference in the ontogeny of inhibitory control, with a delay in bonobo development relative to that of chimpanzees. Bonobos took longer to develop the same skill level shown even among the youngest chimpanzees tested. Controls revealed no evidence for significant species differences in motivation or attention, while a second estimate of subject age (weight) revealed the same pattern of results as above and removal of outliers did not change the results (Appendix 3, Supplemental Analyses).

However, this task appeared to be relatively simple, given that only the pre-weaning bonobos struggled. Since post-weaning individuals of both species performed similarly, the two species could in theory develop social inhibitory control at different rates but have similar skills as adults. To test this, we presented a slightly older group of bonobos and chimpanzees with a social inhibitory task that was cognitively more demanding.

Experiment 3: Social Reversal Learning

In Experiment 3 we evaluated the ability of subjects to adjust to changes in the sharing behavior of two experimenters in a social reversal learning paradigm. 17 bonobos and 11

chimpanzees participated (mean age in years ($\pm$ SEM): chimpanzees, 9.8 ($\pm$ 1.4); bonobos, 10.2 ($\pm$ 1.4), independent samples t-test, $p = \text{NS}$).

Subjects chose between two human experimenters, only one of whom held a concealed food reward, until they learned that one human consistently held the food (to the level of 84% correct, see (Rumbaugh & Pate, 1984)). After reaching this introductory learning criterion subjects immediately received 20 reversal trials where the experimenter hiding the reward was switched. The experimenter who reliably shared food in the introduction now always had no food while the other previously “stingy” experimenter would now always share (Wobber & Hare, 2009). After this switch, we recorded the number of trials in which subjects chose the newly generous experimenter.

As a control for whether the two species were equally engaged in the task, we first assessed performance on the introductory trials. The two species did not differ in the number of trials it took them to reach the 84% correct criterion (independent samples t-test $p = \text{NS}$, Table 3.3). In addition, linear regression analysis showed that the number of trials needed to reach the introductory criterion did not correlate with age in either species.

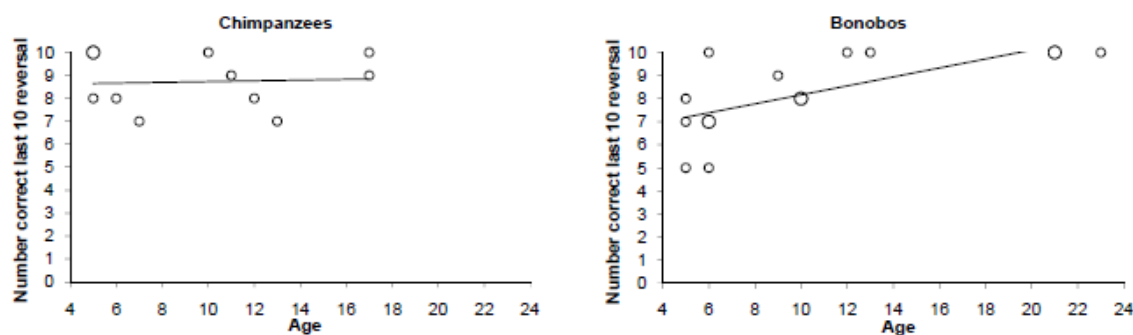


Figure 3.3 Social reversal learning according to species and age, experiment 3. The number of correct choices that subjects made in the last 10 trials of the social reversal learning test in relation to their age is shown. The small circles represent the performance of a single subject while the large circles represent multiple individuals.

In the reversal trials bonobos showed a significant positive relationship between age and performance (linear regression, $r^2 = 0.29$, $p = 0.03$), but chimpanzees did not (linear regression, $r^2 = 0.001$, $p = \text{NS}$) (Figure 3.3). We also performed a 2x2 ANOVA with species and age category as factors, dividing subjects into juveniles and adults (as in Experiment 1). This ANOVA revealed only a weak effect of species ($F(1,27) = 3.58$, $p = 0.07$), with there being a tendency for chimpanzees to outperform bonobos on the 20 trials of the reversal (Table 3.3).

Table 3.3. Performance across species and age groups in the social reversal learning task, experiment 3. The last trial of the introduction represents how many trials it took subjects to learn the introductory association to the criterion of 84% correct. For the reversal, we report performance overall and separated into the first and last ten trials. Age groups are divided into juvenile and adult, as described in the manuscript. Means for each variable are listed with standard error in parentheses.

	Last trial introduction	Reversal, first 10 trials	Reversal, last 10 trials	Reversal overall
Chimpanzee juveniles	17.40 (2.77)	8.40 (1.12)	8.60 (0.60)	17.00 (1.64)
Chimpanzee adults	25.00 (3.72)	9.00 (0.52)	8.83 (0.48)	17.83 (0.87)
Chimpanzee mean	21.50 (2.57)	8.73 (0.56)	8.73 (0.36)	17.45 (0.85)
Bonobo juveniles	22.56 (2.69)	6.89 (0.95)	7.00 (0.71)	13.89 (1.22)
Bonobo adults	16.38 (2.69)	6.75 (0.94)	9.38 (0.32)	16.12 (1.16)
Bonobo mean	19.70 (2.00)	6.82 (0.65)	8.12 (0.49)	14.94 (0.86)

We further examined performance in the reversal by looking at the first and last 10 trials separately, since subjects can have difficulty with the reverse association at first, then solve the inhibitory problem over the course of multiple trials. Regressions showed no correlation between age and performance in the first half of the test session in either species. An ANOVA of performance on the first 10 trials with species and age category as factors showed a near-significant effect of species ($F(1,27) = 3.82$, $p = 0.06$), but no effect of age category, nor a

significant interaction. Chimpanzees performed somewhat better than bonobos on these first 10 trials (Table 3.3).

In contrast, in the last 10 trials of the reversal, bonobos showed a positive relationship between age and performance ($r^2 = 0.35$, $p = 0.01$) while chimpanzees did not ($r^2 = 0.004$, $p = \text{NS}$). An ANOVA of performance on the second 10 trials demonstrated a significant effect of age category ($F(1,27) = 4.85$, $p = 0.04$), but no significant effect of species or interaction. In contrast to the pattern in the first 10 trials, there was no species difference in performance on these latter 10 trials (Table 3.3). Instead, post-hoc tests revealed that adult bonobos significantly outperformed juvenile bonobos on the last 10 trials (Tukey's HSD $p < 0.05$), while there was no difference in performance between adult and juvenile chimpanzees (Tukey's HSD $p > 0.05$) (Table 3.3).

Thus in the first ten trials of the reversal, bonobos of all ages struggled while chimpanzees of all ages performed well. In the latter half of the reversal, younger bonobos continued to have difficulty but adult bonobos adjusted and subsequently raised the species mean for these ten trials to within the range of the performance of the chimpanzees. In short, the juvenile bonobos were slower than the other individuals to adapt to the reversal, performing at a lower level in the latter reversal trials relative to juvenile chimpanzees and to adults of both species. Further, adult bonobos exhibited less social inhibitory control than adult chimpanzees, with a tendency to perform worse during the first ten trials and overall. Results were similar when using weight as a proxy for age or removing outlier individuals, and motivation levels did not differ between the two species or correlate with test performance (Appendix 3, Supplemental Analyses). Subjects who had previously participated in Experiment 2 performed no differently

from the novel subjects in their learning of the initial association or in the reversal (independent samples t-tests).

In sum, Experiment 3 tested an older sample with a relatively challenging cognitive task, and again revealed a developmental delay in bonobos relative to chimpanzees. Our evidence that the delay in the ontogeny of social inhibition in bonobos persists into adulthood resembles differences seen previously when adults of the two species were compared in a non-social inhibition task (Rosati, Stevens, Hare, & Hauser, 2007; Vlamming, Hare, & Call, 2010).

DISCUSSION

Our findings support the hypothesis that developmental delays play a role in producing differences in the social psychology underlying food competition in bonobos and chimpanzees. Inter-individual tolerance in sharing food decreased with age in chimpanzees while bonobos maintained juvenile levels of tolerance into adulthood. Infant bonobos were less capable of inhibiting themselves from begging for food than were same-age chimpanzees, with chimpanzees successful from the youngest age tested. In a social reversal learning task, juvenile and even adult bonobos were more inhibited by their previously learned social associations than chimpanzees, who again showed adult levels of performance even as juveniles. Thus in both tolerance and social inhibition, shifts in the ontogeny of behavior corresponded to distinctions between adults of the two species. Controls ruled out differences in motivation or comprehension of the tasks as plausible explanations of the observed species differences.

The association in bonobos of juvenile levels of tolerance, delayed development of social inhibition and a paedomorphic cranium suggests that a common developmental mechanism might be responsible for the retention of juvenile traits into adulthood. By analogy, populations

of mammals selected for reduced aggression tend to exhibit ontogenetic delays across numerous traits relative to their wild-type ancestors (Hemmer, 1990; Trut, Plyusnina, & Oskina, 2004). A similar process could be responsible for our findings, for example if selection against aggression in bonobos led to delays in the ontogeny of multiple other traits (Gariépy, Bauer, & Cairns, 2001; Hare, Wobber, & Wrangham, in press). This idea does not imply that bonobos are juvenilized globally. Instead, it suggests that juvenilization has occurred in a set of traits that are strongly genetically linked.

Understanding the evolutionary processes by which ontogenetic changes occurred in bonobos may provide insight into our own species' evolution. Herrmann et al. (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007) proposed that the crucial cognitive adaptation of humans relative to other apes is the accelerated development of social skills in infants. While the genetic changes that produce such developmental shifts are not well understood, if we can determine the process by which the ontogeny of bonobos evolved, inferences can be made regarding analogous evolution in our own species.

EXPERIMENTAL PROCEDURES

Experiment 1: Inter-individual tolerance

Subjects in all three experiments were tested at the Tchimpounga Chimpanzee Sanctuary in the Congo Republic and the Lola ya Bonobo Sanctuary in the Democratic Republic of Congo (Appendix 3, Table A3.1 provides a list of subjects' experimental participation. Note: the chimpanzees here were *Pan troglodytes troglodytes* not *Pan troglodytes schweinfurthii* as previously tested (Hare, et al., 2007)). For this experiment we tested 30 chimpanzees (4 to 19 years) and 24 bonobos (4 to 23 years). In all trials subjects were released into the test room

simultaneously, with food placed prior to their release. Each dyad was given three trials of each of three food configuration conditions, with one condition presented per day over the course of three separate days for a total of nine trials. All statistics for this and the subsequent experiments were two-tailed. All tests were videotaped, with behavior scored from this video. See Appendix 3 for additional details regarding the experimental procedures.

Experiment 2: Response Inhibition

Subjects in both species ranged in age from 2 to 7 years, and there were 6 female and 14 male bonobos tested, and 8 female and 12 male chimpanzees. Subjects were given one test session, consisting of three types of trials: warm-up, introduction, and test trials. In the two warm-up trials, all three experimenters held food to introduce the test paradigm and the potentially unfamiliar humans. These were followed by four introduction trials where only two adjacent experimenters held food. Finally, in the 12 test trials the two nonadjacent experimenters always held food while the center experimenter did not. The three human experimenters maintained their position relative to one another throughout the test. Only those individuals taking food in the trial reached towards the food container. Those individuals did so simultaneously in view of the subject, then all three experimenters raised their arms toward the subject simultaneously and closed their fists so that the food was not visible at the time of choice. Performance was scored live by the experimenters, though all tests were also videotaped.

Experiment 3: Reversal Learning

Chimpanzee subjects' ages ranged from 5 to 17 years and bonobo subjects' ages ranged from 5 to 23 years. There were 6 female and 11 male bonobos tested, and 7 female and 4 male chimpanzees. For this experiment, two experimenters again stood in front of the subjects, with the potential to be holding food. In the test trials, both individuals appeared to take food from a

container, but only one individual did so. The two experimenters presented their closed fists to the subject, so that it did not know who was holding food. The same experimenter held food for every trial of the introduction, and in the reversal the other experimenter always held food. The two experimenters always stood in the same position for a given subject's entire test session (with their locations counter-balanced across subjects). Subjects were given a maximum of 40 introduction trials to reach the 84% correct criterion, otherwise their test session was aborted and their performance was not included as part of the results (this occurred for 6 individuals, supplemental to the 28 individuals presented here). Performance was scored live, in addition to being videotaped. Prior to the test trials, we performed a baseline task to ensure that any preferences that subjects possessed for one of the two human experimenters did not impact results in the test. The methods and results of this baseline are discussed in Appendix 3.

Chapter 3 Literature Cited.

- de Waal, F. (1987). Tension regulation and nonreproductive functions of sex in captive bonobos (*Pan paniscus*). *National Geographic Research Reports*, 3, 318-335.
- Durrleman, S., Pennec, X., Trouve, A., Ayache, N., & Braga, J. (2012). Comparison of the endocranial ontogenies between chimpanzees and bonobos via temporal regression and spatiotemporal registration. *Journal of Human Evolution*, 62(1), 74-88. doi: 10.1016/j.jhevol.2011.10.004
- Fruth, B., & Hohmann, G. (2002). How bonobos handle hunts and harvests: why share food? In C. Boesch, G. Hohmann & L. Marchant (Eds.), *Behavioural diversity in chimpanzees and bonobos* (pp. 231-243). Cambridge: Cambridge University Press.
- Gariepy, J., Bauer, D., & Cairns, R. (2001). Selective breeding for differential aggression in mice provides evidence for heterochrony in social behaviours. *Animal Behaviour*, 61, 933-947.
- Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: Harvard University Press.
- Hare, B., Melis, A. P., Woods, V., Hastings, S., & Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology*, 17(7), 619-623.
- Hare, B., Wobber, V., & Wrangham, R. (in press). The self-domestication hypothesis: bonobo psychology evolved due to selection against aggression. *Animal Behaviour*.
- Heilbrunner, S., Rosati, A., Stevens, J., Hare, B., & Hauser, M. (2008). A fruit in the hand or two in the bush? Divergent risk preferences in chimpanzees and bonobos. *Biology Letters*, 4(3), 246-249.
- Hemmer, H. (1990). *Domestication: the Decline of Environmental Appreciation*. Cambridge: Cambridge University Press.
- Herrmann, E., Call, J., Hernandez-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis. *Science*, 317, 1360-1366.
- Kano, T. (1992). *The Last Ape: Pygmy Chimpanzee Behavior and Ecology*. Stanford, California: Stanford University Press.
- Kuroda, S. (1989). Developmental retardation and behavioral characteristics of pygmy chimpanzees. In P. Heltne & L. Marquardt (Eds.), *Understanding chimpanzees* (pp. 184-193). Cambridge, Massachusetts: Harvard University Press.
- Lieberman, D. E., Carlo, J., Ponce de Leon, M., & Zollikofer, C. (2007). A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. *Journal of Human Evolution*, 52, 647-662.

- Palagi, E. (2006). Social Play in Bonobos (*Pan paniscus*) and Chimpanzees (*Pan troglodytes*): Implications for Natural Social Systems and Interindividual Relationships. *American Journal of Physical Anthropology*, 129, 418-436.
- Parish, A., & de Waal, F. (2000). The Other “Closest Living Relative” How Bonobos (*Pan paniscus*) Challenge Traditional Assumptions about Females, Dominance, Intra- and Intersexual Interactions, and Hominid Evolution. *Annals of the New York Academy of Sciences*, 907(1), 97-113.
- Rosati, A., Stevens, J., Hare, B., & Hauser, M. (2007). The evolutionary origins of human patience: Temporal preferences in chimpanzees, bonobos, and human adults. *Current Biology*, 17(19), 1663-1668.
- Rumbaugh, D., & Pate, J. L. (1984). The evolution of cognition in primates: a comparative perspective. In H. Roitblat, T. Bever & H. Terrace (Eds.), *Animal cognition* (pp. 569-587). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sannen, A., Heistermann, M., van Elsacker, L., Moehle, U., & Eens, M. (2003). Urinary testosterone metabolite levels in bonobos: a comparison with chimpanzees in relation to social system. *Behaviour*, 140, 683-696.
- Smith, B., Crummett, T., & Brandt, K. (1994). Ages of eruption of primate teeth: a compendium for aging individuals and comparing life histories. *Yearbook of Physical Anthropology*, 37, 177-231.
- Trut, L. N., Plyusnina, I. Z., & Oskina, I. N. (2004). An experiment on fox domestication and debatable issues of evolution of the dog. *Russian Journal of Genetics*, 40(6), 644-655.
- Vlammings, P., Hare, B., & Call, J. (2010). Reaching around barriers: the performance of the great apes and 3–5-year-old children. *Animal Cognition*, 13(2), 273-285.
- Wobber, V., & Hare, B. (2009). Testing the social dog hypothesis: are dogs also more skilled than chimpanzees in non-communicative social tasks? *Behavioural Processes*, 81, 423-428.
- Won, Y., & Hey, J. (2005). Divergence Population Genetics of Chimpanzees. *Molecular Biology and Evolution*, 22(2), 297-307.

Chapter 4:
Different ontogenetic patterns of testosterone production reflect divergent male reproductive strategies in chimpanzees and bonobos

INTRODUCTION

Investment in reproduction among males can be divided into both the production of gametes and the allocation of energy towards somatic and behavioral strategies that facilitate mating opportunities (Bribiescas, 2001; Muehlenbein & Bribiescas, 2005). The steroid hormone testosterone (abbreviated as T) is particularly important in influencing these latter two elements of male reproductive strategy, increasing muscle mass, enhancing libido, and stimulating aggressive and dominance behaviors in a given season or situation (Bhasin et al., 1996; Cunningham & Huckins, 1979; Ellison, 2003; J. Wingfield, Hegner, Dufty, & Ball, 1990).

While the association between testosterone and male reproductive effort has been well-documented in adults of numerous taxa, our understanding of how development mediates this relationship is less clear. According to life history theory, the production of testosterone across ontogeny should differ between species or individuals to facilitate the optimal allocation of energy toward growth, maintenance, and reproduction across the lifespan (Bribiescas, 2001; Stearns, 1992). Since high levels of testosterone can have a deleterious effect on the immune system (Muehlenbein & Bribiescas, 2005; Zuk, Johnsen, & Maclarty, 1995), production of testosterone may be minimized in situations or life stages where it is not sufficiently advantageous (J. Wingfield, et al., 1990; J. C. Wingfield, Lynn, & Soma, 2001). Accordingly, testosterone levels typically remain low during juvenility, only beginning to increase at puberty in conjunction with reproductive maturation (Archer, 2006; Elmlinger, Kuehnel, Wormstall, & Doeller, 2005; Gesquiere et al., 2005). Despite this general developmental pattern being present across mammals, there may be important species differences in the precise patterns of

testosterone production throughout development that reflect diverging male reproductive strategies in adulthood. This possibility is particularly compelling in light of the growing body of evidence that phenotypic changes between species commonly arise through evolutionary shifts in developmental trajectories (Carroll, 2008; Wobber, Wrangham, & Hare, 2010a, 2010b).

Several studies have begun to investigate whether individual and species-level variation in the ontogeny of androgen production exist in association with differing adult reproductive strategies, using non-human primate models to examine these effects over an extended period of ontogeny. Within mandrills (*Mandrillus sphinx*) and chacma baboons (*Papio hamadryas ursinus*), individual differences in the production of testosterone during puberty have been found to correlate with dominance ranks among adult males (J. Beehner, Bergman, Cheney, Seyfarth, & Whitten, 2006; Setchell & Dixson, 2002). Similarly, in orangutans (*Pongo pygmaeus*), males who retained subadult body size into adulthood (a viable strategy in this species to obtain sneaky mating opportunities without overt physical competition) were found to show smaller increases in testosterone during adolescence than males who developed their body size fully (Maggioncalda, Sapolsky, & Czekala, 1999). In addition, differences between baboon species in the timing and magnitude of the pubertal testosterone increase have been found to reflect inter-specific variation in the length of alpha male tenure and the association between rank and mating success (J. C. Beehner et al., 2009). These results thus support the hypothesis that within and across species, variation in the developmental trajectory of androgen production is central to the relationship between testosterone and reproductive effort among adult males.

These prior studies of testosterone production throughout development have largely focused on the pubertal increase in testosterone levels, since the period of adolescence represents an important transition between an individual's focus on growth and its focus on reproduction.

However, individuals or species may also vary in their production of testosterone even earlier on in life. In a number of species, from humans to yellow baboons (*Papio cynocephalus*) and cotton-top tamarins (*Saguinus oedipus*), males and females show a neonatal elevation in testosterone that lasts for the first few weeks or even months after birth (Andersson et al., 1998; Gesquiere, et al., 2005; Ginther, Carlson, Ziegler, & Snowden, 2002). Though there is considerable debate about the function of this neonatal testosterone elevation (Mann & Fraser, 1996; Sharpe et al., 2003), one possibility is that variation in its duration or magnitude contributes to differences in reproductive capabilities among adult males (Andersson, et al., 1998; Mann, Akinbami, Gould, Paul, & Wallen, 1998).

Here we test the hypothesis that species differences in male mating strategy are associated with variation in the ontogenetic patterns of testosterone production across the entire lifespan. We do so by comparing testosterone levels from infancy into adulthood in two closely-related ape species, bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*). Chimpanzees and bonobos provide an ideal test case for this hypothesis, as they have been found to differ in both their reproductive strategy and in broader aspects of their ontogeny, despite having diverged from one another as recently as 850 kya (Won & Hey, 2005).

Differences in male reproductive strategy between chimpanzees and bonobos appear to derive largely from the increased social gregariousness and sexual receptivity of bonobo females relative to chimpanzee females, presuming that the last common ancestor of the two species was chimpanzee-like (Hare, Wobber, & Wrangham, in press; Wrangham & Pilbeam, 2001). Bonobo females associate with males more regularly than female chimpanzees, and engage in extensive non-conceptive sexual behavior throughout their menstrual cycle while the copulations of female chimpanzees are largely limited to the period of maximal sexual swelling (de Waal, 1987;

Furuichi, 2009; Mulavwa et al., 2010; Reichert, Heistermann, Hodges, Boesch, & Hohmann, 2002). Correspondingly, it has been argued that competition for dominance rank and coercive aggression are less effective means of obtaining conceptive mating opportunities for bonobo males than for male chimpanzees (Furuichi, 2011; Surbeck, Mundry, & Hohmann, 2011; Wrangham, 2002; Wrangham & Pilbeam, 2001). In support of this argument, bonobo males show less frequent and less severe forms of aggression than chimpanzee males in their intra-group aggression, inter-group aggression, and even inter-specific predation (Furuichi & Ihobe, 1994; Hare, et al., in press; Ihobe, 1997; Kano, 1992; Lwanga, Struhsaker, Struhsaker, Butynski, & Mitani, 2011; Muller, 2002; Muller, Kahlenberg, Emery Thompson, & Wrangham, 2007). Bonobos have also been found to show a lesser sex difference in androgen production relative to chimpanzees, lesser increases in male androgen levels when females are cycling than found among male chimpanzees, and a weaker correlation on the whole between basal testosterone level and dominance rank among adult males (Marshall & Hohmann, 2005; Muller & Wrangham, 2004; Sannen, Heistermann, van Elsacker, Moehle, & Eens, 2003; Surbeck, Deschner, Schubert, Weltring, & Hohmann, in press). These two species therefore provide an excellent opportunity to test whether divergent male reproductive strategies in closely-related taxa are associated with broader differences in their endocrine maturation.

In addition to their divergent reproductive strategies, bonobos and chimpanzees have been found to differ in numerous facets of their development. In particular, bonobos exhibit delays in development relative to chimpanzees in aspects of their morphology (Lieberman, Carlo, Ponce de Leon, & Zollikofer, 2007; Shea, 1984), behavior (Kuroda, 1989; Wobber, Wrangham, et al., 2010a, 2010b), and cognition (Wobber, Herrmann, Hare, Wrangham, & Tomasello, submitted). These distinctions provide support for the possibility that the ontogenetic

pattern of testosterone production has also shifted between these two species, given the evidence from numerous taxa that hormones are a central mechanism in facilitating developmental transitions (Ketterson & Nolan, 1992; McGlothlin & Ketterson, 2008).

No evidence exists at present to compare endocrine maturation between bonobos and chimpanzees, since to our knowledge no prior study of bonobo endocrine ontogeny has been performed. Existing studies of endocrine maturation in chimpanzees have consistently found that male testosterone levels increase with age, with these increases coinciding with growth in body weight and testicular volume (Anestis, 2006; Kondo et al., 2000; Marson, Meuris, Cooper, & Jouannet, 1991b; Martin, Swenson, & Collins, 1977; Nadler, Wallis, Rothmeyer, Cooper, & Baulieu, 1987; Seraphin, Whitten, & Reynolds, 2008; Winter, Faiman, Hobson, & Reyes, 1980; Young, Gould, & Smithwick, 1993). In captive populations, male chimpanzees begin to show elevations in testosterone between 6 to 7 years of age (Kondo, et al., 2000; Marson, et al., 1991b; Martin, et al., 1977; Winter, et al., 1980; Young, et al., 1993), with puberty and the onset of spermatogenesis occurring between 7 and 9 years of age (Blank & Murphy, 1991; Marson, Meuris, Cooper, & Jouannet, 1991a). The only existing study of testosterone development among a small sample of wild chimpanzees indicates a similar developmental increase, occurring at a slightly later age (Seraphin, et al., 2008). Despite the relatively large number of studies documenting patterns of testosterone production throughout chimpanzee development, few have incorporated individuals from a broad developmental window (encompassing infancy, juvenility, adolescence, and adulthood). Moreover, these studies have primarily been conducted in laboratory environments, where asocial or minimally social housing conditions may have minimized any effects of dominance rank or social behavior on testosterone production. This study represents one of the first opportunities to study testosterone production in chimpanzees

ranging from infancy into adulthood, utilizing semi free-ranging study populations where individuals live in mixed-age and sex groups closely resembling those found in the wild (Wobber & Hare, 2011).

Our major prediction was that bonobos would show a lesser developmental increase in testosterone production than chimpanzees, given their lesser mating competition as adults and their maintenance of a juvenile phenotype into adulthood in numerous traits (Wobber, Wrangham, et al., 2010a; Wrangham & Pilbeam, 2001). Our alternative hypothesis was that bonobos and chimpanzees would differ little in their ontogeny of testosterone production, given their genetic similarity. We tested these predictions by measuring salivary testosterone levels from infancy into adulthood among bonobos and chimpanzees, making it possible for us to examine the contributions of both neonatal and pubertal testosterone elevations to the overall trajectory of testosterone production in each species. We examined testosterone in both sexes to assess the degree to which male patterns of development diverged from those of females.

MATERIALS AND METHODS

Subjects

Subjects for this research were chimpanzees living at the Tchimpounga Chimpanzee Sanctuary in Pointe Noire, Congo Republic and bonobos living at Lola ya Bonobo in Kinshasa, Democratic Republic of Congo. Both facilities house semi free-ranging ape populations living in mixed age and sex groups that have access to forest enclosures during the day and sleep in dormitories at night. Apes at these sites are provisioned but have access to natural food items in their primary forest enclosures. Although these apes are largely orphans of the bushmeat trade, their behavior patterns and cognitive abilities are typical of captive apes (Wobber & Hare, 2011).

In addition, we have demonstrated that orphans living at these sites show comparable baseline cortisol levels to mother-reared individuals born at the sites, suggesting that the physiological impacts of any early life stress these individuals undergo are minimal (Wobber & Hare, 2011). Further, any effects that these circumstances may have are controlled for in our cross-species comparison since individuals of both species arrive at the sites at a comparable age and are reared in similar circumstances upon arrival according to guidelines of the Pan-African Sanctuary Alliance, of which both sites are members (Farmer, 2002; Wobber & Hare, 2011). Because subjects' exact ages were not known (other than for those individuals born on-site), estimates were made based on comparisons of weight and dental emergence patterns to published values both at the time of the individual's arrival at the sanctuary and at the time of data collection (Grether & Yerkes, 1940; Leigh & Shea, 1996; Smith, Crummett, & Brandt, 1994). These estimates allowed us to be confident of subjects' ages to the year; we also placed individuals in wider age categories (see below), which conferred an even greater degree of certainty in the assignment of individuals to a particular category.

In total, samples were collected from 77 chimpanzees (41 male, 36 female) and 53 bonobos (29 male, 24 female) (Table 4.1). Individuals ranged in age from 1 to 24 years over the three years of sampling (chimpanzees: mean age 8.6 years, median age 7.0 years; bonobos: mean age 8.4 years, median age 7.0 years; there was no species difference in the ages sampled, independent samples t-test). These ages spanned infancy and adulthood in both species but did not include any individuals that could be considered geriatric.

Table 4.1. Characteristics of the subjects that participated in saliva sampling, divided by species and sex. The number of individuals that contributed saliva samples in at least one year and those that contributed samples for multiple (two or three years) are shown for each subgroup. We also show, for each subgroup, the mean age (along with the age range) and the mean number of samples per individual per year (along with the range of samples collected per individual in a given year).

Group	<u>Number of individuals sampled</u>		Age range	Samples per individual in each year
	In at least one year	In multiple years		
Chimpanzee males	41	18	2 to 21 years (mean 9.0 years)	1 to 20 samples (mean 7.3 samples/year)
Chimpanzee females	36	11	2 to 18 years (mean 8.1 years)	1 to 25 samples (mean 9.0 samples/year)
Bonobo males	29	16	3 to 24 years (mean 8.5 years)	1 to 9 samples (mean 5.1 samples/year)
Bonobo females	24	13	1 to 23 years (mean 8.3 years)	1 to 8 samples (mean 4.1 samples/year)

Saliva sampling

Samples for endocrine analysis were collected during the summers of 2007, 2008, and 2009. Each individual was represented by at least one sample, with a range of 1 to 25 samples collected per individual in a given year (chimpanzees: mean 8.0 samples per year, median 8.0 samples, range 1 to 25; bonobos: mean 4.7 samples per year, median 4.0 samples, range 1 to 9), with a total of 1392 samples collected (Table 4.1). A greater number of samples was collected per individual in a given year from chimpanzees than from bonobos (independent samples t-test on the number of samples for each individual in each year, $t(209) = 6.59$, $p < 0.001$). Procedures were taken in the statistical analysis to control for this unbalanced sampling (see below).

Samples for a given individual in a given year were all collected within a 2-month period. Certain individuals were sampled in multiple data collection seasons (29 of the 77 chimpanzees

and 34 of the 53 bonobos), with this repeated sampling controlled for in our statistical analysis (see below). Saliva samples were collected throughout the day (chimpanzees: mean and median hour of sampling 12:00, range 7:42 to 17:10; bonobos: mean and median hour of sampling 11:00, range 6:24 to 16:01). While a circadian pattern in androgen secretion has been demonstrated for chimpanzees (Kutsukake et al., 2009; Muller & Lipson, 2003), the timing of our sample collection largely avoided any potential peaks associated with waking (individuals of both species tend to wake up at 5:30 to 6:00 AM, with previous work showing particularly dramatic declines in human and non-human ape steroid levels immediately after waking (Heintz, Santymire, Parr, & Lonsdorf, 2011)). Nonetheless, samples were collected significantly earlier in the day for bonobos than for chimpanzees (independent samples t-test, $t(1390) = 5.99$, $p < 0.001$), necessitating caution in any comparison of absolute testosterone level between the two species (our comparisons focused instead on the patterns of development within-species). The species difference in the timing of sampling was largely due to the fact that adult bonobos were difficult to sample later in the day, since saliva collection was voluntary on subjects' behalf. In support of the point that bonobo adults were sampled earlier in the day than bonobo juveniles, we found a significant negative relationship in bonobos between age (in years) and the hour that samples were taken (linear regression: $b = -0.167$, $r = 0.329$, $p < 0.001$, $n = 421$) while no such relationship was present in chimpanzees ($b = -0.027$, $r = 0.050$, $p = 0.116$, $n = 971$) (Appendix 4, Figure A4.1). We thus took several measures in our statistical analysis to control for the effects that time of day might have had on testosterone values.

Identical procedures were followed for collection and storage of the saliva samples in all three data collection seasons, as described previously. Saliva collection protocols and radioimmunoassay of testosterone also followed previously published methods (Wobber et al.,

2010). In brief, fifty microliters of 0.1% sodium azide solution was added to each saliva sample immediately after collection to prevent contamination and to allow samples to be kept at room temperature until being returned to the laboratory (Lipson & Ellison, 1989). Salivary testosterone measurements were made in the Reproductive Ecology Laboratory at Harvard University using an I-125 based radioimmunoassay kit (#4100, Diagnostic Systems Laboratories, Webster, TX, USA) with the following modifications: standards were prepared in assay buffer and run at six concentrations from 2 to 375 pg/ml. Samples were added in 100 μ l amounts together with 300 μ l of assay buffer. First antibody (20 μ l) and labeled steroid (50 μ l) were added to each tube to yield a total reaction volume of 470 μ l per tube. After overnight incubation at 4° C, 500 μ l of second antibody was added to each reaction tube. Reaction tubes were subsequently centrifuged for 45 minutes; after aspiration of the supernatant, tubes were counted in a gamma counter for two minutes. In pilot assays using the standard human assay protocol, the ape testosterone values were too high to be readable in the assay range. Therefore, we reduced the sample aliquot to 100 μ l of ape saliva (from 200 μ l for human saliva) in order to be able to read the values on the same standard curve as employed in the human testosterone radioimmunoassay protocol. Assays were counterbalanced according to species, sex, and age. Cross-reactivity of the testosterone RIA kit with other steroids is as follows: 6.6% with 5 α -dihydrotestosterone, 2.2% with 5-androstane-3 β ,17 β -diol, 1.8% with 11-oxotestosterone, 0.9% with androstenedione, and 0.6% with 5 β -dihydrotestosterone. Cross-reactivity with all other steroids was 0.5% or less.

It is important to emphasize that RIA we utilized is highly specific for testosterone, showing extremely low levels of cross-reactivity with other androgens. Unlike analyses of urine which rely on measurements of steroid metabolites, free (unbound) steroids diffuse directly from the blood into the saliva, causing salivary and serum steroid measurements to be highly

correlated both within and outside of humans (Davenport et al., 2003; Ellison, 1989; Kutsukake, et al., 2009; Vittek, Lhommedieu, Gordon, Rappaport, & Southren, 1985). Because the RIA is sensitive to the steroid itself and steroids are identical in structure across mammals, published cross-reactivities of a particular antiserum are therefore the same across species. Moreover, in specifically validating the use of a commercially-available testosterone RIA kit with chimpanzee saliva, a previous study (Kutsukake, et al., 2009) demonstrated that 1) measurements of testosterone from salivary RIA strongly correlate with those obtained from serum, and 2) measurements of salivary testosterone from RIA strongly correlate with those obtained by liquid chromatography-tandem mass spectrometry (LC-MS) (Kutsukake, et al., 2009). It can thus be concluded that RIA successfully binds testosterone in non-human ape saliva, rather than quantifying significant fractions of other androgens or their metabolites. Salivary methods have now been successfully used to quantify steroid levels in numerous non-human primate species (Heintz, et al., 2011; Higham, Vitale, Rivera, Ayala, & Maestripieri, 2010; Kuhar, Bettinger, & Laudenslager, 2005; Pearson, Judge, & Reeder, 2008), making this an exciting direction for future research.

The quality control samples (QC) used for the assays were changed after the data from 2007 were analyzed, so coefficients of variation (CV) are reported separately for this year and for the two subsequent years (2008 and 2009). For assays run in 2007, the average intra-assay CV was 8% and the average inter-assay CV was 16%. For assays run in 2008 and 2009 combined, the average intra-assay CV was 10% and the average inter-assay CV was 15%. There were no significant differences in the pool values between 2008 and 2009 for either the low pool (Mann-Whitney U test given the small sample size, $Z = 0.47$, $p = 0.6$) or the high pool ($Z = 0.78$, $p = 0.4$), suggesting that assay characteristics did not vary significantly between years.

Sampling of body weight

To provide an additional independent measure of growth (because our age measures were only estimates), we also examined the relationship between testosterone and body weight. For this analysis, we were able to obtain weights taken in the same month as saliva sampling for a number of individuals who were younger than 9 years ($n = 55$ weights across the three years of data collection, taken from 42 individuals). However, individuals who were 9 years and older could only be weighed when anesthetized. Thus for these individuals, we utilized weights obtained from an annual health check performed within 6 months of saliva sampling ($n = 23$ weights from 23 individuals). Because individuals who were 9 and older were likely growing less rapidly than the younger age group, this 6-month weight estimate provided the best available proxy for their weights at the time of saliva sampling.

Sampling of dental emergence

To provide yet another independent measure of general maturation, we examined the relationship between testosterone and an individual's level of dental development. Namely, we performed a visual inspection of subjects' tooth emergence, recording the emergence of permanent teeth for the majority of individuals sampled in the hormone analysis ($n = 99$). Based on the previously-documented patterns of dental emergence, which are identical in sequence between chimpanzees and bonobos (Boughner & Dean, 2008; Kuykendall, Mahoney, & Conroy, 1992; Smith, et al., 1994), we created 6 dental categories: no permanent dentition ($n = 11$), first molar (M1) only ($n = 30$), permanent incisors only ($n = 13$), second molar (M2) only ($n = 34$), permanent canine only ($n = 2$), and third molar (M3) emergence/complete adult dentition ($n = 65$). We treated the presence of either a mandibular or maxillary tooth as sufficient for placement of the individual into a given category, and we grouped together both permanent incisors (I1 and

I2) into our “incisors only” category. Because only two individuals who were possible to sample for dental emergence fell into the “permanent canine only” category, we grouped these individuals together with the “M3” category for our statistical analysis. While these dental categories provided numerous classifications for young individuals, they did not provide a way to distinguish between age groups of individuals that were fully dentally developed (e.g., a 10-year-old and a 20-year-old would both be classified as dentally mature).

Statistical analysis

To control for the fact that certain individuals were sampled more frequently than others in any given year, we began by computing an average of each individual’s samples taken in each year (total sample size = 211 averages across 3 years of sampling). To control for any circadian effects on testosterone, we also computed individual averages comprised of samples taken during the early-morning hours (6:00 to 11:00) (n = 164) or during the mid-day (11:00 to 16:00) (n = 183) in any given year. Only four bonobo samples were collected after 16:00, so we excluded this time range in our controls for time of day. Below, we present analyses performed with the overall averages as well as with the averages taken only from early morning samples and averages taken only from the mid-day. Critically, as mentioned above, bonobo adults were difficult to sample in the mid-day and afternoon. Therefore, while the sample size was larger for the mid-day averages than the early morning averages, the early morning averages better represented the full range of bonobo development.

After computing these average values, we log-transformed the averages to normalize the data (since, as is typical with hormonal measures, our data exhibited significant skew). In addition to performing analyses separately based on time of day, we also performed analyses separately for each year of sampling (2007, 2008, and 2009) because the samples in 2007 were

collected in conjunction with a behavioral experiment (Wobber, Hare, et al., 2010). We present our results in a series of tables below and in Appendix 4 to clarify these multiple levels of analysis.

To begin our analyses, we investigated general patterns in the data by performing linear regressions of age and log average testosterone separately by species and by sex (given the prediction that males and females would differ in their values of testosterone, and our hypothesis that the two species would differ in their relationship between age and testosterone). To probe the data further, and to control for repeated sampling across multiple years for any given individual, we used Generalized Estimating Equations (GEE). In the GEE analysis, we were able to ascertain the effects of species, sex, and age, in addition to controlling for the within-subject factor individual (which could range from 1 to 3 based on the number of years that individual was sampled).

In our GEE analyses, we used four different dependent measures to denote development, presented in sequence in the Results section. First, we performed a GEE analysis with age (in years) expressed as a continuous variable. Then, to facilitate the use of post-hoc tests on interactions in the model (for example, between species and age), we divided individuals into four age categories in line with general patterns of aging observed in chimpanzees and bonobos (de Lathouwers & van Elsacker, 2006; Furuichi et al., 1998; A. Pusey, 1990; A. E. Pusey, Oehlert, Williams, & Goodall, 2005): infant (1 to 4 years, $n = 39$), juvenile (5 to 8 years, $n = 84$), subadult (9 to 12 years, $n = 50$), and adult (13 years and above, $n = 38$). Subsequently, we performed our analyses using two empirical measures of growth. We first performed a GEE with body weight entered as a covariate, and then performed a GEE with dental category as our developmental parameter. Full factorial models were performed for all GEE analyses except

where noted otherwise, with post-hoc tests adjusted for multiple comparisons using a Bonferroni correction. These analyses were performed in SPSS 16.0 Graduate Student Version.

Finally, to ensure that we had adequate power to detect significant effects of age on testosterone in each species and sex, we also performed a power analysis using the program GPower (Version 3.1.3).

RESULTS

Analyses with age in years as a continuous measure

Linear regressions between age and log average testosterone value, performed separately for each species and sex, revealed significant positive relationships in male chimpanzees ($b = 0.033$, $r = 0.524$, $p < 0.001$, $n = 70$) and female chimpanzees ($b = 0.019$, $r = 0.277$, $p = 0.049$, $n = 51$), but not male bonobos ($b = 0.003$, $r = 0.072$, $p = 0.614$, $n = 51$) or female bonobos ($b = -0.002$, $r = 0.043$, $p = 0.797$, $n = 39$) (Table 4.2). The regression for female chimpanzees became non-significant when excluding particularly high average values (above 2000 pmol/L, or 3.0 on the log scale) taken from one individual ($b = 0.011$, $r = 0.198$, $p = 0.173$, $n = 49$). These two values were excluded from all further analyses, as well as their respective figures and tables. Though the analysis was performed with log-transformed testosterone values, untransformed values are shown in Figure 4.1 for illustrative purposes, including these female chimpanzee outliers.

The positive relationship between age and log testosterone in male chimpanzees was present in both the early morning samples ($b = 0.030$, $r = 0.565$, $p < 0.001$, $n = 54$) and the mid-day samples ($b = 0.036$, $r = 0.499$, $p < 0.001$, $n = 70$), while the regressions for female chimpanzees and bonobos of both sexes remained non-significant in these two time periods

(Table 4.2). In addition, the positive relationship between age and log testosterone in chimpanzee males was present in all three years of sampling (2007: $b = 0.036$, $r = 0.624$, $p = 0.010$, $n = 16$; 2008: $b = 0.015$, $r = 0.345$, $p = 0.039$, $n = 36$; 2009: $b = 0.049$, $r = 0.570$, $p = 0.013$, $n = 18$), while again no year of sampling revealed a significant age-testosterone relationship in female chimpanzees or bonobos of either sex (Table 4.2). These findings indicate that testosterone increased significantly over development in chimpanzee males but not bonobo males. Even at 2 to 3 years of age, bonobos showed the same range of testosterone values as adult individuals.

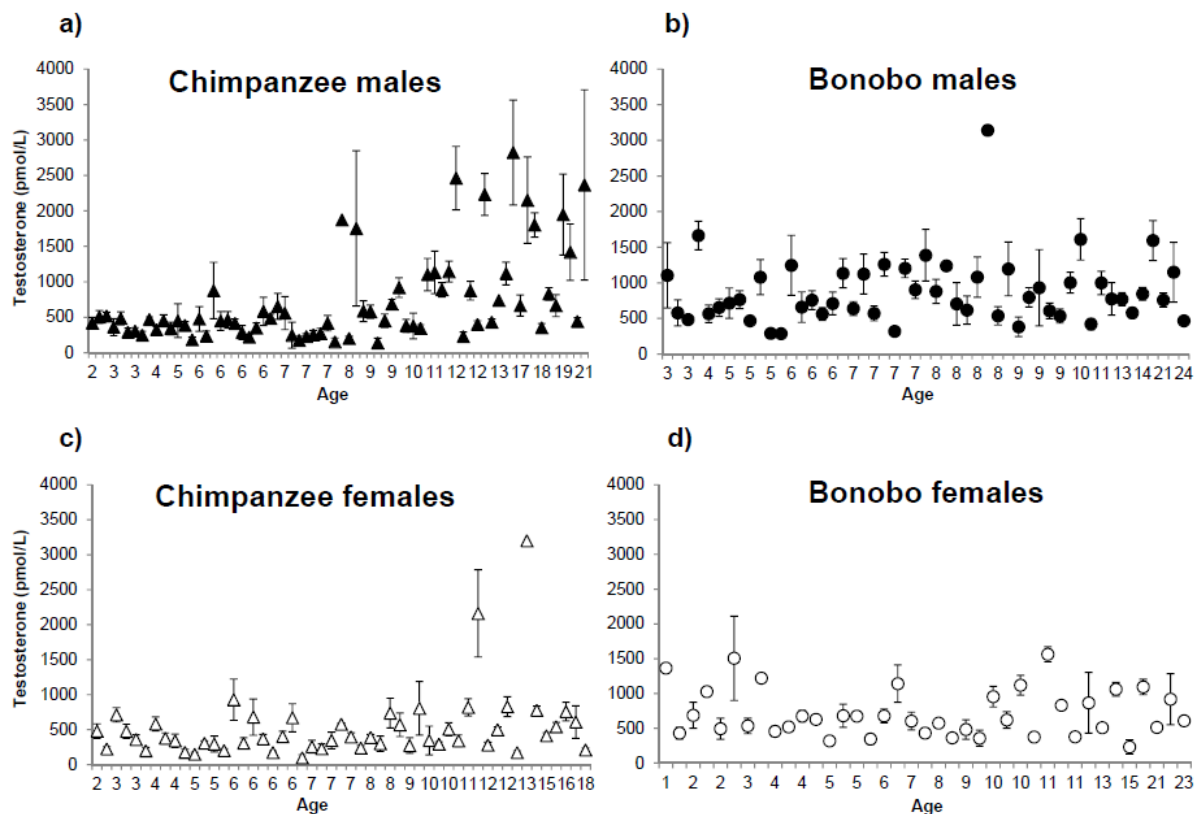


Figure 4.1. Average testosterone levels according to age for a) chimpanzee males, b) bonobo males, c) chimpanzee females, and d) bonobo females. Individual yearly averages and standard errors around those averages are shown, ordering individuals according to increasing age in each species and sex. All graphs are shown on the same scale. Actual testosterone values (in pmol/L) are shown here, though log-transformed values were used for the statistical analyses. Bonobos changed little in testosterone with age in either sex, while in chimpanzees there was a slight increase in testosterone with age in females and a more dramatic increase in testosterone with age among males.

Table 4.2. Regression parameters for the relationship between age (in years) and log testosterone across all data and subsets of the data. Linear regressions were performed separately by species and sex. Regressions were performed with overall log averages, as well as separately with the early morning and mid-day log averages, and separately for each year of sampling. The slope, correlation coefficient (R), p-value (p), and sample size (N) for each regression are shown. In addition, the mean age for the points included in each regression is indicated, since certain individuals were not sampled by every analysis (for example, mid-day averages were not available for several adult bonobos). Significant p-values are indicated in bold. In this table we report results including two outliers within female chimpanzees and then excluding these points. We removed these points from all further analyses, and from all further figures and tables.

	Slope	R	p	N	Mean age
All Samples:					
Male chimpanzees	0.033	0.524	<0.001	70	9.0
Female chimpanzees	0.019	0.277	0.049	51	8.1
Female chimpanzees removing outliers	0.011	0.198	0.173	49	7.9
Male bonobos	0.003	0.072	0.614	51	8.5
Female bonobos	-0.002	0.043	0.797	39	8.3
Early Morning Samples Only:					
Male chimpanzees	0.030	0.565	<0.001	54	9.1
Female chimpanzees (removing outliers)	0.014	0.192	0.254	37	7.9
Male bonobos	0.003	0.054	0.732	42	8.9
Female bonobos	0.008	0.215	0.263	29	9.0
Mid-Day Samples Only:					
Male chimpanzees	0.036	0.499	<0.001	70	9.0
Female chimpanzees (removing outliers)	0.014	0.211	0.151	48	7.9
Male bonobos	-0.001	0.017	0.922	36	7.4
Female bonobos	0.005	0.121	0.549	27	6.9
All Samples, by Year:					
Male chimpanzees					
2007	0.036	0.624	0.010	16	9.6
2008	0.015	0.345	0.039	36	7.7
2009	0.049	0.570	0.013	18	11.0
Female chimpanzees (removing outliers)					
2007	0.015	0.256	0.338	16	8.6
2008	0.008	0.149	0.477	25	7.4
2009	0.045	0.407	0.317	8	8.1
Male bonobos					
2007	-0.002	0.073	0.831	11	8.6
2008	0.008	0.213	0.296	26	8.5
2009	-0.006	0.121	0.681	14	8.3
Female bonobos					
2007	-0.008	0.500	0.254	7	8.3
2008	0.004	0.171	0.470	20	8.6
2009	-0.003	0.139	0.667	12	7.8

To ensure that these effects did not result from our smaller bonobo sample size, we also performed analyses of effect size and power obtained from the linear regressions. With our sample of 70 male chimpanzee data points, the effect size and power of the age-testosterone relationship were quite high ($f^2 = 0.377$, power = 0.999). Assuming a similar effect size were present in male bonobos, the sample size needed to obtain a significant ($p < 0.05$) result with reasonable power (> 0.8) would have been 23 individuals, indicating that our sample size of 51 bonobo male data points was large enough to have detected such a relationship. In addition, even if bonobo males were to show a weaker relationship between age and testosterone (as might be predicted given the known differences in the male dominance rank-testosterone relationship between the two species), our sample size gave us the ability to detect any effect size greater than 0.16. We can thus conclude that our sample size of bonobo males was more than adequate to detect any relationship present between age and testosterone.

To further investigate these patterns and control for repeated sampling across years, we next performed a GEE analysis of log average testosterone with individual as a repeated factor (ranging from 1 to 3 samples per individual, based on the number of years that they participated), species and sex as between-subject factors, and age (in years) as a covariate. This analysis revealed significant effects of species (Wald Chi-Square (1) = 39.581, $p < 0.001$), age (Wald Chi-Square (1) = 12.990, $p < 0.001$), species*age (Wald Chi-Square (1) = 11.070, $p = 0.001$), and sex*age (Wald Chi-Square (1) = 4.295, $p = 0.038$) (Appendix 4, Table A4.1). Bonobos had higher log testosterone values on average than chimpanzees. The main effect of sex was likely non-significant due to the lack of a sex difference among juveniles, hence there being only a sex*age interaction.

Post-hoc tests were not possible for the effects and interactions of the continuous age measure, necessitating our analysis with age categories described below. The main effect of species and the species*age interaction remained significant when performing the GEE analyses with the early morning averages, the mid-day averages, and when performing the analyses separately by sample year, except that the species*age interaction was not significant in 2008 (Appendix 4, Table A4.1). These results substantiated our finding of an increase in testosterone with age in chimpanzees but not in bonobos. We investigated these patterns in more detail by using a categorical age measure.

Analyses using age categories

To examine the main effects and interactions of age, we performed GEE analyses using the age category measure described above. An analysis of log average testosterone with individual as a within-subject factor and species, sex, and age category as between-subjects factors revealed significant effects of species (Wald Chi-Square (1) = 33.469, $p < 0.001$), sex (Wald Chi-Square (1) = 17.082, $p < 0.001$), age category (Wald Chi-Square (3) = 16.341, $p = 0.001$), species*age category (Wald Chi-Square (3) = 11.495, $p = 0.009$), and sex*age category (Wald Chi-Square (3) = 8.154, $p = 0.043$) (Figure 4.2). Post-hoc comparisons for the species*age category interaction revealed that chimpanzee infants had significantly lower testosterone levels than chimpanzee adults (Bonferroni-corrected $p = 0.024$), and significantly lower testosterone levels than bonobos of all ages (Bonferroni-corrected p values < 0.001). The same was true for chimpanzee juveniles, who had lower levels of testosterone than chimpanzee adults (Bonferroni-corrected $p = 0.004$) and lower levels of testosterone than all bonobo groups (Bonferroni-corrected p values < 0.001). All other groups were comparable, with chimpanzee subadults no different from chimpanzee adults, and both of these groups no different from bonobos of any age

group. Importantly, there were no significant differences in levels of testosterone between any bonobo age groups (Appendix 4, Table A4.2). These results demonstrate, in line with the regression analyses, that chimpanzees showed a developmental increase in testosterone production while bonobos did not, with infant, juvenile, subadult, and adult bonobos all possessing similar testosterone levels. Post-hoc tests investigating the sex*age category interaction revealed that overall, adult males had higher testosterone levels than females of all age groups (Bonferroni-corrected p-values $p < 0.05$), and had higher testosterone levels than infant and juvenile males (Bonferroni-corrected p-values $p < 0.01$), with no such age differences present among females.

The main effect of species and the interaction between species and age category remained significant when performing the GEE analyses separately for the early morning averages and the mid-day averages (Table 4.3). When performing the analyses separately by year, the main effect of species was significant in two of the three years (2008 and 2009, but not 2007), while the interaction between species and age category was significant in two of the three years as well (2007 and 2009, but not 2008) (Appendix 4, Table A4.2).

Though the GEE analysis did not reveal a significant 3-way interaction between species, sex, and age category, we performed post-hoc comparisons of these groups in order to determine when sexual dimorphism in testosterone levels emerged in each species. In particular we wanted to assess whether our results supported the previous finding that adult chimpanzees show a greater sex difference in androgen levels than do adult bonobos (Sannen, et al., 2003).

Independent samples t-tests of the sex difference in each species and age category revealed that the only significant difference was in adult chimpanzees ($t(20) = 3.065$, Bonferroni-corrected p-value = 0.048), with no such sex difference among adult bonobos ($t(13) = 1.190$, Bonferroni-

corrected p -value > 1.0) or among any other age group in either species (Appendix 4, Table A4.3). These results suggest that bonobos had a lesser sex difference in testosterone as adults relative to chimpanzees, though it is important to note that this distinction was not strong enough to result in a 3-way species*sex*age interaction in our broader analysis. Our findings indicate that the key mechanism underlying this lesser sex difference in adult bonobos relative to adult chimpanzees is a lesser developmental increase in testosterone levels among bonobo males.

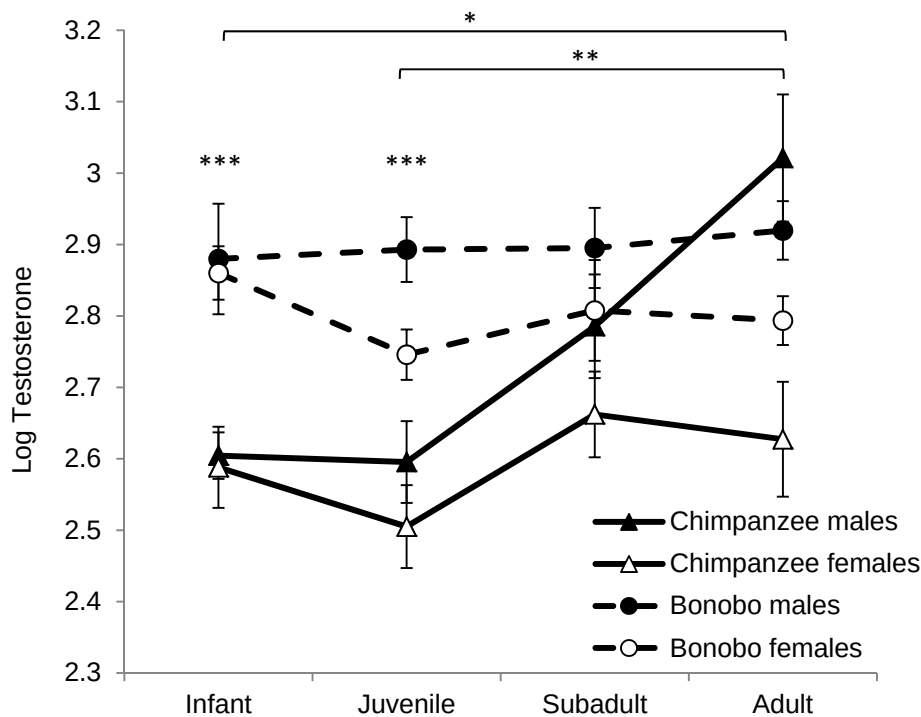


Figure 4.2. Log testosterone levels across development in chimpanzees and bonobos.

Average log testosterone values and sample sizes are shown for each age group, excluding two outliers within female chimpanzees since these were excluded from the statistical analysis. Significant post-hoc tests of the species*age group interaction are indicated, with significant inter-species comparisons shown directly above the data while significant differences between chimpanzee age groups are shown above the comparison bars. There were no significant differences between age groups in bonobos. Bonferroni-corrected p -values for these comparisons are denoted as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Analyses were performed with both sexes pooled; values are shown separately by sex here for illustrative purposes.

Table 4.3. List of significant effects from the Generalized Estimating Equations (GEE) analysis of log testosterone with individual, species, sex, and age category as factors. GEE analyses were performed with the overall log averages, as well as separately for the early morning and mid-day testosterone log averages, and separately for each year of sampling. Though a full factorial model was run for each analysis, here we show only the significant effects and interactions along with their respective Wald chi-square values and p-values. We also report the full-model “Quasi-Likelihood under Independence Model Criterion,” or QIC, and sample size used in each model. Note that smaller QIC values indicate a better model fit.

	Wald chi-square	p	Whole model	
			QIC	N
<i>All Samples:</i>			11.889	209
Species	33.469	<0.001		
Sex	17.082	<0.001		
Age category	16.341	0.001		
Species*Age category	11.495	0.009		
Sex*Age category	8.154	0.043		
<i>Early Morning Samples Only:</i>			10.916	162
Species	7.676	0.006		
Sex	22.846	<0.001		
Age Category	16.281	0.001		
Species*Age category	10.413	0.015		
<i>Mid-Day Samples Only:</i>			13.921	181
Species	43.735	<0.001		
Age Category	20.645	<0.001		
Species*Sex	7.127	0.008		
Species*Age category	11.429	0.010		
Species*Sex*Age category	9.940	0.019		
<i>All Samples, by Year:</i>				
2007			2.038	50
Sex	17.783	<0.001		
Age category	8.673	0.034		
Species*Age category	21.398	<0.001		
Sex*Age category	8.620	0.013		
2008			4.476	107
Species	15.064	<0.001		
Sex	6.761	0.009		
Age category	11.497	0.009		
2009			4.288	52
Species	17.326	<0.001		
Age category	9.313	0.025		
Species*Age category	7.921	0.019		

Analyses using body weight

In our next set of analyses, we used weight as our developmental measure to examine changes in testosterone with general growth in both species. Linear regressions between weight (in kg) and log testosterone, performed separately for each species and sex, revealed a significant positive relationship in chimpanzee males ($b = 0.011$, $r = 0.682$, $p < 0.001$, $n = 24$) but not in chimpanzee females or bonobos of either sex (Appendix 4, Table A4.4, Figure 4.3; note that we continued to remove the high-T chimpanzee female outlier in this analysis). The relationship between weight and testosterone in male chimpanzees was present when including only early morning averages ($b = 0.011$, $r = 0.698$, $p < 0.001$, $n = 21$) and only mid-day averages ($b = 0.015$, $r = 0.668$, $p < 0.001$, $n = 24$), while the regressions for female chimpanzees and bonobos of both sexes remained non-significant in these two time periods. Due to the small sample size of weights available for each species and sex in any given year, analyses could not be performed separately by year. Nonetheless, these regression analyses demonstrate that even when using an empirical measure of growth, chimpanzee males continued to show an increase in testosterone over development while no such increase was present among bonobo males. This stronger relationship between weight and testosterone in chimpanzees was not associated with a stunted growth trajectory in our bonobo sample: chimpanzees were larger on average than bonobos (independent samples t-test of weights used in the testosterone analysis, $t(76) = 3.073$, $p = 0.003$), but both species showed a significant increase in weight with age (regressions between age and weight, chimpanzees: $b = 3.154$, $r = 0.924$, $p < 0.001$, $n = 47$; bonobos: $b = 1.852$, $r = 0.810$, $p < 0.001$, $n = 31$) (see also (Wobber, Hare, & Wrangham, In preparation) for more detailed analyses of body growth in the two species). Thus on the whole, the weight analyses demonstrated significant differences in patterns of testosterone production across maturation

between the sexes as well as between the two species. Males' testosterone increased more than females' with increased weight growth, while chimpanzees showed a greater testosterone increase with growth than did bonobos.

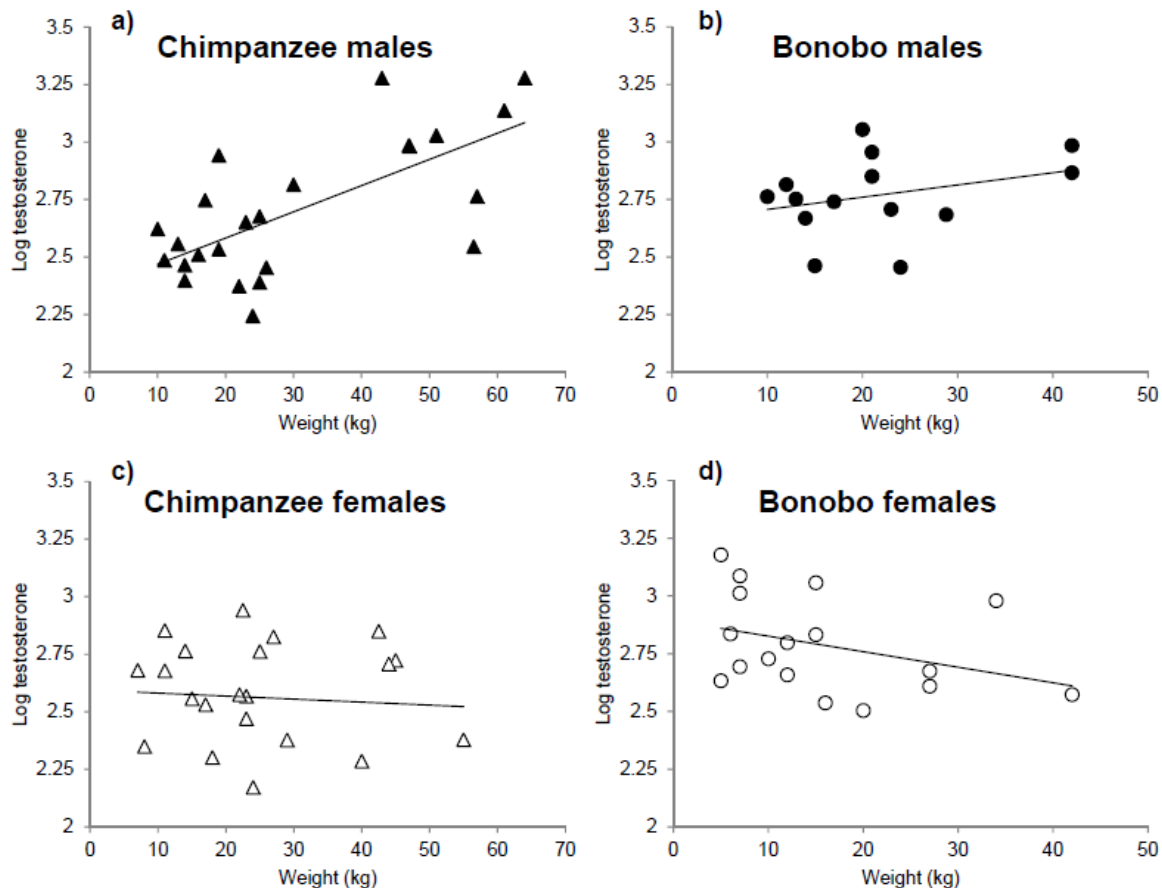


Figure 4.3. Log testosterone levels according to weight (in kg) for a) chimpanzee males, b) bonobo males, c) chimpanzee females, and d) bonobo females. Sample size here was reduced relative to the overall analysis, since we only included weights taken within the same month as saliva sampling, or within 6 months of sampling for individuals who were 9 years and older. Values exclude two outlier points in female chimpanzees. The scale of the y-axis is the same for all graphs; the scale of the x-axis is consistent within-species but was reduced for bonobos since no bonobo exceeded 45 kg. Linear regression trend lines are shown for each graph. Similar to the findings with age, chimpanzee males increased in testosterone as they increased in body mass, while no such increase was present in chimpanzee females or bonobos of either sex.

Analyses using dental category

As a final empirical index of development we performed analyses using the dental category assignments (see Methods). Again, this allowed us to examine the production of testosterone across development independent from any potential bias introduced by our age estimates. For this GEE analysis, we removed the three-way interaction between species, sex, and dental category from the model due to small sample size in certain categories (for example, there was only one bonobo female in the “M2 only” category, though there were 12 bonobos in the “M2 only” category overall). A GEE analysis of log average testosterone with individual, species, sex, and dental category as factors (including all main effects and 2-way interactions) revealed significant effects of species (Wald Chi-Square (1) = 32.250, $p < 0.001$), sex (Wald Chi-Square (1) = 4.133, $p = 0.042$), dental category (Wald Chi-Square (4) = 14.065, $p = 0.007$), and species*dental category (Wald Chi-Square (4) = 23.232, $p < 0.001$) (Figure 4.4, Appendix 4, Table A4.5). Post-hoc tests of the species*dental category interaction revealed that within chimpanzees, individuals with only an M1 (Bonferroni-corrected $p = 0.043$), with incisors only (Bonferroni-corrected $p = 0.001$), and with an M2 only (Bonferroni-corrected $p = 0.001$) had significantly lower testosterone than individuals with all permanent dentition. Notably, chimpanzees of the youngest age group (no permanent dentition) did not differ from the testosterone levels of adults (all permanent dentition), providing some evidence for their being a neonatal elevation in chimpanzee testosterone production. Meanwhile, chimpanzees in all dental groups except the oldest (permanent dentition) had lower testosterone than bonobos of at least one dental group (Appendix 4, Table A4.6). There were no differences in testosterone level between any dental categories within bonobos.

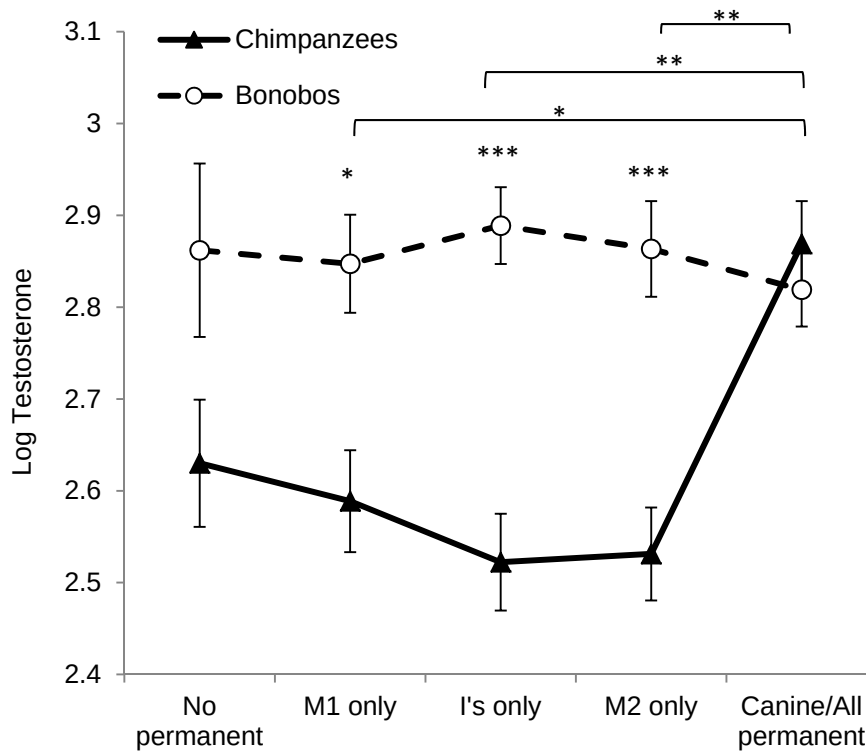


Figure 4.4. Log testosterone levels according to dental development in chimpanzees and bonobos. Average log testosterone and sample size are shown separately by species for each dental category: no permanent dentition, first molar (M1) only, permanent incisors (I's) only, second molar (M2) only, and permanent canine only/all permanent dentition. Values exclude two outlier points in female chimpanzees. Sample size was reduced relative to the overall analysis, since dental emergence was not recorded for all individuals sampled. Significant post-hoc tests of the species*dental category interaction are indicated, with significant inter-species comparisons shown directly above the data while significant differences between chimpanzee dental categories are shown above the comparison bars. There were no significant differences between any dental categories in bonobos. Bonferroni-corrected p-values for these comparisons are denoted as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The interaction between species and dental category was present when looking only at the early morning averages, and the main effect of species was present when looking only at mid-day averages while the species*dental category interaction was only significant at a trend level in this time period (Appendix 4, Table A4.5). Similar to the weight analysis, dental categories were not obtained equally across the three sample years so the sample size was too small to perform separate-year analyses for this factor.

The dental category results provide additional support for the notion that chimpanzees show an increase in testosterone with maturation, while bonobos do not, with these effects particularly pronounced in males. Therefore, regardless of the developmental parameter used (age, weight, or dental stage), bonobos were not found to show a significant change in testosterone level between infancy and adulthood.

DISCUSSION

Our results support the hypothesis that differences in male reproductive strategy between bonobos and chimpanzees are associated with important distinctions in the ontogeny of testosterone production between the two species. In chimpanzees, levels of testosterone increased in both males and females during the transition from juvenility to adulthood, doing so more markedly in males in agreement with previous work (Marson, et al., 1991b; Martin, et al., 1977). In bonobos, by contrast, there was no evidence of maturational increases in testosterone production in either sex. This minimal change in testosterone production with age among bonobos did not reflect stunted growth in our study population, as revealed by analyses of body weight (see also (Wobber, et al., In preparation)). Relative to chimpanzees, bonobos showed a lesser neonatal decline and a lesser pubertal elevation in testosterone, indicating that both developmental periods might be critical in shaping adult reproductive behavior.

It is important to emphasize that we cannot conclude definitely on the basis of our results that bonobos show no pubertal increase in testosterone production, given that this would strongly contradict the general mammalian pattern. Since these saliva samples were collected during the summers of subsequent years (rather than continuously throughout the year), it is possible that bonobo males showed transient testosterone increases (not represented in these samples) as part

of their pubertal maturation. But even if such increases occurred, our results indicate that adolescent and adult bonobos did not sustain heightened levels of testosterone for more than a matter of months. We thus argue that bonobos show a lesser developmental increase in testosterone in association with their lesser degree of male mating competition relative to chimpanzees. Before elaborating on this point, we first review the quality of our data.

Strengths and limitations of the present data set

Aspects of our sample collection method were unlikely to have generated the present pattern of results. For example, the use of cotton as a saliva collection material has been suggested to lead to over-estimation of steroid concentrations, with the use of oral stimulants potentially elevating measurement values as well (Granger, Shirtcliff, Booth, Kivlighan, & Schwartz, 2004; Talge, Donzella, Kryzer, Gierens, & Gunnar, 2005). However, even if some bias were introduced into our data by cotton or oral stimulants, any impacts of these factors on the salivary steroid measures would have similarly influenced the results from all individuals of both species (since the same procedures were used for saliva collection throughout). This therefore could not account for our finding of a developmental transition in testosterone production among chimpanzees but not bonobos.

The observed patterns were also unlikely to have resulted from differential sampling throughout the day, simply reflecting circadian variation rather than meaningful developmental patterns. As discussed in the Methods, there was a bias in the time of day when samples were obtained – bonobo adults were sampled more frequently in the morning, while bonobo juveniles and chimpanzees of all ages were sampled more equally throughout the day (Appendix 4, Figure A4.1). Based on the findings in humans and non-human apes that testosterone levels tend to be higher and more variable in the morning than afternoon (Muller & Lipson, 2003; Van Cauter,

1990), this sampling pattern would predict that bonobo adults should show **higher** and **more variable** testosterone levels than any other group. It is therefore particularly surprising that morning bonobo adult testosterone levels were **comparable** to those taken from juveniles throughout the day (where, in theory, afternoon samples from adults would have further lowered their testosterone levels relative to those of juveniles). It is similarly striking that adult male bonobos' testosterone levels were considerably **less variable** than those of adult male chimpanzees. This effect may partly have been due to a larger sample size of adult chimpanzee males, though we did sample all adult bonobo males living at our study site except for one individual that had a history of biting caretakers. We discuss potential explanations for the lesser inter-individual variability found among adult bonobos below.

Finally, in regards to our finding that testosterone levels were comparable in adult and juvenile bonobos, it is important to note that this pattern is not unprecedented for non-human primates. In several seasonally breeding strepsirrhines, adult male testosterone levels have been observed to drop into the juvenile range outside of the breeding season (Fitch-Snyder & Jurke, 2003; Gould & Ziegler, 2007; von Engelhardt, Kappeler, & Heistermann, 2000). Similarly, such patterns have been documented outside of the breeding season in mandrills, with low-ranking males increasing little in their androgen levels during puberty (Setchell & Dixson, 2002). Bonobos may thus represent the rare case of an aseasonally breeding species where testosterone levels are consistently low among adults. It is possible that this denoted a stable hierarchy among bonobos in our sample, similar to a group of baboons in which a positive relationship between male rank and testosterone was only present when the hierarchy was unstable (Sapolsky, 1993)). However, male chimpanzees have been found to maintain rank-testosterone relationships even during periods of rank stability (Muller & Wrangham, 2004). Moreover, a recent study found

little association between basal testosterone and dominance rank in a group of wild bonobos (Surbeck, et al., in press). Our results thus suggest that the reduced aggression and fluid dominance hierarchy present among bonobos may be accompanied by low, invariant testosterone levels in adult bonobo males.

While the between-species differences that we found in maturational patterns of male testosterone production are easily interpreted in relation to reproductive strategies, the differences between bonobos and chimpanzees in absolute testosterone level cannot be evaluated without data on androgen receptor density. If bonobos have a lesser density of androgen receptors, they may need to produce a greater amount of testosterone relative to chimpanzees to obtain an equivalent metabolic effect. As in humans, there is considerable inter-individual variability in the expression of the androgen receptor gene in both chimpanzees and bonobos (Giovannucci et al., 1997; Hong et al., 2006; Sirugo, Deinard, Kidd, & Kidd, 1997), so it is difficult even to characterize average receptor densities in each species. Further research is thus necessary to illuminate how individual differences in genotype translate to the phenotypic differences observed between bonobos and chimpanzees.

Directions for future research

An intriguing facet of our data was the lesser variability in testosterone production found among adult bonobos relative to adult chimpanzees, with this effect particularly strong among males. We propose that this reduction in adult testosterone variability in fact reflects a crucial element of the bonobo male reproductive strategy. In chimpanzees, aggression and dominance rank are effective strategies for obtaining conceptive mating opportunities (Muller, et al., 2007; Wrangham, 2002). Correspondingly, chimpanzee males show significant rank-dependent variation in testosterone production (Muller & Wrangham, 2004). In our data set as well, it is

likely that rank differences underlay the variation in adult male testosterone among chimpanzees, though we could not test the rank-testosterone correlation directly because the adults sampled were living in multiple social groups. In contrast, the reduced efficacy of aggression and competition for dominance in bonobo males may explain the reduced inter- and intra-individual variability in their testosterone levels (Marshall & Hohmann, 2005; Surbeck, et al., in press), as well as the lesser sex difference in adult testosterone levels we found in bonobos relative to chimpanzees (in line with previous work on urinary androgens, (Sannen, et al., 2003)). Together with lesser variation over the course of development, bonobo males may optimize their immune function and overall life history strategy by elevating testosterone levels only when necessary during puberty and maintaining low testosterone levels otherwise. Future work investigating gonadotropin and adrenal androgen production in bonobos can determine whether these effects are specific to testosterone or instead reflect broader shifts in the features of bonobo ontogeny.

In addition to the minimal change in testosterone shown by bonobos during puberty, we also found no evidence for a decline in their testosterone levels between infancy and juvenility. Conversely, among chimpanzees, there was some signature of this neonatal decline in the dental category analysis (which provided the greatest resolution in grouping young individuals) (Figure 4.4). Assuming that the chimpanzee pattern is the ancestral condition (in line with the patterns of infant testosterone production documented for multiple non-human primates), bonobos thus appear unusual in maintaining neonatal elevations of testosterone throughout infancy and juvenility. It is possible that because genital contacts are an important facet of bonobo social behavior even in infancy (Hashimoto & Furuichi, 2006; Kano, 1989; Kuroda, 1984; Parish & de Waal, 2000; Woods & Hare, 2011), testosterone levels remain high throughout infancy and juvenility to sustain high levels of libido. Alternatively, sexual contacts themselves might elevate

testosterone levels in infant and juvenile bonobos, given the evidence from human males that sexual activity can increase testosterone levels and the finding that frequency of genito-genital rubbing correlates positively with androgen levels among adult bonobo females (Dabbs & Mohammed, 1992; Sannen, Van Elsacker, Heistermann, & Eens, 2005). However, because we did not sample any bonobos younger than 1.5 years, we cannot say whether levels of testosterone among neonatal individuals were even higher than those measured in infants and juveniles. Additional study of endocrine maturation and social behavior in neonatal bonobos is thus warranted.

Overall, our data suggest that differences in male reproductive strategies across species are associated with differences in the developmental patterns of testosterone production. Additional research on the ontogeny of testosterone production in closely-related species is essential to understand how slight variations in developmental trajectory can facilitate and constrain the reproductive strategies pursued by adults. Such inquiry will illuminate the role of hormones in shifting the maturation of a broad array of phenotypes, and will provide insight into the mechanisms by which evolution produces variation across species.

Chapter 4 Literature Cited.

- Andersson, A. M., Toppari, J., Haavisto, A. M., Petersen, J. H., Simell, T., Simell, O., & Skakkebaek, N. E. (1998). Longitudinal reproductive hormone profiles in infants: Peak of inhibin B levels in infant boys exceeds levels in adult men. *Journal of Clinical Endocrinology & Metabolism*, 83(2), 675-681.
- Anestis, S. (2006). Testosterone in Juvenile and Adolescent Male Chimpanzees (Pan troglodytes): Effects of Dominance Rank, Aggression, and Behavioral Style. *American Journal of Physical Anthropology*, 130(4), 536-545.
- Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience and Biobehavioral Reviews*, 30(3), 319-345.
- Beehner, J., Bergman, T., Cheney, D., Seyfarth, R., & Whitten, P. (2006). Testosterone predicts future dominance rank and mating activity among male chacma baboons. *Behavioral Ecology and Sociobiology*, 59(4), 469-479.
- Beehner, J. C., Gesquiere, L., Seyfarth, R. M., Cheney, D. L., Alberts, S. C., & Altmann, J. (2009). Testosterone related to age and life-history stages in male baboons and geladas. *Hormones and Behavior*, 56(4), 472-480. doi: 10.1016/j.yhbeh.2009.08.005
- Bhasin, S., Storer, T., Berman, N., Callegari, C., Clevenger, B., Phillips, J., . . . Casaburi, R. (1996). The Effects of Supraphysiologic Doses of Testosterone on Muscle Size and Strength in Normal Men. *New England Journal of Medicine*, 334, 1-7.
- Blank, M. S., & Murphy, J. R. (1991). Luteinizing-hormone sensitivity to naloxone in maturing male chimpanzees. *Brain Research Bulletin*, 27(2), 241-245.
- Boughner, J., & Dean, M. (2008). Mandibular Shape, Ontogeny and Dental Development in Bonobos (Pan paniscus) and Chimpanzees (Pan troglodytes). *Evolutionary Biology*, 35(4), 296-308. doi: 10.1007/s11692-008-9043-6
- Bribiescas, R. G. (2001). Reproductive ecology and life history of the human male *Yearbook of Physical Anthropology*, Vol 44 (Vol. 44, pp. 148-176). New York: Wiley-Liss, Inc.
- Carroll, S. B. (2008). Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell*, 134(1), 25-36. doi: 10.1016/j.cell.2008.06.030
- Cunningham, G. R., & Huckins, C. (1979). Persistence of complete spermatogenesis in the presences of low intra-testicular concentrations of testosterone. *Endocrinology*, 105(1), 177-186.
- Dabbs, J., & Mohammed, S. (1992). Male and female salivary testosterone concentrations before and after sexual activity. *Physiology and Behavior*, 52, 195-197.
- Davenport, M. D., Novak, M. A., Meyer, J. S., Tiefenbacher, S., Higley, J. D., Lindell, S. G., . . . Suomi, S. J. (2003). Continuity and change in emotional reactivity in rhesus monkeys

- throughout the prepubertal period. *Motivation and Emotion*, 27(1), 57-76. doi: 10.1023/a:1023626425807
- de Lathouwers, M., & van Elsacker, L. (2006). Comparing infant and juvenile behavior in bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*): a preliminary study. *Primates*, 47, 297-293.
- de Waal, F. (1987). Tension regulation and nonreproductive functions of sex in captive bonobos (*Pan paniscus*). *National Geographic Research Reports*, 3, 318-335.
- Ellison, P. (1989). Human Salivary Steroids: Methodological Considerations and Applications in Physical Anthropology. *Yearbook of Physical Anthropology*, 31, 115-142.
- Ellison, P. (2003). Energetics and reproductive effort. *American Journal of Human Biology*, 15(3), 342-351.
- Elmlinger, M., Kuehnelt, W., Wormstall, H., & Doeller, P. (2005). Reference intervals for testosterone, androstenedione and SHBG levels in healthy females and males from birth until old age. *Clinical Laboratory*, 51(11-12), 625-632.
- Farmer, K. (2002). Pan African Sanctuary Alliance: Status and range of activities for great ape conservation. *American Journal of Primatology*, 58, 117-132.
- Fitch-Snyder, H., & Jurke, M. (2003). Reproductive Patterns in Pygmy Lorises (*Nycticebus pygmaeus*): Behavioral and Physiological Correlates of Gonadal Activity. *Zoo Biology*, 22, 15-32.
- Furuichi, T. (2009). Factors underlying party size differences between chimpanzees and bonobos: a review and hypotheses for future study. *Primates*, 50(3), 197-209.
- Furuichi, T. (2011). Female Contributions to the Peaceful Nature of Bonobo Society. *Evolutionary Anthropology*, 20(4), 131-142. doi: 10.1002/evan.20308
- Furuichi, T., Idani, G., Ihobe, H., Kuroda, S., Kitamura, K., Mori, A., . . . Kano, T. (1998). Population dynamics of wild bonobos (*Pan paniscus*) at Wamba. *International Journal of Primatology*, 19(6), 1029-1043.
- Furuichi, T., & Ihobe, H. (1994). Variation in male relationships in bonobos and chimpanzees. *Behaviour*, 130, 211-228.
- Gesquiere, L., Altmann, J., Khan, M., Couret, J., Yu, J., Endres, C., . . . Wango, E. (2005). Coming of Age: Steroid Hormones of Wild Immature Baboons (*Papio cynocephalus*). *American Journal of Primatology*, 67(1), 83-100.
- Ginther, A. J., Carlson, A. A., Ziegler, T. E., & Snowdon, C. T. (2002). Neonatal and pubertal development in males of a cooperatively breeding primate, the cotton-top tamarin (*Saguinus oedipus oedipus*). *Biology of Reproduction*, 66(2), 282-290.

- Giovannucci, E., Stampfer, M. J., Krithivas, K., Brown, M., Brufsky, A., Talcott, J., . . . Kantoff, P. W. (1997). The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 94(7), 3320-3323. doi: 10.1073/pnas.94.7.3320
- Gould, L., & Ziegler, T. (2007). Variation in fecal testosterone levels, inter-male aggression, dominance rank and age during mating and post-mating periods in wild adult male ring-tailed Lemurs (*Lemur catta*). *American Journal of Primatology*, 69(12), 1325-1339.
- Granger, D., Shirtcliff, E., Booth, A., Kivlighan, K., & Schwartz, E. (2004). The “trouble” with salivary testosterone. *Psychoneuroendocrinology*, 29, 1229-1240.
- Grether, W., & Yerkes, R. (1940). Weight norms and relations for chimpanzees. *American Journal of Physical Anthropology*, 27(2), 181-197.
- Hare, B., Wobber, V., & Wrangham, R. (in press). The self-domestication hypothesis: bonobo psychology evolved due to selection against aggression. *Animal Behaviour*.
- Hashimoto, C., & Furuichi, T. (2006). Comparison of behavioral sequence of copulation between chimpanzees and bonobos. *Primates*, 47(1), 51-55.
- Heintz, M. R., Santymire, R. M., Parr, L. A., & Lonsdorf, E. V. (2011). Validation of a Cortisol Enzyme Immunoassay and Characterization of Salivary Cortisol Circadian Rhythm in Chimpanzees (*Pan troglodytes*). *American Journal of Primatology*, 73(9), 903-908. doi: 10.1002/ajp.20960
- Higham, J. P., Vitale, A. B., Rivera, A. M., Ayala, J. E., & Maestripieri, D. (2010). Measuring salivary analytes from free-ranging monkeys. *Physiology & Behavior*, 101(5), 601-607. doi: 10.1016/j.physbeh.2010.09.003
- Hong, K. W., Hibino, E., Takenaka, O., Hayasaka, I., Murayama, Y., Ito, S., & Inoue-Murayama, M. (2006). Comparison of androgen receptor CAG and GGN repeat length polymorphism in humans and apes. *Primates*, 47(3), 248-254. doi: 10.1007/s10329-005-0174-4
- Ihobe, H. (1997). Non-antagonistic Relations Between Wild Bonobos and Two Species of Guenons. *Primates*, 38(4), 351-357.
- Kano, T. (1989). The sexual behavior of pygmy chimpanzees. In P. Heltne & L. Marquardt (Eds.), *Understanding chimpanzees* (pp. 176-183). Cambridge: Harvard University Press.
- Kano, T. (1992). *The Last Ape: Pygmy Chimpanzee Behavior and Ecology*. Stanford, California: Stanford University Press.
- Ketterson, E. D., & Nolan, V. (1992). Hormones and life histories - an integrative approach. *American Naturalist*, 140, S33-S62.

- Kondo, M., Udon, T., Jin, W., Shimizu, K., Funakoshi, M., Itoh, M., . . . Taya, K. (2000). Changes in plasma concentrations of inhibin A and inhibin B throughout sexual maturation in the male chimpanzee. *Endocrine Journal*, 47(6), 707-714.
- Kuhar, C. W., Bettinger, T. L., & Laudenslager, M. L. (2005). Salivary cortisol and behaviour in an all-male group of western lowland gorillas (*Gorilla g. gorilla*). *Animal Welfare*, 14(3), 187-193.
- Kuroda, S. (1984). Interaction over food among pygmy chimpanzees. In R. Susman (Ed.), *The pygmy chimpanzee: evolutionary biology and behavior* (pp. 301-324). New York: Plenum Press.
- Kuroda, S. (1989). Developmental retardation and behavioral characteristics of pygmy chimpanzees. In P. Heltne & L. Marquardt (Eds.), *Understanding chimpanzees* (pp. 184-193). Cambridge, Massachusetts: Harvard University Press.
- Kutsukake, N., Ikeda, K., Honma, S., Teramoto, M., Mori, Y., Hayasaka, I., . . . Hasegawa, T. (2009). Validation of Salivary Cortisol and Testosterone Assays in Chimpanzees by Liquid Chromatography-Tandem Mass Spectrometry. *American Journal of Primatology*, 71(8), 696-706. doi: 10.1002/ajp.20708
- Kuykendall, K. L., Mahoney, C. J., & Conroy, G. C. (1992). Probit and survival analysis of tooth emergence ages in a mixed-longitudinal sample of chimpanzees (*Pan troglodytes*). *American Journal of Physical Anthropology*, 89(3), 379-399. doi: 10.1002/ajp.1330890310
- Leigh, S., & Shea, B. (1996). Ontogeny of body size variation in African apes. *American Journal of Physical Anthropology*, 99(1), 43-65.
- Lieberman, D. E., Carlo, J., Ponce de Leon, M., & Zollikofer, C. (2007). A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. *Journal of Human Evolution*, 52, 647-662.
- Lipson, S., & Ellison, P. (1989). Development of Protocols for the Application of Salivary Steroid Analyses to Field Conditions. *American Journal of Human Biology*, 1(3), 249-255.
- Lwanga, J. S., Struhsaker, T. T., Struhsaker, P. J., Butynski, T. M., & Mitani, J. C. (2011). Primate Population Dynamics Over 32.9 Years at Ngogo, Kibale National Park, Uganda. *American Journal of Primatology*, 73(10), 997-1011. doi: 10.1002/ajp.20965
- Maggioncalda, A., Sapolsky, R., & Czekala, N. (1999). Reproductive Hormone Profiles in Captive Male Orangutans: Implications for Understanding Developmental Arrest. *American Journal of Physical Anthropology*, 109(1), 19-32.
- Mann, D. R., Akinbami, M. A., Gould, K. G., Paul, K., & Wallen, K. (1998). Sexual maturation in male rhesus monkeys: importance of neonatal testosterone exposure and social rank. *Journal of Endocrinology*, 156(3), 493-501.

- Mann, D. R., & Fraser, H. M. (1996). The neonatal period: A critical interval in male primate development. *Journal of Endocrinology*, 149(2), 191-197.
- Marshall, A., & Hohmann, G. (2005). Urinary Testosterone Levels of Wild Male Bonobos (*Pan paniscus*) in the Lomako Forest, Democratic Republic of Congo. *American Journal of Primatology*, 65(1), 87-92.
- Marson, J., Meuris, S., Cooper, R. W., & Jouannet, P. (1991a). Puberty in the male chimpanzee: progressive maturation of semen characteristics. *Biology of Reproduction*, 44(3), 448-455.
- Marson, J., Meuris, S., Cooper, R. W., & Jouannet, P. (1991b). Puberty in the male chimpanzee: time-related variations in luteinizing hormone, follicle-stimulating hormone, and testosterone. *Biology of Reproduction*, 44(3), 456-460.
- Martin, D. E., Swenson, R. B., & Collins, D. C. (1977). Correlation of serum testosterone levels with age in male chimpanzees. *Steroids*, 29(4), 471-481.
- McGlothlin, J., & Ketterson, E. (2008). Hormone-mediated suites as adaptations and evolutionary constraints *Philosophical Transactions of the Royal Society B*, 363(1497), 1611-1620.
- Muehlenbein, M. P., & Bribiescas, R. G. (2005). Testosterone-mediated immune functions and male life histories. *American Journal of Human Biology*, 17(5), 527-558. doi: 10.1002/ajhb.20419
- Mulavwa, M. N., Yangozene, K., Yamba-Yamba, M., Motema-Salo, B., Mwanza, N. N., & Furuichi, T. (2010). Nest Groups of Wild Bonobos at Wamba: Selection of Vegetation and Tree Species and Relationships between Nest Group Size and Party Size. *American Journal of Primatology*, 72(7), 575-586. doi: 10.1002/ajp.20810
- Muller, M. (2002). Agonistic relations among Kanyawara chimpanzees. In C. Boesch, G. Hohmann & L. Marchant (Eds.), *Behavioural diversity in chimpanzees and bonobos* (pp. 112-123). Cambridge: Cambridge University Press.
- Muller, M., Kahlenberg, S., Emery Thompson, M., & Wrangham, R. (2007). Male coercion and the costs of promiscuous mating for female chimpanzees. *Proceedings of the Royal Society of London: B*, 274, 1009-1014.
- Muller, M., & Lipson, S. (2003). Diurnal patterns of urinary steroid excretion in wild chimpanzees. *American Journal of Primatology*, 60(4), 161-166.
- Muller, M., & Wrangham, R. (2004). Dominance, aggression and testosterone in wild chimpanzees: a test of the 'challenge hypothesis'. *Animal Behaviour*, 67, 113-123.
- Nadler, R. D., Wallis, J., Rothmeyer, C., Cooper, R. W., & Baulieu, E. E. (1987). Hormones and behavior of prepubertal and peripubertal chimpanzees. *Hormones and Behavior*, 21(1), 118-131. doi: 10.1016/0018-506x(87)90037-7

- Parish, A., & de Waal, F. (2000). The Other “Closest Living Relative” How Bonobos (*Pan paniscus*) Challenge Traditional Assumptions about Females, Dominance, Intra- and Intersexual Interactions, and Hominid Evolution. *Annals of the New York Academy of Sciences*, 907(1), 97-113.
- Pearson, B., Judge, P., & Reeder, D. (2008). Effectiveness of saliva collection and enzyme-immunoassay for the quantification of cortisol in socially housed baboons. *American Journal of Primatology*, 70(12), 1145-1151.
- Pusey, A. (1990). Behavioural Changes At Adolescence in Chimpanzees. *Behaviour*, 115(3-4), 203-246.
- Pusey, A. E., Oehlert, G. W., Williams, J. M., & Goodall, J. (2005). Influence of ecological and social factors on body mass of wild chimpanzees. *International Journal of Primatology*, 26(1), 3-31. doi: 10.1007/s10764-005-0721-2
- Reichert, K., Heistermann, M., Hodges, J., Boesch, C., & Hohmann, G. (2002). What Females Tell Males About Their Reproductive Status: Are Morphological and Behavioural Cues Reliable Signals of Ovulation in Bonobos (*Pan paniscus*)? *Ethology*, 108, 583-600.
- Sannen, A., Heistermann, M., van Elsacker, L., Moehle, U., & Eens, M. (2003). Urinary testosterone metabolite levels in bonobos: a comparison with chimpanzees in relation to social system. *Behaviour*, 140, 683-696.
- Sannen, A., Van Elsacker, L., Heistermann, M., & Eens, M. (2005). Certain Aspects of Bonobo Female Sexual Repertoire Are Related to Urinary Testosterone Metabolite Levels. *Folia Primatologica*, 76(1), 21-32.
- Sapolsky, R. (1993). The physiology of dominance in stable versus unstable social hierarchies. In W. Mason & S. Mendoza (Eds.), *Primate social conflict* (pp. 171-204). Albany, NY: SUNY.
- Seraphin, S. B., Whitten, P. L., & Reynolds, V. (2008). The influence of age on fecal steroid hormone levels in male Budongo Forest Chimpanzees (*Pan troglodytes schweinfurthii*). *American Journal of Primatology*, 70(7), 661-669. doi: 10.1002/ajp.20541
- Setchell, J. M., & Dixson, A. F. (2002). Developmental variables and dominance rank in adolescent male mandrills (*Mandrillus sphinx*). *American Journal of Primatology*, 56(1), 9-25.
- Sharpe, R. M., Fraser, H. M., Brougham, M. F. H., McKinnell, C., Morris, K. D., Kelnar, C. J. H., . . . Walker, M. (2003). Role of the neonatal period of pituitary-testicular activity in germ cell proliferation and differentiation in the primate testis. *Human Reproduction*, 18(10), 2110-2117. doi: 10.1093/humrep/deg413
- Shea, B. (1984). An allometric perspective on the morphological and evolutionary relationships between pygmy (*Pan paniscus*) and common (*Pan troglodytes*) chimpanzees. In R. Susman

- (Ed.), *The pygmy chimpanzee: evolutionary biology and behavior* (pp. 89-130). New York: Plenum Press.
- Sirugo, G., Deinard, A. S., Kidd, J. R., & Kidd, K. K. (1997). Survey of maximum CTG/CAG repeat lengths in humans and non-human primates: Total genome scan in populations using the Repeat Expansion Detection method. *Human Molecular Genetics*, 6(3), 403-408.
- Smith, B., Crummett, T., & Brandt, K. (1994). Ages of eruption of primate teeth: a compendium for aging individuals and comparing life histories. *Yearbook of Physical Anthropology*, 37, 177-231.
- Stearns, S. (1992). *The Evolution of Life Histories*. Oxford: Oxford University Press.
- Surbeck, M., Deschner, T., Schubert, G., Weltring, A., & Hohmann, G. (in press). Mate competition, testosterone and intersexual relationships in bonobos, *Pan paniscus*. *Animal Behaviour*.
- Surbeck, M., Mundry, R., & Hohmann, G. (2011). Mothers matter! Maternal support, dominance status and mating success in male bonobos (*Pan paniscus*). *Proceedings of the Royal Society B-Biological Sciences*, 278(1705), 590-598. doi: 10.1098/rspb.2010.1572
- Talge, N., Donzella, B., Kryzer, E., Gierens, A., & Gunnar, M. (2005). It's not that bad: Error introduced by oral stimulants in salivary cortisol research. *Developmental Psychobiology*, 47(4), 369-376.
- Van Cauter, E. (1990). Diurnal and ultradian rhythms in human endocrine function - a minireview. *Hormone Research*, 34(2), 45-53. doi: 10.1159/000181794
- Vittek, J., Lhommedieu, D. G., Gordon, G. G., Rappaport, S. C., & Southren, A. L. (1985). Direct radioimmunoassay (RIA) of salivary testosterone - correlation with free and total serum testosterone. *Life Sciences*, 37(8), 711-716. doi: 10.1016/0024-3205(85)90540-5
- von Engelhardt, N., Kappeler, P., & Heistermann, M. (2000). Androgen levels and female social dominance in *Lemur catta*. *Proceedings of the Royal Society of London: B*, 267(1452), 1533-1539.
- Wingfield, J., Hegner, R., Dufty, R., & Ball, G. (1990). The "challenge hypothesis": theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. *American Naturalist*, 136, 829-846.
- Wingfield, J. C., Lynn, S. E., & Soma, K. K. (2001). Avoiding the 'costs' of testosterone: Ecological bases of hormone-behavior interactions. *Brain Behavior and Evolution*, 57(5), 239-251.
- Winter, J., Faiman, C., Hobson, W., & Reyes, F. (1980). The endocrine basis of sexual development in the chimpanzee. *Journal of Reproduction and Fertility, Supplement*, 28, 131-138.

- Wobber, V., & Hare, B. (2011). Psychological health of orphan bonobos and chimpanzees in African sanctuaries. *PLOS One*, 6(6), e17147.
- Wobber, V., Hare, B., Maboto, J., Lipson, S., Wrangham, R., & Ellison, P. (2010). Differential changes in steroid hormones before competition in bonobos and chimpanzees. *Proceedings of the National Academy of Sciences*, 107(28), 12457-12462.
- Wobber, V., Hare, B., & Wrangham, R. (In preparation). Patterns of weight growth in bonobos and chimpanzees from infancy to adulthood.
- Wobber, V., Herrmann, E., Hare, B., Wrangham, R., & Tomasello, M. (submitted). Differences in the early cognitive development of children and great apes.
- Wobber, V., Wrangham, R., & Hare, B. (2010a). Application of the heterochrony framework to the study of behavior and cognition. *Communicative and Integrative Biology*, 3, 1-3.
- Wobber, V., Wrangham, R., & Hare, B. (2010b). Bonobos exhibit delayed development of social behavior and cognition relative to chimpanzees. *Current Biology*, 20(3), 226-230.
- Won, Y., & Hey, J. (2005). Divergence Population Genetics of Chimpanzees. *Molecular Biology and Evolution*, 22(2), 297-307.
- Woods, V., & Hare, B. (2011). Bonobo but not chimpanzee infants use socio-sexual contact with peers. *Primates*, 52(2), 111-116.
- Wrangham, R. (2002). The cost of sexual attraction: is there a trade-off in female *Pan* between sex appeal and received coercion? . In C. Boesch, G. Hohmann & L. Marchant (Eds.), *Behavioural Diversity in Chimpanzees and Bonobos* (pp. 204-220). Cambridge, UK: Cambridge University Press.
- Wrangham, R., & Pilbeam, D. (2001). African apes as time machines. In B. Galdikas, N. Briggs, L. Sheeran, G. Shapiro & J. Goodall (Eds.), *All apes great and small* (Vol. 1, pp. 5-18). New York: Kluwer Academic/Plenum Publishers.
- Young, L. G., Gould, K. G., & Smithwick, E. B. (1993). Selected endocrine parameters of the adult male chimpanzee. *American Journal of Primatology*, 31(4), 287-297.
- Zuk, M., Johnsen, T. S., & Maclarty, T. (1995). Endocrine-immune interactions, ornaments, and mate choice in red jungle fowl. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 260(1358), 205-210. doi: 10.1098/rspb.1995.0081

Chapter 5:
Differential changes in steroid hormones prior to competition in bonobos and chimpanzees

INTRODUCTION

Across species, including humans, males engaged in competition tend to show acute shifts in their levels of steroid hormones such as testosterone and cortisol. These hormones change in a matter of minutes surrounding a competitive event, in anticipation of the competition and in response to its outcome (Booth, Shelley, Mazur, Tharp, & Kittok, 1989; Mazur & Booth, 1998). In humans, men normally demonstrate an increase in cortisol prior to competition (Alix-Sy, Le Scanff, & Filaire, 2008; Filaire, Maso, Sagnol, Ferrand, & Lac, 2001). After the competition male winners tend to maintain their testosterone levels while male losers' testosterone decreases (Elias, 1981; O. C. Schultheiss, Campbell, & McClelland, 1999). In other animals, competing males show similar rapid changes in glucocorticoids and testosterone, as these hormones are thought to mediate energy allocation toward mating effort across species (Bernstein, Rose, & Gordon, 1974; Oyegbile & Marler, 2005; Wingfield, Hegner, Dufty, & Ball, 1990; Wingfield & Sapolsky, 2003). Because competition for overt markers of status and mating opportunities is more relevant to males, these effects are less consistent in females (Booth, Granger, Mazur, & Kivlighan, 2006; Filaire, Alix, Ferrand, & Verger, 2009; Kivlighan, Granger, & Booth, 2005; Suay et al., 1999). Beyond these typical patterns, there is also high variability within and between species in the nature of the hormonal shifts surrounding competition that may be shaped by the psychology underlying competitive behavior.

Two main psychological factors have been implicated in governing the endocrine changes surrounding competition within and between species: implicit power motive and coping style. Implicit power motive, in the human literature, denotes an individual's drive to achieve

high status (see (Stanton & Schultheiss, 2009) for a review). Men with a high power motive are more likely to show increases in testosterone prior to competition and stronger shifts in testosterone and glucocorticoids post-competition according to the outcome (O. Schultheiss et al., 2005; Wirth, Welsh, & Schultheiss, 2006). Implicit power motives may drive between-species differences as well. In a comparison of a territorial and non-territorial mouse species, only the territorial species showed an increase in testosterone after a competitive event, while the non-territorial species showed no significant changes in testosterone levels (Fuxjager & Marler, 2010). Coping style, on the other hand, quantifies how an individual responds physiologically across numerous stressful events, such as competition (Salvador & Costa, 2009). Individuals with a “passive” coping style are more likely to show greater glucocorticoid increases prior to the competition than those with an “active” coping style, who show a less marked increase in glucocorticoids (Koolhaas, de Boer, Buwalda, & van Reenen, 2007). Lines of mice bred for low aggression tend to exhibit passive coping styles, and the associated large glucocorticoid shift, more than lines of mice bred for high aggression (Veenema, Koolhaas, & De Kloet, 2004). These results suggest that appraisal of competition and the corresponding endocrine shifts surrounding competition vary between even closely related species according to the significance of competition in that species’ behavioral ecology.

In turn, humans’ responses to competition may also have been shaped by ecological pressures. Studying humans’ closest living relatives, chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*), can reveal the degree to which humans’ rapid hormonal shifts surrounding competition are unique. In addition, chimpanzees and bonobos differ markedly in their social behavior in the context of both competition and cooperation, providing a direct test of ecology’s influence on competitive behavior and endocrinology. Male chimpanzees exhibit more severe

aggression and more concern for dominance status than male bonobos (Hohmann, 2001; Kano, 1992; Muller & Wrangham, 2009; Wrangham, 1999). Male dominance hierarchies are more rigid and more strongly associated with basal testosterone levels among chimpanzees than among bonobos (Marshall & Hohmann, 2005; Muehlenbein, Watts, & Whitten, 2004; Muller, 2002; Muller & Wrangham, 2004; Vervaecke, de Vries, & van Elsacker, 2000). Thus, in the terms used in human competition research, chimpanzee males may show a stronger “power motive” than bonobo males. In contrast, bonobos are better able to cooperate than chimpanzees, sharing food more readily in the wild and in captive experiments (Fruth & Hohmann, 2002; Hare, Melis, Woods, Hastings, & Wrangham, 2007). Previous studies have shown that bonobos exhibit a rise in cortisol prior to a competition over limited amounts of food, with greater increases in cortisol when that food was visibly difficult to share (implying greater social stress) (Hohmann, Mundry, & Deschner, 2008). Because bonobo conflicts rarely escalate to severe aggression, we might classify bonobos as possessing a passive coping style – similar to the low-aggression mice.

Accordingly we tested the hypothesis that psychological differences in the appraisal of competition in chimpanzees and bonobos are associated with species differences in rapid endocrine shifts surrounding competition. We presented chimpanzees and bonobos with an identical experimental dyadic food competition and measured testosterone and cortisol levels prior to and after the competitive event. We made two separate predictions about how the species difference in endocrinology be manifested. These predictions apply principally to males, though we tested individuals of both sexes (Kivlighan, et al., 2005; Oyegbile & Marler, 2005).

Prediction 1: Bonobo males will show an anticipatory increase in cortisol, which will be more pronounced in situations of higher stress (manifested here as social uncertainty), indicative

of a passive coping style. Cortisol shifts will be less pronounced in chimpanzee males. Chimpanzees do exhibit rapid cortisol changes surrounding anesthesia (Anestis & Bribiescas, 2004), but we predict that in the competitive situation their cortisol will not shift as markedly as that of bonobo males.

Prediction 2: Chimpanzee males will show an anticipatory increase in testosterone and greater sensitivity to the outcome of the competition in both testosterone and cortisol relative to bonobo males, in line with their having a greater power motive.

Alternative hypotheses: Chimpanzees and bonobos will show similar responses to competition, or neither species will show significant endocrine shifts surrounding competition over food.

RESULTS

Prior to all food competitions (detailed below), subject pairs participated in a dominance test. The individual who obtained more food in this test was assigned the status of the “dominant” in the pair. Dominance in this test strongly predicted dominance in the food competitions (Appendix 5, Supplemental Results).

Each individual was tested as a member of only one pair. For each trial of the food competition, a controlled amount of food was placed in a specific configuration in a testing room. The subject pair viewed the placement of the food in an adjacent room, and was then released into the room and allowed to eat the food. After the pair finished eating, the experimenter immediately placed new food for the subsequent trial. 3 food competition trials were presented in sequence on a given day. If the dominant monopolized food on 2 or 3 of the trials on that day, this was scored as a “1” for the behavioral variable *outcome*, in denoting that

food was obtained asymmetrically. If the dominant obtained more food on 1 or 0 of the trials (e.g. individuals shared the food relatively equally or the dominant obtained less), a “0” was scored for *outcome*. Pairs participated in 3 days of testing, thus each individual was represented 3 times in the data set, once for each day of food competition (food configuration varied across days, as described in Appendix 5, Supplemental Methods).

Chimpanzees and bonobos did not differ in their relative frequencies of the *outcome* variable: the dominant monopolized significantly more food approximately 50% of the time in both species (a chi-square test showed that the proportions of *outcome* were not different across the two species). The two species did show significant differences in other more targeted behavioral measures of sharing in this task (Wobber, Wrangham, & Hare, 2010).

In addition to the paired food competitions, each subject was presented with a solo condition that replicated the procedure of the paired conditions exactly except that individuals were tested alone rather than in a pair. This condition served to measure individuals’ baseline hormone levels in the general test situation, without social interaction.

On each day of food competition, saliva samples were taken from both subjects immediately prior to the first trial, before the food was presented but after individuals were placed in their pairing. Samples were then collected again from both subjects 15 minutes after the third trial was finished (Elias, 1981; Gladue, Boechler, & McCaul, 1989). Saliva samples were analyzed for testosterone and cortisol using previously validated radioimmunoassay procedures (see Methods and (Lipson & Ellison, 1989)). The values of testosterone and cortisol were log-transformed to normalize the data and allow the use of parametric statistics.

Statistical analyses

To analyze differences between groups, we performed Generalized Linear Model (GLM) analyses. In all of these analyses, we controlled for the within-subject factor *individual*, since each individual was represented in the data set 3 times. For all models, we examined the main effects, 2-way interactions, and 3-way interactions (where applicable). We controlled for multiple comparisons by using Fisher's Least Significant Difference (LSD) procedure in post-hoc tests.

Cortisol

Pre-test cortisol

We first analyzed anticipatory effects and then looked at effects in response to the outcome of the test. For cortisol, all GLM analyses had *individual* as a within-subject factor and *species*, *sex*, and *outcome* as between-subject factors. A GLM analysis of log pre-test cortisol revealed that bonobos had significantly higher log pre-test cortisol than chimpanzees (Wald Chi-Square (1) = 11.62, $p = 0.001$), and showed an interaction between *sex* and *outcome* (Wald Chi-Square (1) = 6.04, $p = 0.014$), as well as a 3-way interaction between *species*, *sex*, and *outcome* (Wald Chi-Square (1) = 8.91, $p = 0.003$). Post-hoc tests revealed that among bonobo males, log pre-test cortisol was significantly higher when the dominant was going to obtain more food than when the two individuals in a pair were going to share/the dominant was going to obtain less (Fisher's LSD, $p = 0.001$). In contrast, for chimpanzee males there was no difference in cortisol levels across outcomes, nor were there any significant effects among females of either species. Bonobo males also had significantly higher levels of log pre-test cortisol than chimpanzee males when the dominant was going to monopolize more food (Fisher's LSD, $p < 0.001$) (Figure 5.1). Control analyses revealed that these patterns were present equally in dominants and

subordinates, independent of order (e.g. first versus last day of food competitions), and equivalent whether a male was paired with another male or paired with a female (Appendix 5, Supplemental Results).

To ensure that these results did not reflect anticipation of food being presented or baseline cortisol differences between individuals, we examined anticipatory cortisol as it departed from the solo condition (baseline) values of cortisol. Subjects' log pre-test cortisol values were highly correlated with their log pre-solo cortisol values (linear regression, $r^2 = 0.25$, $df = 162$, $p < 0.001$). We stored the unstandardized residuals of this regression as an index of how much an individual's pre-test cortisol value on a given test day departed from what would be predicted based on their pre-solo day cortisol level. If the residual was positive, this represented a value higher than baseline, while if it was negative, this value was lower than baseline.

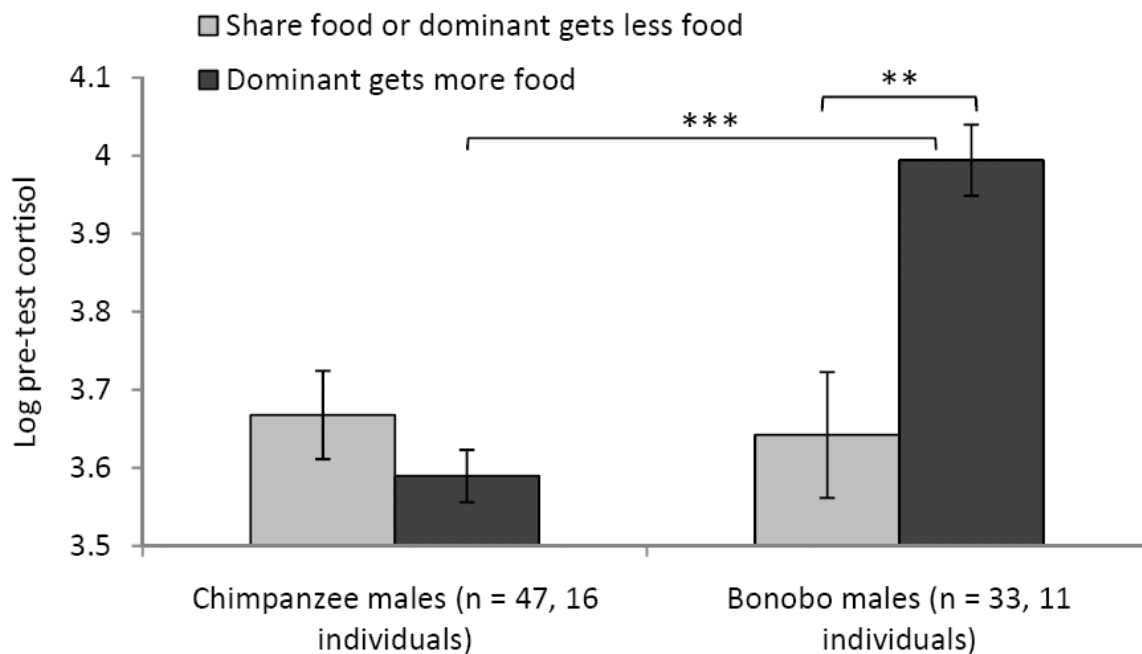


Figure 5.1. Pre-test log cortisol values according to species and outcome of the food competitions, males only. Bars denote standard error of the mean. * denotes $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

A GLM analysis on these residuals demonstrated that the main effects of each factor were not significant but there was the expected significant interaction between *species*, *sex*, and *outcome* (Wald Chi-Square (1) = 6.77, $p = 0.009$). Post-hoc tests showed that bonobo males increased in cortisol relative to their baseline when the dominant was going to obtain more food, and decreased in cortisol relative to baseline when the two individuals in a pair were going to share, with this creating a significant difference between these two outcomes (Fisher's LSD, $p = 0.001$). Again, there were no significant differences in chimpanzee males according to outcome and no differences among females. Bonobo males had significantly higher relative levels of cortisol than chimpanzee males when the dominant was going to monopolize significantly more food (Fisher's LSD, $p = 0.01$) (Appendix 5, Supplemental Results). These residual analyses indicate that bonobo males' cortisol levels were sensitive to their pairing, while those of chimpanzee males were not.

Post-test cortisol

We conducted a similar GLM analysis with the log post-test cortisol values. Bonobos exhibited significantly higher post-test cortisol than chimpanzees (Wald Chi-Square (1) = 25.55, $p < 0.001$), and we found a significant interaction between *species*, *sex*, and *outcome* (Wald Chi-Square (1) = 7.19, $p = 0.007$). Because these results paralleled those obtained using the pre-test cortisol values, we examined the relative contribution of the test events independent of the pre-test effects. We first ran a regression of log post-test cortisol values and log pre-test cortisol values (linear regression, $r^2 = 0.43$, $df = 214$, $p < 0.001$). We then used the unstandardized residuals of this regression to assess how much an individual's post-test cortisol level departed from his or her pre-test cortisol levels.

A GLM analysis of these post-test residuals revealed only a main effect of *species* (Wald Chi-Square (1) = 12.54, $p < 0.001$): bonobos' cortisol tended to increase over the course of the test regardless of outcome, while chimpanzees' cortisol levels did not change significantly.

Our results support a previous finding that anticipation of food competition elevates bonobo cortisol levels and that bonobos' cortisol increases differentially based on the predicted outcome of the competition (Hohmann, et al., 2008). The observed decrease in bonobo males' cortisol prior to sharing suggests that lower levels of arousal in bonobos may in part explain their tendency to voluntarily share food with other individuals (Hare & Kwetuenda, in press). The relative stability of cortisol in chimpanzee males could in theory have occurred either because they do not perform anticipatory appraisals prior to competition, or because such appraisals are not tied to a significant physiological response. We were able to test these alternatives, in addition to testing our main hypotheses regarding species differences in these acute endocrine shifts, with our analysis of testosterone.

Testosterone

Pre-test testosterone

As with cortisol, we began by analyzing anticipatory effects then moved to post-test effects. For testosterone we performed separate analyses by sex, given the known differences in testosterone levels between males and females in humans and other apes, and the prediction from the human literature and our cortisol results that the effects on this hormone would be more pronounced in males (Kivlighan, et al., 2005; Sannen, Heistermann, van Elsacker, Moehle, & Eens, 2003). Thus for testosterone, all GLM analyses had *individual* as a within subject factor and *species* and *outcome* as between-subject factors.

A GLM analysis of log pre-test testosterone showed that bonobo females' log pre-test testosterone was significantly higher than that of chimpanzee females (Wald Chi-Square (1) = 5.43, $p = 0.02$) (these baseline differences in hormone levels are discussed elsewhere (Wobber, Lipson, Hare, Wrangham, & Ellison, submitted)). In males, this analysis demonstrated a significant interaction between *species* and *outcome* (Wald Chi-Square (1) = 5.86, $p = 0.02$). Post-hoc tests in males revealed that among chimpanzees, males in pairs where the dominant was going to obtain more food had higher log pre-test testosterone than males in pairs where the two individuals were going to share (Fisher's LSD, $p = 0.03$). There were no distinctions in log pre-test testosterone across outcomes in bonobo males. As a result, when individuals shared, male chimpanzees had significantly lower log pre-test testosterone levels than bonobo males (Fisher's LSD, $p = 0.02$), with no species difference when the dominant monopolized the food (Figure 5.2). Similar to the cortisol results, there were no effects of dominance status, test day, or pair type on these effects (Appendix 5, Supplemental Results).

Again, we wanted to ensure that these pre-test testosterone values were not simply reflections of individuals' basal testosterone levels. We performed a regression analysis of the log pre-test day testosterone values and the log pre-solo day testosterone values (linear regression, $r^2 = 0.13$, $p < 0.001$, $df = 132$). We used the unstandardized residuals of this regression as an index of how much an individual's pre-test testosterone value on a given test day departed from baseline levels.

We performed a GLM analysis on these residuals and found that in females, there was a main effect of *species* (Wald Chi-Square (1) = 6.70, $p = 0.01$): bonobo females exhibited higher relative testosterone on test days while chimpanzee females did not. Bonobo males' relative testosterone levels also tended to be significantly higher on test days than chimpanzee males'

relative testosterone levels (Wald Chi-Square (1) = 4.22, $p = 0.04$), and we found a significant interaction between *species* and *outcome* in males as well (Wald Chi-Square (1) = 5.24, $p = 0.02$) (Appendix 5, Supplemental Results). Post-hoc tests revealed that male chimpanzees increased in testosterone relative to baseline when the dominant was going to obtain more food, and decreased relative to baseline when individuals were going to share, creating a significant difference between these two outcomes (Fisher's LSD $p = 0.02$). In contrast, for bonobo males there was no difference in relative testosterone levels between the two outcomes. The decrease in chimpanzee males' testosterone when they were going to share led to their relative testosterone levels being significantly lower than bonobo males' relative testosterone levels in these situations (Fisher's LSD $p = 0.003$).

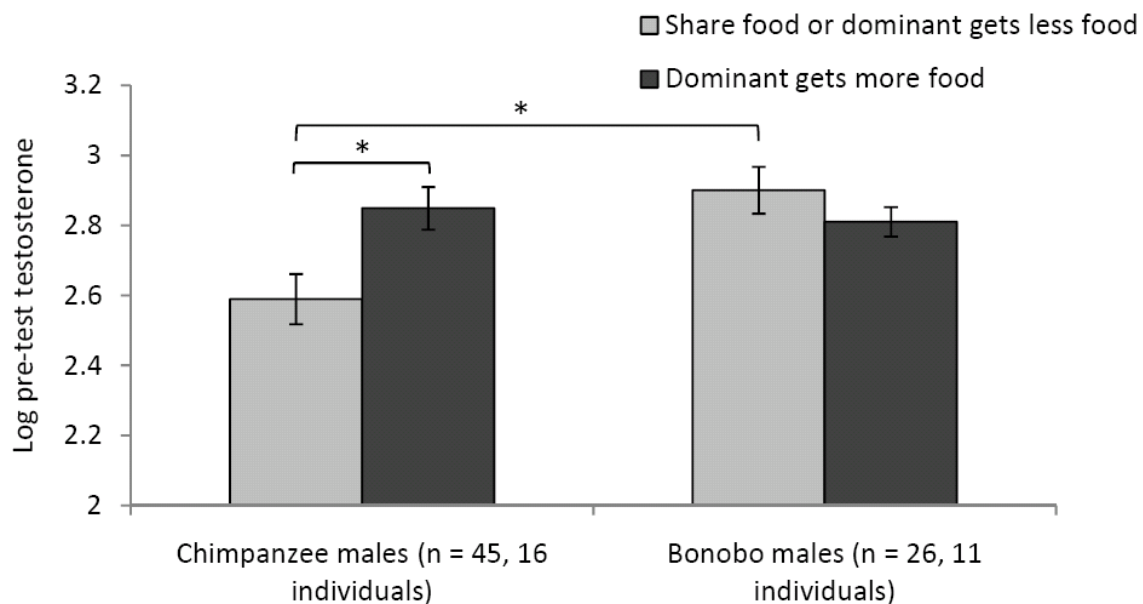


Figure 5.2. Pre-test log testosterone values according to species and outcome, males only. Bars denote standard error of the mean. * denotes $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Post-test testosterone

A GLM analysis of log post-test testosterone values revealed a significant effect of *species* in females (Wald Chi-Square (1) = 15.09, $p < 0.001$) and a significant interaction between *species* and *outcome* in males (Wald Chi-Square (1) = 4.50, $p = 0.03$). Because these results paralleled those found using the pre-test testosterone values, we again examined subjects' changes in response to the events of the competition as a function of their pre-test testosterone levels. Log post-test testosterone and log pre-test testosterone were highly correlated (linear regression, $r^2 = 0.44$, $p < 0.001$, $df = 178$). We used the unstandardized residuals of this regression to denote how much post-test testosterone values departed from pre-test testosterone values.

A GLM analysis of these post-test residuals revealed only a main effect of *species* in females (Wald Chi-Square (1) = 6.54, $p = 0.01$): bonobo females' testosterone values tended to increase over the course of the test, while those of chimpanzee females remained relatively constant. Competition did not significantly impact post-test testosterone levels in males of either species.

In contrast to the cortisol results, where bonobo males showed stronger anticipatory shifts based on outcome than did chimpanzees, the patterns of anticipatory change in testosterone were stronger in chimpanzee males. This refutes the suggestion that the greater cortisol response observed in bonobos might be due to their superior ability to predict the outcome of a food competition based on pairing compared to chimpanzees.

DISCUSSION

These results support the hypothesis that bonobos and chimpanzees differ significantly in endocrine shifts surrounding competition, and support both of our predictions regarding the

nature of that species difference. Bonobo males' cortisol increased relative to baseline prior to a competition where the dominant would obtain more food, and decreased prior to a competition where sharing would occur. Therefore, bonobos appeared to respond to the competition as a social stressor when food would not be shared, exhibiting a passive coping style and an associated large anticipatory shift in glucocorticoids. In the same context, chimpanzee males did not show any shifts in cortisol. Instead, their testosterone changed, showing either an anticipatory increase when the dominant was going to obtain more food or decrease when placed with a partner where sharing would occur. Bonobo males did not exhibit significant shifts in testosterone according to outcome. Thus chimpanzees appeared to view the competition as status-determining, similar to human men with a stronger power motive, with this driving shifts in testosterone.

These data demonstrate that between these two closely related species, there are important differences in the physiological response to competition that are correlated with differences in social behavior and ecology. Our findings provide the first evidence for rapid endocrine changes in association with competition in chimpanzees, and corroborate previous evidence for pre-competition cortisol increases in bonobos (Hohmann, et al., 2008). These results suggest that after the divergence of chimpanzees and bonobos, selection against escalated aggression in bonobo males may have caused them to acquire a passive coping style (analogous to that observed in lines of mice bred for low aggression) (Veenema, et al., 2004; Wrangham & Pilbeam, 2001). Chimpanzees, in contrast, may have retained an ancestral state with strict hierarchies, where individuals possess a high drive to achieve dominance rank, or power motive, and show corresponding large shifts in testosterone (Vervaecke, et al., 2000; Wrangham & Pilbeam, 2001). Future research comparing chimpanzees and bonobos can further reveal the role

of hormones in the morphological, behavioral, and cognitive differences between the two species.

Interestingly, the observed endocrine shifts occurred prior to the competition, rather than after the test. While it is possible that the post-test sampling interval of 15 minutes was too short to observe post-test effects, responses to competition in human men have been observed in that length of time (Elias, 1981; Gladue, et al., 1989). Further, in a previous study, even one hour after a competition over food bonobos did not exhibit any increases in cortisol beyond their anticipatory increases (Hohmann, et al., 2008). It could be that chimpanzees and bonobos react much more slowly than humans, signifying a difference between apes and humans in the speed of endocrine response to wins or losses. Alternatively, we propose that the apes in our experiments anticipated the outcome of competition particularly easily given their mutual familiarity and ability to track each other's tolerance levels (Hare, et al., 2007; Melis, Hare, & Tomasello, 2006). Individuals did not frequently vocalize or engage in aggressive behavior during the competition. This suggests that the actual process of feeding may cause relatively less arousal in apes than the anticipation of feeding competition. In turn, this indicates that the patterns of anticipatory appraisal seen in humans are not unique to our species, but that our species' endocrine shifts in response to the outcomes of even relatively trivial competitions (such as a chess match) are derived (Mazur, Booth, & Dabbs, 1992).

Similar to what is seen in humans, we found the strongest effects of the competition on steroid hormones in males, whereas females did not exhibit any significant patterns. Steroid shifts surrounding competition in women are inconsistent across studies (Bateup, Booth, Shirtcliff, & Granger, 2002; Kivlighan, et al., 2005; van Anders & Watson, 2007). This indicates

that the pattern of minimal response by women to psychological status competitions or stressors may be an ancient hominoid trait.

Overall, the present results suggest that our closest living relatives have the capacity to anticipate and appraise the results of dyadic food competitions and that their physiology changes accordingly. Further, they support the hypothesis that species differences in the ecology of competitive behavior shape the endocrinology of competition, extending this model into non-human primates. These findings suggest that independent mechanisms govern the sensitivity of testosterone and cortisol to competition, and that distinct factors may affect anticipatory versus response shifts in apes and humans. Future species comparisons can continue to illuminate how ecology has shaped species differences in behavioral endocrinology, including the selection pressures acting in human evolution.

METHODS

Subjects

The subjects for this experiment were 24 bonobos (median age 8 years, range 4 to 23 years) living at Lola ya Bonobo Sanctuary in the Democratic Republic of Congo and 33 chimpanzees (median age 7 years, range 5 to 19 years) living at Tchimpounga Chimpanzee Sanctuary in the Congo Republic (there was no species difference in subject age, Mann-Whitney U). Within bonobos, 11 males and 12 females were sampled for steroid analysis, but enough saliva volume for testosterone analysis was only obtainable for 7 of these females. One bonobo male participated in the behavioral testing but did not provide a sufficient volume of saliva to perform either testosterone or cortisol analysis. 16 male and 17 female chimpanzees were

sampled for both cortisol and testosterone. More information about the subjects' living circumstances and rearing histories can be found in Appendix 5, Supplemental Methods.

There were 12 bonobo pairs and 24 chimpanzee pairs tested. Equal numbers of adult and juvenile pairs were tested in each species. The age of the two individuals in a pair was matched as closely as possible. Equal numbers of male-male, male-female, and female-female dyads were tested in each species. Certain chimpanzees participated in repeated pairs, but for the analyses reported here, only the first pair that these subjects participated in was used. The second individual in that subject's repeated pair was still included as a subject, resulting in 24 bonobos and 33 chimpanzees in the sample.

Coding of behavioral variables

All testing was videotaped. Videos of behavior in the test were coded by the first author. For reliability, a randomly chosen 20% of the trials were also coded by a second coder who was blind to the hypotheses of the study. The reliability for the *outcome* measure was excellent (Cohen's kappa = 0.88, $p < 0.001$). *Outcome* was usually consistent within a given pair, in that a dominant would obtain more food (or not) across each of the 3 food competition conditions, but could vary across condition within each pair. Importantly, the scores for *outcome* were the same for both individuals in the pair (the dominant and the subordinate), thus this variable represented asymmetry versus sharing in the distribution of feeding rather than a win or loss.

In each pair, one subject was given the solo condition on a day prior to the three food competition days, and the other member of that pair was given the solo condition on a day after the three food competition days, thus counterbalancing any effect of test experience on the hormone values in the solo condition.

Hormonal sampling

During the 15-minute post-competition interval, subjects remained in the testing room with their partners. Subjects were observed so that they could not ingest any food or fecal matter during this time, making it unlikely that food debris from the test or other contaminants were present in the individuals' mouths at the time of sample collection. In the solo condition, subjects were alone when their pre-test sample was taken, and they waited alone in the testing room for the 15 minute post-test interval.

To control for the effect of time of day on hormone levels, a given pair was always run within the same two-hour time window across all three test sessions. The solo conditions for the individuals in a pair were also run within the same two-hour window. Thus any departure from the solo condition (baseline) values in the test was not due to circadian variation. The number of pairs in each age and sex category tested in the morning and the afternoon was counterbalanced as best as possible. It was not feasible to do this for all pairs due to constraints of the testing facilities. All tests were carried out between 8:00 AM and 4:00 PM. Subjects were awake for several hours prior to the start of the tests, reducing the probability that the high levels of steroids observed in apes upon waking influenced results (Muller & Lipson, 2003). These tests were not physiologically demanding for subjects, making it unlikely that exertion affected the endocrine changes seen. Further, any changes that occurred as a result of being fed would also have been present in both the solo condition and paired conditions, thus these potential effects were controlled for in the residual analyses.

Saliva samples were collected while subjects were in the test rooms, highly familiar rooms that individuals slept in each night. To collect a sample, the experimenter or caretaker first washed and disinfected his/her hands, then poured ground Sweet Tarts candy onto a cotton

round. The experimenter/caretaker then stood next to the mesh of the dormitory, and if the subject approached her, she placed the cotton round inside the subject's lip so that it could suck on the cotton and ingest the Sweet Tarts while the cotton absorbed its saliva. The experimenter held on to the cotton throughout the collection procedure rather than allowing the subject to take the cotton itself, to prevent potential contamination from fecal matter on subjects' hands. Once the cotton round had taken in enough saliva, it was placed into a syringe and squeezed to express the saliva into a test tube. Though using cotton as a collection implement may affect measurements of steroids, cotton has been shown to introduce fairly uniform rates of error across samples (Granger, Shirtcliff, Booth, Kivlighan, & Schwartz, 2004; Smider et al., 2002). This means that while the absolute results presented here might not be comparable to those obtained without stimulation, the comparisons within this subject pool are effective since the method was consistent across subjects. The collection period for any particular sample did not span longer than 20 minutes.

Sweet Tarts were used to stimulate saliva because they have been shown not to alter measurements of cortisol in humans (Smider, et al., 2002; Talge, Donzella, Kryzer, Gierens, & Gunnar, 2005). We performed control analyses on a small sample of human men and women to assess whether ingesting Sweet Tarts affected measurements of testosterone. Among 5 individuals, there was no significant change in testosterone levels in a saliva sample taken prior to Sweet Tarts ingestion and one taken immediately after ingestion of several Sweet Tarts (Wilcoxon signed-ranks test, $p = 0.50$). This suggests that Sweet Tarts have little impact on the measurement of testosterone using this radioimmunoassay procedure.

Fifty microliters of 0.1% sodium azide solution was added to the ape saliva samples immediately after collection to prevent contamination and to allow samples to be kept at room

temperature until they were returned to the laboratory (Lipson & Ellison, 1989). The saliva samples were analyzed in the Reproductive Ecology Laboratory at Harvard University. Salivary testosterone measurements were made using an I-125 based radioimmunoassay kit (#4100, Diagnostic Systems Laboratories, Webster, TX, USA) with the following modifications: standards were prepared in assay buffer and run at six concentrations from 2 to 375 pg/ml. Samples were added in 100 μ l amounts together with 300 μ l of assay buffer. First antibody (20 μ l) and labeled steroid (50 μ l) were added to each tube to yield a total reaction volume of 470 μ l per tube. After overnight incubation at 4° C, 500 μ l of second antibody was added to each reaction tube. Reaction tubes were subsequently centrifuged for 45 minutes; after aspiration of the supernatant, tubes were counted in a gamma counter for two minutes. In pilot assays, the ape testosterone values using the standard aliquot for human assays (200 μ l) were too high to be readable in the assay range. Thus, we used only 100 μ l of the chimpanzee and bonobo saliva for the T assays, with the same standard curve as employed in the human testosterone radioimmunoassay protocol.

Salivary cortisol measurements were made using an I-125 based radioimmunoassay kit (#2000, Diagnostic Systems Laboratories, Webster, TX, USA) with the following modifications: Standards were prepared in assay buffer and run at six concentrations from 35 to 2000 pg/ml. Samples were added in 25 μ l amounts together with 200 μ l of assay buffer. Antibody complex and labeled steroid were diluted 1:2 and added to each tube in 150 μ l amounts to yield a total reaction volume of 525 μ l per tube. After overnight incubation at 4° C, 500 μ l of second antibody was added to each reaction tube. Reaction tubes were subsequently centrifuged for 45 minutes; after aspiration of the supernatant, tubes were counted in a gamma counter for two minutes.

The average intra-assay coefficient of variation (CV) was 8% for testosterone and 8% for cortisol, and average inter-assay CV was 16% for testosterone and 25% for cortisol. Though this inter-assay CV for cortisol is on the higher end of the acceptable range, all of the samples for a given individual were run in the same assay, meaning that any within-individual variation would not have been affected by inter-assay variation. We counter-balanced the individuals whose samples were run in each assay according to species, sex, and age.

Chapter 5 Literature Cited.

- Alix-Sy, D., Le Scanff, C., & Filaire, E. (2008). Psychophysiological responses in the pre-competition period in elite soccer players. *Journal of Sports Science and Medicine*, 7(4), 446-454.
- Anestis, S., & Bribiescas, R. (2004). Rapid changes in chimpanzee (*Pan troglodytes*) urinary cortisol excretion. *Hormones and Behavior*, 45(3), 209-213.
- Bateup, H., Booth, A., Shirtcliff, E., & Granger, D. (2002). Testosterone, cortisol, and women's competition. *Evolution and Human Behavior*, 23(3), 181-192.
- Bernstein, I., Rose, R., & Gordon, T. (1974). Behavioral and environmental events influencing primate testosterone levels. *Journal of Human Evolution*, 3, 517-525.
- Booth, A., Granger, D. A., Mazur, A., & Kivlighan, K. T. (2006). Testosterone and social behavior. *Social Forces*, 85(1), 167-191.
- Booth, A., Shelley, G., Mazur, A., Tharp, G., & Kittok, R. (1989). Testosterone, and Winning and Losing in Human Competition. *Hormones and Behavior*, 23(4), 556-571.
- Elias, M. (1981). Serum Cortisol, Testosterone, and Testosterone-Binding Globulin Responses to Competitive Fighting in Human Males. *Aggressive Behavior*, 7(3), 215-224.
- Filaire, E., Alix, D., Ferrand, C., & Verger, M. (2009). Psychophysiological stress in tennis players during the first single match of a tournament *Psychoneuroendocrinology*, 34(1), 150-157.
- Filaire, E., Maso, F., Sagnol, M., Ferrand, C., & Lac, G. (2001). Anxiety, hormonal responses, and coping during a judo competition. *Aggressive Behavior*, 27(1), 55-63.
- Fruth, B., & Hohmann, G. (2002). How bonobos handle hunts and harvests: why share food? In C. Boesch, G. Hohmann & L. Marchant (Eds.), *Behavioural diversity in chimpanzees and bonobos* (pp. 231-243). Cambridge: Cambridge University Press.
- Fuxjager, M., & Marler, C. (2010). How and why the winner effect forms: influences of contest environment and species differences. *Behavioral Ecology*, 21(1), 37-45.
- Gladue, M., Boechler, M., & McCaul, K. (1989). Hormonal Response to Competition in Human Males. *Aggressive Behavior*, 15(6), 409-422.
- Granger, D., Shirtcliff, E., Booth, A., Kivlighan, K., & Schwartz, E. (2004). The "trouble" with salivary testosterone. *Psychoneuroendocrinology*, 29, 1229-1240.
- Hare, B., & Kwetuenda, S. (in press). Bonobos voluntarily share their own food with others. *Current Biology*.

- Hare, B., Melis, A. P., Woods, V., Hastings, S., & Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology*, 17(7), 619-623.
- Hohmann, G. (2001). Association and Social Interactions Between Strangers and Residents in Bonobos (*Pan paniscus*). *Primates*, 42(1), 91-99.
- Hohmann, G., Mundry, R., & Deschner, T. (2008). The Relationship Between Socio-Sexual Behavior and Salivary Cortisol in Bonobos: Tests of the Tension Regulation Hypothesis. *American Journal of Primatology*, 70(1), 1-10.
- Kano, T. (1992). *The Last Ape: Pygmy Chimpanzee Behavior and Ecology*. Stanford, California: Stanford University Press.
- Kivlighan, K., Granger, D., & Booth, A. (2005). Gender differences in testosterone and cortisol response to competition. *Psychoneuroendocrinology*, 30, 58-71.
- Koolhaas, J. M., de Boer, S. F., Buwalda, B., & van Reenen, K. (2007). Individual variation in coping with stress: A multidimensional approach of ultimate and proximate mechanisms. *Brain, Behavior and Evolution*, 70, 218-226. doi: 10.1159/000105485
- Lipson, S., & Ellison, P. (1989). Development of Protocols for the Application of Salivary Steroid Analyses to Field Conditions. *American Journal of Human Biology*, 1(3), 249-255.
- Marshall, A., & Hohmann, G. (2005). Urinary Testosterone Levels of Wild Male Bonobos (*Pan paniscus*) in the Lomako Forest, Democratic Republic of Congo. *American Journal of Primatology*, 65(1), 87-92.
- Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and Brain Sciences*, 21, 353-363.
- Mazur, A., Booth, A., & Dabbs, J. (1992). Testosterone and chess competition. *Social Psychology Quarterly*, 55, 70-77.
- Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science*, 311(5765), 1297-1300.
- Muehlenbein, M., Watts, D., & Whitten, P. (2004). Dominance Rank and Fecal Testosterone Levels in Adult Male Chimpanzees (*Pan troglodytes schweinfurthii*) at Ngogo, Kibale National Park, Uganda. *American Journal of Primatology*, 64(1), 71-82.
- Muller, M. (2002). Agonistic relations among Kanyawara chimpanzees. In C. Boesch, G. Hohmann & L. Marchant (Eds.), *Behavioural diversity in chimpanzees and bonobos* (pp. 112-123). Cambridge: Cambridge University Press.

- Muller, M., & Lipson, S. (2003). Diurnal patterns of urinary steroid excretion in wild chimpanzees. *American Journal of Primatology*, 60(4), 161-166.
- Muller, M., & Wrangham, R. (2004). Dominance, aggression and testosterone in wild chimpanzees: a test of the 'challenge hypothesis'. *Animal Behaviour*, 67, 113-123.
- Muller, M., & Wrangham, R. (2009). *Sexual Coercion in Primates and Humans: An Evolutionary Perspective on Male Aggression Against Females*. Cambridge, MA: Harvard University Press.
- Oyegbile, T., & Marler, C. (2005). Winning fights elevates testosterone levels in California mice and enhances future ability to win fights. *Hormones and Behavior*, 48(3), 259-267.
- Salvador, A., & Costa, R. (2009). Coping with competition: Neuroendocrine responses and cognitive variables. *Neuroscience and Biobehavioral Reviews*, 33, 160-170. doi: 10.1016/j.neubiorev.2008.09.005
- Sannen, A., Heistermann, M., van Elsacker, L., Moehle, U., & Eens, M. (2003). Urinary testosterone metabolite levels in bonobos: a comparison with chimpanzees in relation to social system. *Behaviour*, 140, 683-696.
- Schultheiss, O., Wirth, M., Torges, C., Pang, J., Villacorta, M., & Welsh, K. (2005). Effects of Implicit Power Motivation on Men's and Women's Implicit Learning and Testosterone Changes After Social Victory or Defeat. *Journal of Personality and Social Psychology*, 88(1), 174-188.
- Schultheiss, O. C., Campbell, K. L., & McClelland, D. C. (1999). Implicit power motivation moderates men's testosterone responses to imagined and real dominance success. *Hormones and Behavior*, 36(3), 234-241.
- Smider, N., Essex, M., Kalin, N., Buss, K., Klein, M., Davidson, R., & Goldsmith, H. (2002). Salivary cortisol as a predictor of socioemotional adjustment during Kindergarten: a prospective study. *Child Development*, 73(1), 75-92.
- Stanton, S., & Schultheiss, O. (2009). The hormonal correlates of implicit power motivation. *Journal of Research in Personality*, 43(5), 942-949.
- Suay, F., Salvador, A., Gonzalez-Bono, E., Sanchis, C., Martinez, M., Martinez-Sanchis, S., . . . Montoro, J. (1999). Effects of competition and its outcome on serum testosterone, cortisol and prolactin. *Psychoneuroendocrinology*, 24(5), 551-566.
- Talge, N., Donzella, B., Kryzer, E., Gierens, A., & Gunnar, M. (2005). It's not that bad: Error introduced by oral stimulants in salivary cortisol research. *Developmental Psychobiology*, 47(4), 369-376.
- van Anders, S., & Watson, N. (2007). Effects of ability- and chance-determined competition outcome on testosterone. *Physiology and Behavior*, 90(4), 634-642.

- Veenema, A. H., Koolhaas, J. M., & De Kloet, E. R. (2004). Basal and stress-induced differences in HPA axis, 5-HT responsiveness, and hippocampal cell proliferation in two mouse lines. *Annals of the New York Academy of Sciences*, 1018, 255-265. doi: 10.1196/annals.1296.030
- Vervaecke, H., de Vries, H., & van Elsacker, L. (2000). Dominance and its Behavioral Measures in a Captive Group of Bonobos (*Pan paniscus*). *International Journal of Primatology*, 21(1), 47-68.
- Wingfield, J., Hegner, R., Dufty, R., & Ball, G. (1990). The “challenge hypothesis”: theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. *American Naturalist*, 136, 829-846.
- Wingfield, J., & Sapolsky, R. (2003). Reproduction and resistance to stress: when and how. *Journal of Neuroendocrinology*, 15, 711-724.
- Wirth, M., Welsh, K., & Schultheiss, O. (2006). Salivary cortisol changes in humans after winning or losing a dominance contest depend on implicit power motivation. *Hormones and Behavior*, 49(3), 346-352.
- Wobber, V., Lipson, S., Hare, B., Wrangham, R., & Ellison, P. (submitted). Differences in the ontogeny of testosterone production in chimpanzees and bonobos.
- Wobber, V., Wrangham, R., & Hare, B. (2010). Bonobos exhibit delayed development of social behavior and cognition relative to chimpanzees. *Current Biology*, 20(3), 226-230.
- Wrangham, R. (1999). Evolution of coalitionary killing. *Yearbook of Physical Anthropology*, 42, 1-30.
- Wrangham, R., & Pilbeam, D. (2001). African apes as time machines. In B. Galdikas, N. Briggs, L. Sheeran, G. Shapiro & J. Goodall (Eds.), *All apes great and small* (Vol. 1, pp. 5-18). New York: Kluwer Academic/Plenum Publishers.

Chapter 6: **Differences in the early cognitive development of children and great apes**

INTRODUCTION

Recent research has established many important cognitive similarities and differences between humans and their closest living relatives, the great apes (Lonsdorf, Ross, & Matsuzawa, 2011; Tomasello, 2009; Tomasello & Call, 1997; Whiten et al., 1999). However, most of this research compares human children to great ape adults, and we know from recent work in developmental biology that many, if not most, important differences between closely related species occur via differences in developmental patterning (Arthur, 2002; Carroll, 2003). What is needed for a fuller and more complete description and explanation, therefore, is a comparison of humans and great apes with respect to their early cognitive ontogenies.

There are several predictions that can be made regarding how human psychological ontogeny should compare to that of other ape species. First, evidence from humans' somatic growth might predict that we exhibit **Delayed Ontogeny** in our psychology. Humans demonstrate a marked slowdown in growth relative to other primates, exhibiting a later age of reproductive maturity and an extended period of juvenility and maternal dependence (Charnov & Berrigan, 1993; Hrdy, 2005; Walker, Hill, Burger, & Hurtado, 2006). Based on this evidence, many have argued that human psychology is characterized by paedomorphosis, or a retention of juvenile characteristics into adulthood, with an extended juvenile period in particular argued to confer additional psychological plasticity (Bjorklund & Green, 1992). Recent findings support this possibility in showing transcriptional "neoteny" of certain genes expressed in the human brain, with peak expression in these genes achieved at later ages in humans relative to other primate species (Liu et al., in press; Somel et al., 2009). A second possibility is that human psychological

ontogeny proceeds at the same pace as that of other primate species, which we will term **Fixed Ontogeny**. Evidence for this prediction again draws from the anthropological literature, where the argument has been put forth that humans' extended juvenile period is essential to accommodate greater skill learning of the complex human foraging environment and social world (Bruner, 1974; Kaplan, Hill, Lancaster, & Hurtado, 2000). This hypothesis presumes that the rate of skill learning is consistent and fixed across species, with only a change in the *duration* of the juvenile period able to allow heightened cognitive complexity among adults. The third and final possibility is that human psychological ontogeny occurs at a more rapid rate than among other primates, indicating **Accelerated Ontogeny**. This possibility is supported by comparative data on brain growth in humans and chimpanzees (*Pan troglodytes*), with humans showing a more marked increase in brain size during juvenility (Leigh, 2004; Lieberman, 2012; Robson & Wood, 2008; Vinicius, 2005). In addition, recent comparisons of children and adult apes have supported the idea that human cognitive development is accelerated relative to other species, identifying early-emerging capacities for social and cultural learning as the feature which enables humans' rapid cognitive ontogeny (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007). It might be the case that some domains of human cognitive development are accelerated while others are delayed or proceed at the same rate in comparison with the cognitive development of other great apes.

In order to discriminate among these alternatives, it is necessary to characterize both the *rate* and *pattern* of cognitive development in humans as compared to other apes across a broad array of cognitive skills. Prior studies have found similarities in sequences of sensorimotor and cognitive and logicomathematical development across a number of primate species (e.g., in Piagetian tasks of object permanence, categorization, and related skills; (Antinucci, 1990;

Langer, 2000; Natale, Antinucci, Spinozzi, & Poti, 1986)). One study has found patterns of socio-cognitive development in three human-reared (enculturated) chimpanzees that differ from patterns found in previous research with human children (Tomasello & Carpenter, 2005). Though these findings suggest potential aspects of convergence and divergence in the sequences of cognitive ontogeny in humans and other apes, they have relied on studies of only a few (at most three) individuals per species. This sample size might be adequate to characterize patterns of cognitive development in non-human primates, if one assumes that patterns of psychological ontogeny in these species are less flexible than those found in humans (Bjorklund & Green, 1992; Bruner, 1974). However, this question can only be fully addressed by testing a large enough sample of non-human ape infants to discriminate developmental differences from individual differences. Moreover, prior studies of comparative development have covered only relatively short periods of ontogeny and employed tasks confined to only one domain (e.g. social cognition or logicomathematical skills). No previous study has looked across the many years of early cognitive development with a broad battery of tasks. Finally, focus has been placed on humans' best-known relative, the chimpanzee, but no previous research has looked in detail at the cognitive development of humans' "other" closest relative, the bonobo (*Pan paniscus*), and how it compares to human development.

Here we provide a broad-scale comparison of social and physical cognitive development in humans and our two closest living relatives, chimpanzees and bonobos. For this comparison, we employed a battery of cognitive tasks validated in previous work (Carpenter, Nagell, & Tomasello, 1998; Herrmann, et al., 2007; Tomasello & Carpenter, 2005; Wobber, Wrangham, & Hare, 2010)(Table 6.1). In Study 1, we compared a cross-sectional sample of 48 human children to 49 same-age chimpanzees and bonobos (hereafter referred to by their genus, *Pan*). In Study 2,

we followed a group of 44 *Pan* infants and juveniles longitudinally over the course of three years to document their patterns of cognitive development in greater detail. We analyzed data in both studies in terms of the *speed* of cognitive development, using individuals' proficiency across tasks to index their general comprehension of the capacities being investigated, and in terms of the *patterns* of cognitive development, extracting information about the inter-relationships between skills using techniques from prior longitudinal studies of human psychological development (Carpenter, et al., 1998; Carpenter, Pennington, & Rogers, 2002).

STUDY 1

Our first study compared a cross-sectional sample of human children to same-age individuals of our closest living relatives, chimpanzees and bonobos (genus *Pan*). We studied an identical age range in both groups, 2 to 4 years, because *Pan* individuals begin to locomote self-sufficiently around 2 years (Doran, 1992; Pontzer & Wrangham, 2006), and can thus be tested in tasks requiring them to independently manipulate objects or move around in space. However, it is worth noting that *Pan* individuals in this age range are typically unweaned, with weaning occurring between 3.5 and 4.5 years of age (Goodall, 1986; Kano, 1992). We thus refer to them here as *Pan* “infants.” In contrast, humans in developed societies are most often weaned by the age of 2 years (Scott, Binns, Oddy, & Graham, 2006; Sellen, 2001), with several studies now supporting a weaning age between 2 and 3 years in natural fertility human societies as well (Kennedy, 2005; Sellen & Smay, 2001). We therefore refer to our human participants as “children.” Though we matched humans and *Pan* by absolute age in this study, in Study 2 we provided a matched sample for relative developmental stage by testing post-weaning *Pan* individuals of a wider age range.

Methods

Subjects

Chimpanzees were tested at the Tchimpounga Chimpanzee Sanctuary in the Republic of Congo and bonobos were tested at Lola ya Bonobo in the Democratic Republic of Congo. Apes at these sites are semi free-ranging but can voluntarily participate in cognitive testing in their dormitories (Wobber & Hare, 2011). In addition, we tested three chimpanzees and one bonobo living at the Wolfgang Koehler Primate Research Center in Leipzig, Germany. Our sample consisted of chimpanzees ($n = 26$, 15 males) and bonobos ($n = 23$, 12 males) ranging from 1.5 years to 4 years of age. For most non-human ape subjects we did not know ages to the month, and so here grouped them only by year of age: 2 years ($n = 15$); 3 years ($n = 20$), 4 years ($n = 14$).

Because the majority of non-human ape subjects were orphans with unknown birth dates, individuals' ages were estimated to the year using weight and dental emergence both upon arrival at the sanctuary and at the time of testing (see Appendix 6, Supplemental Methods). In Study 2, our longitudinal data controlled for any remaining uncertainty in subjects' precise age by examining improvements in performance over a known period of time. To ensure that being orphaned did not significantly impact apes' success in the cognitive tasks, we compared the performance of orphans to mother-reared individuals in the test sample (see Results). Ape subjects had never taken part in any previous cognitive study of this kind, though a few had taken part in previous tests of inhibitory control (Wobber, et al., 2010). Subjects were never food or water deprived for testing and all testing was voluntary.

Children ($n = 48$, 24 males) were tested in the Department of Comparative and Developmental Psychology at the Max Planck Institute for Evolutionary Anthropology (MPI-

EVA) in Leipzig, Germany. To match the ages of the *Pan* sample, we tested 2 year olds ($n = 16$, range: 19 to 23 months, mean: 22.2 months), 3 year olds ($n = 16$, range: 33 to 39 months, mean: 36.4 months), and 4 year olds ($n = 16$, range: 49 to 53 months, mean: 51.8 months). We targeted age groups that were 14 months apart, rather than 12 months, to provide maximal contrast between age groups. No child subject had previously participated in a similar study; therefore, the test situation and test items were novel to all species.

Design

Non-human apes were tested individually in familiar rooms of their dormitories. Children were tested individually in test rooms at the MPI-EVA. All subjects had a caregiver in the testing room or nearby, who did not participate in the test in any way.

Subjects participated in a battery of 14 cognitive tasks, in addition to 3 attentional/motivational control tasks, over the course of multiple test sessions (Table 6.1). Subjects received one testing session (lasting approximately 30 minutes) per day, with subjects receiving anywhere from 3 to 10 test sessions in total depending on their relative motivation to participate in multiple tasks on any given day (see Appendix 6, Supplemental Methods). Individuals always completed a given task in only one testing session, with breaks between sessions only occurring in between tasks. Two chimpanzees and one bonobo, not included in our sample sizes mentioned above, began but did not complete the test battery because they became unmotivated across repeated days of testing.

The order in which tasks were presented was consistent within-genus. Children received the tasks in a slightly different order from *Pan* infants, in line with previous work (Herrmann, et al., 2007) (Appendix 6, Table A6.1).

Table 6.1. The Comparative Developmental Cognitive Battery (CDCB). Tasks were divided into two domains, those assessing social cognition (reasoning about other individuals) and physical cognition (reasoning about objects). All tasks have previously been used with non-human apes and human infants. Tasks were chosen to represent a diverse subset of the basic cognitive skills utilized for more complex processes in the social and physical cognitive domains. Where trial number differed between children and non-human apes, the number of trials presented to apes is indicated in parentheses. Abbreviations for each task that are used in other figures are shown.

Domain	Task	Description	# of Trials
Social cognition	Intention-emulation (IE)	Achieve experimenter's goal, seeing only failed attempt	3
	Social obstacle (SO)	Look to experimenter's face after being teased	3
	Gaze-following around barriers (GFB)	Follow experimenter's gaze geometrically	6
	Social inhibition (SI)	Reach selectively during simulated feeding competition	12
	Gaze-following (GF)	Follow experimenter's gaze into space	10
	Social learning (SL)	Copy action demonstrated by experimenter	1
	Point production (PP)	Direct experimenter to a reward out of her view	4
	Goal understanding (GU)	Understand experimenter's goal from failed attempt	12
	Reputation (Rep)	Discriminate between a generous and a stingy experimenter	2 (4)
	Object permanence (OP)	Track invisibly displaced rewards	6
Physical cognition	Transposition (Tra)	Track visibly displaced reward locations	6
	Number (Num)	Discriminate relative quantities	6
	Tool use (TU)	Use tool to obtain out-of-reach reward	1
	Tool properties (TP)	Choose functional over non-functional tools	12

Procedure

The same experimenter presented the tests to all non-human apes and another experimenter presented the tests to all children. Previous analyses have shown that different experimenters can reliably administer these tasks (see supplemental material in (Herrmann, et al., 2007)). Moreover, this procedure ensured that even if there were any slight differences in the experimenters' behaviors, any within-genus age patterns were not a result of these differences since the same experimenter consistently conducted the study within each genus.

All tasks were videotaped. For 11 of the 14 cognitive tasks and 1 of the 3 attentional/motivational control tasks, results were coded live. Performance on the remaining tasks was scored from video by the first author. For these 5 tasks, as well as 2 of the live coding tasks where performance was not simply choice-based, coders blind to the hypotheses of the study scored videos from a random 20% of individuals in each genus. Inter-observer reliability was assessed using Cohen's kappa for tasks where performance was dichotomous (0/1) and a Pearson correlation for tasks where performance was continuous (e.g. duration in seconds), with values for these analyses and their relative significance levels shown below (Table 6.2)(Martin & Bateson, 1986). Reliability across all 5 tasks was high, with similar values across both children and *Pan* infants suggesting that any differences between genera were unlikely to be due to greater measurement error in one group.

Tasks

The 14 cognitive tasks used here were taken either directly from previous work (Herrmann, et al., 2007; Wobber, et al., 2010) or adapted from prior studies of human-reared infant chimpanzees (Tomasello & Carpenter, 2005) (Table 6.1). Tasks performed identically to previous work are noted below (Herrmann, et al., 2007; Wobber, et al., 2010). For the other tasks

we present short descriptions, with more detailed procedures outlined in Appendix 6, Supplemental Methods. Procedures were identical for children and *Pan* infants, except where mentioned below and in that 1) toys served as the reward for children rather than food and 2) in certain tasks, no mesh barriers separated the child from the experimenter.

Table 6.2. Analyses of inter-observer reliability. A coder blind to the hypotheses of the study coded a randomly selected 20% of trials in tasks where performance was unambiguous (5 cognitive and 2 attentional/motivational control tasks). Results are shown for each task according to genus, along with the sample included in the analysis (for tasks with multiple trials, each trial was entered as an independent value in the reliability analysis). Cohen's kappa tests were performed for tasks scored dichotomously (0/1). Pearson's correlations were performed for tasks scored with continuous measures, with these tasks denoted by asterisks.

Task	<i>Children</i>		<i>Pan infants</i>	
	Cohen's kappa/ Pearson correlation	p-value	Cohen's kappa/ Pearson correlation	p-value
Intention-emulation	K = 1.00	<0.001	K = 0.83	<0.001
Social obstacle	K = 1.00	<0.001	K = 0.67	0.001
Gaze-following around barriers	K = 0.94	<0.001	K = 0.82	<0.001
Gaze following	K = 0.95	<0.001	K = 0.77	<0.001
Social learning	K = 1.00	<0.001	K = 1.00	<0.001
Novel objects*	r = 0.83	0.003	r = 0.98	<0.001
Unsolvable task*	r = 0.90	<0.001	r = 0.96	<0.001

Social cognition

Intention-emulation (IE). This test served to measure whether subjects could infer an experimenter's goal, having never seen her complete the goal but seeing only her failed attempts to achieve it (Bellagamba & Tomasello, 1999; Tomasello & Carpenter, 2005). The experimenter (E1) attempted three times to put together two pieces of PVC pipe, but failed each time. E1 then handed the pieces of PVC pipe to the subject, with the dependent measure for this task the

number of trials where the subject successfully put together the two pieces of PVC pipe (Tomasello & Carpenter, 2005). Subjects received three trials of this task (one per day on three subsequent test days) and were rewarded for handing back the PVC pipes regardless of whether they succeeded in putting them together.

Social obstacle (SO). This task was designed to measure a subject's tendency to look to another individual's face as a cue to his or her intentions (Phillips, Baron-Cohen, & Rutter, 1992; Tomasello & Carpenter, 2005). E1 engaged the subject's attention with a toy and then teasingly pulled the toy away, looking straight ahead for 5 seconds. The dependent measure for this task was whether the subject looked to the experimenter's face in these 5 seconds (Tomasello & Carpenter, 2005). Three trials were presented in sequence, with a short break between trials to re-engage the subject in playing with the toy. Subjects were rewarded after each trial irrespective of their performance in that trial.

Gaze-following around barriers (GFB). This task served to measure whether individuals were able to follow an experimenter's gaze geometrically, requiring the subject to physically move around a barrier to follow this gaze rather than simply re-orienting his or her gaze direction (Moll & Tomasello, 2004; Tomasello & Carpenter, 2005). E1 called the subject's name and subsequently looked behind a barrier, alternating her gaze between the subject and this location while calling the subject's name for 30 seconds. The dependent measure for this task was whether the subject moved its body to look behind the barrier (Tomasello & Carpenter, 2005). Subjects were rewarded after each trial and given a short break prior to the next trial. Three trials per day were performed on two subsequent test days (resulting in 6 total trials). Two different barrier setups were utilized (one for the first day, and one for the second) to diminish potential habituation effects (see Appendix 6, Table A6.1).

Social inhibition (SI). This task was designed to measure individuals' abilities to inhibit their responses in a social situation where they requested rewards from selected human experimenters (Barth & Call, 2006; Herrmann, et al., 2007). Procedures were performed identically to the "social response inhibition" test in Experiment 2 of Wobber et al (2010).

Gaze-following (GF). This test, similar to gaze-following around barriers, measured individuals' abilities to track another's gaze. The experimenter sat across from the subject, called its name, and then looked upwards with her head and eyes for 10 seconds. The dependent measure was whether the subject also looked upwards (Butterworth & Jarrett, 1991; Herrmann, et al., 2007; Tomasello, Hare, & Fogleman, 2001). 10 trials were performed in sequence, with subjects rewarded and given a short break after each trial.

Social learning (SL). In this task, we observed whether subjects imitated the means demonstrated by an experimenter to achieve a goal (Call, Carpenter, & Tomasello, 2005; Gergely, Bekkering, & Kiraly, 2002; Herrmann, et al., 2007; Meltzoff, 1988; Tomasello & Carpenter, 2005). Procedures were performed identically to the "banana/balloon tube" social learning item in Herrmann et al (2007), with only this one trial performed (in contrast to the three social learning trials employed by Herrmann and colleagues).

Point production (PP). This task measured whether individuals would signal the location of a reward to an experimenter if that reward were out of her view, reflecting an understanding of the experimenter's attentional state and an ability to communicate gesturally (Herrmann, et al., 2007; Tomasello & Camaioni, 1997; Tomasello & Carpenter, 2005). Procedures were performed identically to the "attentional state" task of Herrmann et al (2007), with two trials of the "away" condition and two trials of the "towards" condition.

Goal understanding (GU). In this task, subjects needed to interpret an experimenter's intentions and goals in order to find a hidden reward in an object choice paradigm (Braeuer, Kaminski, Riedel, Call, & Tomasello, 2006; Herrmann, et al., 2007). Procedures were performed identically to the "intentions" task of Herrmann et al (2007), with three trials of the "trying" condition followed by three trials of the "reaching" condition, except that two sessions of 6 trials each were presented on two subsequent test days, for a total of 12 trials.

Reputation (Rep). This task measured whether subjects could track other individuals' behavior and base decisions on this information (Hamlin, Wynn, & Bloom, 2007; Melis, Hare, & Tomasello, 2006). Subjects witnessed a demonstration where one ("nice") experimenter attempted to give a reward to a neutral individual but was prevented from doing so by another ("mean") experimenter. Subjects were then presented with a choice between the "nice" and the "mean" experimenters, both of whom were holding a reward. The dependent measure for this task was whether subjects selectively requested a reward from the nice experimenter. Neither experimenter provided a reward upon the subject's request, to prevent learning from affecting decisions in subsequent trials. Two trials were performed for children whereas four were performed with *Pan* infants (as children became unmotivated in piloting when using a greater number of trials).

Physical cognition

Object permanence (OP). This task measured subjects' knowledge of object permanence with a Stage 6 invisible displacement task (Barth & Call, 2006; Herrmann, et al., 2007; Piaget, 1952). Procedures were performed identically to Herrmann et al (2007), except that here we used only two trials of three trial types (single, double adjacent, and double non-adjacent displacements), for a total of 6 trials.

Transposition (Tra). This task also measured individuals' abilities to track hidden rewards, in this case with the reward location being moved in full view of the subject (Barth & Call, 2006; Herrmann, et al., 2007; Sophian, 1984). Procedures were performed identically to Herrmann et al (2007), except that we used only two trials of three trial types (single, double unbaited, and double baited swaps), for a total of 6 trials.

Relative Number (Num). This task measured individuals' ability to discriminate between varying quantities of a reward, with individuals successful if they were able to choose the option providing the larger reward (Hanus & Call, 2007; Herrmann, et al., 2007; Tomonaga, 2008). Procedures were performed identically to Herrmann et al (2007), except that only 6 quantity comparison trials were presented, in the following order: 1:0, 6:3, 6:2, 3:2, 2:1, 4:1.

Tool use (TU). In this task, subjects needed to use a tool to obtain an out-of-reach reward (Herrmann, et al., 2007). Procedures were performed identically to Herrmann et al (2007).

Tool properties (TP). To test whether subjects understood the functional properties of tools, beyond simply being able to use tools, we presented them with an object choice task where they needed to choose between a functional and non-functional tool, each of which was associated with a reward (Hauser, 1997; Herrmann, et al., 2007; Herrmann, Wobber, & Call, 2008). Procedures were performed identically to Herrmann et al (2007), with 3 trials of the "side" condition and three trials of the "ripped" condition presented in sequence in each test session. Subjects received two test sessions of this task on subsequent test days, resulting in a total of 12 trials.

Attentional/motivational controls

Three control tasks were conducted to ensure that any species or age patterns reflected differences in subjects' cognitive abilities rather than differences in their motivation to complete the tasks.

Risk box. This task served to measure subjects' interest in novelty, or general willingness to take risks in an unfamiliar situation (Kagan & Snidman, 2004). This task was presented prior to all of the other tasks, making it the first interaction that subjects had with the experimenter and the general test environment. The experimenter presented the subject with a wooden box with a hole on one side, giving the subject 30 seconds to manipulate the box initially and then placing a reward inside the hole. The dependent measure for this task was whether the subject reached into the hole in the box to obtain the reward, with individuals given 30 seconds to do so. Only one trial was performed.

Unsolvable task. This task provided an index of how interested subjects were in obtaining a reward and how determined they were to independently solve a problem (Miklosi et al., 2003). The experimenter presented the subject with three trials of a task that was solvable, with a reward placed under an upside-down clear box that could be opened by lifting the box off of its lid. For the unsolvable trial, the experimenter placed a reward in the box but then fixed the box to its lid (unbeknownst to subjects), making it impossible to open but visually identical to the solvable situation. The dependent measure for this task was how long subjects would manipulate the box in attempting (unsuccessfully) to obtain the reward, with individuals given a maximum of 1 minute to do so.

Novel objects. This task measured subjects' reactivity to novel objects, quantifying their position on a shy-bold continuum and their general interest in objects that might pertain to the

test (Herrmann, et al., 2007; Kagan & Snidman, 2004). The experimenter sat behind the testing table and placed an object on the table. Two differing objects were used, each of which was presented first as a still object (for 30 seconds) and then as a moving object (for 30 seconds). The dependent measure for this task was the time (out of two minutes total) that subjects spent in close proximity to the table. The camera was positioned such that it captured a pre-specified area of a certain size (140 cm x 110 cm). Thus in coding, the experimenter could record how many seconds subjects spent in this area as a measure of their interest.

Analysis

We began our analyses by examining differences in the rate of cognitive development between children and *Pan* infants, and then examined patterns of performance in each group. Chimpanzees and bonobos were combined for the analyses because the sample size of each species in certain age groups was too small ($n < 4$) to compare individually to children (differences in performance between the two species emerging in adolescence and adulthood are discussed elsewhere, see (Wobber, Herrmann, Hare, Wrangham, & Tomasello, in preparation; Wobber, et al., 2010)).

Rate of cognitive development

To assess the rate of cognitive development between the ages of 2 to 4 years, we computed each subject's average performance for the social and physical domains, as well his or her average performance in the three control tasks. We then performed univariate General Linear Model (GLM) analyses separately for the social domain, the physical domain, and the control tasks with genus (*Homo* vs. *Pan*) and age group (2 years, 3 years, or 4 years) as factors. Post-hoc analyses were controlled for multiple comparisons using a Bonferroni correction.

Table 6.3. Passing criteria used in the emergence analysis, Study 1 and Study 2. To examine patterns of development, we created pass/fail criteria that signified the minimum level of performance in a given task that denoted comprehension of that task. These criteria were defined either from past work or based on statistical relationships. Justifications for the emergence criteria across tasks are provided in Appendix 6.

Domain	Task	Number of trials	Forced choice?	Emergence criterion
Social cognition	Intention emulation	3		Achieve experimenter's goal on 1 or more trials
	Social obstacle	3		Look to experimenter's face on 1 or more trials
	Gaze-follow barriers	6		Follow gaze around barrier on 1 or more trials
	Social inhibition	12	X	Reach to both correct experimenters on 7 or more trials (chance success 33%)
	Gaze-following	10		Follow gaze on 1 or more trials
	Social learning	1		Copy experimenter's action to obtain reward
	Point production	4		Direct experimenter to reward on 1 or more trials
	Goal understanding	12	X	Choose experimenter's targeted container on 9 or more trials (chance success 50%)
	Reputation	2 (4)		Choose nice experimenter on first participating trial
Physical cognition	Object permanence	6	X	Choose correctly once or more in each potential reward location (left, right, middle)
	Transposition	6	X	Choose correctly once or more in each potential reward location (left, right, middle)
	Number	6	X	Choose correctly once or more on each side (left, right)
	Tool use	1		Use tool to obtain reward
	Tool properties	12	X	Choose functional tool on 9 or more trials (chance success 50%)

Patterns of cognitive development

We used several measures to analyze patterns of cognitive development in the two genera. First, to determine the age at which individuals began to succeed in the differing

cognitive tasks, we created an emergence criterion for each task (Table 6.3). These emergence criteria were based on previous research where possible (Carpenter et al, 1998; Tomasello and Carpenter, 2005), and represented the minimum level of performance necessary to be considered comprehension for a given task.

The **age of emergence** (AOE) for each task was calculated as the age group where 50% or more of individuals successfully met the emergence criterion. We then calculated the **order of task emergence** based on the proportion of individuals meeting the emergence criterion in each task (Carpenter, et al., 2002). We ranked the tasks from those where the highest proportion was successful to those where the lowest proportion was successful, within each genus. We used Green's index of consistency (Green, 1956) to determine the degree to which these rank sequences represented stable patterns, both for the overall sequences and separately within the social and physical domains. Next, we investigated emergence relationships between pairs of tasks using the **ordering-theoretic method** (Bart & Airasian, 1974), which allowed us to determine which tasks were necessary precursors to one another and which were logically independent. Again, we performed these calculations using only pass/fail emergence data.

Finally, we performed two types of analysis using the continuous data set consisting of percentage correct in each task (rather than the pass/fail emergence measures). We first determined the **relative proficiency** across tasks in each genus. For this analysis, we ranked tasks *within each individual* based on that individual's relative performance in each (rather than performing these rankings on the group level). We then calculated differences in average within-individual task rank between children and *Pan* infants, using Mann-Whitney tests for this analysis since these data were not normally distributed. Note that within-individual task ranks could be biased by tasks where performance was only measured as pass/fail (e.g., success in the

social learning task would be represented as 100% correct). However, because trial numbers were identical for children and *Pan* infants (except in the Reputation task), any bias introduced by trial number would have been held constant in our comparisons of the two genera. Our second analysis in this area examined **inter-task correlations** in performance, to elucidate the degree to which individuals were consistent in their performance on the whole and to determine whether specific tasks were related in their levels of success.

RESULTS

Rate of cognitive development

A univariate GLM of average performance in the social domain revealed significant effects of genus and age group, as well as a significant interaction between genus and age (genus: $F(1,94) = 335.20$, $p < 0.001$; age group: $F(2,94) = 24.51$, $p < 0.001$; genus*age group: $F(2,94) = 19.82$, $p < 0.001$) (Figure 6.1a). Post-hoc analyses revealed a strong effect of age in humans ($F(2,47) = 33.53$, $p < 0.001$) but not in *Pan* ($p > 0.8$). Human 3- and 4-year-olds both outperformed human 2-year olds (p values < 0.001), with 4-year-olds also outperforming 3-year-olds ($p < 0.05$), whereas there were no differences among any age groups in *Pan*. Humans also outperformed *Pan* at every age (2 years, 3 years, and 4 years; all p -values < 0.01). In sum humans were already more skilled than both chimpanzees and bonobos at socio-cognitive tasks by the age of 2 years, and continued improving rapidly until 4 years while *Pan* infants did not significantly improve in their performance in this age range.

In the physical domain there were also effects of genus and age group on performance, as well as an interaction between the two variables (univariate GLM; genus: $F(1,96) = 62.27$, $p < 0.001$; age group: $F(2,96) = 23.36$, $p < 0.001$; genus*age group: $F(2,96) = 7.52$, $p = 0.001$) (Figure 6.1b). Post-hoc analyses again revealed a significant effect of age in humans ($F(2,47) =$

30.50, $p < 0.001$) but not in *Pan* ($p > 0.1$). Furthermore, 4-year old humans outperformed human 2- and 3-year olds (p values < 0.001), with no differences between age groups in *Pan*. However in contrast to the social domain, humans did not outperform *Pan* in the physical domain at 2 years, becoming detectably more skilled at 3 years ($p < 0.01$) and distinctly more skilled at 4 years ($p < 0.001$). Thus, humans were comparable to chimpanzees and bonobos in their physical cognition proficiency at 2 years, but they quickly began to outperform the other apes in the next one to two years. There were no sex differences in performance in either domain among either humans or *Pan* (univariate GLM analyses with sex as a factor, p values > 0.2).

In the attentional/motivational controls, there were no main effects of genus or age group, nor a significant interaction between the two factors ($p > 0.05$, with a trend-level interaction between genus and age group given the marginal improvement in *Pan* infants) (Figure 6.1c). These results suggest that increases in humans' performance on the cognitive tasks were not tied to increases in attention and motivation, since these measures remained consistent across age in the children and *Pan* infants tested with no significant difference between the genera at any age. These results indicate that differences in performance did not simply represent attentional or motivational biases.

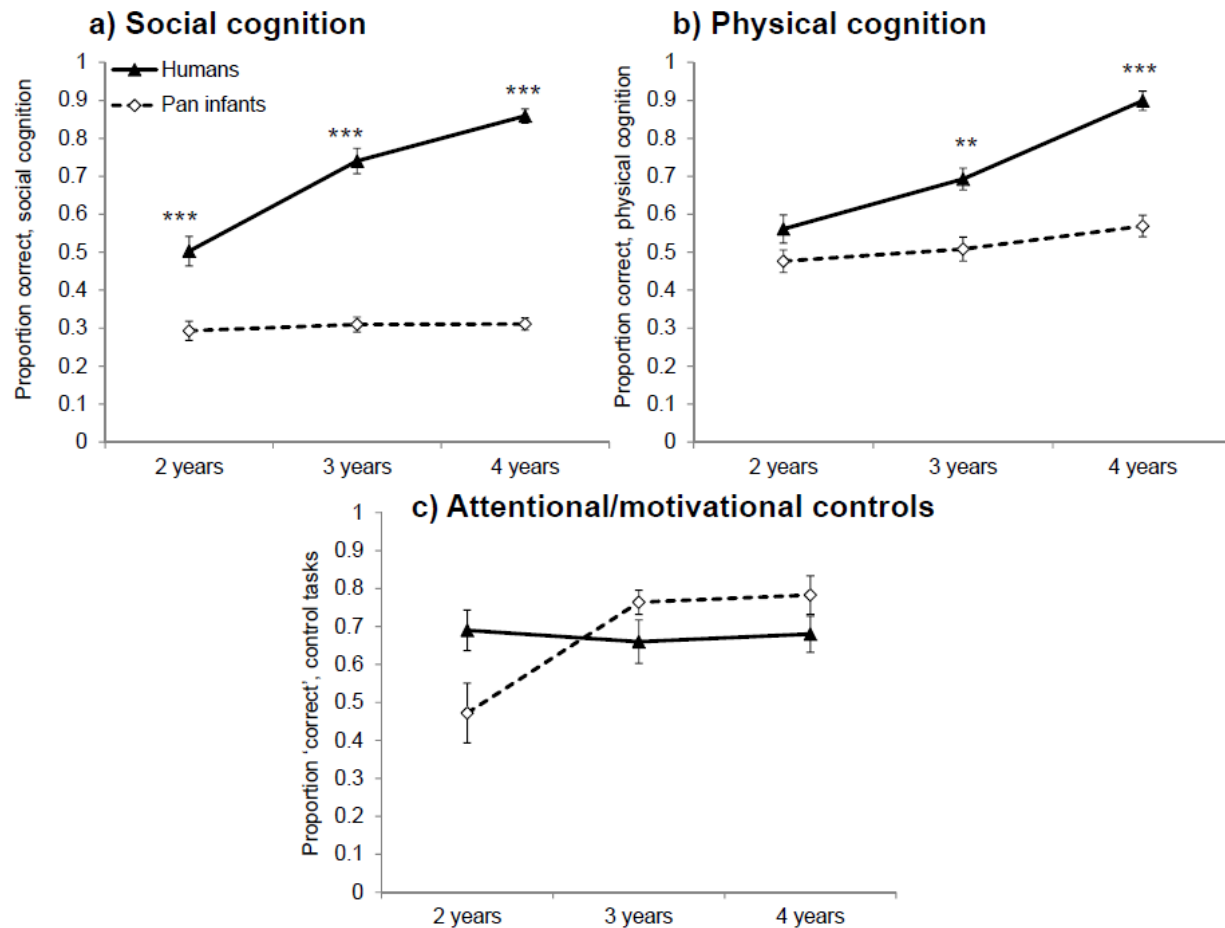


Figure 6.1. Performance in the cross-sectional comparison of human children and *Pan* infants, Study 1. The y-axis denotes mean percentage correct in a) 9 social cognition, b) 5 physical cognition and c) 3 attentional/motivational control tasks, and the x-axis denotes the three age groups (children: 2 years, $n = 16$, 3 years, $n = 16$, 4 years, $n = 16$; *Pan* infants: 2 years, $n = 15$, 3 years, $n = 20$, 4 years, $n = 14$). Bars denote standard error. Significant genus differences (adjusted for multiple comparisons using a Bonferroni correction) are denoted as follows: * $p \leq 0.05$ ** $p \leq 0.01$, *** $p \leq 0.001$. Children's performance in both domains increased from 2 to 4 years, while performance in *Pan* did not. 2 year old humans were only more skilled than the 2 year old *Pan* infants in social cognition but not in physical cognition. At 3 and 4 years, humans were progressively more skilled in both social and physical cognition. There were no significant genus differences in performance on the control tasks either overall or at any age group, nor did performance on these tasks change significantly with age in either genus.

Controls for rearing history

Because the majority of our nonhuman ape subjects were orphans, we wanted to ensure that this factor did not account for the observed differences in performance between *Pan* infants

and human children. A comparison of mother-reared apes living in the African ape sanctuaries ($n = 9$) and mother-reared apes living in a zoo population ($n = 4$) revealed no significant differences between these groups in either social cognition or physical cognition performance (univariate GLM analyses: physical cognition, $p > 0.4$, social cognition, $p = 0.06$ with a trend for sanctuary individuals to perform slightly better than zoo individuals). We therefore combined these two mother-reared groups to compare to the orphans. This sample differs slightly from the mother-reared sample in our previous paper (Wobber & Hare, 2011) since only individuals between 2 to 4 years of age were examined here.

A univariate GLM of performance in the social domain comparing the 13 mother-reared individuals to 13 age- and sex-matched orphans revealed no significant effect of mother-rearing ($p > 0.1$), suggesting that being orphaned at 2-3 years of age does not significantly affect socio-cognitive abilities in sanctuary individuals. Similarly, there were no differences between mother-reared individuals and orphans in performance on the attentional/motivational control tasks (univariate GLM, $p > 0.1$). However, mother-reared individuals did significantly outperform orphans in the physical domain ($F(1,25) = 7.30$, $p = 0.01$). Further investigation revealed no significant differences in performance on any physical cognition task between orphans and mother-reared individuals after correction for multiple comparisons. The only physical cognition task where mother-reared individuals outperformed orphans prior to this correction was object permanence ($t(24) = 2.05$, uncorrected $p = 0.05$), with a trend towards mother-reared individuals performing more skillfully in tool use ($\chi^2(1) = 3.47$, uncorrected $p = 0.06$, $n = 23$) (Figure 6.2). These results thus indicate that orphans perform as well as mother-reared infants on the vast majority of cognitive and attentional/motivational tasks, allowing us to group them together with the mother-reared individuals for our analyses. The results also conform to previous findings that

adult sanctuary orphans perform just as well or better than mother-reared apes in a zoo population (Hanus & Call, 2008; Vlammings, Hare, & Call, 2010; Wobber & Hare, 2011), indicating that they represent a viable population for non-human primate research (though see (J. L. Russell, Lyn, Schaeffer, & Hopkins, 2011) for a contradictory perspective).

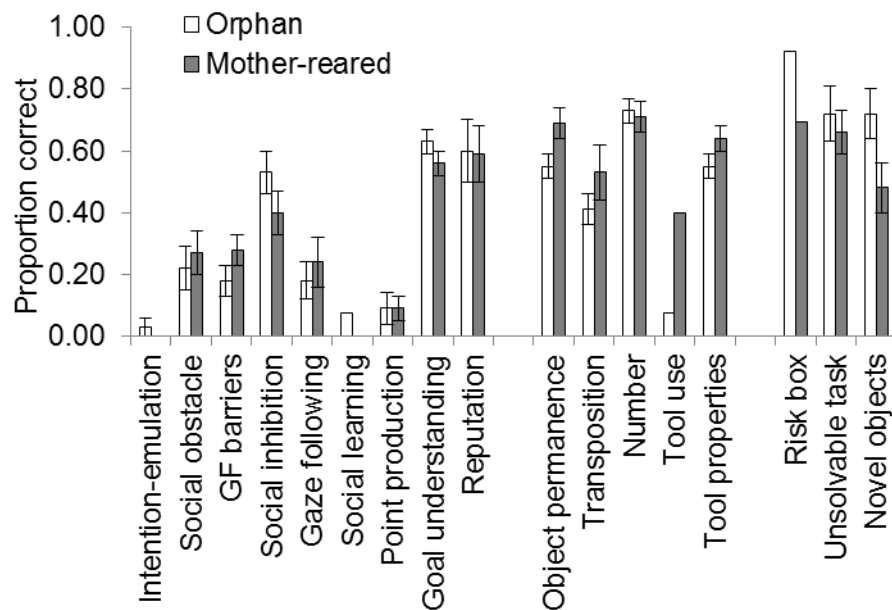


Figure 6.2. Performance across cognitive and attentional/motivational tasks by mother-reared and orphan *Pan* infants, Study 1. Average proportion correct is shown for tasks where the dependent measure was continuous, with bars to represent standard error. For tasks where a success/failure measure was used, proportion of individuals correct is shown (and thus there is no standard error for these tasks). Social tasks are on the left, followed by physical tasks, and then the attentional/motivational controls. Comparisons of performance across each task revealed that mother-reared individuals performed comparably to orphans in 13 of the 14 tasks. They performed significantly better than orphans in only the object permanence task (denoted in the figure with an asterisk), with a trend towards more skillful performance in the tool use task. Thus on the whole sanctuary orphans do not appear to suffer cognitive or motivational deficits relative to mother-reared individuals.

Patterns of development

Age of Emergence

To investigate patterns of cognitive development, we began by examining the age of emergence across the cognitive tasks in humans and *Pan*. Among children, the majority of tasks (9 of 14) had emerged by 2 years of age (Table 6.4). Children also met the emergence criterion for all 14 tasks by the age of 4 years. In contrast, *Pan* infants had met the emergence criterion for only 4 of the 14 tasks by 2 years of age, and did so in only 8 tasks by the oldest age group tested (Table 6.4). Notably, the physical cognition tasks where *Pan* infants had the most difficulty were also those most difficult for children – tool use and tool properties. Moreover, 4 of the earliest-emerging social tasks in children were also early to emerge in *Pan* infants (social obstacle, reputation, and the two gaze-following tasks). However, *Pan* infants struggled with certain social tasks (intention-emulation and goal understanding) where children succeeded even at the earliest age tested. In Study 2, we were able to determine whether these represented consistent differences in skill between children and *Pan* individuals in these tasks or whether proficiency in these areas simply emerged later on in *Pan* development.

Order of Emergence

In addition to the later age of emergence for several tasks in *Pan*, there were also significant differences between children and *Pan* infants in the overall orders of task emergence (Figure 6.3). These differences were present within both the social domain (Appendix 6, Figure A6.1) and, to a lesser extent, the physical domain (Appendix 6, Figure A6.2). The tasks where children performed most skillfully were overwhelmingly in the social domain, with only 1 of the 5 physical cognition tasks present among children's top 50% of skills (Figure 6.3). In contrast, 3 of the 5 physical cognition skills were in the top 50% of skills for *Pan* infants. Within the social domain, both children and *Pan* infants were highly skilled in the gaze-following tasks as well as the social obstacle task, but children's success in comprehending others' goals (measured by the

intention-emulation and goal understanding tasks) was not matched by same-age *Pan*. Again, these results supported the importance of goal understanding in human relative to *Pan* development, in line with recent findings with human infants and adult chimpanzees (Myowa-Yamakoshi, Scola, & Hirata, 2012).

Table 6.4. Average age of emergence (in years) for each task in children and young chimpanzees/bonobos (genus *Pan*), using the cross-sectional data from *Pan* infants, Study 1, and the longitudinal sample of *Pan* infants/juveniles, Study 2. Ages are rounded to the closest year (see Methods). Tasks where the average passing criterion was not reached by the oldest age group in the cross-sectional sample (4 years) are indicated as emerging “>4 years”; tasks that did not emerge by the oldest age in the longitudinal *Pan* sample (8 years) are indicated as emerging “never.” Tasks where *Pan* individuals were least skillful relative to children are highlighted with grey bars – namely, those pertaining to cooperative motivations and understanding others’ goals. Ages of emergence were on average later for the longitudinal *Pan* data relative to the cross-sectional *Pan* data owing to the smaller sample of 2-year-old individuals in the former than the latter.

Domain	Task	Age of emergence (in years)		
		Humans	<i>Pan</i> cross-sectional	<i>Pan</i> longitudinal
Social cognition	Intention emulation	2	>4	Never
	Social obstacle	2	2	3
	Gaze-follow barriers	2	2	3
	Social inhibition	3	3	3
	Gaze-following	2	2	3
	Social learning	3	>4	Never
	Point production	3	>4	4
	Goal understanding	2	>4	7
	Reputation	2	2	2
Physical cognition	Object permanence	2	4	3
	Transposition	2	4	4
	Number	2	3	3
	Tool use	4	>4	7
	Tool properties	3	>4	7

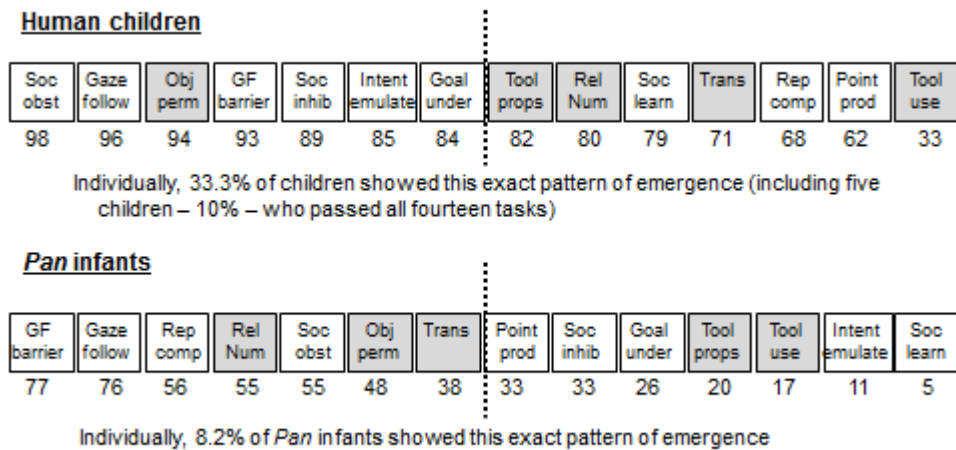


Figure 6.3. Patterns of emergence across cognitive tasks in human children and *Pan* infants, Study 1. Each task is represented by its own box; physical cognition tasks are denoted by gray boxes. The percentage of individuals meeting the passing criterion for each task is shown under its respective box. The dotted line denotes the halfway point among the 14 tasks, with skills to the left of this line those where individuals of each genus could be considered most successful. Individual support for patterns of emergence within each genus is also reported.

Beyond these differences in the order of task emergence, individual patterns of emergence were more consistent in children than they were in *Pan* infants. Overall, 33.3% of children supported the 14-task pattern of emergence, while only 8.2% of *Pan* infants did so (though both proportions were significantly greater than the proportion of individuals expected to match these exact patterns by chance, binomial tests, p values <0.001). No *Pan* individual supported the 14-task pattern found within children, suggesting stronger support for their respective 14-task pattern (support for each pattern by each sample group is summarized in Table 6.5). Similarly, children showed greater levels of individual support for their patterns of social and physical emergence than *Pan* infants did for their respective domain-level patterns (Table 6.5). *Pan* infants showed even less individual support for the patterns found within children,

suggesting significant differences between the genera both in overall average sequence and in levels of inter-individual plasticity (Table 6.5).

Table 6.5. Levels of individual support for patterns of task emergence in children and young chimpanzees/bonobos (genus *Pan*), using the cross-sectional data from *Pan* infants, Study 1, and the longitudinal sample of *Pan* infants/juveniles, Study 2. Within each genus, the group-level order of task emergence was computed for the 14 tasks overall, as well as separately for the 9-task social domain and the 5-task physical domain. Orders of emergence for *Pan* were calculated from both the cross-sectional and longitudinal data, thus they are represented here as distinct entries (see Figure 6.9). Here, the percentage of individuals that matched the exact sequences of emergence within each group is shown. This percentage is shown both for the group's own sequence (highlighted in gray), as well as for the sequences determined for the other groups. Note that percentage support was on the whole higher among children, but that *Pan* individuals better supported their respective emergence patterns than they supported any others. Note also that 10.4% of children passed all 14 cognitive tasks, leading them to support any pattern investigated.

Group	Emergence sequence	Domain		
		Overall	Social	Physical
Humans	Human pattern	33.3	45.8	81.3
	<i>Pan</i> cross-sectional pattern	12.5	33.3	68.8
	<i>Pan</i> longitudinal pattern	12.5	33.3	45.8
<i>Pan</i> cross-sectional	Human pattern	0.0	6.1	30.6
	<i>Pan</i> cross-sectional pattern	8.2	24.5	51.0
	<i>Pan</i> longitudinal pattern	2.0	10.2	42.9
<i>Pan</i> longitudinal	Human pattern	0.0	2.3	2.3
	<i>Pan</i> cross-sectional pattern	0.0	2.3	22.7
	<i>Pan</i> longitudinal pattern	0.0	4.5	31.8

When calculating a measure of scalability that takes chance scaling into account, Green's index of consistency (*I*) (Green, 1956), the only reliable scale was in the 5-task pattern of physical cognition task emergence in children, with the overall 14-task sequences and the 9-task social sequences not meeting the criterion for reliable scalability in either children or *Pan* infants (children: overall, $I = 0.17$, social cognition, $I = 0.04$, physical cognition, $I = 0.50$; *Pan* juveniles:

overall, $I = -0.03$, social cognition, $I = -0.12$, physical cognition, $I = 0.17$, where values of 0.50 or above for I indicate a reasonable degree of scaling consistency (Green, 1956)).

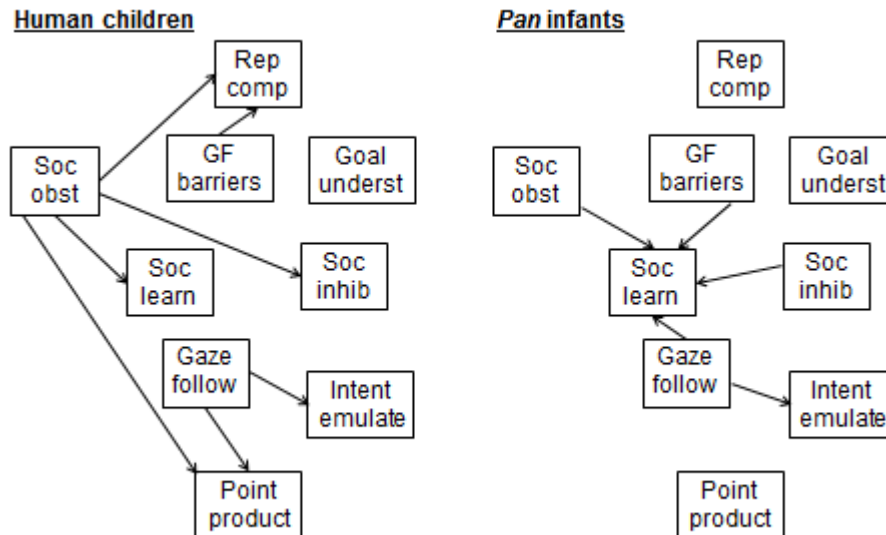


Figure 6.4. Relationships between tasks in the social domain among children and *Pan* infants using the ordering-theoretic method, Study 1. We used a 0% tolerance level to determine logical inter-relationships between pairs of tasks, separating by domain. Arrows denote tasks where one was a logical prerequisite to another, while tasks that are not connected by arrows were logically independent from another.

We also performed an analysis of task emergence using the ordering-theoretic method (Bart & Airasian, 1974) to provide insight into the inter-relationships between pairs of tasks beyond our analysis of the sequence-level patterns. We set a 0% tolerance level to establish these relationships, as performed in past work (Bart & Airasian, 1974; Carpenter, et al., 2002). In children, there were a number of task pairs where one skill was found to be a necessary prerequisite to the other, while there were many fewer task pairs in *Pan* infants showing these logical prerequisite relationships. Within the social domain (Figure 6.4), children's success in social obstacle preceded success in the greatest number of other tasks (four), while their success

in gaze-following preceded success in two other skills. Similarly, in *Pan* infants, success in gaze-following preceded success in two other tasks (Figure 6.4), but success in social obstacle preceded only one other skill. The task preceded by the most other skills (four) for *Pan* infants was social learning. In the physical domain (Figure 6.5), children's success in object permanence was the only task to precede success in any others, preceding three of the other physical cognition skills. There were no logical prerequisites in the physical domain among *Pan* infants, with no relationships consistent enough to meet the 0% tolerance criterion (Figure 6.5). These findings provide further support for the notion that individual patterns of development are more plastic in *Pan* development than in human development, with fewer logical prerequisites and more logical equivalence between tasks. The results also suggest that significant changes in skill inter-relationships during ontogeny are present specifically within the socio-cognitive domain. In particular, Figure 6.6 shows that a number of socio-cognitive tasks logically preceded success in physical cognition tasks among children, while this was not the case in *Pan* infants.

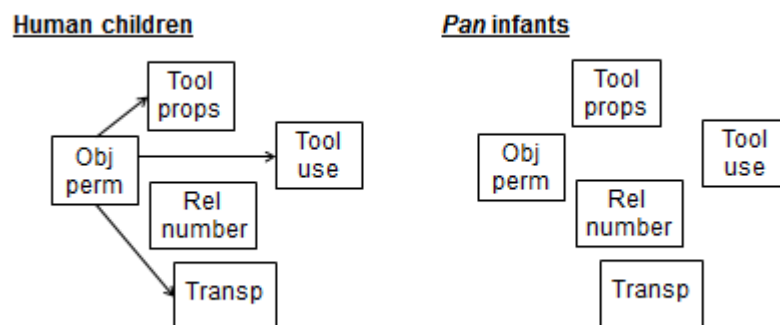
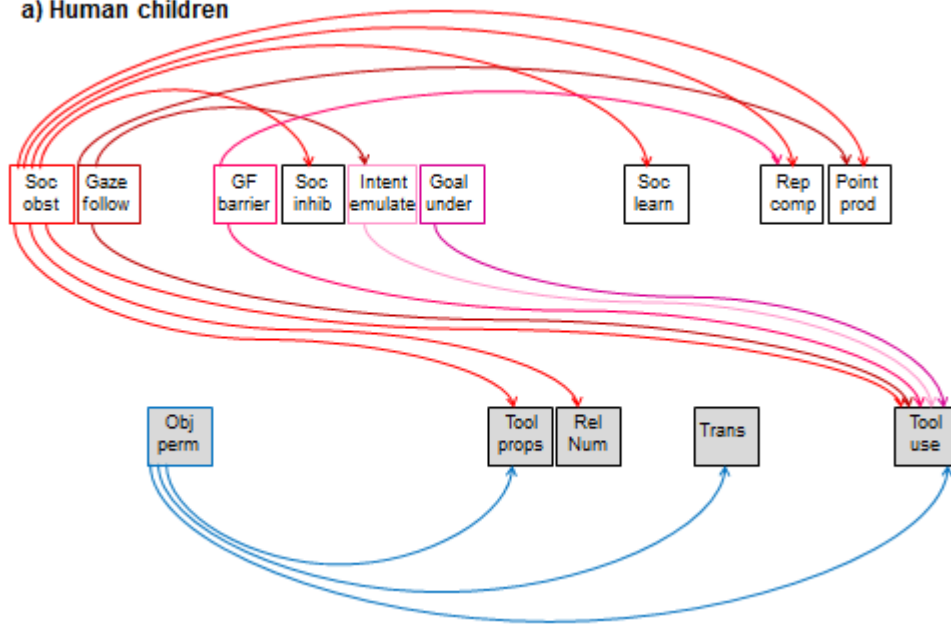


Figure 6.5. Relationships between tasks in the physical domain among children and *Pan* infants using the ordering-theoretic method, Study 1. We used a 0% tolerance level to determine logical inter-relationships between pairs of tasks, separating by domain. Arrows denote tasks where one was a logical prerequisite to another, while tasks that are not connected by arrows were logically independent from another.

a) Human children



b) *Pan* infants

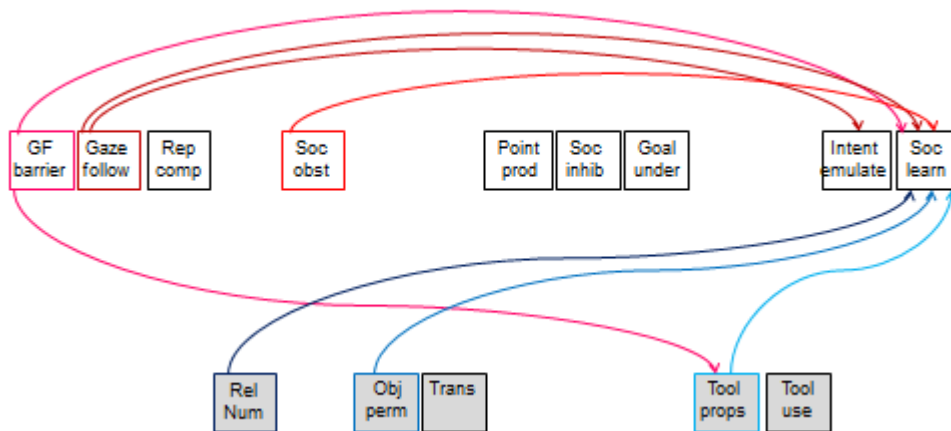


Figure 6.6. Relationships between tasks across both domains among a) children and b) *Pan* infants using the ordering-theoretic method, Study 1. We used a 0% tolerance level to determine logical inter-relationships between pairs of tasks, separating by domain. Physical cognition tasks are denoted by boxes with grey shading. Social tasks that were logical prerequisites to others are colored red/pink, while physical tasks that were logical prerequisites are colored blue (with arrows color-coded in line with each prerequisite task). Tasks that were not prerequisites to any other are outlined in black.

Interestingly, there were two commonalities between children and *Pan* infants in the logical prerequisite relationships. For both genera, success in social obstacle preceded success in social learning, and success in gaze-following preceded success in intention-emulation (achieving another's failed goal). Mainly, these results stemmed from *Pan* infants performing quite poorly on the latter two tasks (social learning and intention emulation). However, it is a compelling suggestion that *Pan* infants, like human children, must begin to seek information about others' attention and intention before inferring their goals (Carpenter, et al., 1998; Tomasello & Carpenter, 2005).

Relative Proficiency

We next used the continuous data set (consisting of proportion correct rather than pass/fail measures) to examine within-individual ranks of task performance. There were a number of differences in within-individual task ranks between children and *Pan* infants in the social domain (Figure 6.7). Children showed significantly lower (better) ranks than *Pan* infants in 4 of the 9 social cognition tasks: social obstacle (Mann-Whitney U, $Z = -6.08$, $p < 0.001$), social learning ($Z = -5.81$, $p < 0.001$), intention emulation ($Z = -5.68$, $p < 0.001$), and gaze-following around barriers ($Z = -2.58$, $p = 0.01$). Meanwhile, *Pan* infants showed significantly lower (better) within-individual ranks for reputation ($Z = 4.79$, $p < 0.001$) and gaze-following ($Z = 2.07$, $p = 0.04$ – though the difference in mean rank for gaze-following was minimal, *Pan* infants' performance was highly consistent and so this led to an overall group-level difference). In contrast, there were no differences in task rank between children and *Pan* infants within the physical domain. These findings indicate that when controlling for differences in absolute skill level (by comparing individuals to their own average task performance), the rankings of the physical tasks were similar between children and *Pan* infants (e.g. the same tasks proved most

difficult for both groups). These results further support the notion that the development of socio-cognitive skills has changed more dramatically between humans and our closest living relatives than has the development of physical cognition skills. It remains open to question whether socio-cognitive development is more plastic across species on the whole, or whether this represents a unique case in human evolution.

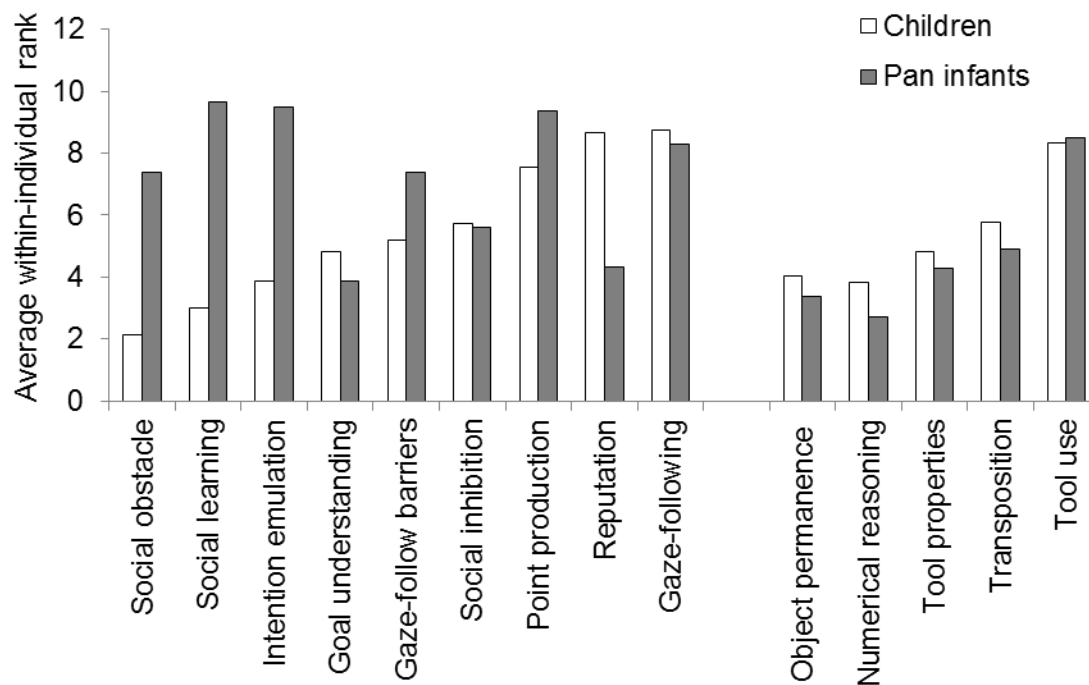


Figure 6.7. Average within-individual task ranks in children and *Pan* infants, Study 1.

Social tasks are on the left of the graph, with physical tasks on the right. Importantly, lower ranks represent better performance (since a subject's best task would be his or her 1st rank task, while that subject's worst task would be rank 14). There were a number of significant differences in task rank between children and *Pan* juveniles in the social domain, noted on the graph, with the significance values denoted as follows: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. Meanwhile, within-individual ranks in the physical domain did not differ between children and *Pan* infants for any task.

Inter-task correlations

Finally, to examine the degree to which individual performance was correlated across tasks, we calculated pairwise task correlations using the continuous performance data set.

Pearson correlations between each of the 14 tasks in each genus revealed 38 significant ($p < 0.05$) relationships in children, but only 5 significant relationships in *Pan* infants (Table 6.6). Among these 5 significant relationships in *Pan*, 1 was between two social tasks, 2 were between two physical tasks, and 2 were cross-domain. Meanwhile, in children, there were 15 significant correlations solely within the social domain, 7 correlations solely within the physical domain, and 16 cross-domain correlations. These results revealed 1) greater within-individual consistency in task performance among children, in addition to the heightened inter-individual consistency demonstrated by the order of emergence analysis, and 2) greater inter-correlation of the social domain with other skills in children relative to *Pan* infants, in line with past work showing a distinct social cognition “factor” in children but not chimpanzees (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2010).

CONCLUSIONS – STUDY 1

The results of our first study provide support for the possibility that humans exhibit **Accelerated Ontogeny** in our psychological development relative to that of other species. In particular, this accelerated ontogeny might result from early-emerging socio-cognitive skills in human children, with our findings demonstrating particular enhancements in children’s capacities for goal understanding and cooperative motivations. Our results also demonstrate greater inter-individual consistency in patterns of skill emergence among children relative to *Pan* infants, contradicting the assumption that patterns of development are more plastic within our species.

Table 6.6. Correlations in performance across cognitive tasks in a) children and b) *Pan* infants, Study 1. Social cognition tasks are divided from the physical cognition tasks by lines, with the social tasks in the upper left quadrant. Correlations between tasks solely within the social domain are shaded light gray, where tasks involving at least one task from the physical domain are shaded in dark gray. Tasks where the correlation did not reach significance are labeled “NS.” Pearson correlation values are shown, with their significance denoted as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Task abbreviations are as listed in Table 6.1 and in the Methods section.

a) Children

Task:	GF	GFB	SI	IE	GU	SL	Rep	PP	OP	TP	Num	Tra	TU
SO	NS	NS	0.548***	NS	NS	0.362*	NS	NS	NS	NS	NS	0.355*	NS
GF		NS	NS	0.369*	NS	NS	NS	0.374*	NS	NS	NS	NS	NS
GFB			0.374*	0.628***	NS	0.377*	NS	NS	0.295*	NS	NS	0.503***	0.368*
SI				0.444**	0.440**	0.673***	NS	NS	0.477**	0.503**	0.424**	0.577***	NS
IE					0.354*	0.353*	0.350*	0.390**	NS	NS	0.448**	0.522***	0.405**
GU						NS	NS	NS	NS	NS	NS	NS	NS
SL							NS	0.333*	0.363*	0.690***	NS	0.365*	NS
Rep								NS	NS	NS	NS	0.349*	NS
PP									NS	NS	NS	0.419**	NS
OP										0.616***	NS	0.448**	0.297*
TP											NS	0.467**	NS
Num												0.447**	0.383**
Tra													0.303*

b) *Pan* infants

Task:	GF	GFB	SI	IE	GU	SL	Rep	PP	OP	TP	Num	Tra	TU
SO	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
GF		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
GFB			NS	NS	NS	NS	NS	NS	NS	0.391**	NS	NS	NS
SI				NS	NS	0.328*	NS	NS	NS	NS	NS	NS	NS
IE					NS	NS	NS	NS	NS	NS	NS	NS	NS
GU						NS	NS	NS	NS	NS	NS	NS	NS
SL							NS	NS	NS	0.355*	NS	NS	NS
Rep								NS	NS	NS	NS	NS	NS
PP									NS	NS	NS	NS	NS
OP										0.490**	NS	0.377**	NS
TP											NS	NS	NS
Num												NS	NS
Tra													NS

STUDY 2

One possibility in interpreting our results from Study 1 was that the slower rate of cognitive development among *Pan* infants was simply due to having matched human and *Pan* individuals based on absolute rather than relative age. As discussed above, *Pan* individuals normally do not wean until 4 years of age, which is later than most estimates of weaning age in human populations (Kennedy, 2005; Sellen & Smay, 2001). Importantly, after they are weaned, young *Pan* individuals are responsible for 100% of their foraging intake, with no active provisioning on behalf of their mothers in stark contrast to typical patterns within our own species (Goodall, 1986; Kaplan, et al., 2000; Kramer & Ellison, 2010). This in turn suggests that *Pan* individuals may undergo a period of rapid cognitive development upon needing to forage for themselves and to independently navigate their social group at 4 years of age.

Thus in the present study, we examined a larger age range of *Pan* individuals spanning 2 to 6 years of age and followed them longitudinally for three subsequent years of testing. This allowed us to track individuals through the period of infancy and juvenility, with our oldest age group (8 years) mapping approximately onto the onset of the chimpanzee adolescent period (with menarche in females and spermatogenesis in males occurring around this time in captive chimpanzees (Coe, Connolly, Kraemer, & Levine, 1979; Marson, Meuris, Cooper, & Jouannet, 1991)). In this study we were able to discriminate between two alternatives: first, that *Pan* juveniles show a period of rapid cognitive maturation after weaning, similar to that found in human children in Study 1, or instead, that *Pan* juveniles continue their modest rate of cognitive development throughout ontogeny.

METHODS

Subjects

Again, chimpanzees were tested at the Tchimpounga Chimpanzee Sanctuary in the Republic of Congo and bonobos were tested at Lola ya Bonobo in the Democratic Republic of Congo. Our sample consisted of 30 chimpanzees (16 males) and 14 bonobos (8 males) that ranged from 2 to 6 years of age in the initial data collection season (2 years: $n = 9$; 3 years: $n = 11$; 4 years: $n = 7$; 5 years: $n = 6$; 6 years: $n = 11$). We followed these individuals for three subsequent years of testing (2008, 2009, and 2010), enabling us to examine development occurring between 2 to 8 years of age (total sizes for each age group across 3 years of testing: 2 years: $n = 9$; 3 years: $n = 20$; 4 years: $n = 27$; 5 years: $n = 24$; 6 years: $n = 24$; 7 years: $n = 17$; 8 years: $n = 11$). To ensure that changes in performance between years were not a result of increased familiarity with the tasks, we also tested a sample of 6 adults (3 chimpanzees, 3 bonobos, mean age = 13.2 years in 2008) in all three years. As in Study 1, owing to small sample size in particular age groups, we combined chimpanzees and bonobos for our analysis (differences between the two species are discussed elsewhere (Wobber, et al., in preparation; Wobber, et al., 2010)).

Design

Subjects participated in the same battery of cognitive tasks described in Study 1. Data was collected from chimpanzees in May/June 2008, June 2009, and June/July 2010. Data was collected from bonobos in July/August 2008, May/June 2009, and July/August 2010. 52 subjects began the longitudinal testing but 8 individuals did not complete it because they were re-introduced into the wild ($n = 5$) or died ($n = 3$).

Procedure

Procedures were identical to those described for Study 1, except that adult subjects did not participate two social cognition tasks (intention emulation and social learning) that involved breakable objects being passed into the test room. Their average social cognition performance thus represents the average of the remaining 7 social cognition tasks. Task abbreviations are the same as those used in Study 1.

Analysis

Again, we began by quantifying the general rate of improvement across tasks, and then investigated patterns of performance.

Rate of cognitive development

To examine performance across the multiple years of testing, we performed repeated measures ANOVAs with test year (2008, 2009, 2010) as a factor. We performed separate ANOVAs for the social domain, the physical domain, and the attentional/motivational controls. We performed these ANOVAs for our infant and juvenile subjects, as well as for the adult control group.

To examine improvement across tasks in more detail, we calculated difference scores between each individual's performance in the last year of testing (2010) and his or her performance in the first year of testing (2008) for each task. This analysis tracked the degree to which apes' performance changed over the course of two years of development, and allowed us to determine the areas where individuals showed the greatest improvement.

Patterns of cognitive development

We again determined an **age of emergence** for each task, using the same individual emergence criteria as Study 1 (Table 6.3). The group-level age of emergence was defined as the

age where 50% of individuals had met the emergence criterion either at or prior to that age (given that longitudinal data were available). Ages of emergence calculated from the longitudinal data were likely to be older than those calculated from the cross-sectional data due to the smaller number of individuals in the youngest age categories (several 2-year-olds began the longitudinal battery but did not finish it, and so are excluded from the analyses for Study 2). In addition to looking at ages of emergence, we examined the **order of task emergence** to determine the degree to which this matched the patterns found cross-sectionally in Study 1. We determined the order of task emergence by ranking tasks according to the proportion of individuals that were ever successful in that task over all three years of testing. We also assessed levels of individual support for these patterns, doing so for the overall sequence and separately for the social and physical domains. Finally, we again determined **inter-task correlations**, to validate our findings from Study 1 that *Pan* juveniles showed fewer inter-relationships in success across tasks.

RESULTS

Rate of cognitive development

A repeated measures ANOVA of *Pan* infant and juveniles' performance in the social domain revealed a significant effect of test year ($F(2,40) = 6.76, p < 0.01$). Post-hoc tests demonstrated that *Pan* infants and juveniles performed better in 2009 than they did in 2008 (Bonferroni-corrected $p = 0.012$) and better in 2010 than they did in 2008 (Bonferroni-corrected $p = 0.006$). A similar effect of year was present in the repeated measures ANOVA for the physical domain ($F(2,41) = 6.71, p < 0.01$), with the only significant improvement in performance between 2008 and 2010 (Bonferroni-corrected $p = 0.003$). Finally, *Pan* infants and juveniles also improved over the course of three years in their performance in the attentional/motivational controls (repeated measures ANOVA, $F(2,34) = 8.37, p = 0.001$), performing "better" (being

more attentive and motivated) in 2010 than 2008 (Bonferroni-corrected $p = 0.006$) and in 2010 than 2009 (Bonferroni-corrected $p = 0.045$). Thus across the social, physical, and attentional domains *Pan* infants and juveniles showed significant improvements in their performance with age from the initial to the final year of testing (Figure 6.8).

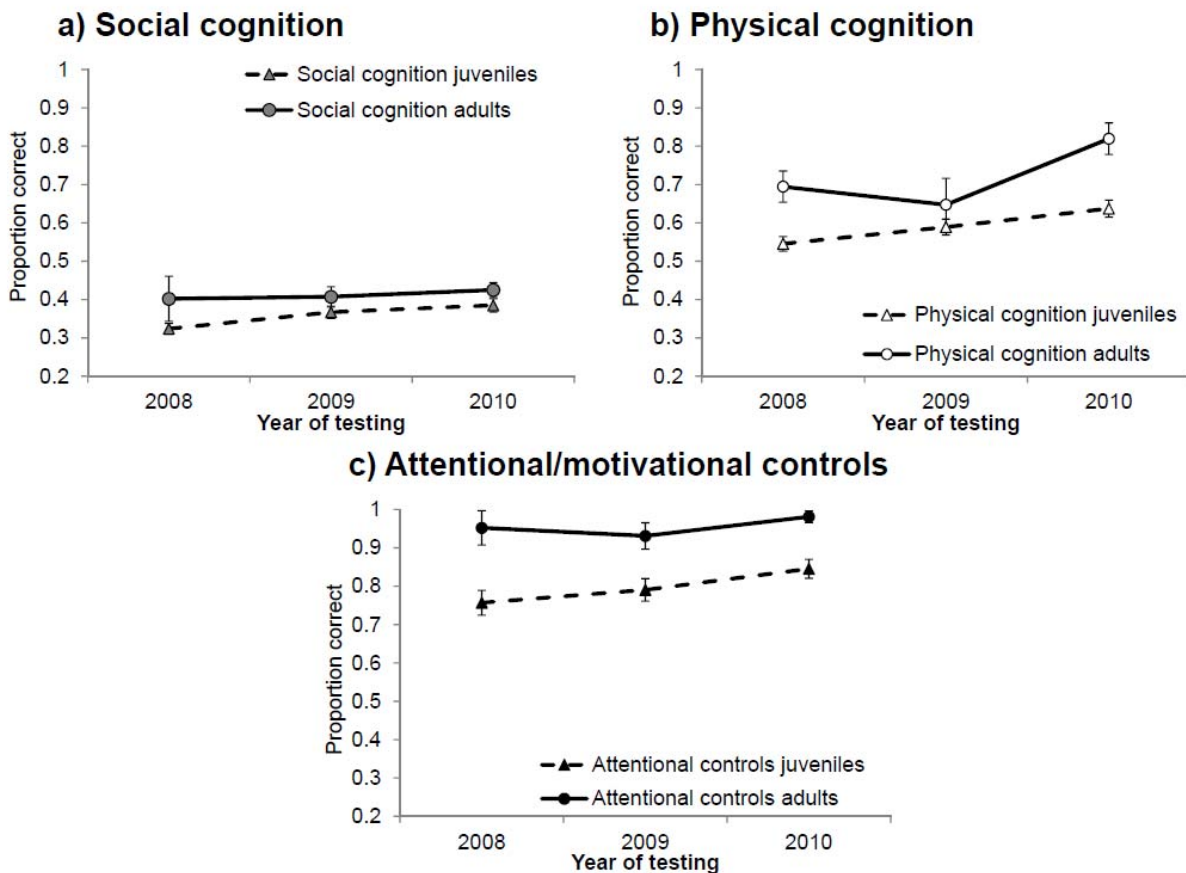


Figure 6.8. Average performance in *Pan* infants/juveniles and *Pan* adults across three years of longitudinal testing, Study 2. The y-axis denotes mean percentage correct in a) 9 social cognition, b) 5 physical cognition and c) 3 attentional/motivational control tasks, and the x-axis denotes the three tests years (2008, 2009, 2010). Average adult performance is shown with circles and solid lines, while average infant/juvenile performance is shown by triangles and dotted lines. Bars denote standard error. Performance in infants and juveniles improved in all three areas across the three years of testing, while performance in the adults did not (in physical cognition, there was a trend effect of year but this was not significant). Yet overall, the degree of improvement among *Pan* infants/juveniles was modest compared to that seen in children (Figure 6.1).

Importantly, we could rule out the possibility that these improvements were due to increased experience with the tasks by using our comparative population of 6 adults tested across all three years, concurrent with the juvenile subjects. For these 6 adults, there was no significant effect of test year in repeated measures ANOVAs for the social domain ($p>0.4$), the physical domain ($p>0.09$), or the attentional/motivational controls ($p>0.1$) (Figure 6.8). These results thus indicate that the changes in performance measured among the younger subjects represented maturational change rather than familiarity-based improvements or variance in task administration across years.

We next analyzed the difference scores for each task, which tracked within-individual changes in performance from 2008 to 2010. We found that *Pan* juveniles improved in 4 of the 5 physical cognition tasks from the first to the last task administration, showing on average a 9.2% increase in performance. Meanwhile, they improved in 7 of the 9 social cognition tasks across this time period, showing a 6.1% improvement on average. Finally, subjects also became more attentive throughout the three years of testing, showing an average 8.9% “improvement” in their performance across the three attentional/motivational control tasks (Appendix 6, Figure A6.3). Though these gains in performance led the effect of test year to be statistically significant in the repeated measures ANOVAs, these gains are modest compared to children’s average improvement of over 30% in both the social and physical domains over the course of 2 years (see Figure 6.1). Thus rather than *Pan* showing a rapid period of cognitive development in juvenility after an initial slow period in infancy, the results of Study 2 suggest that *Pan* cognitive development progresses more slowly than that of humans throughout infancy and juvenility.

Patterns of cognitive development

Age of emergence

Since several tasks in Study 1 were found to emerge after 4 years of age in *Pan* infants (see Table 6.4), we were able to use the results from the present study to ascertain whether *Pan* juveniles became successful at these tasks later on in development. Indeed, *Pan* individuals ultimately succeeded in 4 of the 6 tasks where they did not meet the emergence criterion in Study 1 (Table 6.4). The two tasks where *Pan* juveniles never succeeded, even at the oldest age tested (8 years), were those where children were successful even at 3 years of age: social learning and intention-emulation (Table 6.4). Studies with adult chimpanzees indeed demonstrate that their capacities for imitative learning and cooperatively-minded goal attribution are reduced relative to human children (Horner & Whiten, 2005; Itakura & Tanaka, 1998; Myowa-Yamakoshi, et al., 2012; Tomasello, Carpenter, Call, Behne, & Moll, 2005; Tomasello, Savage-Rumbaugh, & Kruger, 1993). Instead, chimpanzees have been found to show greater success at attributing goals to others in competitively-oriented paradigms (Braeuer, et al., 2006; Hare & Tomasello, 2004). Here, even in our competitively-oriented goal understanding task (taken from Braeuer et al, 2006), *Pan* juveniles only began to succeed at 7 years of age (Table 6.4). This is in striking contrast to human children, where individuals in the first year of life begin to attribute goals to others while also starting to track others' attention and behavior in rapid sequence (Behne, Carpenter, & Tomasello, 2005; Butterworth & Itakura, 2000; Carpenter, et al., 1998; Woodward, 1998). Instead, we found that *Pan* infants were able to track others' attention and behavior by 3 years of age but could only successfully comprehend others' goals much later on in development (Table 6.4). Goal understanding may thus be prioritized in human ontogeny relative to the socio-cognitive ontogeny of any other species, given its important role in imitative learning

(Tomasello, 2009; Tomasello, et al., 2005; Tomasello, Kruger, & Ratner, 1993). This finding highlights the importance of a comparative developmental perspective and prompts future targeted inquiry into the ontogeny of goal understanding in non-human animals.

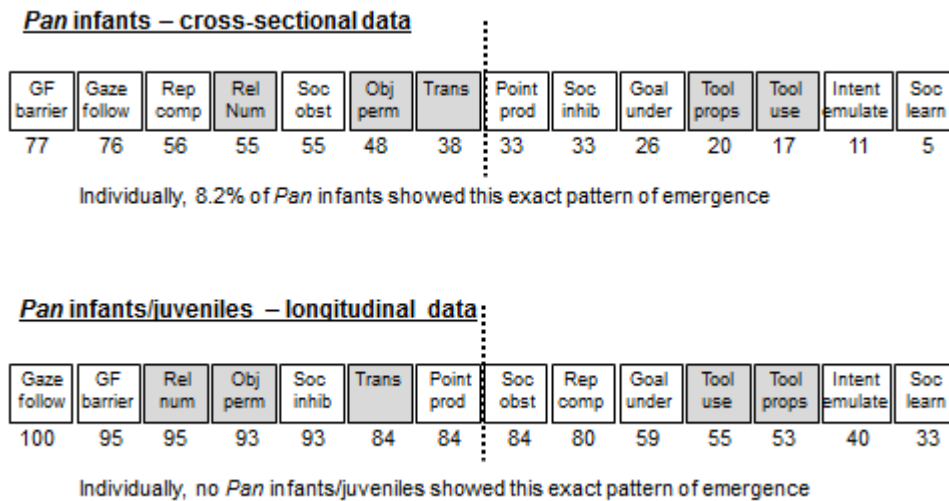


Figure 6.9. Patterns of emergence across cognitive tasks in *Pan* individuals using the cross-sectional data, Study 1, and longitudinal data, Study 2. Each task is represented by its own box, with physical cognition tasks denoted by gray boxes. The percentage of individuals meeting the passing criterion for each task is shown under its respective box – for the longitudinal data, this represents the percentage of individuals who passed the task any time during the three years of testing. The dotted line denotes the halfway point among the 14 tasks, with skills to the left of this line those where individuals of each genus could be considered most successful. Individual support for patterns of emergence within each type of data is also reported. There were few differences in pattern of emergence between the estimates of the cross-sectional and the longitudinal data, with the main changes in the social inhibition task (where individuals initially struggled but eventually succeeded) and the reputation comprehension task (where the reverse pattern was the case).

Order of emergence

The order of task emergence for *Pan* juveniles determined from the longitudinal data mapped fairly closely onto that from the cross-sectional data (Figure 6.9). As would be expected, a higher percentage of individuals met the task emergence criteria over three years of longitudinal testing than did so in one year of cross-sectional testing (though the older age of our

Study 2 sample may also have accounted for this difference). However, the relative ordering of skills was still broadly similar, with 3 of the 5 physical cognition tasks among the best-ranking while the most difficult tasks for *Pan* individuals were the intention-emulation and social learning tasks.

Individual support for the longitudinal sequence of emergence was quite low, with no individuals matching the 14-task pattern. No individual in the longitudinal data matched the pattern of emergence determined by the cross-sectional analyses, nor did any individual in the longitudinal data match the emergence pattern for children (Figure 6.3). This suggests that in fact, individual patterns of development in *Pan* were even more variable than suggested by the cross-sectional sample in Study 1. There was greater support when examining emergence sequences by domain, with 4.5% of individuals matching the 9-task social emergence pattern and 31.8% of individuals matching the 5-task physical emergence pattern (though again both of these levels of support were less than found in Study 1). This both replicated our finding that the general order of success across tasks differed between *Pan* juveniles and human children, and supported our claim that patterns of cognitive development are more consistent across individuals in children than in *Pan* (in contrast to prior assumptions that humans' patterns of development are more plastic (Bjorklund & Green, 1992; Bruner, 1974)).

Inter-task correlations

In Study 2 we found more significant correlations in performance across tasks relative to Study 1 (Table 6.7). Yet similar to Study 1, *Pan* juveniles showed more inter-relationships between tasks in the physical domain than tasks in the social domain. Across the three years of testing, there were in total 23 significant correlations between tasks within a given year (18 positive and 5 negative). Among the 18 positive relationships, 7 were between two social tasks, 5

were between two physical tasks, and 6 were cross-domain (thus 11 of these 18 relationships involved at least one physical cognition task).

Importantly, only one relationship between tasks was significant across all three years of testing – the positive correlation between performance in object permanence and performance in transposition (Table 6.7). This result aligns with previous work demonstrating a distinct spatial cognition factor comprising success in these two tasks among both adult chimpanzees and human children (Herrmann, et al., 2010). One other relationship was significant in two of the three test years – the positive relationship between success in social inhibition and success in social learning. Relatively few *Pan* subjects ever succeeded in the social learning task, indicating that those who did so may have had an exceptional level of social inhibitory control. Such links between inhibitory control and imitative learning provide an exciting avenue for further comparative developmental inquiry, given the growing body of work in developmental psychology investigating the ties between executive function and socio-cognitive skill within our own species (Carlson & Moses, 2001; Carpenter, et al., 2002; Perner, Lang, & Kloo, 2002).

Table 6.7. Correlations in performance across cognitive tasks in *Pan* infants/juveniles across three years of testing, Study 2. Social cognition tasks are divided from the physical cognition tasks by lines, with the social tasks in the upper left quadrant. Correlations between tasks solely within the social domain are shaded light gray, where correlations involving at least one task from the physical domain are shaded in dark gray. Negative correlations are shaded in slightly darker hues. Tasks where the correlation was not significant are labeled “NS.” Pearson correlation values are shown, with their significance denoted as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Task abbreviations are as listed in Table 6.1 and in the Methods section.

a) 2008

Task:	GF	GFB	SI	IE	GU	SL	Rep	PP	OP	TP	Num	Tra	TU
SO	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
GF		0.362*	NS	NS	NS	NS	-0.323*	0.321*	NS	NS	NS	NS	NS
GFB			NS	NS	NS	NS	-0.431**	0.335*	NS	NS	NS	NS	NS
SI				NS	NS	0.404**	NS	NS	NS	NS	NS	NS	NS
IE					NS	NS	NS	NS	NS	NS	NS	NS	NS
GU						NS	NS	-0.356*	NS	NS	NS	NS	NS
SL							NS	NS	NS	NS	NS	0.322*	NS
Rep								NS	NS	NS	NS	-0.400**	NS
PP									NS	NS	NS	NS	0.339*
OP										0.346*	NS	0.368*	NS
TP											NS	NS	NS
Num												NS	NS
Tra													NS

b) 2009

Task:	GF	GFB	SI	IE	GU	SL	Rep	PP	OP	TP	Num	Tra	TU
SO	NS	NS	NS	NS	NS	NS	NS	NS	-0.379*	NS	NS	NS	NS
GF		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
GFB			NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
SI				NS	NS	0.325*	NS	NS	NS	NS	NS	NS	NS
IE					NS	NS	NS	NS	NS	NS	NS	NS	NS
GU						NS	NS	NS	NS	NS	NS	NS	NS
SL							NS	NS	NS	0.333*	NS	NS	NS
Rep								NS	NS	NS	NS	NS	NS
PP									NS	NS	NS	NS	NS
OP										NS	NS	0.515***	NS
TP											NS	NS	NS
Num												NS	NS
Tra													NS

c) 2010

Task:	GF	GFB	SI	IE	GU	SL	Rep	PP	OP	TP	Num	Tra	TU
SO	NS	0.375*	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.304*	NS
GF		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	0.365*	NS
GFB			NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
SI				NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
IE					NS	NS	NS	NS	NS	NS	NS	NS	NS
GU						NS	NS	0.298*	NS	NS	NS	NS	NS
SL							NS	NS	NS	NS	NS	NS	NS
Rep								NS	NS	NS	NS	NS	NS
PP									0.321*	NS	NS	NS	NS
OP										NS	NS	0.339*	NS
TP											NS	NS	0.310*
Num												NS	NS
Tra													NS

CONCLUSIONS – STUDY 2

The results of our second study indicate that the rate of cognitive development in *Pan* remains slow throughout infancy and juvenility relative to the pace in human children. Thus there is no period in *Pan* development, either before or after weaning, when their rate of cognitive development matches that of 2- to 4-year old children. We found further support for the patterns of skill emergence in *Pan* development determined in Study 1, providing additional evidence that these patterns differ significantly from those of human children. We again found a large degree of inter-individual plasticity in patterns of development among *Pan* individuals, particularly within the social domain. To review the major findings for each type of analysis performed in Study 1 and Study 2, we provide a summary table below (Table 6.8). These results indicate significant changes in both the rate and pattern of development between humans and young *Pan*, with notable shifts in the ontogeny of social cognition.

Table 6.8. Key results from each analysis type, Study 1 and Study 2. Divisions are the same as those presented in the Results sections. Using a variety of analytical techniques, these results provide support for the hypothesis that the rate and pattern of cognitive development vary significantly between humans and other apes. They also indicate specific skills that are fundamental to differences in cognitive development between human children and *Pan* juveniles.

Study 1:		
Analysis	Subtype	Result
Rate of development	GLM analyses	Children show accelerated rate of improvement, earlier success in social cognition
Patterns of development	Age of emergence	Cognitive development slowed in <i>Pan</i> , elements of socio-cognitive development missing entirely in <i>Pan</i> infants
	Order of emergence	Children show earlier emergence of social tasks, more individual consistency in patterns of emergence, key differences in understanding others' goals
	Relative proficiency	Children show earlier proficiency in understanding others' intentions and goals, physical cognitive skills more comparable in within-individual rank across children and <i>Pan</i> infants
	Inter-correlations	Children show greater inter-task correlations, particularly in the social domain, suggests greater inter-relationship in skill development in children than in <i>Pan</i> infants
Study 2:		
Analysis	Subtype	Result
Rate of development	Repeated measures ANOVA	Rate of improvement is slow relative to children among <i>Pan</i> infants and juveniles, but is more marked than improvement among <i>Pan</i> adults
Patterns of development	Age of emergence	Several tasks emerge much later on in <i>Pan</i> relative to human development such as goal understanding, tasks requiring cooperative motivations never emerge
	Order of emergence	Relative ordering of skills similar to cross-sectional result, physical cognition skills prioritized, skills pertaining to cooperative motivations least prioritized
	Inter-correlations	More inter-relationships between tasks than cross-sectional sample, but still fewer than children, few relationships solely within the social domain

DISCUSSION

Our findings provide support for the hypothesis that changes in development are central to the psychological distinctions observed between humans and other apes. In particular, we found support for the notion of **Accelerated Ontogeny** in human psychological evolution, with our results pointing to a particularly strong role of socio-cognitive development. In Study 1, children showed more rapid improvement in performance from 2 to 4 years of age than did same-age *Pan* infants, contradicting the predictions of the **Delayed** or **Fixed Ontogeny** hypotheses. Moreover, children outperformed *Pan* already at 2 years in the social but not the physical domain, supporting the notion that trajectories of development may not be similarly linked across domains in humans and other apes. Distinctions between children and *Pan* infants were particularly notable in patterns of socio-cognitive skill emergence, with capacities for goal understanding emerging relatively (and absolutely) earlier on in human than in *Pan* development. Meanwhile, patterns of physical cognition skill emergence were largely similar between children and young *Pan*, though overall levels of proficiency were greater in children. Finally, we found greater inter-individual consistency in patterns of task emergence in children than among *Pan* infants and juveniles. Control analyses ruled out the possibility that the observed differences were due to differing levels of attention and motivation between species or age groups, or to aspects of rearing history in our non-human ape sample. Our longitudinal data from *Pan* infants and juveniles in Study 2 supported the general conclusions drawn from the cross-sectional *Pan* data collected in Study 1, demonstrating continued slow rates of cognitive development throughout *Pan* infancy and juvenility in addition to relatively late proficiency in skills of goal understanding and social learning. These results thus suggest that human cognitive

development is accelerated on the whole in comparison to our closest living relatives, with particularly marked changes in humans' development of socio-cognitive skills.

Our findings provide the first empirical evidence to address certain uniform assumptions present in the psychological and anthropological literatures. The first is that espoused by the **Delayed Ontogeny** hypothesis, in asserting that human cognitive development is slowed in line with our extended juvenile period and lifespan. Instead, we find evidence that the rate of cognitive development from 2 to 4 years of age is dramatically *accelerated* in humans when compared to that of our closest living relatives. Our findings indicate that the length of the juvenile period alone cannot be taken to represent either the speed or the characteristics of cognitive development, which is important for conclusions drawn either about the cognitive capacities of fossil (Potts, 2004; Tattersall, 2009) or living taxa (Leigh, 2004; Pagel & Harvey, 2002).

The second assumption contradicted by our results is that human psychological development is inherently more plastic than that of other species, with non-human cognitive development is fixed across individuals because it is "simpler." Here, we found that in the emergence of basic social and physical cognitive skills that underlie more complex thought, children were much more individually consistent in their ontogenetic patterns than were *Pan* infants and juveniles. Of course, humans' extended juvenility may confer additional flexibility in how these skills are built upon and ultimately employed (Bjorklund & Green, 1992; Kaplan, et al., 2000). We simply argue that in the emergence of these early fundamentals, children may in fact be more consistent than other animals. One possibility is that our tests revealed greater ties in children's development because they were adapted from the developmental psychology literature (see further discussion of this point below). Strong relationships between skills may

exist in *Pan* psychological development that we were not able to capture with the present battery of tasks. Future work investigating species-specific parameters of cognitive ontogeny will help to further investigate this area.

Overall, these results indicate an important role for comparative developmental studies in psychology and anthropology, both for our understanding of human evolution and for our interpretations of how the human mind works (Gomez, 2005; Matsuzawa, 2007; Matsuzawa, Tomonaga, & Tanaka, 2006; Tomasello & Carpenter, 2005; Wobber, in preparation). Crucial for psychologists, we note that theories regarding the way psychology develops have almost entirely been derived from one developmental sequence (our own), leaving it open to question to what degree cognitive development is flexible across species. Here we propose that an acceleration in human development might be facilitated by children's heightened abilities and motivations to learn socially (in line with Herrmann et al, 2007), with this potentially being independent from the effects of language. To determine the relative contributions of skills in social learning and factors such as language, we will need comparative developmental studies of taxa that are more proficient in cooperative learning than chimpanzees but lack language – such as meerkats or dogs (Hare, Brown, Williamson, & Tomasello, 2002; Thornton & McAuliffe, 2006; Topal, Gergely, Erdohegyi, Csibra, & Miklosi, 2009). Meanwhile, studies of deaf children can help to fully test this claim. Finally, data from human societies where parental investment is lower (Henrich, Heine, & Norenzayan, 2010; Kaplan, 1996; Marlowe, 2005) will establish the degree to which this accelerated development is a human universal.

A potential criticism of our approach is the use of conspecific experimenters for the human children but not for the *Pan* juveniles. In favor of this technique, it enabled us to have a much greater degree of precision, performing methods identically between individuals and

between tests, than if we had employed a confederate child demonstrator for our child subjects or a conspecific demonstrator for our ape subjects. Against it, having a conspecific demonstrator in certain social tasks could in theory augment the performance of non-human apes. Although one study of eye tracking did indeed suggest that chimpanzees follow the gaze of a human less readily than the gaze of a conspecific (Hattori, Kano, & Tomonaga, 2010), most social cognition tasks (including the ones utilized here) involve the perception of much less subtle cues.

Numerous studies have demonstrated that non-human apes are able to interpret the actions, intentions, and dispositions of human experimenters, and even to follow their gaze direction reliably in geometric space and around barriers (Braeuer, Call, & Tomasello, 2005; Horner & Whiten, 2005; Y. Russell, Call, & Dunbar, 2008; Warneken, Hare, Melis, Hanus, & Tomasello, 2007). In addition, studies employing both a human and a conspecific demonstrator have found little difference in chimpanzees' performance between the two situations (Hare & Tomasello, 2004; Itakura, Agnetta, Hare, & Tomasello, 1999; Myowa-Yamakoshi, et al., 2012; Tomasello, Call, & Hare, 1998). Thus ample previous research indicates that non-human apes are able to successfully perceive human experimenters as social agents. Further, in the present study, any purported enhancement of children's performance by a conspecific demonstrator would not account for 1) the differing rate of improvement within each genus (since *Pan* 2-year-olds would have been similarly affected by this potential bias as *Pan* 4-year-olds) or 2) children's relatively more skilled performance in the physical cognition tasks (where the role of the experimenter is greatly reduced). Therefore we are confident that both child and *Pan* subjects were given equal opportunities to succeed in these experiments.

Our results indicate that across species, shifts in the trajectories of cognitive ontogeny underlie species differences in adult psychology. Further research is required to elucidate

whether the changes in social cognitive development between *Pan* and humans are unique to our lineage, or whether patterns of social cognitive development are more variable across species in general. Additional study of non-human ape juveniles can target the specific aspects of cognitive development that differ from patterns found in humans, with our broad-scale analyses suggesting these differences will be most significant in the ontogeny of cooperative motivations and the understanding of others' goals. Investigating patterns of psychological ontogeny across species will also help to provide insight into variations in developmental trajectory within our own species, such as in the case of Autism Spectrum Disorders where aspects of socio-cognitive ontogeny are critically altered relative to typically-developing children (Carpenter, et al., 2002; Charman et al., 1997; Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997). On the whole then, greater investigation of the patterns of cognitive ontogeny across multiple domains both within and outside our species will provide important insight into the nature of human psychology.

Chapter 6 Literature Cited.

- Antinucci, F. (1990). The comparative study of cognitive ontogeny in four primate species. In S. Parker & K. Gibson (Eds.), *"Language" and intelligence in monkeys and apes* (pp. 157-171). Cambridge: Cambridge University Press.
- Arthur, W. (2002). The emerging conceptual framework of evolutionary developmental biology. *Nature*, 415(6873), 757-764.
- Bart, W., & Airasian, P. (1974). Determination of the ordering among seven Piagetian tasks by an ordering-theoretic method. *Journal of Educational Psychology*, 66(2), 277-284.
- Barth, J., & Call, J. (2006). Tracking the Displacement of Objects: A Series of Tasks With Great Apes (Pan troglodytes, Pan paniscus, Gorilla gorilla, and Pongo pygmaeus) and Young Children (Homo sapiens). *Journal of Experimental Psychology: Animal Behavior Processes*, 32(3), 239-252.
- Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, 8(6), 492-499.
- Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: comparing 12- and 18-month-olds. *Infant Behavior and Development*, 22(2), 277-282.
- Bjorklund, D., & Green, B. (1992). The Adaptive Nature of Cognitive Immaturity. *American Psychologist*, 47(1), 46-54.
- Braeuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology*, 119(2), 145-154.
- Braeuer, J., Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2006). Making Inferences About the Location of Hidden Food: Social Dog, Causal Ape. *Journal of Comparative Psychology*, 120(1), 38-47.
- Bruner, J. (1974). The nature and uses of immaturity. In K. Connolly & J. Bruner (Eds.), *The Growth of Competence* (pp. 11-47). London: Academic Press.
- Butterworth, G., & Itakura, S. (2000). How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology*, 18, 25-50.
- Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9, 55-72.
- Call, J., Carpenter, M., & Tomasello, M. (2005). Copying results and copying actions in the process of social learning: chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Animal Cognition*, 8(3), 151-163.

- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72(4), 1032-1053. doi: 10.1111/1467-8624.00333
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), 1-143.
- Carpenter, M., Pennington, B. F., & Rogers, S. J. (2002). Interrelations among social-cognitive skills in young children with autism. *Journal of Autism and Developmental Disorders*, 32(2), 91-106.
- Carroll, S. (2003). Genetics and the making of Homo sapiens. *Nature*, 422, 849-857.
- Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology*, 33(5), 781-789.
- Charnov, E., & Berrigan, D. (1993). Why Do Female Primates Have Such Long Lifespans and So Few Babies? or Life in the Slow Lane. *Evolutionary Anthropology*, 1, 191-194.
- Coe, C. L., Connolly, A. C., Kraemer, H. C., & Levine, S. (1979). Reproductive development and behavior of captive female chimpanzees. *Primates*, 20(4), 571-582. doi: 10.1007/bf02373439
- Doran, D. M. (1992). The ontogeny of chimpanzee and pygmy chimpanzee locomotor behavior - a case-study of paedomorphism and its behavioral correlates. *Journal of Human Evolution*, 23(2), 139-157.
- Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature*, 415, 755.
- Gomez, J. C. (2005). Species comparative studies and cognitive development. *Trends in Cognitive Sciences*, 9(3), 118-125.
- Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: Harvard University Press.
- Green, B. (1956). A method of scalogram analysis using summary statistics. *Psychometrika*, 21(1), 79-88.
- Hamlin, K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450(7169), 557-559.
- Hanus, D., & Call, J. (2007). Discrete quantity judgments in the great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, *Pongo pygmaeus*): The effect of presenting whole sets versus item-by-item. *Journal of Comparative Psychology*, 121(3), 241-249.

- Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology*, 18(9), R370-R372.
- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, 298, 1634-1636.
- Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, 68, 571-581.
- Hattori, Y., Kano, F., & Tomonaga, M. (2010). Differential sensitivity to conspecific and allospecific cues in chimpanzees and humans: a comparative eye-tracking study. *Biology Letters*, 6(5), 610-613. doi: 10.1098/rsbl.2010.0120
- Hauser, M. (1997). Artifactual kinds and functional design features: what a primate understands without language. *Cognition*, 64, 285-308.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-+. doi: 10.1017/s0140525x0999152x
- Herrmann, E., Call, J., Hernandez-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis. *Science*, 317, 1360-1366.
- Herrmann, E., Call, J., Hernandez-Lloreda, M., Hare, B., & Tomasello, M. (2010). The Structure of Individual Differences in the Cognitive Abilities of Children and Chimpanzees. *Psychological Science*, 21(1), 102-110.
- Herrmann, E., Wobber, V., & Call, J. (2008). Great apes' (Pan troglodytes, Pan paniscus, Gorilla gorilla, and Pongo Pygmaeus) understanding of tool functional properties after limited experience. *Journal of Comparative Psychology*, 122, 220-230.
- Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Animal Cognition*, 8(3), 164-181.
- Hrdy, S. (2005). Comes the child before man: how cooperative breeding and prolonged postweaning dependence shaped human potential. In B. Hewlett & M. Lamb (Eds.), *Hunter-gatherer childhoods: evolutionary, developmental, and cultural perspectives* (pp. 65-91): Aldine Transaction.
- Itakura, S., Agnetta, B., Hare, B., & Tomasello, M. (1999). Chimpanzee use of human and conspecific social cues to locate hidden food. *Developmental Science*, 2(4), 448-456.
- Itakura, S., & Tanaka, M. (1998). Use of Experimenter-Given Cues During Object-Choice Tasks by Chimpanzees (Pan troglodytes), an Orangutan (Pongo pygmaeus), and Human Infants (Homo sapiens). *Journal of Comparative Psychology*, 112(2), 119-126.

- Kagan, J., & Snidman, N. (2004). *The Long Shadow of Temperament*. Cambridge, MA: Harvard University Press.
- Kano, T. (1992). *The Last Ape: Pygmy Chimpanzee Behavior and Ecology*. Stanford, California: Stanford University Press.
- Kaplan, H. (1996). A Theory of Fertility and Parental Investment in Traditional and Modern Human Societies. *Yearbook of Physical Anthropology*, 39, 91-135.
- Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology*, 9(4), 156-185.
- Kennedy, G. E. (2005). From the ape's dilemma to the weanling's dilemma: early weaning and its evolutionary context. *Journal of Human Evolution*, 48(2), 123-145. doi: 10.1016/j.jhevol.2004.09.005
- Kramer, K. L., & Ellison, P. T. (2010). Pooled Energy Budgets: Resituating Human Energy Allocation Trade-offs. *Evolutionary Anthropology*, 19(4), 136-147. doi: 10.1002/evan.20265
- Langer, J. (2000). The descent of cognitive development. *Developmental Science*, 3(4), 361-388.
- Leekam, S., Baron-Cohen, S., Perrett, D., Milders, M., & Brown, S. (1997). Eye-direction detection: A dissociation between geometric and joint attention skills in autism. *British Journal of Developmental Psychology*, 15, 77-95.
- Leigh, S. R. (2004). Brain growth, life history, and cognition in primate and human evolution. *American Journal of Primatology*, 62(3), 139-164. doi: 10.1002/ajp.20012
- Lieberman, D. E. (2012). *The Evolution of the Human Head*. Cambridge, MA: Harvard University Press.
- Liu, X., Somel, M., Tang, L., Yan, Z., Jiang, X., Guo, S., . . . Khaitovich, P. (in press). Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques. *Genome Research*.
- Lonsdorf, E. V., Ross, S., & Matsuzawa, T. (Eds.). (2011). *The Mind of the Chimpanzee: Ecological and Experimental Perspectives*. Chicago: University of Chicago Press.
- Marlowe, F. W. (2005). Hunter-gatherers and human evolution. *Evolutionary Anthropology*, 14(2), 54-67.
- Marson, J., Meuris, S., Cooper, R. W., & Jouannet, P. (1991). Puberty in the male chimpanzee: progressive maturation of semen characteristics. *Biology of Reproduction*, 44(3), 448-455.
- Martin, P., & Bateson, P. (1986). *Measuring behavior*. Cambridge: Cambridge University Press.

- Matsuzawa, T. (2007). Comparative cognitive development. *Developmental Science*, 10(1), 97-103.
- Matsuzawa, T., Tomonaga, M., & Tanaka, M. (2006). *Cognitive development in chimpanzees*. Tokyo: Springer-Verlag.
- Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science*, 311(5765), 1297-1300.
- Meltzoff, A. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, 24(4), 470-476.
- Miklosi, A., Kubinyi, E., Topal, J., Gacsi, M., Viranyi, Z., & Csanyi, V. (2003). A simple reason for a big difference: wolves do not look back at humans, dogs do. *Current Biology*, 13, 763-766.
- Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, 7(1), F1-F9.
- Myowa-Yamakoshi, M., Scola, C., & Hirata, S. (2012). Humans and chimpanzees attend differently to goal-directed actions. *Nature Communications*, 3, 1-7.
- Natale, F., Antinucci, F., Spinozzi, G., & Poti, P. (1986). Stage 6 Object Concept in Nonhuman Primate Cognition : A Comparison Between Gorilla (*Gorilla gorilla gorilla*) and Japanese Macaque (*Macaca fuscata*). *Journal of Comparative Psychology*, 100(4), 335-339.
- Pagel, M., & Harvey, P. (2002). Evolution of the juvenile period in mammals. In M. Pereira & L. Fairbanks (Eds.), *Juvenile Primates* (2nd ed., pp. 28-37). Chicago: University of Chicago Press.
- Perner, J., Lang, B., & Kloo, D. (2002). Theory of Mind and Self-Control: More than a Common Problem of Inhibition. *Child Development*, 73(3), 752-767.
- Phillips, W., Baron-Cohen, S., & Rutter, M. (1992). The role of eye contact in goal detection: Evidence from normal infants and children with autism or mental handicap. *Development and Psychopathology*, 4, 375-383.
- Piaget, J. (1952). *The Child's Conception of Number*. London: Routledge and Kegan Paul.
- Pontzer, H., & Wrangham, R. (2006). Ontogeny of ranging in wild chimpanzees. *International Journal of Primatology*, 27(1), 295-309.
- Potts, R. (2004). Paleoenvironmental basis of cognitive evolution in great apes. *American Journal of Primatology*, 62(3), 209-228. doi: 10.1002/ajp.20016
- Robson, S., & Wood, B. (2008). Hominin life history: reconstruction and evolution. *Journal of Anatomy*, 212, 394-425.

- Russell, J. L., Lyn, H., Schaeffer, J. A., & Hopkins, W. D. (2011). The role of socio-communicative rearing environments in the development of social and physical cognition in apes. *Developmental Science*, 14(6), 1459-1470. doi: 10.1111/j.1467-7687.2011.01090.x
- Russell, Y., Call, J., & Dunbar, R. (2008). Image scoring in great apes. *Behavioural Processes*, 78(1), 108-111.
- Scott, J. A., Binns, C. W., Oddy, W. H., & Graham, K. I. (2006). Predictors of breastfeeding duration: Evidence from a cohort study. *Pediatrics*, 117(4), E646-E655. doi: 10.1542/peds.2005-1991
- Sellen, D. W. (2001). Comparison of infant feeding patterns reported for nonindustrial populations with current recommendations. *Journal of Nutrition*, 131(10), 2707-2715.
- Sellen, D. W., & Smay, D. B. (2001). Relationship between subsistence and age at weaning in "preindustrial" societies. *Human Nature-an Interdisciplinary Biosocial Perspective*, 12(1), 47-87. doi: 10.1007/s12110-001-1013-y
- Somel, M., Franz, H., Yan, Z., Lorenc, A., Guo, S., Giger, T., . . . Khaitovich, P. (2009). Transcriptional neoteny in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 106(14), 5743-5748. doi: 10.1073/pnas.0900544106
- Sophian, C. (1984). Spatial transpositions and the early development of search. *Developmental Psychology*, 20(1), 21-28.
- Tattersall, I. (2009). Human origins: Out of Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 106(38), 16018-16021. doi: 10.1073/pnas.0903207106
- Thornton, A., & McAuliffe, K. (2006). Teaching in wild meerkats. *Science*, 313(5784), 227-229. doi: 10.1126/science.1128727
- Tomasello, M. (2009). *Why we cooperate*. Cambridge, MA: The MIT Press.
- Tomasello, M., & Call, J. (1997). *Primate Cognition*. New York: Oxford University Press.
- Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, 55, 1063-1069.
- Tomasello, M., & Camaioni, L. (1997). A comparison of the gestural communication of apes and human infants. *Human Development*, 40, 7-24.
- Tomasello, M., & Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development*, 70(1), 1-155.

- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675-735.
- Tomasello, M., Hare, B., & Fogleman, T. (2001). The ontogeny of gaze following in chimpanzees, *Pan troglodytes*, and rhesus macaques, *Macaca mulatta*. *Animal Behaviour*, 61, 335-343.
- Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16, 495-511.
- Tomasello, M., Savage-Rumbaugh, S., & Kruger, A. (1993). Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Development*, 64(6), 1688-1705.
- Tomonaga, M. (2008). Relative numerosity discrimination by chimpanzees (*Pan troglodytes*): evidence for approximate numerical representations. *Animal Cognition*, 11(1), 43-57.
- Topal, J., Gergely, G., Erdohegyi, A., Csibra, G., & Miklosi, A. (2009). Differential Sensitivity to Human Communication in Dogs, Wolves, and Human Infants. *Science*, 325(5945), 1269-1272.
- Vinicius, L. (2005). Human encephalization and developmental timing. *Journal of Human Evolution*, 49(6), 762-776. doi: 10.1016/j.jhevol.2005.08.001
- Vlammings, P., Hare, B., & Call, J. (2010). Reaching around barriers: the performance of the great apes and 3–5-year-old children. *Animal Cognition*, 13(2), 273-285.
- Walker, R., Hill, K., Burger, O., & Hurtado, A. (2006). Life in the Slow Lane Revisited: Ontogenetic Separation Between Chimpanzees and Humans. *American Journal of Physical Anthropology*, 129, 577-583.
- Warneken, F., Hare, B., Melis, A., Hanus, D., & Tomasello, M. (2007). Spontaneous Altruism by Chimpanzees and Young Children. *PLOS Biology*, 5(7), 1414-1420.
- Whiten, A., Goodall, J., McGrew, W., Nishida, T., Reynolds, V., Sugiyama, Y., . . . Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399, 682-685.
- Wobber, V. (in preparation). The value of comparative developmental studies of behavior and cognition to anthropology.
- Wobber, V., & Hare, B. (2011). Psychological health of orphan bonobos and chimpanzees in African sanctuaries. *PLOS One*, 6(6), e17147.
- Wobber, V., Herrmann, E., Hare, B., Wrangham, R., & Tomasello, M. (in preparation). Distinctions between social and physical cognitive evolution illuminated by comparative developmental inquiry in chimpanzees and bonobos.

- Wobber, V., Wrangham, R., & Hare, B. (2010). Bonobos exhibit delayed development of social behavior and cognition relative to chimpanzees. *Current Biology*, 20(3), 226-230.
- Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, 1-34.

Chapter 7: **Discussion**

The results of this thesis suggest several important conclusions for our understanding of human evolution. First, they indicate that behavioral and cognitive differences between species can arise via shifts in behavioral and cognitive development. Second, they indicate that changes in the relationship between physiology and behavior are also crucial in shaping species differences in sociality. Finally, these results provide evidence of paedomorphosis, or retention of juvenile characteristics into adulthood, across numerous traits in bonobos relative to chimpanzees. They support the hypothesis that broad-scale shifts in development were prevalent in the human lineage as well. Here I place the results of each chapter in a broader context.

In Chapter 3, I found that bonobos retain juvenile behavioral and cognitive traits into adulthood, with these shifts in development underlying adult behavioral and cognitive differences between bonobos and chimpanzees (Wobber, Wrangham, & Hare, 2010). In addition to the results discussed in this thesis, I also found that less skillful performance by adult bonobos in tasks assessing physical cognition derived from slower acquisition of these skills throughout bonobo ontogeny (Wobber, Herrmann, Hare, Wrangham, & Tomasello, in preparation). Together with the evidence of paedomorphosis in aspects of the bonobo cranium (Durrleman, Pennec, Trouve, Ayache, & Braga, 2012; Lieberman, Carlo, Ponce de Leon, & Zollikofer, 2007), these results indicate analogous changes in numerous aspects of bonobo development relative to chimpanzees and substantiate the argument of a chimpanzee-like last common ancestor of the two species (Hare, Wobber, & Wrangham, in press; Wrangham & Pilbeam, 2001). More generally, they suggest that changes in behavioral or cognitive complexity between species can arise via changes in the rate of development, contrary to a prevailing assumption in the literature

that shifts in cognitive complexity between species are primarily facilitated by shifts in the duration of the juvenile period (Kaplan, Hill, Lancaster, & Hurtado, 2000; Walker, Burger, Wagner, & Von Rueden, 2006). Future comparative studies of behavioral and cognitive development are necessary to determine the degree to which these findings are generalizable across taxa.

In Chapter 4, I found evidence for a potential mechanism underlying broad developmental differences between bonobos and chimpanzees, namely shifts in the production of steroid hormones throughout development. A growing body of evidence suggests that hormones are critical in shaping life history patterns (Ketterson & Nolan, 1992; McGlothlin & Ketterson, 2008). In line with this research, I found that the shifts in bonobo behavioral, cognitive, and morphological development are accompanied by shifts in the production of testosterone throughout bonobo development relative to chimpanzees. Namely, while chimpanzees showed the mammalian-typical pattern of low testosterone levels during juvenility followed by increases in the transition to adulthood, bonobo testosterone levels were largely constant throughout development, with little elevation in testosterone levels at puberty. These results provide further evidence that ontogenetic pathways have changed substantially between bonobos and chimpanzees. To determine whether these shifts in bonobo testosterone production characterize a broader change in their pubertal maturation, future research can investigate the production of gonadotropins and energetic mediators such as leptin throughout bonobo ontogeny.

In Chapter 5, to build on the conclusions of my developmental studies of endocrinology and of behavior, I directly investigated the interaction between behavior and endocrinology among bonobos and chimpanzees. My results revealed that in addition to the broader ontogenetic shifts between the two species, bonobos and chimpanzees of all ages also differ in their

immediate physiological shifts surrounding competitive interactions. Faced with an identical dyadic feeding competition, male chimpanzees showed significant shifts in their levels of testosterone, while male bonobos instead shifted in their levels of cortisol. Intriguingly, males of both species showed analogous pre-competition shifts in these separate steroids, with decreases in steroid level among partners who ultimately shared equally and increases in steroid level among partners who shared unequally. This suggests that males of both species were equally able to appraise the competitive event, but in doing so may have viewed the competition quite differently, with chimpanzees perceiving the competition as critical to status while bonobos instead viewed the competition as a stressor. Future comparisons of the two species can illuminate to what degree these differences in behavioral neuroendocrinology are specific to the domain of competition or characteristic of numerous aspects of their social behavior. More generally, these results highlight that key differences may exist in the proximate mechanisms influencing behavior across species, generating an exciting direction for future inquiry in studying comparative behavioral and cognitive neuroendocrinology.

Finally, in Chapter 6, I examined the evolutionary origins of human cognitive development through direct comparison of young humans, chimpanzees, and bonobos. Utilizing a broad-scale battery of cognitive tasks, I found that human children showed dramatic acceleration in their rate of cognitive development relative to same-age *Pan*. These gains were particularly notable in the domain of social cognition, where children excelled relative to *Pan* individuals even at the earliest age tested (with performance in the physical domain more comparable among the youngest children and same-age *Pan*). These results suggest two exciting conclusions regarding human evolution. First, they indicate that human cognitive development is accelerated in comparison to other apes despite our species' general slowdown in somatic and

reproductive maturation (Walker, Hill, Burger, & Hurtado, 2006). In line with the findings of Chapter 3, these results also build on existing models suggesting that changes in the length of the juvenile period facilitate increases in cognitive complexity across species (Kaplan, et al., 2000). My results suggest that changes in the rate of cognitive development can also be critical in shaping species differences in cognition, with humans' accelerated rate of cognitive development enabling children to acquire greater skill in any given unit of developmental time relative to other ape juveniles. The second major conclusion that can be drawn from my results is that mechanisms of social learning are key in facilitating this accelerated development in human children, with humans' capacities and motivations to build on the knowledge of others unparalleled by any other species (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007; Tomasello, Carpenter, Call, Behne, & Moll, 2005). Additional inquiry in this area can examine the degree to which this acceleration in cognitive development is present across human societies.

The results of this thesis prompt numerous directions for future research. They suggest that comparative studies of development, using multiple species to directly quantify the evolution of life history patterns, will prove most valuable in combatting the singularity problem inherent to studying human evolution (c.f. (MacLean et al., in press)). They also indicate that studies of bonobos are essential in understanding human origins, whether bonobos are taken as an alternative model for the last common ancestor of humans and the genus *Pan* (Zihlman, Cronin, Cramer, & Sarich, 1978) or as a case of convergent evolution with humans in diverging from a chimpanzee-like progenitor (Hare, et al., in press; Wrangham & Pilbeam, 2001). In addition, the present results highlight the importance of integrating studies of development across multiple traits, to elucidate the degree to which patterns of behavioral and cognitive

ontogeny scale with parameters of somatic and reproductive maturation. Finally, my findings indicate that studies of development will prove instrumental in understanding behavioral and cognitive evolution, particularly in determining the mechanisms underlying behavioral and cognitive change in human evolution.

Chapter 7 Literature Cited.

- Durrleman, S., Pennec, X., Trouve, A., Ayache, N., & Braga, J. (2012). Comparison of the endocranial ontogenies between chimpanzees and bonobos via temporal regression and spatiotemporal registration. *Journal of Human Evolution*, 62(1), 74-88. doi: 10.1016/j.jhevol.2011.10.004
- Hare, B., Wobber, V., & Wrangham, R. (in press). The self-domestication hypothesis: bonobo psychology evolved due to selection against aggression. *Animal Behaviour*.
- Herrmann, E., Call, J., Hernandez-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis. *Science*, 317, 1360-1366.
- Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology*, 9(4), 156-185.
- Ketterson, E. D., & Nolan, V. (1992). Hormones and life histories - an integrative approach. *American Naturalist*, 140, S33-S62.
- Lieberman, D. E., Carlo, J., Ponce de Leon, M., & Zollikofer, C. (2007). A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. *Journal of Human Evolution*, 52, 647-662.
- MacLean, E., Matthews, L., Hare, B., Nunn, C., Anderson, R., Aureli, F., . . . Wobber, V. (in press). How Does Cognition Evolve?: Phylogenetic Comparative Psychology. *Animal Cognition*.
- McGlothlin, J., & Ketterson, E. (2008). Hormone-mediated suites as adaptations and evolutionary constraints *Philosophical Transactions of the Royal Society B*, 363(1497), 1611-1620.
- Stanford, C. (1998). The social behavior of chimpanzees and bonobos. *Current Anthropology*, 39(4), 399-420.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675-735.
- Walker, R., Burger, O., Wagner, J., & Von Rueden, C. R. (2006). Evolution of brain size and juvenile periods in primates. *Journal of Human Evolution*, 51(5), 480-489. doi: 10.1016/j.jhevol.2006.06.002
- Walker, R., Hill, K., Burger, O., & Hurtado, A. (2006). Life in the Slow Lane Revisited: Ontogenetic Separation Between Chimpanzees and Humans. *American Journal of Physical Anthropology*, 129, 577-583.

- Wobber, V., Herrmann, E., Hare, B., Wrangham, R., & Tomasello, M. (in preparation). Distinctions between social and physical cognitive evolution illuminated by comparative developmental inquiry in chimpanzees and bonobos.
- Wobber, V., Wrangham, R., & Hare, B. (2010). Bonobos exhibit delayed development of social behavior and cognition relative to chimpanzees. *Current Biology*, 20(3), 226-230.
- Wrangham, R., & Pilbeam, D. (2001). African apes as time machines. In B. Galdikas, N. Briggs, L. Sheeran, G. Shapiro & J. Goodall (Eds.), *All apes great and small* (Vol. 1, pp. 5-18). New York: Kluwer Academic/Plenum Publishers.
- Zihlman, A., Cronin, J., Cramer, D., & Sarich, V. (1978). Pygmy chimpanzee as a possible prototype for the common ancestor of humans, chimpanzees, and gorillas. *Nature*, 275(5682), 744-746.

Appendix 1: **Distinguishing behavior and cognition**

In discussing studies of behavioral and cognitive development, it is necessary to define what I mean by “cognition.” Here I will use the definition of Tomasello and Call (Tomasello & Call, 1997), that cognition is a subset of behavior. Cognitive processes are those that involve organisms making choices among possible courses of action that involve some kind of mental representation (Tomasello & Call, 1997). Thus not all behaviors are cognitive, such as those that are inflexible, automatic responses. In turn, there are a wide variety of behaviors that do involve cognitive underpinnings. For example, choosing a tree to forage in likely relies on a cognitive process, but chewing one’s food once in the tree would not require underlying cognitive mechanisms. Though cognitive processes in animals have been investigated for over a century, and the ontogeny of cognition has been well-studied in humans, research into the ontogeny of cognition in other animals has been rare.

It is important to draw distinctions between studies of behavior and cognition because there are often critical differences in methodologies used to explore each. Behavior is usually best quantified through observational study, documenting patterns via extended observation.(Martin & Bateson, 1986) However, cognitive abilities are difficult to infer from observation alone, since one must obtain information about an individual’s mental representation or the deliberateness of an action.(Janson & Byrne, 2007; Tomasello & Call, 1997) In this way, experimental methods can help to tease out cognitive processes, in examining how individuals choose to act in specific, controlled situations. Though this often necessitates working with populations where environmental manipulation is possible, such as semi free-ranging (Rawlins & Kessler, 1987; Wobber & Hare, 2011) or captive individuals, it is nonetheless valuable to

document the range of potential capacities present in a given species. In addition, there are a growing number of techniques to perform cognitive experiments in the wild (through playbacks (Cheney & Seyfarth, 1988) or minor environmental modifications (Gruber, Muller, Strimling, Wrangham, & Zuberbuhler, 2009)), allowing a broader range of cognitive abilities to be indexed across multiple living environments. Meanwhile, studies of behavior are also becoming more sophisticated with the proliferation of long-term research sites among primate populations, (Lwanga, Struhsaker, Struhsaker, Butynski, & Mitani, 2011; Wright, 1999) allowing us to better understand typical behavior patterns independent from any influences of a given season or observer.(Sapolsky & Share, 2004) Thus there are a number of exciting opportunities to build on these existing techniques developed for research with adults in extending them to the area of development.

Appendix 1 Literature Cited.

- Cheney, D. L., & Seyfarth, R. M. (1988). Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, 36, 477-486. doi: 10.1016/s0003-3472(88)80018-6
- Gruber, T., Muller, M. N., Strimling, P., Wrangham, R., & Zuberbuhler, K. (2009). Wild Chimpanzees Rely on Cultural Knowledge to Solve an Experimental Honey Acquisition Task. *Current Biology*, 19(21), 1806-1810. doi: 10.1016/j.cub.2009.08.060
- Janson, C., & Byrne, R. (2007). What wild primates know about resources: opening up the black box. *Animal Cognition*, 10(4), 357-367.
- Lwanga, J. S., Struhsaker, T. T., Struhsaker, P. J., Butynski, T. M., & Mitani, J. C. (2011). Primate Population Dynamics Over 32.9 Years at Ngogo, Kibale National Park, Uganda. *American Journal of Primatology*, 73(10), 997-1011. doi: 10.1002/ajp.20965
- Martin, P., & Bateson, P. (1986). *Measuring behavior*. Cambridge: Cambridge University Press.
- Rawlins, R., & Kessler, M. (1987). *The Cayo Santiago Macaques*. New York: SUNY University Press.
- Sapolsky, R. M., & Share, L. J. (2004). A pacific culture among wild baboons: Its emergence and transmission. *Plos Biology*, 2(4), 534-541.
- Tomasello, M., & Call, J. (1997). *Primate Cognition*. New York: Oxford University Press.
- Wobber, V., & Hare, B. (2011). Psychological health of orphan bonobos and chimpanzees in African sanctuaries. *PLOS One*, 6(6), e17147.
- Wright, P. C. (1999). Lemur traits and Madagascar ecology: Coping with an island environment. In C. Ruff (Ed.), *Yearbook of Physical Anthropology 1999, Vol 42* (Vol. 42, pp. 31-72). New York: Wiley-Liss, Inc.

Appendix 2: **Insights into human evolution from taxa distantly related to humans**

Focus on understanding human behavioral and cognitive evolution has largely been placed on one of our closest living relatives, the chimpanzee. This emphasis is understandable, given the genetic proximity of chimpanzees to humans (Mikkelsen et al., 2005; Ruvolo, 1997), the early origins of research on chimpanzees in the wild (Goodall, 1986), and the frequent findings that they share capacities with humans that were once thought to be unique to our species (Melis, Hare, & Tomasello, 2006; Warneken & Tomasello, 2006; Whiten, Horner, & Marshall-Pescini, 2003). However, we can add to our knowledge of human behavioral and cognitive evolution by broadening the range of study taxa.

A first example is better study of humans' "other" closest living relative, the bonobo. Bonobos are equally genetically related to humans as are chimpanzees (Satta, Klein, & Takahata, 2000; Won & Hey, 2005), but have received less research effort – largely due to the difficulties of working in the politically tumultuous Democratic Republic of Congo (the bonobo range is confined entirely within this country). Nonetheless, recent studies of bonobos have made important contributions to our understanding of the behavioral and cognitive capacities that might have been present in the last common ancestor (LCA) of humans and *Pan* (Hare, 2009; Hare & Kwetuenda, 2010; Herrmann, Hare, Call, & Tomasello, 2010; Parish & de Waal, 2000; Surbeck, Mundry, & Hohmann, 2011; White & Wood, 2007). This LCA might have behaved more like a chimpanzee (Wrangham & Pilbeam, 2001), more like a bonobo (De Waal & Lanting, 1997; Zihlman, Cronin, Cramer, & Sarich, 1978), or behaved in a way that combines aspects of modern-day chimpanzees and bonobos. Regardless of which is the case, further exploration of bonobo behavior and cognition can help us to quantify the range of variability within great apes

and, through direct comparisons with chimpanzees, the degree to which such traits can change in short evolutionary time.

Beyond the study of great apes, there are numerous taxa from which we can gain insight into the nature of the human mind. Comparative anthropological research has focused on non-human primates, given humans' membership in the primate order, and indeed we have learned a great deal about the phylogenetic origins of capacities for theory of mind, numerical reasoning, and even skills involved in language (Brannon, 2006; Flombaum & Santos, 2005; Savage-Rumbaugh, Shanker, & Taylor, 2001). However, recent studies from more distant taxa such as birds and even domestic dogs can also clarify important elements of human evolutionary history. Birds in particular provide remarkable opportunities to index differences in behavior and cognition between closely-related species that have differing ecological niches (N. Emery, 2006; Lefebvre, Reader, & Sol, 2004). With their sophisticated capacities in song learning (Doupe & Kuhl, 1999), tool use (Lefebvre, Nicolakakis, & Boire, 2002), memory (Salwiczek, Watanabe, & Clayton, 2010), and even social cognition (Dally, Emery, & Clayton, 2006; N. Emery & Clayton, 2001; Seed, Clayton, & Emery, 2008), additional studies of birds will help us to characterize how evolution shapes trajectories of behavioral and cognitive development (N. J. Emery & Clayton, 2004).

Meanwhile, studies of domestic dogs provide another avenue by which to investigate the origins of human social behavior and cognition. In particular, this is because domestic dogs are genetically distant from our species but have evolved while living in close proximity to humans. As such, research has found that they share a number of socio-cognitive capacities with our species, particularly in the area of cooperative learning (Hare, Brown, Williamson, & Tomasello, 2002; Topal, Gergely, Erdohegyi, Csibra, & Miklosi, 2009; Wobber & Hare, 2009; Wobber,

Hare, Koler-Matznick, Wrangham, & Tomasello, 2009). Studies of dogs thus provide a novel opportunity by which to determine the effects of the human social environment on cognitive capacities independent from the effects of language (though dogs do show some capability in linguistic recognition (Kaminski, Call, & Fischer, 2004)). Integration of the comparative developmental approach across studies of varying taxa will allow us to clarify the mechanisms (both proximate and ultimate) that shape behavioral evolution as well as life history evolution on the whole.

Appendix 2 Literature Cited.

- Brannon, E. (2006). The representation of numerical magnitude. *Current Opinion in Neurobiology*, 16(2), 222-229.
- Dally, J. M., Emery, N. J., & Clayton, N. S. (2006). Food-caching western scrub-jays keep track of who was watching when. *Science*, 312(5780), 1662-1665.
- De Waal, F., & Lanting, F. (1997). *Bonobo: the forgotten ape*. Berkeley: University of California Press.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567-631. doi: 10.1146/annurev.neuro.22.1.567
- Emery, N. (2006). Cognitive ornithology: the evolution of avian intelligence. *Philosophical Transactions of the Royal Society B*, 361, 23-43.
- Emery, N., & Clayton, N. (2001). Effects of experience and social context on prospective caching strategies in scrub jays. *Nature*, 414, 443-446.
- Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science*, 306(5703), 1903-1907. doi: 10.1126/science.1098410
- Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, 15(5), 447-452.
- Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: Harvard University Press.
- Hare, B. (2009). What is the effect of affect on bonobo and chimpanzee problem solving? In A. Berthoz & Y. Christen (Eds.), *The Neurobiology of the Umwelt: how living beings perceive the world* (pp. 89-102): Springer Press.
- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, 298, 1634-1636.
- Hare, B., & Kwetuenda, S. (2010). Bonobos voluntarily share their own food with others. *Current Biology*, 20(5), R230-R231.
- Herrmann, E., Hare, B., Call, J., & Tomasello, M. (2010). Differences in the Cognitive Skills of Bonobos and Chimpanzees. *PLOS One*, 5(8), e12438.
- Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: evidence for "fast mapping". *Science*, 304, 1682-1683.

- Lefebvre, L., Nicolakakis, N., & Boire, D. (2002). Tools and brains in birds. *Behaviour*, 139, 939-973. doi: 10.1163/156853902320387918
- Lefebvre, L., Reader, S., & Sol, D. (2004). Brains, Innovations and Evolution in Birds and Primates. *Brain Behavior and Evolution*, 63, 233-246.
- Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science*, 311(5765), 1297-1300.
- Mikkelsen, T. S., Hillier, L. W., Eichler, E. E., Zody, M. C., Jaffe, D. B., Yang, S. P., . . . Waterston, R. H. (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, 437(7055), 69-87.
- Parish, A., & de Waal, F. (2000). The Other “Closest Living Relative” How Bonobos (*Pan paniscus*) Challenge Traditional Assumptions about Females, Dominance, Intra- and Intersexual Interactions, and Hominid Evolution. *Annals of the New York Academy of Sciences*, 907(1), 97-113.
- Ruvolo, M. (1997). Molecular phylogeny of the hominoids: inferences from multiple independent DNA sequence data sets. *Molecular Biology and Evolution*, 14(3), 248-265.
- Salwiczek, L. H., Watanabe, A., & Clayton, N. S. (2010). Ten years of research into avian models of episodic-like memory and its implications for developmental and comparative cognition. *Behavioural Brain Research*, 215(2), 221-234. doi: 10.1016/j.bbr.2010.06.011
- Satta, Y., Klein, J., & Takahata, N. (2000). DNA Archives and Our Nearest Relative: The Trichotomy Problem Revisited. *Molecular Phylogenetics and Evolution*, 14(2), 259-275.
- Savage-Rumbaugh, E. S., Shanker, S., & Taylor, T. (2001). *Apes, Language, and the Human Mind*. Oxford: Oxford University Press.
- Seed, A., Clayton, N., & Emery, N. (2008). Cooperative problem solving in rooks (*Corvus frugilegus*) *Proceedings of the Royal Society of London: B*, 275(1641), 1421-1429.
- Surbeck, M., Mundry, R., & Hohmann, G. (2011). Mothers matter! Maternal support, dominance status and mating success in male bonobos (*Pan paniscus*). *Proceedings of the Royal Society B-Biological Sciences*, 278(1705), 590-598. doi: 10.1098/rspb.2010.1572
- Topal, J., Gergely, G., Erdohegyi, A., Csibra, G., & Miklosi, A. (2009). Differential Sensitivity to Human Communication in Dogs, Wolves, and Human Infants. *Science*, 325(5945), 1269-1272.
- Warneken, F., & Tomasello, M. (2006). Altruistic Helping in Human Infants and Young Chimpanzees. *Science*, 311, 1301-1303.
- White, F., & Wood, K. (2007). Female Feeding Priority in Bonobos, *Pan paniscus*, and the Question of Female Dominance. *American Journal of Primatology*, 69(8), 837-850.

- Whiten, A., Horner, V., & Marshall-Pescini, S. (2003). Cultural Panthropology. *Evolutionary Anthropology*, 12(2), 92-105.
- Wobber, V., & Hare, B. (2009). Testing the social dog hypothesis: are dogs also more skilled than chimpanzees in non-communicative social tasks? *Behavioural Processes*, 81, 423-428.
- Wobber, V., Hare, B., Koler-Matznick, J., Wrangham, R., & Tomasello, M. (2009). Evidence for two waves of selection on the social skills of dogs. *Interaction Studies*, 10(2), 206-224.
- Won, Y., & Hey, J. (2005). Divergence Population Genetics of Chimpanzees. *Molecular Biology and Evolution*, 22(2), 297-307.
- Wrangham, R., & Pilbeam, D. (2001). African apes as time machines. In B. Galdikas, N. Briggs, L. Sheeran, G. Shapiro & J. Goodall (Eds.), *All apes great and small* (Vol. 1, pp. 5-18). New York: Kluwer Academic/Plenum Publishers.
- Zihlman, A., Cronin, J., Cramer, D., & Sarich, V. (1978). Pygmy chimpanzee as a possible prototype for the common ancestor of humans, chimpanzees, and gorillas. *Nature*, 275(5682), 744-746.

Appendix 3:
Supplemental methods, analyses, tables, and figures for Chapter 3

SUPPLEMENTAL ANALYSES

Experiment 1

Comparisons of food sharing with Hare et al, 2007

In the current study there were no significant differences between species in the number of trials where individuals shared or co-fed when age was not considered as a variable (mean trials ($\pm$ SEM): *sharing*: chimpanzees, 5.87 ($\pm$ 0.68), bonobos, 6.58 ($\pm$ 0.45); *co-feeding*: chimpanzees, 2.53 ($\pm$ 0.65), bonobos, 2.92 ($\pm$ 0.50); independent-samples t-tests, $p = \text{NS}$). This is in contrast to Hare et al (2007), who found a significant species difference in sharing between bonobos and chimpanzees. While the means for sharing and co-feeding are higher in the bonobos than in the chimpanzees in the present study, we believe that the lack of a significant species difference here is due to important distinctions between the sampling and the testing procedure in the two studies. First, the chimpanzees tested by Hare et al (2007) were older than those in the present study. The mean age of the chimpanzees used by Hare et al (2007) was 11.8 years $\pm$ 1.0 SE while the current sample's mean age was 9.3 years $\pm$ 0.8 SE which is marginally younger in absolute terms (independent samples t-test, $t(60)=1.875$, $p = 0.07$). Meanwhile the mean age of the bonobos in Hare et al, 2007 (9.6 years $\pm$ 0.8 SE) was similar to the mean age of the current sample (9.0 years $\pm$ 1.0 SE, independent samples t-test, $p = \text{NS}$). Given that here we demonstrated that chimpanzees share less as adults, the older age of the chimpanzees in Hare et al 2007 would have led to the larger species differences seen there.

Moreover, the food that subjects could share in the Hare et al 2007 was placed outside the testing room and was more difficult to obtain while also protecting one's body (when reaching outside to take food, individuals were exposed to potential aggression from their partner). In contrast, in the current study food was placed inside the testing room (the rationale for this is described below in the Supplemental Methods). We suspect that the younger chimpanzee sample used here and the greater ease with which food could be obtained by subordinate chimpanzees led to more equivalent overall food sharing scores. Finally, in the Hare et al, 2007 paper, sharing was coded using a different definition taking into account the pattern where individuals took turns monopolizing across trials. This pattern did not occur in the current study.

Sharing and age estimates using subjects' weight

While we have confidence in the age estimates made according to the criteria described in the main manuscript, we also used subjects' weights from a recent physical exam as a potential proxy for age at the time of testing. When using these weight estimates (represented as a pair's average weight), the pattern of older chimpanzees being less tolerant was still strongly significant (linear regressions with weight, *sharing*, $r^2 = 0.30$, $p = 0.04$, *co-feeding*, $r^2 = 0.35$, $p = 0.02$). Bonobos' regressions between weight and the sharing behaviors exhibited remained non-significant.

Analysis of non-feeding social behaviors

We analyzed two social behaviors not pertaining to feeding in the task to complement the analyses of sharing behavior. In play behavior, chimpanzees exhibited a decrease with age, with a significant negative relationship between dyad age and the number of trials in which play occurred (linear regression, $r^2 = 0.40$, $p = 0.01$), while

bonobos only showed a marginal decrease in play with increasing age ($r^2 = 0.31$, $p = 0.06$) (Figure A3.1). A 2x2 ANOVA of play behavior with species and age category as factors revealed a significant effect of species ($F(1, 26) = 6.08$, $p = 0.02$) and a significant effect of age category ($F(1,26) = 13.74$, $p = 0.001$), but no interaction. Post-hoc tests showed that bonobos played more than chimpanzees overall, consistent with previous work (Palagi, 2006), and in both species juveniles played more than adults (Table A3.3). In examining sexual behavior, regressions with age were not significant in either species. However, a species x age category ANOVA revealed a significant effect of species ($F(1,26) = 14.49$, $p = 0.001$) on the number of trials where sexual behavior was exhibited (Table A3.3). Bonobos of all ages engaged in more sexual behavior than chimpanzees of all ages.

Sharing depending on food configuration

As expected given past work showing distinctions in sharing between configurations where food was clumped or dispersed (Hare, Melis, Woods, Hastings, & Wrangham, 2007; Melis, Hare, & Tomasello, 2006), there were significant differences across food placement conditions in the behaviors exhibited here. We performed a repeated measures ANOVA with condition (3 levels, according to the differing food configurations as described below) as a within-subject factor and species and age category as between-subjects factors to assess the differences in sharing behaviors and the other social behaviors. For *sharing*, this ANOVA revealed a significant effect of condition ($F(2,46) = 24.62$, $p < 0.001$), and no effect or interaction of species, with the predicted effect of age category also present ($F(1,23) = 4.13$, $p = 0.05$). Post-hoc analyses showed that for both species and all age categories, sharing was least likely in the

condition where there was one, monopolizable food source (CS), with this significantly less than the condition where there were two food piles (DD) (Tukey's HSD, $p < 0.001$), and significantly less than the condition where there was one divisible pile of food (CD) (Tukey's HSD, $p < 0.001$). For *co-feed*, the results were similar – there was a significant effect of condition ($F(2,46) = 11.21$, $p < 0.001$), and a significant effect of age category ($F(1,23) = 15.67$, $p = 0.001$), but no interaction or effect of species. Post-hoc tests revealed that co-feeding was significantly less likely in the CS condition than in the CD condition (Tukey's HSD, $p < 0.001$), but that there were no other differences between conditions. This suggests that in both species, across ages, subjects had the most difficulty sharing or co-feeding when the food was highly monopolizable (presented in two large pieces, CS).

Repeated measures ANOVAs of the two social behaviors revealed no differences across conditions in play, though there were significant main effects of species ($F(1,23) = 6.24$, $p = 0.02$) and age category ($F(1,23) = 13.17$, $p = 0.001$) on play as discussed in the main paper. In sexual behavior as well, condition did not affect the amount exhibited, and there was the expected main effect of species ($F(1,23) = 14.49$, $p = 0.001$).

Experiment 2

Controls for species differences in motivation

In order to test if the effects we observed might simply be due to motivational factors differing between individuals, we examined the performance of subjects on 4 introduction trials, which were presented prior to the test trials. In these trials, food was held by two adjacent experimenters. Apes are more skilled with these adjacent choices than the nonadjacent choices presented in the test trials (Barth & Call, 2006; Call, 2001).

There was no species difference in performance on the introductory trials, nor any relationship between age and performance in either species (Table 3.2). Thus even the youngest bonobos were as skilled as older apes at solving the simpler introductory trials, implying that their difficulty in the test was only in avoiding the middle experimenter and making it unlikely that the age effects we observed were due to motivational differences between ages and species.

We also utilized session time as a proxy for motivation to complete the test, with the assumption that subjects who were less motivated would take longer to choose on a given trial. There was a slight difference between species in the length of the test session (mean in seconds ($\pm$ SEM): chimpanzees, 581 ($\pm$ 43), bonobos, 823 ($\pm$ 100), $t(37) = -2.19$, independent samples t-test, $p = 0.04$), but test session length did not correlate with performance in either species (linear regressions, $p = \text{NS}$). In addition, when partialing out the effect of session length on the relationship between age and performance, the correlation in chimpanzees remained nonsignificant (partial correlation = 0.05, $p = \text{NS}$) while the correlation in bonobos remained significant (partial correlation = 0.58, $p = 0.009$). Therefore, overall there was little evidence that the differences in performance were due to motivation factors.

Outliers in age and performance

Two chimpanzees could be considered outliers in this data set since they chose correctly on 0 of the 12 test trials, so we performed separate analyses after removing these two individuals. The correlation between age and performance in chimpanzees remained non-significant when removing these two individuals ($r^2 = 0.11$, $p = \text{NS}$).

Experiment 3

Baseline preferences within and between species.

In 10 baseline trials, subjects could choose to take food from either of the two experimenters (E1 or E2), as both displayed a food reward in their open hand. Individuals of both species tended to prefer E1 to E2, in that they chose this experimenter significantly more than would be predicted by chance (mean trials choosing E1 ($\pm$ SEM): chimpanzees, 7.81 (0.50), bonobos, 7.71 (0.32); one-sample t-test, chimpanzees, $t(10) = 5.62$, $p < 0.001$; bonobos $t(16) = 8.21$, $p < 0.001$). In both species, E1 was the first author (V.W.), who had administered other tests to subjects so was more familiar in the food giving context than the E2s at either sanctuary. However, performance on the baseline trials did not differ between the two species (independent samples t-test, $p = \text{NS}$) and did not correlate with age in either species (linear regressions, $p = \text{NS}$). Further, performance on these trials did not correlate with performance in the introduction trials or the reversal trials (linear regressions). This suggests that any preferences that individuals had for one experimenter did not differ across species or ages, or affect performance on the test.

Controls for species differences in motivation

Similar to the previous experiment, here we used session time as an estimate of subjects' motivation to take part in the test. There was no species difference in the average time to complete a session (mean in seconds ($\pm$ SEM): chimpanzees, 1,061 ($\pm$ 97), bonobos, 1,129 ($\pm$ 88), independent samples t-test, $p = \text{NS}$), suggesting that subjects were equally interested in the experiment. There was no correlation between session time and performance on the reversal in either species (linear regressions, $p = \text{NS}$). There was a marginal correlation between session time and the number of trials taken to reach the

introductory 84% correct criterion, simply because performing more trials took more time ($r^2 = 0.14$, $p = 0.06$). After partialling out the effect of session length, the correlation in bonobos between age and overall reversal performance remained significant (partial correlation $p = 0.66$, $p = 0.009$), as did the correlation with age and performance on the last 10 reversal trials (partial correlation = 0.69, $p = 0.006$) while the correlations in chimpanzees remained nonsignificant (partial correlations, $p = \text{NS}$). Thus, it is unlikely that simple differential motivation across species or ages influenced the patterns found here.

Outliers in age and performance

It was possible that the bonobo correlation was driven by the fact that a few older individuals were used, so the correlation was re-analyzed after removing the three bonobos over 20 years of age. When the remaining 14 individuals were re-analyzed, the correlation between age and total reversal performance in bonobos did not remain significant, but the correlation between age and the last 10 trials of the reversal was still nearly significant ($r^2 = 0.27$, $p = 0.06$), suggesting that this correlation was influenced but not completely determined by these older individuals.

Performance and age estimates using subjects' weight

When using weight as the independent variable, the correlation with performance on the last 10 trials of the reversal was even stronger in bonobos than that between performance and age (linear regression, $r^2 = 0.47$, $p = 0.003$), though weight did not correlate with overall reversal performance ($r^2 = 0.17$, $p = 0.11$). The correlation between weight and performance on the last 10 trials also remained significant when removing the

3 oldest bonobos ($r^2 = 0.41$, $p = 0.02$). There was no significant relationship between weight and performance in chimpanzees ($p = \text{NS}$).

SUPPLEMENTAL TABLES AND FIGURES

Table A3.1. Subject list, experiments 1-3. Subjects were tested at Tchimpounga Chimpanzee Sanctuary and Lola ya Bonobo Sanctuary. Tests were carried out in the summers of 2006, 2007, and 2008, leading to subjects participating in these experiments at differing ages.

Subject	Species	Sex	Experimental participation	Age in Exp. 1	Age in Exp. 2	Age in Exp. 3
Agnagui	Chimpanzee	F	1,2,3	5	4	5
Binda	Chimpanzee	F	1	5		
Blake	Chimpanzee	M	1,2	5	5	
Bumbo	Chimpanzee	M	1	17		
Cheri	Chimpanzee	F	1,3	13		13
Christophe	Chimpanzee	M	1,3	17		17
Diba	Chimpanzee	F	1	12		
Elikia	Chimpanzee	M	1	19		
Gao	Chimpanzee	F	2		4	
Imphondo	Chimpanzee	F	1	13		
Kiki	Chimpanzee	F	1,3	17		17
Kimenga	Chimpanzee	M	2		3	
Koyamba	Chimpanzee	M	1,2	4	3	
Kuzanza	Chimpanzee	M	1,2	4	3	
Likabou	Chimpanzee	F	1,2	6	5	
Loufoua	Chimpanzee	M	2		5	
Loufoumbou	Chimpanzee	M	2		3	
Lounama	Chimpanzee	F	1	7		
Maku	Chimpanzee	F	1	6		
Mandzi	Chimpanzee	M	2		3	
Marcelle	Chimpanzee	F	3			5
Maya	Chimpanzee	M	1	13		
Mbifani	Chimpanzee	F	3			7
Mbolo	Chimpanzee	F	1	11		
Moka	Chimpanzee	M	2		4	
Mpilinitu	Chimpanzee	M	1,3	6		6
Mvouti	Chimpanzee	F	1,2	5	4	
Ngoro	Chimpanzee	F	2		2	
Nzambi	Chimpanzee	F	1,2,3	5	4	5
Nzeke	Chimpanzee	M	2		4	
Outounda	Chimpanzee	M	1,2	6	5	
Pembele	Chimpanzee	F	1	15		
Petit Pere	Chimpanzee	M	1	6		
Petit Prince	Chimpanzee	M	1	7		
Ramsay	Chimpanzee	F	1,3	11		11
Taliane	Chimpanzee	M	2		7	
Tavich	Chimpanzee	M	1,2	6	6	
Tchibanga	Chimpanzee	M	3			10
Tchivingina	Chimpanzee	F	2		2	
Tsere	Chimpanzee	M	1,3	12		12
Tsowa	Chimpanzee	F	1	6		
Vitika	Chimpanzee	M	2		7	
Wolo	Chimpanzee	M	1	10		
Yoko	Chimpanzee	M	1	11		

Table A3.1 (Continued)

Api	Bonobo	M	1,2,3	7	7	7
Bandundu	Bonobo	F	1,3	10		10
Bili	Bonobo	M	1,2,3	5	4	5
Boende	Bonobo	M	2		5	
Boyoma	Bonobo	M	2		3	
Dilolo	Bonobo	M	1,2	6	5	
Etumbe	Bonobo	F	1,3	21		21
Isiro	Bonobo	F	1,3	9		9
Kalina	Bonobo	F	1	9		
Kasongo	Bonobo	M	2		4	
Kikongo	Bonobo	M	1,2,3	6	5	6
Kikwit	Bonobo	M	1,3	9		9
Kisantu	Bonobo	F	1	9		
Kubulu	Bonobo	M	1,2,3	5	4	5
Likasi	Bonobo	F	1,2,3	6	5	6
Lisala	Bonobo	F	1,2	6	5	
Lodja	Bonobo	F	1,2	4	3	
Lomami	Bonobo	M	3			8
Lukaya	Bonobo	F	1	6		
Luozi	Bonobo	M	2		3	
Mabali	Bonobo	M	2		4	
Makali	Bonobo	M	1,3	23		23
Malou	Bonobo	F	1,2,3	5	5	5
Maniema	Bonobo	M	1,2	6	6	
Manono	Bonobo	M	1,3	13		13
Max	Bonobo	M	1,3	21		21
Mbandaka	Bonobo	M	2,3		6	6
Mixa	Bonobo	M	1,3	8		8
Muanda	Bonobo	F	1,2	4	3	
Nioki	Bonobo	F	1	8		
Opala	Bonobo	F	3			12
Tembo	Bonobo	F	1	9		
Tshileng	Bonobo	F	2		3	
Vanga	Bonobo	M	2		2	
Yolo	Bonobo	M	2		4	

Table A3.2. Performance of individual subject pairs, experiment 1. The average age (in years) of a pair is shown, as is the number of trials (out of 9 total) where subjects shared food and where subjects co-fed on a pile simultaneously. The number of trials in which chimpanzee pairs shared or co-fed decreased with age, while this decrease was not present in bonobos.

Pair	Species	Average age of pair	Number of trials share	Number of trials co-feed
Kuz-Bla	Chimpanzee	4	9	1
Agn-Koy	Chimpanzee	4	9	7
Mpi-Bin	Chimpanzee	5	7	6
Nza-Lik	Chimpanzee	5	8	5
Mvo-Mak	Chimpanzee	5	2	2
Pri-Tav	Chimpanzee	6	9	7
Out-Per	Chimpanzee	6	5	2
Lou-Tso	Chimpanzee	6	8	3
Wol-Mbo	Chimpanzee	10	7	2
May-Ram	Chimpanzee	12	6	1
Dib-Imp	Chimpanzee	12	2	0
Tse-Pem	Chimpanzee	13	3	1
Bum-Yok	Chimpanzee	14	6	0
Kik-Che	Chimpanzee	15	2	0
Eli-Chr	Chimpanzee	18	5	1
Lodj-Mua	Bonobo	4	7	3
Bili-Dil	Bonobo	5	9	3
Kubu-Mal	Bonobo	5	8	6
Api-Mani	Bonobo	6	6	2
Kiko-Lik	Bonobo	6	7	6
Lisa-Luk	Bonobo	6	4	3
Kikw-Kis	Bonobo	9	5	0
Kali-Ban	Bonobo	9	5	3
Mano-Nio	Bonobo	10	8	2
Temb-Max	Bonobo	15	5	3
Maka-Mix	Bonobo	15	8	3
Isir-Etu	Bonobo	15	7	1

Table A3.3. Social behaviors exhibited across species and age groups in the tolerance test, experiment 1. Scores are out 18 total opportunities to exhibit these behaviors (9 test trials and 9 pre-test periods). Age groups are divided into juvenile and adult, as described in the manuscript. Means for each variable are listed, separated by species, with standard error in parentheses.

	Play	Sexual behavior
Chimpanzee juveniles	2.38 (0.80)	0.25 (0.16)
Chimpanzee adults	0.00 (0.00)	0.00 (0.00)
Chimpanzee mean	1.27 (0.52)	0.13 (0.09)
Bonobo juveniles	7.67 (2.44)	3.50 (1.48)
Bonobo adults	0.83 (0.83)	2.33 (0.72)
Bonobo mean	4.25 (1.61)	2.92 (0.80)

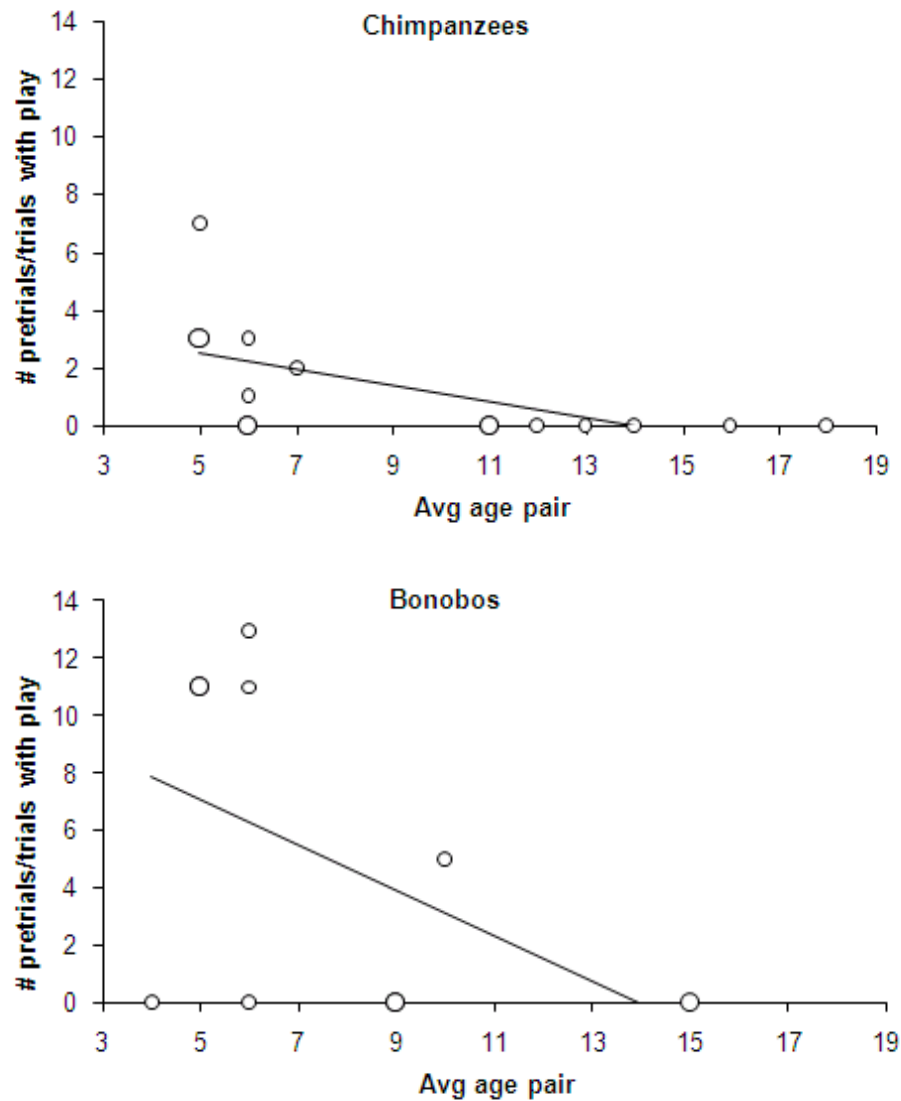


Figure A3.1. Play behavior across species and ages, experiment 1. Values here are across the 9 pre-trials and 9 test trials of experiment 1, thus there were a total of 18 possibilities to play. Small circles represent one dyad while large circles represent multiple dyads with the same behavioral score.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

General Procedures

Subjects

These three experiments were carried out with bonobos at the Lola ya Bonobo sanctuary in Kinshasa, Democratic Republic of Congo, and chimpanzees at Tchimpounga chimpanzee sanctuary in Pointe Noire, Republic of Congo. Nearly all of these apes were born in the wild and came to the sanctuary after being confiscated at an early age from the bushmeat trade (a few were born at the sanctuary and raised there by their mothers). Correspondingly, all subjects were unrelated.

At both of these sanctuaries the chimpanzees and bonobos live in social groups of 10 to 20 individuals that have access to large areas of primary tropical forest (15-40 hectares) during the day and sleep in dormitories at night. Subjects were tested in the rooms of these dormitories. They were never food deprived for testing and water was available throughout the tests. Subjects could choose to stop participating in the testing at any time by refusing to approach the experimenter or sitting by the exit of the building. All individuals were motivated to participate in experiment 1, in experiments 2 and 3 there were some individuals who met abort criteria put in place to ensure that subjects were motivated (see below).

Because experiments 1-3 were performed at the same two locations, some individuals participated in multiple experiments. Supplemental Table 3.1 below provides a list of the subjects, which experiments they participated in, and how old they were at the time of participation. Among both ape populations, ages are only estimates since individuals entering the sanctuaries are orphans of largely unknown origin. The present

estimates were made based on weight and dental data taken at the time of the test and/or at the time of an individual's arrival at the sanctuary.

Experiment 1

Subjects

Experiment 1 was carried out with 24 bonobos (forming 12 pairs) and 30 chimpanzees (15 pairs). The mean age of the juvenile group of bonobos did not differ significantly from that of the juvenile chimpanzees (mean in years ($\pm$ SEM): bonobos, 5.4 ($\pm$ 0.3); chimpanzees, 5.6 ($\pm$ 0.2); independent samples t-test, $p = \text{NS}$) nor did the mean age of the adult bonobos differ significantly from the mean age of the adult chimpanzees (mean in years ($\pm$ SEM): bonobos, 14.0 ($\pm$ 2.0); chimpanzees, 13.6 ($\pm$ 0.8); independent samples t-test, $p = \text{NS}$). The average difference between the ages of each member of a dyad was similar between bonobos and chimpanzees (mean difference in age in years ($\pm$ SEM): bonobos, 1.5 ($\pm$ 0.2); chimpanzees, 1.5 ($\pm$ 0.1); independent samples t-test, $p = \text{NS}$). Equal numbers of pairs of each sex combination (male-male, male-female, and female-female) were tested in each species, resulting in 4 pairs of each sex combination in bonobos and 5 pairs of each in chimpanzees.

Procedure

Subjects were always tested in pairs. Pairs were tested in 3 separate food sharing conditions, each of which occurred on a separate day. Subjects received 3 subsequent trials of each condition on the day that condition was presented. This resulted in a total of 9 sharing trials per pair. The order in which the conditions were presented was counterbalanced across pairs according to species, age category, and sex.

The three different conditions manipulated the shareability of the food, as has been done in previous food sharing experiments (Hare, et al., 2007; Melis, et al., 2006). Food was placed inside the room, rather than outside the room on a platform (Hare, et al., 2007), because the design of the housing facilities did not allow a platform of comparable size to that used previously to be placed outside the testing room without concrete walls blocking one individuals' access to the food. Therefore, the food piles were placed inside the room, so that in the dispersed condition the piles could be at the distance of 3 meters utilized in prior work (Hare, et al., 2007). Chimpanzees received bananas and bonobos received apples, according to the relative motivation for these items (each species highly preferred the food employed). The conditions were as follows:

Dispersed-divisible (DD): two piles of food were placed 3 meters apart inside the test room. For the chimpanzees, each pile consisted of a banana sliced into 8 pieces (thus a total of 16 slices were placed, 8 in one pile and 8 in another pile). For the bonobos, an apple was cut into 32 pieces and each pile consisted of 8 of these standard size pieces (thus a total of 16 apple slices were placed, 8 in one pile and 8 in another).

Clumped-divisible (CD): one food pile was placed in the testing room, with 16 pieces of banana/apple (of the same-sized slices mentioned above) all placed in this one pile.

Clumped-solid (CS): one food pile was placed in the testing room, but rather than being small pieces of sliced banana/apple there were either two whole bananas or two quarters of an apple (with each quarter being the same amount as the 8 pieces placed in the other trials).

Food slices were placed so that no slices were touching each other, and so that the slices were always within an approximately 50 cm circle.

The procedure of the testing days was as follows. Subjects were kept in a room adjacent to the test room so they could witness the experimenter placing the food. After the food was placed in the test room, the subjects were videotaped for one minute to capture any anticipatory behaviors (see behavioral codes below). They were then released into the testing room, with the entire test filmed as well. Each trial was considered finished when both subjects finished eating their food. The next trial began (with the same one minute anticipation period) immediately after the previous one was finished.

Videos of the test trials and the one-minute anticipatory period (pre-trials) were coded by the first author and a randomly chosen 20% of the trials were coded for reliability by a coder blind to the hypotheses of the study. The codes were simply presence/absence (0/1) codes. Reliability scores for each behavioral variable are listed below. The behavioral definitions were as follows:

Sharing: both individuals in the dyad obtained food at any point during the trial. This could be obtaining a piece of the food in the pile or simply a scrap that the other individual had dropped (Cohen's kappa = 1.00, $p < 0.001$).

Co-feeding: both individuals in the dyad simultaneously fed on the food pile (Cohen's kappa = 0.75, $p < 0.001$).

Play: contact between individuals that resulted in laughter (Cohen's kappa = 0.68, $p < 0.001$). The reliability for this measure was low because the behavior was not exhibited frequently; there were only four cases of disagreement, all where chasing

occurred between the two individuals but there was no laughter so the coders differentially scored the behavior.

Sociosexual behavior: genital-genital contact between the two subjects (Cohen's kappa = 0.88, $p < 0.001$).

The results of the behavioral coding were represented as the number of trials in which a given behavior occurred. For the two social behaviors (play and sexual behavior), coding was performed during both the test trials and the pre-test anticipation periods. Thus, these behaviors could occur a maximum of 18 times for a given pair (in 9 pre-trials and 9 trials).

Analyses

Subjects' weights were used as another proxy for age in the control analyses, since ages are only estimates. These weights were taken when individuals were anesthetized for health examinations. The weights used in the analysis were at the longest 6 months from the date that subjects participated in the behavioral experiment. One bonobo was not included in the weight analysis because he was not put under anesthesia and thus could not be weighed. All statistics were two-tailed in this and the subsequent experiments, unless otherwise stated.

Experiment 2

Experiment 2 was carried out with 20 bonobos and 20 chimpanzees. The identities of the three experimenters that took part in this experiment were consistent within subjects, but varied between subjects. A caretaker or familiar individual always served as the middle experimenter (adding to the difficulty of bypassing this individual), whereas two less familiar experimenters served as the outer individuals. Subjects

received peanuts, raisins, or apples as a reward, depending on their relative motivation for these items.

As a proxy for motivation to complete the test, session times were measured from video (i.e., subjects that were unmotivated would take longer to choose and thus would have longer test sessions). We recorded session time as the time between when the experimenters first reached down for the food in the first warm-up trial and when the subject was given its last piece of food on the 12th test trial. One chimpanzee was not included in this analysis because its session video was corrupted thus length could not be recorded. Subjects' weights were not available for use in this study because the experiment was conducted more than 6 months from when subjects had last been weighed.

Experiment 3

Experiment 3 was carried out with 17 bonobos and 11 chimpanzees. Subjects each received one test session, which consisted of baseline trials, introduction trials, and reversal trials all presented subsequently. In every type of trial, the subject was presented with a choice between two human experimenters who stood 2 meters apart in front of the mesh separating them from the subject. To begin each trial, the experimenters took food (or pretended to, depending on the condition) from a bucket or bag placed in a central location out of the subject's reach. They then returned to their positions 2 m apart, and a caretaker approached and presented the subject with a small piece of food equidistant from the two experimenters at the mesh so that the subject would begin the trial in the center. Once the subject was centered, the experimenters lifted up their arms toward the mesh to allow the subject to choose one of their hands as a potential location of food and

held their arms there either until the subject chose or 30 seconds had passed. A subject's choice was coded when it protruded something (e.g. a finger, a piece of straw) through the mesh toward one of the experimenters. Since the experimenters were 2 m apart, choices were unambiguous. If a subject had not chosen either individual in 30 seconds, the trial was coded as a "no choice." If a subject failed to choose on more than 3 trials throughout the session, the session was aborted. The data from those individuals was not included in the analyses.

To begin each session, the subject received 10 baseline trials. These trials were designed to assess any pre-existing preference the subject might have for one of the experimenters. In these trials, both experimenters presented their open hands to the subject and both were holding food. The subject was given the piece of food by the experimenter it chose. Whichever individual was chosen less during the baseline became the initial reward-holding individual in the introduction test trials.

In both the introduction and reversal test trials, the experimenters presented closed fists to the subject, and only one individual was holding food. On each trial both individuals reached into the bucket or bag to simulate having food, thus the subject did not know who had food and had to use its learning from previous trials in order to choose correctly. It is unlikely that subjects were able to visually discern who took the piece of food from the bucket or bag, since they often chose the incorrect experimenter. Subjects were given the reward if they chose correctly. If subjects chose incorrectly, the incorrect experimenter opened her hand to show it was empty and the correct experimenter opened her hand and showed the subject that she had held the food.

Subjects received trials of the introduction until they obtained the criterion of choosing correctly on at least 84% of trials. This 84% criterion did not include the first trial, since this trial served simply to show the subject where the food was located (following the procedure used in past reversal learning paradigms (Rumbaugh & Pate, 1984)). Subjects had to receive at least 10 trials of the introduction. Subjects then received 20 trials of the reverse association (again after one signal reversal trial where performance was not included in the total score). These methods were identical to those described elsewhere (Wobber & Hare, 2009).

The experimenters remained on the same side throughout the test session, through the baseline and all of the test trials, but the side on which each stood was counterbalanced across subjects. The same two experimenters were used across subjects of each species, but the identity of one experimenter changed between the two species (one, the first author, remained constant across the two species). Session times were coded from video, and measured as the time between the presentation of the first baseline trial and the subject's last choice on its 20th reversal trial. Peanuts were used as rewards for all subjects.

Subjects' weights were the same as those used in experiment 1, since again these experiments were carried out within 6 months of the health examinations when subjects were anesthetized and weighed.

Appendix 3 Literature Cited.

- Barth, J., & Call, J. (2006). Tracking the Displacement of Objects: A Series of Tasks With Great Apes (*Pan troglodytes*, *Pan paniscus*, *Gorilla gorilla*, and *Pongo pygmaeus*) and Young Children (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes*, 32(3), 239-252.
- Call, J. (2001). Object permanence in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and children (*Homo sapiens*). *Journal of Comparative Psychology*, 115(2), 159-171.
- Hare, B., Melis, A. P., Woods, V., Hastings, S., & Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology*, 17(7), 619-623.
- Melis, A. P., Hare, B., & Tomasello, M. (2006). Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Animal Behaviour*, 72, 275-286.
- Palagi, E. (2006). Social Play in Bonobos (*Pan paniscus*) and Chimpanzees (*Pan troglodytes*): Implications for Natural Social Systems and Interindividual Relationships. *American Journal of Physical Anthropology*, 129, 418-436.
- Rumbaugh, D., & Pate, J. L. (1984). The evolution of cognition in primates: a comparative perspective. In H. Roitblat, T. Bever & H. Terrace (Eds.), *Animal cognition* (pp. 569-587). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wobber, V., & Hare, B. (2009). Testing the social dog hypothesis: are dogs also more skilled than chimpanzees in non-communicative social tasks? *Behavioural Processes*, 81, 423-428.

Appendix 4:
Supplemental methods and analyses for Chapter 4

SUPPLEMENTAL TABLES AND FIGURES

Table A4.1. List of significant effects from the Generalized Estimating Equations (GEE) analysis of log testosterone with individual, species and sex as factors and age (in years) as a covariate. GEE analyses were performed with the overall log averages, as well as separately for the early morning and mid-day testosterone log averages, and separately for each year of sampling. Though a full factorial model was run for each analysis, this table shows only the significant effects and interactions from each model, along with their respective Wald chi-square values and p-values. We also report the full-model “Quasi-Likelihood under Independence Model Criterion,” or QIC, and sample size used in each model. Note that smaller QIC values indicate a better model fit. The two outlier points from female chimpanzees were removed for this and all subsequent statistical analyses.

Table A4.1

	Wald chi-square	p	Whole model	
			QIC	N
<i>All Samples:</i>			11.623	209
Species	39.581	<0.001		
Age	12.990	<0.001		
Species*Age	11.070	0.001		
Sex*Age	4.295	0.038		
<i>Early Morning Samples Only:</i>			10.927	162
Species	11.277	0.001		
Age	13.169	<0.001		
Species*Age	4.763	0.029		
<i>Mid-Day Samples Only:</i>			14.073	181
Species	44.210	<0.001		
Age	9.243	0.002		
Species*Age	6.805	0.009		
<i>All Samples, by Year:</i>				
2007			2.153	50
Species	18.961	<0.001		
Age	5.815	0.016		
Species*Age	12.918	<0.001		
2008			4.468	107
Species	7.721	0.005		
Age	4.945	0.026		
2009			3.507	52
Species	27.649	<0.001		
Age	6.155	0.013		
Species*Age	8.970	0.003		

Table A4.2. Significant post-hoc comparisons in the GEE analysis of log testosterone with individual, species, sex, and age category as factors. P-values were adjusted for multiple comparisons using a Bonferroni correction. All possible pairwise comparisons were investigated for both the species*age category and the sex*age category interactions. Only those where significant differences were present are shown here, with all other comparisons non-significant. The group in each comparison for which log testosterone values were lower is shown in the left column of the table.

Species*Age Category		
Significant post-hoc comparisons		Bonferroni-corrected p value
Chimpanzee infants	Chimpanzee adults	0.024
Chimpanzee infants	Bonobo infants	<0.001
Chimpanzee infants	Bonobo juveniles	<0.001
Chimpanzee infants	Bonobo subadults	<0.001
Chimpanzee infants	Bonobo adults	<0.001
Chimpanzee juveniles	Chimpanzee adults	0.004
Chimpanzee juveniles	Bonobo infants	<0.001
Chimpanzee juveniles	Bonobo juveniles	<0.001
Chimpanzee juveniles	Bonobo subadults	<0.001
Chimpanzee juveniles	Bonobo adults	<0.001
Sex*Age Category		
Significant post-hoc comparisons		Bonferroni-corrected p value
Infant males	Adult males	0.011
Infant females	Adult males	0.001
Juvenile males	Adult males	0.006
Juvenile females	Subadult males	0.005
Juvenile females	Adult males	<0.001
Subadult females	Adult males	0.014
Adult females	Adult males	0.002

Table A4.3. Analysis of sex differences in log testosterone, separated by species and age group. The three-way interaction between species, sex, and age category was not significant in our broader GEE analysis, but we wanted to determine the age at which a sex difference in testosterone level emerged in each species based on prior findings in the literature that chimpanzees are more sexually dimorphic in adult androgen production than are bonobos. Because we were only interested in the sex difference within species and age category, rather than performing all possible pairwise comparisons of this three-way interaction we compared the two sexes in each age group of each species. We therefore adjusted the p-values of these independent samples t-tests only for the number of comparisons performed in this hypothesis-driven post-hoc analysis (eight). P-values that were above 0.05 after Bonferroni correction are denoted in the table as “NS.”

Independent samples t-tests of log testosterone in each sex

Group where sexes compared	T-test	Bonferroni-corrected p value
Infant chimpanzees	t(19) = 0.252, p = 0.8	NS
Juvenile chimpanzees	t(46) = 1.183, p = 0.2	NS
Subadult chimpanzees	t(26) = 1.238, p = 0.2	NS
Adult chimpanzees	t(20) = 3.065, p = 0.006	0.048
Infant bonobos	t(16) = 0.199, p = 0.8	NS
Juvenile bonobos	t(34) = 1.787, p = 0.08	NS
Subadult bonobos	t(19) = 0.970, p = 0.3	NS
Adult bonobos	t(13) = 1.190, p = 0.3	NS

Table A4.4. Regression parameters for the relationship between weight (in kilograms) and log testosterone across the overall data and subsets of the data. Linear regressions were performed separately by species and sex. Regressions were performed with the overall log averages, as well as separately for the early morning and mid-day testosterone log averages. Weight analyses could not be run separately by year due to small sample size ($n < 3$) in certain groups. The slope, correlation coefficient (R), p-value (p), and sample size (N) for each regression are shown. The mean weight for the points involved in each regression is also listed, to provide an indication of the characteristics of that subset of the data. Significant p-values are indicated in bold.

	Slope	R	p	N	Mean weight
<i>All Samples:</i>					
Male chimpanzees	0.011	0.682	<0.001	24	30.6
Female chimpanzees	0.004	0.069	0.761	22	25.0
Male bonobos	0.005	0.307	0.285	14	21.6
Female bonobos	0.005	0.354	0.164	17	15.7
<i>Early Morning Samples Only:</i>					
Male chimpanzees	0.011	0.698	<0.001	21	27.2
Female chimpanzees	-0.004	0.135	0.560	21	24.1
Male bonobos	0.012	0.552	0.063	12	22.1
Female bonobos	0.004	0.268	0.400	12	18.3
<i>Mid-Day Samples Only:</i>					
Male chimpanzees	0.015	0.668	<0.001	24	30.6
Female chimpanzees	0.001	0.037	0.875	21	25.0
Male bonobos	-0.014	0.480	0.097	13	21.1
Female bonobos	-0.006	0.244	0.381	15	14.2

Table A4.5. List of significant effects from the Generalized Estimating Equations (GEE) analysis of log testosterone with individual, species, sex, and dental category as factors.

GEE analyses were performed with the overall log averages, as well as separately for the early morning and mid-day testosterone log averages. Dental category analyses could not be run separately by year due to small sample size. For the analyses of dental category, we also removed the three-way interaction from the model due to small sample size – thus each model contained all main effects and 2-way interactions. Only the significant effects and interactions are shown here, along with their respective Wald chi-square values and p-values, except for the species*dental category interaction in the mid-day samples which was significant only at a trend level but is reported here. We also report the full-model QIC along with the sample size used in that model.

	Wald chi-square	p	Whole model	
			QIC	N
<i>All Samples:</i>			7.982	155
Species	32.250	<0.001		
Sex	4.133	0.042		
Dental category	14.065	0.007		
Species*Dental category	23.232	<0.001		
<i>Early Morning Samples Only:</i>			7.784	119
Sex	4.461	0.035		
Dental category	38.459	<0.001		
Species*Dental category	12.664	0.013		
<i>Mid-Day Samples Only:</i>			10.812	134
Species	43.436	<0.001		
Species*Dental category	8.865	0.065		

Table A4.6. Significant post-hoc comparisons in the GEE analysis of log testosterone with individual, species, sex, and dental category as factors. P-values were adjusted for multiple comparisons using a Bonferroni correction. All possible pairwise comparisons were investigated for the species*dental category interaction. Only those where significant differences were present are shown here, with all other comparisons non-significant. The group in each comparison for which log testosterone values were lower is shown in the left column of the table.

<i>Species*dental category</i>		
Significant post-hoc comparisons		Bonferroni-corrected p value
Chimpanzee no permanent	Bonobo incisors only	0.029
Chimpanzee M1 only	Chimpanzee Canine/M3	0.043
Chimpanzee M1 only	Bonobo M1 only	0.027
Chimpanzee M1 only	Bonobo incisors only	0.001
Chimpanzee M1 only	Bonobo Canine/M3	0.017
Chimpanzee incisors only	Chimpanzee Canine/M3	0.001
Chimpanzee incisors only	Bonobo M1 only	0.001
Chimpanzee incisors only	Bonobo incisors only	<0.001
Chimpanzee incisors only	Bonobo M2 only	0.053
Chimpanzee incisors only	Bonobo Canine/M3	<0.001
Chimpanzee M2 only	Chimpanzee Canine/M3	0.001
Chimpanzee M2 only	Bonobo M1 only	<0.001
Chimpanzee M2 only	Bonobo incisors only	<0.001
Chimpanzee M2 only	Bonobo M2 only	0.005
Chimpanzee M2 only	Bonobo Canine/M3	<0.001

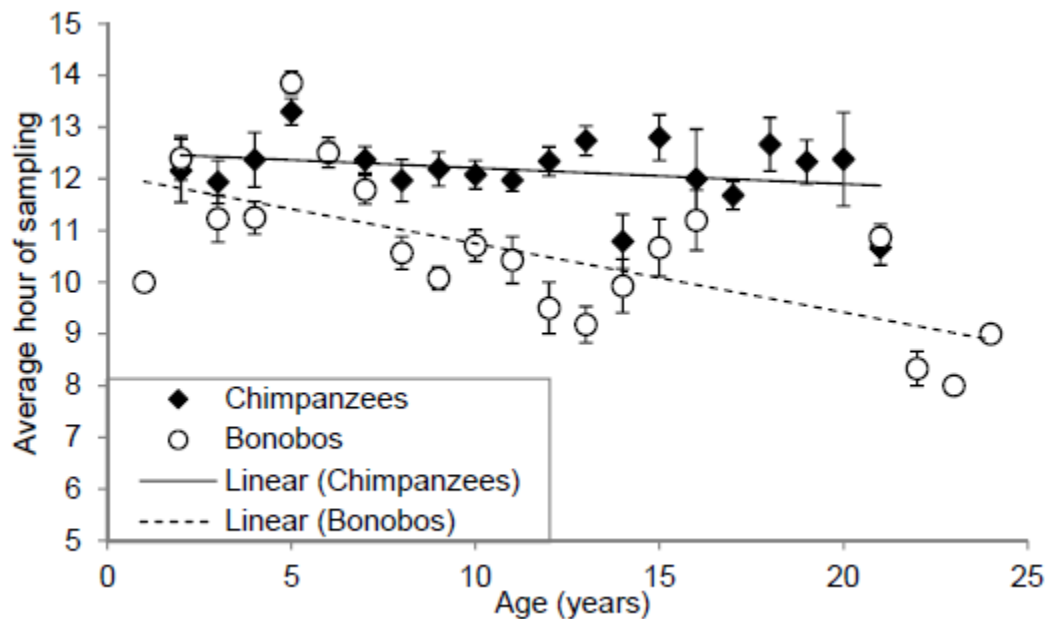


Figure A4.1. Average hour at which saliva samples were taken in chimpanzees and bonobos, according to age (in years). The average hour of sampling was computed for each age year sampled in each species, with bars denoting standard error for that age. Linear trend lines are shown for each species. Samples were taken equally throughout the day for chimpanzees of all ages (leading the average hour of sampling to be near 12:00). In contrast, samples were collected earlier in the day for bonobo adults than for bonobo juveniles, with a linear regression between age and hour of sample significant in bonobos but not chimpanzees. Procedures were taken in the statistical analysis to control for any potential bias introduced by this differential sampling across hours of the day.

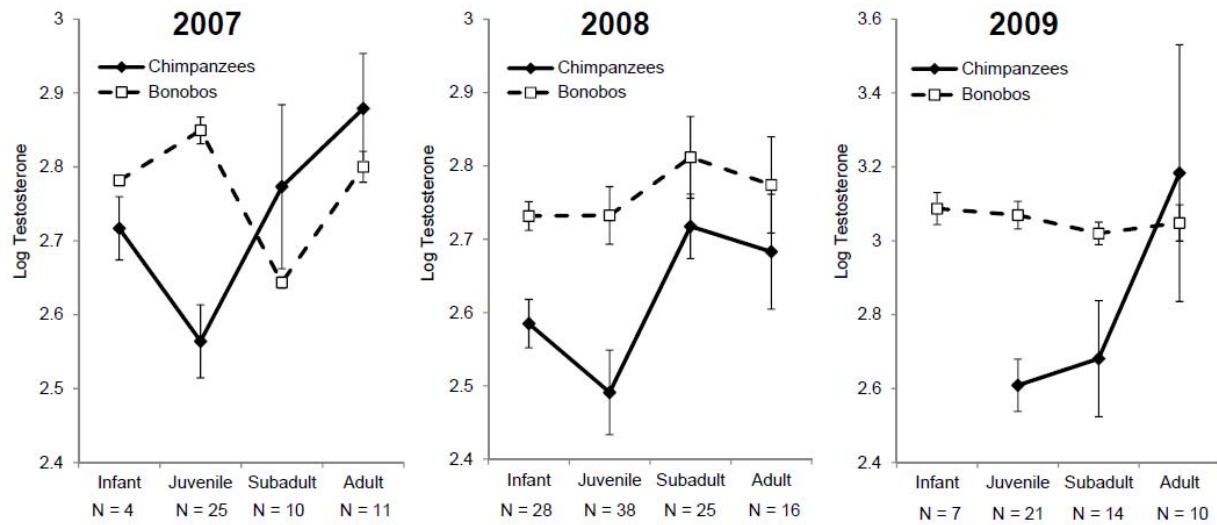


Figure A4.2. Log testosterone levels across development in chimpanzees and bonobos, separated by sample year. Average log testosterone values and sample sizes are shown for each age group in each year, excluding two outlier points from female chimpanzees. Note that the scale of the y-axis differs in the 2009 graph relative to the other years, since the chimpanzee adult samples and bonobo samples on the whole were higher in this year than in prior years. As mentioned in the manuscript, quality control values did not differ significantly between 2008 and 2009, suggesting that the differences in testosterone level between 2008 and 2009 did not reflect variation in assay characteristics.

Appendix 5:
Supplemental methods, analyses, tables, and figures for Chapter 5

SUPPLEMENTAL METHODS

Subjects

These experiments were carried out with bonobos at the Lola ya Bonobo sanctuary in Kinshasa, Democratic Republic of Congo, and chimpanzees at Tchimpounga chimpanzee sanctuary in Pointe Noire, Republic of Congo. At both sanctuaries, individuals live in social groups that have access to large areas of primary tropical forest (15-40 hectares) during the day and sleep in dormitories at night. Nearly all of these apes are wild-born orphans (except for those born at the sanctuary), arriving at the facility after being confiscated by local governments. Thus individuals have experienced some early life trauma in being separated from their mother and potentially kept in deprived conditions. However, our preliminary research suggests that sanctuary individuals show fewer behavioral indications of negative welfare than zoo apes and that the orphans show no cognitive impairment relative to mother-reared infants. Though the effects of early life experience on these individuals' hormonal pathways cannot be known, the two populations tested in this experiment are comparable because both have the same rearing histories and highly similar living environments at the sanctuaries.

Procedure

Dominance test

Prior to the food competitions, subjects participated in a test to assess relative dominance in a feeding context in a given pair. Similar to previous work (Hare, Melis, Woods, Hastings, & Wrangham, 2007; Melis, Hare, & Tomasello, 2006), the two individuals in a pair were brought into a test room and fed by caretakers 2 meters apart at the mesh wall of the room. As they were

being fed, the experimenter placed a large piece of food at the mesh equidistant between the two individuals. Whichever individual obtained this piece of food was scored as the “winner” of that trial. This procedure was repeated 8 times, and the individual that obtained the food on more of these 8 trials was scored as the dominant in that dyad. If both individuals obtained food on an equal number of trials, the dominant was assigned based on who obtained more food in the food competitions.

Dominants in this test also tended to monopolize more food in the food competitions. There was a significant relationship between the number of trials (out of 8) where an individual monopolized food in this test and the number of conditions (out of 3 test days) where that individual monopolized food in the food competitions (we performed a Kendall’s Tau ordinal by ordinal analysis for this comparison. $\text{Tau} = 0.33$, $p < 0.001$, $n = 328$).

Food competitions

Three food competition conditions varying the monopolizability of the food were utilized, following the procedures used in previous experiments (Hare, et al., 2007; Melis, et al., 2006). Each condition was presented on a separate day, with individuals receiving three trials of the same condition on a given day (resulting in 9 total food competition trials over the 3 conditions). The order in which the conditions were presented was counterbalanced across species, sex, and age. Each species was tested with its most preferred food to produce similar levels of motivation: bananas were used with the chimpanzees and green apples with the bonobos. The relative amount of food used in each condition was similar across the two species. The conditions were as follows:

Dispersed-divisible: two piles of food with 8 pieces in each were placed approximately 3 m apart inside the test room, with the piece size standardized as each banana was cut up

into 8 pieces (chimpanzees) and each apple was cut up into 32 pieces (bonobos) (thus a total of 16 food slices was placed in this condition, 8 in each of the two separate piles).

Clumped-divisible: one food pile was placed in the testing room, with 16 pieces (of the same size as the previous condition) all placed in this one pile.

Clumped-solid: one food pile was placed in the testing room, but rather than being small pieces there were simply two whole bananas (chimpanzees) or two quarters of an apple (with each banana or quarter-apple being the same amount as the 8 pieces placed in the other trials).

Subjects were placed in pairs prior to the pre-test saliva sample collection, and were kept in an adjacent room to the test room. They did not know the configuration in which the food was going to be presented on a given day when the pre-test saliva sample was taken, but did know their pairing and could see that food was present when the experimenters brought food into the dormitory building. After the pre-test saliva sample was taken, subjects witnessed the placement of the food from the adjacent room, and the pair was then videotaped for one minute to capture any behaviors exhibited in anticipation of the food competition. Subjects were then released into the test room, and their behavior in the test trial was videotaped as well. The trial was considered finished when both subjects finished eating. Preparation for the next trial began immediately after a given trial ended. After the third and last trial, upon finishing their food subjects waited in the testing room for 15 minutes for collection of saliva samples. During these 15 minutes they sat in their pairing and were not given any additional food, then the post-test saliva sample was taken while they were still in the room with the other individual.

In the solo condition subjects underwent the same procedure, being released into the test room for 3 trials, with a one-minute anticipation period before each trial after the food was

placed, except that subjects were alone in the test room during the pre-test saliva collection, the food presentations, the 15 minutes following the test, and the post-test sample collection.

SUPPLEMENTAL RESULTS

Cortisol

Controls for Anticipatory Results

As mentioned in the main manuscript, anticipatory effects in cortisol represented shifts relative to baseline values, rather than differences in basal cortisol levels between individuals (Figure A5.1). To control for whether there were differential anticipatory effects in cortisol based on the dominance status of the two individuals in the pair, we used the factor *domsub* which assigned dominants and subordinates based on the results of the dominance test described above. Performing analyses on the species-level for only males (since the main anticipatory patterns were not present in females), we ran a Generalized Linear Model (GLM) analysis on pre-test log cortisol with *individual* as a subject factor, and *outcome* and *domsub* as between-subject factors. This analysis revealed no effects or interactions in chimpanzees. In bonobo males, the main effect of *outcome* was still significant (Wald Chi-Square (1) = 10.75, $p = 0.001$), and there was a significant main effect of *domsub* (Wald Chi-Square (1) = 4.81, $p = 0.03$), in that subordinate bonobo males had higher cortisol than dominant bonobo males, but the interaction between *domsub* and *outcome* was not significant (Figure A5.2). This suggested that dominants and subordinates showed similar endocrine shifts in anticipation of the test.

Another factor that might have impacted cortisol shifts besides dominance status was the number of times the individuals in the pair had been tested. Since this test occurred over the course of multiple sessions, individuals may have reacted more in later sessions, after

experiencing sharing (or a lack thereof) with a given partner. Alternatively, individuals may have reacted more in earlier pairings due to the unfamiliarity of being paired in a dyad with the other individual. To assess this, we incorporated the factor *order*, denoting the first, second, or third test session. We performed a GLM with males only, split by species, of log pre-test cortisol with *individual*, *outcome*, and *order* as factors and found no effects or interactions in chimpanzees. In bonobos, the effect of *outcome* remained significant (Wald Chi-Square (1) = 16.214, $p < 0.001$), yet there was no significant effect of *order* and no interaction between *order* and *outcome* (Figure A5.3). Thus, bonobos showed equal changes in cortisol regardless of the number of times they had been tested with their partner.

A final factor that may have influenced anticipatory cortisol was the type of pair individuals were in – namely, whether they were competing against another male or a female. Unfortunately, this categorization was highly skewed in terms of the *outcome* variable – very few chimpanzee males paired with females shared the food equally. Thus we removed the *outcome* variable to assess whether *pairtype* alone predicted any differences in pre-test cortisol. A GLM of log pre-test cortisol with *individual* and *pairtype* as factors revealed no significant effect of *pairtype* in either species (Figure A5.4). As such, it did not appear that pre-test cortisol was altered simply by being partnered with another male versus a female.

Testosterone

Controls for anticipatory results

Similar to the cortisol results, the anticipatory shifts in testosterone represented departures from baseline levels (Figure A5.5). To examine whether dominance status influenced the effects in anticipatory testosterone, we again used the factor *domsub* as assigned by the results of the dominance test. Performing analyses on the species-level for only males, we ran a

GLM analysis on log pre-test testosterone using *individual*, *outcome*, and *domsub* as factors. In chimpanzee males, there was a significant effect of *domsub* (Wald Chi-Square (1) = 6.460, $p = 0.010$), in that subordinates had higher T than dominants, and the main effect of *outcome* was also significant, with T higher in both individuals when the dominant monopolized more food (Wald Chi-Square (1) = 12.004, $p = 0.001$). However, there was no interaction between *domsub* and *outcome*. Thus, this effect of outcome was equally present in both dominant and subordinate chimpanzees. In bonobo males, there were no significant effects or interactions (Figure A5.6). It is notable that dominant chimpanzee males had lower testosterone than the subordinate chimpanzee males, as this contradicts the typical finding that dominance is positively correlated with testosterone in captive and wild male chimpanzees (Anestis, 2006; Muller & Wrangham, 2004). This was likely because our dominance measure was only on the dyadic scale, with the larger group hierarchy potentially showing a stronger relationship with dominance than these potential dyadic overlaps. It was not possible to construct a group-level hierarchy because our subjects came from numerous different social groups living at the sanctuaries.

As with the cortisol analyses, we also examined the potential effects of order of the testing day on the anticipatory T values. To assess this, we performed a GLM analysis for males only, split by species, of pre-test log testosterone with *individual*, *outcome*, and *order* as factors. We found the predicted effect of *outcome* in chimpanzee males (Wald Chi-Square (1) = 4.621, $p = 0.03$), but no effect of *order*, nor any interaction between *order* and *outcome* in either species (Figure A5.7).

Finally, we wanted to assess whether pair type impacted males' pre-test testosterone. Again, we had to remove the outcome variable because this was skewed according to pair type, and simply examined whether there were any differences in males paired with other males versus

males paired with females. We performed a GLM analysis of log pre-test testosterone separately by species in males only, with *individual* and *pairtype* as factors, and found no effect of *pairtype* in males of either species (Figure A5.8). These results suggest that males' differential T based on outcome was not confounded merely by the sex of their partner and was instead sensitive to the identity of that partner.

SUPPLEMENTAL FIGURES

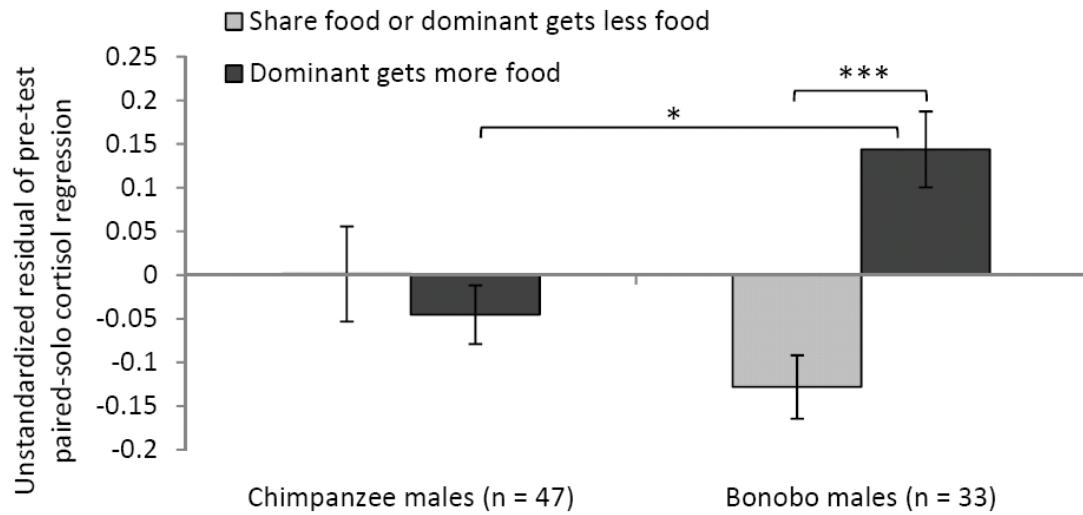


Figure A5.1. Pre-test cortisol values according to species and outcome, males only. These values are expressed as residuals of the log pre-test values relative to the log pre-solo (baseline) values. Bars denote standard error of the mean. * denotes $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

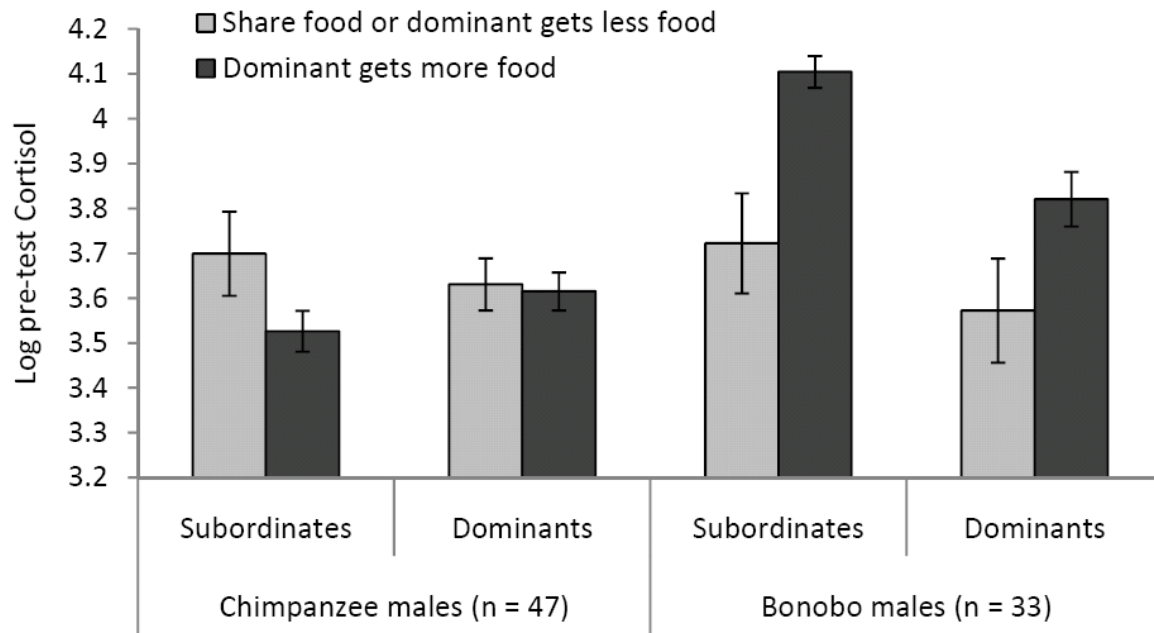


Figure A5.2. Pre-test log cortisol values according to species, outcome, and dominance status in males only. Bars denote standard error.

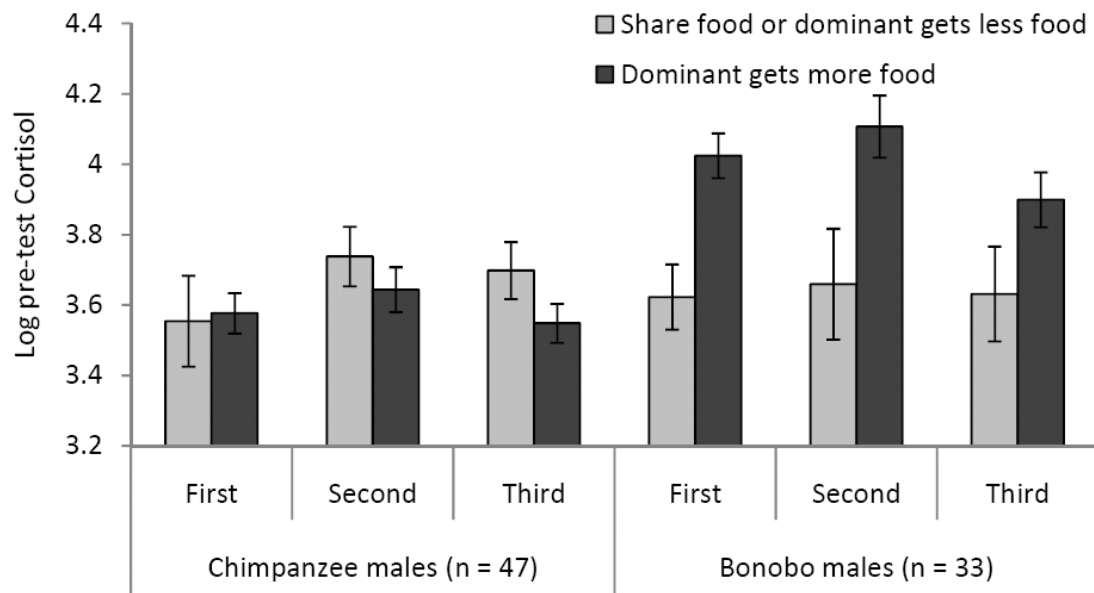


Figure A5.3. Pre-test log cortisol values according to species, outcome, and order in males only. Bars denote standard error.

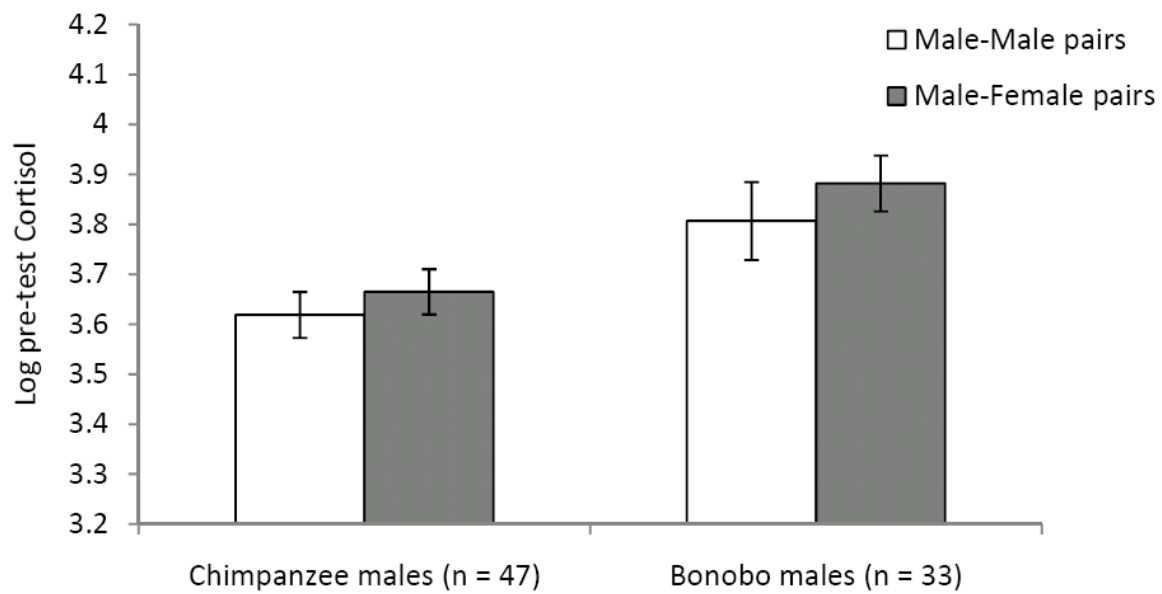


Figure A5.4. Pre-test log cortisol values according to species and pair type in males only. Bars denote standard error.

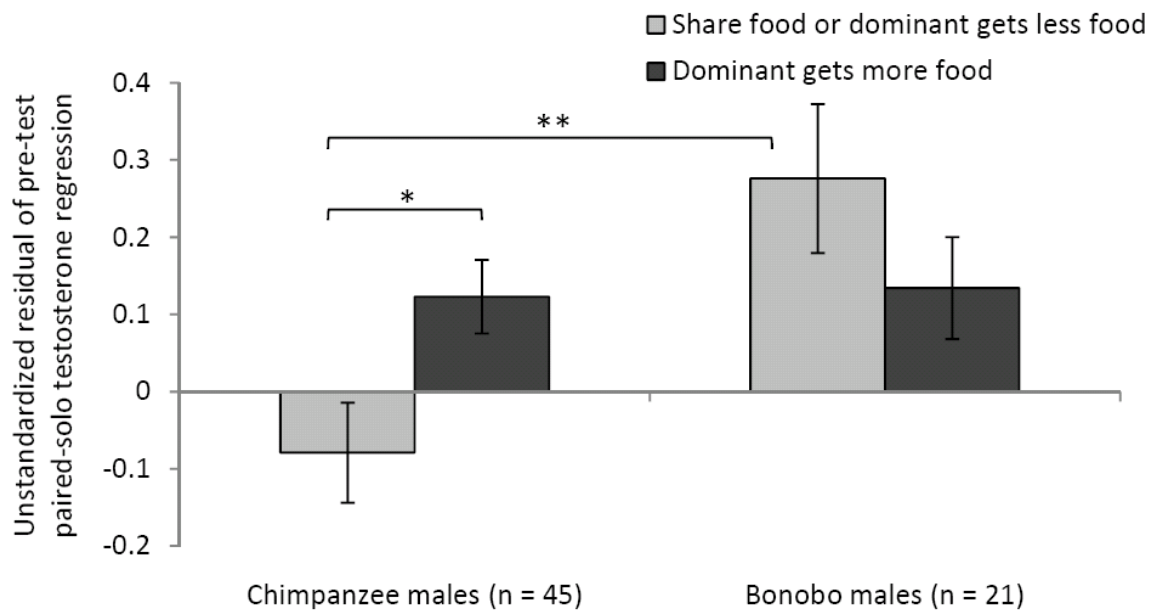


Figure A5.5. Pre-test testosterone values according to species and outcome, males only. These values are expressed as residuals of the pre-test values relative to the pre-solo values. Bars denote standard error of the mean. The bonobo sample size in this analysis is smaller because some bonobos completed the food competitions but did not produce enough saliva in the solo condition (baseline) to measure testosterone. * denotes $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

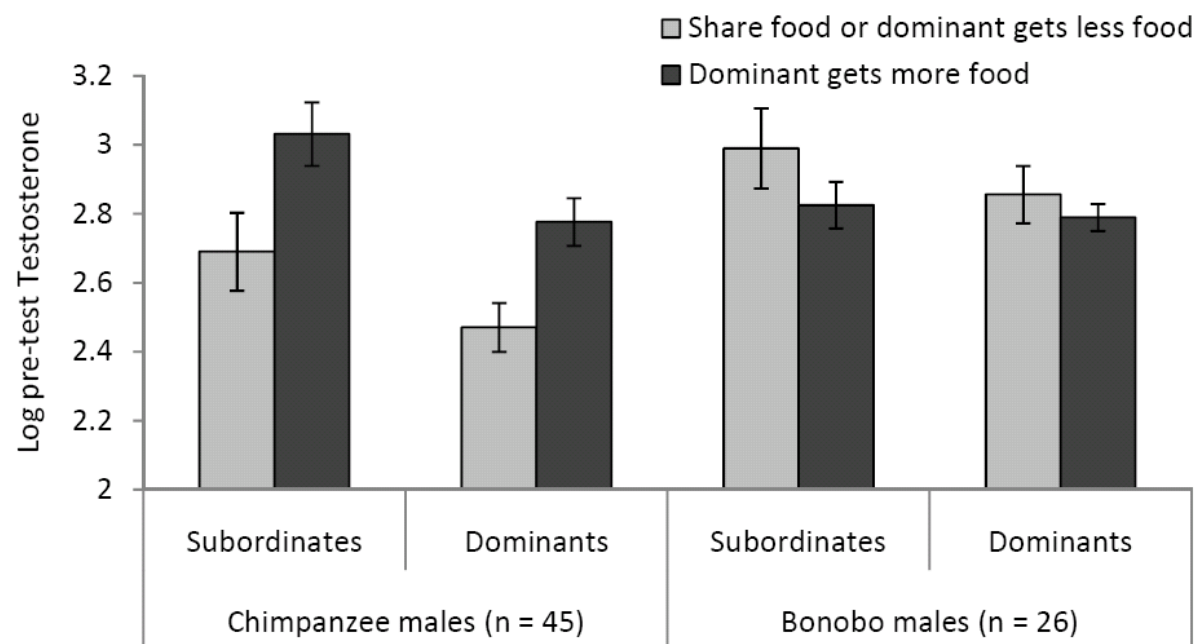


Figure A5.6. Pre-test log testosterone values according to species, dominance status and outcome, males only. Bars denote standard error.

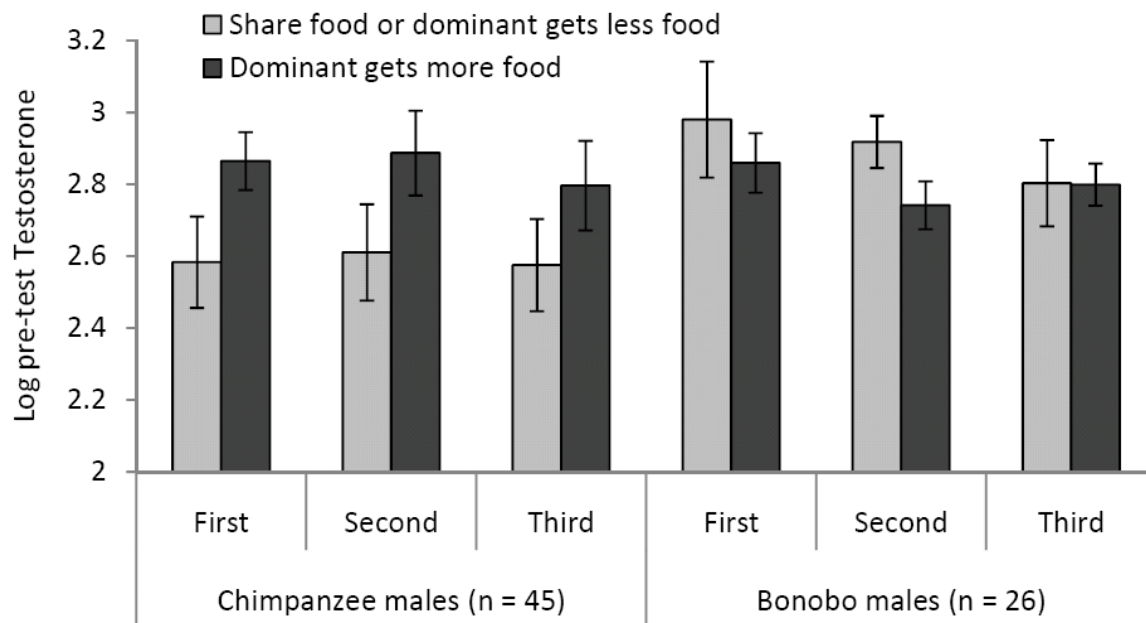


Figure A5.7. Pre-test log testosterone values according to species, outcome, and order in males only. Bars denote standard error.

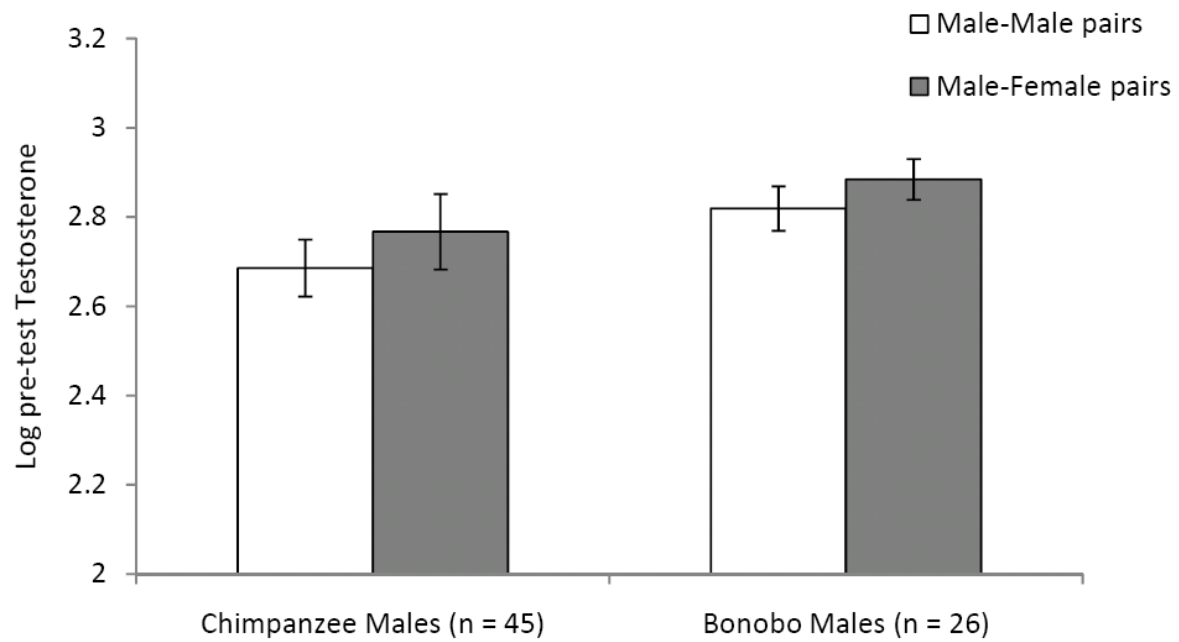


Figure A5.8. Pre-test testosterone values according to species and pairtype, males only. Bars denote standard error.

Appendix 5 Literature Cited.

- Anestis, S. (2006). Testosterone in Juvenile and Adolescent Male Chimpanzees (Pan troglodytes): Effects of Dominance Rank, Aggression, and Behavioral Style. *American Journal of Physical Anthropology*, 130(4), 536-545.
- Hare, B., Melis, A. P., Woods, V., Hastings, S., & Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology*, 17(7), 619-623.
- Melis, A. P., Hare, B., & Tomasello, M. (2006). Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Animal Behaviour*, 72, 275-286.
- Muller, M., & Wrangham, R. (2004). Dominance, aggression and testosterone in wild chimpanzees: a test of the 'challenge hypothesis'. *Animal Behaviour*, 67, 113-123.

Appendix 6:
Supplemental methods, analyses, tables, and figures for Chapter 6

SUPPLEMENTAL METHODS

Subjects

Chimpanzees and Bonobos (genus Pan):

Study 1 and Study 2 were performed with semi free-ranging bonobos living at Lola ya Bonobo in Kinshasa, Democratic Republic of Congo, and chimpanzees living at the Tchimpounga Chimpanzee Rehabilitation Center in Pointe Noire, Republic of Congo. We also tested 4 mother-reared *Pan* infants (3 chimpanzees and 1 bonobo) living at the Wolfgang Koehler Primate Research Center in Leipzig, Germany to supplement our control analyses of mother-reared individuals (see Supplemental Results). *Pan* subjects lived in mixed-age social groups of 10 to 20 individuals with multiple adult males and adult females. These groups have access to large areas of primary tropical forest (15-40 hectares) during the day. At night the apes sleep in dormitories (12 m²-160 m²) where they are fed, in addition to being fed in their enclosures throughout the day. Tests were performed in these dormitories.

Nearly all of the apes at Lola ya Bonobo and Tchimpounga are orphans of the bushmeat trade, having arrived at these sites as infants. Both chimpanzees and bonobos typically arrive at these sites at the age of 2 to 3 years. Upon arrival, an individual's age is estimated by the veterinarians on staff, who have over 10 years of experience working with infant chimpanzees and bonobos. Further, detailed weight and dental records are kept for each individual. Using this data, we estimated individual ages based on weight and dental data upon arrival, between arrival and testing, and at the time of testing, comparing these measures to the published age estimates for captive chimpanzees and bonobos in previous work (Grether & Yerkes, 1940; Lieberman,

Carlo, Ponce de Leon, & Zollikofer, 2007; Shea, 1983; Smith, Crummett, & Brandt, 1994).

With these multiple measures we were confident in these ages to the year. Several of our subjects were born at these sites, thus their exact ages were known. We utilized these individuals' weight and dental emergence patterns to validate the estimates for our orphan individuals.

Our preliminary data (see Supplemental Results) and previously published work demonstrate that apes at these sites represent valuable populations for non-invasive research, matching or exceeding measures of psychological health in comparison to other captive apes (Wobber & Hare, 2011). Upon arrival, individuals are cared for according to procedures specified by the Pan-African Sanctuary Alliance (PASA), of which these two sites are members (Cox, Rosen, Montgomery, & Seal, 2000). Infants are placed in quarantine from conspecifics for approximately 2-6 months but receive consistent human care during this time. Individuals are then placed in a peer group with other young apes where their interactions with humans are limited outside the feeding context. These young apes are ultimately integrated into groups that include individuals of older age. Our preliminary data suggests that baseline cortisol levels do not differ between mother-reared individuals and orphans living at these sites, providing an initial indication that early life experiences have not created chronic stress for these ape orphans. Further, we have found that the rate of aberrant, or stereotyped, behaviors among adults at these sites is lower than that seen in a zoo population (Wobber & Hare, 2011). Finally, as discussed in the main Results section of Chapter 6 and the Supplemental Results section below, our data indicates that orphans at the sanctuaries have comparable cognitive abilities to mother-reared individuals at these sites and mother-reared individuals living in a highly enriched captive environment at the Leipzig Zoo. This evidence is in line with our previous research finding that adults at these sites perform as well or even better on cognitive tests as apes in a zoo

environment (Hanus & Call, 2008; Vlammings, Hare, & Call, 2010). These results suggest that these apes are not affected by their early life experiences, at least in terms of their general social behavior and cognitive abilities.

Subjects were never food or water deprived for testing and all testing was voluntary. Individuals could indicate their desire to stop participating in the test by refusing to take food from the experimenter or sitting by the door of the dormitory that led to the forest enclosure, at which point they were released. Because we wanted individuals to be fully attentive and motivated when participating in the task, this led to individuals taking varying amounts of time to finish the test battery. Using a fixed number of test days might have forced individuals to continue participating after their attention had waned. Thus the number of days taken to complete the battery ranged from 7 to 10 for chimpanzees and bonobos. Individuals always completed a given task in only one testing session, with breaks between sessions only occurring between tasks.

Human children:

Human children were tested in the Department of Comparative and Developmental Psychology at the Max Planck Institute for Evolutionary Anthropology (MPI-EVA) in Leipzig, Germany. Similar to the *Pan* subjects, children had not participated in a similar cognitive study, and testing was stopped when individuals were not motivated. Correspondingly, children completed the test battery in 3 to 4 days, again with the order of tasks remaining consistent even if testing needed to be stopped because the child was unmotivated to participate. All children who began the test battery completed it.

Procedure

Subjects were tested individually, and received one testing session (approximately 30 minutes) per day. Individuals were tested with their mother or caregiver in the test room or nearby, to assure their comfort. These caregivers remained in the room but did not participate in the test in any way. Presence of a caregiver in the room (as opposed to simply being nearby) did not affect *Pan* infants' performance in either physical cognition (univariate GLM, $p > 0.9$) or social cognition (univariate GLM, $p > 0.1$) tasks. Mother-reared *Pan* infants were tested either in the same room as their mother, or in an adjacent room. If an infant was in the same room as its mother and she attempted to take part in the tasks, the mother was distracted by a second experimenter (e.g. by being given small pieces of food) a few meters away from the infant so that she could not influence her infant's performance. For human children, parents were instructed not to help their children in any way, and to look straight ahead if children looked to their faces for guidance. If parents forgot these instructions and cued the child in some way, that trial was thrown out and repeated at the end of the task. Parents were not informed regarding the objectives of the study until all testing was completed.

For apes, the majority of tasks were presented on a wooden table (80 cm x 39 cm) with a sliding platform (78 cm x 35 cm) made of either wood or plastic affixed to the top of the table (see Table A6.1). For children, the majority of tasks were presented on a Plexiglas platform (80 cm x 40 cm) that could be slid across a larger testing table. This platform had a transparent window (80 cm x 40 cm) attached to the front, with three holes (12 cm in diameter) in this window. Cups or other objects (such as tools) were placed on top of the sliding platform. Individuals could make a choice between two options by touching one of the cups. Choices were almost exclusively made by touching a cup with a finger, though touching the cup with one's

tongue, genitals, or a piece of straw was also considered a choice when certain *Pan* individuals preferred this method over using their finger to indicate. Individuals were familiarized with this choice procedure in warm-up trials (described below). As noted in Chapter 6, reliability coding was performed for any task where the results were not simply choice-based (Table 6.2).

Tasks were presented in a constant order for both chimpanzees and bonobos. Children received the tasks in a slightly different order from the *Pan* infants/juveniles, in line with previous work (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007) (Table A6.1). Certain individuals did not complete every task in the battery. Reasons for this included: individuals being scared of the test materials (one bonobo in tool properties and one chimpanzee in social referencing), failures in video recording (one chimpanzee and one child in gaze-following around barriers), declining to choose in one or more test trials (three children in goal understanding, nine children in social inhibition, ten children in tool properties, two children in numerical reasoning, and two children in gaze-following), or maternal interference with the test for object-based tasks with *Pan* infants (when an object needed to be passed into the test room, only mother-reared infants with their mother in an adjacent room could participate in the task, since mothers in the same room would frequently take the object away from the infant). Note that we chose not to include subjects who only partially completed a task, to ensure that comparisons of performance were based on equal numbers of trials. If a subject failed to complete only one task from a domain, its average was computed for the remaining tasks from that domain (e.g., the average was computed from 4 physical cognition tasks if an individual did not complete tool properties). If a subject did not participate in two or more tasks from a domain, its performance for that domain was not included (though that subject could still be represented for the other domain – this was the case for 1 human subject and 4 *Pan* subjects represented in

the social but not physical domain and 2 *Pan* subjects represented in the physical but not the social domain).

Tasks

Individuals in both experiments participated in a battery of 14 cognitive and 3 attentional/motivational control tasks. Each of these tasks was validated by numerous prior experiments (Carpenter, Nagell, & Tomasello, 1998; Herrmann, et al., 2007; Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2010; Herrmann, Hare, Call, & Tomasello, 2010; Tomasello & Carpenter, 2005; Wobber, Wrangham, & Hare, 2010). We performed only a subset of the tasks used by Herrmann et al (Herrmann, et al., 2007) because some of these tasks (for example, additional numbers) would have been too difficult for infants to complete and we wanted to ensure that subjects remained motivated across tasks. The procedures for children were identical to those utilized for the *Pan* subjects, except where noted below and in that toys served as the reward rather than food and that in certain tasks, no mesh separated the child from the experimenter. One experimenter (E1) administered all tasks to a given subject, though in certain tasks additional experimenters were required.

Warm-ups

To familiarize subjects with the standard object choice paradigm, where a subject chooses between one of several locations where a reward can be hidden, individuals received warm-up trials prior to participating in any object choice tasks. In these trials, a reward was visibly hidden under a cup in full view of the subject. Subjects needed to touch the correct cup in order to receive the reward. *Pan* subjects first received warm-up trials with only two cups, then with three cups; for human subjects, only three-cup warm-up trials were performed because humans' baseline performance on these trials was much higher. Thus, *Pan* subjects were given

additional experience with the reward-finding procedure to ensure that all individuals were equally competent before moving onto the object choice tasks in the test battery.

Social cognition

Intention-emulation

This test served to measure whether subjects could infer the experimenter's goal, having never seen her complete the goal but seeing only her failed attempts to achieve that goal (Bellagamba & Tomasello, 1999; Tomasello & Carpenter, 2005). On each trial, E1 picked up two pieces of PVC pipe, a tube with a curved end and a T-joint piece. In view of the subject, E1 attempted to put the two pieces of pipe together so that straight end of the tube fit into the joint of the T-shaped piece. These two specific joints were painted/colored red, to enhance the attribution of the experimenter's goal. Each time that E1 attempted to put the two pieces together, however, she failed to do so, making a disappointed grunting vocalization upon failure. E1 attempted to put the tubes together 3 times, and then handed the objects to the subject. The subject was allowed to manipulate the objects for 1 minute, with the dependent measure being whether the subject completed the experimenter's desired goal (putting the two tubes together), even though it had never seen this goal accomplished. Because *Pan* subjects rarely put the two tubes together after the experimenter's demonstration, we considered individuals (both *Pan* and human) successful if they put together any two ends of the tubes, even if they did not put together the experimenter's intended ends (those colored red). Subjects were not rewarded for success. Three trials of this test were presented, one trial per day for three subsequent test days. The percentage of trials where subjects succeeded in putting the two tubes together served as the dependent measure for this task. For children, the tubes were slightly larger and were a different color from those given to apes, but otherwise the procedure of the test was identical.

Social referencing

This task was designed to measure individuals' tendency to look to another individual's face as a cue to his or her intentions. We utilized a teasing task to measure this tendency (Phillips, Baron-Cohen, & Rutter, 1992; Tomasello & Carpenter, 2005). E1 sat in front of the subject, obtained its attention and then held out a small toy toward the subject as if initiating play. She played with the subject and the toy for a few seconds, to ensure that the subject was interested in, and not afraid of, the toy. E1 then teasingly pulled the toy away, holding the toy on the ground and looking straight ahead for 5 seconds. The dependent measure was whether the subject looked at E1's face during these 5 seconds, to infer E1's intentions for withdrawing the toy. This measure was coded from videotape. Because E1 held the toy on the ground, this allowed the best possible differentiation of whether the subject looked toward her face or looked downward toward the object. 3 trials were performed in sequence. The percentage of trials where subjects looked toward the experimenter's face was used as the measure of success in this task. For children, the procedure was identical except that E1 and the child sat at table, so instead of lowering the object to the ground upon teasing, E1 held it on the table. This again allowed differentiation of whether the child looked downward toward the object or upward at E1's face. The toys used for the children and *Pan* infants differed, but were chosen to be novel and interesting.

Gaze-following around barriers

This task served to measure whether individuals were able to follow an experimenter's gaze geometrically, requiring the subject to physically move around a barrier to follow this gaze rather than simply re-orienting his or her gaze direction (Moll & Tomasello, 2004; Tomasello & Carpenter, 2005). For this task, E1 sat across from the subject, called its name, and then looked

to a specific location behind a barrier for 30 seconds. During these 30 seconds, she alternated her gaze between the subject and the specific location, and made excited noises (as appropriate for the species). After 30 seconds, the subject was given a reward, regardless of the outcome of the trial. The subject was then given a short break before proceeding to the next trial. 3 trials per day were performed on two subsequent days (resulting in 6 total trials). Two barrier setups were utilized (one on the first day, and one on the second). In the first setup, E1 looked under the testing table. In order to follow her gaze, the subject needed to physically move its body and lower its head under the table. In the second setup, the table was placed on its side next to E1, 50 cm away and at a 45 degree angle. E1 looked behind this table, and in order for the subject to follow her gaze, it needed to move to the far side of the table or climb above the table and look down. The percentage of trials where subjects successfully moved around the barrier was recorded as the measure of success in this task.

For children, the procedures were identical except that there was no mesh between the experimenter and the child, so the child could move their body completely behind the barrier when following the experimenter's gaze. Because this was the case, performance was scored live for children (while it was scored from video for the *Pan* infants and juveniles). For the second setup, a larger barrier was used for children rather than turning the testing table on its side. Children could move entirely behind this barrier, making successful performance unambiguous. As such, a trial in this setup was ended if the child successfully moved around the barrier.

Social inhibition

This task was designed to measure individuals' abilities to inhibit their responses in a social situation where they begged for food from a human experimenter (Wobber, et al., 2010). Three human experimenters stood or sat shoulder to shoulder in front of the subject. To start

each trial, the reward-holding experimenters (which ones depended on the trial type) reached into a dish containing rewards that was placed in front of the middle experimenter, out of reach but in full view of the subject. The non reward-holding experimenters did not reach into the dish. All three experimenters then showed subjects their hands palm-up so that the subject could easily see who held the rewards. The experimenters then closed their hands and extended their arms toward the subject, offering their closed fists. Thus, though the subjects could have clearly seen who had reached toward the dish and taken rewards, they could not see the rewards at the time of choice.

Subjects received two familiarization trials so that they could become accustomed to receiving rewards from the experimenters. In these trials, all three experimenters took a reward and the subject was allowed to choose each of the three in sequence. Subsequently, subjects received four introductory trials, where one of the side experimenters and the middle experimenter took rewards. In these trials, the subject was then allowed to make up to two choices; if they chose the experimenter not holding the reward they were not able to continue choosing. Their performance was recorded as correct if they were able to correctly choose both experimenters holding rewards, and incorrect if they did not receive both rewards.

After these introductory trials, subjects were immediately presented with 12 test trials. In the test trials, only the two side experimenters held rewards; thus, the middle experimenter did not have a reward, not even reaching toward the reward dish. Again, the subject was allowed to choose twice if it chose correctly, but if it chose incorrectly on its first or second choice (choosing the middle experimenter) then it was not permitted to continue choosing. Correct choices were the trials on which the subject was able to bypass touching the hand of the middle experimenter. Percentage correct of these 12 test trials served as the dependent measure.

Gaze-following

This test, similar to gaze-following around barriers, measured individuals' abilities to track another's gaze. However, in this paradigm the dependent measure was only whether the subject re-oriented to follow the experimenter's line of sight – a more basic form of gaze following than the barrier paradigm. For this test, E1 sat across from the subject, called its name, and then looked upwards with her head and eyes for 10s. 10 trials were performed in sequence, to capture subjects' initial response and to measure whether they eventually habituated to E1's repeated gaze, as is typical for both children and adult chimpanzees (Tomasello, Hare, & Fogleman, 2001). E1 scored from the video whether subjects looked upwards on each trial. Percentage of trials where subjects did so served as the dependent measure.

Social learning

In this task, we observed whether subjects imitated the means demonstrated by an experimenter to achieve a goal (Call, Carpenter, & Tomasello, 2005; Gergely, Bekkering, & Kiraly, 2002; Herrmann, et al., 2007; Meltzoff, 1988; Tomasello & Carpenter, 2005). A reward was placed in the center of a 30 cm long transparent Plexiglas tube. The reward was trapped in the tube, such that a specific force had to be applied to remove it from the tube. E1 demonstrated for the subject how to get the reward out by banging one end of the Plexiglas tube on a table or the floor. After the successful demonstration, E1 handed an identical tube with a reward inside to the subject. Apes were given 2 minutes and children 1 minute to solve the task, in line with previous research (Herrmann, et al., 2007). E1 coded the dependent measure live, determining whether the subject obtained the reward using the demonstrated means (e.g. as opposed to using another means to obtain the reward or failing to obtain the reward altogether).

Point production

This task measured whether individuals would signal the location of a reward to an experimenter if that reward were out of her view, reflecting an understanding of the experimenter's attentional state and an ability to communicate gesturally (Herrmann, et al., 2007; Tomasello & Camaioni, 1997; Tomasello & Carpenter, 2005). For this task, two locations were clearly marked on the ground or on two small platforms, 2 meters apart. For *Pan* subjects, the procedure was as follows. Each trial began with a second experimenter (E2) entering the testing area and placing a piece of food on one of these locations, while E1 was outside the testing area. E1 then entered the testing area from the side opposite where the food was placed. She stood at the pre-determined location where there was no food (e.g. if the food were placed on the subject's left, the experimenter stood at the designated location on the right), 2 meters away from where the food was placed. She did not look toward the food when entering, looking straight ahead. She then positioned herself in one of two ways:

Away: E1 stood with her back turned toward the food, facing the opposite direction. To obtain the food, the subject needed to approach the experimenter so that she could see the subject in the direction she was facing. The subject then needed to direct the experimenter toward the food by gesturing. An acceptable gesture was extending an arm or a piece of straw toward the food.

Towards: In this case the experimenter stood facing the food, though she looked straight ahead so she could not see the food on the ground. To obtain the food, the subject needed to gesture towards it while looking at the experimenter, thus indicating a desire to communicate the location of the food to her (rather than simply reflecting the subject's desire to obtain the food by reaching).

4 trials were performed of this task (2 trials of the first condition and then 2 trials of the second condition). The percentage of trials where subjects successfully showed the experimenter the location of the food (whether through simply gesturing in the Towards condition or obtaining her attention and then gesturing in the Away condition) served as the dependent measure for this task.

There were three minor changes in the procedure with children, as performed in previous research (Herrmann, et al., 2007). First, E1 hid the reward and E2 served as the individual for whom the subject had to produce a communicative gesture (a role reversal in comparison to the ape procedure). Second, the reward was hidden inside of one of two boxes, rather than on the ground. These boxes were placed at a height where children could not obtain the toy themselves and needed the experimenter's help to obtain the item. Finally, the context changed somewhat: to start each trial, E1 and the child played together with a toy that consisted of two separate parts (such as a helicopter with a detachable rotating propeller). E1 then took one part of the object (in this example, the helicopter), placed it in one of the boxes, and left the room. This left the child with the remaining part of the object (the propeller), but he or she needed the additional object to play with the toy. E2 then entered the room and stood in one of the pre-determined positions, as described above. To obtain the object, the child needed to direct E2 to the correct location. If the child did so successfully, he or she was rewarded by being allowed to play with the 2 parts of the toy until the next trial.

Goal understanding

In this task, subjects needed to interpret an experimenter's intentions and goals in order to find food in an object choice paradigm (Braeuer, Kaminski, Riedel, Call, & Tomasello, 2006; Herrmann, et al., 2007). E1 hid a reward in one of two metal containers with lids for this task;

subjects were familiarized with these containers prior to the test with warm-up trials. In each warm-up trial, E1 placed the two cups, with their lids off and sitting next to them, in the center of the table. She placed food in one of the cups in view of the subject, covered both cups with their lids, and moved them to the corners of the platform. Subjects were allowed to choose one cup; if they chose the correct cup they were rewarded and if they did not they were shown the location of the reward. If the subject did not choose correctly in both of the two warm-up trials, the experimenter continued to present warm-up trials until the subject had chosen correctly at least once on each side.

For the test trials, E2 sat behind the testing table while E1 stood next to the table. To start each trial, E1 placed the cups into the center of the table after showing the subject that they were empty. She then showed the subject the reward, placed a plastic occluder (80 cm x 40 cm) in front of the table to block the subject's view of the cups, and hid the reward in one of the cups. In this way the subject knew the reward was hidden but did not know in which container. The subject also knew that E2 had witnessed the hiding of the reward. E1 then placed the lids on the cups, slid them to the corners of the table closest to the subject, and removed the occluder. E1 stepped away from the table and E2 then performed one of the following actions:

Trying: E2 tried to unscrew the lid of the container where the reward was hidden but did so unsuccessfully, making a vocalization to denote that she was struggling to open the container. She attempted to open the container three times then returned to her original position, staring straight ahead.

Reaching: With the apes, E2 sat behind bars so that she could not reach the containers. For children, E2 sat far enough from the table that she could not reach the containers. In both cases, E2 reached toward the container where the reward was hidden, extending her arm and

hand and making a vocalization to denote that she was attempting to reach the container. She reached toward the container three times, and left her arm extended toward the container while the subject made its choice.

After the demonstration by E2, E1 pushed the sliding platform forward, enabling the subject to make a choice. If the subject chose the correct container, it was given the reward, while if it chose the wrong container it was shown the location of the reward but not given the food. 6 trials (3 of the trying condition followed by 3 of the reaching condition) were performed per day on two subsequent test days, resulting in 12 total trials. Percentage correct out of these 12 trials served as the dependent measure.

Reputation

This task measured whether subjects could track other individuals' behavior and base decisions on this information (Hamlin, Wynn, & Bloom, 2007; Melis, Hare, & Tomasello, 2006). Three experimenters participated in this task. E1 always served as the “nice” or “mean” experimenter, with this counterbalanced across subjects.

To begin the task, the “nice” experimenter attempted to give a reward (a peanut or toy) to the “neutral” experimenter. When the nice experimenter did so, the “mean” experimenter stole the reward away from the nice experimenter and either ate it or simulated eating it, making noises to show her pleasure at stealing the food (with ape subjects). The nice experimenter then made angry vocalizations toward the mean experimenter, gently hitting and pushing her. The neutral experimenter also made these vocalizations, showing his or her disappointment at not receiving the reward from the nice experimenter. The nice experimenter attempted to give a reward to the neutral experimenter 10 times (for apes) or 3 times (for children). The nice and mean experimenters then left the test area and returned with an equal number of rewards (10

peanuts for the apes and one toy for children) in their hands. For the ape subjects, the experimenters stood or crouched, 2 m apart, at the mesh and the neutral experimenter gave the subject a piece of food at the mesh equidistant between the two experimenters, to center the subject so that its location could not bias its choice. For children, the experimenters sat on opposite ends of the testing table, equidistant from the child. For both apes and children, the two experimenters simultaneously extended their arms toward the subject, offering the reward(s) in their palms face up. They held this position for 20 seconds, allowing the subject to approach either or both of them. The subject was not permitted to take a reward from either individual, so that it was not reinforced for choosing either experimenter.

For apes, 4 trials (each with 10 demonstrations) were performed. For children, only 2 trials were performed (piloting showed that children became uninterested across trials given the low level of experimenter-child interaction in the test procedure, and so only this many trials could be performed while maintaining high levels of motivation). The experimenters coded live which individual the subject approached (apes) or reached towards (children) first. Subjects did not approach/reach within the 20-second interval for all trials. Thus the dependent measure for this task was the percentage of trials *where subjects made a choice* that they chose the nice experimenter. Thus if subjects only made a choice on 3 of the 4 trials, choosing the nice experimenter once, their percentage correct was scored as 0.33 rather than 0.25. This procedure controlled for decreasing interest in the task, as several ape subjects were frustrated that their begging behavior did not yield any reward and refrained from choosing on later trials. However subjects in this instance did approach to take the reward from the neutral experimenter, indicating that they were still attentive to the situation but simply did not choose to approach the nice or mean experimenter.

Physical cognition

Object permanence

This task measured subjects' knowledge of object permanence, or understanding that objects continue to exist after they disappear from view, with a Stage 6 invisible displacement task (Barth & Call, 2006; Herrmann, et al., 2007; Piaget, 1952). To begin the test, E1 placed three opaque plastic square containers in a row face down on the sliding platform. Before each trial, she would flip the containers up so that the subject could see they were empty then flip them down in sequence to start the trial. E1 then placed an additional small opaque cup on the far left side of the platform and placed a reward under it in view of the subject. She moved this cup under one (or more) of the larger square containers in one of the following patterns:

Single displacement: E1 moved the small cup under only one larger container, leaving the reward underneath this larger container.

Double adjacent displacement: E1 moved the small cup under two larger containers that were adjacent to one another (the left and middle containers, or right and middle containers), leaving the reward underneath one of the larger containers.

Double nonadjacent displacement: E1 moved the small cup under the two larger containers that were not adjacent to one another (the left and right containers), leaving the reward underneath one of them.

After finishing the movement of the small cup, the experimenter showed the subject that this cup was empty then slid the platform forward so that the subject could choose. If subjects understood that the reward moved under the small cup and was left under one of the larger containers, they should have chosen one of the larger containers where the small cup moved. In the single displacement condition, individuals could only make one choice. In the double

displacement conditions there were two potential locations of the reward, thus individuals were allowed to make two choices if necessary. If, on their first choice, they chose a larger container where the small cup had not traveled, they were not permitted to make a second choice. Individuals were considered correct if they chose the correct location on their first choice, or if they chose two potentially correct locations, obtaining the reward on the second choice. Six total trials were performed in sequence, 2 of each type, with percentage correct out of these 6 trials serving as the dependent measure.

Transposition

This task also measured individuals' abilities to track hidden rewards, this time with the reward location being moved in full view of the subject (Barth & Call, 2006; Herrmann, et al., 2007; Sophian, 1984). E1 began by placing three cups in a row on the platform. She hid the reward under one of these cups in view of the subject then visibly swapped the positions of the cups. She performed the following three types of swaps:

Single transposition: E1 swapped the location of the baited cup with one empty cup.

Double unbaited transposition: E1 swapped the location of the baited cup with one empty cup, and then swapped the position of the two empty cups.

Double baited transposition: E1 swapped the location of the baited cup with one empty cup, and then swapped the location of the baited cup with the other empty cup.

After finishing the cup swaps, E1 pushed the sliding platform forward and the subject was allowed to make its choice. It received the reward if it chose the baited cup first. Six total trials were performed in sequence, 2 of each condition, with percentage correct out of these 6 trials serving as the dependent measure.

Relative Number

This task measured individuals' ability to discriminate between varying quantities of a reward, with individuals successful if they were able to choose the option providing the larger reward (Hanus & Call, 2007; Herrmann, et al., 2007; Tomonaga, 2008). For this task E1 used two white trays with lids. E1 presented subjects with two warm-up trials to familiarize them with these trays. In these trials, E1 placed the two trays, uncovered, in the center of the testing table, and then placed a piece of food in one of the trays in view of the subject. She then covered the trays and moved them to the two corners of the table, allowing the subject to choose one of the two. If the subject did not choose correctly in both of these trials, the experimenter continued to present trials until the subject had chosen correctly at least once on each side.

In the test trials, E1 placed the two dishes in the center of the table and occluded the subject's view of the table (using the plastic occluder described above). She baited each dish with a certain number of reward items, placed the lids on top of the two containers and removed the occluder. When the subject was attentive, she removed the lids of the two containers simultaneously, and paused for a few seconds to allow the subject to view the contents of the containers. She then pushed the containers to the two corners of the table and slid the platform forward so that the subject could make a choice. Subjects were given the contents of the dish that they chose. To be credited with a correct response subjects had to choose the container with the larger quantity of the reward. Six trials were performed in sequence, one each of the following numerical comparisons, presented in the following order: 1:0, 6:3, 6:2, 3:2, 2:1, 4:1. Percentage correct out of these 6 trials served as the dependent measure.

For children, the trays for this task were red, rather than white, and were slightly larger in size than the trays used with *Pan* subjects.

Tool use

In this task, subjects needed to use a tool to obtain an out-of-reach reward (Herrmann, et al., 2007). E1 placed a reward on the table far enough away that the subject could not reach it with its hands. She then placed a wooden stick (30 cm in length) on the table next to the reward, and moved away from the testing table. Ape subjects had 2 minutes and children had 1 minute to use the stick to obtain the out-of-reach reward (Herrmann, et al., 2007). If the subject correctly obtained the reward by a means other than using the stick (e.g. shaking the table) the trial was repeated. If the subject broke the stick it was given, it was provided with a second; if it broke the second stick the test was considered finished with failure as the result.

Tool properties

To test whether subjects understood the functional properties of potential tools we presented them with an object choice task where they needed to choose between a functional and non-functional tool, each of which was associated with a food reward (Hauser, 1997; Herrmann, et al., 2007; Herrmann, Wobber, & Call, 2008). For each trial, E1 placed two tools on the table behind the occluder, then removed the occluder, paused to allow the subject to view the tools, and pushed the sliding platform forward so the subject could make its choice. Two conditions were presented:

Side: Two equally-sized pieces of burlap (20 cm x 15 cm) were placed on the two sides of the table. On one side, the reward was placed on top of the piece of burlap, while on the other side the reward was placed next to the burlap. The reward was only obtainable by pulling the burlap piece where the reward was on top.

Ripped: E1 placed one rectangular piece of burlap (20 cm x 15 cm) on one side of the table and on the other side she placed two smaller pieces of burlap (11 cm x 15 cm and 8 cm x 15

cm). These two smaller pieces formed a rectangle the same size as the larger burlap rectangle (11 cm long and 8 cm long with a 1 cm gap between them), so that size preferences could not influence a subject's choice. E1 placed one reward on the larger piece of burlap and one on the ripped piece of burlap that was further from the subject. If the subject pulled the piece of ripped burlap that was closest to its hand, the food reward would remain sitting on the further piece of ripped burlap since these two pieces were not connected.

When the subject touched one of these tools, the other one was removed and subjects were allowed to successfully obtain the reward or experience the non-functionality of the incorrect option. A subject was considered correct if it touched the functional tool first. Six trials were performed per day, 3 of the side condition followed by 3 of the ripped condition, for two subsequent days. This resulted in 12 total trials, and percentage correct over these 12 trials served as the dependent measure. Colored cloth was used in the place of burlap with children.

Attentional/motivational controls

Three control tasks were conducted to ensure that any species or age patterns reflected differences in subjects' cognitive abilities rather than differences in their motivation to complete the tasks. These tasks are described below. For both apes and children, the risk box task was presented prior to any cognitive tasks and the unsolvable task was presented after gaze-following around barriers. The novel objects task was designed to create an unfamiliar situation for the subject, thus this task was always administered at the beginning of a test session. In apes, this task was presented in the test session following unsolvable task and in children, this task was presented in the session following reputation.

Risk box

This task served to measure subjects' interest in novelty, or general willingness to take risks in an unfamiliar situation (Kagan & Snidman, 2004). This task was presented prior to all of the other tasks, making it the first interaction that subjects had with the experimenter and the general test situation. For this task, the experimenter presented the subject with a wooden box with a hole on one side, positioned such that it was dark inside the box and thus potentially risky to place one's hand inside the hole. To begin the test, the experimenter placed the box on the table within the subject's reach and gave it 30 seconds to manipulate the box. After 30 seconds, the experimenter placed a reward inside the hole in the front of the box in view of the subject and presented the box to the subject, again for 30 seconds. The dependent measure for this task was whether or not the subject reached into the hole in the box in the reward condition.

Unsolvable task

This task provided an index of how interested subjects were in obtaining food and how determined they were to independently solve a problem (Miklosi et al., 2003). To begin the task, the experimenter placed a reward inside a clear plastic box with a lid, placing the box upside-down on top of its lid in the center of the testing table. To obtain the reward subjects simply needed to lift up the box, leaving the reward sitting on the lid. Subjects were presented with this same situation three times. For the fourth presentation, the experimenter sealed the box, out of the subject's view, by pushing a button on the lid that vacuum sealed the lid closed. Thus while the problem appeared to still be easily solvable, it was now unsolvable (assuming that subjects could not detect that pushing the button would re-open the lid). Subjects were allowed to manipulate the box for one minute, with the dependent measure being the number of seconds that they actively manipulated the box (converted to a percentage as the number of seconds out of 60

seconds total). If subjects successfully opened the box in the unsolvable portion of the task, the task was concluded and their performance for this task was not included in the analysis.

Novel objects

This task measured subjects' reactivity to novel objects designed to draw interest (by moving or making noise). We utilized this task to quantify subjects' reactions on a shy-bold continuum, in addition to their general interest in objects that might pertain to the test (Herrmann, et al., 2007; Kagan & Snidman, 2004). For this task, the experimenter sat behind the testing table and placed an unfamiliar object in the center of the table. She left this object sitting on the table for 30 seconds, and then moved the object from one side of the table to the other for the next 30 seconds. Afterward, she removed this first object and replaced it with another unfamiliar object that was more interesting than the first. Again, she left this object sitting still for 30 seconds, and then for the next 30 seconds the object moved (these objects were either wind-up toys or electric cars that could move independently). The camera was positioned such that it captured a pre-specified area of a certain size (140 cm x 110 cm). Thus in coding, the experimenter could measure subjects' general interest in the object by recording how many second subjects spent in this area (converted to a percentage as the number of seconds out of 120 seconds total).

Analysis

Emergence criteria

For the emergence criteria, we divided the 14 cognitive tasks into two categories: those that were "forced choice," where subjects needed to make some response in order for the task to continue (where non-responses would be considered aborted trials), and those where "no choice" was an acceptable response (Table 6.3). In the forced choice tasks where there were enough trials

that individuals could perform successfully above chance levels (according to a binomial test), this level of performance served as the emergence criterion (social inhibition, goal understanding, and tool properties). In the forced choice tasks where there were too few trials for this to be the case (object permanence, transposition, number, reputation), other emergence parameters were designated. Finally, in tasks where “no choice” was an acceptable response, simply performing the desired behavior once or more was considered sufficient for the emergence criterion (Table 6.3).

SUPPLEMENTAL RESULTS

Study 1

Task by task rate of development analysis

In addition to the analyses performed on the domain level, we also performed analyses individually for each cognitive task. The results for each task are described below.

A GLM analysis was performed for each task with genus (*Homo* vs. *Pan*) and age group (2 years, 3 years, 4 years) as factors. Post-hoc tests were performed to determine whether there were genus differences at each age, and whether there were significant increases with age in each genus (using a Bonferroni correction for multiple comparisons). For 2 tasks (social learning and tool use) performance was not continuous, consisting of merely success or failure. For these tasks, chi-square tests were performed to compare results between genera overall and at each age group, and Kendall’s Tau ordinal by ordinal comparisons were performed to assess whether there were increases with age in each genus (again performing a Bonferroni correction to correct for these multiple comparisons). The results of these analyses are summarized in Table A6.2.

Social cognition

Intention-emulation. Human children outperformed *Pan* infants overall ($F(1,84) = 145.24, p < 0.001$) and separately at every age (2 years: $p < 0.05$, 3 years: $p < 0.001$, 4 years $p < 0.001$). There was also a significant interaction between genus and age group ($F(2,84) = 4.70, p = 0.01$). This derived from an improvement in performance with age among humans ($F(1,47) = 14.79, p < 0.001$), with 3- and 4-year olds outperforming 2-year olds (p values < 0.01), while there was no improvement with age in *Pan* infants.

Social referencing. Children outperformed *Pan* infants overall ($F(1,90) = 121.39, p < 0.001$) and separately at every age (2 years: $p < 0.01$, 3 years: $p < 0.001$, 4 years $p < 0.001$). There was no significant improvement with age in either genus.

Gaze-following around barriers. Children outperformed *Pan* infants overall ($F(1,93) = 85.00, p < 0.001$) and separately at every age (2 years: $p < 0.05$, 3 years: $p < 0.001$, 4 years $p < 0.001$). Children improved slightly in their performance with age ($F(1,45) = 6.43, p < 0.05$), though there were no differences between age groups in children that were significant after multiple comparisons. There was no improvement with age in *Pan* infants.

Social inhibition. Children outperformed *Pan* infants overall ($F(1,84) = 32.35, p < 0.001$) and there was a significant interaction between genus and age group ($F(2,84) = 9.80, p < 0.001$). Humans outperformed *Pan* only at 3 and 4 years (3 years: $p < 0.001$, 4 years $p < 0.001$). Accordingly, there was a significant improvement with age in humans ($F(1,37) = 46.62, p < 0.001$), with 3- and 4-year olds outperforming 2-year olds (p values < 0.001), while there was no improvement with age in *Pan* infants.

Gaze-following. Children outperformed *Pan* infants overall ($F(1,94) = 27.28, p < 0.001$), and separately only at 2 years ($p < 0.01$), not at 3 or 4 years. There was no significant improvement with age in either genus.

Social learning. Children outperformed *Pan* infants overall ($\chi^2(1, n = 90) = 50.23, p < 0.001$), and at 3 and 4 years (3 years: $p < 0.001$, 4 years $p < 0.001$). Humans improved significantly with age (Kendall's tau = 0.59, $p < 0.001$), while *Pan* infants did not.

Point production. Children outperformed *Pan* infants overall ($F(1,92) = 33.78, p < 0.001$) and separately only at 4 years (4 years $p < 0.001$). There was also a significant interaction between genus and age group ($F(2,92) = 5.50, p < 0.01$), but while children improved slightly in their performance with age neither genus showed a significant effect of age after correcting for multiple comparisons.

Goal understanding. Children outperformed *Pan* infants overall ($F(1,91) = 68.85, p < 0.001$) and separately at 3 and 4 years (3 years: $p < 0.001$, 4 years: $p < 0.001$). There was no significant improvement with age in either genus.

Reputation. Children did not significantly outperform *Pan* infants in this task overall or at any age, however there was a significant interaction between genus and age group ($F(2,94) = 5.40, p < 0.01$). *Pan* infants slightly improved in performance with age, but neither genus showed a significant effect of age after correcting for multiple comparisons.

Physical cognition

Object permanence. Children significantly outperformed *Pan* infants overall ($F(1,95) = 68.01, p < 0.001$), and separately at every age group (2 years: $p < 0.01$, 3 years: $p < 0.001$, 4 years: $p < 0.001$). Humans improved significantly with age ($F(2,45) = 12.41, p < 0.001$), with 4-year-olds outperforming 2-year-olds ($p < 0.001$), while *Pan* infants did not improve with age.

Transposition. Children significantly outperformed *Pan* infants overall ($F(1,95) = 20.09$, $p < 0.001$), and separately at 4 years ($p = 0.01$). Children improved significantly with age ($F(2,47) = 19.01$, $p < 0.001$), with 4-year-olds outperforming 2-year-olds ($p < 0.001$), while *Pan* infants did not improve with age.

Number. Children significantly outperformed *Pan* infants overall ($F(1,94) = 20.09$, $p < 0.001$), and there was a significant interaction between genus and age group ($F(2,94) = 3.87$, $p = 0.03$). Children significantly outperformed *Pan* infants only at 4 years ($p < 0.001$). Children improved significantly with age ($F(2,45) = 12.96$, $p < 0.001$), with 4-year-olds outperforming 2-year-olds ($p < 0.001$), while *Pan* infants did not improve with age.

Tool use. Children did not perform significantly better than *Pan* infants overall (χ^2 test, $p > 0.4$), only doing so at 4 years of age ($p < 0.05$). Accordingly, humans improved significantly with age (Kendall's tau = 0.36, $p < 0.05$), while *Pan* infants did not.

Tool properties. Children significantly outperformed *Pan* infants overall ($F(1,83) = 108.48$, $p < 0.001$), and separately at all ages (2 years: $p < 0.01$, 3 years: $p < 0.001$, 4 years: $p < 0.001$). Children improved significantly with age ($F(2,37) = 10.63$, $p < 0.01$), with 3- and 4-year-olds outperforming 2-year-olds (p values < 0.05), while *Pan* infants did not improve with age.

Attentional/Motivational Controls

Risk box. Children did not perform significantly differently than *Pan* infants overall (χ^2 test, $p > 0.7$), reaching into the box significantly more only at 3 years of age ($p < 0.04$). There was no significant change in performance with age in either genus.

Unsolvable task. Children spent significantly longer attempting to solve the unsolvable task relative to *Pan* infants overall ($F(1,89) = 4.74$, $p = 0.03$), though these differences were not

significant in any age category and there were no significant changes in performance with age in either genus.

Novel objects. Children spent significantly less time in proximity to the novel objects than *Pan* infants overall ($F(1,92) = 8.12, p < 0.01$), though these differences were not significant in any age category and there were no significant changes in performance with age in either genus.

Study 2

Task by task rate of development analysis

For Study 2, a repeated measures ANOVA was performed for each task, separately for infants/juveniles and for adults, with year as a factor. Again, for 2 tasks (social learning and tool use), performance was not continuous thus we performed a McNemar's test to check for differences in performance between years. The results are summarized in Table A6.3.

Social cognition

Intention-emulation. *Pan* infants/juveniles improved in their performance across years in this task ($F(2,34) = 4.86, p = 0.01$). Adults did not participate in this task (as described in the Study 2 methods in Chapter 6).

Social referencing. There was no change in performance in this task across years in either *Pan* infants/juveniles ($p > 0.5$) or *Pan* adults ($p > 0.9$).

Gaze-following around barriers. There was no change in performance in this task across years in either *Pan* infants/juveniles ($p > 0.5$) or *Pan* adults ($p > 0.6$).

Social inhibition. There was no change in performance in this task across years in either *Pan* infants/juveniles ($p > 0.9$) or *Pan* adults ($p > 0.2$).

Gaze-following. There was a trend for performance to improve across test years in *Pan* infants/juveniles ($F(2,42) = 2.94$, $p = 0.06$), but no change in performance across years in *Pan* adults ($p > 0.4$).

Social learning. McNemar's tests revealed a significant improvement in social learning performance among *Pan* infants/juveniles from 2008 to 2010 ($p = 0.04$), but no differences between other task years. Adults did not participate in this task.

Point production. There was no change in performance in this task across years in either *Pan* infants/juveniles ($p > 0.1$) or *Pan* adults ($p > 0.7$).

Goal understanding. *Pan* infants/juveniles improved in their performance across years in this task ($F(2,41) = 10.28$, $p < 0.001$). Meanwhile, adults did not change in their performance across test years ($p > 0.6$).

Reputation. There was no change in performance in this task across years in either *Pan* infants/juveniles ($p > 0.2$) or *Pan* adults ($p > 0.4$).

Physical cognition

Object permanence. *Pan* infants/juveniles did not change in their object permanence performance across test years ($p > 0.1$), but there was a trend toward improvement in performance across years among *Pan* adults ($F(2,4) = 6.13$, $p = 0.06$).

Transposition. *Pan* infants/juveniles did not change in their transposition performance across test years ($p > 0.2$), but there was a significant effect of year in *Pan* adults ($F(2,4) = 22.00$, $p = 0.007$), mainly reflecting a dip in adults' performance in 2009 relative to performance in 2008 and 2010 rather than systematic improvement across test years.

Number. There was no change in performance in this task across years in either *Pan* infants/juveniles ($p > 0.9$) or *Pan* adults ($p > 0.6$).

Tool use. McNemar's tests revealed a significant improvement in tool use among *Pan* infants/juveniles from 2008 to 2010 ($p = 0.006$), but no differences between other task years. Performance among adults did not change across the three task years.

Tool properties. *Pan* infants/juveniles improved in their performance across years in this task ($F(2,41) = 23.39$, $p < 0.001$). Meanwhile, adults did not change in their performance across test years ($p > 0.5$).

Attentional/Motivational Controls

Risk box. McNemar's tests revealed no significant changes in performance across task years among *Pan* infants/juveniles. The test could not be computed for adults because performance was consistent across individuals in all three years of testing (with all adults putting their hands into the hole of the box in all test years).

Unsolvable task. There was a significant effect of year on time spent trying to solve the unsolvable task in *Pan* infants/juveniles, with increases in each subsequent test year ($F(2,34) = 5.02$, $p = 0.01$). Meanwhile, *Pan* adults did not change in their performance across test years ($p > 0.3$).

Novel objects. There was no change in performance in this task across years in either *Pan* infants/juveniles ($p > 0.1$) or *Pan* adults ($p > 0.1$).

SUPPLEMENTAL TABLES AND FIGURES

Table A6.1. Order of cognitive tasks, Studies 1 and 2. Pictures of the setup of each task with apes are shown. Tasks with stars were used in both Studies 1 and 2 and were presented in an identical order. Task order differed slightly between *Pan* infants and children but was consistent within genus. Some tasks were presented in multiple test sessions; the first session where they were presented is depicted in the table.

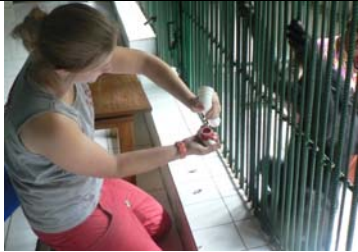
Task	Order: <i>Pan</i> infants and juveniles	Order: children	Set-up
Intention emulation	1	1	
Social referencing*	2	2	
Gaze-following around barriers*	3	3	 

Table A6.1 (Continued)

Social inhibition*	4	5	
Object permanence*	5	7	
Transposition*	6	8	
Gaze-following*	7	11	
Relative numbers*	8	14	

Table A6.1 (Continued)

Tool use*	9	9	
Tool properties*	10	10	
Social learning	11	12	
Point production*	12	13	
Goal understanding*	13	4	

Table A6.1 (Continued)

Reputation*	14	6	
-------------	----	---	---

Table A6.2. Comparisons of performance by task across genus and age, Study 1. GLM analyses were performed for each task with genus and age group as factors. For tasks with a success/failure measure, chi-squared and Kendall's tau analyses were performed. Significant results with Bonferroni-corrected post-hoc tests are indicated as follows: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

Category	Test	Significant increase with age?		Humans outperform apes?		
		Humans	<i>Pan</i>	2 years	3 years	4 years
Social cognition	Intention-emulation	Yes***		Yes*	Yes***	Yes***
	Social referencing			Yes**	Yes***	Yes***
	Gaze-following around barriers	Yes*		Yes*	Yes***	Yes***
	Social inhibition	Yes***			Yes***	Yes***
	Gaze following			Yes**		
	Social learning	Yes***			Yes***	Yes***
	Point production					Yes***
	Goal understanding				Yes***	Yes***
Physical cognition	Reputation					
	Object permanence	Yes***		Yes**	Yes***	Yes***
	Transposition	Yes***				Yes**
	Number	Yes***				Yes***
	Tool use	Yes*				Yes*
	Tool properties	Yes**		Yes**	Yes***	Yes***
Attentional/motivational controls	Risk box				Yes*	
	Unsolvable task					
	Novel objects					

Table A6.3. Comparisons of performance across test years in *Pan* infants/juveniles and *Pan* adults, Study 2. A repeated measures ANOVA with year as a factor was performed for each task. For tasks with a success/failure measure, we performed McNemar's tests to discriminate whether proportions of success differed between test years. Here, because there were a number of results that were significant at a trend level, significance levels with Bonferroni-corrected post-hoc tests are indicated as follows: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Category	Test	Significant increase across years?	
		<i>Pan</i> juveniles	<i>Pan</i> adults
Social cognition	Intention-emulation	Yes***	n/a
	Social referencing		
	Gaze-following around barriers		
	Social inhibition		
	Gaze following	Yes*	
	Social learning	Yes**	n/a
	Point production		
	Goal understanding	Yes***	
Physical cognition	Reputation		
	Object permanence		Yes*
	Transposition		
	Number		
	Tool use	Yes***	
	Tool properties	Yes***	
Attentional/motivational controls	Risk box		
	Unsolvable task	Yes***	
	Novel objects		

Human children

Soc obst	Gaze follow	GF barriers	Soc inhib	Intent emulate	Goal underst	Soc learn	Rep comp	Point product
98	96	93	89	85	84	79	68	62

Individually, 45.8% of children showed this exact pattern of emergence
(including twelve children – 25% – who passed all nine tasks)

Pan juveniles

GF barriers	Gaze follow	Rep comp	Soc obst	Point product	Soc inhib	Goal underst	Intent emulate	Soc learn
77	76	56	55	33	33	26	11	5

Individually, 24.5% of *Pan* juveniles showed this exact pattern of emergence

Figure A6.1. Patterns of emergence across social cognitive tasks in human children and *Pan* infants, Study 1. Each task is represented by its own box. The percentage of individuals meeting the passing criterion for each task is shown under its respective box. Individual support for patterns of emergence within each genus is also reported.

Human children

Obj perm	Tool props	Rel number	Transp	Tool use
94	82	80	71	33

Individually, 81.3% of children showed this exact pattern of emergence
(including ten individuals – 20.8% – who passed all five tasks and two
individuals – 4.2% – who passed none of the tasks)

Pan juveniles

Rel number	Obj perm	Transp	Tool props	Tool use
55	48	38	20	17

Individually, 51.0% of *Pan* juveniles showed this exact pattern of emergence
(including one individual – 2.0% – who passed all five tasks and five
individuals – 10.2% – who passed none of the tasks)

Figure A6.2. Patterns of emergence across physical cognitive tasks in human children and *Pan* infants, Study 1. Each task is represented by its own box. The percentage of individuals meeting the passing criterion for each task is shown under its respective box. Individual support for patterns of emergence within each genus is also reported.

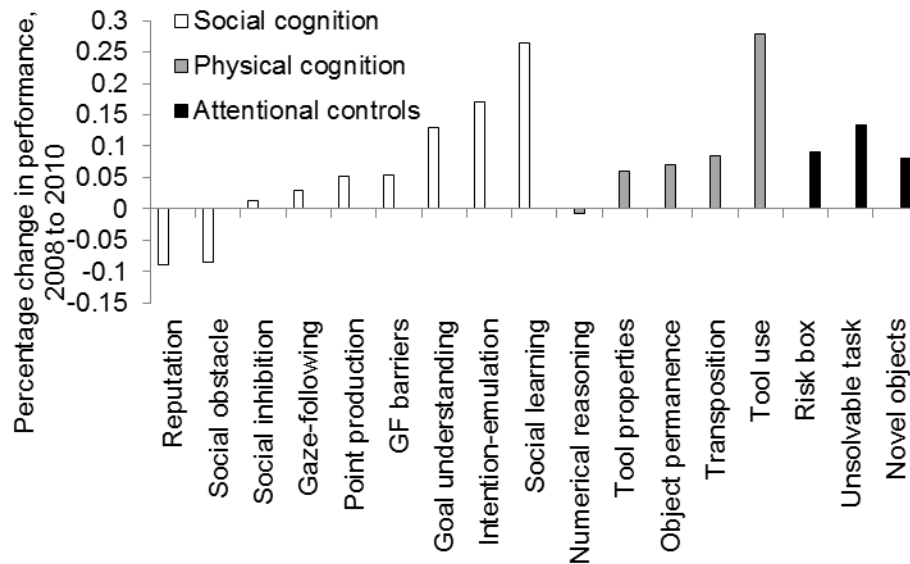


Figure A6.3. Degree of improvement in the social cognitive, physical cognitive, and attentional/motivational controls across test years in *Pan* infants/juveniles, Study 2. Social cognition tasks are shown first (white bars), followed by physical cognition tasks (grey bars), and then control tasks (black bars). Average change in performance across test years, calculated as intra-individual difference between performance in 2010 and that in 2008, is shown on the y-axis. Task labels are provided on the x-axis.

Appendix 6 Literature Cited.

- Barth, J., & Call, J. (2006). Tracking the Displacement of Objects: A Series of Tasks With Great Apes (*Pan troglodytes*, *Pan paniscus*, *Gorilla gorilla*, and *Pongo pygmaeus*) and Young Children (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes*, 32(3), 239-252.
- Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: comparing 12- and 18-month-olds. *Infant Behavior and Development*, 22(2), 277-282.
- Braeuer, J., Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2006). Making Inferences About the Location of Hidden Food: Social Dog, Causal Ape. *Journal of Comparative Psychology*, 120(1), 38-47.
- Call, J., Carpenter, M., & Tomasello, M. (2005). Copying results and copying actions in the process of social learning: chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Animal Cognition*, 8(3), 151-163.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), 1-143.
- Cox, D., Rosen, N., Montgomery, C., & Seal, U. (2000). *Chimpanzee sanctuaries: guidelines and management workshop report*. Paper presented at the Conservation Breeding Specialist Group (SSC/IUCN), Apple Valley, MN.
- Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature*, 415, 755.
- Grether, W., & Yerkes, R. (1940). Weight norms and relations for chimpanzees. *American Journal of Physical Anthropology*, 27(2), 181-197.
- Hamlin, K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450(7169), 557-559.
- Hanus, D., & Call, J. (2007). Discrete quantity judgments in the great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, *Pongo pygmaeus*): The effect of presenting whole sets versus item-by-item. *Journal of Comparative Psychology*, 121(3), 241-249.
- Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology*, 18(9), R370-R372.
- Hauser, M. (1997). Artifactual kinds and functional design features: what a primate understands without language. *Cognition*, 64, 285-308.

- Herrmann, E., Call, J., Hernandez-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis. *Science*, 317, 1360-1366.
- Herrmann, E., Call, J., Hernandez-Lloreda, M., Hare, B., & Tomasello, M. (2010). The Structure of Individual Differences in the Cognitive Abilities of Children and Chimpanzees. *Psychological Science*, 21(1), 102-110.
- Herrmann, E., Hare, B., Call, J., & Tomasello, M. (2010). Differences in the Cognitive Skills of Bonobos and Chimpanzees. *PLOS One*, 5(8), e12438.
- Herrmann, E., Wobber, V., & Call, J. (2008). Great apes' (Pan troglodytes, Pan paniscus, Gorilla gorilla, and Pongo Pygmaeus) understanding of tool functional properties after limited experience. *Journal of Comparative Psychology*, 122, 220-230.
- Kagan, J., & Snidman, N. (2004). *The Long Shadow of Temperament*. Cambridge, MA: Harvard University Press.
- Lieberman, D. E., Carlo, J., Ponce de Leon, M., & Zollikofer, C. (2007). A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. *Journal of Human Evolution*, 52, 647-662.
- Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science*, 311(5765), 1297-1300.
- Meltzoff, A. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, 24(4), 470-476.
- Miklosi, A., Kubinyi, E., Topal, J., Gacsi, M., Viranyi, Z., & Csanyi, V. (2003). A simple reason for a big difference: wolves do not look back at humans, dogs do. *Current Biology*, 13, 763-766.
- Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, 7(1), F1-F9.
- Phillips, W., Baron-Cohen, S., & Rutter, M. (1992). The role of eye contact in goal detection: Evidence from normal infants and children with autism or mental handicap. *Development and Psychopathology*, 4, 375-383.
- Piaget, J. (1952). *The Child's Conception of Number*. London: Routledge and Kegan Paul.
- Shea, B. (1983). Allometry and Heterochrony in the African Apes. *American Journal of Physical Anthropology*, 62(3), 275-289.
- Smith, B., Crummett, T., & Brandt, K. (1994). Ages of eruption of primate teeth: a compendium for aging individuals and comparing life histories. *Yearbook of Physical Anthropology*, 37, 177-231.

- Sophian, C. (1984). Spatial transpositions and the early development of search. *Developmental Psychology*, 20(1), 21-28.
- Tomasello, M., & Camaioni, L. (1997). A comparison of the gestural communication of apes and human infants. *Human Development*, 40, 7-24.
- Tomasello, M., & Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development*, 70(1), 1-155.
- Tomasello, M., Hare, B., & Fogleman, T. (2001). The ontogeny of gaze following in chimpanzees, *Pan troglodytes*, and rhesus macaques, *Macaca mulatta*. *Animal Behaviour*, 61, 335-343.
- Tomonaga, M. (2008). Relative numerosity discrimination by chimpanzees (*Pan troglodytes*): evidence for approximate numerical representations. *Animal Cognition*, 11(1), 43-57.
- Vlammings, P., Hare, B., & Call, J. (2010). Reaching around barriers: the performance of the great apes and 3–5-year-old children. *Animal Cognition*, 13(2), 273-285.
- Wobber, V., & Hare, B. (2011). Psychological health of orphan bonobos and chimpanzees in African sanctuaries. *PLOS One*, 6(6), e17147.
- Wobber, V., Wrangham, R., & Hare, B. (2010). Bonobos exhibit delayed development of social behavior and cognition relative to chimpanzees. *Current Biology*, 20(3), 226-230.