Age-dependent cognitive inflexibility in great apes

Héctor Marín Manrique[1,2] and Josep Call[1]

1Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103, Leipzig, Germany

2 University of the Balearic Islands, Ctra Valldemossa KM 7.5, 07122, Palma de Mallorca, Spain

Correspondence:
Héctor Marín Manrique
Departament de Psicologia Bàsica. University of the Balearic Islands
Edifici Guillem Cifre de Colonya, Ctra Valldemossa KM 7.5, 07122, Palma de Mallorca, Spain
e-mail: hector.marin@uib.es; phone: +34 971259899; fax: +34 971173190

Word count : 4114 (references not included)

The ability to suppress and/or change behaviour on the basis of negative feedback, often conceptualised as cognitive flexibility, has rarely been investigated in nonhuman great apes across a broad age range. Twenty-five chimpanzees, eight bonobos, seven orang-utans and three gorillas, whose ages ranged from five to forty-eight years, were presented with a transparent Plexiglas rectangular box horizontally attached to their cage mesh. A squared container 7.5 cm^2 fixed inside the apparatus contained a food reward (i.e. grape). While the container rested on its central position the grape was not accessible. To retrieve the grape the subjects needed to grasp the handle connected to the reward container and displace it sideways to reach one of the lateral access windows. Subjects were intensively trained to displace the handle to a specific side (right or left, depending on the group) to later reverse the rewarded side during the test. Performance in this reversal task did not significantly differ between species. However, a U-shape relation between age and perseverative responding (i.e. moves to the previously rewarded side) was observed, extending findings with humans to their closest living primate relatives.

Keywords: aging, cognitive flexibility, great apes, perseverative responding, reversal task.

The ability to suppress and/or change behaviour on the basis of negative feedback is essential to adapt in a changing environment. This ability, conceptualised as cognitive flexibility, belongs to the so-called executive control function and relies on the integrity of the prefrontal cortex (Miller, 2000). One of the most used tasks to study cognitive flexibility in humans is the Wisconsin Card Sorting Test (WCST) (Anderson, Damasio, Jones, & Tranel, 1991; Berg, 1948; Milner, 1963; Nagahama et al., 1996). The WCST measures the ability to learn to focus on a particular stimulus dimension (e.g., colour) and shift to another dimension (e.g., shape) as a function of changes in the reward contingencies. Proficiency in this task also requires generating hypothesis and replacing them as soon as they no longer predict reward delivery.

Cognitive flexibility in humans is negatively affected by aging (Albert & Moss, 1999; Haaland, Vranes, Goodwin, & Garry, 1986; Libon, Malamut, Swenson, Sands, & Cloud, 1994; Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998). A sample of 95 elderly healthy subjects was investigated by Haaland et al. (1986) using a modified version of the WCST. The oldest group (80 to 87 years) formed fewer categories and accumulated more errors than their younger counterparts (i.e. 64 to 68 years). Moreover, shrinkage of the prefrontal cortex has been associated with age-related increases in perseveration (Raz et al., 1998), which could provide a neural substrate for the deficits observed in elderly people.

Although numerous studies have tested cognitive flexibility in nonhuman primates (e.g., Amici, Aureli, & Call, 2008; Izquierdo, Newman, Higley, & Murray, 2007; Rygula, Walker, Clarke, Robbins, & Roberts, 2010), only a handful of studies have investigated its age-related deficits. This paucity of results is particularly surprising given that frontal cortical development follows a similar developmental pattern in human and nonhuman primates (Goldman-Rakic, 1987) and some models of human frontal cortical dysfunction have been evaluated and tested on nonhuman primates (Decamp and Schneider, 2004; Lewis, Hayes, Lund, & Oeth, 1992). Moore ,Killiany, Herndon, Rosene, and Moss (2005) developed the Conceptual Set Shifting Task (CSST), a test analogous to the WCST, to explore age-related cognitive deficits in rhesus monkeys. In the CSST the monkeys face a touch screen in which three stimuli appear that differ along two dimensions, their colour (red, green, and blue) and their shape (triangle, star, and circle). In some trials the target dimension is the colour (i.e. red) and subjects are rewarded for touching the appropriate colour. After 10 consecutive correct responses to the colour, the rewarded dimension, changes to shape (i.e. triangle). In order to succeed subjects need to form a conceptual set (colour) and then shift to a new conceptual set (shape) on the basis of feedback alone. Aged adult monkeys evidenced more problems both when forming the initial concept and later shifting to a different concept than young adults (Moore et al., 2003, 2005, 2006). Bonté, Flemming, and Fagot (2011) used a virtually identical task and reported similar findings in baboons. However, the onset age of the deficits greatly differed from one species to the other. While rhesus displayed an increase in perseverative responding at the age of twelve, baboons showed comparable levels of perseverative responding by eight years of age.

Weed, Bryant, and Perry (2008) also studied rhesus monkeys' cognitive flexibility in relation to age. They used an adaptation of the Cambridge Neuropsychological Test Automated Battery (CANTAB) measuring attentional set-shifting and perseverative responding to compare performance of juvenile (mean age 2.3 years) and adult (mean age 10.3 years) monkeys. Monkeys were trained to respond to one of two stimuli simultaneously presented on a touch screen. Once they mastered this simple discrimination they were required to reverse their initial response and choose the alternative non-reinforced stimulus exemplar (simple reversal) or else to identify which new stimulus from an array of different stimuli was predictive of reward delivery. The new stimulus sharing a dimension with the previous reinforced stimulus (intra-dimensional shift, IDS), or belonging to a new dimension (extra-dimensional shift, EDS). Juvenile macaques' performed worse than adults in the simple reversal task as well as in the tasks requiring an IDS, an IDS reversal, or an EDS.

Thus, contrary to the other two studies, Weed et al. (2008) found that younger individuals performed worse than older individuals. These two sets of studies, however, are not the only ones that have produced mixed results as a function of age. Picq (2007) tested lemurs in a Set Shifting Task using an apparatus with 6 corridors connected to a chamber containing a food reward. In any single trial only one corridor led to the reward. The authors measured lemurs' reversal discrimination based on spatial cues by changing the location of the corridor that had been associated with the reward (IDS). Additionally, Picq (2007) measured lemurs' ability to switch corridors on the basis of a visual cue (i.e. light) (EDS). Aged subjects committed more perseverative errors than younger subjects in the EDS and IDS tasks, although it did not reach statistical significance in the latter task. In contrast, Trouche, Maurice, Rouland, Verdier, and Mestre-Frances (2010) found that young adult lemurs made significantly more perseverative errors than older individuals in a three-panel runaway maze after the original rewarded location was no longer rewarded. Trouche et al. (2010) argued that young adult lemurs' higher levels of anxiety compared to older individuals translated into a significantly larger number of attempts at opening the wrong gate.

Although methodological differences between the studies reviewed above may contribute to explain the mixed results (e.g., some tests relied more heavily on a memory component than others), another plausible explanation might be that the relationship between age and cognitive flexibility is non-linear. To shed more light on the topic of the effects of age on cognitive flexibility in nonhuman primates we developed a reversal task with a strong motor component and a minimum contribution of complex perceptual information and memory loads. The task consisted of, displacing laterally an encapsulated baited box until reaching a window where the bait inside could be extracted. The baited box was fixed inside a rectangular transparent apparatus attached horizontally to the subjects' cage. Subjects were intensively trained to move the handle in one direction (i.e. right) to gain access to a grape to later change the rewarded side during the test. We administered this task to a relatively large sample of great apes belonging to all species ranging in age from 5 to 48 years. Despite the existence of a slight maturational *decalage* between species (with gorillas and orangutans being the fastest and slowest to mature, respectively), they all share similar developmental and life history trajectories characterized by a slow development and a long lifespan (Parker, 1999). More specifically, all ape species possess a long period of immaturity and maternal dependency followed by a reproductive period beginning at about 8 years of age in females and a lifespan of 40 to 50 years. Such similarities between species justify our decision to pool together all the species to obtain a reasonable large sample with a continuous age distribution that can be used to examine in detail the relation between age and motor control.

Methods

Subjects

Twenty-five chimpanzees (*Pan troglodytes*), eight bonobos (*Pan paniscus),* three gorillas (*Gorilla gorilla)* and seven orangutans (*Pongo abelii*) housed at the Wolfgang Köhler Primate Research Center (WKPRC) in the Leipzig Zoo participated in the study (see Table 1 for details). There were 11 males and 32 females ranging in age from 5 to 48 years. Subjects were housed in social groups of 6-18 individuals and spent the day in indoor (175-430 m^2) or outdoor enclosures (1400-4000 m^2), depending on the season. Both enclosures were spacious and naturally designed, equipped with climbing structures and enrichment devices to foster extractive foraging activity. All tests were conducted in special testing cages (5.1-7.3 m^2) interconnected by lockable doors. Subjects were provided with fresh fruits, vegetables, eggs, cereals, leaves and meat (once a week) distributed in three main meals (7.30 am, 1.30 pm and 5 pm). Some more food was dispensed between 7.30 am and 1.30 pm (mainly fresh fruit) and at 3.30 pm, as part of the enrichment program. Our experiments never interfered with the daily feeding routine. Water was available *ad libitum* during the experiments.

Ethical Note: Tests adhered to ethical principles for non-invasive research in compliance with the European and World Associations of Zoos and Aquariums (EAZA and WAZA) ethical guidelines. The zoo keepers called in the apes right before starting the test and they entered the testing room through a door connected with their indoor enclosure. Subjects were separated from the rest of their group only for the duration of the test, and were allowed to abandon the experiment at any sign of distress. When infants were tested their mothers were always sitting next to them in an adjacent cage, where visual as well as partial physical contact was still possible.

Table 1

Apparatus

The apparatus consisted of a transparent Plexiglas rectangular box (63 cm long x 9 cm side length) horizontally attached to the subjects' cage mesh. A squared container 7.5 cm^2 fixed inside the apparatus contained a food reward (i.e., grape). While the container rested on its central starting position the grape was visible but not accessible. To retrieve the grape the subject needed to grasp the handle connected to the reward container (see Figure 1) and displace it sideways 24 cm to reach one of the lateral access windows (Ø=4.3 cm), where the grape became accessible. A locking device situated 5.6 cm from each of the lateral windows permitted the experimenter to block and unblock each solution. A black painted surface (8.7 x 6.5 cm) prevented the apes from seeing the locking mechanism. When in place, this locking device stopped the sliding container before reaching the access window.

Figure 1

Procedure

Subjects were assigned to one of two groups in the training phase. One group was trained to displace the handle rightwards (right-then-left: N=23) and the other group was trained to displace the handle leftwards (left-then-right: N=20). In order to complete training, subjects had to displace the handle to the correct side for a total of 100 trials. The apes could accumulate a maximum of 15 grapes in 20-minute daily sessions. Thus, a minimum of 7 sessions was always required to reach the training criterion. Once this criterion was reached, subjects advanced to the test phase in which they had to displace the handle in the opposite direction of training to obtain the grape. Subjects received a maximum of two 20-min sessions in which they could accumulate up to 10 grapes. Throughout the experiment, the experimenter removed the grape every time the reward container became blocked (i.e. after false moves) and waited a few seconds before re-baiting the apparatus for the next trial. This was done to make mistakes more salient for the subjects. However, this procedure could not be followed in some cases because subjects became mildly agitated due to the removal of the reward.

Data scoring and analysis

All trials were videotaped. To assess inter-observer reliability, an observer who was unaware of the study's hypothesis scored whether the subject moved the handle to the left or to the right for 20% of the trials. Inter-observer reliability was excellent (Cohen's kappa=1, N=168). Our main dependent measure was the number of errors during the training and test phase. We distinguished between pre- and post-solution errors. Pre-solution errors consisted of the number of incorrect trials before the occurrence of the first correct trial. Post-solution errors consisted of the number of incorrect trials after the first successful trial. Due to the different number of trials administered during training and testing (100 vs. 10) and to enable a fair comparison between phases, we only considered the number of post-solution errors until subjects accumulated10 correct trials.

Our independent variables were experimental phase (training, test), species, and chronological age (measured in years). We analysed the data using two-tailed non-parametric statistics. The binomial test was employed to detect side biases in moving handle. Wilcoxon test allowed us to assess the difference between phases and the Kruskal-Wallis test was run to investigate differences in performance between species.

Results

Prior to training subjects failed to displace the handle towards the correct side above chance levels (Binomial test: P=0.55, N=43). Moreover, they showed no preference for displacing the handle toward a particular side (Binomial test: P=0.36, N=43). Subjects required an average of 103.5 (SEM=0.8, Median=102) trials to reach the training criterion of 100 correct trials.

Figure 2 presents the number of pre- and post-solution errors during the training and testing phases. Subjects committed significantly more pre-solution errors during testing compared to training (Wilcoxon test: z=5.34, P<0.001, N(1 tie)=42, Figure 2a). In contrast, there was only a trend for post-solution errors (Wilcoxon test: z=1.92, P=0.054, N(19 ties)=18, Figure 2b). Three young orang-utans (Suaq, Tanah, Raaja) and three adult chimpanzees (Corrie, Natascha, Jeudi) were not included in this last analysis because they failed to solve the task after the reversal.

Figure 2

Overall, there were no significant differences between species in the frequency of pre-solution errors during training (Kruskal-Wallis test: χ^2_3=2.82, P=0.42, N=43) or testing (Kruskal-Wallis test: χ^2_3=0.39, P=0.94, N=43). Similarly, there were no significant differences between species in the frequency of post-solution errors during training (Kruskal-Wallis test: χ^2_3=6.60, P=0.086, N=43) or testing (Kruskal-Wallis test: χ^2_3=1.82, P=0.61, N=37). Therefore, we pooled all species together in subsequent analyses.

Figure 3 presents the number of pre- and post-solution errors during testing as a function of age. There was a significant U-shaped relation between pre-solution errors and age (R^2=0.162, $F_{2,40}$=3.86, P=0.029, $\hat{Y}=0.0221*age^2 -0.8178*age + 17.08$, Figure 3a). The same relation still held after subtracting the number of errors during training from pre-solution errors during testing (as a way to control for general error proneness during training) (R^2=0.155, $F_{2,40}$=3.66, P=0.035, $\hat{Y}=0.0301*age^2 -1.2708*age + 18.04$). In contrast, there was no relation between post-solution errors and age before ($F_{2,34}$=0.43, P=0.655, Figure 3b) or after controlling for training errors ($F_{2,34}$=2.11, P=0.136). Similarly, there was no relation between pre- or post-solution errors during the training and age ($F_{2,40}$<1.34, P>0.28).

Figure 3

Discussion

The number of pre-solution errors during the test phase was particularly high in the youngest and oldest individuals of our sample. Thus, age was a reliable predictor of perseverative responding. The highest frequencies of errors pre-solution were observed in a 5-year-old male chimpanzee (Kofi) and a 30-year-old female chimpanzee (Natascha), with 33 and 31 errors, respectively. Moreover, the only subjects who failed the test (after the reversal) were either younger than 7 or older than 27 years of age. In contrast, age did not predict post-solution errors in the test phase (and neither pre- or post-solution errors during training), with all of them being much less frequent than pre-solution errors during test across the entire age range. We found no evidence of species differences in pre- or post-solution errors during the training or the testing phase.

Results of the present study fit well with previous findings in monkeys (Bonté et al., 2011; Moore et al., 2003, 2005, 2006; Picq, 2007; Weed et al., 2008; Zeamer et al., 2011). Aged subjects committed more pre-solution errors than their young adult counterparts. Additionally, they mirrored Weed's et al. (2008) findings showing that the youngest subjects were also impaired in their ability to inhibit a previously rewarded response compared to young adults. However, we have to be cautious before drawing general conclusions as the tasks used differed between studies. While Weed et al. (2008) and Zeamer et al. (2011) presented subjects with a simple discrimination reversal task equivalent to the one used in the current study, the remaining authors employed shift response set tasks (i.e. extra-dimensional shift). As we discuss next, performing these two types of task seems to recruit different areas of the prefrontal cortex.

Wise, Murray, and Gerfen (1996) proposed a model that links different types of cognitive/behavioural flexibility to different prefrontal cortical areas of the monkey's brain. The model distinguishes two types of processing: a lower-order processing, allowing for a shift in response within a dimension (intra-dimensional shift) and a higher-order processing, allowing shifts in response from one stimulus's dimension to another (extra-dimensional shift; i.e., from colour to shape). The former stimulus processing would simply assign a positive or negative valence to the whole stimulus. The second type of processing, however, would imply treating the different dimensions of a stimulus separately, and assigning a positive or negative valence to each of these dimensions also separately.

To solve the reversal task employed in the current study subjects had to assign a positive or negative valence to the whole response (i.e. moving the handle right or left). No rules or categories needed to be formed. According to Wise's model, this task would fall into the lower-order processing category. Since subjects hardly made any regressive (post-solution) error during the reversal, the perseverative responding observed here seems due to incapacity to stop a previously rewarded response, rather than to inability to produce a new alternate behaviour. It is conceivable, however, that a more complex behavioural change would throw different results, in the form of more regressive (post-solution) errors.

The idea of two different types of processing responsible for the intra-dimensional (reversal learning) and extra-dimensional response shifts has received some empirical support. Studies with monkeys indicate that different prefrontal cortex sub-regions are involved in different types of cognitive flexibility (Dias, Robbins, & Roberts, 1996, 1997). The dorsolateral prefrontal cortex would be responsible for the response shifts from one stimulus dimension to a different stimulus dimension, as lesions of this brain structure impair set-shifting but spare the capacity to learn a simple reversal. On the other hand, lesions of the orbital prefrontal cortex (OFC) hinder the learning of a reversal (intra-dimensional shift) but spare the extra-dimensional shifts of response (Dias et al., 1996, 1997). Lesion studies with rats support the same functional and structural distinction. Thus, if the damage produced by the lesions is limited to the prelimbic area the rodents can still learn and reverse their learning, but fail in extra-dimensional shifts. The opposite is true for damage to the orbitofrontal cortex, which impairs reversal learning but spares extra-dimensional response shifts (see Ragozzino, 2007 for a review). It is important to note that lesions of these brain structures do not affect acquisition, but specifically impair the shifts of response from one dimension to another, or from one stimulus exemplar to another along the same dimension. In other words, errors occur when there are changes in the relation value established between stimulus-response-outcome (Ragozzino, 2007).

There are at least two other studies that investigated how nonhuman ape species overcome prepotent responses. The first study employed the classical Piagetian A-not-B error task, in which subjects were rewarded for finding a food item hidden under one of three cups on three consecutive trials and then the food item was moved to a different cup in full view of the subject (e.g., Barth & Call, 2006, see also MacLean et al., 2014). One can see the similarity between this task and the current one as in both cases the response became prepotent after being rewarded multiple times. Just like in the current study, the four great ape species performed at a similar level. Reaching directly for a food item placed behind a transparent barrier also constitutes a prepotent response that does not require any formal training. Vlamings, Hare, and Call (2010) took advantage of this reaching prepotent response and presented a task in which subjects had to inhibit reaching directly for the food from the front and instead make a detour to grab the food from behind. Unlike the results of the present study, Vlamings et al. (2010) found that orang-utans outperformed all the other great ape species. One possible explanation for this difference is that these two tasks tap onto different aspects of inhibitory control. While the detour reaching tasks does not require any formal training to reveal its effects, the A-not-B error task is initially neutral and requires several trials to create the prepotent response. The label "inhibitory control" is associated with a variety of tasks in the literature that may rely on different cognitive processes and possibly different brain substrates. Thus, equating the behavioural results obtained through them might be misleading and future studies are needed to map out the relations between various tasks that are considered to measure inhibitory control.

The most relevant finding of the present study is perhaps that juvenile subjects committed more pre-solution errors during the test (but not during training) than young adults. To our knowledge, there is only one other study with primates reporting similar findings (Weed et al., 2008). Human and nonhuman primate prefrontal cortical maturation seems to follow a similar pattern (Goldman-Rakic, 1987). Frontal lobe maturation in human progresses in a back-to-front direction, beginning in the primary motor cortex and ending in the prefrontal cortex, that does not reach full maturity until early adulthood (Gogtay, et al., 2004). Paralleling this maturational pattern, adolescent performance in several tasks relying on the prefrontal cortex is not yet at adult levels (Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001). Also, an inverted U-shape relation between age and inhibitory control has been reported in humans (Dempster, 1992). Therefore, it is not surprising that juveniles in our sample had more problems than the young adults to learn the reversal.

To our knowledge, this is the first study to address how aging affects cognitive inflexibility in nonhuman apes. The strength of our results resides in the simplicity of the task employed, virtually devoid of memory demands; and the sample size, large enough to avoid forming age clusters that could bias or superimpose a specific shape to our results. By pooling the data from the four great ape species for statistical analysis we assumed that the life cycle of the four ape species was similar. This assumption, however, is based on only a handful of available studies. Wobber, Wrangham, and Hare (2010) reported slight differences in the ontogeny of inhibitory social control between chimpanzees and bonobos. More specifically, pre-weaning bonobos had more difficulty to refrain from begging from a particular experimenter compared to both post-weaning bonobos and pre-weaning chimpanzees. Moreover, Wobber et al. (2010) also found a positive relationship between age and performance in a social reversal task in bonobos but not in chimpanzees. Taken together these findings suggest that social inhibitory control might develop earlier in chimpanzees compared to bonobos. Future studies are needed to investigate the developmental trajectories and the relationship between social and non-social inhibitory control.

Several studies have assigned similar longevity about 60 years to chimpanzees and orangutans (Hakeem, Sandoval, Jones, & Allman, 1996; Herndon, Tigges, Anderson, Klumpp, & McClure, 1999; Wich et al. 2004), although it is true that orangutans seem to have a slower life history and hence, a little advantage over the chimpanzees. As for the gorillas few data are available but they seem to have the shortest life span (close to 50 years), which fits well with the idea of leaf-eaters having shorter life spans compared to fruit-eaters (Hakeem et al., 1996). No reliable data were found for the bonobos. We are aware that this constitutes a limitation of our study and we encourage our colleagues to run similar studies with apes that include subjects of all ages. Together, these data might allow us to produce a function that accurately predicts subjects’ performance on the basis of age.

Acknowledgements

We would like to thank Pilar Andrés for her comments on the manuscript and Lou Haux for reliability coding.

References

Albert, M., & Moss, M.B. (1999). Cognitive profiles of normal aging. In: A. Peters, & J. Morrison (Eds.), *Cerebral cortex*, (pp 1-20). Plenum Press: New York.

Amici, F., Aureli, F., & Call, J. (2008). Fission-Fusion Dynamics, Behavioral Flexibility, and Inhibitory Control in Primates. *Current Biology*, *18*, 1415-1419.

Anderson, S.W., Damasio, H., Jones, R.D., & Tranel, D. (1991). Wisconsin Card Sorting Test performance as a measure of frontal lobe damage. *Journal of Clinical and Experimental Neuropsychology*, *13*, 909-922.

Anderson, V.A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology*, *20*, 385-406.

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*, *32*(3), 239-252.

Berg, E.A. (1948). A simple objective test for measuring flexibility in thinking. *Journal of General Psychology*, *39*, 15-22.

Bonté, E., Flemming, T., & Fagot, J. (2011). Executive control of perceptual features and abstract relations by baboons (*Papio papio*). *Behavioral Brain Research*, *222*, 176-182.

Decamp. E., & Schneider, J.S. (2004). Attention and executive function deficits in chronic low-dose MPTP-treated non-human primates. *European Journal of Neuroscience*, *20*(5), 1371-1378.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*, 45–75

Dias, R., Robbins, T.W., & Roberts, A.C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature*, *380*, 69–72.

Dias, R., Robbins, T.W., & Roberts, A.C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sorting Test: restriction to novel situations and independence from "on-line" information. *Journal of Neuroscience*, *17*, 9265–9297.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., & Vaituzis, A.C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Journal of Neuroscience, 101*(21), 8174-8179.

Goldman-Rakic, P.S. (1987). Circuitry of primate prefrontal cortex and regulation of behaviour by representational memory. In P.F. Mouncastle (Eds.), *Handbook of physiology(*pp 373–417). Washington, DC: The American Physiological Society.

Hakeem, A., Sandoval, G.R., Jones, M., & Allman, J. (1996). Brain and life span in primates. In J.E. Birren, & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (pp 78-104). San Diego: Academic Press.

Haaland, K.Y., Vranes, L.F., Goodwin, J.S., & Garry, P.J. (1986). Wisconsin Card Sort Test performance in a healthy elderly population. *Journal of Gerontology*, *42*(3), 345-346.

Herndon, J.G., Tigges, J., Anderson, D.C., Klumpp, S.A., & McClure, H.M. (1999). Brain weight throughout the life span of the chimpanzee. *Journal of Comparative Neurology*, *409*(4), 567 – 572.

Izquierdo, A., Newman, T.K., Higley, J.D., & Murray, E.A. (2007). Genetic modulation of cognitive flexibility and socioemotional behavior in rhesus monkeys. *Proceedings of the National Academy of Sciences*, *104*(35), 14128-14133.

Lewis, D.A., Hayes, T.S., Lund, J.S., & Oeth, K.M. (1992). Dopamine and the neural circuitry of primate prefrontal cortex: implications for schizophrenia research. *Neurospychopharmacology*, *6*(2), 127-34.

Libon, D.J., Malamut, B.L., Swenson, R., Sands, L.P., & Cloud, B.S. (1994). Age, executive functions, and visuospatial functioning in healthy older adults. *Neuropsychology*, *8*, 38-43.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology*, *9*, 90 –100

Miller, E.K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, *1*, 59–65.

MacLean, E.L., Hare, B., Nunn, C.L., Addessi, E., Amici, F., Anderson, R.C., …Zhao, Y. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences*, *111*, E2140-E2148.

Moore, T. L., Killiany, R. J., Herndon, J, G., Rosene, D. L., & Moss, M. B. (2003). Impairment in abstraction and set shifting in aged rhesus monkey. *Neurobiology of Aging*, *24,* 125–134.

Moore, T. L., Killiany, R. J., Herndon, J. G., Rosene, D. L., & Moss, M. B. (2005). A non-human primate test of abstraction and set shifting: An automated adaptation of the Wisconsin Card Sorting Test. *Journal of Neuroscience Methods*, *146,* 165–173.

Moore, T. L., Killiany, R. J., Herndon, J. G., Rosene, D. L., & Moss, M. B. (2006). Executive system dysfunction occurs as early as middle-age in the rhesus monkey. *Neurobiology of Aging*, *27*, 1484-1493.

Nagahama, Y., Fukuyama, H., Yamauchi, H., Matsuzaki, S., Konishi, J., Shibasaki, H., & Kimura, J. (1996). Cerebral activation during performance of a card sorting test. *Brain*, *119*, 1667–1675.

Parker, S.T. (1999). The life history and development of great apes in comparative perspective. In S.T. Parker, R.W. Mitchell, & H.L. Miles (Eds.), *The mentalities of gorillas and orang-utans* (pp 43-69). New York: Cambridge University Press.

Picq, J.L. (2007). Aging affects executive functions and memory in mouse lemur primates. *Experimental Gerontology*, *42*, 223-232

Ragozzino, M.E. (2007). The contribution of the medial prefrontal cortex, orbitofrontal cortex and dorsomedial striatum to behavioral flexibility. *Annals of the New York Academy of Sciences*, *1121*, 355-375

Raz, N., Gunning-Dixon, F.M., Head, D., Dupuis, J.H., & Acker, J.D. (1998). Neuroanatomical correlates of cognitive aging: evidence from structural magnetic resonance imaging. *Neuropsychology*, *12*(1), 91-114.

Rygula, R., Walker, S.C., Clarke, H.F., Robbins, T.W., & Roberts, A.C. (2010). Differential contributions of the primate ventrolateral prefrontal and orbitofrontal cortex to serial reversal learning. *Journal of Neuroscience*, *30*(43)-14552-14559.

Trouche, S.G., Maurice, T., Rouland, S., Verdier, J.M., & Mestre-Frances, N. (2010). The three-panel runway maze adapted to *Microcebus murinus* reveals age-related differences in memory and perseverance performances. *Neurobiology of Learning and Memory*, *94*, 100–106.

Vlamings, P.H.J.M., Hare, B., & Call, J. (2010). Reaching around barriers: the performance of the great apes and 3-5 year-old children. *Animal Cogniton*, *13*, 273-285.

Weed, M.R., Bryant, R., & Perry, S. (2008). Cognitive development in macaques: Attentional set-shifting in juvenile and adult rhesus monkeys. *Neuroscience*, *157*, 22-28.

Wich, S. A., Utami-Atmoko, S. S., Mitra Setia, T., Rijksen, H. R., Schürmann, C., van Hooff, J. A. R. A. M., & van Schaik, C. P. (2004). Life history of wild Sumatran orangutans (Pongo abelii). *Journal of Human Evolution*, *47*, 385–398.

Wise, S.P., Murray E.A., & Gerfen, C.R. (1996). The frontal cortex—basal ganglia system in primates. *Critical Reviews in Neurobiology*, *10*, 317–356.

Wobber, V., Wrangham, R. & Hare, B. (2010). Bonobos exhibit delayed development of social behavior and cognition relative to chimpanzees. *Current Biology*, *20*, 226–230.

Zeamer, A., Decamp, E., & Clark, K. (2011). Attention, executive functioning and memory in normal aged rhesus monkeys. *Behavioural Brain Research*, *219*(1), 23-30.

Table 1. Subjects included in the study.

Subject	Species	Gender	Age (years)	Rearing history	Experimental group
Ulla	Chimpanzee	Female	33	Nursery	Right-Left
Pia	Chimpanzee	Female	10	Mother	Left-Right
Annet	Chimpanzee	Female	9	Nursery	Left-Right
Riet	Chimpanzee	Female	33	Nursery	Left-Right
Natascha	Chimpanzee	Female	30	Nursery	Left-Right
Corrie	Chimpanzee	Female	34	Nursery	Left-Right
Sandra	Chimpanzee	Female	17	Mother	Left-Right
Lome	Chimpanzee	Male	8	Mother	Right-Left
Patrick	Chimpanzee	Male	12	Mother	Right-Left
Kara	Chimpanzee	Female	5	Mother	Right-Left
Kofi	Chimpanzee	Male	5	Mother	Right-Left
Robert	Chimpanzee	Male	35	Nursery	Right-Left
Fraukje	Chimpanzee	Female	34	Nursery	Right-Left
Dorien	Chimpanzee	Female	30	Nursery	Right-Left
Tai	Chimpanzee	Female	8	Mother	Left-Right
Frodo	Chimpanzee	Male	16	Mother	Left-Right
Fifi	Chimpanzee	Female	16	Mother	Right-Left
Alexandra	Chimpanzee	Female	9	Nursery	Left-Right
Alex	Chimpanzee	Male	8	Nursery	Right-Left
Jahaga	Chimpanzee	Female	16	Mother	Left-Right
Gertruida	Chimpanzee	Female	16	Mother	Right-Left
Jeudi	Chimpanzee	Female	28	Mother	Left-Right
Frederike	Chimpanzee	Female	40	Mother	Right-Left
Brigitta	Chimpanzee	Female	48	Mother	Right-Left

Daza	Chimpanzee	Female	28	Unknown	Left-Right
Joey	Bonobo	Male	26	Nursery	Left-Right
Kuno	Bonobo	Male	12	Nursery	Right-Left
Yasa	Bonobo	Female	11	Mother	Right-Left
Luisa	Bonobo	Female	5	Mother	Right-Left
Gemena	Bonobo	Female	9	Mother	Right-Left
Fimi	Bonobo	Female	6	Mother	Left-Right
Lexi	Bonobo	Female	15	Nursery	Right-Left
Jasongo	Bonobo	Male	24	Mother	Left-Right
Dokana	Orang-utan	Female	18	Mother	Left-Right
Padana	Orang-utan	Female	11	Mother	Left-Right
Pini	Orang-utan	Female	20	Mother	Right-Left
Kila	Orang-utan	Female	8	Mother	Right-Left
Raaja	Orang-utan	Female	7	Mother	Right-Left
Suaq	Orang-utan	Male	5	Mother	Right-Left
Tanah	Orang-utan	Female	5	Mother	Left-Right
Kibara	Gorilla	Female	6	Mother	Left-Right
Abeeku	Gorilla	Male	15	Mother	Left-Right
Kumili	Gorilla	Female	10	Mother	Right-Left

Figure captions

Fig 1. Apparatus from the apes' perspective. The white arrow at the centre signals the grasping handle. Black arrows left and right signal the windows where the grapes can be retrieved. Black stripes lateral to each window prevent the ape from detecting the lockable device in the reversal of the task.

Fig 2. Frequency of errors pre-solution (a) and (b) post-solution in the training and test phases. The line represents the median, the bottom and top of each box represents the 25th and 75th percentile, the whiskers show the minimum and maximum values that are not considered outliers (i.e., values > 1.5*IQR from the 25th or 75th percentile) which are in turn represented by circles.

Fig 3. Frequency of errors pre-solution (a) and (b) post-solution as a function of age.