The Success Rate in a Complicated Spatial Memory Test Is Determined by Age, Sex, Life History and Search Strategies in Cynomolgus Monkeys

HUDA S. DARUSMAN[1,4], OTTO KALLIOKOSKI[1], DONDIN SAJUTHI[2], STEVEN J. SCHAPIRO[1,3,] ALBERT GJEDDE[5,6,7,8] and JANN HAU[1]

Departments of [1]Experimental Medicine and [5]Neuroscience and Pharmacology,
Faculty of Health Science, University of Copenhagen, Copenhagen, Denmark;
[2]Primate Research Center, Bogor Agricultural University, Bogor, Indonesia;
[3]Department of Veterinary Sciences, The University of Texas, MD Anderson Cancer Center, Bastrop, TX, U.S.A.;
[4]Department of Anatomy, Physiology and Pharmacology, Faculty of Veterinary Medicine,
Bogor Agricultural University, Bogor, Indonesia;
[6]Center for Functionally Integrative Neuroscience, University of Aarhus, Aarhus, Denmark;
[7]Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, MD, U.S.A.;
[8]Department of Neurology and Neurosurgery, McGill University, Montréal, Québec, Canada

Abstract. *In a retrospective analysis of data from three studies using a delayed response task in cynomolgus monkeys, we examined the subjects' search patterns and success rates. Twenty-seven monkeys of both sexes, divided into three age groups, were tasked with retrieving two food items hidden in an array of six identical opaque cups. Although the task was challenging for all subjects, generating a high level of guesswork, evidence of common behaviors when approaching the spatial memory test were found. The search patterns employed by the monkeys suggest the use of landmark cues, adaption in response to failure and chronological memory recall. These strategies appeared to be shared by most subjects, however, the overall success rate appeared to also depend on individual characteristics including age, gender and whether the subject had been born in caged captivity or not. By elucidating some of the underlying cognitive mechanisms, these findings may serve to refine interpretation of future studies using similar delayed response tasks in non-human primates.*

Correspondence to: Huda S. Darusman, Department of Experimental Medicine, Panum Institute, building 16.1, Blegdamsvej 3B, DK 2200 Copenhagen, Denmark. Tel: +45 35327012, Fax: +45 35327399, e-mail: hudada@sund.ku.dk

Delayed response tasks are utilized for assessing working memory in non-human primates (1). These tests assess the subject's ability to retain information that varies in time and/or in content, which is the type of memory that is often affected in dementias (2). Consequently, delayed response tasks have been frequently used in studying memory decline and dementias in aging primates (1, 3). In addition, the complex behavior exhibited by non-human primates in these tests can be used to infer underlying cognitive processes, making them valuable tools for studies in cognitive neuroscience and comparative psychology (4-7).

The present report collates the results from three delayed response task studies in a retrospective analysis, focusing on aspects of the data that were beyond the scope of the original studies. To assess the working memory in cynomolgus monkeys (*Macaca fascicularis*) of differing ages, we utilized a type of delayed response task where the subjects were assessed to see whether they remembered the location of food items hidden in two of six identical opaque cups arranged in a straight line. Part of the data used herein have previously been published in a pair of studies exploring impaired delayed response task performance in aging cynomolgus monkeys and its association to biomarkers associated with Alzheimer's disease (8, 9). Additional data come from a study that relates delayed response task performance of old monkeys to clinical magnetic resonance imaging findings (10). The overall performance in the task referred to as the memory load test (MLT), which we will be focusing on in the present study, is summarized in Table I.

When analyzing the data of the aforementioned studies, it was noted that individuals within an age group would differ greatly in

Table I. *Summary of overall test performance in the Memory Load Test (MLT). With subjects having been tested in different numbers of trials, data have also been calculated as weighted by number of trials. Retrieval accuracies are given as averages with 95% confidence intervals. A subject could, at most, recover two baits in a single trial (thus yielding a retrieval accuracy of 2). The response frequency would decrease with subjects not providing full answers (designating two cups) in a trial.*

Age group	Young	Middle-aged	Old
Retrieval accuracy in tests without delay (baits per test)	Unweighted: 1.25 (1.07-1.44) Weighted by no. of trials: 1.24 (1.23-1.25)	Unweighted: 1.13 (1.01-1.25) Weighted by no. of trials: 1.13 (1.12-1.13)	Unweighted: 0.88 (0.79-0.96) Weighted by no. of trials: 0.85 (0.84-0.85)
Retrieval accuracy in tests with 30 s. delay (baits per test)	Unweighted: 1.05 (0.84-1.26) Weighted by no. of trials: 1.05 (1.03-1.07)	Unweighted: 1.01 (0.87-1.14) Weighted by no. of trials: 1.01 (1.00-1.02)	Unweighted: 0.72 (0.65-0.79) Weighted by no. of trials: 0.71 (0.71-0.72)
Average response frequency	100%	100%	Unweighted: 98.1% Weighted by no. of trials: 96.8%

their performance, begging the question whether other characteristics might also influence test performance. Non-human primate studies often feature small experimental groups – it is fairly common that the subjects are identified by name – and the cognitive studies will often discuss individual personalities. With a fairly large sample size, we decided to use this opportunity to investigate potential sex differences and whether the type of breeding facility the subjects originated from would influence test performance. Moreover, food item baits in certain cup positions would be retrieved more often than others and incorrect answers conformed to patterns shared by all subjects. Our hope was that insights into these answering patterns might be further used to distinguish between errors of different types, and thus, might serve to refine the framework for interpreting results from similar tests. With monkeys' and apes' performance in (static) delayed response tasks being comparable (11), we aimed to compare and contrast the findings in cynomolgus monkeys to trends/strategies previously discussed in great apes. Of particular interest was whether the monkeys use landmark cues to locate baits, the underlying mechanisms leading to clustering of answers and whether the monkeys are able to approach the challenge of retrieving baits in a systematic fashion.

We hypothesized that the cynomolgus monkeys would exhibit common strategies – underlying the spatial response biases – in solving the delayed response task. Since delayed response task performance tends to decline with age (12), we further hypothesized that the strategies would differ between age groups. Lastly, we hypothesized that the sex and life-history of the subjects would influence the overall performance, in addition to the previously-found effects of age.

Materials and Methods

Working memory was assessed using a delayed response task similar to what has been previously used in apes (13-15), termed the [working] memory load test (MLT). In this test, six identical opaque cups (7 cm in diameter) were arranged on a 55×20 cm tray. The cups were placed upside-down on the tray, always going from left-to-right

(directions are given from the experimenter's point of view throughout), and small pieces of fruit (sliced apple or guava) were placed under two of the cups. The experimenter made sure that the subject was oriented toward the test apparatus and was watching as the cups were baited. Either immediately after the baiting, or following a 30-second delay, the tray was placed within the subject's reach, allowing the monkey to designate a cup by touching it. If a baited cup was chosen, the subject received the fruit within; if an empty cup was designated, the experimenter showed the cup to be empty. The tray was left in front of the subject until two answers had been provided (or the monkey moved away from the front of the testing cage, signaling a loss of interest in the test), at which point the trial was terminated. The full testing procedure can be seen in films that have been supplied as online supplements to this article. The test was conducted in randomized modules consisting of all 30 possible trial combinations (15 possible bait combinations with or without a 30-second delay). Subjects would receive one module's worth of testing per day – no more. The baiting process is illustrated in Figure 1.

Originally experimentally-naïve to delayed response tasks, all subjects were trained in three types of tests (for information on the two memory tests not covered in this study we refer to Darusman *et al*. (8)). Training consisted of building up the test procedure by increments. In the first step, the subjects were trained to collect a bait from the testing tray, habituating to the testing equipment and experimenter. Once successful, the subject was trained to retrieve a bait by designating a cup, but where the bait was only partially hidden by the cup. Finally, the subject was trained to retrieve a fully hidden bait from a cup. At the time of testing, subjects voluntarily sat down and faced the experimenter for the baiting procedure. If a subject at any point during testing lost interest and would not observe the baiting process, testing was stopped, to be resumed another day. When the subjects reached a stable performance level in the tests and showed strong motivation for participating in the task, data collection commenced.

The MLT data were collected from 27 cynomolgus monkeys of both sexes, consisting of both island-bred, wild-caught animals and animals born in caged captivity (Table II). The cage-bred animals originated from the Primate Research Center (PRC) IPB breeding facility (Bogor, Indonesia), whereas the island-bred animals were collected from the PRC natural habitat breeding facility on Tinjil island. Inside the PRC breeding facility, the animals were housed in social groups (uni-male-multi-female in composition) with access

Figure 1. *The baiting process started with the experimenter showing the baits (A), arranging the cups (B) and presenting the test array by sliding the tray to the subject (C). The subject touched the first cup (D) and received the bait (E) when picking the first baited cup; after having consumed the bait, the subject touched the second cup and received the bait when successfully choosing the second baited cup. The tray was withdrawn from the subject and all the cups opened (E) in full view of the subject, indicating the test was finished.*

to both indoor and outdoor areas. As an integrated standard operating procedure, all animals had to be quarantined for a month before being brought into the testing facilities to minimize the risk of their bringing in pathogens. The subjects were caged in pairs approximately one month prior to being quarantined. Since the quarantine itself took 30 days, the subjects were in essence habituated to their new environment – caging, cage-mates, diet, staff, and daily routines – for two months prior to training/testing. In the testing environment, the subjects were paired by age, adapting the age ratio of Moss *et al*. (16) from humans to macaques. The age groups were "Young" (4-9 years old), "Middle aged" (10-19 years old) and "Old" (above 20 years old). A finer age-subdivision, although desirable, was not possible. Whereas the ages of the cage-bred animals were known, the island-bred animals' (approximate) ages had to be determined through dental scaling (17, 18).

An implicit assumption underlying a subset of the analyses in this study is that the time delay employed in half of the trials influences the behavior of the subjects indirectly – and only indirectly – by affecting memory retention. The subjects are less likely to correctly recall the bait placement when a delay is employed; but it is the remembering, or not remembering, that influences the subjects' behavior. As the effect is of an indirect nature, we can choose to omit the time delay as an explanatory variable when analyzing answering patterns/behavior (however, not when analyzing performance); the time delay is only one of many unaccounted-for factors that may influence memory retention. Previously (8), inter-observer reliability had been estimated from a random subset (n=90) of videotaped tests scored by two independent observers and found to be high (Cohen's kappa: 0.97).

Most analyses were carried out at the level of the individual; however, for studying the clustering of answers in the MLT, all trials were considered independent observations, the hypothesis thus being that the clustering is, in all subjects, due to the same underlying mechanisms. To analyze which factors contributed to the overall retrieval accuracy of baits in the test, a general linear mixed effects model (GLMM) was applied. The retrieval accuracy (an ordinal variable) was modeled with a generalized logit link, subject identities were accounted for as random effects (with an unstructured covariance type), and age [classes], origin (cage-bred or island-bred), sex and directionality of answers were provided as explanatory variables. Finally, the interaction of these variables with the time delay was tested, as essentially there were two levels of difficulty – with and without a time delay – and it was conceivable that some individuals would be better at tackling the more difficult tests. To accommodate this exploratory model structure, Satterthwaite approximations were used for the degrees of freedom and robust covariances were used for the tests of fixed effects. For testing whether adjacent answers affected the rate of missing both baits, a similar GLMM was constructed. Here, a nested structure for the individuals was used, accounting for individuals, sex and age as a random effect. The double failure was considered a binary outcome and the effect of providing adjacent answers was put forth as an explanatory variable.

Results

Analyzing the main factors influencing the retrieval accuracy in the trials (Table III), an effect of age, previously discussed by Darusman *et al*. (8, 9), was found as well as the
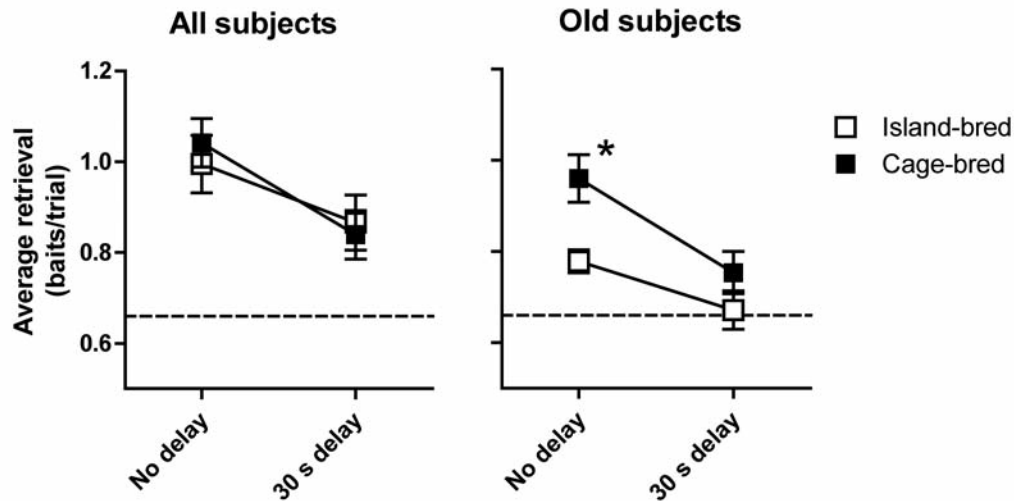
Figure 2. *Performance difference of subjects depending on life-history. The effect of a delay on the number of recovered baits for all subjects (n=15 and n=12 for island-bred and cage-bred subjects, respectively) and old subjects alone (n=7 and n=8 for island-bred and cage-bred subjects, respectively). The asterisk denotes a significant difference between the groups at a level of p<0.05 as tested by multivariate analysis of variance (MANOVA). The dashed line represents the chance level recovery.*
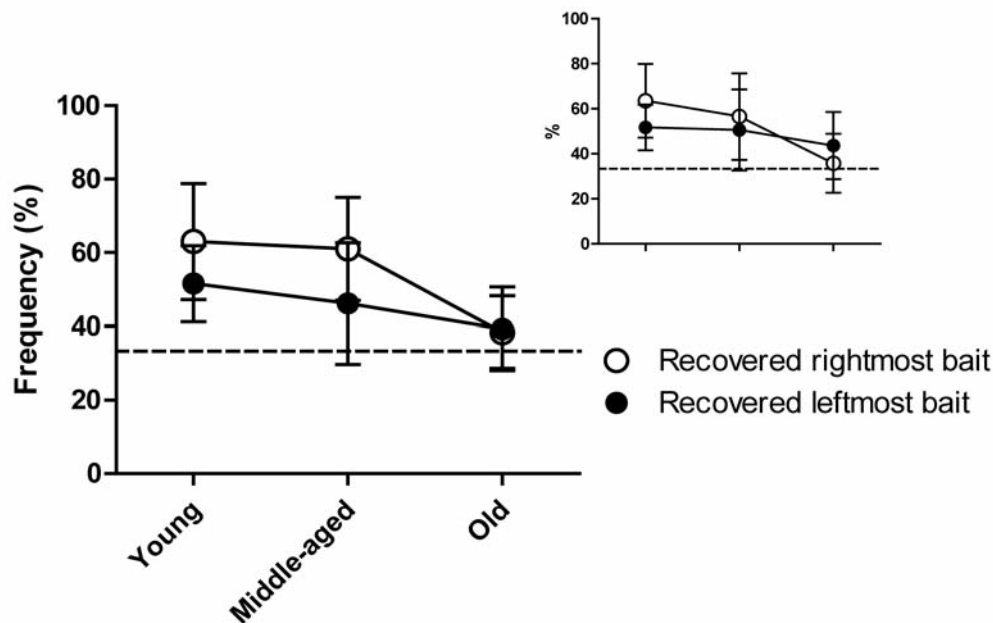

Figure 3. *Average fraction of correctly-retrieved baits by relative placement. The dashed line represents the chance level of retrieving a bait. The error bars represent standard deviations. The more illustrative weighted data (where individuals are weighted by the number of trials they have carried out – refer to Table II) is shown in the main graph, with the un-weighted data inset in the smaller graph.*

previously established (Table I) effect of adding a delay to a trial. In addition, an effect of sex was found, where female subjects performed better than their male counterparts, and cage-bred subjects performed better in the trials with no delay than did island-bred monkeys. The latter effect, particularly prominent in the old subjects (Figure 2), probably affected all trials, but was more difficult to distinguish due to a floor-effect, with some subjects approaching chance levels, in the trials where a 30-second delay was utilized.

Table II. *Characteristics of the subjects. The age groups are defined as follows: Young: 4-9 years of age; Middle-aged: 10-19 years of age; Old: more than 20 years of age. Cage-bred animals originated from the Primate Research Center (PRC) breeding facility, whereas island-bred animals were collected from the PRC natural habitat breeding facility on Tinjil island.*

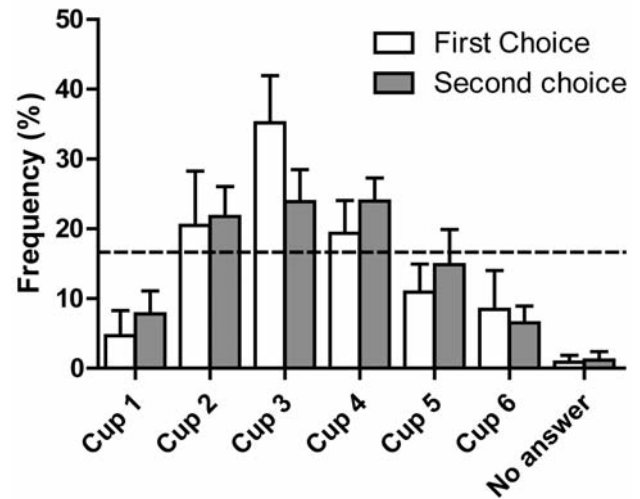| ID (Tattoo)	Age group	Gender	Number of trials	Background
1. FA9103	Young	Female	120	Island-bred
2. C2538	Young	Female	120	Cage-bred
3. T3615	Middle-aged	Female	120	Island-bred
4. T3619	Middle-aged	Female	120	Island-bred
5. I1166	Old	Female	150	Cage-bred
6. I1112	Old	Female	150	Cage-bred
7. C0744	Young	Male	120	Cage-bred
8. C3852	Young	Male	120	Cage-bred
9. T3051	Middle-aged	Male	120	Island-bred
10. T2895	Middle-aged	Male	120	Island-bred
11. T3311	Old	Male	150	Island-bred
12. T3296	Old	Male	120	Island-bred
13. T3281	Old	Male	30	Island-bred
14. T3235	Old	Male	30	Island-bred
15. T3295	Young	Male	30	Island-bred
16. T3285	Young	Female	30	Island-bred
17. T3287	Middle-aged	Male	30	Island-bred
18. C5357	Middle-aged	Female	30	Cage-bred
19. T3107	Old	Male	30	Island-bred
20. C5545	Old	Female	30	Cage-bred
21. T2800	Old	Male	30	Island-bred
22. T3232	Old	Male	30	Island-bred
23. C0168	Old	Female	30	Cage-bred
24. C1980	Old	Female	30	Cage-bred
25. 11085	Old	Female	30	Cage-bred
26. 10749	Old	Female	30	Cage-bred
27. 9661	Old	Male	30	Cage-bred

Figure 4. *Cup preference in the memory load test. The bars represent average frequencies (with inter-individual 95% confidence interval (CI)) for 27 subjects, totaling 1980 trials. The dashed line represents the average frequencies expected if subjects were choosing at random.*
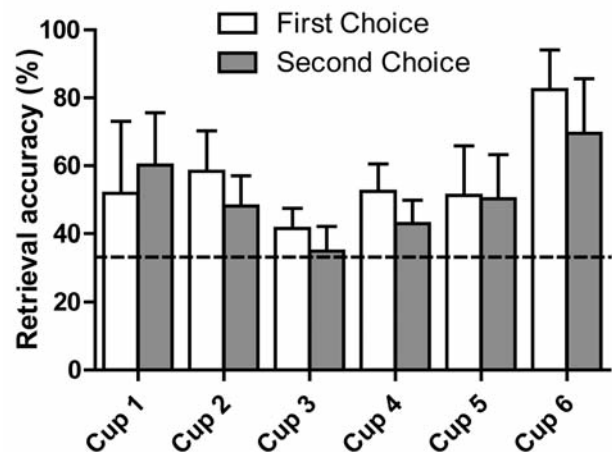

Figure 5. *Retrieval accuracy in the memory load test. The bars represent the average percentage (with inter-individual 95% confidence interval (CI)) of correct answers per chosen cup. The dashed line represents chance level for both choices.*

Finally, an effect of the direction of search – whether the two answers were provided going from left to right, or the opposite – was found to influence the overall retrieval accuracy. During the tests, the monkeys were, regardless of age, better at retrieving the last baited cup – the rightmost of the two baits (Figure 3). A subject who intent on retrieving both baits would, however, probably benefit from recreating the baiting process from memory, chronologically, moving from left to right, picking the baits in the order they were placed in the cups. The present data suggest such a strategy may indeed exist among the tested monkeys. Left-to-right chosen search-pairs were unevenly distributed across the age groups (χ^2=7.3, p<0.05), occurring more frequently among the young monkeys than among the middle aged ones; this was, in fact, the predominant strategy in young monkeys (54% left-to-right chosen pairs in contrast to the 47% of old and middle-aged monkeys). Even though many of the cases have occurred by chance – a monkey choosing truly at random is after all as

likely to go right-to-left as the opposite – there is good reason to assume that some of the data were due to a more deliberate, systematic, memory recall process: The left-to-right search approach yielded significantly better results in retrieving baits (Table III), the difference being an increased retrieval rate of both baits in a trial by approximately 40% (Table IV). Worth noting is that no interaction between directionality and age

Table III. *General Linear Mixed Model (GLMM) describing the factors affecting retrieval accuracy. Note that Age is offered as a multinomial rather than a linear predictor (i.e. the age ranges of Table II are used).The statistical significance (Sign.) is, for each factor, presented as a p-value estimated from a calculated F-statistic presented with between and within group degrees of freedom (df1 and df2, respectively). For the directionality of the effect we have offered an explanation in words rather than displaying the exponential coefficients for increased transparency.*

Factor	F	df1	df2	Sign.	Effect: "Retrieval accuracy…"
Corrected model	36.2	11	34	$p<0.001$	
Age	54.1	2	10	$p<0.001$	"… decreases with age."
Sex	30.1	1	13	$p<0.001$	"… is higher in females than in males."
Origin	2.50	1	12	$p=0.139$	-
Direction	4.84	1	1.97	$p<0.05$	"… is higher going left to right, than the opposite."
Age × Delay	0.59	2	1.97	$p=0.556$	-
Sex × Delay	0.95	1	1.97	$p=0.331$	-
Origin × Delay	5.87	1	1.97	$p<0.05$	"… is lower in wild-caught subjects when no delay is applied."
Direction × Delay	0.57	1	1.97	$p=0.450$	

was found (data from a separate GLMM are not shown). That is: although predominant as a strategy among the young monkeys, choosing baits in a left-to-right order appears to have improved the retrieval accuracy of monkeys, regardless of age.

Whereas the overall test performance was contingent on the subject's age, sex, background/life history and ability to approach the challenge in a systematic fashion, there were distinct patterns to the answers shared by all subjects, suggesting common strategies were employed. All monkeys showed a clear bias toward the middle cups in the trials, regardless of bait positions. Since the baits were perfectly equally distributed across the cups in each module, correct answers, interspersed with perfectly random guesses should align along the dashed line in Figure 4. As this is clearly not the case, it would seem the monkeys gravitated towards the middle cups when choosing. The central tendency in cup preference was not positively correlated with retrieval accuracy however (Figure 5) – in more than a third (average: 35%) of the trials, the subjects chose cup number three with their first guess, but less than half of these choices were found to be correct (average: 42%). This can be contrasted with cup number six, which was initially chosen less than once in every ten trials (average: 8.9%), but these choices turned out to be correct more than four times in five (average: 82%).

In addition to the central tendency, the monkeys tended to provide clustered answers: More than four out of five (82%) of all provided answer-pairs were adjacent cups, which can be contrasted with only one third (33%) of all bait positions being adjacent. Whereas this adjacent-answers bias naturally made the monkeys better at obtaining both baits when they were in adjacent positions, they failed at retrieving any of the baits just as often as they would with non-adjacent baits ($F_{1,1.98}=0.22$, $p=0.64$). This indicates that the monkeys were no better at recalling adjacent positions than they were at recalling non-adjacent positions; the difference simply lies

Table IV. *Contingency table of success rate by directional strategy. All subjects have been pooled for an overview. Directions are given relative to the experimenter and the baiting process was always carried-out going left to right.*

Strategy	Retrieved no bait	Retrieved one bait	Retrieved both baits
Right to left	216 (21%)	656 (65%)	142 (14%)
Left to right	191 (20%)	594 (62%)	181 (19%)

Table V. *The effect of answering correctly on answer adjacency. Only data from trials with non-adjacent cup placements are shown/analyzed.*

	Adjacent answer	Non-adjacent answer
First answer is incorrect	572 (87%)	82 (13%)
First answer is correct	452 (71%)	186 (29%)

in the answering strategy. The central tendency in guessing (Figure 4) will naturally increase the odds for adjacent answer-pairs. But if we assume that the guesses were independent of one another, the resulting frequency of adjacent answer-pairs accounted for is still only 45%. Clearly the answers were not independent of one another and the high level of answer adjacency cannot be attributed solely to the central tendency. This strategy may instead in part be a simple search pattern where the monkey picks the neighboring cup following a mistaken answer. Looking at only the non-adjacent bait positions, when the monkeys provided a correct answer on the first attempt, the incidence of non-adjacent answers approximately doubled (Table V; Relative risk 95% confidence interval (CI): 1.84-2.94; $\chi^2=54$,
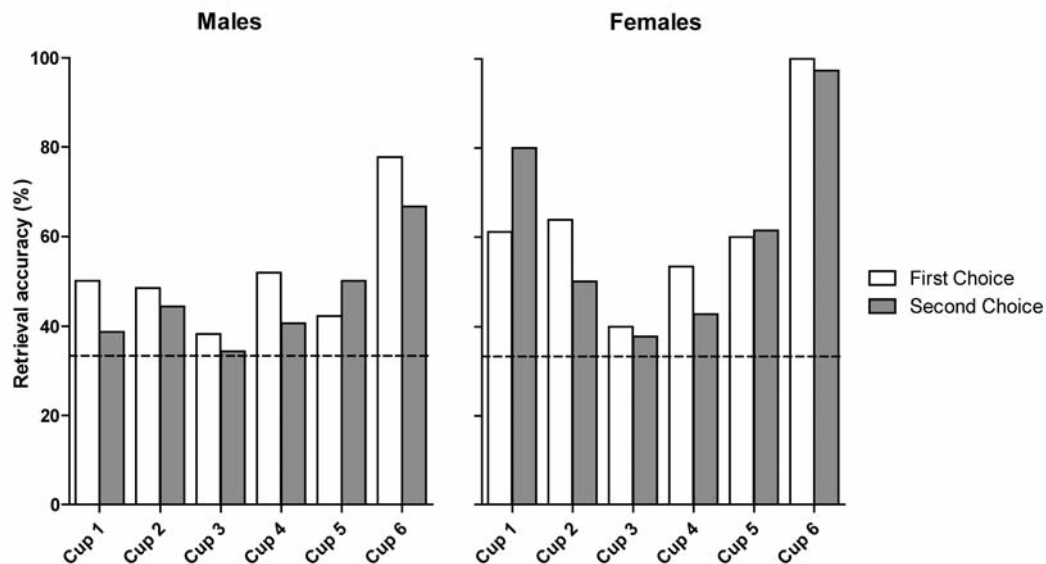
Figure 6. *Median retrieval accuracy of subjects sub-divided by sex. The pattern clearly differs between sexes, with the retrieval accuracy increasing towards both ends of the array in a clear U-shape for the females, whereas a clear increase can only be seen on the right-hand side for the males. The dashed line represents chance level.*

$p<0.001$). Since correct answers in part abolished adjacent answers (for non-adjacent baits), it is fair to assume that they were, at least in part, due to a natural search pattern where the monkeys adjust their initial guesses. In the case of a correct first answer, however, the subjects would still choose adjacent cups in 71% of the (non-adjacent bait) trials, which is considerably higher than the expected 45% (accounting for central tendency).

A complete data set, comprising the raw data from all 1,980 trials, has been made available under a free-to-use license (refer to the section on Additional Online Materials).

Discussion

While all subjects performed at higher-than-chance levels, it is clear that the task was challenging for cynomolgus monkeys of any age. Even the best performing subjects could not reliably retrieve both baits under the best of conditions, and with an added 30-second delay, the old monkeys' performance barely exceeded chance levels. It is clear from the data that not only the age of the subjects predicted their test performance; sex and life history seemed to have a significant influence as well.

A study highlighting sex-differences in rhesus monkeys solving spatial tasks has found, much like the present study, that females would out-perform their male counterparts (19). This was hypothesized to be due to a superior ability of the females in relating baited positions to (local) landmark cues. Although this study is the only one we are aware of

investigating the phenomenon in primates, the theory is supported by studies in rodents and humans, where it has been shown that males tend to rely highly on distal cues for spatial orientation, whereas females tend to be more flexible, but primarily use local cues (20). In the present study, the high retrieval accuracy of baits placed at the ends of the array (cups 1 and 6 – refer to Figure 5) mirrors previous findings (15) in great apes. As with apes, we suggest that this pattern emerges due to the subjects' using landmark cues to locate the hidden baits – *i.e.* the end cups are more easily remembered/retrieved by their not being flanked on both sides by other cups or by relating their positions to the edges of the tray. When examining the retrieval accuracy in the present study by sex (Figure 6), we find support for the hypothesis proposed by Herman and Wallen (19), also in cynomolgus monkeys. Whereas all monkeys excelled at retrieving baits from the last-baited cup of the array (cup 6), the females were also proficient at retrieving baits placed at the other extreme end of the array (cup 1); their male counterparts were not. This, in turn, can be seen as a symptom of the male subjects being worse at utilizing local landmark cues, which may help explain the overall performance difference between sexes.

Why cage-bred monkeys would perform better overall than did the island-bred monkeys can be debated. An explanation could be that cage-bred monkeys were more accustomed to working with an experimenter, having a long-standing history of interacting with personnel during their daily feeding and cage cleanings. They would also be more

adjusted to the tools and treats used, to the environment, to being subjected to arbitrary rules *etc*., which may have made them ultimately perform better. The effect of enculturation has been discussed in length with greater apes and, although the concept is a source of great controversy and the term itself may not be a good fit for monkeys (21), there is some evidence for monkeys being better-equipped for tackling human-designed games if they have lived in close proximity to people (22, 23). In addition, the island-bred subjects can essentially be considered wild-caught, and may thus have found the testing experience – in particular, interacting with an experimenter – to be much more stressful than did the cage-bred subjects. They may furthermore have been subjected to prolonged stress, if we assume that they adapted poorly to captivity. Acute stress has been shown to negatively impact cognitive test results in monkeys (24-26) and chronic stress is well-known to affect the hippocampus (27), which plays a crucial role in solving delayed response tasks.

Whereas the overall performance differed greatly between monkeys of different sex, of different ages and from different breeding facilities, the strategies and resulting answering patterns were remarkably ubiquitous. The central tendency in cup preference – although most striking – appears to be the result of guesswork. When not remembering the location of a bait, but prompted to make a choice, it seems that the subjects would choose a centrally placed cup. This may be simply due to convenience, as the central cups would be physically nearest to the subject when being centered across from the tray. Alternatively, if the subjects found it, in some trials, hard to distinguish individual cups, but rather treated the entire array as an object; it would be natural for them to aim for the center of the "object."

It appears, however, that the guesswork exhibited by the monkeys was not without some rationale. The high incidence of adjacent answers mirrors findings from similar studies, primarily involving great apes. Call (13) noted that, when prompted to provide multiple answers – to retrieve two hidden baits from a linear array of possible caches – great apes and human children were more likely to provide adjacent answers than could be explained by chance. It was suggested that the pattern arose from the subjects' poor inhibition control – the positive trait of the baited caches "carrying over" to their neighbors. This was disputed by Beran and collaborators (14) who, with a larger array, could not find evidence to support the theory in chimpanzees. Instead, adjacent answers were explained as the results of incorrect memory recall. The merits of this reasoning were acknowledged by Barth and Call (28), who nevertheless pointed-out that the disproportionate number of adjacent answers probably had a multi-factorial explanation as poor memory recall alone could not explain the answering patterns.

Providing mostly adjacent answers did not, in an obvious way, improve the monkeys' overall retrieval accuracy in the trials. The pattern does, however, betray a conscious strategy: The fact that the frequency of adjacent answers was conditional on whether the first answer was a success suggests that the monkeys were able to improve on an incorrect answer by guessing at an adjacent cup. This helps explain the clustering of answers seen in these tests. However, even accounting for this adaptive search pattern and the central tendency in answers, the frequency of adjacent answers (for non-adjacent baits) was far higher than could be expected by chance. The remaining unaccounted-for adjacent answers cannot be explained using the memory hypothesis of Beran *et al*. (14) but may well be an expression of poor inhibition control.

The central tendency in answers combined with the disproportionately high number of adjacent answers suggests a high level of impulsivity. The monkeys tended to answer rashly, reaching for cups that were physically proximal to them once the tray was placed within their reach; a majority of first answers were provided within the first two seconds (unfortunately only a subset of the trials were filmed, thus preventing a formal analysis). The rashness is a well-known characteristic of the cynomolgus monkey: Amici *et al*. (4) previously reported that cynomolgus monkeys would, when challenged with a simple task of retrieving a food item from behind a translucent screen, unlike many closely-related primates, more often than not, attempt to go through the screen, rather than first finding an appropriate hole or door allowing them to reach the bait. They also found that cynomolgus monkeys were poor at postponing gratification. The answering strategy employed by the subjects in the present study may also, in part, be related to the social structure of the cynomolgus monkeys. The animal caretakers identified most of the studied subjects to be socially subordinate, based on their behavior. Living in a highly despotic, strict dominance hierarchy (29), alpha males/females have first pick when it comes to food items available to the colony. This, in turn, may force the lower ranked animals – comprising the majority of the colony – to adjust their strategy in retrieving the available food. Priority may be given to acting fast, possibly at the expense of accurate decision making. In future studies, it would be interesting to see if social rank, as well as sex, age and breeding facilities, have an influence on the overall success in similar cognitive tests. It is also worth noting that the monkeys were not overly penalized for guessing in the MLT paradigm, as they were more likely to obtain at least one bait per trial, than not (pure guesswork would yield an average of 0.66 baits per trial). It can be speculated whether this may have inflated the level of guesswork by making the subjects somewhat complacent.

With the array of cups being visible throughout the trial – as opposed to similar tests where the array is hidden in the interim between baiting and challenge – non-mnemonic

solutions to the test can be utilized. The subject could place itself, spatially, in front of one of the baited cups, thus eliminating the need for correctly recalling its position (30). As previously concluded by Zola and collaborators (31) in a delayed non-matching to sample task, however, no evidence for a reliance on similar techniques could be found in the cynomolgus monkeys. The purported prominent use of landmark cues does not, in itself, rule out the possibility that the subjects also utilized egocentric or non-mnemonic strategies. But, even though the positioning of the subjects during the individual trials was not recorded, we would, nevertheless, expect a high reliance on these to produce a much more uniform retrieval accuracy, at least for the first chosen cups, than is evident from Figure 5 and 6. In addition, it was noted that subjects would frequently move laterally to the array in the testing cage during trials, thus making non-mnemonic strategies reliant on their own position hard to implement.

Data further suggest that the acts of remembering (/recalling) and retrieving should be separated for a fruitful discussion. The cups at the ends of the array were baited just as often as the central cups, but the monkeys seemed to direct their guesses toward the ends much less frequently. This is in stark contrast to these cups being associated with the highest retrieval accuracy. In order to retrieve a hidden bait, the subject needed to first recall its position and then correctly identify the remembered position in the array. Evidence suggests that it is the latter task that is easier with the cups at the extreme ends of the array. A possible interpretation of the data is thus that the monkeys do not find it easier remembering the outermost cups – they may even have more trouble remembering them – but when they do, the outermost cups are more easily (correctly) retrieved due to landmark cues.

Furthermore, since the directionality of search determines the retrieval accuracy, we theorize that the monkeys may be capable of recreating the baiting process from memory for improved results. Humans remember the order in which events occurred, using multiple cognitive mechanisms, while it has been suggested that adult rhesus monkeys, rely heavily on a mechanism that appears to encode order of occurrence relative to other events, rather than in absolute time (32). With aging, this mechanism, suggested to be controlled by the hippocampus (33), deteriorates. This agrees with our findings (Figure 3) where young and middle-aged subjects, but not old subjects, seem to gravitate toward retrieving the last hidden bait. Serial position effects, where the subjects adapt to take advantage of recent memories being stronger, have previously been noted in macaques (34-36).

In conclusion, the cynomolgus monkeys exhibited common strategies in solving the MLT. Impulsive guesswork was highly prevalent in our studies but, looking beyond this, we found evidence of the use of landmark cues, of adaptive answering and possibly evidence of systematic (/chronological) memory recall.

Whereas we can expect these strategies to have influenced the overall MLT success rate, individual characteristics – age, sex and which breeding facility the subjects originated from – appeared to exert a clearer effect on the monkeys' overall performance. Clearly, subject characteristics and life history need to be considered when managing test variation and assay capability and when interpreting data from cognitive tests on cynomolgus monkeys.

Ethical Standards

Conflicts of Interest

The Authors declare that they have no conflicts of interest.

Acknowledgements

Additional Online Materials

A complete data set comprising all trials has been made available online under a free-to-use license together with films demonstrating the MLT procedure. They can be accessed using the links listed below.
Complete data set: http://dx.doi.org/10.6084/m9.figshare.1038346
Testing procedure without delay, example 1:
http://dx.doi.org/10.6084/m9.figshare.1038348
Testing procedure without delay, example 2:
http://dx.doi.org/10.6084/m9.figshare.1038349
Testing procedure with delay, part 1:
http://dx.doi.org/10.6084/m9.figshare.1038354
Testing procedure with delay, part 2:
http://dx.doi.org/10.6084/m9.figshare.1038355

References

1   Nagahara AH, Bernot T and Tuszynski MH: Age-related cognitive deficits in rhesus monkeys mirror human deficits on an automated test battery. Neurobiol Aging *31(6)*: 1020-1031, 2010.

2   Rodriguez JS and Paule MG: Working Memory Delayed Response Tasks in Monkeys. *In*: Methods of Behavior Analysis in Neuroscience, 2nd ed (Buccafusco JJ ed.). Boca Raton, CRC Press, 2009.

3   Hara Y, Rapp PR and Morrison JH: Neuronal and morphological bases of cognitive decline in aged rhesus monkeys. Age *34(5)*: 1051-1073, 2012.

4 Amici F, Aureli F and Call J: Fission-fusion dynamics, behavioral flexibility, and inhibitory control in primates. Curr Biol *18(18)*: 1415-1419, 2008.

5 Call J: Understanding apes to understand humans: the case of object-object relations. *In*: Towards a Theory of Thinking. Springer, pp. 215-230, 2010.

6 Heuer E and Bachevalier J: Effects of Selective Neonatal Hippocampal Lesions on Tests of Object and Spatial Recognition Memory in Monkeys. Behav Neurosci *125(2)*: 137-149, 2011.

7 Tsujimoto S and Postle BR: The prefrontal cortex and oculomotor delayed response: a reconsideration of the "mnemonic scotoma". J Cogn Neurosci *24(3)*: 627-635, 2012.

8 Darusman H, Call J, Sajuthi D, Schapiro S, Gjedde A, Kalliokoski O and Hau J: Delayed response task performance as a function of age in cynomolgus monkeys (*Macaca fascicularis*). Primates *55(2)*: 259-267, 2013.

9 Darusman HS, Sajuthi D, Kalliokoski O, Jacobsen KR, Call J, Schapiro SJ, Gjedde A, Abelson KS and Hau J: Correlations between serum levels of beta amyloid, cerebrospinal levels of tau and phospho tau, and delayed response tasks in young and aged cynomolgus monkeys (*Macaca fascicularis*). J Med Primatol *42(3)*: 137-146, 2013.

10 Darusman HS, Pandelaki J, Mulyadi R, Sajuthi D, Putri IA, Kalliokoski OH, Call J, Abelson KS, Schapiro SJ and Gjedde A: Poor memory performance in aged cynomolgus monkeys with hippocampal atrophy, depletion of amyloid Beta 1-42 and accumulation of tau proteins in cerebrospinal fluid. In Vivo *28(2)*: 173-184, 2014.

11 Amici F, Aureli F and Call J: Monkeys and apes: are their cognitive skills really so different? Am J Phys Anthropol *143(2)*: 188-197, 2010.

12 Hertzog C, Dixon RA, Hultsch DF and MacDonald SW: Latent change models of adult cognition: are changes in processing speed and working memory associated with changes in episodic memory? Psychol Aging *18(4)*: 755-769, 2003.

13 Call J: Object permanence in orangutans (*Pongo pygmaeus*), chimpanzees (Pan troglodytes), and children (Homo sapiens). J Comp Psychol *115(2)*: 159-171, 2001.

14 Beran MJ, Beran MM and Menzel CR: Spatial memory and monitoring of hidden items through spatial displacements by chimpanzees (Pan troglodytes). J Comp Psychol *119(1)*: 14-22, 2005.

15 Hribar A and Call J: Great apes use landmark cues over spatial relations to find hidden food. Anim Cogn *14(5)*: 623-635, 2011.

16 Moss MB, Moore TL, Schettler SP, Killiany R and Rosene D: Successful *vs*. Unsuccessful Aging in the Rhesus Monkey. *In*: Brain Aging: Models, Methods, and Mechanisms. Riddle DR (ed.). Boca Raton, CRC Press, pp. 21-38, 2007.

17 Bowen W and Koch G: Determination of age in monkeys (Macaca irus) on the basis of dental development. Lab Anim *4(1)*: 113-123, 1970.

18 Swindler DR: Primate dentition: an introduction to the teeth of non-human primates. Cambridge University Press, 2002.

19 Herman RA and Wallen K: Cognitive performance in rhesus monkeys varies by sex and prenatal androgen exposure. Horm Behav *51(4)*: 496-507, 2007.

20 Sandstrom NJ: Sex differences in use of visual cues by rhesus monkeys performing a spatial learning task: Comment on "Cognitive performance in rhesus monkeys varies by sex and prenatal androgen exposure" by Herman and Wallen. Horm Behav *52(2)*: 139-142, 2007.

21 Bering JM: A critical review of the "enculturation hypothesis": the effects of human rearing on great ape social cognition. Anim Cogn *7(4)*: 201-212, 2004.

22 Fredman T and Whiten A: Observational learning from tool using models by human-reared and mother-reared capuchin monkeys (Cebus apella). Anim Cogn *11(2)*: 295-309, 2008.

23 van de Waal E and Bshary R: Contact with Human Facilities Appears to Enhance Technical Skills in Wild Vervet Monkeys (Chlorocebus aethiops). Folia Primatol *81(5)*: 282-291, 2010.

24 Arnsten AFT and Goldman-Rakic PS: Noise stress impairs prefrontal cortical cognitive function in monkeys – Evidence for a hyperdopaminergic mechanism. Arch Gen Psychiat *55(4)*: 362-368, 1998.

25 Lyons DM, Lopez JM, Yang C and Schatzberg AF: Stress-level cortisol treatment impairs inhibitory control of behavior in monkeys. J Neurosci *20(20)*: 7816-7821, 2000.

26 Lyons DM, Yang C, Eliez S, Reiss AL and Schatzberg AF: Cognitive correlates of white matter growth and stress hormones in female squirrel monkey adults. J Neurosci *24(14)*: 3655-3662, 2004.

27 Uno H, Tarara R, Else JG, Suleman MA and Sapolsky RM: Hippocampal Damage Associated with Prolonged and Fatal Stress in Primates. J Neurosci *9(5)*: 1705-1711, 1989.

28 Barth J and Call J: 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). J Exp Psychol Anim Behav Process *32(3)*: 239-252, 2006.

29 Riddick NV, Czoty PW, Gage HD, Kaplan JR, Nader SH, Icenhower M, Pierre PJ, Bennett A, Garg PK and Garg S: Behavioral and neurobiological characteristics influencing social hierarchy formation in female cynomolgus monkeys. J Neurosci *158(4)*: 1257-1265, 2009.

30 Tinklepaugh OL: An experimental study of representative factors in monkeys. J Comp Psychol *8(3)*: 197, 1928.

31 Zola SM, Squire LR, Teng E, Stefanacci L, Buffalo EA and Clark RE: Impaired recognition memory in monkeys after damage limited to the hippocampal region. J Neurosci *20(1)*: 451-463, 2000.

32 Templer VL and Hampton RR: Cognitive mechanisms of memory for order in rhesus monkeys (*Macaca mulatta*). Hippocampus, 2012.

33 Fortin NJ, Agster KL and Eichenbaum HB: Critical role of the hippocampus in memory for sequences of events. Nature Neurosci *5(5)*: 458-462, 2002.

34 Hampton RR: Rhesus monkeys know when they remember. Proc Natl Acad Sci *98(9)*: 5359-5362, 2001.

35 Hampton RR and Hampstead BM: Spontaneous behavior of a rhesus monkey (*Macaca mulatta*) during memory tests suggests memory awareness. Behav Processes *72(2)*: 184-189, 2006.

36 Tanaka A and Funahashi S: Macaque monkeys exhibit behavioral signs of metamemory in an oculomotor working memory task. Behav Brain Res *233(2)*: 256-270, 2012.