

TOOLS • TREATS • TOYS

©Copyright Marusha Dekleva 2011

Dekleva, Marusha

TOOLS, TREATS, TOYS

What human and non-human primates recall about their past and plan for their future

PhD-Dissertation, Utrecht University, 2011

ISBN

978-90-393-5615-9

Cover design: Claude Maba Kudivwila, Gesa Webering, Brigitte Goossens and Marusha Dekleva

Layout: Dan Christensen and Marusha Dekleva

Printing: Ipskamp Drukkers, Enschede, NL

TOOLS • TREATS • TOYS

**What human and non-human primates recall about their past and
plan for their future**

GEREEDSCHAPPEN • BELONINGEN • SPEELTJES

**Wat mensen en apen zich herinneren van hun verleden en wat ze
voor hun toekomst uitstippelen**

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag
van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit
van het college voor promoties in het openbaar te verdedigen op woensdag
21 september 2011 des middags te 12.45 uur.

door
Marusha Dekleva
geboren op 16 januari 1979
te Ljubljana, Slovenië

Promotor: Prof. dr. B.M. Spruijt

Co-promotor: Dr. E.H.M. Sterck

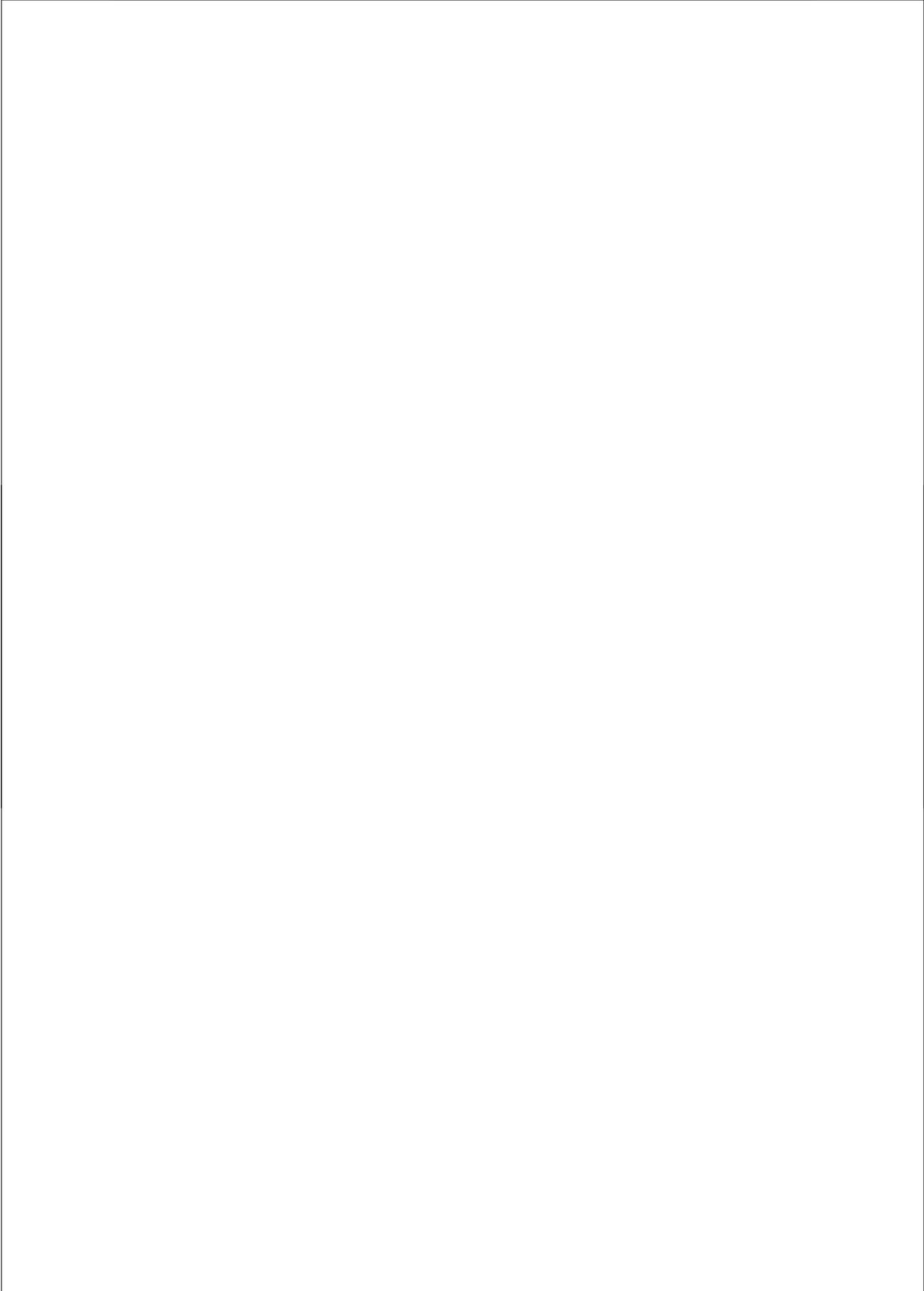
This thesis was accomplished with financial support from The Netherlands Organisation for Scientific Research (NWO).

Mojim staršem

Til oldefar

Mpo na Claude

“It's a poor sort of memory that only works backwards.”
-Lewis Carroll



SUMMARY ENG/NL 9/17

INTRODUCTION 25

TOOLS 45

Transport of Tools for Future Use by Long-tailed Macaques (*Macaca fascicularis*)

TREATS 71

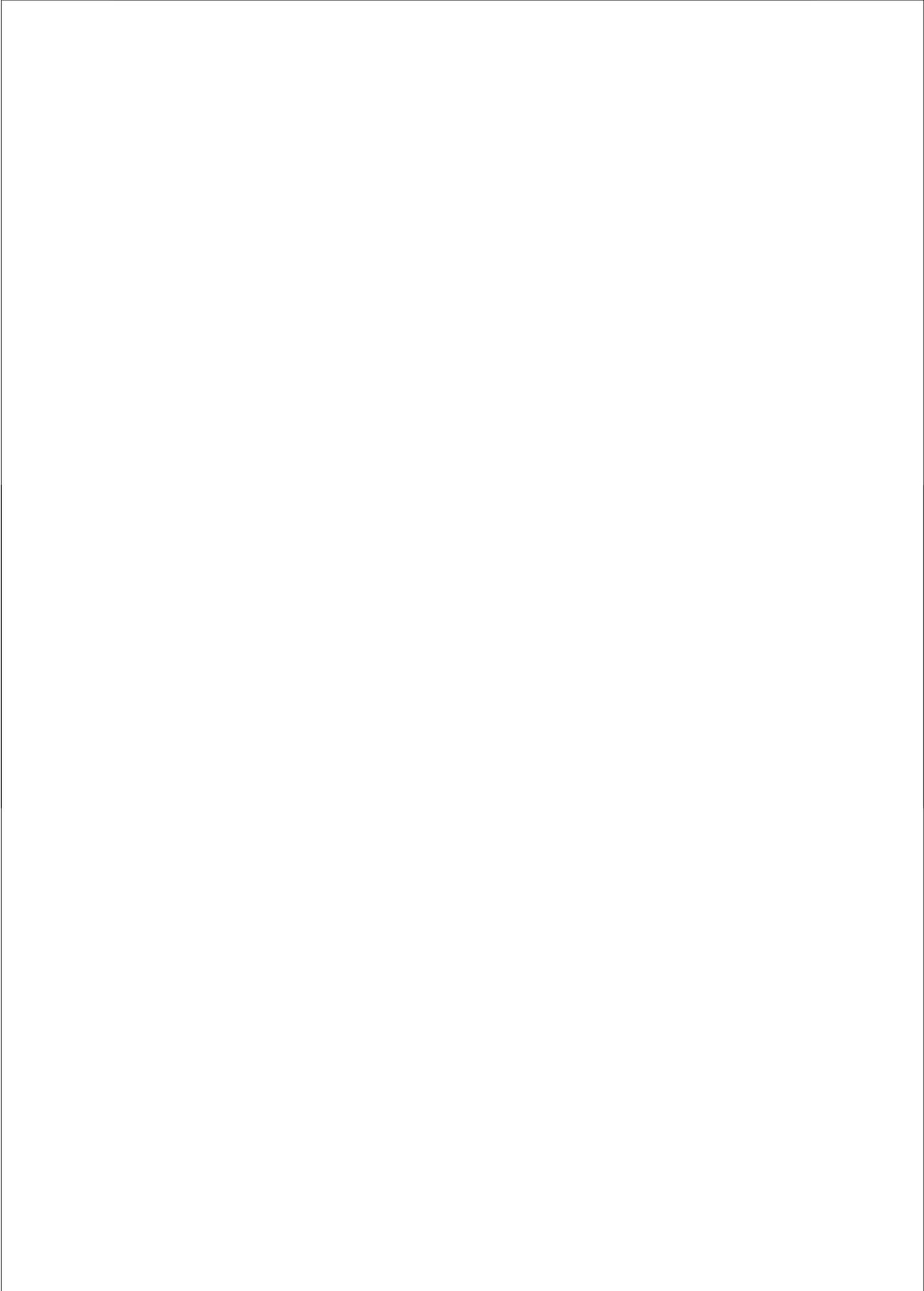
Failure at a What-Where-When Task Results in Obtaining Rewards by Means of a Location-Based Association Strategy in Chimpanzees (*Pan troglodytes*)

TOYS 93

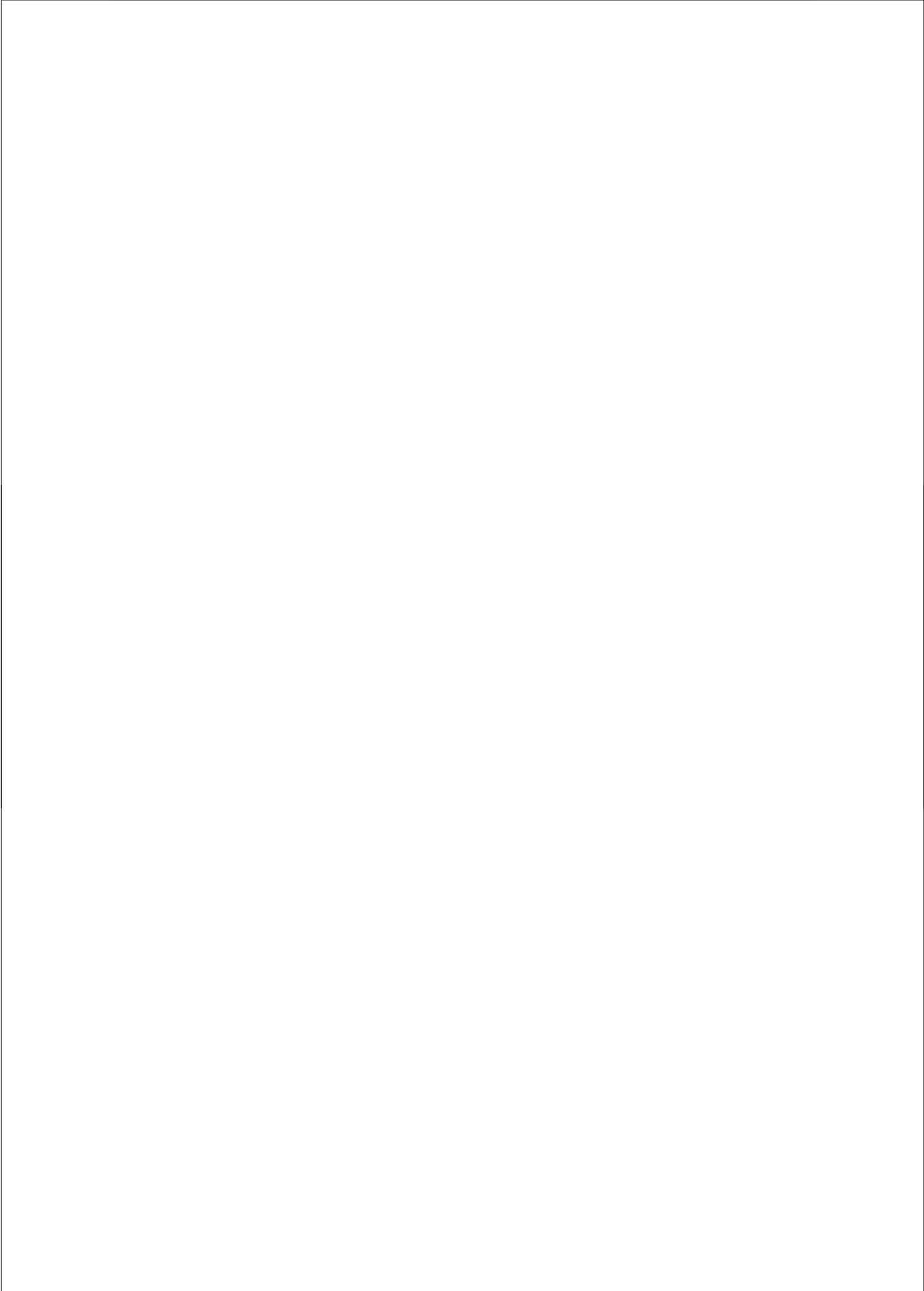
The Developmental Onset of Integrated What-Where-When Memory and Theory of Mind in Children

DISCUSSION 119

ETC... 135



Summary ENG



INTRODUCTION

Mental or cognitive time travel enables humans to remember details of past events by means of episodic memory, as well as to prepare for potential future events by means of future planning. These retrospective and prospective abilities, respectively, require language to be assessed and have therefore been considered unique to humans since they were first described in the 1970s.

However, these abilities may not require language to be expressed and may be relied on by other animal species in various degrees. Applying a behavioural paradigm would enable investigating the levels to which these abilities may be present in other species, and help to reveal the evolutionary history of these abilities. In fact, in the past 20 years it has been shown that certain species of corvids and rodents have some prospective and retrospective abilities as revealed by their behaviour in carefully designed experiments. These behavioural paradigms have circumvented the inaccessible subjective aspects of human mental time travel, and have enabled the comparative testing of both human and non-human primates.

Surprisingly, the prospective and retrospective abilities of different primate species have only been investigated by a few studies and not systematically, hindering a comprehensive comparison across the different species. As non-human primates are our closest phylogenetic relatives and have some advanced cognitive capacities, it is relevant to systematically address their abilities also in the domain of past and future knowledge.

In this thesis, I experimentally examined prospective and retrospective abilities in different primate taxa to obtain a more comprehensive understanding of the shared and unique aspects of different species complex cognitive abilities from a behavioural perspective. I utilized methods successfully used in other species and attempted to fill in gaps in our knowledge concerning the presence of these abilities in primates. I took both a comparative and a developmental approach and examined the prospective abilities of a representative species of monkeys, the retrospective abilities of a representative species of great apes, as well as the ontogeny of the retrospective abilities in children. Investigating what different species of primates know about their environment and how they deal with complex problems concerning their past and their future, may eventually lead to unravelling the evolutionary history of prospective and retrospective abilities in primates.

TOOLS

In this chapter, I presented a series of experiments investigating the prospective abilities of long-tailed macaques (*Macaca fascicularis*) by requiring them to transport tools for future use. This method has previously been successfully used with several great ape species and therefore, my study allows for a comparison of the prospective abilities of a representative monkey species with those of great apes.

In the first experiment, I tested whether macaques would select a tool and after a delay, transport and use it to obtain rewards. I manipulated the spatial and temporal distance between tool provision and reward availability in several tests, however, all six macaques failed to transport the tool in all tests. In the second experiment, I trained five of the macaques to transport the tool to obtain immediate rewards, and then tested their tool transport abilities after increasing delays. Three of the macaques were able to transport and use the tool after a delay of 5 minutes; one was successful after 10 minutes and even transported the tool on a few trials after a 20 minute delay. In the third experiment, I tested the ability of two of the macaques to transport novel tools for future use. The best performer from the second experiment successfully selected and transported novel tools.

The macaques were unable to solve the task in the first experiment, and failed to meet the criteria for planning in non-humans. They may have been unable to foresee the appropriate sequence of behaviours required to solve the task. After training, the animals displayed flexibility by successfully transporting tools after delays. Remarkably, one macaque also generalized this behaviour by transporting appropriate novel tools, and several attempted to use non-provided objects or their tails to obtain the rewards. The tool transport training provided the macaques with the correct sequence of behaviours to solve the task, and additionally seemed to open new avenues for them to flexibly display the shaped behaviour, indicating that macaques may have some prospective abilities, although inferior to those of great apes.

TREATS

While we know that great apes possess some prospective abilities and can plan for certain future events, their retrospective abilities remain sparsely investigated. In this chapter, I examined the retrospective abilities of a representative great ape species namely, the chimpanzee (*Pan troglodytes*). I used a behavioural paradigm requiring the accurate integration of several features, namely the what, where and when, of past food hiding events. This retrospective ability, also known as what-where-when or episodic-like memory is most conclusively established by the so-called three step method. In this method, three food types of different palatability and different rates of disappearance are hidden in trial unique locations and can be obtained by the chimpanzees after three different retention intervals.

The chimpanzees observed the baiting of two locations amongst four, and chose one after a given delay (15min, 1h or 5h). I used two combinations of food types, a preferred and a less preferred food that disappeared at different rates. The individuals had to base their choices on the time elapsed since baiting, and on their memory of which food was where. Depending on the trial, they could recover either their preferred food or the one that remained present. Evaluating the chimpanzees' ability to integrate the what-where-when features of the food hiding events, all eight individuals did not adjust their choices based on what was hidden where, nor integrate

the time passed since hiding and thus, did not show what-where-when memory in the domain of food. In fact, they failed to use any feature of the baiting events to guide their choices. Nonetheless, their choices were not random, but the result of a simple, developed location-based association strategy. The chimpanzees choose locations that delivered most rewards earlier in the testing sequence and obtained a fair amount of food.

The animals' failure to remember the what-where-when in the presented set-up may be due to the complexity of the task, rather than an inability to form episodic-like memories, as they did not use any form of memory to solve this task. The extent of the chimpanzee retrospective abilities as reflected by their behaviour in this what-where-when task, thus, remains inconclusive. However, their behaviour reveals an interesting strategy of how animals, when exposed to a multitude of changing features through-out the testing sequence (what, where and when), focused on the stable locations of the foods and, by means of a simple strategy, obtained apparently sufficient rewards.

TOYS

In this chapter, I examined the development of retrospective abilities in three and five year old children. I designed a story test based on the principles of the three step what-where-when method originating from the corvid studies and used with chimpanzees in my previous study. What-where-when or episodic-like memory is comparable to the content requirement of episodic memory in humans, however, episodic memory is said to additionally require auto-noetic awareness. Auto-noetic awareness is the ability to consciously re-experience a past episode including the knowledge that it is oneself who has experienced the remembered event. Whether what-where-when memory relies on auto-noetic awareness is however unclear, because this cannot be assessed in non-verbal species. Auto-noetic awareness develops in children between three and five years and is thought to mediate the co-development of episodic memory and Theory of Mind, which is the ability to distinguish own beliefs and thoughts from those of others.

I evaluated the children's performance on a battery of Theory of Mind tests in order to compare its developmental onset with what-where-when memory. This is to my knowledge the first study to look at the development of what-where-when memory with a three step method in children, and to relate this developmental onset to that of Theory of Mind abilities. The what-where-when story test included several questions that allowed me to examine whether the children could retrieve the what, where and when components separately, or in an integrated manner.

Around the age of four, the children passed the "what was where" question, thus, before passing the Theory of Mind tests. Both the "when" question and the integrated "what-where-when" question were, however, first passed after the Theory of Mind tests around the age of six. Performance on the Theory of Mind tests did not predict performance on the what-where-when memory task. Thus suggesting that

performance on this what-where-when task may not rely on Theory of Mind, which episodic memory was previously shown to rely on. However, the relationship between what-where-when and episodic memory remains to be directly examined. The tested what-where-when memory task may be especially demanding for children due to its reliance on understanding of the temporal feature, which may not develop before the age of five or six.

DISCUSSION

The aim of the present thesis was to examine the prospective and retrospective abilities of human and non-human primates from a comparative and a developmental perspective. Placing my findings in context with what was previously known about the prospective and retrospective abilities of primates seems to suggest that primate taxa may differ in their abilities. Although with such a small number of studies conducted and only a few individuals tested in each study, one should be cautious with reaching premature conclusions.

My results show the **children** begin passing the three step what-where-when test around the age of five and the failure of younger children seems to be caused by an inability to integrate the “when” feature. Previous findings have established that around the age of five children also begin passing verbal episodic memory tests. Concerning their prospective abilities, one study found that while three year olds were able to transport a tool in order to obtain immediate rewards, children were not able to transport a tool to obtain rewards in the future until the age of four. Taken together these findings suggest that temporal displacement is the demanding and late developing feature for children concerning both retrospection and propection. Future testing is required to examine how children further develop their prospective abilities over increasing periods and how and when they develop the ability to formulate complex plans.

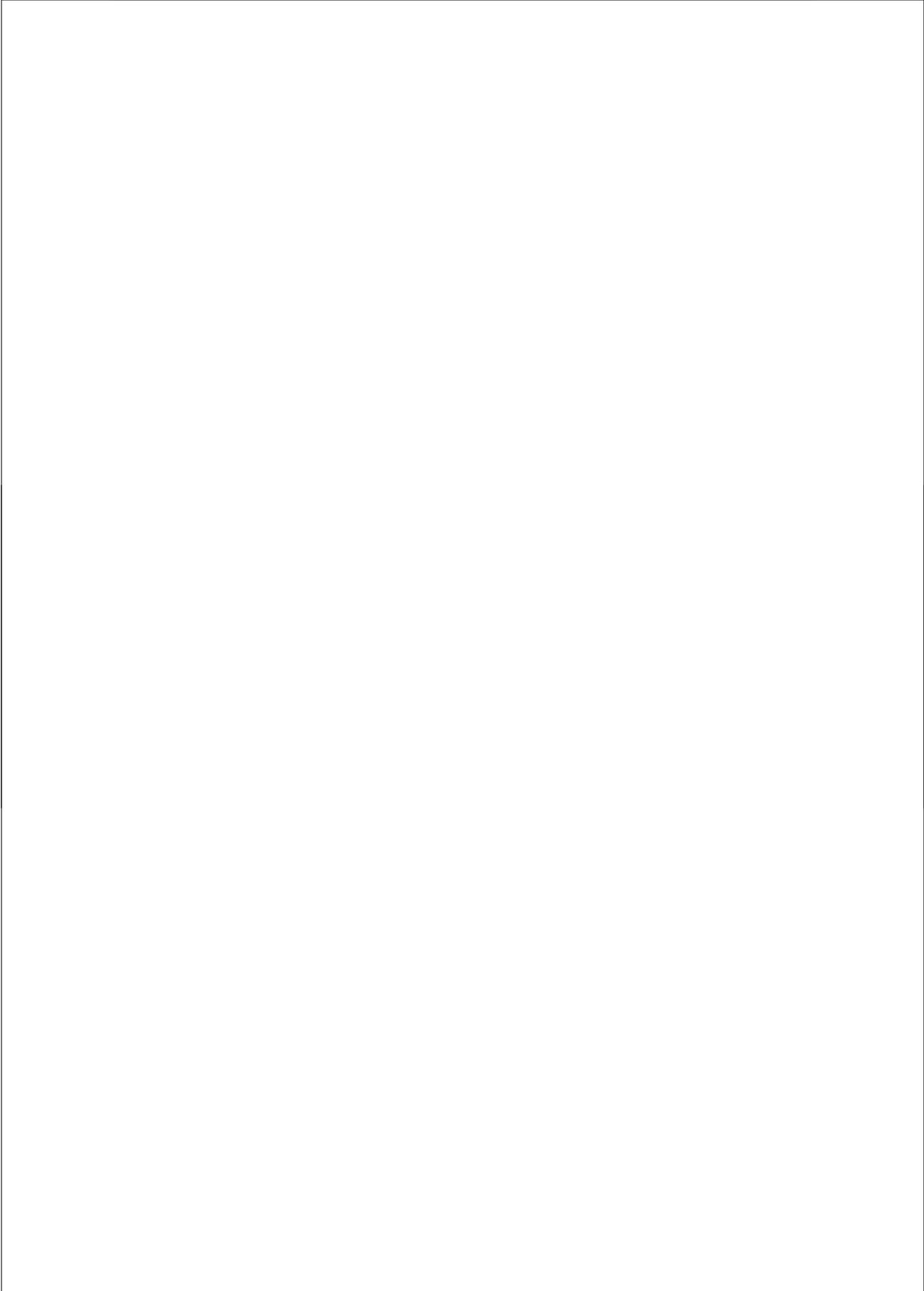
Evaluating the retrospective abilities of great **apes**, chimpanzees, tested with the three step method in my study did not show what-where-when memory and did not use their long-term memory altogether. Thus, the extent of their retrospective abilities is at present inconclusive. Nevertheless, previous findings have established that chimpanzees can remember information for extended periods and can remember baited locations for up to three days. Concerning propection, studies have shown that great apes can select and transport tools even after delays of 14 h and can transport novel tools to use then an hour later to obtain rewards. Based on current data, apes may have better prospective than retrospective abilities. However, this may be a consequence of more research effort invested into systematically investigating their prospective rather than their retrospective abilities and should be addressed by future studies.

Concerning the prospective abilities of **monkeys**, the long-tailed macaques I tested failed to meet the criteria for planning, despite reduction of the temporal

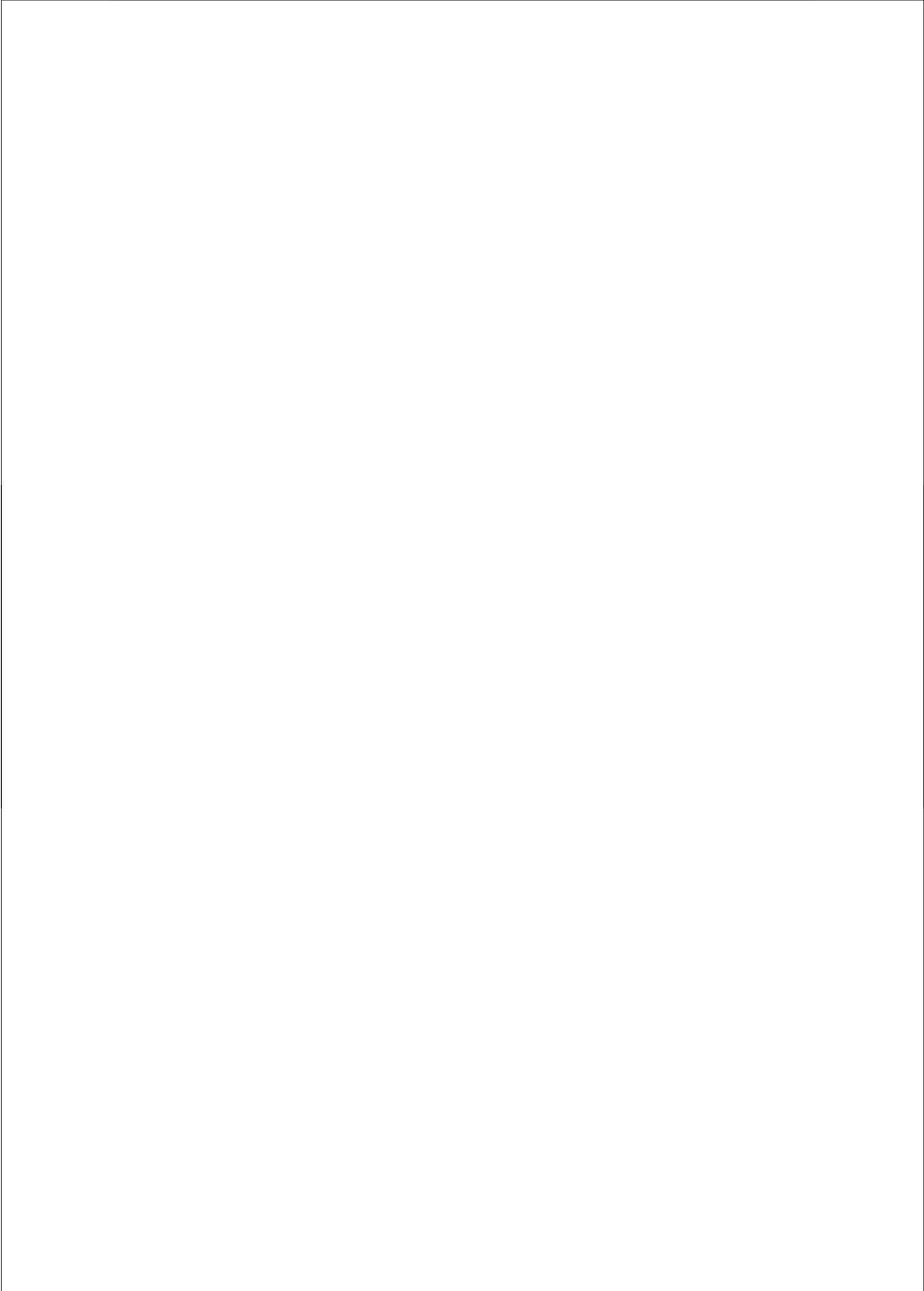
component used with apes from 1 hour to 5 minutes. However, once experience with tool transport was obtained, the macaques successfully completed the task. Interestingly, the best performing monkey was, after training, successful on just a few trials after a 20 min delay, suggesting that macaques' potential prospective abilities could be limited to a much smaller time window than those of apes, scrub jays and humans. Prior studies have shown that rhesus macaques failed to integrate the temporal component in both prospective and retrospective tests, however due to methodological issues these results should be reassessed. Current data suggests that monkeys seem to have limited prospective and retrospective abilities, possibly due to difficulties with integrating the "when" feature. Further studies are required to establish whether monkeys are able to show some, although simpler, future oriented behaviours, as suggested by my findings, and whether their performance would improve in tasks where less complex actions are required to bridge time gaps.

Concerning the prospective and retrospective abilities of different primate taxa as reflected in their behaviour it seems that both monkeys and young children have difficulty with the "when" feature. Children of about five and apes can integrate the temporal displacement of a future event, while my study suggests that long-tailed macaques may have simpler prospective abilities limited to a shorter time frame compared to apes and humans. My results also suggest that children of about five can integrate the what-where-when of a past event, while younger children can only remember the "what" and "where" features of past events. Concerning the retrospective abilities of apes, the present results are inconclusive on their ability to match the children on remembering integrated what-where-when memories, while it currently seems that monkeys may only be able to remember the "what" and "where" features of past events. Further testing is required before we can confirm the indicated differences between the primate taxa and further examine what are their causes. The studies presented in this thesis have merely scratched the surface of examining the prospective and retrospective abilities of human and non-human primates, and the debate on animal time travel is far from finished.

Future studies should move the focus from searching for an existence proof of human-like abilities in other species, towards an understanding of which components of cognitive processes are shared among species and which are species unique. Following this path will allow for an examination of the selection pressures leading to the evolution of the advanced cognitive abilities of humans. Additionally, this will open the door to acknowledging that other species may be sentient about their own experiences, either in ways similar to ours or ways unique to each species. The future does not lie in understanding how human-like the cognitive abilities of other animals are, but in what ways human, corvid, rodent and primate abilities are alike and in what ways they are different.



Summary NL



INLEIDING

Mentaal of cognitief reizen door de tijd stelt mensen in staat om zich details te herinneren van gebeurtenissen uit het verleden door middel van het episodisch geheugen. Daarnaast wordt het gebruikt voor het voorbereiden van acties en gebeurtenissen die in de toekomst zullen plaatsvinden. De vaardigheden om terug te kijken in het verleden (retrospectief) en vooruit te kijken naar de toekomst (prospectief) vereisen het gebruik van taal en worden daarom beschouwd als typisch menselijk vanaf het moment dat deze voor het eerst beschreven werden in de jaren '70.

Deze vaardigheden zouden echter ook in verschillende gradaties bij dieren kunnen voorkomen zonder dat daar een taal voor vereist is. Met gedragsstudies kan onderzocht worden in welke mate deze vaardigheden aanwezig zijn en ze kunnen inzicht geven in het evolutionaire ontstaan ervan. De afgelopen 20 jaar hebben verschillende zorgvuldig opgezette gedragsstudies aangetoond dat enkele kraaiachtigen en sommige knaagdieren prospectieve en retrospectieve eigenschappen vertonen. Deze onderzoeken hebben ertoe bijgedragen dat het mentaal reizen door de tijd niet meer alleen een menselijke eigenschap wordt gezien. Hierdoor is het mogelijk om vergelijkend onderzoek te doen naar deze eigenschappen bij zowel mensen als apen.

Tot op heden is er slechts weinig onderzoek gedaan naar de prospectieve en retrospectieve vaardigheden bij verschillende apensoorten, daarnaast zijn de studies niet systematisch waardoor er geen vergelijking gemaakt kan worden tussen de onderlinge soorten. Omdat apen onze naaste fylogenetische verwanten zijn en hoge cognitieve capaciteiten hebben, is het relevant om het vermogen om terug te kijken in het verleden en vooruit te kijken naar de toekomst systematisch te onderzoeken.

Voor dit promotieonderzoek heb ik vanuit een gedragsmatig perspectief door middel van experimenten de prospectieve en retrospectieve vaardigheden in verschillende soorten primaten onderzocht om een beter beeld te krijgen van de gedeelde en unieke aspecten bij deze soorten op het gebied van deze complexe en cognitieve vaardigheden. Er zijn methodes gebruikt die succesvol bleken bij andere soorten en ik heb geprobeerd de hiaten in de kennis over de aanwezigheid van deze vermogens bij primaten te vullen. Er is zowel een vergelijkende als een ontwikkelingsgerichte benadering gebruikt, daarbij is er gekeken naar de prospectieve vaardigheden van een gewone aap, de retrospectieve vaardigheden van een mens-aap en de retrospectieve ontwikkeling bij kinderen. Onderzoek naar wat de verschillende apensoorten weten over hun leefgebied en hoe deze omgaan met complexe problemen uit hun verleden en toekomst zouden uiteindelijk kunnen leiden tot meer inzicht in de evolutionaire geschiedenis van de prospectieve en retrospectieve capaciteiten bij primaten.

GEREEDSCHAPPEN

In dit hoofdstuk wordt een serie experimenten beschreven die samen een onderzoek vormen naar de prospectieve vermogens van de Java-aap (*Macaca Fascicularis*) door hen te leren gereedschappen met zich mee te nemen voor later gebruik. Bij verschillende soorten mensapen was deze methode eerder succesvol gebleken en daarom kon er met behulp van dit onderzoek een vergelijking worden gemaakt tussen een gewone aap en mensapen.

In het eerste experiment is er getest of zes makaken een gereedschap zouden uitkiezen die ze zouden meenemen om het elders en na een tijdsinterval voor een beloning te gebruiken. De ruimtelijke afstand en het tijdsinterval tussen het pakken van het gereedschap en de beloning in de verschillende testen werd gemanipuleerd. Echter, geen van de zes apen nam het gereedschap met zich mee. In het tweede experiment werden de apen getraind om een stuk gereedschap mee te nemen waarbij ze het gereedschap meteen konden gebruiken om een beloning te krijgen. De apen werden getest op hun vermogen om het gereedschap met zich mee te brengen wanneer het tijdsinterval voor de beloning verhoogd werd. Drie van de apen waren in staat het gereedschap mee te nemen en na een pauze van vijf minuten te gebruiken, van hen kon één aap het zelfs met een tussenpose van 10 minuten en gebruikte het gereedschap zelfs in enkele testen pas na 20 minuten. In het derde experiment werd onderzocht of twee apen een geschikt nieuw stuk gereedschap konden selecteren en met zich mee zouden nemen voor toekomstig gebruik. De aap die het beste presteerde bij het tweede experiment bleek dit te kunnen.

In het eerste experiment slaagden de apen er niet in om een stuk gereedschap mee te brengen voor een latere beloning, hierdoor bleken ze niet in staat te zijn om te kunnen plannen. Het zou echter mogelijk kunnen zijn dat ze de juiste volgorde van de handelingen die ze moesten verrichten niet konden overzien. Nadat de apen getraind waren om een stuk gereedschap met zich mee te nemen, bleken ze dit ook flexibel te kunnen doen wanneer de tijdsinterval voor de beloning verhoogd werd. Er was één aap die dit concept kon generaliseren naar een andere taak waarbij ze een geschikt gereedschap moesten uitzoeken, de andere apen probeerden voorwerpen uit hun verblijf te gebruiken of zelfs hun staart bij het verkrijgen van hun beloning. Door middel van de training om het gereedschap mee te brengen leerden de apen de juiste volgorde van gedragingen die nodig zijn om een beloning te kunnen bemachtigen, hierna bleken ze dit gedrag flexibel te kunnen toepassen wat erop zou kunnen wijzen dat makaken enige vooruitziende vaardigheden bezitten, al zijn die dan van een lager niveau dan bij de mensapen.

BELONINGEN

Het is bekend dat mensapen enkele prospectieve vaardigheden bezitten en toekomstige gebeurtenissen vooruit kunnen plannen. Over hun retrospectieve competenties is echter weinig bekend. In dit hoofdstuk zijn de retrospectieve vaardigheden van een mensaap, de chimpansee (*Pan troglodytes*), onderzocht. Er is een gedragskundig paradigma gebruikt waarin verschillende eigenschappen accuraat geïntegreerd moeten worden. Dit houdt voor het huidige onderzoek in dat er voedsel wordt verstopt en dat de dieren weten *wat* het voedsel was, *waar* het verstopt was en *wanneer* het verstopt werd. Deze retrospectieve vaardigheid wordt ook wel het wat-waar-wanneer, of episodisch-like geheugen genoemd. Het wordt vastgesteld door de zogenaamde drie stappen methode. Met deze methode worden drie soorten voedsel van verschillende smaken verstopt op een unieke lokatie, in de loop van de tijd verdwijnt het voedsel echter en elke smaak verdwijnt in een ander tempo. De chimpansees kunnen het voedsel op drie verschillende momenten bemachtigen.

In het experiment werden er twee soorten voedsel gebruikt, één lekkere en één minder lekkere, beide soorten verdwenen in een verschillend tempo. Daarnaast waren er vier lokaties waar het voedsel verstopt kon worden. De chimpansees observeerden hoe twee van de vier lokaties van voedsel werden voorzien, vervolgens mochten ze na een bepaald tijdsinterval (15 minuten, 1 uur of 5 uur) een lokatie aanwijzen. De individuen moesten hun keuze baseren op de tijd die verstreken was en hun geheugen over welke soort voedsel waar verstopt was. Afhankelijk van de trial konden ze hun favoriete beloning vinden of de beloning die nog over was.

Geen van de acht chimpansees bleken hun keuze te baseren op wat waar verstopt was, daarnaast hebben ze de verstreken tijd ook geen rol laten spelen bij hun keuze. Hiermee is in dit experiment niet aangetoond dat chimpansees het wat-waar-wanneer geheugen gebruiken bij het zoeken naar voedsel. Sterker nog, de chimpansees hadden geen enkel element van het voedsel-verstop experiment gebruikt om hun keuze op te baseren. Hun keuzes waren desondanks niet random, maar gebaseerd op een simpele associatie die gebaseerd was op de lokatie van het voedsel. De chimpansees kozen die plekken uit waar ze op een eerder tijdstip in de testen succesvol waren geweest, met deze strategie hadden ze redelijk veel voedsel gevonden.

Doordat de chimpansees geen enkele vorm van geheugen hebben gebruikt bij de experimenten, lijkt het niet toepassen van het wat-waar-wanneer paradigma eerder voort te komen uit de complexiteit van de testen dan een gebrek aan het episodisch-like geheugen. De mate waarin chimpansees retrospectieve vaardigheden vertonen blijven in dit onderzoek onbekend. Desondanks brengt het gedrag een interessante strategie aan het licht, namelijk dat dieren die worden blootgesteld aan veel elementen tijdens de testen (wat-waar-wanneer) zich gaan richten op de stabiele locaties van de beloningen en door middel van een simpele associatie een aanzienlijke hoeveelheid beloningen kunnen vinden.

SPEELTJES

In dit hoofdstuk is de ontwikkeling van retrospectieve vaardigheden bij kinderen van drie en vijf jaar onderzocht. Er werd een verhaal-test ontwikkeld die gebaseerd was op de drie stappen wat-waar-wanneer methode die zijn oorsprong vindt in het onderzoek bij kraaiachtigen en die eerder in het promotieonderzoek gebruikt is bij de studies aan de chimpansees. Wat-waar-wanneer of episodisch-like geheugen is vergelijkbaar met het episodische geheugen bij mensen, waarbij er wordt aangenomen dat dit laatste alleen samen kan gaan met het zelfbewustzijn.

Zelfbewustzijn is het vermogen bewust een episode uit het verleden te herbeleven, inbegrepen de wetenschap dat men het zelf is die het herinnerde heeft beleefd. Of het wat-waar-wanneer geheugen zich verlaat op zelfbewustzijn is niet duidelijk omdat het niet kan worden vastgesteld bij niet-sprekende soorten. Zelfbewustzijn ontwikkelt zich bij kinderen tussen het derde en vijfde levensjaar, er wordt van gedacht dat het samenhangt met de ontwikkeling van het episodisch geheugen en de Theory of Mind, wat het vermogen is om de eigen gedachten en overtuigingen te onderscheiden van die van anderen.

In het experiment is gekeken naar de prestaties van kinderen bij een serie Theory of Mind testen en de relatie met het wat-waar-wanneer geheugen. Dit is de eerste studie waarin de drie stappen methode bij kinderen gebruikt wordt om onderzoek te verrichten naar de ontwikkeling van het wat-waar-wanneer geheugen en de samenhang met het de ontwikkeling van Theory of Mind.

De wat-waar-wanneer verhaal test bevatte verschillende vragen waarmee onderzocht kon worden of kinderen de wat, waar en wanneer componenten afzonderlijk of op een geïntegreerde manier konden vinden. Rond een leeftijd van vier jaar slaagden de kinderen voor de “wat was waar?” vraag. Bij kinderen van rond de zes jaar slaagden de kinderen voor de “wanneer” vraag en de geïntegreerde “wat-waar-wanneer” vraag. Een goede uitslag van de Theory of Mind test was geen garantie voor een goed verloop in de wat-waar-wanneer geheugen test, wat er op kan duiden dat de prestaties van de wat-waar-wanneer test niet samenhangen met Theory of Mind, terwijl het episodisch geheugen deze samenhang wel eerder vertoonde. De relatie tussen wat-waar-wanneer en het episodisch geheugen zou echter nader onderzocht moeten worden. De onderzochte wat-waar-wanneer geheugentest was misschien te veel gevraagd voor jonge kinderen omdat er van een begrip van tijd wordt uitgegaan, dit ontwikkelt zich pas rond een leeftijd van zes jaar.

DISCUSSIE

Het doel van dit proefschrift was om vanuit een vergelijkend en evolutionair perspectief de prospectieve en retrospectieve competenties van mensen en niet-humane primaten te onderzoeken. In het licht van eerdere onderzoeken naar deze

competenties, lijkt het huidige proefschrift aan te tonen dat verscheidene soorten primaten verschillen vertonen in prospectieve en retrospectieve vaardigheden. Er moet echter voorzichtig worden omgegaan met voorbarige conclusies, aangezien dit proefschrift slechts een klein aantal studies met slechts enkele individuen omvat.

De resultaten laten zien dat **kinderen** de wat-waar-wanneer driestapstest rond de leeftijd van vijf jaar kunnen begrijpen, en dat jongere kinderen daar – waarschijnlijk door het niet begrijpen van de ‘wanneer’ opdracht – niet in slagen. Eerdere bevindingen hebben vastgesteld dat kinderen rond de leeftijd van vijf jaar ook succesvol worden in verbale onderzoeken naar episodisch geheugen. Een onderzoek naar de prospectieve vaardigheden van kinderen heeft aangetoond dat kinderen van drie jaar oud een gereedschap mee kunnen nemen als ze daar direct een beloning voor krijgen, ze kunnen dit echter niet als ze dit gereedschap pas in de toekomst nodig hebben totdat ze ongeveer vier jaar oud zijn. Al deze bevindingen suggereren dat met betrekking tot prospectieve en retrospectieve vaardigheden het verplaatsen in de tijd de zwaarstwegende component is en zich pas laat ontwikkelt bij kinderen. Verder onderzoek moet uitwijzen hoe kinderen hun prospectieve vaardigheden verder ontwikkelen in de loop der tijd, evenals hoe en wanneer zij de mogelijkheid ontwikkelen om complexe (toekomst)plannen te maken.

In de studie naar de retrospectieve capaciteiten van **mensapen**, gebruikten de chimpansees geen wat-waar-wanneer geheugen bij de driestaps methode, ze maakten zelfs geen gebruik van hun lange termijn geheugen, waardoor er geen uitsluitel werd gegeven over de retrospectieve vaardigheden bij chimpansees. Desalniettemin hebben eerdere studies aangetoond dat chimpansees informatie over een langere periode kunnen onthouden en dat een plek waar voer verstopt ligt zelfs drie dagen onthouden wordt. Onderzoeken naar prospectieve vaardigheden laten zien dat mensapen zelfs na 14 uur tussenpauze gereedschappen kunnen selecteren en transporteren en ook voor hen onbekende gereedschappen kiezen om die na een uur pauze te gebruiken om een beloning te verkrijgen. Met de huidige kennis lijkt het zo te zijn dat mensapen betere prospectieve dan retrospectieve kwaliteiten bezitten, hierbij moet vermeld worden dat hun prospectieve kwaliteiten breder onderzocht zijn en dat de nadruk bij toekomstig onderzoek meer op het retrospectieve vlak zou moeten liggen.

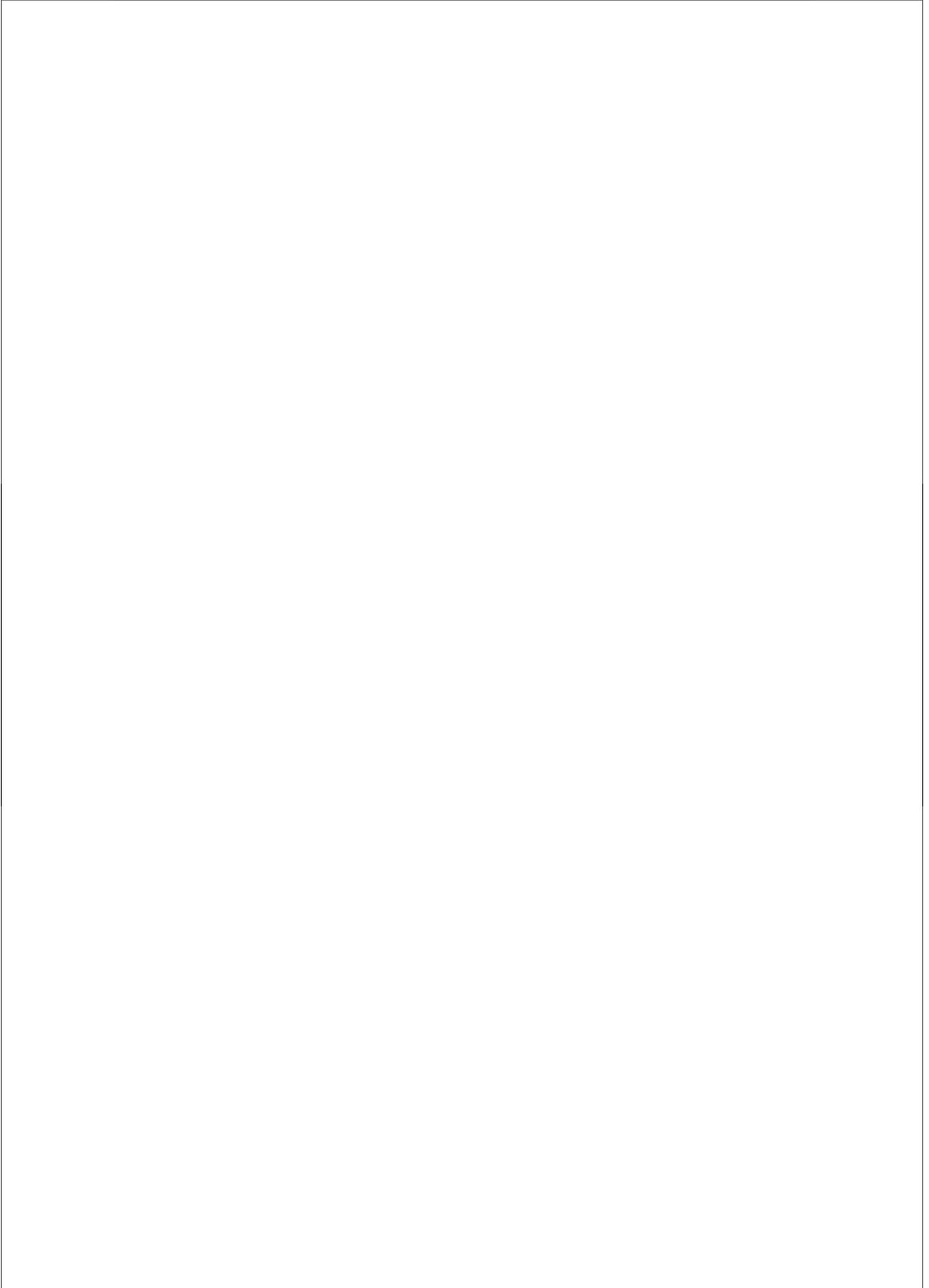
Ten aanzien van de prospectieve vermogens van **oude wereld apen** haalden de apen het criterium niet voor de ‘planning’ testen, ondanks het achterwege laten van de tijdscomponent – een interval van 5 minuten tot een uur – die bij de mensapen werd gebruikt. Echter, toen de dieren getraind waren om een gereedschap met zich mee te brengen, werden ze succesvol in het voltooiën van de test. De best presterende aap was, interessant genoeg, in staat om na 20 minuten interval de test af te maken, wat wellicht suggereert dat de potentiële prospectieve competenties van makaken binnen een korter tijdsbestek vallen dan die van gaaien, mensapen en mensen. Eerdere onderzoeken hebben aangetoond dat rhesus apen niet goed zijn in het integreren van de tijdsfactor bij prospectieve en retrospectieve testen, hoewel dit door methodologische problemen opnieuw beoordeeld dient te worden. Huidige data suggereren dat gewone apen beperktere prospectieve en retrospectieve eigenschappen

hebben, dit komt waarschijnlijk doordat ze moeite hebben met het integreren van de component 'wanneer'. Nadere onderzoeken zijn nodig om vast te stellen of gewone apen wellicht enige mate van of simpele toekomstgerichte gedragingen vertonen, zoals dit proefschrift suggereert en of ze beter zullen presteren als er minder complexe taken nodig zijn om de tijdsduur te overbruggen.

Bij de onderzoeken naar prospectieve en retrospectieve vaardigheden bij verschillende apensoorten, lijkt het dat zowel apen als jongere kinderen moeilijkheden met de 'wanneer' component hebben. Kinderen van rond de vijf jaar en mensapen kunnen het verplaatsen naar een toekomstige gebeurtenis begrijpen, terwijl het huidige onderzoek suggereert dat Java-apen versimpelde prospectieve vaardigheden hebben die gelimiteerd zijn aan een kortere tijdsduur dan mensapen en mensen. Het huidige proefschrift suggereert ook dat kinderen van ongeveer vijf jaar het wat-waar-wanneer van een gebeurtenis uit het verleden kunnen begrijpen, terwijl jongere kinderen zich alleen het 'wat' en 'waar' herinneren. De resultaten van de retrospectieve vaardigheden bij mensapen geven in het proefschrift geen uitsluitel en kunnen niet worden vergeleken met het onderzoek aan kinderen. Bij gewone apen lijkt het aannemelijk dat ze zich alleen het 'wat' en 'waar' uit het verleden kunnen herinneren. Ook is er nader onderzoek nodig om de aangegeven verschillen tussen de verscheidene apensoorten kunnen bevestigen en hun oorzaken vast kunnen stellen. De onderzoeken in dit proefschrift zijn slechts een kleine opening naar het inzicht omtrent prospectieve en retrospectieve capaciteiten van mensen en niet-humane primaten en de discussie aangaande het reizen door de tijd van dieren is nog maar net begonnen.

Toekomstig onderzoek zouden zich moeten focussen op de overeenkomsten van componenten van cognitieve processen bij primaten, in plaats van het zoeken naar de bewijzen van menselijke vermogens bij primaten. Dit zou uiteindelijk kunnen leiden tot een onderzoek naar de evolutionaire selectiedruk die tot de verstrekkende cognitieve processen binnen de mens hebben geleid. Daarnaast zou dit de deur kunnen openen naar de bevestiging dat andere soorten zich mogelijk bewust zijn van hun ervaringen, ofwel op een wijze die op de onze lijkt, ofwel op een soort-eigen manier. Het is belangrijk om te begrijpen op wat voor een manier de cognitieve vaardigheden van mensen, kraaiachtigen, knaagdieren en primaten op elkaar lijken, en op welke wijze zij verschillen en niet om te beoordelen hoe menselijk de cognitieve vaardigheden van dieren wel niet zijn.

Introduction



Humans can re-experience unique personal episodes and imagine future events, by mental or cognitive time travel into their past and future (Gilbert & Wilson, 2007; Suddendorf & Corballis, 1997; Tulving, 2005). These retrospective and prospective abilities, known as episodic memory and future planning respectively (Tulving, 1972; Tulving, 2005), enable fast and flexible responses to present and future environmental demands. Re-experiencing past events enables a re-evaluation of the event and its features in the present when new information becomes available. Features or relations between these features, which may have been perceived as irrelevant before, can now be assigned new values. This re-evaluation enables the extraction of new associations and information from past experiences. The re-combination of known features also enables the envisioning of future scenarios, even for events that were never experienced before. Mental time travel thus enables highly flexible responses to both current and future demands and doubtlessly has great adaptive value for humans. It has been stated that humans are the only species able of such mental time travelling (Roberts, 2002; Tulving, 1972; Tulving, 1984; Tulving, 2005; Nelson & Fivush, 2000) and, therefore, that this ability may represent a discontinuity in the evolution of cognition (Suddendorf & Busby, 2003; Suddendorf & Corballis, 1997; Suddendorf & Corballis, 2007a).

Mental time travel in humans is defined as a process relying on consciousness and is assessed by verbal tests (Tulving, 2005). As consciousness is inaccessible to us in non-human animals (Clayton et al., 2003a; Clayton & Dickinson, 1998; Clayton & Dickinson, 1999b; Clayton et al., 2001a) and referential language is uniquely human (Christiansen & Kirby, 2003), the hypothesis of human uniqueness is although plausible (Suddendorf & Busby, 2003; Tulving, 2002), also unfalsifiable. This definition of mental time travel thus a priori excludes the possibility to investigate this capacity in non-human animals. Instead of anthropocentrically claiming human uniqueness, a broader biological approach entails empirically examining how animals deal with complex problems concerning their past and future. Following this approach will expand our understanding of how species differ in their prospective and retrospective abilities and whether similar abilities or their precursors are present in other species as well as the evolutionary history of these abilities.

To study episodic memory and future planning in non-humans, a useful approach has been formulated focusing on observable behaviours and thereby circumventing the inaccessible subjective aspects of human mental time travel (Correia et al., 2007; Mulcahy & Call, 2006; Raby et al., 2007b; Clayton & Dickinson, 1998; Clayton et al., 2001b). This approach allows testing whether other species are able to remember details of certain episodes and plan for some future events. Animals may benefit from the ability to remember features of experienced episodes in terms of foraging behaviour and social interactions (Schwartz & Evans, 2001; Clayton et al., 2001a; Dere et al., 2006). In the foraging context, animals may benefit from remembering the locations of recently depleted resources to effectively minimize revisits (Schwartz & Evans, 2001; Suddendorf & Corballis, 2007a; Janmaat et al., 2006) or, in the case of food hoarding species, the locations of their caches (Clayton et al.,

2001a). In the social domain keeping track of who did what to whom could be used in alliance formation (Clayton et al., 2001a; Cheney & Seyfarth, 1992) or when keeping track of the social interactions of an unfamiliar individual to guide one's own behaviour in subsequent interactions with this individual (Paz-y-Miño et al., 2004). Also species with a fission-fusion social structure, where individuals depending on the activity switch between merging (fusion) into large groups and splitting (fission) into smaller sub-groups, may require increased monitoring across space and time of the comings and goings of conspecifics (Barrett et al., 2003). Tool using animals may benefit from transporting or manufacturing tools when these are not currently required, but will be in the future (Boesch & Boesch, 1984). The study of behaviours reflecting prospective and retrospective abilities of non-human animals is rapidly gaining empirical support (Crystal, 2009; Eacott & Easton, 2010; Osvath & Osvath, 2008; Raby et al., 2007b; Raby & Clayton, 2009; Salwiczek et al., 2010; Zentall, 2005) and is casting doubt on the hypothesis of human uniqueness in this regard.

The ability to recall features of past events and anticipate future ones has mostly been investigated in corvids and rodents (Crystal, 2009; Eacott & Easton, 2010; Salwiczek et al., 2010). Considering this ability is believed so fundamental to human cognition, surprisingly little research has focused on prospective and retrospective abilities of non-human primates. Furthermore, the few primate studies have each used different definitions of the studied concepts and did not always adequately exclude alternative explanations of the animals' behaviour in terms of simpler strategies. Therefore, it is difficult to make a comprehensive comparison of the findings across these studies (Hampton & Schwartz, 2004; Schwartz & Evans, 2001; Martin-Ordas et al., 2010). As non-human primates are our closest phylogenetic relatives and have some advanced cognitive capacities (Call, 2003; Tomasello & Call, 1997; Byrne, 2002), it is relevant to systematically address their abilities also in the domain of past and future knowledge. The present thesis addresses the presence of prospective and retrospective abilities in two species of non-human primates as well as the ontogeny of retrospective abilities in children in order to obtain a better understanding of the differences and similarities between primate taxa and to study the possible evolutionary history of these abilities in primates.

I study the ability of long-tailed macaques (*Macaca fascicularis*) to prepare for a future event with a method previously only used with great apes, in order to compare the prospective abilities of a representative monkey species with those of apes. I also examine the ability of chimpanzees (*Pan troglodytes*) to form and recall features of past, trial unique episodes, with a behavioural method that enables exclusion of simpler explanations of the animals' performance. This will enable a comparison of the retrospective abilities of a representative great ape species to the abilities of other animals tested with this paradigm. Lastly, I test three and five year old children's abilities to remember details of episodes with the behavioural test used with the chimpanzees, to examine this tests requirements on children in the age range before and after they develop the ability to verbally report their mental time travels. This will expand our understanding of how and to what extent human and non-human

primates can deal with complex problems in the context of past and future episodes as well as the levels of similarities between their abilities.

In the following, I first present an overview of the taxonomy of the mammalian memory system in order to present the characteristics of episodic memory as it is described in humans, as well as its association with prospective abilities. I then explain the criteria used to investigate episodic-like or what-where-when memory in animals. In order to better understand the requirements of episodic memory I give a short overview of its ontogeny and the abilities it relies on in human children. Following this I summarize animal studies on what-where-when memory and planning for a future need, with an emphasis on primates. Lastly, I present how this thesis will expand our knowledge of what human and non-human primates recall about their past and plan about their future.

WHAT IS EPISODIC MEMORY?

The mammalian memory system (Figure 1.1) is based on its different features divided into several interconnected subsystems (Miyashita, 2004; Squire, 1992; Squire, 2004). The separation of the long-term and short-term memory system is globally based on the length of time the information is stored. Although there is no clear boundary between short and long-term memory even within a particular species, information retained for several minutes is considered to be subserved by long-term memory systems (Hampton & Schwartz, 2004; Schwartz & Evans, 2001). Long-term memory is further subdivided based on how the stored information is retrieved. Memories remembered consciously are stored in the declarative system, while those that do not require conscious awareness for retrieval are stored in the nondeclarative system (Cohen & Squire, 1980; Schacter & Tulving, 1994). Declarative memories provide greater behavioural flexibility as they can be voluntarily triggered top-down from the frontal lobes, while nondeclarative ones are triggered bottom-up through perception (Miyashita, 2004). The human declarative system is, based on the kind of information it stores, divided into the episodic and the semantic memory system (Tulving, 1972; Tulving, 1983). Thus, while other memory systems are present amongst all mammals, it is only the episodic memory system that is considered unique to humans and its presence in other species is currently debated. The episodic memory system stores details of personally experienced one-time events or particularities, while general facts about the world or regularities are contained in the semantic memory system without the context in which these facts were first learned (Tulving, 1972; Tulving, 2001; Tulving, 2002). Facts, such as “people receive presents on their birthdays”, are stored in the semantic memory system, while detailed information about the last birthday party held, e.g. “the location and people present”, is re-experienced by accessing the episodic memory system. The disassociation of the episodic and semantic memory system is in humans supported by neuropsychological evidence from case studies of impaired subjects and experiments with healthy individuals (Tulving, 2001; Vargha-

Khadem et al., 2001; Gardiner et al., 2002). In addition, the episodic memory system ontogenetically develops later and deteriorates earlier in old age compared to the semantic system (Tulving, 2005; Wheeler et al., 1997).

Humans can report remembering features of experienced episodes by re-experiencing the episode, an ability referred to as mental time travel (Suddendorf & Corballis, 1997). Humans can also report on mentally travelling into the future and imagine future events by pre-experiencing events (Roberts & Feeney, 2009). Mental or cognitive time travel is, thus, a bidirectional ability with a retrospective component oriented towards the past and a prospective component oriented towards the future. This association between prospective and retrospective cognitive time travelling is in humans confirmed by shared cognitive resources; such as similar loss of detail in past and futures reports by healthy adults (D'Argembeau & Linden, 2004), similar developmental onsets (Busby & Suddendorf, 2005), and similarities in brain activation imaging studies of both healthy subjects (Addis et al., 2007; Okudaa et al., 2003; Szpunar & McDermott, 2008) and amnesic subjects (Klein et al., 2002; Tulving, 2002).

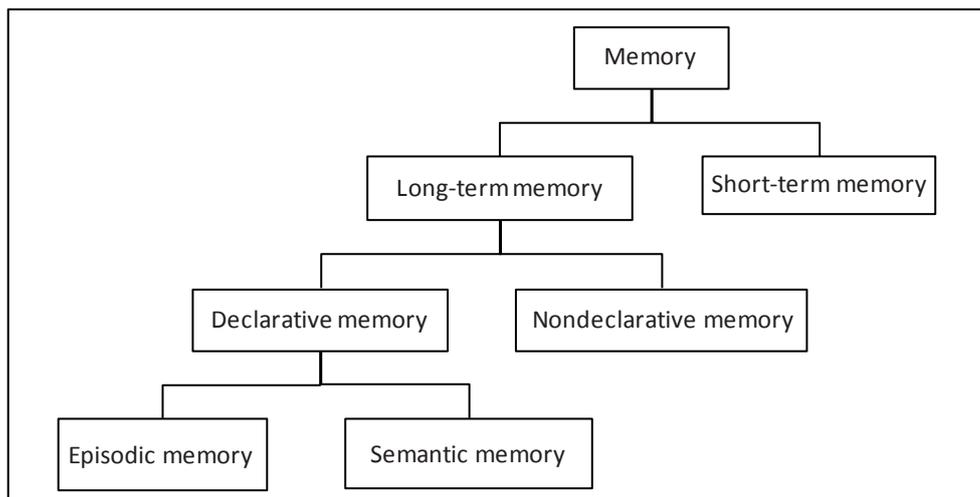


Figure 1.1: Overview of the taxonomy of the mammalian memory system (adapted after (Squire, 2004)).

WHAT IS EPISODIC-LIKE OR WHAT-WHERE-WHEN MEMORY?

Tulving's (1972) originally defined episodic memory as a system that "receives and stores information about temporally dated episodes or events, and temporal spatial relations among these events". He later modified this definition in order to clearly distinguish episodic memory from semantic memory, which could enable knowing what happened, where and when, without specifically remembering the given episode during which the information was first learned (Tulving, 2002). Therefore, he added that episodic memories also require both auto-noetic consciousness, that is, the

conscious re-experiencing of the remembered event including the knowledge that it is oneself who has experienced the event (Tulving, 1985; Roberts & Feeney, 2009; Wheeler, 2000) as well as *cronesthesia*, that is, the subjective time when an event took place in one's own past (Tulving, 2005). The presence of these subjective abilities is in human episodic memory studies confirmed by language (Conway, 2009; Yonelinas, 2002; Klein et al., 2002; Tulving, 2002; Dere et al., 2008).

Several different criteria have been proposed to study whether animals could have similar retrospective abilities in the absence of reporting them by language. Without an agreed upon operational definition of non-human consciousness, different studies have made vastly different assumptions as to the nature of episodic memory in non-humans (Clayton et al., 2003a; Eacott & Easton, 2010; Hampton & Schwartz, 2004). Such vastly different approaches make it difficult to obtain a comprehensive overview of animals' retrospective abilities. In this thesis, I have chosen to apply the criteria proposed by Clayton and colleagues (Clayton & Dickinson, 1998; Clayton et al., 2001b; Clayton et al., 2009b). Their approach draws on the fact that animals can retain perceived information about events and locations, as well as mark the passage of time using interval and circadian timing mechanisms (Roberts, 2002; Shettleworth, 1998). Clayton and colleagues outline animals' prospective and retrospective abilities in behavioural terms, thereby circumventing the subjective and phenomenological aspects of episodic recollection in humans. Furthermore, their approach enables controlling for the use of semantic memory and simpler rule learning or associative explanations of the animals' behaviour. This approach is centred on the animals' behaviour and is based on Tulving's original definition of episodic memory (Tulving, 1972). Therefore, the tested ability is referred to as **episodic-like memory** (Clayton & Dickinson, 1998). To establish the presence of **episodic-like memories** in animals, three behavioural criteria have been proposed: *content*, *structure* and *flexibility* (Clayton et al., 2003a). The *content criterion* specifies that the animals should recall the what, where and when of a trial unique episode. The "when" criterion, is especially important, as no two events can occur at the same point in time. Accurate retrieval of the when component is claimed to demonstrate the retrieval of a specific episode, formed following singular exposure, and in turn excludes the potential use of semantic memory. The *structure criterion* involves integration of the what, where and when into a single unified representation, while the *flexibility criterion* refers to the flexible deployment of **episodic-like memories** in novel situations not predicted at the point of encoding, a defining feature of the declarative memory system (Clayton et al., 2003a).

Episodic-like memory is assessed by behaviour demonstrating the retrieval of what-where-when features of a unique past episode. While it does not measure the conscious re-experiencing of the remembered information, the term does imply that such awareness could be present. A somewhat more neutral term **what-where-when memory** has been suggested as the more precise term describing the measured abilities in animals (Suddendorf & Busby, 2003). In the present thesis, I predominantly

use the term **what-where-when memory**, however, the two terms essentially describe the same ability.

ONTOGENY OF EPISODIC MEMORY IN CHILDREN

In order to understand the cognitive demands of episodic memory, it is useful to consider the ontogeny of this and other abilities that share the same period of developmental onset. Declarative memory is considered a late developing memory system (Nelson, 1995; Schacter & Tulving, 1994) with the ability to report on episodic memories gradually developing between the ages of four and six years (McCormack & Hoerl, 1999; Suddendorf & Busby, 2005; Perner, 2000; Roberts, 2002). Around the age of four children also begin developing the prospective ability to plan for future needs, events (Atance & O'Neill, 2005; Russell et al., 2010; Tulving, 2005) and actions (Table 1.1) (Suddendorf et al., 2011). This development is a complex gradual process influenced by the development of several other abilities, such as overall memory maturation, understanding of time and differentiation between past and future events (Friedman, 2005; Friedman & Lyon, 2005; Atance & O'Neill, 2005; Friedman, 2003), language and the concept of self (Nelson & Fivush, 2004). Language is a valuable tool in this regard as verbal reports allow us to examine the children's ability to remember and understand their memories of past events (Atance & O'Neill, 2005). However, their inability to accurately report on their knowledge and understanding should not be mistaken for proof that before the age of four these abilities are not present, albeit in a simplified form. Children under the age of four may, for example, not sufficiently master language as a representational system (Nelson, 1992; Nelson & Fivush, 2004) and may be unable of passing verbal test on episodic memory. In fact, there is no clear cut evidence or criteria for when infants become able to form conscious declarative memories (Perner, 2000). To accurately assess episodic memories in pre-verbal children non-verbal paradigms with adequate controls should be developed in adults and then tested on pre-verbal children (Salwiczek et al., 2010; Suddendorf & Busby, 2003). This would require that the definition of human episodic memory would be expanded to include non-subjective, behavioural elements. Behavioural what-where-when tests could be considered as candidates for non-verbal tests of episodic memory (Salwiczek et al., 2010).

Verbal tests with children have revealed that the development of the understanding of the concept of self, as demonstrated by the development of Theory of Mind abilities between the ages of three and five, is of particular importance for the development of episodic memory (Naito, 2003; Perner, 2000; Perner et al., 2007; Perner & Ruffman, 1995). Theory of Mind is the conceptual system that underlies the human ability to impute mental states such as what we know and want, to both others and ourselves (Perner, 2000). In the context of episodic memory ontogeny particular emphasis is placed on certain Theory of Mind tasks. These tasks tap into the ability to understand how different types of sensory experience lead to knowledge and into the

ability to connect a given memory to its original event (Perner et al., 2007; Perner & Ruffman, 1995; Tulving, 2002) and identify features required to form episodic memories. A study with three to six year old children directly compared the development of one of such Theory of Mind tasks, called the see-know task, which assesses the understanding between perceiving and knowing, and episodic memory measured by free recall. The free recall method is routinely used to measure episodic memory competence in humans and measures accurate retrieval of previously presented items with open ended questions, such as “name the previously presented pictures” (Perner & Ruffman, 1995; Tulving, 1985). In the see-know task the children either observe or are told about a hidden object’s location and are tested on their ability to report how they obtained knowledge about the object’s location (i.e. seeing or hearing). A significant association between passing the see-know task and free recall was found, even after controlling for the influences of semantic memory and verbal intelligence, demonstrating that acquisition of these Theory of Mind elements is developmentally related to episodic memory, but not to semantic memory (Perner & Ruffman, 1995). Recently, a study assessed episodic memory by manipulating informational conditions to either permit or prevent the formation of episodic memories. This was done by allowing children to either see which picture cards they were placing in a box, or having them do this blindfolded and only later showing them what was on the cards (Perner et al., 2007). The authors’ reasoning is that this seeing or direct condition enables remembering items by episodic memory as the event can be re-experienced; while the indirect condition prevents this, as children cannot visually re-experience an event they did not witness due to the blindfold. Recall accuracy of directly experienced pictures was found to increase with Theory of Mind competence, but not with the recall of indirectly experienced pictures indicating a specific developmental relationship between the Theory of Mind and episodic memory (Perner et al., 2007). Episodic memory competence improves as children improve on Theory of Mind tasks, both when episodic memory is measured with free recall or with the direct experience method. Episodic memory and Theory of Mind development, as measured by verbal reports, seem specifically linked to the growing ability to introspect an ongoing experience and interpret it as representing an actual past event (Perner, 2000).

The ontogeny and relationship between what-where-when memory and Theory of Mind have not yet been investigated in children. Such a study could elucidate whether the behavioural what-where-when memory development is dependent on similar Theory of Mind abilities as the phenomenological episodic memory.

WHAT IS KNOWN ABOUT WHAT-WHERE-WHEN MEMORY IN ANIMALS?

Several different methods have been used to study what-where-when memory in animals (Table 1.1), reflecting the many different interpretations of the concept

(Crystal, 2009; Dere et al., 2006; Hampton & Schwartz, 2004; Schwartz & Evans, 2001). Many of the methods do not adequately control for potential alternative explanations of the animals behaviour in terms of simpler abilities, which is essential when examining the presence of complex cognitive abilities. Consideration of such different levels of interpretation is especially important when establishing whether non-linguistic species have cognitive skills similar to episodic memory, which is in humans defined in subjective terms and assessed by language. Exclusion of alternative explanations is, therefore, more prominent in animal literature compared to human literature mentioned in the previous section. In the following, I give an overview of the methods used to study what-where-when memory in animals, with a focus on primates.

Great apes have shown the ability to retain information for extensive periods of time (years) following training (Beran, 2004; Bearn et al., 2000) and the ability to distinguish several baited locations after single exposure with retention intervals of 30 minutes and 1h (Gibeault & Macdonald, 2000; MacDonald & Agnes, 1999; Menzel, 1973; Martin-Ordas et al., 2010). Their ability to remember several features of an event have been tested with the following two methods: the free recall method and the unique event method (Menzel, 1999; Schwartz et al., 2002; Schwartz et al., 2005; Schwartz et al., 2004). The free recall method (Tulving, 1984) has been used with a lexigram trained chimpanzee, who observed an experimenter hiding food or objects in a large outside area (Menzel, 1999; Menzel, 2005). Sixteen hours later the chimpanzee, on her own account, obtained the attention of a naive caretaker and with the aid of the lexigram communicated the objects' locations. Because the chimpanzee elicited the caretakers assistance on her own account and thus in the absence of potential retrieval cues, it has been argued that this method is comparable to the free-recall method used in humans (Hampton & Schwartz, 2004). The presented method clearly tests long-term memory of what and where features after single exposure. It is less clear whether it addresses the when feature and whether the features are structured into a unified event (Clayton et al., 2003a), as the animal is not required to recall how long ago the event took place or distinguish between different events occurring at different times. The chimpanzee's success could be explained by remembering what is where instead of recalling the hiding event and its behaviour could be guided by memory updates of the locations and their content (Schwartz et al., 2005). Therefore, this method does not meet the requirements for establishing what-where-when memory in animals.

The unique event method is based on a different definition than the free-recall method, of what may constitute episodic-like memory in non-humans. Schwartz and colleagues propose that animals should be tested about their knowledge of single, unique events, and that the animal's response should concern the past and not the present situation (Schwartz et al., 2002; Schwartz et al., 2005; Schwartz et al., 2004). They have shown that a single gorilla (*Gorilla gorilla gorilla*) could, by handing over the correct picture card, indicate which food type it had received from which caretaker (Schwartz et al., 2002) as well as which novel actions, persons or objects it had recently observed (Schwartz et al., 2004). According to the authors, this method

demonstrates knowledge of unique past events, as at the time of questioning the events have already transpired (the food had been consumed). Additionally, this method enables testing the “who” feature (by asking which caretaker had given the food), which may be an important feature for primates living in complex social systems. However, this method also does not test the integration of the when feature (Clayton et al., 2003a) and, in addition, the gorilla could be returning the card that presents the most familiar stimuli, without remembering the actual event (Schwartz et al., 2005). If the task was solved based on familiarity of the pictures, then it cannot be considered a memory task, as familiarity is a process distinct from remembering, both psychologically (Jacoby, 1991) and neurobiologically (Aggleton & Brown, 1999; Ranganath et al., 2004; Yonelinas, 2002). Thus, both of the above methods (free recall method and unique event method) are inadequate as tests of what-where-when memory, because in both cases the positive results could be explained by alternative simpler strategies. In addition, both involve testing subjects who require extensive training in order to communicate with the experimenters by means of lexigrams or picture cards.

Useful insight on how to better study what-where-when memory in animals has come from the study of a corvid, the scrub jay (*Aphelocoma californica*), that naturally caches and recovers different types of food (Clayton et al., 2003a; Clayton & Dickinson, 1998; Clayton & Dickinson, 1999b; Clayton et al., 2001a; deKort et al., 2005; Clayton et al., 2001b). In their classical experiment, the scrub jays cached two food types that differed in both their perishability and their palatability (Clayton & Dickinson, 1998). This method required the animals to distinguish between recovering two different foods after two retention intervals of different lengths, therefore, I refer to it as the two step method. The jays in the experimental group learned that their preferred wax worms degrade a few days after caching, while their less preferred peanuts always remained fresh. When the jays recovered their caches some hours after storing the foods they preferentially searched for the wax worms, but when they were first given access to their caches after several days, they preferentially recovered the peanuts. Jays in the control group, who never experience the degraded worms, continued to recover the preferred worms as these remained fresh after both retention intervals (Clayton & Dickinson, 1998). This test produced the first evidence of what-where-when memory in non-human animals (Crystal, 2009), by showing that the jays could remember which foods they had cached where and when. Clayton and colleagues have since carried out several additional experiments to reinforce that the jays integrate the content of the episodes into a single memory and that they can flexibly update information about the decay rate of the foods, confirming the content, structure and flexibility criteria for episodic-like memory in scrub jays (Clayton & Dickinson, 1999c; Griffiths & Clayton, 2001; Clayton et al., 2009a; deKort et al., 2005; Clayton et al., 2001a; Clayton et al., 2001b; Clayton et al., 2003b; Clayton & Dickinson, 1999a; Clayton et al., 2003a).

In one of the follow-up experiments, an additional food type and interval were added to exclude a potential rule based interpretation of the jays’ behaviour (Clayton

et al., 2001b) in the two step set-up. In the two step set-up, the jays could have searched for the worms after the short interval and avoided them after the long interval, indicating that they had only encoded the where and when of the task, but not the what. Therefore, the content and structure criteria may not be met. In this follow-up set-up the animals had to adjust their searches during food recovery in six different conditions, as three food types are cached in two different combinations and can be recovered after three different time intervals. I refer to this as the three step method. The birds were successful in distinguishing the caches contents after all three time intervals and showed remarkable flexibility, by adjusting their searches based on which of two stored foods remained present (Clayton et al., 2001a). These results confirmed that the jays remembered temporal information about when the foods were cached relative to recovery, as well as which foods were cached where and excluded the possibility that they were guided by simple rule learning (Clayton et al., 2001a). This three step method confirms the content and structure criteria for what-where-when memory.

It has been proposed that in order to adequately exclude that animals are solving the task by relying on semantic knowledge, they should not only be tested with trial unique episodes, as is the case with the scrub jays, but also with unexpected questions about previously presented information (Zentall, 2005; Zentall, 2006; Zentall et al., 2001). However, the delayed-matching-to-sample paradigm proposed to test such unexpected questions uses retention intervals of a few seconds (Zentall, 2006), which is inconsistent with episodic memory being part of the long-term memory system (Crystal, 2009). The method has, however, been instrumental in showing that both pigeons (*Columba livia*) (Zentall et al., 2001) and in an adapted version dolphins (*Tursiops truncatus*) (Mercado et al., 1998; Mercado et al., 1999) can report their previous actions and behaviours indicating declarative memory competence. Nevertheless, the food caching and recovering method remains a better test of what-where-when memory as it allows testing for information retained over extended periods of time.

The food caching and recovering method has been adapted and replicated in several studies. However, these studies usually replicate the original two step set-up (Clayton & Dickinson, 1998) and thus, do not confirm the integrated structure of the what-where-when or the flexibility of the memories, although some do include additional conditions confirming content and structure (Zhou & Crystal, 2009; Martin-Ordas et al., 2010).

Table 1.1: Overview of species tested on retrospective and prospective abilities including the used methods.

Studies vary in the number of individuals used. S indicates the tested individuals successfully passed the criteria as specified in each study, U indicates they were unsuccessful. Note that the criteria and definitions used greatly vary between the studies.

Retrospective abilities		Prospective abilities	
Birds			
Scrub jays	2 and 3 step WWW with food: S (Clayton et al., 2003a; Clayton & Dickinson, 1998; Clayton et al., 2001b; deKort et al., 2005)s	Satiation method: S (Correia et al., 2007; Raby et al., 2007)	
Magpies	2 step WWW with food: S (Zinkivskay et al., 2009)		
Black capped chickadees	2 step WWW with food: S (Feeney et al., 2009)		
Rodents			
Rats	2 step WWW with food: S (Babb & Crystal, 2005; Babb & Crystal, 2006a; Roberts et al., 2008; Zhou & Crystal, 2009; Naqshbandi et al., 2007) 2 step WWW with food: U (Bird et al., 2003; McKenzie et al., 2005) 2 step WWW with objects: S (Kart-Teke et al., 2006) what-where-which: (Eacott et al., 2005) WWW with odours: S (Ergorul & Eichenbaum, 2004)	Satiation method: U (Naqshbandi & Roberts, 2006)	
Mice	2 step WWW with objects: S (Dere et al., 2005b; Dere et al., 2005a)		
Volves	2 step WWW with mating opp.: S (Ferkin et al., 2007)		
Primates			
Squirrel monkeys		Satiation method: S (Naqshbandi & Roberts, 2006)	
Long-tailed macaques		Spoon test: U (<i>this thesis, chapter Tools</i>)	
Rhesus macaques	2 step WWW with food: U (Hampton et al., 2005)	Satiation method: U (Paxton & Hampton, 2009)	
Gorilla	Unique events method: S (Schwartz et al., 2002; Schwartz et al., 2005; Schwartz et al., 2004)		
Orangutans	2 step WWW with food: S (Martin-Ordas et al., 2010)	Spoon test: S (Osvath & Osvath, 2008; Mulcahy & Call, 2006)	
Bonobos	2 step WWW with food: S (Martin-Ordas et al., 2010)	Spoon test: S (Mulcahy & Call, 2006)	
Chimpanzees	2 step WWW with food: S (Martin-Ordas et al., 2010) 3 step WWW with food: U (<i>this thesis, chapter Treats</i>) Free recall method: S (Menzel, 1999; Menzel, 2005)	Spoon test: S (Dufour & Sterck, 2008; Osvath & Osvath, 2008) Exchange method: U (Dufour & Sterck, 2008) Spontaneous planning: S (Osvath, 2009)	
Humans	Free recall method: S at app.4 years (Perner & Ruffman, 1995) Direct experience method: S at app. 4 years (Perner et al., 2007) 3 step WWW with objects: S at app. 5 years (<i>this thesis, chapter Toys</i>)	Future directed action: S at app.4 years (Suddendorf et al., 2011)	

The what-where-when approach has been successfully replicated with rats (*Rattus sp.*) searching for different flavoured food pellets in a radial maze test (Babb & Crystal, 2005; Babb & Crystal, 2006b; Babb & Crystal, 2006a) after initial negative results in set-ups where the rats cached food themselves (Bird et al., 2003; McKenzie et al., 2005). Studies with rats have also carefully examined the integration of the “when” feature. The initial radial maze studies (Babb & Crystal, 2005; Babb & Crystal, 2006a) used a modified two step approach therefore the task could be solved on the basis of memory strength. The rats’ responses could have been based on differences in memory strength for what was encountered where and thus strong after the short retention interval and weak after the long retention interval. If this was the case, the rats were not showing integrated what-where-when memory in this set-up (Roberts et al., 2008). Follow up studies revealed that rats did not seem to use the “when” information but “how long ago” as a clue as to which food was to be found where (Roberts et al., 2008), but in a set-up where “how long ago” was made irrelevant they were able to use the absolute point in time the event happened, i.e. time of day (Crystal, 2009; Zhou & Crystal, 2009). Thus, rats may preferentially use relative cues, but are also able to remember the absolute time when of an event occurred.

Two species of food-hoarding birds have also shown successful performance with the two step what-where-when method: magpies (*Pica pica*) and black capped chickadees (*Poecile articipillus*) (Feeney et al., 2009; Zinkivskay et al., 2009). Also an ecologically inspired set-up with meadow voles (*Microtus pennsylvanicus*) used the two step method to demonstrate the voles’ ability to remember where and how long ago they have encountered a female in a particular reproductive state (Ferkin et al., 2007). The ability to remember the what-where-when has also been applied to a serial odour task with rats (Ergorul & Eichenbaum, 2004), as well as to a novel object exploration task in both rats (Kart-Teke et al., 2006) and mice (*Mus musculus*) (Dere et al., 2005b; Dere et al., 2005a). However, the novelty tasks may be solved by a familiarity process and therefore, may not be conclusively demonstrating that the animals are recalling integrated what-where-when features of experienced events. One novel object study has successfully controlled for solving the task just based on recognition of familiar objects by placing rats in the long arm of a T-maze with different backgrounds, and tested their choice for the two side arms before the rats could see the objects placed at the ends of those arms. Rats showed integrated memory for which objects were previously encountered where not at different points in time, but under different spatial backgrounds or contexts, demonstrating so called what-where-which memory (Eacott et al., 2005). The authors propose that studies could utilise the “which context” feature as a more general alternative to the “when” feature, as the particular context in which an object was encountered can serve as the distinguishing marker for this event.

Two separate studies have also applied the two step method to study episodic-like memory in non-human primates (Hampton et al., 2005; Martin-Ordas et al., 2010). Rhesus macaques (*Macaca mulatta*) were tested in an open field study and successfully remembered the locations of the two test foods even with intervals of 25 hours, but

they failed to switch to the less preferred food after the long retention interval (Hampton et al., 2005). The macaques were thus successful in remembering what was where, but failed with the when aspect (Hampton et al., 2005). This failure could, however, be attributed to certain aspects of the experimental set-up. Each test trial consisted of a study phase followed by two test phases, one and 25 hours after the study phase. In the study phase, the macaques received eight opportunities to recover both of the hidden foods and remember their locations, while in each test phase they received one such opportunity. Thus, in each trial the animals could recover their preferred food nine times, and only on the tenth encounter, which was 25 hours after the study phase, was their preferred food degraded. Perhaps the macaques were over trained on searching for their preferred food and were unable to switch to the less preferred one after the long delay, because of the fewer opportunities (one out of ten per trial) to learn that the preferred food had by then degraded. Further testing is needed to assess whether macaques are able to form and recall what-where-when memories in the domain of food. Rhesus macaques and chimpanzees have previously shown declarative memory competence by collecting more information (Call & Carpenter, 2001) or avoiding memory tests when unable to remember the presented stimuli (Hampton, 2001). A recent study examined what-where-when memory of rhesus macaques with a computer task (Hoffman et al., 2009). The authors argue that the tested macaques integrated the what, where and when information of the presented events, however, as each feature was examined with a separate question, further testing is required to confirm that the three features were in fact integrated into a single event by the animals (Martin-Ordas et al., 2010). Additionally, the paradigm relied on short retention intervals examining the animals working memory, while episodic memory is part of the long-term memory system.

Three species of great apes were also recently tested with food hiding events (Martin-Ordas et al., 2010), with the two step what-where-when method and an additional condition examining the memories structure. The apes were successful at choosing the location where the preferred perishable food was hidden after five minutes, and also switched to the location containing the less preferred non-perishable food after one hour. In a second test, the apes were able to distinguish between two separate baiting tables, baited at two different times, with the same two foods. The animals choose the perishable frozen juice from the table baited recently and the non-perishable grape from the table that was baited a while ago, indicating they distinguished which baiting table was baited when and will in consequence either contain the perishable food (the recently baited table) or only the non-perishable food (the table baited a while ago) (Martin-Ordas et al., 2010). As only two foods and two time intervals were tested, the same rule learning explanation as mentioned with the original scrub jay set-up cannot be excluded (Clayton et al., 2001b), although the rule would be employed with remarkable flexibility in the set-up with two different tables.

A three step set-up with a third food and time interval as was used with corvids, would conclusively exclude a rule learning strategy and confirm both the

content and the structure of what-where-when memories in primates. The ability of primates has not yet been investigated with a three step method.

WHAT IS KNOWN ABOUT PLANNING FOR FUTURE EVENTS IN ANIMALS?

Humans can mentally pre-experience or simulate the future consequences of their current decisions or actions, even of decisions they have never faced before (Gilbert & Wilson, 2007; Suddendorf et al., 2009a), by mental time travel into the future (Tulving, 2002; Tulving & Markowitsch, 1998). A crucial aspect of human planning is the ability to anticipate a need not currently experienced as well as a future oriented response that is not triggered by an innate behaviour, fixed action pattern or a learned association (Suddendorf & Busby, 2005; Tulving, 2005). Examples of planning by humans include packing appropriate items for a holiday destination, by considering the different potential weather conditions that could be encountered, or a more complex emergency evacuation of compounds, which involves coordination between several individuals as well as the consideration of several different outcomes.

Several authors have supported the Bisschof-Köhler hypothesis (Roberts, 2002; Suddendorf & Corballis, 1997), which states that the behaviour of animals is controlled only by their current motivational states and, therefore, they are unable to anticipate future motivational states (Suddendorf & Corballis, 1997; Suddendorf & Corballis, 2008). Although animals do, for example, migrate and hibernate, these behaviours are not driven by the animal's anticipation of the changing of temperature or coming of the seasons, but are innate responses (Paxton & Hampton, 2009). Animals can also anticipate future events based on their present motivational state, as rats and pigeons can learn to accurately time the interval between a cue and the delivery of food (Roberts, 1981; Roberts, 2002; Roberts et al., 1989). Rats for example also show anticipatory behaviour prior to expected positive events as transfers to enriched housing or mating opportunities (Harst et al., 2003). However, such responses are acquired through training and conditioning and do not require the ability to imagine the future. Similar criteria to those used with studies on what-where-when memory have been proposed for studying the anticipation of a future need, namely content, structure and flexibility (Clayton et al., 2003a). The content criteria requires that the animals anticipate the what-where-when of a future event, which is in turn structured into an integrated event and relies on flexible deployment of information (Clayton et al., 2003a). A small number of studies have examined the ability of animals to plan for future events; however, none of the described methods meets all three proposed criteria of content, structure and flexibility. At this early stage of testing animals' abilities to plan for a future event, studies focus on empirically evaluating the validity of the Bisschof-Köhler hypothesis, and thus on whether animals can anticipate a future need. The evaluation of the content, structure and flexibility of the anticipated future is to be confirmed in future studies. In the following overview I have for clarity grouped the methods used to study the prospective abilities of animals into two groups: the

satiation method and the spoon test method, but note that different studies use variations of these methods (Table 1.1).

The satiation method requires that animals choose between two options or two behaviours that will have different consequences for them in the future. These studies typically control that the need satisfied by the animal's future behaviour is absent at the time of choice (Osvath & Osvath, 2008). In other words, the animal's current need is different from the need they will experience in the future. This method then measures the animals' ability to distinguish between two behavioural options leading to different outcomes in the future. Scrub jays will preferentially cache food at the locations where they expect to be hungry the next time this location will be accessible to them and they even cache a specific food type at the location where they expect to be without this food type in the future (Raby et al., 2007b). Additionally, scrub jays are able to cache a food type they are currently satiated on (which would normally cause them to avoid it), when they have experienced that this food type will be again desirable in the future when they will be allowed to recover their caches (Correia et al., 2007). Furthermore, only scrub jays who have in the past pilfered others' caches, and not birds who have never been thieves themselves, will re-cache foods if they were observed by another individual while caching (Emery & Clayton, 2001). This suggests that the jays may anticipate the potential theft of their caches based on their experience as thieves. These scrub jay studies are considered as the strongest evidence of future planning in non-human animals (Eacott & Easton, 2010; Shettleworth, 2007a), however, they may be difficult to replicate with non-caching species.

Rats (*Rattus norvegicus*) and two monkey species, namely rhesus macaques (*Macaca mulatta*) and squirrel monkeys (*Saimiri sciureus*), were tested with the satiation method and two different options based on a similar principle (Naqshbandi & Roberts, 2006; Paxton & Hampton, 2009). The animals were tested on the ability to reverse their initial preference for a large amount of a desirable, but thirst inducing food, as this choice lead to a longer period without water. The choice for a smaller amount of the thirst inducing food on the other hand leads to an earlier return of water. Rats and rhesus macaques were unable to reverse their choice for the larger amount of food, except, in the case of rhesus macaques, when access to water was provided immediately after their choice (Naqshbandi & Roberts, 2006; Paxton & Hampton, 2009). However, success in this condition does not satisfy a future need, but an immediate need. In a separate study, two squirrel monkeys were able to reverse their choice from the large to the small amount of the food which was then followed by a short interval without access to water. Their success, however, may be attributed to the history of these particular individuals with extensive testing with inhibition paradigms (Naqshbandi & Roberts, 2006). In addition, the animals' reversal of choice was a gradual response after several trials, suggesting they may have learned the consequences of their choices and not planned for a future need (Shettleworth, 2007a). Summing up, the satiation method has led to successful results with the scrub jays, but has led to mixed results with primates.

The abilities of great apes to plan for a future need have been successfully tested with the so-called spoon test method. In this method the animal needs to make provisions in the present for a future need. The spoon test is based on the suggestion of Tulving (2005), who proposed that tests for future planning in animals could be based on the idea described in an Estonian children's tale. In the story, a girl attended a party in her dream, but was unable to enjoy the provided dessert, because she did not bring a spoon. The following night, the girl placed a spoon under her pillow. Based on this story, the animals should be tested on their ability to select a tool which serves no purpose at the time of choice and bring it after a delay to a location where the tool can be used to obtain rewards. Crucially, it is the animal's failure to initially obtain the rewards without the tool that should result in them bringing the tool on subsequent trials. Several studies have shown that orangutans (*Pongo pygmaeus*), bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*) are able to select and keep a tool for one hour in order to use it to obtain treats (Dufour & Sterck, 2008; Mulcahy & Call, 2006; Osvath & Osvath, 2008). These studies have been criticized to lack sufficient controls on promoting the animals to bring the tools or transfer effects between different tests (Suddendorf & Corballis, 2008; Suddendorf et al., 2009b), as well as for low consistency of the individual animals' success rates (Suddendorf, 2006; Shettleworth, 2007a). Additional conditions should be included to amend this. In a variation on the spoon test, both chimpanzees and orangutans were successful in selecting appropriate novel objects that fulfilled the function of a tool to use them an hour later to obtain treats (Osvath & Osvath, 2008). This condition with novel tools, selected in advance and with no visual access to the apparatus, is particularly convincing as evidence for great ape foresight, as it suggests the animals may be able to learn the relevant affordances of the tools, rather than just demonstrating simple associative responses (Osvath & Osvath, 2008). Additional evidence of great ape foresight has recently been reported in an observational study documenting the ability of a chimpanzee to spontaneously collect and save appropriate stones to later throw them at zoo guests in dominance displays (Osvath, 2009). However, caution should be taken when considering such anecdotal reports, as it is not documented how this behaviour started. It is therefore possible that the chimpanzee began collecting and storing the stones after initial reinforcement from the public's reactions to his displays and not based on previous failure and the envisioning of future encounters.

Lastly, a single study used an interesting approach based on the same principle as the spoon test (Dufour & Sterck, 2008). The exchange method incorporates a social component into the spoon test, where the animals need to select and keep objects to trade them an hour later with a human partner in order to obtain a reward (Dufour & Sterck, 2008). However, the tested chimpanzees were unsuccessful, possibly because the connection between the need to keep a specific object to exchange it for a reward in the future was too abstract. The spoon test method with a tool provides a more direct connection between the selected tool and its future application, which may be crucial for the animals to consider the tool as a solution to an anticipated problem.

The spoon test has been demonstrated as a useful method to examine the abilities of great apes to prepare for a future need, despite some of its mentioned weaknesses. The prospective abilities of monkeys have however, not yet been examined with the spoon test method.

AIMS AND OVERVIEW OF THESIS

Mental time travel, into the past by retrospection and into the future by prospection, has been stated to be unique to humans (Roberts, 2002; Suddendorf & Corballis, 1997; Suddendorf & Corballis, 2007a; Tulving, 1983; Tulving, 2002). This perspective is somewhat anthropocentric, as mental time travel in humans is defined as a subjective experience and requires verbal competence to be assessed. In order to empirically investigate the prospective and retrospective abilities of non-human animals, behavioural criteria have been formulated (Clayton et al., 2003a). Based on these, episodic-like or what-where-when memory and planning for a future need have been investigated primarily in birds and rodents (Crystal, 2010; Eacott & Easton, 2010; Salwiczek et al., 2010) and only to a limited degree and unsystematically in non-human primates (Hampton & Schwartz, 2004; Martin-Ordas et al., 2010; Schwartz & Evans, 2001). This is striking, since these abilities are considered to be the hallmark of human cognition and may have precursors or be present in non-human primates. In order to obtain a more comprehensive understanding of these capacities and to unravel their evolution, it is imperative that the range of the tested non-human primates is expanded and evaluated based on the criteria used in other animal studies. Likewise, the abilities of humans should be assessed with these behavioural paradigms in order to form a comparative basis in terms of prospective and retrospective behaviours.

The two main objectives of the present thesis are, firstly, to expand the understanding of these abilities in two representative species of non-human primates, namely the retrospective abilities of chimpanzees and the prospective abilities of long-tailed macaques. Secondly, to compare the performance of three and five year old children on both the what-where-when memory test as well as on Theory of Mind tasks in order to investigate the developmental relation between the two. The thesis thus takes both a comparative and a developmental perspective.

In **Tools**, I present a series of experiments investigating the ability of long-tailed macaques to transport tools for future use. This is the first study to use the spoon test method (Tulving, 2005) on a monkey species. In the first experiment, I test the macaques on a simplified version of the spoon test that has been previously successfully used with great apes (Dufour & Sterck, 2008; Mulcahy & Call, 2006; Osvath & Osvath, 2008). I then study whether macaques, trained to transport tools to immediately obtain rewards, are able to extend this behaviour, and transport tools after delays of increasing lengths. To further examine the animals' flexibility, I also test if they can transport novel tools for future use.

In **Treats**, I investigate the ability of chimpanzees to form and recall integrated memories of what-where-when events. I use the three step method (Clayton et al., 2003b) where three food types of different palatability and different rates of disappearance are hidden in trial unique locations and can be obtained by the chimpanzees after three different retention intervals. I evaluate the chimpanzees' performance with regard to integrating the what-where-when features of the food hiding events as well with respect to a developed location-based association strategy. This is the first study to examine the abilities of chimpanzees to form what-where-when memories with a three step method, which controls for a rule learning explanation of the animals performance.

In **Toys**, I use a developmental approach in examining the ability of three and five year old children to remember details of past events. In order to examine the children's development of what-where-when memories I designed a story test based on the principles of the Clayton et al. (2001b) three step method. Additionally, I evaluate the children's performance on a battery of Theory of Mind tests in order to compare the developmental onset of the two abilities. This is the first study to look at the development of what-where-when memory with a three step method in children and to relate this developmental onset to that of Theory of Mind abilities.

In the **Discussion**, I summarize the obtained results and discuss the implications for future research. Mental time travel, with its past and future oriented components, was originally defined and long assumed a uniquely human faculty. The proposal some 20 years ago of behavioural criteria describing these abilities has made it possible to investigate the presence of these and similar abilities in other species. Although the flexibility and the diversity of contexts in which humans apply this ability is remarkable and probably unattainable for other species, further comparative and developmental research is necessary to understand the crucial components and necessary conditions for non-human animals displaying behaviour similar to our prospective and retrospective abilities. Eventually this may lead to a much more nuanced understanding of how humans came to have and excel at mental time travel.

TOOLS

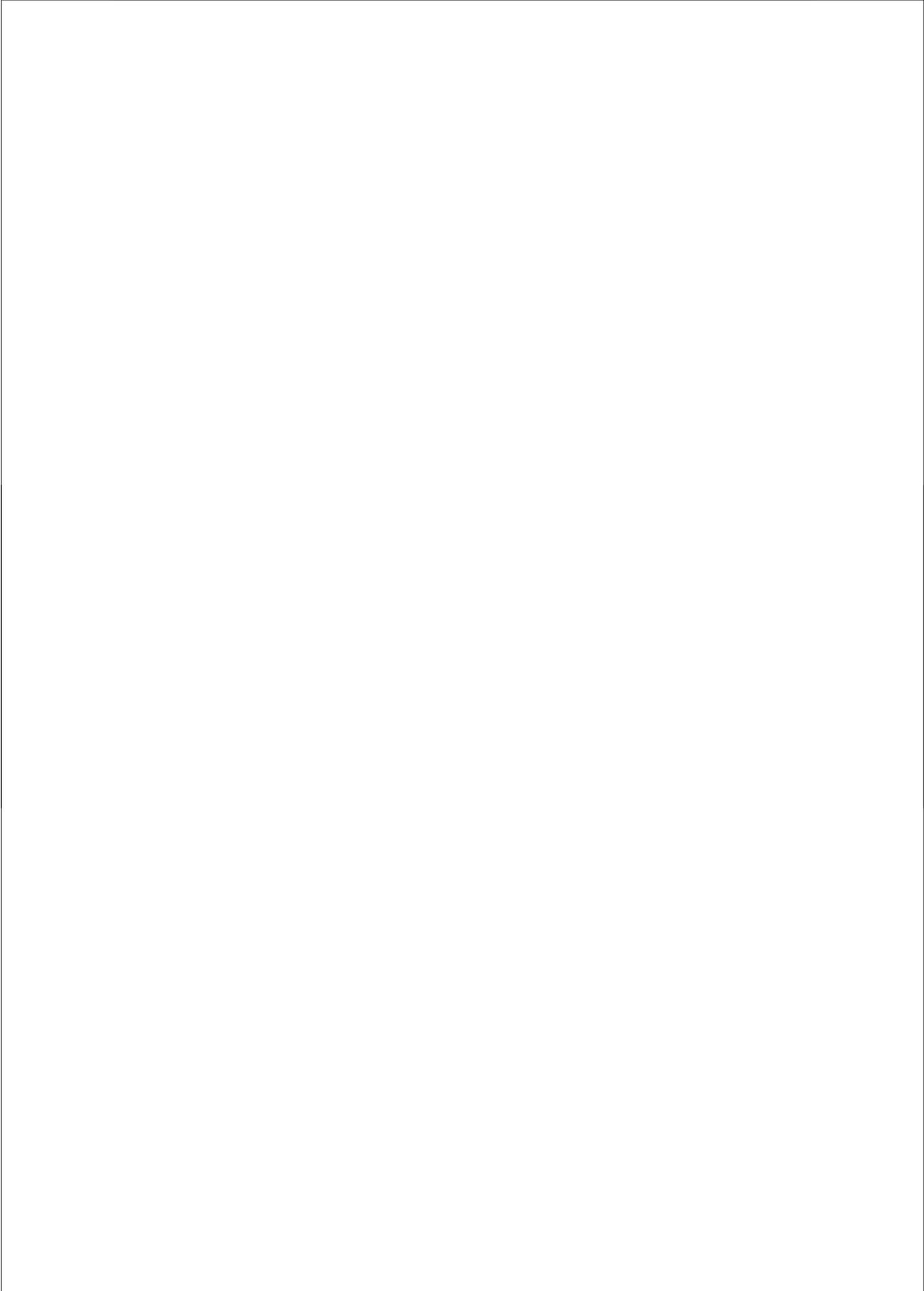
Transport of Tools for Future Use by Long-tailed Macaques (*Macaca fascicularis*)

ABSTRACT

Future planning has been considered a uniquely human ability. However, certain animals can foresee some of their future needs. We tested the ability of long-tailed macaques (*Macaca fascicularis*) to transport and use a rake in the future in three experiments. In the first experiment, we tested whether monkeys would select a rake and after a delay, transport and use it to obtain rewards. We manipulated the spatial and temporal distance between rake provision and reward availability in several tests, however all six macaques failed to transport the rake in all tests. In the second experiment, we trained five of the macaques to transport the rake to obtain immediate rewards, and then tested them with increasing delays. Three of the macaques were able to transport and use the rake after a delay of 5 minutes; one was successful after 10 minutes and also transported the tool on a few trials after a 20 minute delay. In the third experiment we tested the ability of two of the macaques to transport novel tools for future use. The best performer from the second experiment successfully selected and transported novel tools. The macaques failed to plan in the first experiment, potentially because they were unable to foresee the appropriate sequence of behaviours to solve the task. After training, the animals displayed flexibility by successfully transporting tools after delays. Remarkably, one macaque also generalized this behaviour by transporting appropriate novel tools and several attempted to use non-provided objects or their tails. Our results show which information and experience macaques need in order to display tool transport for future use.

Under revision by Behavioural Process as:

*Dekleva M., van den Berg L., Spruijt B. M. and Sterck E. H. M. Take it or leave it: Transport of tools for future use by long-tailed macaques (*Macaca fascicularis*).*



INTRODUCTION

Humans can imagine future events as well as reconstruct past ones by means of mental time travel (Tulving, 2002; Tulving & Markowitsch, 1998). This complex cognitive ability enables optimal decision-making in the present by anticipating future consequences in a variety of domains. The extent to which other animals are also able to plan for future needs remains controversial (Crystal, 2009; Suddendorf & Corballis, 1997; Clayton et al., 2003a; Roberts, 2002). The Bischof-Köhler hypothesis maintains that non-human animals, are unable to disassociate themselves from the present to anticipate a future need which they are not currently experiencing (Suddendorf & Corballis, 1997; Paxton & Hampton, 2009). However, this view does not mean that animals are unable to learn to press a lever to obtain delayed rewards, nor that they are unable to display species specific innate behaviours with future benefits such as nest building or hibernation (Roberts & Feeney, 2009). Such behaviours are either fixed responses or are gradually conditioned and reinforced to secure a current need, but they are not the result of anticipating the future, e.g. birds building a nest probably do not foresee that this will serve for laying and hatching of eggs. In order to show that animals are able to foresee solutions to a future task and thus plan, their behaviour should reveal that they are, based on previous unreinforced experience, able to anticipate a need not currently experienced and act in the present to secure it in the future (Suddendorf & Corballis, 2010). Several studies with scrub jays, a food caching corvid, have shown that these animals are able to preferentially cache food at locations where they expect to be hungry the following morning, and even that they can cache a given food type they are currently satiated on, as they learn that in the future that food type will be the desirable one (Raby et al., 2007b; Correia et al., 2007). Scrub jays have also been extensively tested on their abilities to recall what-where-when or episodic-like memories of food caching episodes (Clayton, 1998; Clayton et al., 2001b; Clayton et al., 2003a) an ability that in human subjects shares cognitive resources with planning (Busby & Suddendorf, 2005; Okudaa et al., 2003). Contrasting results have been obtained in two somewhat similar satiation studies with monkeys, where the animals needed to reverse their choice of a large amount of a thirst inducing food, as this was followed by a longer interval without access to water than if they chose the smaller amount of the food. While two squirrel monkeys could reverse their choice (Naqshbandi & Roberts, 2006) rhesus macaques seem unable to do so (Paxton & Hampton, 2009).

Tulving has proposed that future planning could be tested with the spoon test (Mulcahy & Call, 2006; Roberts, 2002; Suddendorf & Busby, 2003). This is based on an Estonian folktale in which a girl attended a party in a dream but was unable to enjoy the dessert because she did not bring a spoon. The following night the girl therefore placed a spoon under her pillow. Tests inspired by this story have examined the abilities of great apes to transport tools for future use. Orangutans and bonobos selected and transported tools for use 14 hours later and stopped bringing the tool once the test apparatus was no longer made available, despite being rewarded when

they brought the tool (Mulcahy & Call, 2006). This indicates that they anticipated the future function of the tool, and stopped bringing the tool when its function was no longer fulfilled. Similarly, chimpanzees also brought a tool in order to use it following a delay of an hour (Dufour & Sterck, 2008). In addition, chimpanzees and orangutans have been shown to reject a favoured small reward and instead choose a tool they could use an hour later to obtain a large reward, as well as to select and transport novel tools (Osvath & Osvath, 2008). Outside a controlled experimental setting, a male chimpanzee was reported to spontaneously collect and store stones in order to later throw them at zoo visitors in dominance displays (Osvath, 2009). This behaviour is interpreted as evidence of advanced consciousness and the ability to mentally represent an upcoming event (Osvath, 2009). As the subjects in all of the experimental studies of great apes received some training and were tested with repeated trials, it is difficult to determine whether they solve the tasks by associative learning or by foresight alone (Suddendorf & Corballis, 2008). In addition, primate studies of tool use acquisition and affordances have particularly stressed the importance of the influence of additional factors crucial to the animal's success. Individual differences in rearing history, prior experimental experience and possibly also motivation, social status and dexterity have produced contrasting results about the ability of certain species to use tools and understand their functions, in some cases even within the same species (Shettleworth, 1998; Tomasello & Call, 1997; Seed & Call, 2010). Therefore, detailed studies are necessary to understand the characteristics of future oriented behaviour in primates, their flexibility and their limits instead of merely focusing on all or nothing responses (Sterck & Dufour, 2007).

Non-human primates are, together with corvids, exceptional as they use tools in a range of categories, whereas the majority of other animal species are limited to using tools in a single category (Bentley-Condit & Smith, 2010; Beck, 1980). Several species of macaques, baboons and capuchins use tools with great skill in captivity and have occasionally been observed to use tools in the wild (Malaijvitnond et al., 2007; Tomasello & Call, 1997; Watanabe et al., 2007). For example, three macaque species have been observed in the wild to use leaves for both grooming and feeding, while six species held in laboratory settings have been reported to use sticks for grooming and feeding by pulling, extracting and sweeping rewards within reach (Tomasello & Call, 1997). Although it is difficult to identify what exactly causes such discrepancy between wild and captive populations, it remains attractive to study exactly how and what monkeys can learn about the application of tools in captive populations.

Several studies have shown Japanese macaques (*Macaca fuscata*) are able to generalize tool use in novel contexts. Once these monkeys mastered rake use, they were also able to obtain rewards with novel tools (Ishibashi et al., 2000). Japanese macaques were able to solve a sequential rake task (Hihara et al., 2003) and in a separate study, after the animals learned to extract food from a pipe with sticks they also learned to throw stones into the pipe to push the foods out the far end (Tokida et al., 1994). Initial tool use in these studies was acquired over repeated trials and thus, most likely by means of trial and error learning. However, following the large number

of trials required to master the initial task, the monkeys showed remarkable flexibility and increased speed when learning to use the tool in novel solutions (Hihara et al., 2003). In addition, while learning how to use the rake, monkeys show growth in brain areas corresponding to those activated in the human brains during tool usage (Quallo et al., 2009).

Based on such findings of macaque behavioural flexibility in rake tasks, we aimed to investigate whether long-tailed macaques would perform similarly to apes in tasks involving planned tool transport for future use. When faced with a choice between a less favoured immediate reward and a preferred delayed reward under comparable conditions, long-tailed macaques show self control (by waiting about 30 seconds), while pigeons and rats show impulsiveness (Tobin et al., 1996). Recent studies with long-tailed macaques have in addition shown that they are able to delay gratification for up to 10 minutes, an indispensable prerequisite for making future oriented choices (Pelé et al., 2010). On this basis, long-tailed macaques are an interesting species of monkeys in which to study tool transport for future use. The present study addresses two major questions. In our first experiment we investigated the conditions under which macaques would select and transport a familiar tool for later use. Six animals were trained to use a simple raking tool. We then exposed these macaques to a series of conditions where we progressively reduced the spatial and temporal distance between tool provision and availability of rewards. By manipulating features of the set-up, we aimed to determine which information or experience was necessary to elicit the target behaviour of tool transport. In our second experiment we addressed the animals' ability to generalize the transport of tools. Five of the macaques were trained to transport the rake for immediate use and were then tested on their ability to generalize such trained transport. We first introduced and later extended the delay between tool provision and reward availability and measured the animals' success rate. Finally in the third experiment, we examined whether two of the animals would also be able to generalize to an additional context and transport novel tools for future use.

MATERIALS & METHODS

Subjects

The six participating animals (see Table 2.1) were socially housed in a group of eighteen individuals (fifteen females and three males). The group was socially stable and housed at the Ethology station at Utrecht University, the Netherlands. This study was approved by the Animal Experimentation Committee of Utrecht Medical Centre under study number 2008.I.07.05.

Table 2.1: Subjects and their participation in training and testing in all three experiments.

Each individual's age and gender are given, together with the tests and training stages they participated in. Tests 1.1-1.4 refer to experiment 1, tests 2.1-2.3 refer to experiment 2, and 3 refers to experiment 3.

Subject	Gender and age	Participated in tests	Rake trained	Transport trained
Anastasia	F (4 years)	1.1-1.4, 2.1	x	x
Ekzekwo	M (8 years)	1.1, 1.2, 1.4	x	
Era	F (20 years)	1.1-1.4, 2.1	x	x
Icetea	F (12 years)	1.1-1.4, 2.1	x	x
Ophelia	F (3 years)	1.1, 1.2, 1.4, 2.1-2.3, 3	x	x
Zargasso	M (4 years)	1.1-1.4, 2.1-2.3, 3	x	x

Enclosure and husbandry regime

The group's enclosure consisted of two inside rooms (referred to as the home cage and the test cage) and an outside area, connected with bridges and trap doors. Both inside rooms could be further divided into sub-compartments. A drawing of the test cage and its sub-compartments is shown in Figure 2.1. Outside the training and testing sessions the animals had access to all areas. During training and testing the experimenters could limit the animals' access to the different rooms and sub-compartments. Water was available *at libitum*. The animals were fed monkey chow daily approximately 30 min after the training. Because the chow was dispersed across the home cage in liberal amounts on a daily basis, the animals had persistent access to food while in the home cage. The animals only received treats (e.g. fruits, breakfast cereals and sprinkles) as part of the training and testing. All training and testing was performed by two experimenters who were equally familiar to the animals. Only positive reinforcement was used and the animals could chose not to enter the test cage.

Data analysis

The animal's behaviour was scored directly during each test trial as well as videotaped. For the analyses, the number of occurrences of each relevant behaviour was obtained from the score sheets. Latencies and durations of behaviours were obtained from videos. We analyzed each individual's behaviour separately, reasoning that the success of one individual would be of interest. Statistical analyses were performed in SPSS 16. Due to small sample sizes we used non-parametric tests. We corrected for multiple comparisons with the standard Bonferroni correction. All statistical tests used a critical alpha of 0.05 and were two-tailed, unless otherwise stated.

EXPERIMENT 1: SELECTION AND TRANSPORT OF A FAMILIAR TOOL

In this experiment we investigated the factors that would facilitate the selection and transport of a tool to satisfy a future need. We tested subjects proficient in rake use. In the first test we aimed to replicate the test settings previously used with great apes (Mulcahy & Call, 2006; Dufour & Sterck, 2008; Osvath & Osvath, 2008), with a modified

temporal component. In two further tests we manipulated different features of the set-up to see how they would affect the macaques' success at rake transport. Finally, we assessed the macaques' preference for the rake across different contexts to determine whether the animals only showed a preference for the rake in the context in which it was previously trained.

Raking training

Prior to testing six macaques were individually trained to rake in treats on a test table. The rake was an L-shaped metal sticks approx 45 cm long, the test table was approx 60x120 cm. Training progress was measured in three stages. In the sessions, the experimenter demonstrated the use of the tool, encouraged the monkey to touch the tool and assisted it to manipulate the tool and obtain the treats. Stage one was completed once the macaque successfully raked in treats placed in front of the tool. Stage two was completed once the macaque repositioned the tool behind the next treat in order to obtain it. Stage three was completed when such repositioning was achieved on two consecutive sessions. Each training session lasted as long as the animal's motivation could be maintained (approx 10 min). The animals were initially trained once and eventually twice per day. Additionally, the animals were habituated to the tools and the distracter objects by placing several copies of each inside their enclosure on repeated occasions during weekends and holidays when no training took place. The distracter objects (see test 1.1 for explanation of their function) were a gray plastic tube (10 cm in length, 2 cm in diameter) and a metal plate (10 in length, 4 cm in width). The animals received continuous training between the different tests of the experiment to maintain their familiarity with the task.

All subjects completed stage one, which was the stereotyped raking in of the treats placed in front of the rake, in their first or second session. Subjects completed stage two, the repositioning of the rake between their third and 21st session. Stage three, which was the consistent repositioning of the rake on two consecutive sessions, was completed between the seventh and the 34th session. The individuals greatly differed in the number of sessions they needed to achieve raking proficiency. All completed stage one in few sessions, but required a large number of sessions between the times they first managed to reposition the rake until they consistently repositioned it in two consecutive sessions, indicating the difficulty of the task. The six macaques passed stage three on average in 21 sessions and achieved tool proficiency.

Test 1.1: Selection and transport

In the first test we investigated whether the animals would preferentially select the rake over familiar distracter objects, transport the rake from the selection compartment to the waiting compartment and then after 5 minutes transport the rake once more to the test compartment to obtain treats (see Figure 2.1). Importantly, the test table with rewards was not visible from the selection compartment and the animals were not hungry at the time of rake selection. As the animals were familiar with the use of the tool, we expected that the repeated experience of failure to obtain

rewards would lead them to solve the task. If they were able to solve this task, it would indicate that macaques were able to use foresight to anticipate a need or motivation not currently experienced and take action in the present to satisfy the need in the future. If the task was solved based on a previously unreinforced experience (the failure to obtain rewards) associative learning could not be considered a likely explanation.

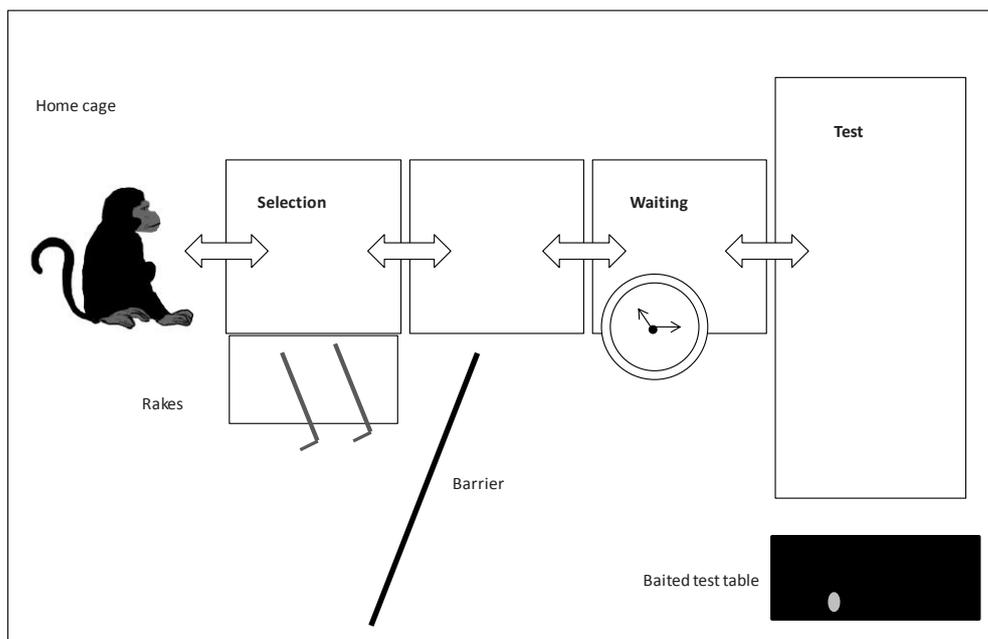


Figure 2.1: The test cage.

A drawing of the different sub-compartments of the test cage and the activity that took place in each one. The drawing is not to scale.

Test 1.1: Methods

The entire social group was restricted to the home cage from where a test animal individually entered the selection compartment of the test cage (Figure 2.1). Once the subject was in the selection compartment, the experimenter placed two rakes and four distracter objects (two of each kind) on a shelf in front of this compartment. Objects were presented in duplicates to insure that the animal had a back-up copy of each object, as it had to pull the objects inside through a mesh wire which required some manipulation. The objects were presented in a pseudo-randomized array and within reach of the subject.

A barrier prevented visual access from the selection compartment to the part of the test compartment where testing took place. After 2 minutes, the shelf and the remaining objects were removed and the subject was given access to the waiting compartment. If the monkey did not enter the waiting compartment within 5 minutes the trial was scored as incomplete. From incomplete trials, we scored subjects'

relevant behaviour in the selection compartment. If a subject entered the waiting compartment, it received a small treat and remained there for 5 minutes while the experimenter left the area. Following this delay, the experimenter returned with the test table and gave the subject access to the test compartment. The experimenter continually placed a treat on the test table approximately every 10 seconds, which the subject could only obtain if they had brought the rake. If the treat was not obtained within 10 seconds, it remained on the table and a new treat was placed next to it. The trial ended when all the treats were consumed or after 4 minutes. Each animal was tested twice a day (with a minimum of 2 hours between the two trials) on four consecutive days followed by a two day break, one day of training and four consecutive testing days. In total, all animals received sixteen trials. We considered an object to be selected when the animal pulled it inside the selection compartment. For the analysis, we counted the number of trials in which the rake was selected from the total trials, and the number of trials in which the rakes were transported to the waiting and testing compartments from completed trials. We tested all six individuals.

Test 1.1: Results and discussion

None of the animals selected the rake in more of the trials than expected by chance (Chi-square test, expected chance: 25%, $n=16$ trials, all individuals exact $p > 0.143$, Figure 2.2). They did show some interest towards the rakes as five out of six animals (all except Ekzekwo) selected them on at least one trial. All animals selected the tube more frequently than the rake. The tube was possibly preferred over the rake as it was plastic and chewable, an activity consistently performed by all individuals. On average, the animals entered the waiting compartment and completed 78% of trials (sd 42%). None of the animals ever brought the rake to either the waiting or the test compartment in any of the completed trials. The animals did transport the tube to the waiting and testing compartments, however, none tried to use it to obtain treats.

The animals were thus able to transport objects, but perhaps not with the intention of using them later on. The macaques appeared unable to connect two temporally separate events (the rake presentation and the test table presentation) and thus failed to select and transport the rake. Their failure could be caused by their inability to anticipate a future motivational state or by the inability to connect the large temporal and spatial distance between the tool provision and test table provision. Importantly, if the animals were to solve the task based on the memory of their own personal previous failure they had to during tool selection remember the experience of not being able to obtain treats in the test compartment, as well as understand that their failure was caused by the missing tool. We addressed this in our second test.

Test 1.2: Effect of cueing on selection and transport

In the second test, we investigated whether the animals were more likely to select and transport the rake if we showed them the baited test table immediately prior to object selection. Providing them with this cue could assist them in connecting the rake presentation with the temporally separate test table presentation in the test

compartment and affect their ability to select and transport the rake. Specifically, the presentation of the table could trigger the memory of their previous failure and stimulate the animals to take the rake with them, if they were indeed aware that the missing rake was the cause of their failure in the test compartment.

Test 1.2: Methods

The testing procedure was similar to test 1.1 except that prior to the presentation of objects in the selection compartment, the test table with treats was shown to the subject in front of the selection compartment. The remainder of the procedure was the same as in test 1.1. The animals received a total of sixteen trials presented in the same way as in test 1.1., except for one subject (Era) who received fifteen trials as one trial was terminated due to interference from a neighbouring group. For the analysis, we again counted the proportion of total trials in which the rake was selected, and the number of completed trials in which the rakes were transported to the waiting and testing compartments. We tested all six individuals.

Test 1.2: Results and discussion

We again assessed if the animals selected the rakes in more trials than expected by chance (Figure 2.2). Only three of the animals (Anastasia, Era and Ophelia) selected the rake in at least one trial. Two of them (Anastasia and Era) did not select the rake above chance levels (Chi square, $n=16$ trials, both animals exact $p > 0.143$), while the third (Ophelia) initially appeared to select the rake above chance (Chi square, $n=16$ trials, exact $p=0.001$). However, she selected a large number of all objects and thus failed to show object selectivity. On average, the animals entered the waiting compartment and completed 67% (sd 21%) of trials. None of the animals transported the rake to either the waiting or the test compartment in any completed trial.

The sight of the baited test table had no effect on the subjects' success rate of selecting and transporting the rake and was not effective in helping them to connect rake presentation with the final test table presentation in the test compartment. While in the test compartment, the animals did show motivation to obtain the rewards, by repeatedly reaching for them. As the cueing had no effect on the animals' success rate, we theorized that bringing the rake provision and test table availability closer together spatially and temporally could assist the animals in solving the task. We addressed this option in the following test.

Test 1.3: The run-back

In the third test, we simplified the set-up by placing two rakes inside the waiting compartment, eliminated the delay in the waiting compartment and provided the animals with continuous access to the rakes while the test table was present. Rake presentation and test table presentation were now occurring virtually simultaneously, thus simplifying the task considerably. We predicted that if the animals did not transport the rake on their first entry to the test compartment, that they would return and fetch it after seeing the rewards placed on the test table. The animals did not need

to use foresight to solve the task in this set up. A realization that moving the tool from one compartment to the other would solve the problem was sufficient.

Test 1.3: Methods

The test procedure was similar to test 1.1 except that the subject was given access to the waiting compartment after 1 minute in the selection compartment where no objects were provided. A minute after entering the waiting compartment, two copies of the rake were inserted directly into this compartment, the door to the test compartment was opened and the test table was provided. The experimenter continued to place treats on the test table every 10 seconds, for 5 minutes and the subject had continuous access to both the waiting compartment with the rakes and the test compartment. All animals again received a total of 16 trials presented in the same way as in test 1.1. As a measure of interest in the rakes, we measured the number of completed trials in which the animals manipulated the rakes in the waiting compartment. Any rake oriented behaviour such as touching, biting, licking or transporting was considered manipulation. We scored the number of trials in which the animals transported the rake to the test table and the number of completed trials in which the animal returned to obtain the rake once the table was presented. We tested four individuals as the other two had not yet proceeded to this test stage.

Test 1.3: Results and discussion

The subjects manipulated the rakes in the waiting compartment in 42% of the trials on average (sd 24%), displaying some interest in them (Figure 2.2). On average, the animals entered the waiting compartment and completed 95% (sd 6%) of the trials. However, none of the animals brought the rakes to the test compartment in any of the completed trials. In addition, none of the animals returned to fetch the rake while the table was presented.

Although the two events, rake presentation and test table presentation, occurred virtually at the same time, the animals still failed to transport the rake. The macaques were unable to solve the task of tool transport, even in the absence of a delay. They were unable to use the memory of their previous failure to obtain the rewards and were unable to fetch the rakes when confronted with their failure to obtain the treats.

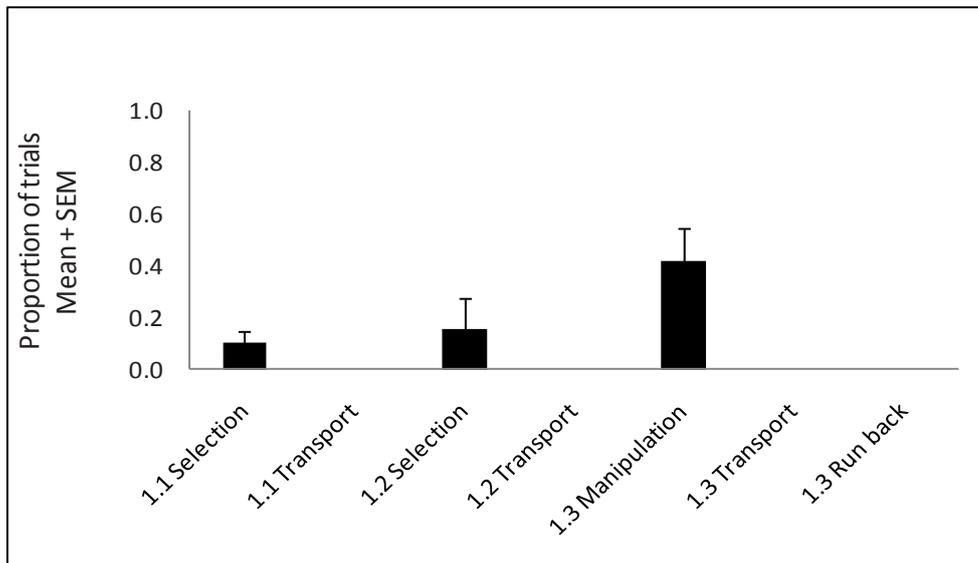


Figure 2.2: Results from tests 1.1, 1.2. and 1.3.

For all tests Mean+SEM of all individuals are shown. For tests 1.1 and 1.2 the proportion of total trials in which the monkeys selected the rake is shown and the proportion of completed trials in which the monkeys transported the rake to the waiting and testing compartments. For test 1.3 the proportion of completed trials in which the rake was manipulated in the waiting compartment is shown, together with the proportion of trials in which the rake was transported to the test compartment and the proportion of completed trials in which the monkeys fetched the rake.

Test 1.4: Context dependent tool preference

In the final test we wished to determine if the animals' failure to transport the rake could ultimately be attributed to a context specific preference for the rake. If the animals only showed interest in manipulating the rake in the context in which they were originally trained to it, i.e. on top of the baited table, and showed no interest in manipulating the rake in a different context, this would explain their inability to solve the novel problem of rake transport. Their failure to transport the tool could then be attributed to a lack of interest in the rake outside the original context. We presented the rakes and distracters in three different contexts: on the shelf in front of the selection compartment, on a non-baited test table in front of the test compartment and on the baited test table in front of the test compartment.

Test 1.4: Methods

We compared the animals' responses to two copies of the three objects (rake, tube and plate), across three different contexts. Each subject received two trials per day, with a total of twelve trials in each context. For the first context (on the shelf in front of the selection compartment) we used the animals' responses from the first twelve trials from test 1.1. For the remaining two contexts (non-baited and baited test table) we presented the objects on the test table in front of the test compartment. We measured the subjects' latency to interact with the rake and the duration of rake manipulation in

the three contexts within the first 2 minutes of presentation. We tested all six individuals.

Test 1.4: Results and discussion

We compared the duration of rake manipulation within the first 2 minutes of presentation across the three presented contexts. The three contexts had a significant effect on duration of rake manipulation (Friedman test, all individuals n=12 trials, exact p=0.001) and all values remained significant after Bonferroni correction. The duration of rake manipulation was largest in the last context; the baited table in the test compartment (Figure 2.3). A similar pattern was found when we analyzed the animals' latency to interact with the rakes across the three contexts: there was a significant difference between contexts (Friedman test, all individuals n=12 trials, exact p=0.001) with the latencies being lowest in the context of the baited table. All values remained significant after Bonferroni correction.

All animals demonstrated a context specific preference for the rake, reflected both in the duration of manipulation and the latency to manipulate the rake. Only in the context where the rake was a direct means to an end and actually fulfilled the role of a tool did the animals manipulate it more. Interestingly, this preference was also shown in the animals' latency to interact with the tool. Thus, as soon as a treat was placed on the test table, the animals exhibited a preference for the rake. The rake was not an object of preferred interest outside the context in which the animals were trained to use it, possibly because it did not fulfil the function of a tool outside this context. This lack of interest in the rake outside the context, in which they were originally trained to use it, may have prevented the animals from solving the tool transport task as they failed even to manipulate the rake when it was presented in a different context.

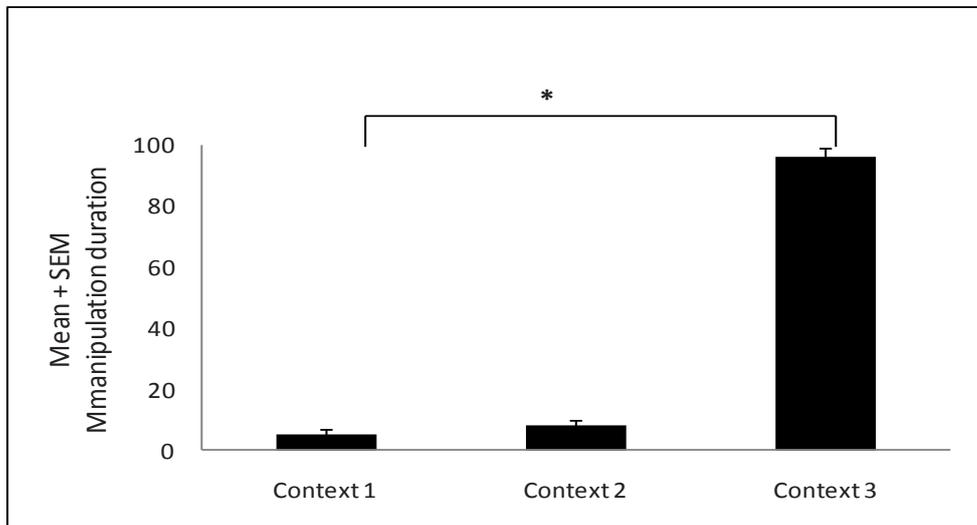


Figure 2.3: Results from test 1.4.

Mean+SEM duration of rake manipulation for all individuals in all three contexts is shown. The first context was on the shelf in front of the selection compartment, the second was on a non-baited test table in front of the test compartment, and the third was on a baited test table in front of the test compartment. Significant difference between the three contexts is indicated by *.

EXPERIMENT 2: TOOL TRANSPORT AFTER DELAYS

In our second experiment we first trained the macaques to transport the rakes to obtain immediate access to rewards, as we were unable to elicit such transport behaviour in the first experiment. We then examined how well the animals were able to generalize this learned association. We first introduced and later increased the delay over which the animals had to keep the tool in three tests. In this way we could determine if the animals were unable to solve the task in the previous experiment because they were unable to foresee the correct behavioural sequence and we could also examine the flexibility of the learned transport behaviour.

Rake transport training

We trained the animals to transport the rake from progressively greater distances towards the baited test table to obtain treats. Subjects were individually trained in three stages. Initially, the subjects were rewarded for any appropriate manipulation of the rake. The animal was verbally encouraged and when necessary assisted by the experimenter to manipulate the rake and obtain rewards. Each training session lasted as long as the animal's motivation could be maintained (approx 10 min). To complete the first stage the subject had to push out a partially inserted rake onto the test table and obtain treats. To complete stage two it had to transport the rake from the waiting compartment to the table in front of the test compartment. Finally, to complete the third stage it had to transport the tool from the selection compartment to the test

compartment. The animal was considered proficient once it directly transported the rake from the selection to the test compartment and used it on at least 5 out of 6 trials. All subjects successfully completed such rake transport tests, prior to each of the remaining tests. The animals continued to receive tool transport training in between all tests to maintain their interest in the task.

All animals completed stage one in their first session. The training of one subject (Ekzekwo) was terminated, since after fifteen sessions he never progressed to stage two. This individual also needed the highest number of sessions to acquire rake use in the first experiment. The remaining five individuals completed stage two between their third and fifth session. The animals required between six to twelve sessions to complete stage three. On average the five animals completed the last stage and achieved proficiency in seven sessions.

All of the subjects (excluding Ekzekwo) progressed faster through the stages of transport training than in initial raking training. This fast progression, compared to the animals poor performance in experiment 1, indicates that once the necessary sequence of behaviours was broken up into smaller rewarded actions the animals were able to learn them quickly and link them together over increasing distances. However, the animals still required several sessions to achieve proficiency, indicating the difficulty of the required actions, and emphasizing the importance of progressing in small increments for successful learning.

Test 2.1: Transport after 5 minute delay

Here we examined whether the animals would transport the rake from the selection compartment to the waiting compartment and after a 5 minute delay transport it on to the test compartment. This would demonstrate whether the subjects were now able to transport a tool when forcibly delayed and thus apply the learned behaviour under the novel conditions.

Test 2.1: Methods

The procedure was essentially similar to test 1.1, except that two copies of the rake were inserted into the selection compartment and the door to the waiting compartment was opened after 5 seconds. If an animal did not enter within 5 minutes, the animal was released back to the group and the trial was scored as incomplete. Each animal had to wait for 5 minutes in the waiting compartment before the door to the test compartment was opened. The trial ended 1 minute after the experimenter started placing a reward on the table every 10 seconds. After the trial was completed, the animal was released back into the selection compartment for an inter-trial interval of at least 10 seconds and then given another trial. Each subject was given a day of training and then three consecutive days of tests, with five trials in a row each day. This procedure was repeated over two separate weeks (with four daily trials in a row the second week) with further transport training in between. Each subject thus received a total of 27 trials. We scored the number of times the rake was transported to the waiting and testing compartments in the completed trials. We also determined

whether the animals were holding the rake when the door to the waiting compartment was opened and the duration of rake manipulation as a proportion of the total time spent in the selection compartment. Lastly, we also scored whether the animals were holding the rake when the door to the test compartment was opened or whether they picked up it up after the door was opened in order to transport it to the test compartment. We tested five individuals (since Ekzekwo did not reach stage 2 of rake transport training).

Test 2.1: Results and discussion

On average, the animals completed 67% (sd 21%) of the trials. One animal (Ophelia) transported the rake to the waiting compartment significantly more often than expected by chance (Binomial test, chance level 50%, 15 of 18 trials, exact $p=0.008$), a finding that remained significant after Bonferroni correction (Figure 2.4). Although the remaining individuals did transport the rake in several trials, these values did not reach significance (Binomial test, all individuals exact $p>0.523$).

We investigated which behaviours predicted whether an animal brought the rake from the selection to the waiting compartment. We considered two possibilities: holding the rake at the time that door to the waiting compartment was opened, and amount of time the rakes were manipulated in the selection compartment. For the first behaviour (holding the rake when the door was opened), we counted the number of trials in which each individual was: manipulating the rake when the door was opened and then brought it with them; not manipulating the rake but did bring it with them; manipulating the rake but then did not bring it with them, and neither manipulating the rake nor brought it with them. A Chi-square test for a 2 by 2 cross table, showed a significant relationship between the two behaviours for two of the individuals (Anastasia: $n=27$, $\chi^2=13.595$, exact $p=0.001$, Era: $n=27$, $\chi^2=6.998$, exact $p=0.030$). The value for only one of these individuals (Anastasia) remained significant after Bonferroni correction. This indicates she was more likely to bring the rake if she was manipulating it when the door was opened and more likely to leave it behind if she was not manipulating it when the door was opened. For the remaining individuals no relationship between the two behaviours was found (Chi-square test, all individuals exact $p>0.127$). For the second potential predictive behaviour of rake transport (amount of time the rakes were manipulated in the selection compartment), we compared the average proportion of time spent manipulating the rake in the selection compartment between trials where the animals brought the tool to the waiting compartment (tool manipulated on average 66% of the time, sd 30%) and those where they did not bring it (tool manipulated on average 24% of the time, sd 24%). We found that there was a significant difference between these trials for the group of all 5 individuals (Wilcoxon test, $n=5$, $Z=-2.023$, exact one-tailed $p=0.03$). Of the two behaviours we considered as predictors for bringing the rakes, rake manipulation time while in the selection compartment was of influence for all individuals, while manipulating the rake at the moment the door was opened was only a significant predictor for one individual.

We further investigated whether bringing a rake to the waiting compartment was related to then transporting it and using it in the test compartment (Figure 2.4). On average, the animals brought a rake to the test compartment on 97 % (sd 4%) of the trials in which they had it in the waiting compartment. The animals always used the rake when they brought it to the test compartment. We analyzed the number of trials in which the rake was brought from the waiting to the test compartment and found that three animals transported the rake more often than expected by chance (Binomial test, Anastasia: 12 trials out of 13, exact $p=0.003$, Ophelia: 14 trials out of 15, exact $p=0.001$, Zargasso 11 trials out of 11, exact $p=0.001$), even after Bonferroni correction. Three individuals (Anastasia, Era and Icetea) were successful in the first trial, while the other two (Ophelia and Zargasso) were first successful in their second trial. The highest number of successful consecutive trials was five, a score achieved by two animals (Anastasia and Zargasso). The values for the remaining two animals did not reach significance, even though they always transported and used the rakes in the test compartment, due to the low number of trials in which they transported a rake to the waiting compartment (Binomial test, both animals 5 trials out of 5, exact $p=0.062$).

Taken together, three of the animals were able to transport the tool from the waiting compartment to the test compartment after a 5 minute delay and were more likely to transport the tool to the waiting compartment the longer they spent manipulating the tool in the selection compartment. On average all animals transported the rake from the selection compartment and then on to the test compartment after a 5 minute delay in 53% (sd 19%) of completed trials. If the animals were solving the task based on a learned associative sequence we would expect them to leave the rake in the waiting compartment when tested with a 5 minute delay, as this delay considerably interrupted the trained sequence of actions. It is possible that the sight of the baited test table prompted them to transport the tool from the waiting to the test compartment. We cannot exclude this option. The animals did not solve the task by means of foresight as they were trained to transport the rake, but they did apply the trained behaviour in a novel context of a delay, which is difficult to achieve by associative learning alone when immediate rewards do not follow the target behaviour. Three of the subjects were now able to transport tools over a delay.

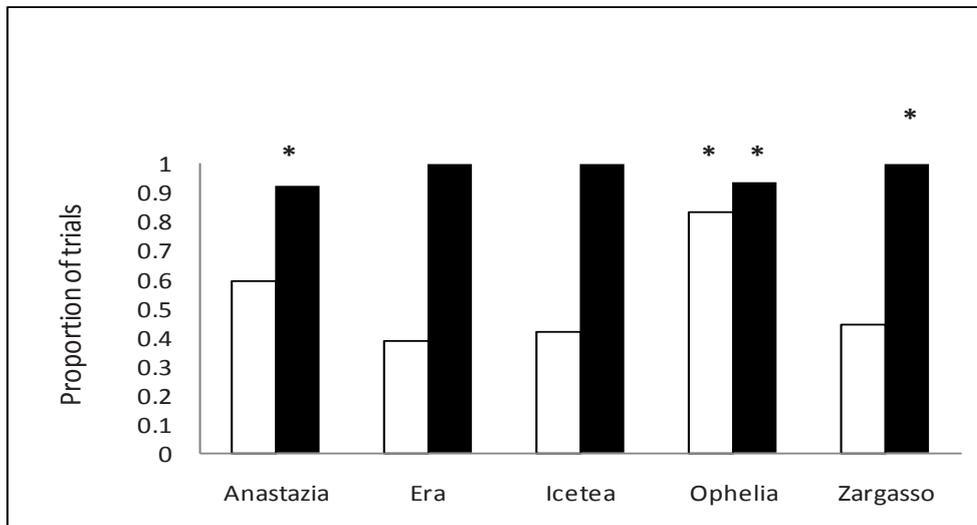


Figure 2.4: Results of test 2.1.

The proportion of completed trials for each of the five monkeys in which the tool was transported to the waiting compartment (open bars), and the proportion of these trials in which the tool was then transported to the testing compartment (black bars). Values significantly different from random choice are indicated by *.

Test 2.2: Transport after 10 min delay

In the second test we increased the time spent in the waiting compartment to 10 minutes and kept the rest of the procedure the same as in test 2.1.

Test 2.2: Methods

Each subject received three trials in a row each day and was tested on four consecutive days. We scored the number of times the rake was transported to the waiting and test compartments. We tested the two best performing individuals from the previous test. The third best performer from the previous test was not tested due to her new offspring.

Test 2.2: Results

On average, the two animals completed 71% (sd 41%) of trials. Of these we counted the number of trials in which they brought at least one rake with them. Ophelia transported the rake to the waiting compartment significantly more often than expected by chance (Binomial test, chance level 50%, 12 out of 12 trials, exact $p=0.001$), also significant after Bonferroni correction. Zargasso did not transport the rake to the waiting compartment in more trials than expected by chance (Binomial test, 4 out of 5 trials, exact $p=0.375$). This non-significant result can be attributed to the fact that he only completed 5 trials. Ophelia also transported and used the rake in the test compartment significantly above chance levels (Binomial test, 12 out of 12 trials, exact $p=0.001$), while the values for Zargasso failed to reach significance (Binomial test, 4 out of 5 trials, exact $p=0.375$). Again, although Zargasso did transport the rake to

the waiting and to the test compartment, the number of completed trials was too low to reach significance. Both individuals always used the rake if they brought it to the test compartment.

Test 2.3: Transport after 20 min delay

Similarly to the previous test we now increased the time spent in the waiting compartment to 20 minutes and kept the reminder of the procedure as described in test 2.1.

Test 2.3: Methods

Each subject received two trials in a row per session, with two sessions per day on four consecutive days. The two trials were separated by an inter-trial interval of at least 10 seconds, and the two daily sessions by at least 2 hours. We counted the number of times the rake was transported to the waiting and test compartments. We tested two individuals.

Test 2.3: Results and discussion

On average, the two animals completed 34% (sd 13%) of the trials. From these completed trials, we counted the trials in which they brought at least one rake with them to the waiting compartment. Zargasso failed to transport the rake to the waiting or test compartment in any of the 7 completed trials. Ophelia successfully transported the rake to the waiting and testing compartment in 3 out of 4 completed trials. Due to the low number of trials this value did not exceed chance level (Binomial test, $n=4$, exact $p=0.625$). Ophelia always used the rake when she brought it to the test compartment. Although Ophelia did successfully transport the rake to the waiting compartment and then on to the test compartment after a 20 minute delay, the number of completed trials was too low to reach significance.

The results from Ophelia suggest that macaques may be able to continue transporting the rake even when the trained sequence of behaviours is interrupted by a delay of 20 min, suggesting flexibility in solving the task; however further testing is required to confirm this indication. Tool transport after a 20 min delay might be the limit of the macaques' ability. We addressed the flexibility of tool transport behaviour further in the next experiment.

EXPERIMENT 3: NOVEL TOOL SELECTION AND TRANSPORT

In the final experiment we examined whether the animals could discriminate between novel functional tools and novel non-functional objects and then transport them to the baited test table after a 5 minute delay.

Experiment 3: Methods

While the subject was in the waiting compartment, two novel objects were simultaneously presented on a shelf outside it. Both objects were shown to the animal from three different sides before being placed on a shelf outside the waiting compartment for 1 minute. One of the novel objects was a functional tool with which a treat could be raked in, while the other was a non-functional object. There were twelve different objects in each category (a full list of the objects is provided in the appendix) and the objects were presented to both individuals in the same sequence. The novel functional tools differed from each other in several dimensions such as shape, colour and texture. After one minute the shelf was removed and the experimenter left the area for 5 minutes. Following the delay the subject was given access to the test compartment where the test table was presented and treats were placed on the table every 10 seconds for 2 minutes. The test table was thus first visible to the animals following the delay. The treats could only be obtained if the animal brought the functional novel tool. The subjects received four trials with an inter-trial interval of at least 10 seconds, on three consecutive days. In total each animal received twelve trials with unique novel tools and non-functional objects. On the first testing day the animals received one trial with a familiar tool and distracter to accustom them to the general procedure before starting the trials with the novel objects. We considered the animal to have selected an object when it was pulled into the waiting compartment. We scored the number of times this occurred for both the functional and the non-functional tools. We also noted which objects were transported to the test compartment and used to obtain treats. We tested two individuals.

Experiment 3: Results and discussion

On average, the two animals completed 79% (sd 18%) of trials. Ophelia selected both objects in all 8 of the completed trials (Figure 2.5), but used the functional tool significantly more often than non-functional objects in the test compartment (Cochran's test, $n=8$ trials, exact $p=0.008$). In fact, she never attempted to obtain the treats with a non-functional object. Zargasso selected the non-functional tools significantly more often than the functional ones (Cochran's test, $N=11$ trials, exact $p=0.016$), however he only brought objects of either type to the test compartment on 6 trials (Figure 2.5). He tried to obtain treats with both types of objects: four times with a functional tool and two times with a non-functional object. No significant difference was found between the functional and non-functional objects he used in the test compartment (Cochran's test, $n=6$ trials, exact $p=0.500$). Both macaques always used the objects they brought to the test compartment. Ophelia was successful in transferring the behaviour to novel tools and used them to obtain treats, while Zargasso was not.

Ophelia was able to generalize the relevant features of the tool necessary to solve the task of obtaining the treats. At the time of selection the test table was not visible, and Ophelia always selected both objects, suggesting that the selection was

driven by object neophilia. However, the fact that she never attempted to obtain the treats with the non-functional object indicates that when faced with unreachable treats, she recognized which features of the novel tools were essential to solve the task. This was less clear to Zargasso, who on two trials attempted to obtain treats with non-functional objects. These results expand our knowledge of the flexibility macaques are able to show when transporting novel tools for future use.

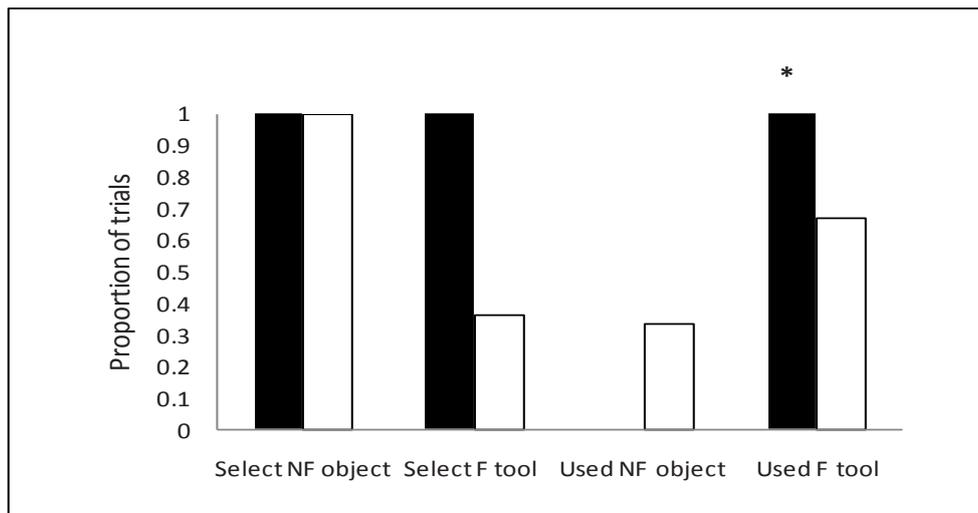


Figure 2.5: Results of Experiment 3.

The bars on the left show the proportion of completed trials in which objects were selected. The bars on the right show the relative frequency with which the two types of objects were used in the test compartment. Black bars represent data from Ophelia, while open bars represent data from Zargasso. F stands for functional tools and NF for non-functional objects. Significant difference between transporting F and NF objects is indicated by *.

GENERAL DISCUSSION

This study aimed to investigate tool transport for future use in long-tailed macaques. Our first experiment revealed that the tested animals were unable to transport the tool from the provided compartment to the baited test table to obtain treats. All six animals failed both when tested with a 5 minute delay between tool provision and access to the test table (tests 1.1 and 1.2), as well as when the two events occurred simultaneously (test 1.3). In fact results from test 1.4 revealed that all animals showed a preference for rake manipulation only when the rake was presented on the test table next to the rewards, indicating that the tested animals were not able to plan for their future need of the rake. In the second experiment, where the macaques were trained to transport the rake to immediately obtain rewards, we examined the flexibility of this learned behaviour. Three of the five macaques successfully transported the rake after a 5 minute delay and one of the two tested animals was even successful after a delay of 10 minutes and on a few trials with a 20 minute delay. In the third experiment the best

performer from the second experiment was also able to extend this behaviour to novel functional tools. Together these results demonstrate that the macaques were able to learn to solve the tool transport problem and apply the learned behaviour in a novel context.

The macaques' performance in our first experiment could have resulted from their inability to anticipate a future need or from their inability to solve the presented task. Solving the task involved connecting two temporally and spatially separate events: the provision of tools and the provision of rewards. In order to connect these two events, the animal had to remember its previous inability to obtain rewards in the test compartment when subsequently presented with tools in the selection compartment. It further had to realize that its previous failure was caused by the absence of the rake. As they were tested twice a day, the minimal requirement of the animal was that they had to remember what happened in the morning session during the afternoon session. Additionally, as they were repeatedly tested on several days, they could also recall this repeated experience. A possible explanation for the macaques' failure to transport the tools could therefore be a general inability to connect the necessary sequence of behaviours to achieve tool transport when forcibly delayed. Similar results have been obtained in a study with rhesus macaques, where their ability to anticipate a future need was investigated by means of a satiety test (Paxton & Hampton, 2009), where a choice for a preferred large amount of thirst inducing food was contingent with a long delay without water, while the choice for a smaller amount of food led to a short delay without water. The macaques were unable to reverse their initial preference for the larger amount of food, except when the choice for the smaller reward was immediately followed by the return of water. The authors argue that the animals' behaviour is caused by their inability to learn behaviour outcome associations across long delay intervals, as would be expected from traditional accounts of operant learning (Paxton & Hampton, 2009). The results from tests 1.1 and 1.2 in our study support this finding. However, results from test 1.3 showed that our macaques were also unable to transport the tool even when there was no delay, as tool provision and test table presentation directly followed each other. This finding differs from that obtained in the rhesus macaque study (Paxton & Hampton, 2009), where the animals did learn to choose for the smaller amount of food when the water was returned immediately after the choice. This demonstrates that the animals in our set-up failed not only due to the delay between the two key events, but also because they were unable to perform the required sequence of behaviours to achieve tool transport.

We present two potential explanations that may have influenced the macaques' performance in the first experiment. The first explanation may be that the animals were able to anticipate their future motivational state of wanting to obtain the treats, but that this motivation was not strong enough to warrant the required investment towards solving the task. In order to ensure that the animals' actions were not based on a current motivational state but a future one, it was important that they were not hungry while selecting the tool. Despite this, monkeys would readily consume

treats when given the opportunity. Perhaps tests using more exceptional or desirable rewards could have motivated the animals to reveal a better performance, however, we observed throughout the first experiment that the monkeys attempted to obtain treats in the test compartment and continually observed the experimenters placing rewards on the test table. Although we cannot measure the extent of the animals' motivation to obtain the treats, we do not consider a lack of motivation the key reason behind the macaques' failure to solve the presented task.

The second explanation behind the animals' behaviour may be that they were "over trained" to use the rake and were unable to display tool transport. The animals were focused on the learned motion and therefore may have failed to consider any other manipulation of the rake outside the trained context. During the rake training all six individuals learned to sweep rewards towards themselves when the tool was placed behind a reward in their first or second session. However, they all required several sessions to consistently lift and place the rake behind the next reward in two consecutive sessions. In fact, the animals required on average 21 sessions to achieve repositioning on two consecutive sessions. A similar rake use acquisition pattern was previously observed with Japanese macaques (Ishibashi et al., 2000). The repositioning of the tool requires the tool to first be moved away from oneself in order to then bring the reward within reach. This differs from the more intuitive and frequently encountered pulling-towards-oneself movement to obtain a far away object, which could account for the different speeds of acquisition (Ishibashi et al., 2000). Because of the large number of sessions required to achieve consistent repositioning, a possible consequence could have been overtraining of this particular action, which in turn could have inhibited manipulation of the tools in any other way in the first experiment. We cannot exclude the potential effect of tool use training on the animals' failure to transport the rakes, however the training was a necessary step in our study.

Regardless of the reason behind the animals' performance in the first experiment, it was striking that in all 158 completed trials, no individual ever made any attempt to solve the task. The animals have been accustomed to having food provided by their caretakers through their lifetime, therefore they may have required experience to learn that they now were able to do something themselves to secure the treats in the future. This passiveness contrasts with behaviours observed during the second and third experiment. Although the animals were now successful at transporting the rake after delays, they still occasionally did not bring the rake. In these trials we observed three individuals (Era, Ophelia and Zargasso) attempting to obtain the treats with objects present in the enclosure (a piece of straw, a small stick, banana peel) or even their own tail. Although these attempts were always unsuccessful, they are interesting as they show that the animals were proactively attending to the presented problem by trying to solve the task as well as indicating an ability to search for alternative solutions that should be further investigated in future studies.

Our study was based on previous studies with great apes, warranting a comparison of the results. In several studies bonobos, chimpanzees and orangutans

successfully selected and transported a familiar tool in order to use it an hour later (Mulcahy & Call, 2006; Dufour & Sterck, 2008; Osvath & Osvath, 2008). The results of the great apes studies are interpreted as demonstrating the abilities of great apes to anticipate a future need not currently being experienced and foresee the appropriate solution to the problem. Our macaques failed to transport the tool even under simplified circumstances with either a 5 min delay or even with no delay in test 1.3. In fact the macaques were only able to transport the tool after this sequence of behaviours was shaped in step-wise progressive training sessions with intermediate rewards.

It remains difficult to establish whether our results reflect a fundamental difference between the abilities of great apes and monkeys, especially given the small sample sizes in all the above mentioned studies and in ours. Factors such as prior experimental experience and rearing history can contribute to vast differences between individuals' abilities to manipulate tools and need to be properly accounted for before conclusions can be made about the evolution of the involved cognitive capacities. Nevertheless, our results show that long-tailed macaques fail to plan even when tested with a simplified version of a task that great apes can pass. These results indicate limits in planning for a future need and warrant further investigation in several macaque species with larger sample size and controlled experimental histories of tool use and rearing conditions.

Our results from the second experiment underscore that once the macaques learned how to solve the presented transport problem, they were also able to solve it in a novel context. Their 53% success rate in test 2.1 was comparable to that of the great apes (approximately 40%) (Mulcahy & Call, 2006; Dufour & Sterck, 2008), although of course the task for the macaques was simplified and they had received training. These findings show that the macaques were able to expand the tool use to a different context from the one they in which they were originally trained, and to do so with a larger delay than would be traditionally used in associative learning paradigms (Roberts, 2002; Shettleworth, 1998). The flexibility of the macaques is also reflected in their ability to solve the task with various novel tools in the third experiment. The macaques average success rate of 63% in experiment 3 is somewhat lower than that of great apes (84% in a similar set-up with a longer delay) (Osvath & Osvath, 2008). The authors argued that the apes solved the novelty task by learning the relevant affordances of the tools rather than by responding to the simple perceptual relations (Osvath & Osvath, 2008). In our opinion more careful experiments need to be carried out before such conclusions can be drawn, for example where the animals are asked to choose between novel objects that differ in relevant and irrelevant dimensions (shape or colour). Interestingly, the best performers in experiments two and three were the three youngest individuals in our sample. Sub-adult individuals generally show higher levels of curiosity about their environment and are more prone to interact with unfamiliar objects as a part of their development (Reader & Laland, 2001; Visalberghi et al., 2003). This higher level of curiosity could have contributed to the increased manipulation and flexibility of the three to four year old macaques in our sample.

Taken together, the results show that the macaques were not able to solve the tool transport problem by using foresight, however whether this was due to a lack of anticipation of a future need remains unanswered. The most parsimonious interpretation of our results is that the animals were unable to foresee the correct sequence of behaviours required to solve the task in the absence of training. The transport training seemed to have opened new avenues for the animals to display the shaped behaviour in novel contexts and with novel tools, as well as to proactively try to solve the task with other objects or body parts in the absence of the rake. Future studies with tasks involving novel tools, sequential tool tasks, tasks involving delays between tool provision and tool selection, as well as a control condition where tool transport is no longer required to obtain rewards, will reveal more about the limits and flexibility of tool transport capacity in long-tailed macaques. The abilities of macaques in the domain of future planning might not match those of great apes, but they provide valuable insight into the evolution of complex cognitive abilities that should be further investigated to elucidate which components are necessary to achieve similar behaviours instead of merely focusing on all or nothing responses.

ACKNOWLEDGEMENTS

We are grateful to Henk Westland for his practical assistance with the set-up and his advice with the monkeys. We are thankful to Valerie Dufour and Will Swaney for helpful comments on the manuscript and Han de Vries for advice on the statistical analysis.

APPENDIX**List of novel tools and non-functional objects.**

Trial	Novel functional tools	Non-functional objects
1	Plain round wooden brush, hard (30 cm in length, 0.5 cm in diameter)	Grey screw-on bottle top (3 cm in diameter 1.5 cm in height)
2	Black triangular plastic poster frame, hard (32 cm in length, 1 cm in height)	Square yellow stick (10 cm in length, 1 cm in width)
3	Blue plastic bottle cleaner with white brush, hard (30 cm in length, 3 cm in diameter)	Metal bolt screw (3.5 cm in length, 1 cm in diameter)
4	Beige metal triangular heavy stick, hard (35 cm in length, 2 cm in height)	Red hollow plastic tube, hard (4 cm in length , 2.5 cm in diameter)
5	Black and green hose, bendable (40 cm in length, 1.5 cm in diameter)	White plastic tube, soft (5 cm in length, 2 cm in diameter)
6	Aluminium blind slat, bendable (35 cm in length, 3.5 cm in width)	Green flat plastic net, hard (11 cm in length, 5 cm in width)
7	Yellow plastic square stick, bendable (33 cm in length, 2 cm in width)	Dutch flag toothpick (6 cm in length, 3.5 cm in width)
8	Thin silver metal stick, hard (35 cm in length, 3mm in diameter)	Half of crock coaster (9 cm in length, 3 cm in width)
9	Metal spatula with blue rubber top, hard (25 cm in length, 4 cm in width)	Transparent plastic cocktail muddler (6 cm in length)
10	Transparent plastic ruler, hard (2 cm in length, 3 cm in width)	Green wooden stick (7 cm in length, 0.5 cm in height)
11	White plastic straw with blue/green lines, soft (25 cm in length, 0.5 in diameter)	Orange plastic net, soft (8 cm in length, 6 cm in width)
12	Light green square plastic stick, bendable (30 cm in length, 3 cm in width)	Natural knotted rope (20 cm in length)

TREATS

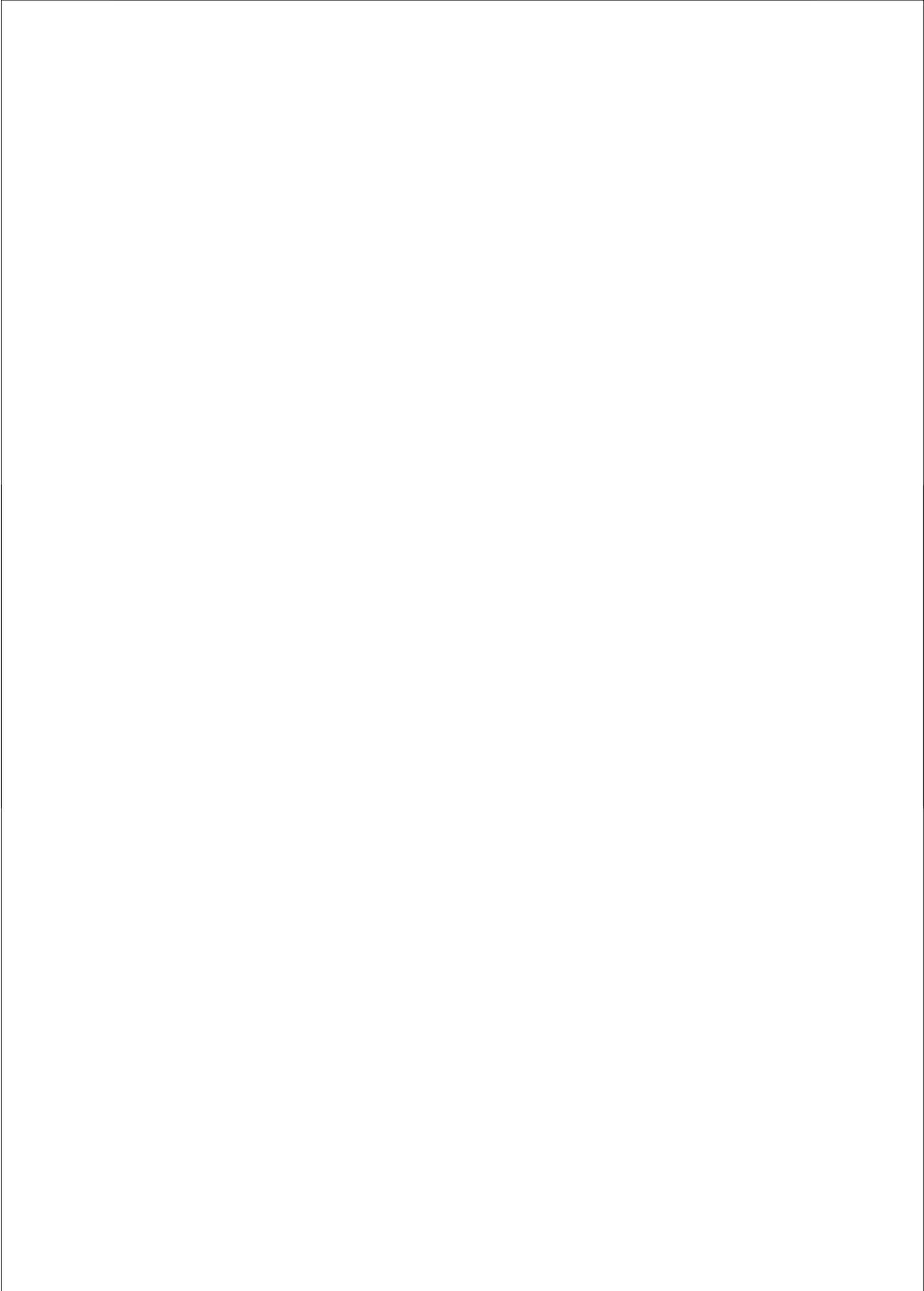
Failure at a What-Where-When Task Results in Obtaining Rewards by Means of a Location-Based Association Strategy in Chimpanzees (*Pan troglodytes*)

ABSTRACT

Recollecting the what-where-when of an episode, or episodic-like memory, has been established in corvids and rodents. In humans, a linkage between remembering the past and imagining the future has been recognised. While chimpanzees can plan for the future, their episodic-like memory has hardly been investigated. We tested chimpanzees (*Pan troglodytes*) with an adapted food-caching paradigm. They observed the baiting of two locations amongst four and chose one after a given delay (15min, 1h or 5h). We used two combinations of food types, a preferred and a less preferred food that disappeared at different rates. The subjects had to base their choices on the time elapsed since baiting, and on their memory of which food was where. They could recover either their preferred food or the one that remained present. All animals failed to obtain the preferred or present foods above chance levels. They were like-wise unsuccessful at choosing baited cups above chance levels. The subjects, thus, failed to use any feature of the baiting events to guide their choices. Nonetheless, their choices were not random, but the result of a developed location-based association strategy. Choices in the second half of the study correlated with the rewards obtained at each location in the first half of the study, independent from the choices made for each location in the first half of the study. This simple location-based strategy yielded a fair amount of food. The animals' failure to remember the what-where-when in the presented set-up may be due to the complexity of the task, rather than an inability to form episodic-like memories, as they even failed to remember what was where after 15 minutes.

Published as:

Dekleva M., Dufour V., de Vries H., Spruijt B.M. and Sterck E.H.M (2011). Chimpanzees (Pan troglodytes) Fail a What-Where-When Task but Find Rewards by Using a Location-Based Association Strategy. PlosOne, 6, e16593.



INTRODUCTION

The conscious re-experience of past events and the anticipation of future ones is ascribed to the ability of mental time travel into the past and future (Suddendorf & Corballis, 2007a). This advanced ability enables our own species the use of detailed knowledge from past personal experiences to meet current demands. Over the last ten years, innovative research has shown that, like humans, some animal species can remember certain features of experienced episodes as well as prepare for anticipated future events, albeit in a more limited manner (Crystal, 2009). Behavioural paradigms investigating episodic-like memory, a simplified version of the human episodic memory system, have led researchers to question the belief that animals are unaware of their past and future and might be “stuck in time” (Roberts, 2002).

The episodic memory system forms together with the semantic system the declarative subdivision of the human long-term memory system (Vargha-Khadem et al., 2001; Squire, 2004). Both systems are characterised by the conscious retrieval of stored information. Whereas the semantic system stores facts and concepts acquired over several exposures, the episodic system consists of events formed after single exposure. Tulving (Tulving, 1972) originally defined episodic memory as storing detailed information about the temporal and spatial features of a unique episode (the what-where-when), but later added, that such memories were episodic when accessed by conscious re-experiencing of the encoding event, or so called mental time travelling (Tulving & Markowitsch, 1998). Investigating episodic memory in animals is constrained by their inability to verbally communicate their re-experience of an episode. Researchers have shown that animals remember important characteristics of experienced episodes, and, thus possess a system similar to the human episodic memory called episodic-like memory (Clayton et al., 2003a; Crystal, 2009; Hampton & Schwartz, 2004; Eacott & Easton, 2010) or what-where-when memory (Suddendorf & Busby, 2003). While these results fit Tulving’s original definition of episodic memory (Tulving, 1972), such behavioural paradigms cannot determine whether the animals experience conscious recollection (Clayton et al., 2003a; Clayton et al., 2001a; Clayton & Dickinson, 1998).

Using a food-caching task, Clayton and Dickinson (Clayton & Dickinson, 1998) tested western scrub-jay’s (*Aphelocoma californica*) ability to form episodic-like memories. The animals cached two foods, preferred but perishable wax worms, and less favoured but not perishable peanuts. When recovering their caches, they tended to search for wax worms if only a short time had passed, but switched to the less-preferred peanuts if a long time had elapsed since caching. Thus they were successful at distinguishing the type of food they cached, its location and how long ago they made their caches (Clayton & Dickinson, 1998). Various studies have shown that other corvids: magpies (*Pica pica*) (Zinkivskay et al., 2009), black-capped chickadees (*Poecile atricapillus*) (Feeney et al., 2009), as well as mammals: rats (*Rattus sp.*) (Kart-Teke et al., 2006; Babb & Crystal, 2006a; Babb & Crystal, 2006b; Ergorul & Eichenbaum, 2004; Bird et al., 2003), mice (*Mus musculus*) (Dere et al., 2005b), and meadow voles

(*Microtus pennsylvanicus*) (Ferkin et al., 2007) are able to recall the what-where-when of similar events. However, the jays' behaviour in the Clayton and Dickinson study (Clayton & Dickinson, 1998) could be alternatively explained by a simple learned rule: search for the preferred wax worms if little time has passed, but avoid the worms if long time has passed since caching (Clayton et al., 2001b). To exclude this option, a third food and time interval were added to the set-up. Jays now had to adjust their searches based on the what-where-when in six different conditions. They showed remarkable flexibility and appropriately adjusted their searches (Clayton et al., 2001b). Clayton and colleagues have since reinforced and extended these results through further experimental investigations. They showed that the jays also can integrate the content of what-where-when episodes into a single memory; that they can flexibly update information about the decay rate of the foods (Clayton et al., 2003a; Clayton et al., 2001b); that they are sensitive to who observes them while caching (Dally et al., 2006); and that they can even plan for future needs (Raby et al., 2007b).

While considerable effort has been invested in the study of corvids' and rodents' episodic memory, few attempts have been made to test episodic-like memory in primates. This is surprising, as apes are known to possess advanced cognitive abilities (Call & Tomasello, 2008) and show long term memory (Beran, 2004; Menzel, 1973). Additionally, experimental and observational reports have shown that chimpanzees and other great apes are able to plan for the future (Mulcahy & Call, 2006; Dufour & Sterck, 2008; Osvath & Osvath, 2008; Osvath, 2009). Several studies with human subjects have established that planning and episodic memory share neural resources (Suddendorf & Corballis, 2007a). The few existing studies of primates episodic memory are burdened with potential alternative explanations of their findings and all have utilised different approaches (Dere et al., 2006). This makes a comprehensive comparison of their findings difficult. The mnemonic ability of a single lexigram-proficient chimpanzee was examined by means of a free-recall paradigm (Menzel, 1999). Sixteen hours after an item was hidden in a large outside area, the chimpanzee obtained the attention of a naïve caretaker and, with the aid of the lexigram, led him to the items. The time interval used is impressive in length, and, informative about this animal's long term mnemonic ability. However, as the structure of the memory content is not tested (Clayton et al., 2003a), the results say little about the animals ability to form and recall episodic-like memories. An alternative approach to episodic memory in non-humans was examined by Schwartz and colleagues (Schwartz et al., 2005). They tested a single gorilla (*Gorilla gorilla gorilla*) that indicated which food it had recently received from which caretaker by handing over the correct picture card (Schwartz et al., 2002). The same gorilla was also tested with novel actions, persons and objects as well as with the temporal order of events. In all contexts, the gorilla returned the correct card above chance levels (Schwartz et al., 2004; Schwartz et al., 2005). These findings demonstrate knowledge of unique past events. Whether the animal truly recalled the details of the event or simply returned the card representing the most recent, and thus familiar, event remains open to discussion (Schwartz et al., 2005). In monkeys, an adapted version of the Clayton and

Dickinson (Clayton & Dickinson, 1998) set-up examined the what-where-when components of memory with rhesus macaques (*Macaca mulatta*) (Hampton et al., 2005). The macaques were able to remember the locations of two foods for up to 25 hours, but failed to recognize that only the less preferred food was palatable after a long delay. More recently, three great ape species (*Pan paniscus*, *Pan troglodytes* and *Pongo pygmaeus*) were likewise tested on the what-where-when features of food hiding events by adapted paradigms of the Clayton lab (Martin-Ordas et al., 2010). In their first experiment the authors showed that the apes were able to remember where and when two types of foods were hidden by selectively choosing the perishable food item after a short interval, but switching to the non-perishable food following a long delay. However, the findings of this experiment could be explained by the same rule-based learning mentioned above (Clayton et al., 2001b). To examine whether the three components (i.e. what, where and when) were structured into a single memory the authors further showed that the animals encoded two baiting events in an integrated fashion (Martin-Ordas et al., 2010). However, these findings can also be explained by the animals following the same rule, admittedly with impressive flexibility. To exclude rule learning, the experimental set-up could be complemented with an additional food and time interval, as was used successfully with corvids (Clayton et al., 2001b).

In the present study we adapted the extended three food and time interval paradigm of Clayton (Clayton et al., 2001b) to test memory of what-where-when in chimpanzees (*Pan troglodytes*). This paradigm is the only one that allows a comparison of behaviour between a control and a test group, while evaluating an animal's behaviour across multiple conditions and facilitating the exclusion of simple rule-based learning or familiarity processes. We determined the chimpanzees' preferences for three foods, trained them to retrieve a preferred food from four potential locations, and showed them that the three foods had different rates of disappearance over time. We tested whether the animals integrated this knowledge, which would allow them to retrieve their preferred food, or the food remaining present in six different food and time combinations. We predicted the animals would choose the preferred food on trials where both foods remained present, but switched to the less preferred food on trials where the preferred food had disappeared. If the animals adjusted their behaviour depending on the amount of time that passed since hiding, while remembering where each food type was hidden it would show they were able to remember the what-where-when of a given food hiding event.

MATERIALS AND METHODS

Ethics

All training and testing was conducted as a part of the chimpanzee enrichment program of the Biomedical Primate Research Centre (Rijswijk, the Netherlands) and involved only positive reinforcement. The research activities were fully integrated into

the daily routine and required no additional manipulation of individuals. The chimpanzees were not deprived of food and water at any stage. In addition, the individuals could choose not to participate in any individual trial, yet all individuals were rewarded with a small treat at the end of each trial, regardless of their level of participation in the task. The individuals showed a constant willingness to interact with the researcher, indicating that the tasks improved the well-being of the chimpanzees and that our efforts to minimize any potential suffering were successful. Therefore, the research offered positive stimulation for the individuals and follows the Weatherall Report recommendations for good welfare. The study was conducted in compliance with all relevant Dutch laws and in agreement with international and scientific standards and guidelines. Due to the non-invasive character of the study and absence of potential discomfort no additional permission from the institutes animal experiment committee was required. By definition following the Dutch Animal Experimentation Act, an animal experiment is undertaken with a scientific purpose and affects animal welfare. This study was not considered an animal experiment because animal welfare was not affected (enhanced if anything). This was so assessed by the Biomedical Primate Research Centre Animal welfare officer.

Subjects

Nine individuals from three chimpanzee groups (named P, F and D) housed at the Biomedical Primate Research Centre (Rijswijk, the Netherlands) participated in this study (Table 3.1). The groups consisted of test individuals as well as several additional animals that were not tested due to inconsistent participation during the training stages. The participating animals had been living together for a minimum of 2 years and all three groups were socially stable. Because all training and testing involved only positive reinforcement, the individuals could choose not to participate in any individual trial. This is reflected in the different number of trials completed between the subjects.

The three social groups were housed in similar enclosures with both inside and outside compartments. The outside compartment consisted of one large area (size approx. 87 m³). The inside compartment consisted of two large “play rooms” (for groups P & F approx. 75 m³, for group D app. 57 m³) as well as several individual cages (size approx. 1.6 m³) arranged in a row on one side of the play rooms. The groups had no visual access to each other’s inside compartments. All chimpanzees had been trained to enter the individual cages as a part of their daily routine for feeding, cleaning and veterinary procedures. The outside compartments were connected onto a yard, arranged next to each other in a row allowing limited visual access between the social groups. Training and testing was performed in all compartments of the enclosures, depending on the stage of the study.

The animals were fed three times a day on a diet of chow, bread, fruit & vegetables. In the morning and afternoon the feeding took place in the individual cages, while at midday they were fed in the common parts of the enclosures. Water was available *ad libitum* throughout the study.

Table 3.1: Study subjects age, group affiliation and rearing history.

P, F and D stand for names of the three different social groups the subjects belonged to. T denotes test group and C control group.

Name	Age	Groups		Rearing history	
		Social	Test/Control	Hand reared	Mother reared (until age)
Claus	14	P	T	X	
Emanuel	17	P	T		X (2y)
Freek	14	P	T		X (1.5y)
Linda	22	F	T	X	
Marlis	26	F	T	X	
Paul	14	P	T		X (2y)
Rene	14	P	T	X	
Denis	23	D	C		X (7mo)
Regina	40	D	C		X (unknown, wild born)

Study design

The study began with a pre-training stage. We established that the animals preferred apple sauce and yoghurt over red bell peppers respectively. In the first of the two training stages we trained the animals to point to one of four distinctly marked locations in order to receive its content. For the second training stage, the subjects were divided into a test and a control group. Test animals were shown that the preferred apple sauce and yoghurt disappeared at different rates, while the less preferred red bell peppers always remained present. Control animals experienced that all three test foods always remained present. On each test trial the researcher showed the animal in which out of four differently coloured locations, two foods were hidden. After a delay of 15 min, 1h or 5h the animal was asked to point out which cup they wanted to receive. On each trial two foods were hidden in fixed combinations of either apple sauce and red bell peppers, or yoghurt and red bell peppers. The control group always experienced that both of the hidden foods were present at recovery, while the test group experienced that apple sauce disappeared after 1 hour and the yoghurt after 5 hours. Following the rationale of the Clayton (Clayton et al., 2001b) study we predicted that the animals in the control group would always choose the preferred food out of the two, while the test group would choose the preferred food on trials where both foods were present, but switch to less preferred food on trials where the preferred food had disappeared. Table 3.2 gives an overview of the predictions for the control and test group at 15 minutes, 1 hour and 5 hours. The choices of the control animals would demonstrate the animals ability to form long term memories for what is where, while the switch in choices of the test group would demonstrate their ability to integrate the what, where and when of the baiting episodes.

Table 3. 2: Overview of the predicted choices for the test and control group.

Time interval	Apple sauce & Red bell peppers		Yoghurt & Red bell peppers	
	Control group	Test group	Control group	Test group
15 min	Apple sauce	Apple sauce	Yoghurt	Yoghurt
1 h	Apple sauce	Red bell peppers	Yoghurt	Yoghurt
5 h	Apple sauce	Red bell peppers	Yoghurt	Red bell peppers

Pre-training procedure: Food preferences

The animals were tested on their preferences between three foods: apple sauce diluted with water 2:1, low fat natural yoghurt diluted with water 1:1 with three tablespoons of fruit syrup per litre and red bell peppers in 2x2 cm pieces. The foods were presented in two fixed combinations: apple sauce and red bell peppers or yoghurt and red bell peppers, which were pseudo-randomly varied between the daily sessions. This pre-training stage was conducted in the mornings following the animals' breakfast while they were already separated in individual cages. The foods were presented in familiar red paper cups (0.16l) containing one kind of food: one piece of red bell pepper, one spoonful of the yoghurt, or the apple sauce solution.

At the start of each session the animals received a taste of the two foods to establish their motivation to feed and to ensure they knew between which two foods they were choosing. Following consumption, the researcher presented two cups containing the foods. The two cups were tilted towards the animal and approximately 30-40 cm apart. The animals indicated their choice by reaching out for one cup. They received the chosen cup and could consume the food. In case an individual did not indicate its choice the researcher removed the cups and presented them again a few seconds later. If the animal remained unresponsive the session was terminated. Only data from sessions where the animals completed six trials were used for analysis.

On each day a combination of the same two foods was presented in six trials, with the side of presentation counterbalanced. Each food combination was tested on five to eight daily sessions, with all nine animals completing a minimum of 30 trials (mean= 36.7, SD= 3.5) per combination. The individuals did not complete the same number of trials. In order to demonstrate a significant preference of one food over the other, we set as a criterion that the animals should chose the preferred food in at least 70% of the trials. A binomial test with 21 positive choices out of 30 (70%) would show a significance of $p = 0.046$. Based on these choices we determined their preferences between apple sauce, yoghurt, and red bell peppers.

Training procedure part one: Pointing

This part of the training was conducted while the animals were separated in their individual cages following breakfast. Two locations were baited, out of initially two and later four possible locations. The four locations were marked by cup holders of different colours and patterns that were located in the corners of an upright held square grid. At this stage we used foods other than those in the final test, i.e. bread or

cookies as the preferred food and carrots as the less preferred food. These foods were selected based on the recommendations of the animals' caretakers.

Each session commenced after the animal consumed a small piece of both foods. The researcher first showed the animals that the cup holders were empty by tipping them upside down. Then, two holders were baited, each with a cup containing one food. The remaining two holders were treated with the same hand motions as the baited ones; however empty cups were inserted into them. During the baiting the researcher used verbal cues to encourage the animals to pay attention to the baiting. In case the animals looked away or were otherwise distracted, baiting at that location was repeated. The four holders were then covered with opaque lids. The metal grid was then held upright close enough for the animals to reach for, but not touch, one of the holders. The animals were, if necessary, verbally encouraged to make a choice and received the cup they indicated. If the animals did not point at any location the researcher stepped back and presented the holders again after a few seconds. If the animal remained unresponsive the trial was scored as no choice and was excluded from analysis. When individually separated, group members could see which cup holder their neighbours chose. This information could influence their choices. To prevent such visual cues the metal grid was either presented inside a large box (group P) or the researcher positioned herself at an angle so her back would function as a visual barrier (groups F and D).

Each training day consisted of one session with four trials per individual. All four locations were overall, baited approximately equally often in a pseudo-random order (mean number of baiting events per cup= 25, SD= 2.31). The animals were trained on ten to thirteen days with four cups, and all completed a minimum of 40 trials (mean= 50.00, SD=4.28). For the animals to pass this training stage we set as a criterion that they should make at least 60% choices for the cup containing the preferred food. At 60% of correct choices, a Chi-square test with 40 trials (minimum completed) and expected choice of 25% would show a significance of $p < 0.001$. The animals were then considered proficient at indicating the one location out of four that contained their preferred food, as well as at understanding the connection between the baiting and the choosing.

Training procedure part two: Temporal properties of food

This part of the training was conducted in front of the entire social groups in the late morning. Depending on the cleaning routine, that restricted which parts of their enclosure the animals could access, the foods were either presented in front of the outside, or the inside enclosures. This ensured that as many individuals as possible witnessed the presentation of the foods.

The researcher first encouraged the individuals to come into the appropriate room by calling their names. All animals received a small amount of the test food to ensure they knew which food it was. A large amount of the test food placed in cups was left in front of the enclosures in plain view. The researcher left and returned after the predetermined time intervals of 15 min, 1hour or 4 hours and gave the animals the

cups. These either still contained the foods or were empty depending on the time interval and whether the animals belonged to the test or the control group.

For the members of the control group all three foods were always present upon recovery. For the members of the test group, however, the cups' content was manipulated so that the yoghurt and apple sauce were either present or absent at recovery. This manipulation was achieved by the following: prior to presentation the cups were either filled with the yoghurt or apple sauce solution, or left empty, but always covered with cling-film secured on top of each cup. The top of the cling film was covered with a layer of either apple sauce or yoghurt. This ensured that upon visual inspection the cups appeared full. These cups were presented to the animals and left in front of their enclosures. Once the designated time interval passed the researcher removed the cling film with her back to the animals' enclosure, so the subjects were unable to see her actions and then distributed the cups to the present animals. The animals would thereby experience that the foods disappeared immediately prior to them receiving the cups. The researcher would, for example, prepare full cups of apple sauce for the 15 minute interval and empty cups for the 1 and 4 hour intervals. The test group members were, thus, able to experience that, depending on the time interval, a given food could either be present or disappear, while the control animals experienced the same procedure but never experienced that the foods disappeared.

On each day one food was presented during one time interval and given to the animals after approximately 15 min (range 0:09- 0:14), 1 hour (range 0:51-1:14) or 4 hours (range 3:46-4:09). The foods were first presented in a descending order of time intervals (4 hours, 1 hour and 15 minutes) and later ascending (15 minutes, 1 hour, 4 hours) order. First the red bell peppers were presented, second the yoghurt and thirdly the apple sauce. After this all three foods were presented again, each at 4 hours, 1 hour and 15 minutes.

Testing procedure

The testing was performed while the animals were separated in their individual cages. Each session consisted of two parts; food hiding and food recovery. During both parts two persons were present; MD and a familiar animal caretaker. One person was hiding or recovering the food, the other was videotaping the trial. The two persons always switched roles within a session, to ensure that the person performing the recovery was unaware of the actual location of the foods and thus unable to cue the animals.

The hiding and recovery procedures were essentially the same as the procedure described above for pointing training. Metal grids with four new distinctly coloured holders were used. Once the foods were hidden in front of each animal, the grid was placed in front of that individual's cage. The foods were hidden for approximately 15 min (range 0:08-0:21), 1h (range 0:56-1:24) or 5 h (range 4:58-5:38). Due to logistical reasons the long time interval was somewhat longer than the one under the training of temporal properties. We reasoned that if the animals successfully learned the foods temporal properties during training, then extending this interval should not influence the direction of the animals' responses.

The animals remained in the individual cages for the 15 minute and 1 hour intervals, but were released and re-entered the individual cages on the 5 hour intervals. We could not ensure that they entered the same individual cage at recovery as they occupied during the hiding. Therefore, we gave each animal's grid an individual colour code, which was a large coloured paper placed in the middle of their metal grid. Before each hiding, the subjects were given a small paper of the same colour with honey or peanut butter to attract their attention to the colour. During the 5 hour intervals the animal's grids remained in front of the cages in which the animals were during the food hiding. Once the animals re-entered the cages for the food recovery the grids were moved so that each animal was situated in front of its own grid. During the 15 minute and 1 hour intervals the grids remained in front of the same individual cages, as the animals were not released.

For the members of the test group the yoghurt and apple sauce disappeared from the baited locations after the 1 and 5 hour intervals. This was achieved in a different way than during training (see also Figure 3.1). Four coloured holders were permanently attached to the same locations on the grid. Two stacked identical plain cups were inserted into each of these holders during food hiding. Two holders were baited with two stacked cups each, where one of the cups contained the test food, while the other was empty. The remaining two holders were baited with two empty stacked cups each. The animals observed the hiding of the foods in the stacked cups, into the holders, and the application of opaque lids. However, as the cups were stacked one inside the other, the animals just saw the baiting of the four holders with two empty cups and two cups containing food. After the hiding, the researcher placed the grid in front of each individual's cage and turned her back to the animals. She then (always in the same sequence) removed one of the stacked cups from each of the holders. Depending on the time interval she either removed the cups containing the food, or the empty cups. For example in a 15 minute interval four empty cups were removed, and thus both foods were still present at recovery. However, at a 1 hour interval (Figure 3.1) the cup containing the apple sauce was removed and upon recovery only the red bell peppers remained.

After the retention interval the animals were presented with their metal grids and verbally encouraged to reach for one location. They then received the inserted plain cup and its content. If the animals would not indicate a choice the grid was removed and presented again after a few seconds. If they remained unresponsive the trial was scored as no choice trial and excluded from the analysis.

At the beginning to the testing period we first familiarized the animals with the testing procedure. Each animal received one hiding and recovery session at each of the three time intervals. During this familiarization we used the same foods as during pointing training.

The 15 minute and 1 hour intervals took place, either in the morning or in the afternoon, following the morning or evening meal. The hiding in the 5 hour intervals was performed in the morning and the foods were recovered in the afternoon, both following the animal's feeding time. The animals received a maximum of three testing

sessions each day. Each animal received between seven and twelve trials of both food combinations in each of the three time intervals (mean=9.52, SD=1.30). Four test trials (from four different individuals) were excluded from the analysis due to researcher error during testing, where incorrect foods were present during recovery. For each of the possible six food and time combinations the position of the foods was counterbalanced. All four locations were baited approximately equally often with both foods for each animal (mean baiting events per cup= 28.5 SD=2.53). The sequence in which the food and time interval were presented was pseudo-randomized.

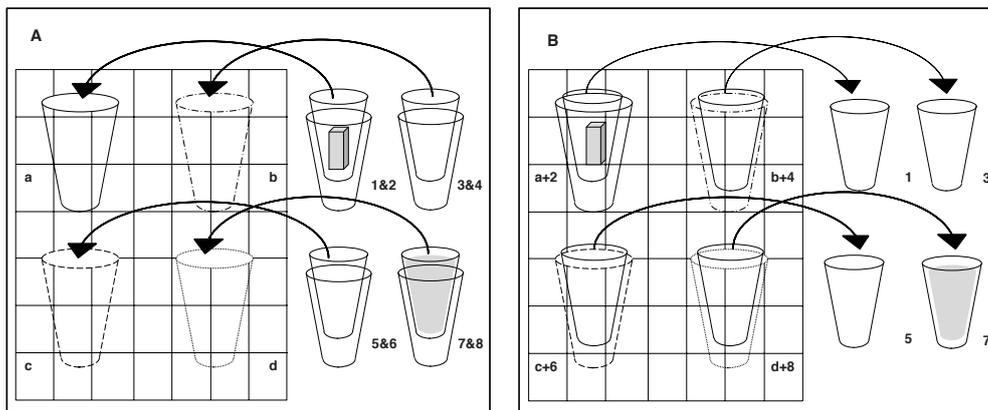


Figure 3.1: Manipulating the temporal properties of the foods during testing.

Panel A shows on the left the four different cup holders (a, b, c & d) attached to the metal grid. On the right are the eight identical plain cups that were inserted into holders a-d in front of each subject. Two stacked cups were inserted into each cup holder. Only two of them contained food. In this example, a 1 hour apple sauce and red bell peppers trial, cup two contained red bell pepper and cup eight contained apple sauce. The remaining 6 cups were empty. Panel B shows that four cups were removed from the holders with the researcher's back to the subject, immediately after the food hiding. One of the plain cups was removed from each holder. In this example, upon recovery, holder a still contained the red bell pepper, while the apple sauce from holder d had disappeared.

Analysis

During training and testing we manually scored the choices the animals made and whether they consumed the obtained food. During the training of the foods temporal properties we scored the group members present, both, when the foods were presented and when they were given. All test trials were videotaped as well as recorded manually after each trial. Data were entered into Excel from the observation sheets and, in case of inconsistencies, confirmed from videotapes. The data were analyzed with SPSS 16 and MatMan 1.1 (Noldus Information Technology, Wageningen) (Vries et al., 1993). Each animal's behaviour was tested individually, reasoning that just one animal's success would be of importance. We used Chi square tests and row-wise matrix correlation tests (Vries, 1993). All statistical tests were two-sided at a critical alpha of 0.05. We used the standard Bonferroni correction for multiple comparisons.

RESULTS

Pre-training: Food Preferences

We scored the animal's choices in the two different food combinations: apple sauce and red bell peppers, and yoghurt and red bell peppers. All individuals chose apple sauce over red bell peppers in more than 70% of the trials (mean= 96.60%, SD=4.77). All individuals except one (Marlis) chose yoghurt over red bell peppers in more than 70% of the trials (mean =87.73%, SD=18.89; excluding Marlis: mean = 93.14%, SD =10.33). Marlis was hereafter excluded from the analysis. None of the animals showed a significant preference for a particular side in either food combination (Binomial test, all individuals $p>0.31$). The animals were thus not guided by a side bias in their choices. All remaining individuals (i.e. except Marlis) showed a significant preference for apple sauce over red bell peppers and for yoghurt over red bell peppers.

Training part one: Pointing

All individuals chose the locations containing the preferred food in more than 60% of the trials (mean=81.55, SD=9.68). The less preferred food was chosen on average in 5.34 % of the trials (SD=4.27) while the two empty cups were chosen on average in 13.06% of the trials (SD=7.99). This demonstrates that all eight individuals were able to clearly discriminate and indicate the location of the preferred food.

Additionally, to assess the animal's motivation for choosing a particular cup, we scored whether they consumed the obtained foods. When obtained, the preferred food was consumed on 100% of the trials, while the less preferred food was consumed in 14.3% of the trials. This further indicated that the animals distinguished between the qualities of the different rewards and wanted to obtain the preferred food.

Training part two: Temporal properties of food

We determined the percentage of trials where the individuals were present, both when a particular food was placed outside their cage, and when it was handed out (mean = 92.28%, SD=13.39). Furthermore, we looked at each individual's presence on the informative trials, defined as those where any of the foods had disappeared. The animals received a total of 5 informative trials in each food and time combination. All individuals were present on minimum 80% of the informative trials. For apple sauce that was at 1 hour: mean = 97.50, SD= 7.07 and 4 hours: mean = 97.50, SD= 7.07. For yoghurt that was at the 4 hour interval: mean = 97.50, SD= 7.07. We assumed that the animals were given sufficient opportunity to learn the temporal properties of the presented foods.

Testing: what-where-when choices

We hypothesized that the animals' behaviour in the test could be guided by three different choice strategies, which we consider in turn. The first strategy we investigated was whether the animal's choices were guided by the principles of the paradigm; so they in each of the six time and food conditions successfully integrated

which food was hidden where as well as the time passed since hiding (Table 3.3). We examined whether the animals made more of such what-where-when choices than expected by chance (25%). We performed the analysis on three different levels: each animal's choices in all of the conditions, at each of the three time intervals irrespective of the food combination and their choices in each time and food condition separately.

When all the food and time conditions were pooled together, none of the animals made significantly more what-where-when choices than expected by chance (exact Chi-square test, $df=1$, all individuals $p>0.23$). Next, to investigate whether the length of the retention interval influenced the animal's success rate; we pooled the number of what-where-when choices made at each of the three time intervals irrespective of the food combinations. One of the animals (Linda) made in the 1 h condition significantly more what-where-when choices (exact Chi-square test, $\chi^2=5.07$, $df=1$, $p=0.03$). However, after a Bonferroni correction for multiple comparisons this value was no longer significant. All the other animals were unsuccessful at all three time intervals (exact Chi-square test, $df=1$, all $p>0.11$). Lastly, we looked at the what-where-when choices made in each time and food condition separately. Three different animals (Freek at 15 min, Linda at 1 h and Claus at 5 h) made significantly more what-where-when choices, all in the apple sauce and red bell pepper condition (statistics in Table 3.3). However, after a Bonferroni correction, none of the values remained significant. All other animals were unsuccessful (exact Chi-square test, $df=1$, all $p>0.13$) in all food and time combinations. Thus, none of the animals made significantly more choices for their present or preferred food in any of the food and time combinations and, thus, all chimpanzees failed to pass the criteria of the food-caching paradigm.

Testing: what-where choices

The second potential strategy we considered, involved determining whether the animals were successful at making what-where choices, considering either food type rewarding (Table 3.3). Success at these choices would indicate that the animals were in each trial remembering either of the two baited locations, but disregarding the temporal properties of the foods. As two cups were baited in each trial the animals chance success rate of making what-where choices was 50%, however, in trials where only one food remained present these choices would not result in a reward.

We investigated whether the animals were choosing the two baited cups more often than expected by chance (50%) when pooled together for all food and time intervals. One of the animals (Freek) made significantly more what-where choices (exact Chi-square: $\chi^2=4.898$, $df=1$, $p=0.036$; all other individuals exact Chi-square test, $df=1$, all $p>0.08$). However, this value did not remain significant following a Bonferroni correction. We also tested the number of what-where choices the animals made at each of the three time intervals regardless of the food combinations. None were successful above the chance level (exact Chi-square test, $df=1$, all $p>0.06$). All animals also failed to make significantly more what-where choices in each food and time combination separately (exact Chi-square test, $df=1$, all $p>0.07$). The animal's choices were, thus,

not guided by the distinction of which cups were baited and which were left empty in each trial.

As a measure of the animals' interest in the different food types we scored, whether or not they consumed the obtained foods. The preferred foods (apple sauce and yoghurt) were consumed in 98.5% of the obtained trials and red bell peppers were consumed in 97.2% of the obtained trials. Thus, during the testing stage the animals consumed any food they obtained regardless of preference, indicating they, in contrast to their behaviour during training, no longer discriminated between the qualities of the rewards.

Table 3.3: The number of trials and the what-where-when and what-where choices per animal in each time and food combination.

WWW stands for what-where-when choices and WW for what-where choices. WWW choices resulted in obtaining either the present or preferred food (according to the paradigm's predictions for each combination). WW choices were those made for either of the two baited cups, regardless of whether the food was still present at recovery. The animals had a 25% chance of making the correct WWW choice, for the WW choices this chance was 50%. Percentages are given in brackets. Significant values before Bonferroni correction are indicated by footnotes. None of the values remained significant after the Bonferroni correction.

Food and time combination		Test group						Control group	
		Claus	Eman.	Freek	Linda	Paul	Rene	Denis	Regina
Apple sauce & Red Bell Peppers (15 min)	# Trials	11	9	11	7	11	11	11	12
	WWW choices	2 (18)	1 (11)	6 (55 ¹)	3 (43)	4 (36)	3 (27)	3 (27)	2 (17)
	WW choices	4 (36)	4 (44)	8 (73)	5 (71)	5 (45)	5 (45)	6 (55)	5 (42)
Yoghurt & Red Bell Peppers (15 min)	# Trials	11	11	11	7	11	11	10	11
	WWW choices	3 (27)	2 (18)	1 (9)	1 (14)	4 (36)	3 (27)	3 (30)	4 (36)
	WW choices	6 (55)	4 (36)	7 (64)	3 (43)	7 (64)	4 (36)	5 (50)	7 (64)
Apple sauce & Red Bell Peppers (1 h)	# Trials	8	8	8	11	8	7	9	9
	WWW choices	3 (38)	0 (0)	2 (25)	6 (55 ²)	2 (25)	2 (29)	1 (11)	3 (33)
	WW choices	7 (88)	3 (38)	5 (63)	8 (73)	2 (25)	2 (29)	6 (67)	5 (56)
Yoghurt & Red Bell Peppers (1 h)	# Trials	10	10	10	8	10	10	10	10
	WWW choices	1 (10)	5 (50)	3 (30)	3 (38)	3 (30)	3 (30)	3 (30)	5 (50)
	WW choices	4 (40)	5 (50)	6 (60)	6 (75)	5 (50)	5 (50)	6 (60)	7 (70)
Apple sauce & Red Bell Peppers (5 h)	# Trials	9	9	9	8	9	8	8	8
	WWW choices	5 (56 ³)	2 (22)	4 (44)	2 (25)	3 (33)	3 (38)	1 (13)	1 (13)
	WW choices	7 (78)	4 (44)	5 (56)	5 (63)	5 (56)	4 (50)	3 (38)	2 (25)
Yoghurt & Red Bell Peppers (5 h)	# Trials	10	10	10	8	10	9	9	10
	WWW choices	3 (30)	3 (30)	3 (30)	2 (25)	3 (30)	1 (11)	3 (33)	4 (40)
	WW choices	6 (60)	5 (50)	7 (70)	4 (50)	5 (50)	4 (44)	4 (44)	6 (60)

¹ exact Chi-square test: $\chi^2=5.12$, $df=1$, $p=0.034$

² exact Chi-square test: $\chi^2=5.12$, $df=1$, $p=0.034$

³ exact Chi-square test: $\chi^2=4.48$, $df=1$, $p=0.049$

Testing: Location-based choices

The third possible strategy we examined was, that the animals' behaviour was directed by the formation of an association between a location and its potential to yield a reward. We called these the location-based choices. We considered both foods as a reward. Such choices would not be based on any information recalled from each baiting event, but simply on the different reward qualities of the four distinct locations.

First, we tested whether each individual showed a preference for a specific location, regardless of its content (Table 3.4), considering their choices from the entire testing period. All but one animal (Regina) showed a clear location preference (Table 3.4), as the number of choices they made for each location was significantly unequal, also after the Bonferroni correction. We also checked whether the number of rewards the animals obtained at each location differed (Table 3.4). All but one animal (Regina) were unevenly rewarded at each location (Table 3.4). After a Bonferroni correction, this remained significant for three individuals (Emanuel, Claus and Paul). As the animals were choosing certain locations more often than others, they also obtained more rewards from these locations.

We were interested in whether this relationship between chosen and rewarded location could be a result of certain decision rules. We considered two possibilities: the win-stay lose-shift strategy and an associative learning process across the first half of the testing sequence.

We tested whether the animals based their choices on a win-stay, lose-shift strategy across the entire testing period, regardless of the time and food combinations. We counted for each animal how many times they performed the following behaviours: win-stay (if the chosen cup was rewarded irrespective of food type, the following choice is for the same cup), win-shift (if the chosen cup was rewarded, the following choice is for a different cup), lose-stay and lose-shift. By means of a Chi-square test for a 2x2 cross table we tested whether the chimps behaved consistently according to this win-stay, lose-shift strategy. For one of the animals (Emanuel) we did find a significant relationship (exact Persons Chi-square test: $\chi^2= 5.34$, $df=1$, $p=0.039$), however, after a Bonferroni correction this value did not remain significant. No significant relationship was found for any of the other animals (exact Chi-square test, $df=1$, all individuals $p>0.16$).

We then proceeded to investigate whether the number of times each location was chosen in the second half of the testing sequence, was influenced by which locations yielded food in the first half of the training sequence. To this end we counted for each chimpanzee the number of times each location was chosen in the first and second half of the testing sequence, as well as the number of times each location was rewarded in the first and second half of the testing sequence (Table 3.4).

We used a row-wise matrix correlation (Vries, 1993) to test whether the number of choices for each location in the second half was related to the number of choices made for each location in the first half of the testing sequence. We obtained a positive non-significant Kendall's τ_{rw} correlation of 0.25 ($p=0.17$). We then computed a partial row-wise correlation between the choices in the first and second half, controlled for the number of rewards obtained in the first half, and found that the previous positive correlation completely disappeared (and even became negative): partial Kendall's τ_{rw} controlled for rewards obtained in the first half = -0.20 ($p = 0.24$). Thus, the cup locations' choices in the second half were made independently from the choices made in the first half of the testing sequence (Figure 3.2).

We then investigated whether the number of choices made in the second half was related to the number of rewarded choices obtained in the first half. We found a significant positive Kendall's tau_{rw} correlation of 0.40 ($p = 0.023$). Next, we computed a partial row-wise matrix correlation to see whether this correlation remained when we controlled for the number of choices made in the first half. Indeed, the correlation remained virtually the same: partial Kendall's tau_{rw} = 0.38 ($p = 0.024$). This shows that, it was indeed the rewards obtained in the first half of the testing sequences, and not the location choices themselves, which influenced the number of location choices in the second half of the testing sequence (Figure 3.2).

We used a row-wise matrix correlation to investigate whether the location choices made in the second half of the training were correlated to the location choices in the first half of the testing. The correlation resulted in a Kendall's tau_{rw} = -0.089 ($p=0.652$), demonstrating that the animals were not choosing the same locations under testing as under training.

Table 3.4: The number of times each animal chose and was rewarded at each location, separated for the first and second half of the study.

Significant values after Bonferroni corrections are indicated by*.

Name		Location 1		Location 2		Location 3		Location 4		exact Chi-square test df=3
		1st half	2nd half	1st half	2nd half	1st half	2nd half	1st half	2nd half	
Claus	Chosen	4	0	7	0	5	0	14	29	$\chi^2=72.46, p=0.001^*$
	Rewarded	0	0	4	0	2	0	7	12	$\chi^2=35.96, p=0.001^*$
Eman.	Chosen	13	19	14	1	1	0	1	8	$\chi^2=36.40, p=0.001^*$
	Rewarded	4	7	3	0	0	0	0	4	$\chi^2=14.45, p=0.002^*$
Freek	Chosen	13	7	14	10	0	0	3	12	$\chi^2=22.42, p=0.001^*$
	Rewarded	6	5	7	3	0	0	2	7	$\chi^2=10.27, p=0.017$
Linda	Chosen	5	4	11	17	7	1	2	2	$\chi^2=28.14, p=0.001^*$
	Rewarded	4	3	5	8	3	0	1	0	$\chi^2=14.0, p=0.003$
Paul	Chosen	23	19	3	0	3	1	1	9	$\chi^2=69.07, p=0.001^*$
	Rewarded	10	6	1	0	2	1	0	5	$\chi^2=21.56, p=0.001^*$
Rene	Chosen	5	1	11	24	8	2	4	1	$\chi^2=43.0, p=0.0000^*$
	Rewarded	2	0	5	7	3	1	2	0	$\chi^2=13.60, p=0.003$
Denis	Chosen	5	4	17	12	4	7	3	5	$\chi^2=20.68, p=0.001^*$
	Rewarded	2	2	10	5	3	4	1	3	$\chi^2=10.80, p=0.013$
Regina	Chosen	8	6	19	6	3	7	0	11	$\chi^2=9.47, p=0.023$
	Rewarded	4	4	12	2	2	2	0	6	$\chi^2=7.0, p=0.068$

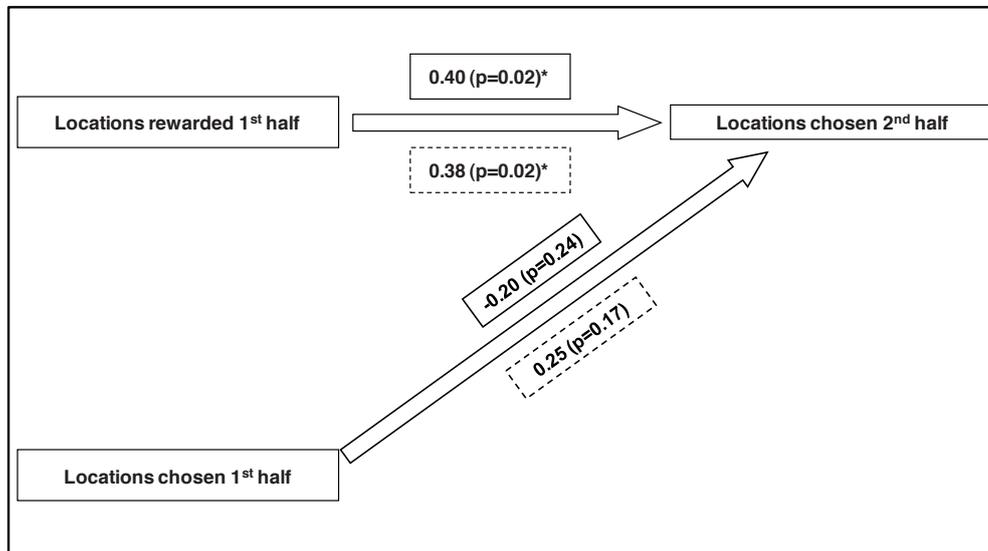


Figure 3.2: The relationship between chosen and rewarded locations.

The locations chosen in the second half of the testing sequence depend on the locations rewarded in the first half of testing, not on the locations chosen in the first half of the sequence. Values in the closed line boxes are Kendall's τ_{nw} correlations; values in the dashed line boxes are partial Kendall's τ_{nw} correlations. * indicates significant values.

Location based choices in training

We revisited the pointing training trials to see whether a location-based choice pattern was already visible then. We tested whether each animal chose all four locations equally often, based on the total choices made for each location. Exact Chi-square tests showed that two of the animals (Rene and Emanuel) were not choosing all of the locations equally often; only one of which (Rene) remained significant after a Bonferroni correction (Emanuel: $\chi^2=8.15$, $p=0.044$; Rene: $\chi^2=17.69$, $p=0.001$). The remaining six animals did not preferentially choose one of the four locations (exact Chi-square test, $df=3$, $p>0.18$).

We also tested whether each chosen cup was rewarded equally often for each animal. Exact Chi-square tests showed that only for one animal (Rene) the four locations were not rewarded equally often; however, this value did not remain significant after a Bonferroni correction (Rene: $\chi^2=9.00$, $p=0.030$). The remaining seven animals were rewarded equally often at each location (exact Chi-square test, $df=3$, $p>0.44$). Thus, just one out of eight animals showed a location based preference during training and none of the animals received significantly more rewards at any location.

DISCUSSION

The objective of this study was to investigate chimpanzees' episodic-like memory by means of a what-where-when food-caching paradigm (Clayton et al., 2001b). All individuals failed to pass the success criteria for demonstrating episodic-like memory in our set-up. The chimpanzees, however, did develop a location-based association strategy, based on the experienced reward quality of the four locations. Through association and not episodic memory they were able to locate, not where rewards were hidden, but at which location they had a higher probability of finding them. This behaviour reveals an interesting strategy of how the animals, when exposed to a multitude of changing features through-out the testing sequence (what, where and when), focused on the stable locations of the cup holders and, by means of a simple strategy, obtained apparently sufficient rewards.

All eight individuals successfully demonstrated a clear order of preference between the test foods, were able to reliably point to the cup they wanted to receive, and witnessed the different foods' rate of disappearance. In the testing phase three of the animals appeared to make more what-where-when choices than expected by chance, but each individual at a different time interval. None of these values remained significant after the Bonferroni correction. All significant what-where-when choices occurred in the apple sauce and red bell pepper condition. This may be attributed to a stronger difference in preference between these foods (compared to the difference between yoghurt and red bell peppers), which may have motivated them to pay more attention to these foods' locations. Nonetheless, to fulfil the success criteria for episodic-like memory the same animal would need to exceed the chance level of what-where-when choices in at least two time intervals. Success at least two time intervals would demonstrate a switch in choice strategy based on the presence of the food types. The fact that the animals failed to make more what-where-when choices than expected by chance, means they failed to integrate the unique trial locations of both foods (what is where), together with the time passed since caching (when), and adjust their choices accordingly for either the preferred food types (after the short interval) or the present food types (after the long interval). This choice strategy poses the highest cognitive requirements to the animals. Importantly, it is precisely the complexity of the task that is essential to conclusively demonstrate the presence of this advanced cognitive capacity (Clayton et al., 2001b). Previous work showed that great apes can solve a less complex paradigm involving what-where-when choices (Martin-Ordas et al., 2010). However, the demand on the flexibility and adjustment of behaviour is higher in our settings than in the previous study. Additionally, successful performance in this previous study may be ascribed to rule learning (Clayton et al., 2001b). From the current literature, including present work, none of the tested great apes or other primate species matched the response of corvids as tested by Clayton and colleagues (Clayton et al., 2001a).

Several explanations could account for our chimpanzees' failure. Firstly, the animals may have failed to obtain the knowledge about the temporal disappearance of

the test foods, or failed to integrate this with their what-where knowledge. Although of potential influence, we do not believe this to be the main explanation of our results. All of the animals showed a poor performance even at the two 15 minutes conditions in which none of the foods disappeared. Also, the two control animals (Denis and Regina) never experienced the foods temporal disappearance and still failed to make what-where choices above chance level. Two other potential explanations for the animals' failure are that they either do not possess the necessary cognitive ability, or that the executive demand imposed by our set-up was too high. In order to distinguish between these two alternatives, we first determined whether the animals' choices were based on any of the information given to them during the food hiding in each trial. This will illuminate which information provided by the set-up the animals were able to utilize.

We examined whether the animals were basing their choices on the 'what was hidden where' information, by looking at the so-called what-where choices. One animal did appear to make more choices for the two baited cups when all six conditions were considered together, however, the value did not remain significant after the Bonferroni correction. None of the other animals were successful above chance levels. This indicates that the animals were not basing their choices on the what-where information in each trial. Such poor performance contradicts other studies of chimpanzee long-term memory, in which chimpanzees were shown to remember the location of at least one food even up to 3 days (Martin-Ordas et al., 2010; Menzel, 1999). Our animals were also out-performed by rhesus macaques, who were able to remember what is where for up to 25 hours (Hampton et al., 2005). Again, none of the eight animals in our study performed above random chance even on remembering what is where for 15 minutes, a time interval that should not have exceeded the species' mnemonic capacity (Beran, 2004; Menzel, 1999; Menzel, 1973). The animals were thus not utilizing the information provided during the food hiding as a cue for their searches during recovery. We parsimoniously suggest that the testing procedure placed too high a demand on our subjects. This could be a result of several factors. Our procedure assumes that the animals, at the least, understand that they can recover foods from the locations where they observed foods being hidden. In fact, during pointing training when the food hiding was immediately followed by the recovery, the animals were successful at indicating the location of their preferred food. However, at this stage the animals only had to remember the location of one food (the preferred one) and this information only needed to be stored in their working memory, as retrieval was immediate. In testing, when food hiding and recovery were separated by intervals, the animals' success level dropped. Possibly, the combination of the time intervals and the need to distinguish the location of two foods between four options, may have limited the quality of the encoding of where the foods were hidden. Due to logistical reasons the animals only received one habituation trial per time interval of the testing procedure. Increasing the amount of such habituation trials could facilitate better understanding that the hiding locations were reliable cues for the locations of the foods during recovery. Finally, in contrast to other primates studies (Hampton et al., 2005; Martin-Ordas et al., 2010; Menzel, 1999; Schwartz et al., 2005) our animals

had previously only participated in one behavioural study and were therefore naive to the concept of “working” for food. This suggests that the testing of such complex abilities may require a large amount of training and experience with similar testing procedures. However, it is essential that the amount of training is appropriately balanced so that the animals test performance reveals their intrinsic capacities and not a trained response. The tested animals thus failed to use the information provided under the food hiding procedure to guide their behaviour when making their choices. They show no evidence of integration of the what-where-when elements and consequently do not show episodic-like memory in our study. Whether or not chimpanzees are indeed able to form episodic-like memories in the domain of food will need to be established in future studies. Future testing should ensure that the animals attend to the hiding procedure and that the necessary prerequisites for memory formation are present.

We further examined the pattern of the animals’ choices, to determine whether it was different from random choice. We considered a potential choice strategy based on the static locations of the four holders. We found that seven out of eight animals developed a significant location-based preference. For three of these individuals this also coincided with a higher rate of rewards at these locations, despite the fact that all four locations were baited approximately equally often. The animals’ initial preference could have been influenced by the cups colour or position relative to the subjects’ eye level or hand used for pointing. This location preference was further self reinforced, as persistent choices for a given location resulted in relatively more rewards obtained there. Interestingly, these location-based preferences developed during the testing phase. We found that a win-stay lose-shift strategy did not reliably explain the development of these preferences. We considered a more general association-based strategy. When looking at the number of each animal’s location choices in the first and second half of the testing sequence, we found no significant correlation between the number of times the chimpanzees chose each of the four locations in the first and second half of the study. In fact, when we controlled for the influence of the rewards obtained at each location in the first half the result was even a negative, non-significant correlation. This means that the animals were not choosing the same locations in the first and second half of the study, indicating a certain shift in the choices the animals made in the second half of the testing sequence, compared to the first half. Indeed, we found that the rewards the animals obtained in the first half influenced the choices made in the second half, even when we corrected for the choices made in the first half of the study. Consequently, choices in the second half depended on the number of rewards obtained at these locations in the first half of the test phase, but were independent from the number of cup location choices made in the first half. Thus, the animals’ behaviour is best explained by a location-based associative learning strategy. The animals formed associations about the potential of the different locations to yield rewards. This knowledge about the reward values of each location was formed through several experiences in the first half of the testing sequence, and then used in the second half of the testing sequence to guide their choices. On average this strategy

yielded, per individual, rewards in 45% of the trials received. Considering that the animals were tested with several trials per day, they obtained about one reward per day. Additionally, we noticed that the animals readily consumed any food they obtained during testing, indicating that, in contrast to their behaviour during training, they disregarded their food preferences. Since the chimpanzees were obtaining fewer rewards during the testing phase, the value of any food may have increased compared to the training trials. Given that this simple strategy resulted in a fair amount of obtained rewards, it is likely that the more difficult strategy in which the what, where and when had to be remembered, was not called upon by the animals.

In conclusion, we aimed to examine the chimpanzees' episodic-like memory by means of a what-where-when food-caching paradigm. Altogether, our chimpanzees showed a much poorer performance compared to scrub jays on a similar task (Clayton et al., 2001b) or compared to rodents, monkeys and apes on a simplified version of the task (Babb & Crystal, 2006a; Hampton et al., 2005; Martin-Ordas et al., 2010). Nonetheless, we maintain that none of the to-date present work, excluding the one on scrub jays, validates the demanded criteria for demonstrating episodic-like memory in primates. In other great ape studies (Martin-Ordas et al., 2010; Menzel, 1999; Schwartz et al., 2005) the animals response can be explained by more parsimonious explanations than the capacity to flexibly integrate the what, where and when elements. While chimpanzees are known to possess most of the cognitive tools required to a-priori solve episodic-like memory tasks, evidence remains slim and our results stress that whenever simpler alternative strategies can be satisfactorily used, chimpanzees may well rely on these. Given their natural skills in food-caching and recovery, corvids, such as scrub jays may have a head start to successfully and flexibly solve this type of task. Further research with carefully designed set-ups will be required to detect the potential for similar skills in non-corvid species.

ACKNOWLEDGMENTS

We thank the Biomedical Primate Research Centre for generously providing the facilities and the animal care taking staff for their assistance during training and testing. We are grateful to Ted Houser and Steven Hrotic for commenting on the manuscript and Jarmila Domažlicka for assistance with the graphics.

TOYS

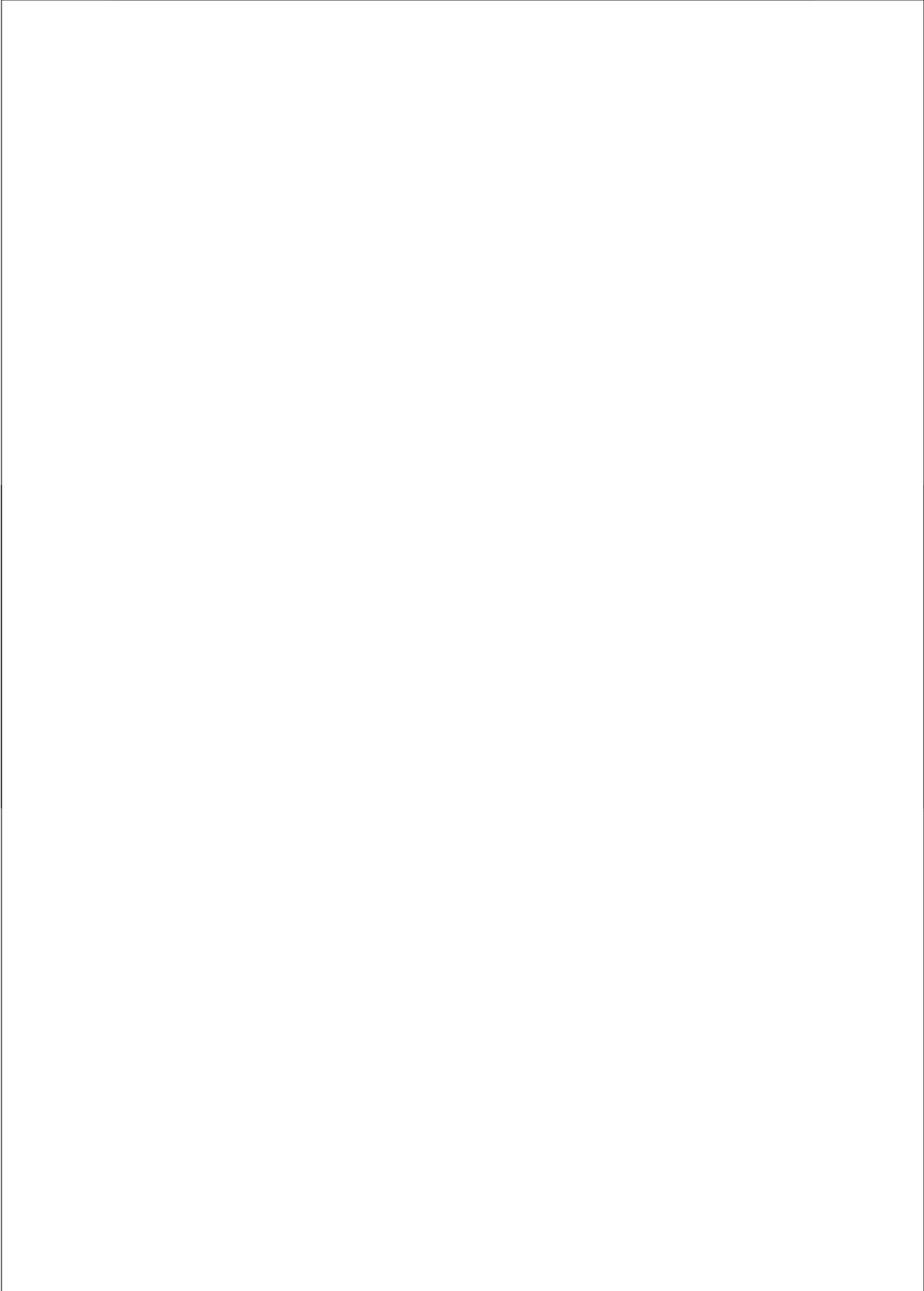
The Developmental Onset of Integrated What-Where-When Memory and Theory of Mind in Children

ABSTRACT

What-where-when (WWW) or episodic-like memory, the ability to recall what happened, where and when, has been demonstrated in corvids and rodents while primates have so far shown poorer performance. The WWW memory test is based on the criteria outlined by the food caching paradigm originating from scrub jay studies. WWW memory is comparable to the content requirement of episodic memory (EM) in humans, however recalling EM is said to additionally require auto-noetic awareness. Whether WWW memory relies on auto-noetic awareness is however unclear, because this cannot be assessed in animals. Auto-noetic awareness develops in children between three and five years and is thought to mediate the co-development of EM and Theory of Mind (ToM) which is the ability to distinguish own beliefs and thoughts from those of others. We designed a story test analogous to the food caching paradigm to test the WWW memory abilities of three and five year old children. This test included several questions that allowed us to examine whether the children could retrieve the what, where and when components separately or in an integrated manner. To investigate the developmental relationship of WWW memory with ToM we included four different ToM tests. We found that children could pass the what-is-where question before passing the ToM tests, but that the when question and the integrated what-where-when question were first passed after the ToM tests. Moreover, performance on ToM tests did not predict performance on the questions of the WWW memory task. These results suggest that the what-where-when memory test is cognitively demanding and that passing it relies on the development of a mature time-related skill not present until around the age of six, but not necessarily on the use of auto-noetic awareness.

Submitted to Developmental Psychology as:

Dekleva M., de Vries H., Kempes M. M., Dufour V., Spruijt B. M. and Sterck E. H. M. Does the Developmental Onset of Integrated What-Where-When Memory Occur After Theory of Mind in Children?



INTRODUCTION

The ability of animals to remember several features of a past event such as what happened when and where is known as episodic-like or what-where-when (WWW) memory (Clayton & Dickinson, 1998; Suddendorf & Busby, 2003). The definition of WWW memory is closely linked to the human ability to re-experience features of past events, called episodic memory (EM). However, whether WWW memory and EM are two similar or two different abilities remains to be established (Holland & Smulders, 2011). Tulving (1972) originally defined EM as the ability to store and recall detailed information about the temporal and spatial features of a unique episode (the what, where and when), but later added that episodic memories also should be accompanied by conscious re-experience or “autonoetic awareness” (Tulving & Markowitsch, 1998). In humans such awareness can be assessed by language, however, in animals researchers are limited to the observation of behaviour, as they cannot access the animals’ mental states. In order to investigate whether animals have cognitive abilities similar to those of humans, WWW memory has been defined based on Tulving’s 1972 definition (Clayton & Dickinson, 1998). According to this definition WWW memory and EM both refer to similar contents of memories but may differ in the re-experiencing aspect, as behavioural WWW tests do not measure whether the animals re-experience the content of their memories, although they may well be able to do so. In children the ability to report on re-experiencing an event (EM) as something one knows has happened to one self (autonoetic awareness) gradually develops around the age of four (Perner & Ruffman, 1995; Suddendorf & Corballis, 2007a; Suddendorf & Busby, 2005). Note however that already 2.5 year olds know something about the past as they are able to recount details about earlier events, such as a trip to the zoo (Fivush & Hamond, 1990). Around the age of four children also begin passing Theory of Mind (ToM) tasks, and are able to distinguish their own beliefs and thoughts from those of others (Perner & Lang, 1999). Both EM and certain ToM tasks rely on the development of autonoetic awareness and the ability to remember events as experienced. Therefore, children before and after the age of four, when the development of autonoetic awareness takes place, make excellent candidates for examining the cognitive requirements of the WWW memory test and in how far these requirements are similar or differ from those of ToM tasks.

Studies of WWW memory are typically based on criteria outlined by the food caching paradigm originally designed to test scrub jays (Clayton & Dickinson, 1998; Clayton et al., 2001b). By adapting these criteria to the behaviours of different species, WWW memory has been successfully demonstrated in several species of birds and rodents (Babb & Crystal, 2006b; Crystal, 2009; Eacott & Easton, 2010; Zinkivskay et al., 2009; Babb & Crystal, 2005; Ferkin et al., 2007). In the scrub jay set-up, the jays cache different food types for which they exhibited different levels of preference and which have different decay rates. To demonstrate memory of temporal information about when the foods were cached relative to recovery, as well as which food types were cached where, the animals had to recover after a short retention interval their

favourite of the two cached foods, but had to switch to the less preferred one after a long retention interval, as their favourite one had by then decayed. To confirm that the jays were indeed recalling integrated what-where-when information and not just searching for their preferred food after the short retention interval, and avoiding it after the long retention interval, a set-up with three different food types that all have different decay rates (Clayton et al., 2001b) was also tested. Two of the food types were cached at a time, and could be recovered after either a short or a long retention interval. The tested scrub jays were successful after all three time intervals and showed remarkable flexibility and adjusted their searches based on which of the two stored foods remained present in all six combinations (Clayton et al., 2001b).

Surprisingly, when non-human primates were tested on their ability to form and recall WWW memories, they have so far shown limited success. Great apes can pass a simpler version of the WWW memory test with two foods and two time intervals (Martin-Ordas et al., 2010), however, when chimpanzees were tested on the more demanding version of the task with three foods and three time intervals their competence may have reached a threshold and they used a location-based association strategy to find rewards (Dekleva et al., 2011). Amongst monkeys, rhesus macaques successfully remembered what was where, but failed to show the integration of the when component, when tested with two food types with different decay rates (Hampton et al., 2005). This difference in success rate between scrub jays and primates raises the question of whether the WWW memory test could be differentially demanding for scrub jays and non-human primates. Scrub jays may have a better WWW memory than non-human primates because of an adapted advantage due to their natural propensity to cache and recover food, and their performance might collapse if tested outside the caching context. Conversely, primates' WWW memory abilities might be better expressed in more relevant contexts for them. Cognitive abilities in primates have evolved in response to demands of complex social interactions as competition and cooperation (Humphrey, 1976; Byrne, 2002), therefore primates might reveal better WWW memory in the social domain. In humans, episodic memory relies on similar cognitive resources as planning for a future need (Busby & Suddendorf, 2005; Okudaa et al., 2003). Scrub jays successfully cache foods where they expect to be hungry in the future (Raby et al., 2007b; Correia et al., 2007). Interestingly, great apes also show planning for a future need for tools one hour in advance (Mulcahy & Call, 2006; Dufour & Sterck, 2008; Osvath & Osvath, 2008) and in the case of a single orangutan and bonobo even 14 h in advance (Mulcahy & Call, 2006). The success of scrub jays on both planning and WWW tasks indicate that the WWW memory test may be considered a behavioural test of EM, provided that the jays can extend their abilities to novel contexts. If great apes were to succeed in a WWW memory test in a social context, we might have a better picture of the equivalence of the WWW memory and EM tests. A recent study has reported that human adults used an EM strategy (i.e., re-experience an earlier event) to solve a WWW task and suggested that the WWW test might be a good test of EM in adults (Holland &

Smulders, 2011). Alternatively, the question of equivalence between WWW memory tests and EM tests can be studied by looking at children's response in WWW tests.

In children, the development of EM (sometimes referred to as autobiographical memory or considered a part of it) does not take place in isolation, but is dependent on the development of language, on the memory systems of early life (Nelson & Fivush, 2004) and on the developing understanding of the concept of time (Friedman, 2003; Friedman, 2005; Friedman & Lyon, 2005; Friedman, 2007; Grant & Suddendorf, 2010). Particular importance has been attached to the co-development of EM and various ToM abilities (Perner & Ruffman, 1995; Melinder et al., 2006; Naito, 2003; Perner et al., 2007), which children also begin passing around the age of four. ToM is considered a collection of interrelated abilities which mature around the same period and not a singular capacity (Naito, 2003; Sterck & Begeer, 2010). These different aspects of ToM are assessed by various tests. Several studies have found a significant relation between the performances in EM and different ToM tasks, in three to six year old children such as: (a) the see-know task (Perner & Ruffman, 1995), which assesses children's understanding of the link between perceiving and knowing, (b) the modality-specificity test (Perner et al., 2007), which assesses the understanding of which aspects of given objects can be perceived with which sense and (c) the unexpected content and transfer tasks, which both assess false belief (Naito, 2003). These studies suggest that the developmental relationship between ToM abilities and EM is mediated by the development of auto-noetic awareness, i.e., remembering events as personal experiences, which takes place around the age of four. Both self-awareness and the understanding of time are crucial aspects of EM in humans. Regarding ToM abilities in animals, chimpanzees do not show understanding of false belief, but may understand other's perception and goals (Call & Tomasello, 2008; Krachun et al., 2009). Scrub jays take into account what others see and may be aware of other's knowledge states, but these abilities have so far only been tested in the context of food caching (Clayton et al., 2007). Whether the success of the scrub jays on WWW tasks, planning and ToM pre-requisites in the food caching domain is evidence for their ability to distinguish their own current experiences from past experiences (EM), as is the case in humans, is still debated (Suddendorf & Corballis, 2007a; Clayton & Emery, 2009). Thus, we cannot exclude that the WWW memory task may well be a behavioural equivalent of EM. On the other hand, we cannot exclude the possibility that a WWW memory task measures the performance of integrating what, where, when information without reliance on auto-noetic awareness, as this cannot be directly measured in animals. Since children eventually develop both full-fledged EM and ToM, they are excellent candidates to examine the developmental relationship between WWW memory and ToM.

The present study tested children of three and five years on a battery of ToM tests, and a newly designed WWW memory story test. This WWW test was based on the same criteria as the three food item what-where-when scrub jay test (Clayton et al., 2001b). The relevant information was verbally explained to the children and acted out with props. The children could either verbally report their answers or point at

provided cards to give their answers. Because the children could verbally report on their knowledge we could separately assess their abilities to remember the what, where and when of hiding events. Moreover, the task contained a what-where-when question which assessed whether the children were able to integrate the what, where and when information. ToM was tested with the unexpected transfer test (Wimmer & Perner, 1983), the unexpected content test (Perner et al., 1987), the source of knowledge test (Wimmer et al., 1988) and the modality specificity test (O'Neill et al., 1992). Our main interest concerns the cognitive requirements of the integrated WWW task in relation to the ToM tests. Depending on the possible cognitive requirements for both tasks, we consider the following two hypotheses about the developmental onset of WWW memory and ToM:

- (1) If the WWW memory task relies mainly on the use of auto-noetic awareness, we expect that only children who pass the relevant ToM tests will be able to pass the integrated WWW memory question, and also that the outcome of this question can be predicted from the outcome on ToM tests.
- (2) If the WWW memory task is demanding because it requires the integration of memorized spatial and temporal features rather than auto-noetic awareness, then we expect that performance on ToM tests will not be associated with performance on the integrated WWW memory question and that the two abilities develop independently.

METHODS

Subjects

We tested 61 children of two age groups with an approximately balanced gender distribution (Table 4.1). For analysis, we only used the data of children that answered all questions in each of the four rounds of the WWW test and those that answered all questions in each of the four TOM tests.

Table 4.1: The total number of tested children in each age group and their gender distribution.

For both TOM and ELM the number of children per age group that answered all test questions and was used for analysis is given. B indicates number of tested boys, G number of tested girls.

Age range (years)	Number tested	Data used for analysis (all questions answered)	
		WWW task	ToM tests
3-4	31 (14B, 17G)	19 (10B, 9G)	30 (14B, 16G)
5-6	30 (15B, 15G)	28 (15B, 13G)	30 (15B, 15G)

Design

All the children were recruited from day care centers and primary schools in the vicinities of Utrecht and Amsterdam, the Netherlands and had Dutch as their native language. For all participants parental consent was obtained prior to testing. They

were tested individually in a separate room at their day care institution or school. Two experimenters were present; one interacted with the child throughout all tests, while the other taped the trial and scored the results. The entire study was conducted by three experimenters that exchanged these roles. For a given child all the tests were conducted on the same day.

The children were tested with one WWW memory test and four different ToM tests. The WWW test consisted of two major parts: a learning phase and a testing phase. Both of these parts included retention intervals during which the ToM tests were conducted. Half of the TOM tests were conducted in the retention intervals of the learning phase and the other half during the testing phase retention intervals. The order of presentation for each test was balanced across the sample.

What-where-when memory

This ability was tested by means of a two part story test. In the learning phase, the children were introduced to three toy characters. The three toys had a set order of importance and they displayed different rates of disappearance over time. In the test phase the children were asked to remember which toys were hidden in which location and after different retention intervals were given the opportunity to choose one location and obtain the toy inside, if it had remained present. The aim of this test was to determine the children's ability to integrate which toy was hidden where and how much time had passed since the hiding. The understanding of the temporal aspect was demonstrated by their successful choosing between the toys that are most important in the intervals where no toys disappeared and those that remained present in the time intervals where the most important toys had disappeared.

In the learning stage the three different plush toy characters (called *foofies*) that lived together in a castle were introduced. The three *foofies* had a set order of importance, which was explained to the children. We chose to set the importance order based on a fixed hierarchy and not on personal preferences, as the pilot study revealed that the children's toy preferences are unstable over time. The most important character was the princess who lived on the top floor of the castle. Just below her lived the slightly less important elephant. At the bottom of the castle there were the two least important characters, the two dragons. The *foofie* castle, thus, had four inhabitants in total: one princess, one elephant and two identical dragons. The child was asked to repeat the importance hierarchy of the *foofies*, and helped by the experimenter if unable to. Lastly an evil wizard was introduced, who did not live in the castle. The wizard had magic powder and wanted to make all the inhabitants of the castle disappear, so he could inhabit it. The experimenter asked the child to help them compete against the wizard and save all the castle inhabitants. The experimenter had obtained some of the wizard's magic powder so that they could first learn how it worked before competing against the wizard. For the learning of the magic powders effect on the toys' temporal properties, the *foofies* were removed from the castle and one copy of each character: a princess, an elephant and a dragon were placed under a sheet and the magic powder was applied to them. This was repeated three times with

retention intervals of 5 min, 15 min or 30 min. Following each interval a bell rung and the child and the experimenter removed the sheet and looked at what effect the magic powder had on the three toys. In order to help the children realize how much time has passed a modified clock was used. The amount of colored spots on the clock increased with the amount of time passed since hiding. After the 5 minute retention interval all three *foofies* remained present, after 15 minutes the princess was gone and after 30 minutes both the princess and the elephant had disappeared, leaving only the dragon. The child was asked to indicate which *foofie* had disappeared after each interval and, if required, was helped by the experimenter. After the last interval, the effect of the wizard's magic powder on all characters was repeated for the child with the help of a picture book.

In the what-where-when testing stage of the story, the child and the experimenter competed against the wizard to save all the *foofies* and return them to the castle. The test had four different rounds where two *foofies* at a time were hidden and following retention intervals of either 5 min, 15 min or 30 min, the child could save the most important or the still present *foofie*. The *foofies* were hidden in set pairs of a more important and a less important character. Thus, one pair was the princess and a dragon and the other was the elephant and the dragon. The princess and the dragon were hidden for both 5 and 15 minutes while the elephant and the dragon were hidden for 15 and 30 minutes. In both pairs, both *foofies* were present after the short interval, but only the less important dragons remained present after the long interval (Table 4.2). Therefore, within each pair of toys, at each retention interval a different toy had to be recovered, leading to four toys in total. Before the hiding the child was instructed to remember well where the wizard was hiding the different *foofies* and that they would only be able to save the one *foofie* that had not disappeared. The saved *foofie* was, when recovered, placed on its shelf in the *foofie* home castle. The *foofies* were in each round hidden in a four-box tower. To minimize a potential location bias two different box towers were used, each in two rounds. Each box was marked by a picture of a familiar object on its front. In each round the box towers were marked with a different picture of a familiar object on its top and were referred to as e.g. the mouse tower. The child was asked to name the pictures on each box, and the one on top of the tower. The two *foofies* were then hidden by the wizard in two different boxes while the child was watching. First, the wizard placed the *foofies* in two of the four boxes. Next, to ensure the child saw the content of all four boxes, it looked inside each box and then closed it. The same action was used for the baiting of all four boxes except that the wizard placed nothing in two of boxes. Once all four boxes were closed, the child was asked to repeat the content of all four boxes and if necessary the child could look inside the boxes again. Lastly, the wizard applied his magic powder to the tower and the child was instructed to wait until a bell indicated that the retention interval was over.

After each retention interval, the child was asked four questions while sitting in front of the box tower:

1. The temporal properties (TP) question: What happened before with the *foofies* when the clock had so many spots (while pointing at the modified clock)?
2. The what was where (WW) question: Which *foofie* was hidden in which box in this tower?
3. The what, where when (WWW) question: Which box do you want to open to save a *foofie* from the wizard?
4. The justification of the what, where, when (JWWW) question: Why do you want to open this box?

The children could indicate their answers verbally or point at pictures of the three *foofies* and the wizard that were placed on the floor in front of them. Subsequently, the child opened the chosen box.

Table 4.2: The correct choices for each of the tested character and time combinations in the WWW task and the importance position in the hierarchy for each character.

Interval	Princess and Dragon	Elephant and Dragon
5 minutes	Princess (importance rank 1)	Not tested
15 minutes	Dragon (importance rank 3)	Elephant (importance rank 2)
30 minutes	Not tested	Dragon (importance rank 3)

The hiding position of each of the *foofies* was balanced for each child, so that two different boxes were baited in each test round and the predicted choice was in each round hidden in a different location. The four test rounds were presented in two set orders so that the two *foofie* combinations were interchanged and the two 15 min intervals did not follow each other. The two orders were balanced across the sample. Order 1 was: princess and dragon at 15 min, elephant and dragon at 30 min, princess and dragon at 5 min and elephant and dragon at 15 min. Order 2 was: elephant and dragon at 15min, princess and dragon at 5 min, elephant and dragon at 30 min and princess and dragon at 15 min.

We scored the answers to each question in each round separately, as a pass under the following conditions:

1. The TP question: when the characters that had disappeared in the given retention interval were named.
2. The WW question: when both toy and box combinations were correct.
3. The WWW question: when the indicated box contained the predicted character for the given round.
4. The JWWW question: when the correct character for the given round was mentioned and the WWW question was answered correctly in the given round.

We evaluated the performance of each age group on each question both in each test round separately and across all four rounds. Performance on each question across the four rounds was scored on a scale from 0 to 4. Where possible (the WW question and the WWW question), we assessed the performance against chance levels. The chance level of answering the WW questions correctly, naming the correct locations of both *foofies* was 0.0825 based on the following calculation. The probability of correctly

naming the first *foofie's* location: correct foofie (2/3) x correct box (1/4). The probability of correctly naming the second *foofie's* location: correct foofie (2/2) x correct box (1/3) or correct foofie (1/2) x correct box (1/3). I.e. $2/3 \times 1/4 \times (2/2 \times 1/3 + 1/2 \times 1/3) = 0.0825$. The chance level of answering the WWW question correctly was 1/4, as there were four boxes.

Theory of Mind

This ability was assessed by four different tests, the unexpected transfer test, the unexpected content test, the source of knowledge test and the modality specificity test. Previous studies have identified these four tests as correlated to EM development (Naito, 2003; Perner et al., 2007; Perner & Ruffman, 1995).

Unexpected transfer test

The children received two different versions of an unexpected transfer false belief test (Wimmer & Perner, 1983). Both were sketch plays acted out with puppets and props by the experimenter. In both the protagonist placed an object in one location and while the protagonist was absent, this object was relocated. In the Barbie story, the object was a bar of chocolate, in the Maxi and Ann story this was a ball. In both stories, the protagonist returned after the relocation and the children were then asked the following three questions:

1. The false belief question: Where will the protagonist look for the object?
2. The true belief question: Where is the object really?
3. The memory question: Where did the protagonist originally place the object?

In order to correctly answer the false belief question the children need to understand that due to the protagonist's absence during the object relocation, the protagonist now had a false belief about the objects location, which differs from the child's own knowledge. The true belief question assesses their ability to understand and follow the story and extract the relevant information about the object relocation. The memory question assesses the children's ability to remember the correct sequence of events. In order to minimize transfer effects, the two versions of the test were administrated at the two different parts of the session, one during the WWW learning phase and the other during the WWW testing phase. The order of presentation was balanced across the sample. The children had to correctly answer all three questions to receive one point for each story test. A significant agreement between performance on the two tests was found ($Kappa=0.765$, $p<0.001$, $N=60$). A total unexpected-transfer score was made based on the performance in both story tests, with a range of 0-1. Children had to pass both story tests to receive a score of 1 of the total unexpected transfer score.

Unexpected content test

We tested the children with two different versions of the unexpected content or "Smarties" test (Perner et al., 1987). The children were presented with a familiar container (a chocolate box or a milk carton) and asked what they thought was inside. The container was then opened to reveal an unexpected content, namely a toothbrush

or buttons. These were then placed back inside and the container was closed. The children were then asked the test question for false belief: What would your mum think is inside this container? If the child did not answer, the question was rephrased as: What would your mum think is inside if we told her we just got it from the supermarket? This false belief task evaluates the child's ability to evaluate another person's false belief in relation to their own true belief. In contrast to the previous test, here the child is first given the opportunity to experience itself that it held a false belief about the content of the container before being asked to evaluate the belief of another person.

Each child received both versions of this test, as presented above. One version during the learning phase of the WWW test and the second during the testing phase, with the order balanced across the sample. The children obtained one point when they correctly answered the test question in each version of the test. A significant agreement between the two test versions was found (Kappa=0.765, $p < 0.001$, $N=60$). A total unexpected-content score was made based on passing both tests, with a range of 0-1.

Source of knowledge test

The test measures the children's ability to report on how they have acquired a particular piece of knowledge (Wimmer et al., 1988). The test consists of three different trials, the see, listen and no information trial. In each trial a different small toy was placed under a differently colored box, however, the child's access to this information was manipulated. In the see trial, the child observed the experimenter performing the action. In the listen trial, the child was facing away from the experimenter, but the experimenter said out loud what was being placed under the box. In the no information trial, the child was facing away from the experimenter and received no information about what the experimenter was doing. The same object and box combinations were used for all three trials: a dwarf was placed under a red box in the see trial, a horse under a green box in the listen trial and Snow White was placed under a blue box in the no information trial. After each trial the child was first asked what they thought was under the box. The test question followed: How come you know or do not know this? We balanced the order of the trials across the sample; however, we never started with the no information trial, since the children in the pilot study found this condition the most difficult. The children received one point for each correct answer to the test question per trial. No significant difference was found between the number of children that passed each of the three trials (Cochran's $Q=1.75$, $p=0.417$, $N=60$). A total source of knowledge score was made ranging from 0-1. The children had to pass at least two of the three trials to receive a 1 in their total source-of-knowledge score.

Modality specificity test

This test measures the children's ability to estimate which modality (seeing or feeling) should be used to find out about the color or the texture of an object (O'Neill et al.,

1992). The children were introduced to a modified cardboard box with two openings on two different sides. They were able to see but not feel the content of the box through the top opening, and feel the content without seeing it through the side opening. They received familiarization trials with both looking and feeling the content of the box to ensure they understood how the box worked. Following this, all children received two see and two feel trials in alternation. In the see trials the experimenter showed the child two objects that felt the same, but looked different (two socks of different colors, and two plastic turtles with different colors). The child was then asked to turn away from the experimenter, who placed one of the two objects inside the box. The child was then asked the test question: To find out which sock/turtle is inside the box would you like to see from the top or feel from the side? We also asked the children to clarify why they would like to use the indicated opening. This was used to control that the children understood the task, as they had a 50% chance of naming the correct modality. Similarly, in the feel trials the two objects looked the same, but were stuffed with a soft or hard content (two toy bears with a soft and a hard content, and two tubes of toothpaste one empty the other filled with plaster).

The order of the see and feel trials was interchanged and balanced across the sample with half the children receiving one see trial first and the other half receiving one feel trial first. The order of presentation of the modality specificity test was counterbalanced across the sample with the source of knowledge test, so that half of the children received one in the learning phase of the testing session and the other in the testing phase of the test session. The children received one point in each trial when they answered both what and why questions correctly. We compared the children's scores in each trial and found agreement between both see trials ($Kappa=0.763$, $p<0.001$, $N=60$) as well as both feel trials ($Kappa=0.676$, $p<0.001$, $N=60$). Agreement was also found between the number of children that passed both see trials and both feel trials ($Kappa=0.575$, $p<0.001$, $N=60$). A total modality-specificity score ranging from 0 to 1 was made, where children had to pass at least three of the four trials to receive a 1.

Data analysis

All sessions were videotaped as well as recorded manually during each session. Data were entered into Excel from the score sheets and, in case of inconsistencies, confirmed from videotapes. The data were analyzed with SPSS 16. All statistical tests were two-sided at a critical alpha of 0.05, unless otherwise stated.

RESULTS

We first present the comparisons of the performance of both age groups on all questions of the WWW test and the four ToM tests. Subsequently we present an overall comparison, between the children's scores on all ToM tests and all WWW task questions using, for each, a binary pass-fail measure.

Comparisons between age groups

What-where-when memory

For each of the four questions we first evaluate the performance of each age group separately and then compare the two age groups to each other. Where appropriate (WW and WWW questions), we also evaluate the performance against chance levels. Among the three year olds between 21 and 32 % of children passed the **temporal properties (TP) question** in the four rounds, while among the five year olds 39 to 57% passed (Table 4.3). For both age groups no significant difference was found between their performance on each of the four rounds (Cochran's test, $df=3$, three year olds: $N=19$, $Q=1.04$, $p=0.79$, five year olds: $N=28$, $Q=6.34$, $p=0.096$). A significant difference was found between the two age groups' overall performance on all four rounds on this question (Mann-Whitney U Test, one-tailed: $Z=-3.34$, $N=47$, $p<0.001$). The five year olds were significantly better at remembering the temporal disappearance property of the three *foofies* they had learned in the learning stage.

Between 26 to 63% of the three year olds passed the **what is where (WW) question** while between 57 and 82% of the five year olds did (Table 4.3). For both age groups no significant difference was found between their performance on each of the four rounds for this question (Cochran's test, $df=3$, three year olds: $N=19$, $Q=5.80$, $p=0.12$, five year olds: $N=28$, $Q=3.85$, $p=0.28$). A significant difference was found between the two age groups' overall performance (Mann-Whitney U Test, one-tailed: $Z=-2.77$, $p=0.003$, $N=47$). The five year olds were significantly better at remembering the correct locations of both foofies across the four rounds. We also evaluated the performance on the WW question against the chance level of correctly naming both foofie and a box combination, which is 0.0825 (see methods). The three year olds passed the WW question more often than expected by chance in three of the four rounds (Binomial test, $N=19$, both princess and dragon rounds and the elephant and dragon 15 min round $p<0.001$, elephant and dragon 30 min round $p=0.034$. After Bonferroni correction for multiple comparisons, the value from the princess and dragon 30 min round was no longer significant.) The five year olds passed the WW question in all four rounds (Binomial test, all 4 rounds: $p<0.001$, $N=28$. All values remained significant after Bonferroni correction.) Thus, except for the 3 year olds in the elephant and dragon 30 min round, both age groups remembered what was where above chance levels.

Between 26 and 53 % of the three year olds passed the **what, where, when (WWW) question** while the range among five year olds was 29 to 61% (Table 4.3) in the four rounds. For both age groups no significant difference was found between the 4 rounds on their performance on this question (Cochran's test, $df=3$, three year olds: $N=19$, $Q=3.77$, $p=0.29$, five year olds: $N=28$, $Q=7.48$, $p=0.058$). No significant difference was found between the two age groups on their overall performance (Mann-Whitney U Test, one-tailed: $Z=-0.90$, $p=0.168$, $N=47$). The two age groups showed similar

performance on remembering which foofie was hidden where and how long ago this happened.

We also examined the children's performance on the WWW question against randomly choosing the correct box, out of the four possible ones. We found that three year olds failed to pass the question more often than expected by chance in all four rounds (Binomial test, N=19, princess and dragon 5 min round: exact $p=0.018$, princess and dragon 15 min round: exact $p=1.00$, elephant and dragon 15 min round: exact $p=0.058$, elephant and dragon 30 min round: exact $p=0.35$. After correction for multiple comparisons, none of the values remained significant.). Thus, the three year olds did not succeed above chance levels in any of the four rounds.

The five year olds passed the WWW question significantly more often than expected by chance in both princess and dragon rounds (Binomial test, princess and dragon 5 min round: exact $p=0.002$; princess and dragon 15 min round: exact $p<0.001$, both tests N=28) and the elephant and dragon 30 min round (Binomial test, exact $p=0.002$, N=28). All values remained significant after correction for multiple comparisons. Their performance in the elephant and dragon 15 min round failed to exceed chance levels (Binomial test, exact $p=0.80$, N=28). They passed both princess and dragon rounds, indicating they successfully switched from the most important character to the one remaining present in that toy combination. Thus the five year olds successfully integrated the toys temporal properties as well as their importance hierarchy in the two princess and dragon rounds. The five year olds however showed a surprising failure in the elephant and dragon 15 min round. Although they could rather successfully integrate the importance hierarchy with temporal properties in the elephant and dragon 30 min round, they showed a poor performance in the elephant and dragon 15 min round (see Table 4.3). We will return to this inconsistent finding in the Discussion.

For the **justification of what, where and when (JWWW) question** the range of the correct answers among the three year olds was between 16 to 32%, and 29 to 46 % among five year olds (Table 4.3). So, in any of the four rounds only a minority of the three year olds and of the five year olds was able to justify its correct answer. For both age groups no difference was found between the four rounds on their performance on this question (Cochran's test, $df=3$, three year olds N=19, $Q=2.40$, $p=0.49$, five year olds: N=28, $Q=3.29$, $p=0.35$). A significant difference was found between the two groups overall performance (Mann-Whitney U Test, one-tailed: $Z=-1.83$, $p=0.034$, N=28). This shows that the five year olds showed more consistency between correctly choosing the box with the correct foofie as well as being able to justify their choice, compared to the three year olds.

The children showed poor consistency on answering all four questions (TP, WW, WWW and JWWW question), within each round and none of the children in either age group passed all questions in all four rounds. Between 0 and 2 three year olds passed all four questions within one round, while the five year olds performed somewhat better as between 4 to 11 children passed all questions within one round. The highest

consistency was shown among the five year old children in the princess and dragon 5 minute round, where 11 children passed all four questions.

Table 4.3. Results on the questions in the WWW memory task.

Percentage of children within each age group that passed each question in each round and the mean and sd of their performance across all 4 rounds. Mean individual scores range between 0 and 4. TP: temporal properties question, WW: what was where question, WWW: the what-where-when question, JWWW: the justification of the what-where-when question. Significantly better than chance performance, after Bonferroni correction ($p < 0.01$) is indicated by *. Double arrow indicates significant difference between the two age groups. For three year olds $N=19$, while for five year olds $N=28$.

Round	TP		WW		WWW		JWWW	
	3 yr	5 yr						
Princess & Dragon 5min	21	71	63*	82*	53	54*	32	46
Princess & Dragon 15min	32	39	53*	71*	26	61*	16	43
Elephant & Dragon 15min	21	50	58*	68*	47	29	16	29
Elephant & Dragon 30min	32	57	26	57*	37	54*	21	32
Mean score across 4 rounds (sd)	1.05 (0.85)	2.18 (1.06)	2.00 (0.94)	2.78 (0.79)	1.63 (1.21)	1.96 (1.17)	0.84 (1.07)	1.50 (1.26)

Theory of mind:

Unexpected transfer

We compared the performance of the two age groups based on the total score from both unexpected transfer story tests. A significant difference between the age groups was found (Chi square test for a 2x2 table, $\chi^2=29.43$, $p < 0.001$, $N=60$) with five year olds outperforming the three year olds (Table 4.4).

Unexpected content

The performance of the two age groups based on their total unexpected-content score showed a significant difference between the age groups (Chi square test for a 2x2 table, $\chi^2=24.09$, $p < 0.001$, $N=60$) with more five year olds passing both tests compared to three year olds (Table 4.4).

Source of Knowledge

Comparing the performance of both age groups based on their total source-of-knowledge score, we found a significant difference between the two age groups (Chi square test for a 2x2 table, $\chi^2=39.09$, $p < 0.001$, $N=60$) with five year olds outperforming the three year olds (Table 4.4).

Modality specificity

A comparison of the total modality-specificity-score showed a significant difference between the two age groups (Chi square test for a 2x2 table, $\chi^2=20.0$, $p < 0.001$, $N=60$), with a better performance by the five year olds (Table 4.4).

Table 4.4: Results of the ToM tests.

For each ToM test the number of children in each age group that passed each test and the mean and sd obtained from the total score from each test. For the modality specificity test the mean and sd for both trials within each modality (see and feel) are given. Double arrow indicates significant difference between the age groups.

ToM tests		3 year olds (N=30)	5 year olds (N=30)
Unexpected transfer	Barbie story	6	27
	Maxi& Ann story	6	26
	Unexpected-transfer score mean (sd)	0.13 (0.38)	0.83 (0.38) ←→
Unexpected content	Chocolate box	7	26
	Milk cartoon	7	25
	Unexpected-content score mean (sd)	0.17 (0.38)	0.80 (0.41) ←→
Source of Knowledge	See	2	23
	Listen	1	26
	No information	1	23
	Source-of-Knowledge score mean (sd)	0.03 (0.18)	0.83 (0.38) ←→
Modality specificity	Both see trials mean (sd)	0.06 (0.25)	1.13 (0.94)
	Both feel trials mean (sd)	0	1.13 (0.86)
	Modality-specificity score	0	0.50 (0.51) ←→
	mean (sd)		

Comparisons between performance on WWW and ToM

Developmental age of onset

To compare the children's performance on the two tested abilities, we made dichotomized (pass/fail) scores for all of the four ToM tests and the four WWW questions. For the ToM tests we used the total scores from each test as these were all in the range of 0 to 1. Thus, to score a pass in both the unexpected transfer and unexpected content test, both versions of each test had to be passed. The source of knowledge test was scored as a pass when at least two out of the three trials were answered correctly. The modality specificity test was scored as a pass when both test questions were answered correctly in three out of the four trials. By applying these strict rules for passing each test, we aimed to avoid scoring chance performance as a pass.

For the four WWW questions we used the following criteria. All four questions from the WWW test (TP, WW, WWW and JWWW) were scored as a pass based on the same criteria. Each question had to be correctly answered in two rounds with the same toy combinations (either princess and dragon or elephant and dragon or both). Answering each question correctly in both rounds with the same toy combination indicates that for the TP and WW questions the children had remembered the correct information enabling them to switch between the most important and the still present toy in the two given rounds. Answering both the WWW and JWWW questions correctly in both rounds with the same toy combinations indicates the switch from most important to still present toy has been made. By performing such a switch between the

toys the children demonstrate successful integration of the what, where and when information and we therefore considered this as passing criteria.

We used binary logistic regression analysis to fit logistic curves of predicted probabilities of passing each test with age as predictor. We plotted these on a graph and interpolated or when necessary extrapolated the curves for each test for ages from three to seven (see Appendix). For each test, we identified the age range where the predicted probabilities of passing the test were between 0.4 and 0.6 (Figure 4.1). Within this range we also identified the age at which half of the children would pass the test (i.e., predicted probability is 0.5) and refer to this as the *midway* age of developmental onset for the given test or question (indicated by vertical marks in the figure). In this way, we could compare the age ranges within which the children begin to develop the tested abilities. Note that for three of the WWW test questions (TP, WWW and JWWW) the age for which the predicted probability of passing the test is 0.5 is beyond the age range of the children in our sample.

Figure 4.1 allows us to compare the tested abilities in two ways. Firstly, we can order the abilities based on the likely midway age of onset and secondly we can compare the widths of the onset ranges. Considering the order of midway onset, we see that between the ages of four and five the children begin passing the WW question, as well as the UC, UT and SOK tests. The midway onset of the MS test is around 5.5 years of age. The midway onset of the TP, WWW and JWWW questions lie between 6 and 6.5 years of age. Development of the abilities necessary to pass the WW questions seems to take place prior to any of the abilities to pass the different ToM tests, while abilities necessary to pass the TP, WWW and JWWW questions only seem to be present about one year after ToM has developed.

The range of onset of each of the abilities may be related to the specificity of each test. Test outcomes with a wide range indicate that the abilities necessary to pass these tests are subject to large individual variation. Such variation may arise because these tests measure the development of several abilities that need to be present in order to succeed in these tests. Abilities with a more narrow range of onset indicate less individual variation, perhaps because these tests are more specific, i.e. measuring a single ability. Looking at the range, we see that all four ToM tests show a narrow range, while the four WWW task questions show a wider range of onset. Among the four WWW task questions the WW question has the narrowest range while the WWW question shows the largest range. We will discuss the possible cause of this very large range for the WWW question below.

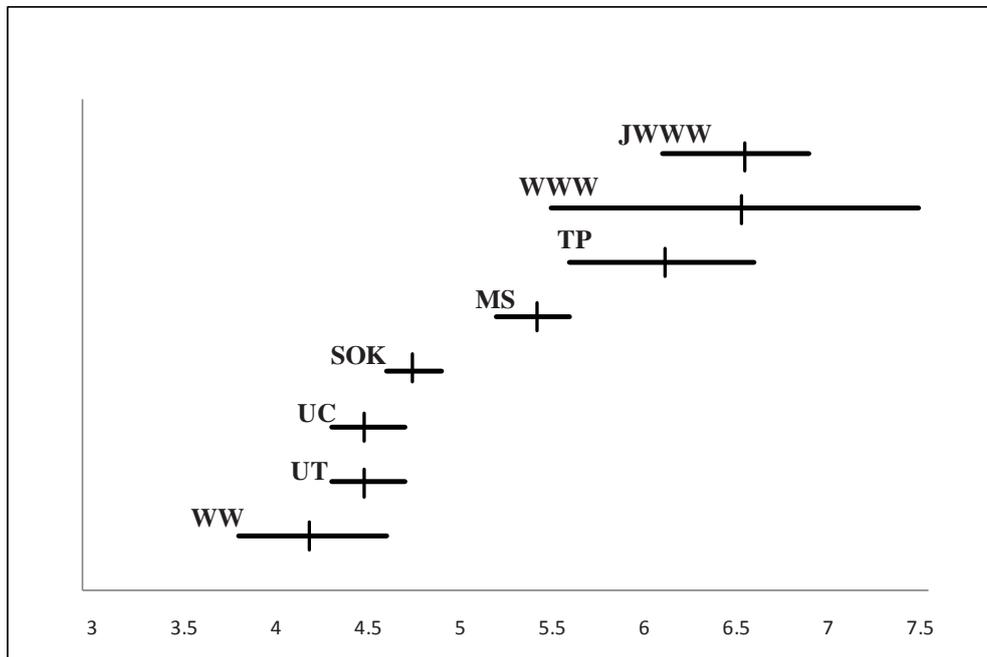


Figure 4.1: Age range of passing WWW and ToM.

Comparison of age range of passing the WWW task questions and the ToM tests. For each of the ToM tests (UT: unexpected transfer, UC: unexpected content, SOK: source of knowledge, MS: modality specificity) and all questions from the WWW test (TP: temporal properties, WW: what was where, WWW: what-where-when question and JWWW: the justification of the correct answer to the WWW question), the age range is presented where the predicted probabilities were between 0.4 and 0.6. The vertical lines indicate the age for which the predicted probability of passing the tests is 0.5, i.e., the midway age of developmental onset.

Independence between WWW questions and ToM tests

To further investigate the relation between performance on the ToM tests and the WWW task questions we compared the dichotomized scores of each child on the four ToM tests and the three WWW task questions that developed after ToM with Chi-square tests for 2x2 cross tables. Performance on the TP question was just significantly related to performance on the SOK test ($\chi^2= 4.81$, $df=1$, $p=0.049$) but not to performance on any of the other ToM tests (all other tests $p > 0.33$). None of the ToM tests was significantly related to the children's performance on the WWW question (all ToM tests $p > 0.36$). Performance on the JWWW question was significantly related to performance on the SOK test ($\chi^2= 6.37$, $df=1$, $p=0.023$) and not to any of the remaining ToM tests (all remaining ToM tests $p > 0.49$). Overall, amongst the four ToM tests used, performance on the SOK test was the only one that could predict performance on both the TP and the JWWW questions, while none of the ToM tests were related to performance on the WWW question. However, after Bonferroni correction for multiple comparisons these two predictive associations were no longer significant.

DISCUSSION

The aim of our study was to examine the cognitive requirements of the WWW memory task, to compare the developmental onset of WWW memory and ToM in three and five year old children and to examine the possible interdependence between these abilities. To investigate the WWW memory capacities we designed a story test based on the principles of a food caching paradigm that has been successfully used with scrub jays (Clayton et al., 2001b). The development of ToM abilities was assessed with four different tests, which were previously reported to be connected to the development of auto-noetic awareness and to EM in children (Melinder et. al 2006, Naito 2003, Perner and Ruffman 1995, Perner et al. 2007).

Our study confirms previous findings that ToM develops around the age of four (Perner and Lang, 1999). We also found significant differences between the 3 year and the 5 year olds, on all four ToM tests. Regarding the WWW task, we found that performance on the integrated WWW (what, where, when) question did not differ significantly between the two age groups while performance on the TP (temporal properties), WW (what is where) and JWWW (justification of what, where, when) questions differed.

Comparing the performance on WWW memory and ToM tests, we found that the age at which children could answer correctly the WWW question, in which they had to integrate the memorized what, where and when of toy hiding events, was higher compared to when they passed the ToM tests. This suggests that the developmental onset of an integrated WWW memory capacity occurs after the development of ToM. The separate component abilities required to correctly answer the integrated WWW question showed different developmental relationships with the ToM tests. The WW question was passed earlier than the ToM tests, whereas the TP question was passed only after the ToM tests were passed.

Finally, the performance on the TP, WWW and JWWW questions was independent from the performance on the ToM tests. Therefore, the outcome of the ToM tests could hardly predict the outcome on the TP, WWW and JWWW questions. This suggests that auto-noetic awareness, which forms the link between the co-developing abilities of ToM and EM, is probably not what makes this integrated WWW task demanding for children between three and five years. Rather, the late developmental onset of correctly answering the TP question suggests that it is specifically the understanding of the temporal component that makes the WWW task demanding.

We will discuss these findings under the following headings: (a) ToM tests: Comparisons between 3 and 5 year olds; (b) The WWW task: Comparisons between 3 and 5 year olds; (c) Does WWW memory develop after ToM? (d) Independence of WWW memory and ToM.

Theory of Mind tests: Comparisons between 3 and 5 year olds

The results we obtained in the four ToM tests all showed that the examined abilities develop between the ages of three and four, as has been previously established (Melinder et al., 2006; Naito, 2003; Perner & Ruffman, 1995; Perner et al., 2007). For both age groups, we found similar performance on both the unexpected transfer (UT) and the unexpected content (UC) tasks, both of which are false belief tests. In these two tasks the highest number of three year olds passed, compared to the other ToM tests, indicating that the cognitive abilities necessary to pass false belief tasks are already present in around 20% of three year olds. We observed a lower performance amongst the three year olds on the source of knowledge (SOK) test and the modality specificity (MS) test, indicating that the children had more difficulty in reflecting on how they came to know a certain fact (the source of knowledge test) and by which means they could obtain relevant information (the modality specificity test). The five year olds showed a high performance on the two false belief tests as well as on the source of knowledge test, while their performance was lower on the modality specificity test. Possibly this was due to the fact that the children had to specify in each trial both “what do you want to do” and “why do you want to do this”. The why question was introduced, because we only tested two modalities (seeing and feeling) and the children could therefore score 50% correct answers by simply naming the same modality throughout the test. By asking the children to explain why they wanted to see or feel the object in the box, the children had to verbalize their reasoning. This verbalization may have resulted in lower scores, since the children may have had difficulty explaining their choices. We found that in all four tests the five year olds outperformed the three year olds, indicating the development of ToM abilities in the tested age range as expected.

The WWW task: Comparisons between 3 and 5 year olds

We found that overall both age groups had difficulties with answering three of the four questions in our newly designed WWW test. Regarding the **WW question**, which measured if the children could remember the location of both toys for the specified length of time (5 min, 15 min or 30 min), we found that the average performance of the five year olds was higher than the three year olds. This question was relatively easy as both age groups performed well above chance in nearly all four rounds. While the five year olds were successful in all four rounds, the three year olds were successful in all but the temporally demanding elephant and dragon 30 min round.

In the **TP question**, the children were asked to report on how the magic powder had affected the presence of the three toy characters after each time interval (with the help of a clock). The five year olds outperformed the three year olds, although naming the still present toys after each time interval was difficult for both age groups. Knowing which characters had disappeared could be based on recalling the specific past event from the learning phase or on remembering facts that were acquired during the learning phase.

For the **WWW question** we found no significant difference in overall performance between the two age groups with rather low success percentages in both age groups. This indicates that integration of the what, where and when information was challenging for both age groups. The three year olds showed some success only in the round where the correct choice concerned choosing the most important character and shorter delays (i.e. the princess and dragon 5 min), however their performance did not reach significance after the Bonferroni correction. The five year olds were successful in the princess and dragon 5 min round and additionally in both of the two cognitively more demanding rounds (the princess and dragon 15 min round and the elephant and dragon 30 min round), where the correct choice was the box that still contained the hidden toy. Surprisingly, the five year olds performance on one of the temporally less demanding rounds, the elephant and dragon 15 min round, did not differ from chance. This failure is surprising as the five year olds were successful in the temporally more demanding elephant and dragon 30 min round. Their failure may be due to the way the children were taught the temporal properties of the three characters in the learning phase. Here all three characters were placed under a cover three times and recovered after 5 min, 15 min and 30 min. We noticed that the children were more engaged in which toys had disappeared than in which ones remained present, although the experimenter made sure the children did notice which toys remained present following each interval. During the learning phase, the princess had disappeared after 15 min while the elephant and the dragon remained present. Therefore, in the 15 min elephant and dragon test round the five year olds had to remember that both toys were still present and that therefore the box with the elephant, the more important toy, should be chosen. The fact that the princess had disappeared during the learning phase, which they found most exciting, now should be disregarded, because in this test round the princess was not hidden at all. Thus, a majority of the five year olds could successfully integrate the temporal properties of the toys with their hiding location and importance rank in three out of the four rounds.

In the **JWWW question** we examined the consistency of the children in naming both the box with the correct toy as well as naming the correct reason for choosing the given box within each round. The children's generally low scores indicate that they had a poor understanding of why they had chosen the correct box, or that they had difficulty with verbalizing their knowledge. The five year olds did however outperform the three year olds.

Overall, the results on our WWW task indicate the difficulty of this task for both age groups. Nevertheless, five year olds did outperform the three year olds on three of the four questions. The length of the retention intervals used may have influenced the level of the children's performance. However, as WWW memory is part of the long-term memory system, intervals of sufficient length should be applied. Better performance may be shown by older children, especially on the TP, WWW and JWWW questions as well as on overall consistency between the questions within the same round.

Does WWW memory develop after ToM?

A comparison between the children's performance on the two tested capacities showed that the midway age of onset for the WW question occurred before ToM, while for the TP, WWW and JWWW questions the midway age of onset was after successful performance on ToM tests. In order to compare the children's performance on the different ToM tests and WWW task questions, we dichotomized the obtained scores. Although we used strict criteria for this dichotomization, the results of this comparison should be interpreted with caution, as we will discuss below.

The midway age of onset for the WW question takes place just before the age when children start passing the UC, UT and the SOK tests (see Fig. 4.1). 50% of children from around four years can thus accurately remember what was where, even after intervals of 30 min. The ability to remember locations of hidden objects appears to be not demanding for young children. This has been recently shown also by three year olds with shorter intervals (Hayne & Imuta, 2011). This study on episodic memory in 3 and 4 year old children used a hide-and-see task, where the children had to search for three different toys hidden one after each other in three separate locations of their home. Both age groups showed similar performance in finding the hiding locations of the three toys.

Passing the TP, WWW and JWWW questions show midway age of onset between 6 and 6.5 years of age, well after passing the ToM tests (see Fig. 4.1). This suggests that the TP, WWW and JWWW questions were more demanding than the ToM tests. We attributed part of the relatively weak performance of the five year olds on the WWW question to misdirected attention during the learning phase resulting in a low performance in the not very demanding elephant and dragon 15 min test round. Their low performance in this round contributes to the late age of midway onset in our sample for this question. If the five year olds would have performed similarly as in the remaining three rounds, the midway age of onset for the WWW question would be about 1 year earlier. Therefore, the midway developmental onset time of the WWW question should be interpreted with caution. This uncertainty is also reflected in the broad age range of onset of the WWW question in Figure 4.1. Notwithstanding, even five year old children had difficulty with remembering the "when" information (TP question). The midway age of onset for the TP question takes place just after the children turn six. This suggests that it is particularly the temporal information that is difficult for the children to grasp. Previous studies have shown that children before the age of five do not seem to possess a mature time concept (McCormack & Hoerl, 2007; Povinelli et al., 1999; McCormack & Hoerl, 2005). Indeed, certain properties and processes concerned with temporal memory are only first understood by children between the ages of 8 and 12 years (Friedman, 2007).

The "when" information in our task was particularly demanding as the children had to remember a previously learned durational effect (the magic powder works differently on each toy dependent on the length of the interval), as well as use this information to determine which of the boxes will contain the correct toy. The integrated WWW question can thus be seen as tapping into the ability to reason about

durations in combination with memorized hiding locations of toys and their importance ranks. A somewhat similar understanding of time has been investigated with so-called temporal reasoning tasks, where children are tested on their ability to make inferences about the current state of the world by considering the temporal order between two relevant events. Consistent with our findings, children are not able to pass such tasks before the age of five (Povinelli et al., 1999; McCormack & Hoerl, 2005; McCormack & Hoerl, 2007).

Our study included a large number of tests and questions, asking a lot of the children's concentration and executive functions. It is therefore possible that the obtained results are influenced by this general cognitive demand of the test session. However, as the results obtained in the ToM tests are in line with those previously published, we believe this is not the main reason of the rather poor performance on the WWW test by the studied children.

ToM and WWW memory are independent

As they developmentally follow each other, we cannot exclude the possibility that ToM and WWW memory may both rely on auto-noetic awareness. Therefore, we tested whether passing of each of the four ToM tests could predict passing the TP, WWW or JWWW question, which have a midway developmental onset after that of the four ToM tests. Only the SOK test significantly predicted performance on just two of the WWW memory task questions, the TP and the JWWW questions ($p=0.047$ and $p=0.023$, respectively). However, after correction for multiple comparisons these predictive associations were no longer significant. Thus, WWW memory appears to be independent of ToM. This suggests that the auto-noetic awareness of EM, which co-develops with ToM and is developed in five year olds (Perner & Ruffman 1995; Naito 2003), is not predictive of a good performance on the WWW question. Apparently other cognitive abilities, which do not form a necessary requirement for ToM, such as understanding of time and reasoning about temporal features in relation to object properties, are required for WWW memory in children. This, however, does not exclude the possibility that episodic memory can be used to solve WWW tasks. Recently, a hide-and-seek task with adults has shown that they reported using a recollective strategy while solving a what-where-when task (Holland & Smulders, 2011). The authors suggest that their WWW task may be a good test for assessing EM in adults.

Suggestions for future research

Since our study shows that young children are not able to pass the WWW memory question due to poor time understanding, the cognitive demands of a similar WWW task should be further examined with children older than five. Older children can be directly asked how they retrieved the what, where and when information and whether they re-experienced the trial-unique hiding events or remembered these as singular facts.

During our test we noticed that some children would attempt to retrieve toys for which they had a personal preference (which could also change during the test session) and disregarded the taught importance hierarchy of the toys. Future studies should include a control for this by asking the children to repeat the importance hierarchy of the toys in the testing phase.

A limitation of the current study is that it did not include one of the tests currently used to measure the performance of episodic memory in children, such as the free recall test (Perner et al., 2007). Unfortunately, due to methodological difficulties, we were not able to complete such a test in our study satisfactorily. Studies that include both a WWW task and an EM task are necessary to examine the possible link between these two tasks directly (cf. Holland & Smulders, 2011).

Our results suggest that future tests with non-human primates may be conducted in the physical domain; however, particular attention should be paid to the method of teaching the primates about the temporal properties of the foods.

Finally, a completely non-verbal WWW task for children would, although much more time consuming, allow for a direct comparison with animal studies. Especially a completely non-verbal phase to teach the children the temporal properties of objects would give insight into whether young children could use such temporal properties in recovering hidden objects without an explicit understanding of the concept of time.

Conclusion

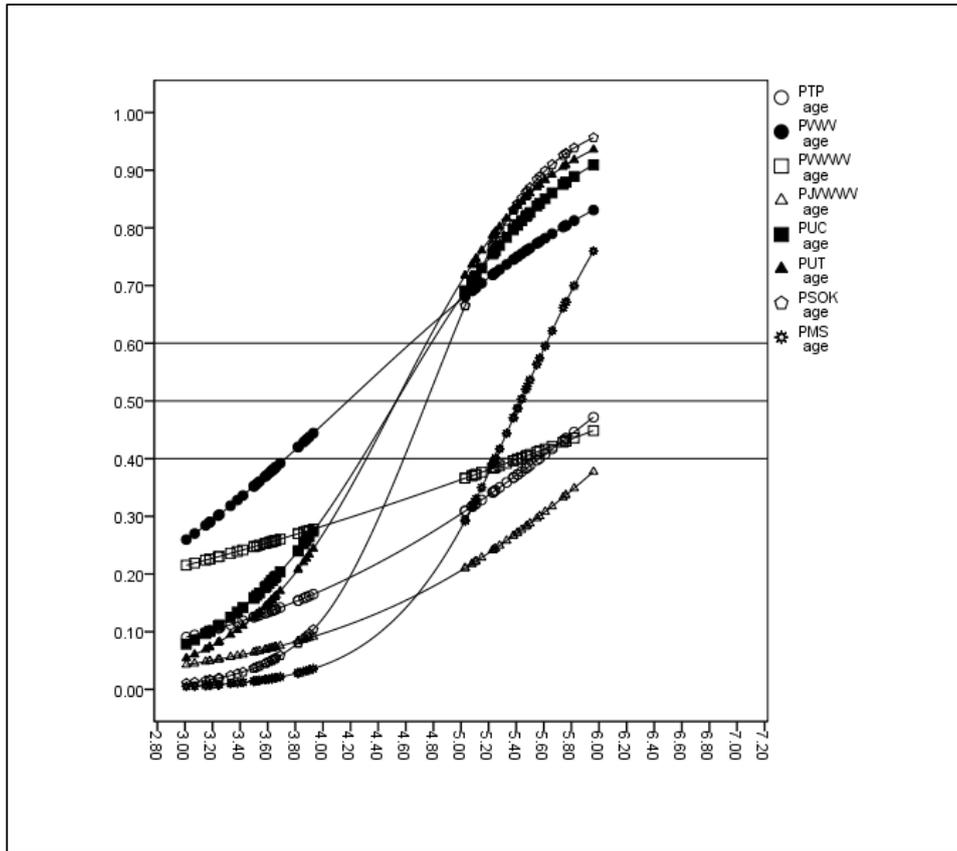
In conclusion, children begin to develop integrated WWW memory abilities at about 5 to 6 years of age. Whereas the three year olds only successfully passed the WW question in three of the four rounds and failed on the WWW question, a majority of the five year olds, but certainly not all, was able to integrate the toys temporal properties with their hiding locations and importance ranks in three of the four test rounds on the WWW question. Our study shows that children develop integrated WWW memory abilities after they have developed ToM, following their fifth birthday. Thus, the presented WWW task appeared to be cognitively more demanding than the ToM tests. Moreover, the WWW memory abilities as measured in our newly designed WWW task seem independent of ToM. This suggests that the understanding of time and reasoning about duration in combination with other known or remembered facts, rather than reliance on auto-noetic awareness, seem to make this task demanding for children.

ACKNOWLEDGEMENTS

We are grateful to Linda Kooij, Iris Mulder and Joost Kunst for collecting the data and to the participating parents, teachers and children for volunteering for this study.

APPENDIX

Binomial regression graph.



Binomial regression graph showing the predicted probabilities of passing each ToM test and the four questions from the WWW task. P stands for predicted probability, TP: temporal properties, WW: what was where, WWW: what-where-when question and JWWW: justification of what-where-when question, UT: unexpected transfer, UC: unexpected content, SOK: source of knowledge, MS: modality specificity.

Binomial regression analysis results.

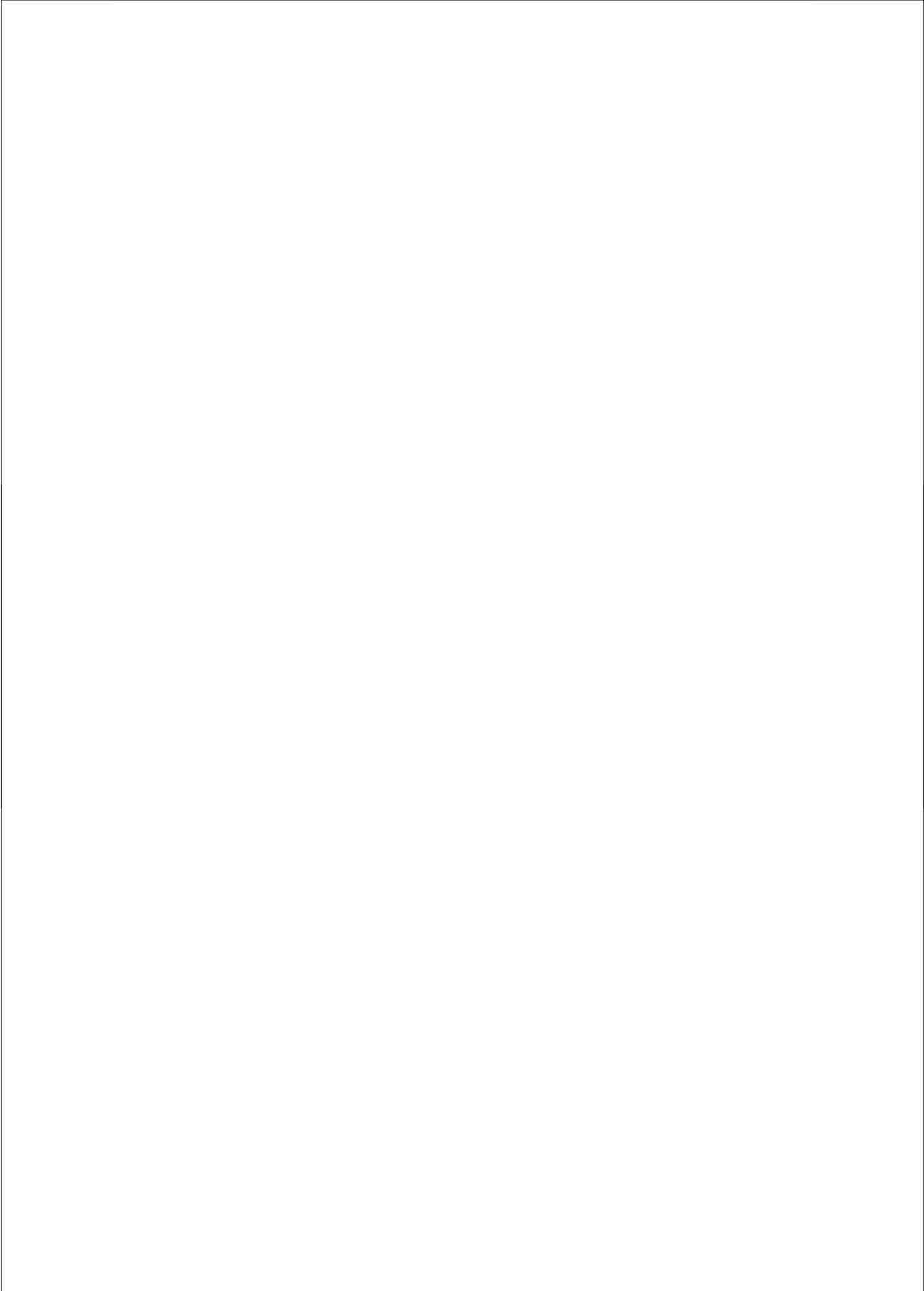
		Hosmer-Lemeshow goodness-of-fit test			Likelihood ratio test of age as predictor		Nagelkerke R square
		Chi-square	df	p	Chi-square (df=1)	p	
WWW task	TP	7.91	7	0.34	3.76	0.053	0.11
	WW	5.21	7	0.64	7.16	0.007	0.19
	WWW	8.31	7	0.31	1.18	0.278	0.03
	JWWW	5.29	7	0.63	3.69	0.055	0.12
ToM tests	UC	9.04	8	0.34	26.54	0.000	0.48
	UT	5.03	8	0.76	32.60	0.000	0.56
	SOK	5.70	8	0.68	44.55	0.000	0.70
	MS	3.90	8	0.87	22.90	0.000	0.47

The Hosmer and Lemeshow goodness-of-fit tests (Hosmer & Lemeshow, 2000) indicate that for each of the logistic curves the logistic regression model fits sufficiently well, since none of the associated *p* values is significant.

The significance of age as predictor is tested by means of the likelihood ratio statistic; this statistic is the difference between -2 log likelihood of the model including age as predictor and -2 log likelihood of the model without age. This statistic has a chi-square distribution with 1 df. The *p* values are significant for the WW question and for all four ToM tests and almost significant for the TP question and JWWW question. Age is not a significant predictor for failing or passing the WWW question. This is reflected in the broad age range (5.5 – 7.5 years) where the predicted probabilities of passing the WWW question lie between 0.4 and 0.6 (see Fig. 1).

The Nagelkerke R square value indicates the proportion of variance explained by age. This value is low for each of the separate questions in the WWW task. The values are much higher for the four ToM tests.

Discussion



The overall aim of the thesis was to examine the prospective and retrospective abilities of human and non-human primates as reflected in their behaviour. While auto-noetic awareness, a crucial aspect of mental time travel as defined in humans, cannot be assessed in non-verbal species, behavioural features of prospective and retrospective abilities can be tested on both human and non-human primates, and of course other animals. I experimentally examined these complex cognitive abilities from a behavioural perspective in different primate taxa in order to obtain a more comprehensive understanding of the shared and unique aspects of these abilities. I utilized methods successfully used in other species and attempted to fill in gaps in our knowledge concerning the presence of these abilities in three different primate taxa, by testing representative species of monkeys and apes as well as humans. Investigating what different species of primates know about their environment and how they deal with complex problems concerning their past and their future, will eventually lead to unravelling the evolutionary history of prospective and retrospective abilities in primates (Byrne, 2002; Martin-Ordas et al., 2010; Mulcahy & Call, 2006; Raby & Clayton, 2009).

While the prospective abilities of great apes have been successfully tested with the spoon test method (Osvath & Osvath, 2008; Mulcahy & Call, 2006; Dufour & Sterck, 2008), it has not been used to examine the prospective abilities of monkeys. In **Tools**, I present the first study to use a simplified version of the spoon test with a representative monkey species, and examine the planning capacities of long-tailed macaques. The macaques in my study were unable to solve the spoon task and failed to meet the criteria for planning in non-humans. However, after shaping to transport the tool for immediate use the macaques did display some flexibility by transporting tools over delays, and one also transported appropriate novel tools. The transport shaping provided the macaques with the correct sequence of behaviours to solve the task, and additionally seemed to open new avenues for them to flexibly display the shaped behaviour, indicating that macaques may have some prospective abilities, although inferior to those of great apes.

The three step what-where-when method is the most conclusive method to study what-where-when memory in animals and has been successfully used to show that scrub jays can integrate the spatial and temporal features of an episode (Clayton et al., 2001a; Clayton & Dickinson, 1998; Clayton et al., 2001b; Salwiczek et al., 2010). While both monkeys and great apes have been tested with the less conclusive two step what-where-when method (Martin-Ordas et al., 2010; Paxton & Hampton, 2009), the three step set-up has not been previously used with primates. In **Treats**, I adapted the three step what-where-when method and tested the retrospective abilities of a representative great ape species, namely the chimpanzee. Surprisingly, all eight individuals did not adjust their choices based on what was hidden where and also did not integrate the time passed since hiding and thus, did not show what-where-when memory in the domain of food. Instead, the animals found rewards based on a developed location-based association strategy and choose locations that delivered most rewards earlier in the testing sequence.

While ontogeny of verbal episodic memory has been investigated in children and found to coincide with Theory of Mind occurring between the ages of three and five (Melinder et al., 2006; Naito, 2003; Perner, 2000; Perner et al., 2007; Perner & Ruffman, 1995), the ontogeny of the behavioural what-where-when memory has not been investigated. In **Toys**, I investigated what-where-when memory and Theory of Mind in children of three and five years. The children passed the “what was where” question before passing the Theory of Mind tests around the age of four. Both the “when” question and the integrated “what-where-when” question were, however, first passed after the Theory of Mind tests around the age of six. Performance on the Theory of Mind tests did not predict performance on the what-where-when memory task, suggesting that performance on this what-where-when task may not rely on Theory of Mind which episodic memory was previously shown to rely on. However, the relationship between what-where-when and episodic memory remains to be directly examined. The tested what-where-when memory task may be especially demanding for children due to its reliance on understanding of the temporal feature, which may not develop before the age of five or six.

These results on monkey, ape and children retrospective and prospective abilities enable a comparison between different species and first suggestions on the possible evolutionary history of these abilities in primates.

COMPARISON OF PROSPECTIVE AND RETROSPECTIVE ABILITIES OF HUMANS, APES AND MONKEYS

Placing my findings in context with what was previously known about the prospective and retrospective abilities of primates seems to suggest that primate taxa may differ in their abilities. Although with such a small number of studies conducted and only a few individuals tested in each study, we should be cautious with reaching premature conclusions. Due to the scarce information, I make a very broad distinction between the abilities of humans, apes and monkeys and sketch a preliminary outline of the similarities and differences (Figure 5.1).

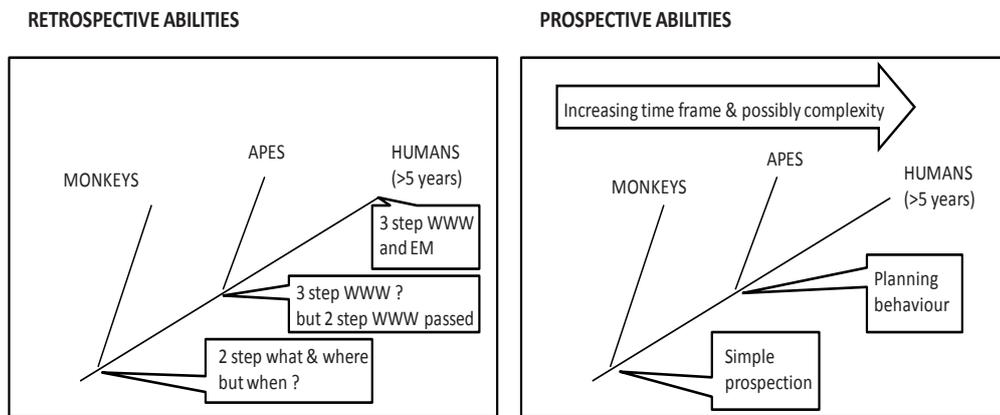


Figure 5.1. Preliminary phylogeny of prospective and retrospective abilities of monkeys, apes and humans, based on current knowledge.

WWW stands for what-where-when memory and EM for episodic memory. ? Indicates current findings are inconclusive.

Children begin passing the three step what-where-when test around the age of five and the failure of younger children seems to be caused by integrating the “when” feature (this thesis, chapter Toys). Around the age of five they also begin passing verbal episodic memory tests (Perner, 2000; Perner et al., 2007). Additionally, children tested with the spoon test were able to transport a tool and used it to obtain rewards in the future by the age of four, but were able to solve a similar task involving immediate rewards at the age of three (Suddendorf et al., 2011). This suggests that temporal displacement is the demanding and late developing feature for children concerning both retrospection and prospection. Future testing is required to examine how children further develop their prospective abilities over increasing time frames and also how and when they develop the ability to formulate complex plans.

Evaluating the retrospective abilities of great **apes**, based on the studies that rely on the what-where-when method, great apes can pass the two step set-up (Martin-Ordas et al., 2010). Chimpanzees, tested with a three step method (this thesis, chapter Treats) did not show what-where-when memory and did not use their long-term memory altogether, thus the extent of their retrospective abilities is at present inconclusive. Nevertheless, chimpanzees can remember information for extended periods of time (Bearn et al., 2000; Beran, 2004) and can remember baited locations for up to three days (Menzel, 1999; Menzel, 2005). Concerning planning great apes have successfully passed several spoon test studies, including with delays of 14 h and with novel tools (Mulcahy & Call, 2006; Osvath & Osvath, 2008). Based on current data, apes may have better prospective than retrospective abilities. However, this may be a consequence of more research effort invested into systematically investigating their prospective rather than their retrospective abilities and should be addressed by future studies.

Amongst **monkeys**, only rhesus macaques have been investigated on both what-where-when memory and future planning (Hampton et al., 2005; Paxton & Hampton, 2009). In both studies, the animals failed to integrate the “when” feature (Hampton et al., 2005; Paxton & Hampton, 2009). However, due to methodological issues (see Introduction chapter, this thesis) the retrospective abilities of rhesus macaques should be reassessed. Nevertheless, rhesus macaques are at a minimum able to remember what is where for 25 hours (Hampton et al., 2005). Regarding prospection, two squirrel monkeys did pass the satiation test (Naqshbandi & Roberts, 2006), however due to a gradual increase in correct responses, this is suggestive of learning rather than foresight (Shettleworth, 2007a). The long-tailed macaques I tested (this thesis, chapter Tools) failed to meet the criteria for planning with the spoon test, despite reduction of the temporal component used with apes from 1 hour to 5 minutes. Current data suggests that monkeys seem to have limited prospective and retrospective abilities, possibly due to difficulties with integrating the “when” feature. In my study, the best performing monkey was after training successful on just a few trials after a 20 min delay, suggesting that macaques’ potential prospective abilities could be limited to a much smaller time window than those of apes, scrub jays and humans (Correia et al., 2007; Mulcahy & Call, 2006; Raby et al., 2007b; Tulving, 2005). However, in my study, before training the macaques also failed in the no delay condition, suggesting that at least in this case they were unable to perform the appropriate sequence of behaviours and did not fail due to the temporal displacement alone. Once experience with tool transport was obtained, they completed the task. In delayed gratification tasks, where animals need to refrain from consuming smaller rewards in order to later obtain larger rewards (Dufour et al., 2007; Pelé et al., 2010; Tobin et al., 1996), macaques fare better. Delayed gratification is considered an indispensable prerequisite for making future oriented choices, and long-tailed macaques are capable of delaying gratification for up to 10 minutes (Pelé et al., 2010), indicating some prospective competence. These tasks may represent circumstances where less complex actions are required to bridge time gaps. Further studies are required to establish whether monkeys are able to show some, although simpler, future oriented behaviours as suggested by my findings and those of delayed gratification studies.

Concerning our current knowledge on prospective abilities of different primate taxa as reflected in their behaviour (Figure 5.1), it seems that both monkeys and young children have difficulty with the “when” feature, while children of about five and apes can integrate the temporal displacement of a future event and show planning behaviour. Monkeys have so far shown simpler prospective abilities, which require further examination. Regarding retrospection, while children of about five can also integrate the what-where-when of a past event, the current results are inconclusive on whether apes can also achieve this and monkeys may only be able to integrate “what was where”, but also this result is currently inconclusive (Figure 5.1). Based on the current state of knowledge, apes show better performance than monkeys on both prospective and retrospective abilities, while children of about five years show

comparable prospective abilities with apes and show better performance than apes concerning retrospection.

The picture on prospective and retrospective abilities of primates is far from complete and further testing is required with current methods both amended with relevant controls and expanded to include more species on one hand and on the other examine the tested abilities in further detail revealing differences in time frames and complexities between different primate species. Below I suggest how our knowledge of different primates' abilities may be expanded both from a practical perspective concerning suggestions for future prospective, retrospective and comparative studies and from a theoretical perspective regarding essential changes concerning our current thinking about these abilities.

SUGGESTIONS FOR FUTURE STUDIES ON PROSPECTION

Future studies on prospection with primates should focus on the abilities required to pass the spoon test and may be improved with controls for simpler behaviours. In addition, the influence of prior experience with tools necessary to solve the task as indicated by my results needs to be further examined in both apes and monkeys. Lastly, I suggest how the spoon test may be used to expand our knowledge about the complexity of primate prospective abilities and their limitations concerning the time frames within which they are used, enabling better comparisons between the abilities of human and non-human primates.

The results obtained with the spoon test method in monkeys and apes may be strengthened by including several controls to insure, the animals are solving the task by means of planning. To exclude prompting cues, the animals should be tested in different physical locations than those used under tool use and transport training. The animals should also not have visual access to the test table from the waiting compartment, a condition I was not able to test due to logistical limitations and was also lacking in some great ape studies (Osvath & Osvath, 2008). All primates could additionally be tested on whether the tool is a conditional reinforcer, as was examined in one great ape study where tool transport was rewarded, but the tool was no longer required to obtain the rewards (Mulcahy & Call, 2006). If tool transport would not drop in this case, it would indicate that the tool was transported due to its association with a reward and foresight of a future task was not applied. Lastly, a flexible response in a novel setting also supports the argument that the animals' behaviour is not merely an associative response. I used a novel tool condition to examine the flexibility of the learned transport behaviour of macaques (this thesis, chapter Tools). When great apes passed this condition in a similar set-up, the authors argued they had learned the relevant affordances of the tool rather than responded to simple perceptual relations (Osvath & Osvath, 2008). However, to convincingly confirm this both apes and macaques should be further tested on their ability to distinguish between novel tools and distracter objects that differ from each other only in one aspect which is either

functionally relevant (shape, size) or irrelevant (colour). This would further illuminate how well the animals are able to foresee the future function of the selected tools. If they indeed can distinguish the tools relevant affordances, then they should choose indiscriminately between tools that differ in irrelevant aspects and show a preference for the novel tools that differ in functionally relevant aspect.

Future studies should specifically address the role of experience with tool use and potential learning necessary to solve the spoon task with monkeys and apes. Although the tool using skills of macaques may be inferior to those of chimpanzees, captive macaques are frequent tool users (Tomasello & Call, 1997) and their prospective abilities may be further investigated with tools. However, captive apes usually have access to enrichment objects and tools and may gradually learn to transport these by reinforcement outside testing contexts. If such gradual learning indeed takes place, the current differences in the training required to pass the spoon test, between monkeys and apes may be overestimated and warrants further attention. Additionally, wild living long-tailed macaques were recently reported using both pounding and sophisticated axe hammers (Malaivijitnond et al., 2007; Gumert et al., 2009), making them interesting candidates for future studies on planning. Perhaps these individuals would exhibit prospective abilities with time frames similar to those of great apes even in absence of specific transport training. The spoon test may also be used to further expand the range of primate species tested on prospection, such as capuchins and baboons who are also tool users (Tomasello & Call, 1997; Visalberghi & Tomasello, 1998) and human adults and children.

To further examine the complexity of primate prospection they could be tested with sequential tool tasks where one tool is necessary to obtain another tool, which in turn can be used to obtain rewards and where the provision of both is separated by delays. Additionally, primates should be tested on their ability to plan for a specific event in the future, an ability humans exhibit daily as well as on the limits of the time frames within which they can plan (Clayton et al., 2001a; Tulving, 2005). Implementation of these suggestions will lead to a more detailed understanding of the differences in prospective abilities between different primates and pinpoint which features of prospection influence such differences.

SUGGESTIONS FOR FUTURE STUDIES ON RETROSPECTION

Future testing on the retrospective abilities of primate may be continued with the three step what-where-when food paradigm, despite the inconclusive results of my study. Monkeys and apes have previously shown they can work in similar settings, as three species of great apes were successful in a two-step what-where-when study (Martin-Ordas et al., 2010), rhesus macaques succeeded on the what and where features (Hampton et al., 2005), and in my study the chimpanzees passed the training trials. Primates can remember for extended periods of time where foods were hidden, and potentially integrate the time passed since hiding (Martin-Ordas et al., 2010).

Several methodological aspects may be adjusted in future studies with the three step what-where-when memory to enhance the animals' performance. Firstly, future studies may include more training trials to reinforce that the foods can be retrieved from the trial unique locations in which they were hidden. This step should increase performance in the intervals where the preferred food remains present and reveal whether they can use the observed baiting locations as clues for where the foods can be found in each trial. Secondly, the number of used locations, in combination with the length of the retention intervals could be adjusted to minimize poor encoding quality and memory decay.

A major challenge of testing what-where-when memory will be the integration of the temporal component into the set-up. In my set-up, the chimps could learn that foods disappeared after different time intervals by observation. However, I was not able to determine whether the animals were able to integrate this temporal information, as they did not succeed on any of the test trials. Support for the fact that they can integrate temporal features comes from the great ape study with the two step set-up (Martin-Ordas et al., 2010). The apes avoided the locations where the preferred frozen juice was hidden after the long retention interval, but showed a preference for these locations after the short interval. The animals did not receive specific training with the frozen juice, but had experience with snow and large ice-blocks (Martin-Ordas et al., 2010). This certainly indicates that great apes have some knowledge about food in the temporal domain. The food depletion set-up that I and other primate studies used may mimic situations primates encounter in the wild, where other animals may deplete food sources, however such depletion is probably not very predictable and difficult to incorporate into ones decision making. Wild primates know the locations of different food sources (Normand et al., 2009; Normand & Boesch, 2009) and may even have knowledge of factors influencing the timing of fruit availability (Janmaat et al., 2006). Future studies could test the animals in a set-up where foods showed a temporal increase in palatability, mimicking fruit ripening. In the "food disappearing" set-ups, the animals need to, after the temporally more demanding longer intervals, choose the less preferred foods. In the proposed "food ripening" set-up, the animals would after the temporally longer intervals be able to obtain more valued (e.g. sweeter) rewards, which may lead to successful performance even with a three step what-where-when set-up.

SUGGESTIONS FOR FUTURE COMPARATIVE STUDIES IN HUMANS

As humans have episodic and what-where-when memory, they are excellent candidates for comparing these abilities. A comparative approach has several benefits. Adults and children can be tested on both abilities, allowing an assessment of whether the two tests rely on similar or different abilities. These insights can then be used in testing animals and pre-verbal children to attain whether successful performance on what-where-when tasks can be considered an indication of some features of episodic

memory. Lastly, by better understanding the abilities required for passing the what-where-when task in humans, we may also obtain insights into which features of this test are particularly demanding and why certain animal species pass the what-where-when task and others fail.

Several possible interpretations of the relationship between what-where-when memory and episodic memory in children have been suggested. In case children would start passing both tasks around the same age of four, this may suggest that the two are highly related abilities and that the what-where-when task may be a valid non-verbal test for episodic memory (Salwiczek et al., 2010). If children started passing what-where-when tests prior to episodic memory tests, this could either suggest that the what-where-when task measures some, but not all, aspects of episodic memory or that the task is indeed measuring non-verbal aspects of episodic memory which develop earlier than suggested by verbal tests (Salwiczek et al., 2010). However, these suggested interpretations seem somewhat circular in their reasoning as passing a what-where-when test both prior to and around the same age as episodic memory is suggested as indication that the what-where-when task may be considered a non-verbal episodic memory task. Unfortunately, I was not able to satisfactorily complete an episodic memory test in children and directly compare their what-where-when and episodic memory performance (this thesis, chapter Toys). Future studies should address the relationship between what-where-when and episodic memory directly, as my results merely indicate that what-where-when memory may not rely on Theory of Mind, whereas episodic memory performance was previously shown to rely on Theory of Mind performance. Therefore, whether the what-where-when test is a good alternative for a non-verbal episodic memory test in children remains inconclusive. My study indicates that children start passing the what-where-when memory late between five and six years, due to the demanding requirement of temporal understanding (McCormack & Hoerl, 2005; McCormack & Hoerl, 2007; Povinelli et al., 1999). Therefore, if young pre-verbal children are assessed on this test, the “when” feature should be taught in a more implicit observational way, as is the case with animals. Future tests could also examine children’s abilities to utilize different temporal clues and compare using “how long ago” and “when” information. When a rat study manipulated conditions so that presence of a preferred reward was either contingent on the time of the day (when) or the duration of the tested interval (how long ago), or both, the rats were initially successful only in the how long ago condition, suggesting a qualitative difference from human episodic memory (Roberts et al., 2008). However, a follow up study revealed that when how long ago information was irrelevant, rats were able to use the absolute when of the events to successfully recover foods (Zhou & Crystal, 2009; Crystal, 2009). Dependent on which temporal cues the children of different ages are able to utilise in a what-where-when task we can determine whether they are using temporal cues equivalent to those used in episodic memory or not.

A recent study compared verbal and non-verbal episodic memory in children with a modified hide-and-peek task inspired by the scrub jay tests (Hayne & Imuta, 2011), but used much lower requirements than my study and the jay studies (Clayton

et al., 2003a; Clayton et al., 2001b). Three year olds reported less information on the verbal recall and showed poorer performance on retrieving three toys in the correct temporal order in which they were hidden than the four year olds, but both age groups could accurately retrieve what was hidden where. Thus, as was the case in my study, the three year olds had difficulty with the temporal component. The authors conclude that three year olds showed rudimentary episodic recall and that reliance on verbal reports may underestimate their ability (Hayne & Imuta, 2011). However, this test did not require the integration of the temporal feature as the children did not have to adjust their choices based on the time passed since hiding, but rather long-term memory for locations of several hidden toys. Therefore, the relationship between episodic and what-where-when memory still needs further examination, with tests posing the same cognitive requirements as those used with animals.

A recent study with adults found that in adults' performance on the episodic memory task predicted performance on the what-where-when task (Holland & Smulders, 2011). This suggests that the what-where-when task may be used as a behavioural test of episodic memory in species and individuals that have auto-noetic awareness (Holland & Smulders, 2011). Whether what-where-when and episodic memory rely on similar abilities in other species, where auto-noetic awareness cannot be assessed, still remains to be answered. It would be interesting to test adults on both what-where-when and episodic memory, but manipulate the what-where-when task so that it could be solved by either remembering or knowing features of past events (Tulving, 1972; Tulving, 1983). This could elucidate the difference in behavioural performance as a consequence of using either episodic memory (remembering) or semantic memory (knowing), and behavioural performance could be compared to that of pre-verbal children, impaired adults and animals. Tests with older children and adults could also include direct questions about whether the hiding events are remembered or re-experienced, to examine the effect of auto-noetic awareness on their behavioural performance.

SUGGESTIONS FOR FUTURE RESEARCH DIRECTIONS

I now turn to essential theoretical issues that should be addressed to enable a more comprehensive and broader investigation of prospective and retrospective abilities across different species in the future. When studying the prospective and retrospective abilities of human and non-human animals, several differences in how they are currently investigated stand out. Firstly, the abilities are differently defined, so human studies are focused on verbally reported abilities, and animal studies are focused on behavioural aspects (Holland & Smulders, 2011). Secondly, as a result of the different definitions, the test methods also differ. Thirdly, human studies do not always explicitly measure the subject's use of subjective abilities, but often assume them (Eacott & Easton, 2010). This may lead to overestimation of the abilities measured in human subjects (Dere et al., 2008; Dere et al., 2006). Furthermore, when behavioural

abilities are investigated with animals, many different aspects and approaches are used, hindering between species comparisons. Lastly, while in human studies use of complex subjective abilities is sometimes assumed, in animal studies a variety of simpler alternative explanations of the animal's behaviours need to be excluded before more complex explanations can be accepted (Clayton et al., 2003a; Clayton et al., 2009b; Dere et al., 2008). All of these factors essentially result in different standards for establishing the presence and levels of these abilities between human and non-human animals (Dere et al., 2006; Raby et al., 2007a) and therefore following this research direction will always result in finding qualitative differences between humans and animals. This is not productive when trying to take a broader biological perspective to examine the similarities and differences in these abilities between different species.

Future studies may better enable a broader research direction on prospective and retrospective abilities by considering the following three essential issues. Namely, future studies should emphasize the behavioural aspects of mental time travel, especially planning behaviour, further the complex abilities involved in this behaviour should be addressed by means of triangulation and lastly a range of prospective and retrospective behaviours differing in complexity should be recognized.

Firstly, both human and non-human studies should focus on observable features of mental time travel, and here planning behaviour may be the better focus point. Certainly, by focusing on the behavioural aspects of what-where-when memory and future planning, as previously described in literature and as I have followed in this thesis, a step in this direction has already been made. What remains to be done is to unravel the exact range of behaviours that the human mental time travel influences. After all, while subjective cognitive abilities, as auto-noetic awareness, may well be the reason why humans can travel through time further and with greater flexibility than other animals, it is the behaviours that these cognitive abilities result in, that lead to adaptive advantages (Raby & Clayton, 2009). Pinpointing the behaviours it influences is far from straight forward, as for example indicated by Suddendorf and Corballis (2007b) who list 15 different hypothesis on the characteristics and functions of mental time travel in humans. Also identifying the precise behaviours effected by absent or damaged episodic memory system seems elusive. Individuals with impaired episodic memory, either through injury or neurodegenerative disease, become unable to experience the continuity of their own lives, as they cannot remember any personally experienced events or imagine any future ones (Klein et al., 2002; Tulving, 2002). Nevertheless, such individuals and young children can persist in stable and provisioned environments (Tulving, 2005) and their episodic memory limitations do not easily translate into given behaviours or actions. Additionally, episodic memories can be unreliable, as shown by reports on false memories of never experienced events (Loftus, 1997; Loftus & Pickrell, 1995) and the creation of false memories in laboratory settings (Roediger & McDermott, 1995; Stadler et al., 1999). Likewise, people often remember the contents of conversations or witnessed events differently (Kinsbourne, 2005; Schacter & Addis, 2007). This susceptibility of episodic memory for inaccuracy

suggests that the adaptiveness of mental time travel may not lie in accurately recording the personal history of individuals. Instead, it has been suggested that the adaptive value of this cognitive skill may be found in its prospectively oriented component (Suddendorf & Corballis, 1997; Suddendorf & Corballis, 2007a; Tulving, 2005; Schacter & Addis, 2007; Schacter et al., 2007). While behaviour resulting from re-experiencing details of past episodes is not directly observable in humans, as it probably supports a wide range of activities indirectly, the prospective component of mental time travel, future planning is reflected in directly observable behaviour. Therefore, future behavioural comparative studies should focus on planning and compare the complexity and flexibility shown by different species. While episodic memory in all likelihood serves as the storage of the building blocks, we rely on when imagining different potential future scenarios, these building blocks may not be retrieved as full and accurate representations of past events and are potentially intertwined with semantic memories as well. Therefore, when testing animals on the accurate retrieval of integrated what-where-when features of unique events we may be focusing on too narrow an aspect of mental time travel. Perhaps by tapping into the animals' ability to flexibly combine features or combinations of features of past events into future scenarios, we will obtain a more complete understanding of the function of mental time travel and its evolution.

Secondly, to address the complex abilities involved in behaviours supported by mental time travel a single definition and unified criteria of these behaviours should be agreed on and then tested in each species with different methods each meeting these criteria. By the process of triangulation (Heyes, 1993; Seed & Call, 2010), where the presence of the same skill is confirmed by several different methods, the presence of a complex skill can be confirmed and the endless cycle of applying simpler alternative explanations of animals behaviours may be broken. Therefore, the behavioural prospective and retrospective abilities of both humans and animals should be examined by means of triangulation, which will on the one hand allow a better understanding of the skill's complexity and on the other enable a more coherent comparison of the behaviours reflecting time travel found in humans and animals.

Thirdly, future research should be shifted from searching for the existence proof of human abilities amongst animals towards an examination of a range of prospective and retrospective abilities and their components, where some may be shared amongst different species and some may be species specific (Shettleworth, 2007b). Thus, instead of focusing on how human-like the abilities of other species are, we should define a range of abilities that species may use when facing complex problems concerning their future or their past. In this line, it has been proposed that different types of future oriented behaviours may be distinguished based on not requiring a sense of future, such as fixed action patterns and learned associations, and a range of those requiring a sense of future, such as prospective working memory, semantic future thinking and episodic future thinking (Raby & Clayton, 2009). Prospective working memory is here defined as the ability to consider responses over short delays, semantic future thinking is the ability to form future scenarios in the

absence of mentally experiencing oneself in the future event while episodic future thinking requires the mental experience of oneself in the event (Raby & Clayton, 2009). Examining which types of prospective behaviours are common and which are unique, will expand our understanding about how different species are equipped to deal with complex problems.

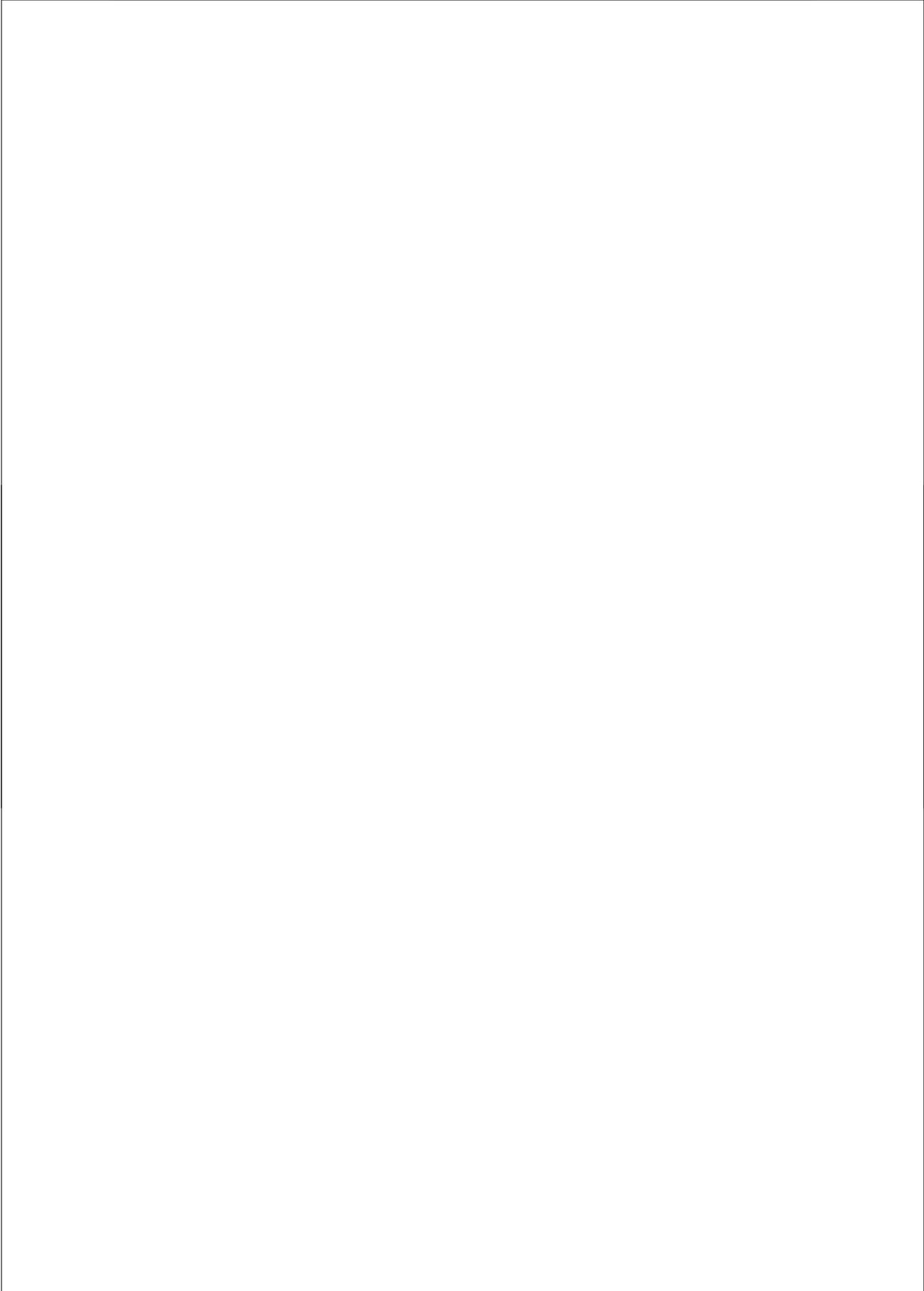
To sum up, future studies will benefit by clearly defining the observable features of mental time travel in terms of behaviours and then examining the complex abilities underlying these behaviours by several methods that each fulfil the same criteria and by not focusing on just one ultimate ability, but a range of behaviours that differ in complexity and may vary in how common they are amongst different animal species.

CONCLUDING REMARKS

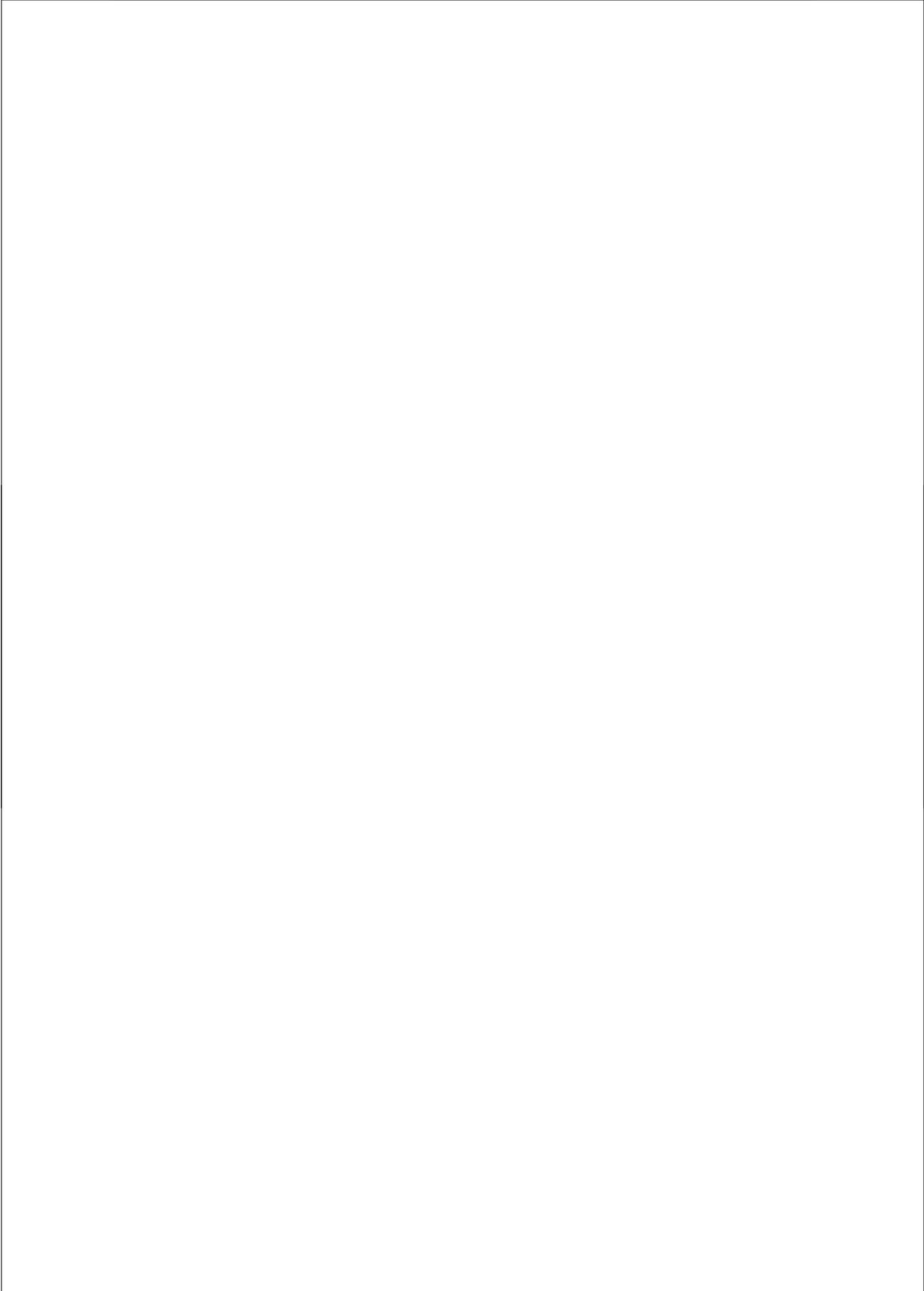
The aim of the present thesis was to examine the prospective and retrospective abilities of human and non-human primates from a comparative and a developmental perspective. Concerning the prospective and retrospective abilities of different primate taxa as reflected in their behaviour it seems that both monkeys and young children have difficulty with the “when” feature. Children of about five and apes can integrate the temporal displacement of a future event, while my study suggests that long-tailed macaques may have simpler prospective abilities limited to a shorter time frame compared to apes and humans. My results also suggest that children of about five can integrate the what-where-when of a past event, while younger children can only remember the “what” and “where” features of past events. Concerning the retrospective abilities of apes the current results are inconclusive on their ability to match the children on remembering integrated what-where-when memories, while it currently seems that monkeys may only be able to remember the “what” and “where” features of past events. Further testing is required before we can confirm the indicated differences between the primate taxa and further examine what are their causes. The studies presented in this thesis have merely scratched the surface of examining the prospective and retrospective abilities of human and non-human primates and the debate on animal time travel is far from finished.

One important challenge for future studies will be to address the stated disparity between human and animal abilities. Two important points should be recognized. Firstly, instead of claiming human uniqueness of certain cognitive abilities, the cognitive abilities of both human and non-humans should first be empirically examined and assessed based on their similarities as well as dissimilarities. In this process similarities should, however, not be confused with equivalence (Premack, 2007) and both quantitative and qualitative differences in abilities between human and non human animals may be recognized. In the case of mental time travel, even if a given animal species is found to have both what-where-when memory and planning, we will still not know whether this species is capable of mental time travel as

described in humans, as correlation does not imply causation. Nevertheless, the species will still have some noteworthy prospective and retrospective abilities, used in the contexts these abilities are adaptive to this particular species. Therefore, the empirical examination of the cognitive abilities should move its focus from searching for an existence proof of human-like abilities in other species, towards an understanding of which components of cognitive processes are shared among species (Shettleworth, 2007b). By following this path, we will also be able to examine and understand due to which selection pressures the advanced abilities of humans evolved. Secondly, this will open the door to acknowledging that other species may be sentient about their own experiences, either in ways similar to ours or ways unique to each species. The future does not lie in understanding how human-like the cognitive abilities of other animals are, but in what ways human, corvid, rodent and primate abilities are alike and in what ways they are different.



ETC...



ACKNOWLEDGMENTS

A good friend told me once, and has been repeating it ever since, that in life you can travel either by train or by boat. If you chose to travel by train, the tracks are laid and someone else has already built the end-station. On your journey, the train might break down and cause a delay, you may need to wait for a replacement or even walk along the tracks to the next station. Nevertheless, in the end, bound by the tracks laid by someone before you, you reach your final destination.

Travelling by boat on the other hand, is a very different story. Your journey begins at a port, with a sturdy boat, precious cargo, a reliable crew and a plan to reach a far away land. Whether you actually arrive there depends on a multitude of factors, most of which are beyond your control. If the wind is favourable, you can sail, if it is too weak, you stand still and must find additional supplies to feed your crew. If you encounter a storm, you may need to throw some of your precious cargo overboard and risk you will reach your final destination empty handed. You may be shipwrecked, attacked by pirates or burn down after a lighting strike. Your crew may stage a mutiny or fall ill to scurvy, leaving you alone to maintain and sail the ship. No matter what, when disaster strikes, you as the captain are still responsible for the ship, the cargo and the crew. You may want to crawl under the deck and hide out in your cabin, but ultimately you are the only one who can bring the ship to safe harbour. As sailing is dangerous, it is of course also exciting. You may discover treasures and wonders along the way, embark on new islands and continents, meet fascinating new people and essentially write your very own story and possibly become famous by selling your memoirs.

My PhD experience was certainly a voyage by ship and for better or worse, I was the designated captain. However, I could have not sailed the dangerous oceans and reached the far away promised land on my own accord. This voyage of discovery was a joint effort and there are many people whose contribution I wish to acknowledge. First of all, in case I forget to mention someone by name, I wish to express my collective gratitude to everyone whom I encountered on this journey for all assistance, advice, wisdom and encouragement you have extended me. I am humbled and grateful.

To Liesbeth I am thankful for giving me this opportunity and entrusting me the position as captain of my PhD ship as well as securing the finances for this ambitious expedition. Your creativity throughout the years was admirable. You also always let me pursue my own ideas. Most of all I am thankful for your relentless patience with the endless corrections of my dyslexic, comma-less manuscripts. To Berry I am grateful for the guidance, encouragement, personal and scientific support and the occasional drop of wisdom. I greatly appreciate you always made time for our numerous discussions on everything from the appropriate behaviour of evil wizards to consciousness and your favourite, the amygdala. You have been instrumental in keeping both my ship in one piece and me on a safe course. Thank you.

At the very beginning of my journey, I met the energetic and incredibly talented French captain Valérie. Your energy and enthusiasm were contagious and I felt honoured and privileged that we could sail our ships together through the unpredictable waters of Rijswijk. Our friendship has throughout the years been a cherished resource. Your kindness never ceases to amaze me. Under your advisement, I made the first detour on my journey and sailed to the bay of Chimpanzees to conduct my first experiments. The chimpanzees were amazing, fascinating and cheeky. I am appreciative of all the time they invested in staring at my numerous cups trying to find the precious applesauce, just to be rewarded with a peanut. I also greatly enjoyed working alongside and learning from all their experienced and dedicated caretakers: Wouter, Mariska, Erik, Antoine and Ruud. A special thanks to the very enthusiastic vet Merei for all the time you took to answer my endless questions on the best care of the chimps and for convincing your girls to spend their weekend gift-wrapping paper cups so my experiment could start on time. I was also blessed with great office mates both fun and always helpful, so ladies Marit, Marjolein, Sylvia and Capucine it was great fun, and thanks for teaching me Dutch pop song lyrics. For all the fun lunches and great company on the commute also thanks to Rob and Fred. You all kept me going when things got hairy and the waves were coming high.

As the chimps were relocated to the Beekse Bergen Island, I sailed my ship further to the bay of rhesus macaques. Here a very special thanks goes to my crew of sailors Jeannette, Noa, Christa and Coby, thanks for your hard work and commitment. Unfortunately, the seas in this bay never calmed down and after much deliberation and attacks of both seasickness and scurvy, I decided it was time to recuperate and departed ahead of plans as the ship was breaking at its seams. To Annet, Alwin, André, Fuad, Peter and Jan thank you for helping me with my research. I also wish to acknowledge the efforts of all the rhesus macaques that had worked so hard with us and bravely sat day after day observing which cups contained their favoured dried apples.

The next stop was the island of Ethology station in the bay of Utrecht. I had visited this island before and was extremely happy to return to its welcoming sea. Henk, Karlijn, Jaantje, Moris, Mikey, Max, Lisette, and the occasional visitors Brigitte and Ido made it a supportive and fun working environment. Dear Henk, thank you for your invaluable assistance and advice (on the human and non-human primates) and especially for your politically incorrect black humour. It is rare to meet a colleague who can manoeuvre it so masterfully and bilingually. Being able to share this with you was always like a very welcome breath of fresh air. I am also grateful for all the times I spent healing my migraines on your sofa in the honourable company of the only slightly snoring Maxi. Lastly, thanks for your hard work on my summaries. Dear Karlijn, thank you for your great humour and down to earth approach to the numerous problems I brought to you, the countless cups of coffee and all your help with manoeuvring the local tongue. I could not have wished for a better colleague. Dear Dr. Ido, we already met on my first stay on the Ethology station island and your calm, composed and critical approach has taught me tremendously about how to deal with

the everyday monkey crises and their various sabotage attempts. An essential skill... when the rascals literally eat your camera cables. And my very dear Dr. Brigitte, I am forever grateful to have met a kindred spirit in you. Without your advice, humour, excellent company, friendship, support and so much more, the many great storms might have succeeded in tearing my ship apart. Thank you for keeping me going on the days the winds were so still, it seemed the boat would never move again. And thank you for so many more things, that there is no space to express here. My crew was also greatly enriched by a young inhabitant of this island, Lisette. It was a pleasure working with you, knowing that both the monkeys were well taken care of and that you could so diligently follow the ever-evolving protocol of my experiments. Thank you for all your hard work and patience. Lastly a profound word of gratitude to the long-tailed macaques without the hard work of which, there would not be much to write about. All the hours they spent learning to manoeuvre tools are greatly appreciated. I was very happy when we, after much negotiation, could agree that they would work for fruit loops, of all things.

From here I also sent a small expedition to the Pavilion Island of Delta Phenomics. Firstly I am grateful to Berry for his support in letting me come on land. My thanks go to Raymond, Niek, Johanneke and Suzanne for your help and advice. A word of thanks also to the young and brave Lennart for taking it upon himself to communicate back and forth and tried to implement my ambitious ideas on an everyday basis.

The last stop on my journey was a surprise encounter, the Children Island. Here a thanks goes to the biggest optimist I ever met, Maaïke, your energy was infectious and all your practical advice was a key to our success. Another key were my hardworking sailors Joost, Linda and Iris. You relentlessly packed and un-packed suitcases full of toys and dragged them around from child to child to collect all the data. Thank you for your hard work and your essential assistance with the local tongue. Without you, the workload would have been overwhelming. I am also grateful to the many children who participated in all our crazy games, as well as their parents and the numerous schools and kindergartens, and Noldus that allowed us to test on their premises.

On my journey I also frequently sailed by the Kruijt Island. Simon, Marie-José, Johan, Jorg, Inge, Thijs, Anne, Esther, Zjef, Adriano, Madeleine, Serge, Sonja, Greg, Miranda it was a pleasure to have you as colleagues. Matt and Ulf, you were both unpredictable and sometimes over the top but always good company. Steven, I still miss your excellent humour and drinking tea with you. Steven and Will, thanks for making my manuscripts so much more eloquent than I ever could. Sanne, thanks for all the hugs and pep talks on the darkest of days. Charlotte, I greatly appreciate sharing the frustrations and joys of research with you and the great fun we had on our trip to France. Marco, you have been a great office mate, and essential support system at sharing the odd experiences one can encounter with the local tribe of the Dutch. Ellen, a tremendous thanks for being a great colleague and friend, you have been there for me in a hundred different ways. Dear Han, you were vital in calculating the correct

direction of my ship on its final leg of the journey. Thanks for the numerous hours you spent pondering with me over my data, searching for the parallels between applesauce and princess, thinking of creative ways to analyse my data and then explaining it all to me. Without your support, when the supplies on my ship were running dangerously low, it was very possible that my ship was heading straight for the dark Bermuda triangle. A humble, thank you.

At the very end of my journey, another very talented French captain came to my rescue, Cécile. Dear girl, thank you so much for helping me through the very last storm, for inspiring me to find my way. You are an amazing and beautiful person and one of my PhD heroes.

To all my friends, spread all over the world, thanks for bearing with me over the last few years. Anne-Christine, Anja, Jade, Sebastian, Mette, Irene, Jaja, Gesa, Andreja, Milena you are my worldwide support system. Every time I see each one of you, it feels like it was just yesterday we last saw each other. Your friendship is an amazing asset in my life. Thank you for keeping me sane and listening to all my monkey, science and miscellaneous rants. Ted thank you for your dedicated effort to sharpen and improve my manuscripts. To my parents, thanks for your support, for teaching me how to walk, how to sail and how to keep going through the toughest of storms. To my brother, thank you for being the best big brother. To my grandmother, thanks for teaching me how to read and write. To my great grand dad, thanks for so much more than words could ever express.

To my precious Claude, thank you for catching me and pulling me out of the water the many times I fell overboard. Thank you for always believing in me, even on days the wind blew me completely off course. Thanks for understanding, caring, supporting, listening, disagreeing and occasionally agreeing, for enduring. Thanks for being the creative mastermind behind my posters, presentations and the cover page of this book. Thank you for reminding me there is more to life than work. Please never stop. I am forever grateful you were there for me through it all in these past years and every day took me just the way I am.

So in the end of it all, I have sailed my ship, cargo and crew to a safe harbour. It has been challenging, exciting, frustrating, and hectic, humbling, stimulating, crazy, hilarious and educational once-in-a-lifetime journey. It was perhaps not the journey or the final destination I envisioned at the start, but such is the destiny of a ship's captain. You do not always know where you will end up dropping your anchor. Although exhausted from the very long journey and the many storms, I also feel immensely enriched by the experience and the many fabulous human and non-human primates I met on my journey. And just as soon as I catch my breath and dry my clothes I am excited to see where my next voyage will take me and Claude.

REFERENCES

- Addis, DR, Wong, AT & Schacter, DL** (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration *Neuropsychologia*, 45, 1363-1377.
- Aggleton, JP & Brown, MW** (1999). Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behavioral and Brain Sciences*, 22, 425-444.
- Atance, CM & O'Neill, DK** (2005). The emergence of episodic future thinking in humans *Learning and Motivation*, 36, 126-144.
- Babb, SJ & Crystal, JD** (2005). Discrimination of what, when, and where: Implications for episodic-like memory in rats. *Learning and Motivation* 36, 177-189.
- Babb, SJ & Crystal, JD** (2006a). Discrimination of what, when, and where is not based on time of day. *Learning & Behaviour*, 34, 124-130.
- Babb, SJ & Crystal, JD** (2006b). Episodic-like memory in the rat. *Current Biology*, 16, 1317-1321.
- Barrett, L, Henzi, P & Dunbar, R** (2003). Primate cognition: from 'what now?' to 'what if?' *Trends in Cognitive Sciences*, 17, 494-498.
- Bearn, MJ, Pate, J, Richardson, WK & Rumbaugh, DM** (2000). A chimpanzee's (*Pan troglodytes*) long-term retention of lexigrams. *Animal Learning & Behavior*, 28, 201-207.
- Beck, BB** (1980). *Animal tool behaviour: the use and manufacture of tools by animals*. New York: Garland STPM Publishing.
- Bentley-Condit, VK & Smith, EO** (2010). Animal tool use: current definitions and an updated comprehensive catalog. *Behaviour* 147, 185-221.
- Beran, MJ** (2004). Long-term retention of the differential values of Arabic numerals by chimpanzees (*Pan troglodytes*). *Animal Cognition*, 7, 86-92.
- Bird, LR, Roberts, WA, Abroms, B, Kit, KA & Crupi, C** (2003). Spatial memory for food hidden by rats (*Rattus norvegicus*) on the radial maze: studies of memory for where, what and when. *Journal of Comparative Psychology*, 117, 176-187.
- Boesch, C & Boesch, H** (1984). Mental Map in Wild Chimpanzees" An Analysis of Hammer Transports for Nut Cracking. *Primates*, 25, 160-170.
- Busby, J & Suddendorf, T** (2005). Recalling yesterday and predicting tomorrow. *Cognitive Development*, 20, 362-372.
- Byrne, RW** (2002). Evolutionary Psychology and Primate Cognition. In: *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition* (Ed. by Bekoff, M, Allen, C & Burghardt, GM), pp. 393-398. Cambridge, Massachusetts: MIT Press.
- Call, J** (2003). Beyond learning fixed rules and social cues: abstraction in the social arena. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 358, 1189-1196.
- Call, J & Carpenter, M** (2001). Do apes and children know what they have seen? *Animal Cognition*, 4, 207-220.

- Call, J & Tomasello, M** (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12, 187-192.
- Cheney, D & Seyfarth, RM** (1992). *How Monkeys See the World* Chicago: University of Chicago Press.
- Christiansen, MH & Kirby, S** (2003). Language evolution: consensus and controversies. *Trends in Cognitive Sciences* 7, 300-308.
- Clayton, NS** (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, 395, 272-274.
- Clayton, NS, Bussey, TJ & Dickinson, A** (2003a). Can animals recall the past and plan for the future? *Nature Reviews Neuroscience*, 4, 685-691.
- Clayton, NS, Dally, JM & Emery, NJ** (2007). Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist. *Philosophical Transactions of the Royal Society B*, 362, 507-522.
- Clayton, NS & Dickinson, A** (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, 395, 272-274.
- Clayton, NS & Dickinson, A** (1999a). Memory for the content of caches by scrub jays (*Aphelocoma coerulescens*). *Journal of Experimental Psychology: Animal Behaviour Process*, 25, 82-91.
- Clayton, NS & Dickinson, A** (1999b). Motivational control of caching behaviour in the scrub jay, *Aphelocoma coerulescens*. *Animal Behaviour* 57, 435-444.
- Clayton, NS & Dickinson, A** (1999c). Scrub jays (*Aphelocoma coerulescens*) remember the relative time of caching as well as the location and content of their caches. *Journal of Comparative Psychology*, 113, 403-416.
- Clayton, NS & Emery, NJ** (2009). What Do Jays Know About Other Minds and Other Times? In: *Neurobiology of "Umwelt": How Living Beings Perceive the World, Research and Perspectives in Neurosciences* (Ed. by Berthoz, A & Christen, Y), pp. 109-125. Berlin Heidelberg: Springer-Verlag
- Clayton, NS, Griffiths, DP, Emery, NJ & Dickinson, A** (2001a). Elements of episodic-like memory in animals. *Philosophical Transactions of the Royal Society B*, 356, 1483-1491.
- Clayton, NS, Russell, J & Dickinson, A** (2009a). Are Animals Stuck in Time or Are They Chronesthetic Creatures? *Topics in Cognitive Science*, 1, 59-71.
- Clayton, NS, Salwiczek, LH & Dickinson, A** (2009b). Episodic memory. *Current Biology*, 17, R189-R191.
- Clayton, NS, Yu, KS & Dickinson, A** (2001b). Scrub jays (*Aphelocoma coerulescens*) form integrated memories of the multiple features of caching episodes. *Journal of Experimental Psychology: Animal Behaviour Processes*, 27, 17-29.
- Clayton, NS, Yu, KS & Dickinson, A** (2003b). Interacting Cache Memories: Evidence for Flexible Memory Use by Western Scrub-Jays (*Aphelocoma californica*). *Journal of Experimental Psychology: Animal Behavior Processes*, 29, 14-22.
- Cohen, N & Squire, L** (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science*, 210, 207-210.

- Conway, MA** (2009). Episodic memories *Neuropsychologia*, 47, 2305-2313.
- Correia, SrPC, Dickinson, A & Clayton, NS** (2007). Western Scrub-Jays Anticipate Future Needs Independently of Their Current Motivational State. *Current Biology*, 17, 856-861.
- Crystal, JD** (2009). Elements of episodic-like memory in animal models. *Behavioural Processes*, 80, 269-277.
- Crystal, JD** (2010). Episodic-like memory in animals. *Behavioural Brain Research*, 215, 235-243.
- D'Argembeau, A & Linden, MVd** (2004). Phenomenal Characteristics Associated with Projecting Oneself Back into the Past and Forward into the Future: Influence of Valence and Temporal Distance. *Consciousness and Cognition*, 13, 844-858.
- Dally, JM, Clayton, NS & Emery, NJ** (2006). The behaviour and evolution of cache protection and pilferage. *Animal Behaviour*, 72, 13-23.
- Dekleva, M, Dufour, V, Vries, Hd, Spruijt, BM & Sterck, EHM** (2011). Chimpanzees (*Pan troglodytes*) Fail a What-Where-When Task but Find Rewards by Using a Location-Based Association Strategy. *PLoS ONE*, 6, e16593.
- deKort, SR, Dickinson, A & Clayton, NS** (2005). Retrospective cognition by food-caching western scrub-jays. *Learning and Motivation* 36, 159-176.
- Dere, E, Armin Zlomuzica, Huston, JP & Silva, MADS** (2008). Animal episodic memory. In: *Handbook of Episodic Memory* (Ed. by Dere, E, Easton, A, Nadel, L & Huston, JP), pp. 155-184. Amsterdam: Elsevier.
- Dere, E, Huston, JP & Silva, MADS** (2005a). Episodic-like memory in mice: Simultaneous assessment of object, place and temporal order memory. *Brain Research Protocols*, 16, 10-19.
- Dere, E, Huston, JP & Silva, MADS** (2005b). Integrated memory for objects, places, and temporal order: Evidence for episodic-like memory in mice. *Neurobiology of Learning and Memory*, 84, 214-221.
- Dere, E, Kart-Teke, E, Huston, JP & Silva, MADS** (2006). The case for episodic memory in animals. *Neuroscience and Biobehavioral Reviews*, 30, 1206-1224.
- Dufour, V, Pele', M, Sterck, EHM & Thierry, B** (2007). Chimpanzee (*Pan troglodytes*) Anticipation of Food Return: Coping With Waiting Time in an Exchange Task. *Journal of Comparative Psychology*, 21, 145-155.
- Dufour, V & Sterck, EHM** (2008). Chimpanzees fail to plan in an exchange task but succeed in a tool-using procedure. *Behavioural Processes*, 79, 19-27.
- Eacott, MJ & Easton, A** (2010). Episodic memory in animals: Remembering which occasion. *Neuropsychologia*, 48, 2273-2280.
- Eacott, MJ, Easton, A & Zinkivskay, A** (2005). Recollection in an episodic-like memory task in the rat. *Learning & Memory*, 12, 221-223.
- Emery, NJ & Clayton, NS** (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature*, 414, 443-349.
- Ergorul, C & Eichenbaum, H** (2004). The Hippocampus and Memory for "What," "Where," and "When". *Learning & Memory*, 11, 397-405.

- Feeney, MC, Roberts, WA & Sherry, DF** (2009). Memory for what, where, and when in the black-capped chickadee (*Poecile atricapillus*). *Animal Cognition* 12, 767-777.
- Ferkin, MH, Combs, A, delBarco-Trillo, J, Pierce, AA & Franklin, S** (2007). Meadow voles, *Microtus pennsylvanicus*, have the capacity to recall the “what”, “where”, and “when” of a single past event. *Animal cognition*, 11, 147-159.
- Fivush, R & Hamond, NR** (1990). Autobiographical memory across the preschool years: Towards reconceptualizing childhood amnesia. In: *Knowing and remembering in young children* (Ed. by Fivush, R & Hudson, JA). New York: Cambridge University Press.
- Friedman, WJ** (2003). Arrows of time in early childhood. *Child Development*, 74, 155-167.
- Friedman, WJ** (2005). Developmental and cognitive perspectives on humans’ sense of the times of past and future events. *Learning and Motivation*, 36, 145-158.
- Friedman, WJ** (2007). The Development of Temporal Metamemory. *Child Development*, 78, 1472-1491.
- Friedman, WJ & Lyon, TD** (2005). Development of Temporal-Reconstructive Abilities. *Child Development*, 76, 1202-1216.
- Gardiner, JM, Ramponi, C & Richardson-Klavehn, A** (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory*, 10, 83-98.
- Gibeault, S & Macdonald, SE** (2000). Spatial Memory and Foraging Competition in Captive Western Lowland Gorillas (*Gorilla gorilla gorilla*). *Primates*, 41, 147-160.
- Gilbert, DT & Wilson, TD** (2007). Propection: Experiencing the Future. *Science*, 317, 1351-1356.
- Grant, J & Suddendorf, T** (2010). Young children's ability to distinguish past and future changes in physical and mental states. *British Journal of developmental psychology* 28, 853-870.
- Griffiths, DP & Clayton, NS** (2001). Testing episodic memory in animals: A new approach. *Physiology & Behavior*, 73, 755– 762.
- Gumert, MD, Kluck, M & Malaijijitnond, S** (2009). The Physical Characteristics and Usage Patterns of Stone Axe and Pounding Hammers Used by Long-Tailed Macaques in the Andaman Sea Region of Thailand. *American Journal of Primatology*, 71, 594-608.
- Hampton, RP** (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Science*, 98, 5359-5362.
- Hampton, RR, Hampstead, BM & Murray, EA** (2005). Rhesus monkeys (*Macaca mulatta*) demonstrate robust memory for what and where, but not when, in an open-Weld test of memory. *Learning and Motivation*, 36, 245-259.
- Hampton, RR & Schwartz, BL** (2004). Episodic memory in nonhumans: what, and where, is when? *Current Opinion in Neurobiology*, 14, 192-197.

- Harst, JEvd, Fermont, PCJ, Bilstra, AE & Spruijt, BM** (2003). Access to enriched housing is rewarding to rats as reflected by their anticipatory behaviour *Animal Behaviour* 66, 493-504.
- Hayne, H & Imuta, K** (2011). Episodic Memory in 3- and 4-Year-Old Children. *Developmental Psychobiology*, 53, 317-322.
- Heyes, CM** (1993). Anecdotes, training, trapping and triangulating: do animals attribute mental states. *Animal Behaviour*, 46, 177-188.
- Hihara, S, Obayashi, S, Tanaka, M & Iriki, A** (2003). Rapid learning of sequential tool use by macaque monkeys. *Physiology & Behavior*, 78, 427-434.
- Hoffman, ML, Beran, MJ & Washburn, DA** (2009). Memory for “What”, “Where”, and “When” Information in Rhesus Monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, 35, 143-152.
- Holland, SM & Smulders, TV** (2011). Do humans use episodic memory to solve a What-Where-When memory task? *Animal Cognition*, 14, 95-102.
- Hosmer, D & Lemeshow, S** (2000). *Applied Logistic Regression*. New York: Wiley.
- Humphrey, NK** (1976). The social function of intellect. In: *Growing points in ethology* (Ed. by Bateson, PPG & Hinde, RA), pp. 303-317. Cambridge, England: Cambridge University Press.
- Ishibashi, H, Hihara, S & Iriki, A** (2000). Acquisition and development of monkey tool-use: Behavioral and kinematic analyses. *Canadian journal of physiology and pharmacology*, 78, 958-966.
- Jacoby, LL** (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513-541.
- Janmaat, KRL, Byrne, RW & Zuberbuhler, K** (2006). Evidence for a spatial memory of fruiting states of rainforest trees in wild mangabeys. *Animal Behaviour*, 72, 797-807.
- Kart-Teke, E, Silva, MADS, Huston, JP & Dere, E** (2006). Wistar rats show episodic-like memory for unique experiences. *Neurobiology of Learning and Memory*, 85, 173-182.
- Kinsbourne, M** (2005). A continuum of self-consciousness that emerges in phylogeny and ontogeny. In: *The Missing Link in Cognition: Origins of Self-Reflective Consciousness* (Ed. by Terrace, H & Metcalfe, J), pp. 142-156. Oxford: Oxford University Press.
- Klein, SB, Loftus, J & Kihlstrom, JF** (2002). Memory and Temporal Experience: the Effects of Episodic Memory Loss on an Amnesic Patient's Ability to Remember the Past and Imagine the Future. *Social Cognition*, 20, 353-379.
- Krachun, C, Carpenter, M, Call, J & Tomasello, M** (2009). A competitive nonverbal false belief task for children and apes. *Developmental Science*, 12, 521-535.
- Loftus, E** (1997). Creating False Memories. *Scientific American*, 277, 70-75.
- Loftus, EF & Pickrell, JE** (1995). The formation of false memories *Psychiatric Annals*, 25, 720-725.

- MacDonald, SE & Agnes, MM** (1999). Orangutan (*Pongo pygmaeus abelii*) spatial memory and behavior in a foraging task. *Journal of Comparative Psychology*, 113, 213-217.
- Malaivijitnond, S, Lekprayoon, C, Tandavanitt, N, Panha, S, Cheewatham, C & Hamada, AY** (2007). Stone-Tool Usage by Thai Long-Tailed Macaques (*Macaca fascicularis*). *American Journal of Primatology*, 69, 227-233.
- Martin-Ordas, G, Haun, D, Colmenares, F & Call, J** (2010). Keeping track of time: evidence for episodic-like memory in great apes. *Animal Cognition*, 13, 331-340.
- McCormack, T & Hoerl, C** (1999). Memory and temporal perspective: The role of temporal frameworks in memory development. *Developmental Review*, 19, 154-182.
- McCormack, T & Hoerl, C** (2005). Children's Reasoning About the Causal Significance of the Temporal Order of Events. *Developmental Psychology*, 41, 54-63.
- McCormack, T & Hoerl, C** (2007). Young children's reasoning about the order of past events. *Journal of Experimental Child Psychology*, 98, 168-183.
- McKenzie, TLB, Bird, LR & Roberts, WA** (2005). The effects of cache modification on food caching and retrieval behavior by rats *Learning and Motivation*, 36, 260-278.
- Melinder, A, Endestad, T & Magnussen, S** (2006). Relations between episodic memory, suggestibility, theory of mind, and cognitive inhibition in the preschool child. *Scandinavian Journal of Psychology*, 47, 485-495.
- Menzel, C** (2005). Progress in the Study of Chimpanzee Recall and Episodic Memory. In: *The missing link in cognition: Origins of self-reflective consciousness* (Ed. by Terrace, HS & Metcalfe, J), pp. 188-224. New York: Oxford University Press.
- Menzel, CR** (1999). Unprompted recall and reporting of hidden objects by a chimpanzee (*Pan troglodytes*) after extended delays. *Journal of Comparative psychology* 113, 426-434.
- Menzel, EW** (1973). Chimpanzee Spatial Memory Organization. *Science, New Series*, 182, 943-946.
- Mercado, E, Murray, SO, Uyeyama, RK, Pack, AA & Herman, LM** (1998). Memory for recent actions in the bottlenosed dolphin (*Tursiops truncatus*): Repetition of arbitrary behaviors using an abstract rule. *Animal Learning & Behavior*, 26, 210-218.
- Mercado, E, Uyeyama, RK, Pack, AA & Herman, LM** (1999). Memory for action events in the bottlenosed dolphin. *Animal Cognition*, 2, 17-25.
- Miyashita, Y** (2004). Cognitive Memory: Cellular and Network Machineries and Their Top-Down Control. *Science* 306, 435-441.
- Mulcahy, NJ & Call, J** (2006). Apes Save Tools for Future Use. *Science*, 312, 1038-1040.
- Naito, M** (2003). The relationship between theory of mind and episodic memory: Evidence for the development of auto-noetic consciousness. *Journal of Experimental Child Psychology*, 85, 312-336.

- Naqshbandi, M & Roberts, WA** (2006). Anticipation of Future Events in Squirrel Monkeys (*Saimiri sciureus*) and Rats (*Rattus norvegicus*): Tests of the Bischof-Köhler Hypothesis. *Journal of Comparative Psychology* 120, 345-357.
- Nelson, CA** (1995). The ontogeny of human memory: A cognitive neuroscience perspective. *Developmental Psychology*, 31, 723-738.
- Nelson, K** (1992). Emergence of autobiographical memory at age 4. *Human Development*, 35, 172-177.
- Nelson, K & Fivush, R** (2000). Socialization of memory. In: *The Oxford handbook of memory* (Ed. by Tulving, E & Craik, FIM), pp. 283-295. New York: Oxford University Press.
- Nelson, K & Fivush, R** (2004). The Emergence of Autobiographical Memory: A Social Cultural Developmental Theory. *Psychological Review* 111, 486-511.
- Normand, E, SDB & Boesch, C** (2009). Forest chimpanzees (*Pan troglodytes verus*) remember the location of numerous fruit trees. *Animal Cognition*, 12, 797-807.
- Normand, E & Boesch, C** (2009). Sophisticated Euclidean maps in forest chimpanzees. *Animal Behaviour*, 77, 1195-1201.
- O'Neill, DK, Astington, JW & Flavell, JH** (1992). Young Children's Understanding of the Role That Sensory Experiences Play in Knowledge Acquisition. *Child Development*, 63, 474-490.
- Okudaa, J, Fujii, T, Ohtake, H, Tsukiura, T, Tanji, K, Suzuki, K, Kawashima, R, Fukuda, H, Itoh, M & Yamadori, A** (2003). Thinking of the future and past: the roles of the frontal pole and the medial temporal lobes *NeuroImage*, 19, 1369-1380.
- Osvath, M** (2009). Spontaneous planning for future stone throwing by a male chimpanzee. *Current Biology*, 19, R190.
- Osvath, M & Osvath, H** (2008). Chimpanzee (*Pan troglodytes*) and orangutan (*Pongo abelii*) forethought: self-control and pre-experience in the face of future tool use. *Animal Cognition*, 11, 661-671.
- Paxton, R & Hampton, RR** (2009). Tests of planning and the Bischof-Köhler hypothesis in rhesus monkeys (*Macaca mulatta*). *Behavioural Processes*, 80, 238-246.
- Paz-y-Miño, CG, Bond, AB, Kamil, AC & Balda, RP** (2004). Pinyon jays use transitive inference to predict social dominance. *Nature*, 430, 778-781.
- Pelé, M, Dufour, V, Micheletta, J & Thierry, B** (2010). Long-tailed macaques display unexpected waiting abilities in exchange tasks. *Animal Cognition*, 13, 263-271.
- Perner, J** (2000). Memory and theory of mind. In: *The Oxford Handbook of Memory* (Ed. by Tulving, E & Craik, FIM), pp. 297-312. New York: Oxford University Press.
- Perner, J, Kloo, D & Gornik, E** (2007). Episodic Memory Development: Theory of Mind is Part of Re-experiencing Experienced Events. *Infant and Child Development*, 16, 471-490.
- Perner, J & Lang, B** (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, 3, 337-345.

- Perner, J, Leekam, S & Wimmer, H** (1987). 2-Year-olds difficulty with false belief- The case for a conceptual deficit. *British Journal of developmental psychology*, 5, 125-137.
- Perner, J & Ruffman, T** (1995). Episodic memory and Autoneotic Consciousness: Developmental Evidence and a Theory of Childhood Amnesia. *Journal of Experimental Child Psychology*, 59, 516-548.
- Povinelli, D, Landry, A, Theall, L, Clark, B & Castille, C** (1999). Development of young children's understanding that the recent past is causally bound to the present. *Developmental Psychology*, 35, 1426-1439.
- Premack, D** (2007). Human and animal cognition: Continuity and discontinuity. *Proceedings of the National Academy of Sciences* 104, 13861-13867.
- Quallo, MM, Price, CJ, Ueno, K, Asamizuya, T, Cheng, K, Lemon, RN & Iriki, A** (2009). Gray and white matter changes associated with tool-use learning in macaque monkeys. *Proceedings of the National Academy of Science USA*, 106, 18379-18384.
- Raby, C, Alexis, D, Dickinson, A & Clayton, N** (2007a). Empirical evaluation of mental time travel. *Behavioral and Brain Sciences*, 30, 330-331.
- Raby, CR, Alexis, DM, Dickinson, A & Clayton, NS** (2007b). Planning for the future by western scrub-jays. *Nature*, 445, 919-921.
- Raby, CR & Clayton, NS** (2009). Prospective cognition in animals. *Behavioural Processes* 80, 314-324.
- Ranganath, C, Yonelinas, AP, Cohen, MX, Dy, CJ, Tom, SM & D'Esposito, M** (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes *Neuropsychologia*, 42, 2-13.
- Reader, SM & Laland, KN** (2001). Primate innovation: sex, age and social rank differences. *International Journal of Primatology*, 22, 787-805.
- Roberts, S** (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, 7, 242-268.
- Roberts, WA** (2002). Are animals stuck in time? *Psychological Bulletin*, 128, 473-489.
- Roberts, WA, Cheng, K & Cohen, JS** (1989). Timing Light and Tone Signals in Pigeons *Journal of Experimental Psychology: Animal Behavior Processes*, 15, 23-35.
- Roberts, WA & Feeney, MC** (2009). The comparative study of mental time travel. *Trends in Cognitive Sciences*, 13, 271-278.
- Roberts, WA, Feeney, MC, MacPherson, K, Petter, M, McMillan, N & Musolino, E** (2008). Episodic-Like Memory in Rats: Is It Based on When or How Long Ago? *Science*, 320, 113-115.
- Roediger, HL & McDermott, KB** (1995). Creating false memories remembering words not presented in lists. *Journal of Experimental Psychology: Learning Memory and Cognition*, 21, 803-14.
- Russell, J, Alexis, D & Clayton, N** (2010). Episodic future thinking in 3- to 5-year-old children: The ability to think of what will be needed from a different point of view. *Cognition*, 114, 56-71.

- Salwiczek, LH, Watanabe, A & Clayton, NS** (2010). Ten years of research into avian models of episodic-like memory and its implications for developmental and comparative cognition. *Behavioural Brain Research*, 215, 221-234.
- Schacter, DL & Addis, DR** (2007). Constructive memory: The ghosts of past and future. *Nature reviews*, 445, 27.
- Schacter, DL, Addis, DR & Buckner, RL** (2007). Remembering the past to imagine the future: the prospective brain. *Nature reviews*, 8, 657-662.
- Schacter, DL & Tulving, E** (1994). What are the memory systems of 1994? In: *Memory systems 1994* (Ed. by Schacter, DL & Tulving, E), pp. 1-38. Cambridge, MA: MIT Press.
- Schwartz, BL, Colon, MR, Sanchez, IC, Rodriguez, IA & Evans, S** (2002). Single-trial learning of “what” and “who” information in a gorilla (*Gorilla gorilla gorilla*): implications for episodic memory. *Animal Cognition*, 5, 85-90.
- Schwartz, BL & Evans, S** (2001). Episodic Memory in Primates. *American Journal of Primatology*, 55, 71-85.
- Schwartz, BL, HoVman, ML & Evans, S** (2005). Episodic-like memory in a gorilla: A review and new findings. *Learning and Motivation*, 36, 226-244.
- Schwartz, BL, Meissner, CA, Hoffman, M, Evans, S & Frazier, LD** (2004). Event memory and misinformation effects in a gorilla (*Gorilla gorilla gorilla*). *Animal Cognition*, 7, 93-100.
- Seed, AM & Call, J** (2010). Problem-Solving in Tool-Using and Non-Tool-Using Animals. In: *Encyclopedia of Animal Behavior* (Ed. by Breed, MD & Moore, J): Academic Press.
- Shettleworth, SJ** (1998). *Cognition, Evolution and Behavior*. Oxford University Press.
- Shettleworth, SJ** (2007a). Planning for breakfast. *Nature*, 445, 825-826.
- Shettleworth, SJ** (2007b). Studying mental states is not a research program for comparative cognition. *Behavioral and Brain Sciences*, 30, 332-333.
- Squire, LR** (1992). Memory and the Hippocampus - a Synthesis from Findings with Rats, Monkeys, and Humans. *Psychological Review*, 99, 195-231.
- Squire, LR** (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory* 82, 171-177.
- Stadler, MA, Roediger, HL & McDermott, KB** (1999). Norms for word lists that create false memories. *Memory and Cognition*, 27, 494-500.
- Sterck, EHM & Begeer, S** (2010). Theory of Mind: Specialized capacity or emergent property? *European Journal of Developmental Psychology*, 7, 1-16.
- Sterck, EHM & Dufour, V** (2007). First test, then judge future-oriented behaviour in animals. *Behavioral and brain sciences*, 30, 299-351.
- Suddendorf, T** (2006). Foresight and Evolution of the Human Mind. *Science*, 312, 1006-1007.
- Suddendorf, T, Addis, DR & Corballis, MC** (2009a). Mental time travel and the shaping of the human mind. *Philosophical Transactions of the Royal Society B*, 364, 1317-1324.

- Suddendorf, T & Busby, J** (2003). Mental time travel in animals? *Trends in Cognitive Sciences*, 7, 391-396.
- Suddendorf, T & Busby, J** (2005). Making decisions with the future in mind: Developmental and comparative identification of mental time travel. *Learning and Motivation* 36, 110-125.
- Suddendorf, T & Corballis, MC** (1997). Mental time travel and the evolution of the human mind. *Genetic, Social & General Psychology Monographs*, 123, 133-168.
- Suddendorf, T & Corballis, MC** (2007a). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and brain sciences* 30, 299-351.
- Suddendorf, T & Corballis, MC** (2007b). Mental time travel across the disciplines: The future looks bright. *Behavioural and Brain Sciences*, 30, 335-351.
- Suddendorf, T & Corballis, MC** (2008). New evidence for animal foresight? *Animal Behaviour*, 75, e1-e3.
- Suddendorf, T & Corballis, MC** (2010). Behavioural evidence for mental time travel in nonhuman animals. *Behavioural Brain Research* 215, 292-298.
- Suddendorf, T, Corballis, MC & Collier-Baker, E** (2009b). How great is great ape foresight? *Animal Cognition*, 12, 751-754.
- Suddendorf, T, Nielsen, M & Gehlen, Rv** (2011). Children's capacity to remember a novel problem and to secure its future solution. *Developmental Science*, 14, 26-33.
- Szpunar, KK & McDermott, KB** (2008). Episodic future thought and its relation to remembering: Evidence from ratings of subjective experience *Consciousness and Cognition*, 17, 330-334.
- Tobin, H, Logue, AW, Chelonis, JJ & Ackerman, KT** (1996). Self-control in the monkey *Macaca fascicularis*. *Animal Learning and Behavior*, 24, 168-174.
- Tokida, E, Tanaka, I, Takefushi, H & Hagiwara, T** (1994). Tool-using in Japanese macaques: use of stones to obtain fruit from a pipe. *Animal Behaviour*, 47, 1023-1030.
- Tomasello, M & Call, J** (1997). *Primate Cognition*. New York: Oxford University Press.
- Tulving, E** (1972). *Episodic and semantic memory*. New York: Academic Press.
- Tulving, E** (1983). *Elements of episodic memory*. Oxford: Clarendon Press.
- Tulving, E** (1984). Precis of Elements of episodic memory. *The behavioral and brain sciences*, 7, 223-268.
- Tulving, E** (1985). Memory and consciousness. *Canadian Psychology*, 26, 1-12.
- Tulving, E** (2001). Episodic Memory and Common Sense: How Far Apart? *Philosophical Transactions: Biological Sciences*, 356, 1505-1515.
- Tulving, E** (2002). Episodic memory: from mind to brain. *Annual Reviews in Psychology* 53, 1-25.
- Tulving, E** (2005). Episodic memory and auto-noesis: Uniquely human? . In: *The Missing Link in Cognition* (Ed. by Terrace, HS & Metcalfe, J), pp. 4-56. New York: Oxford University Press.

- Tulving, E & Markowitsch, HJ** (1998). Episodic and Declarative Memory: Role of the Hippocampus. *Hippocampus*, 8, 198-204.
- Vargha-Khadem, F, Gadian, DG & Mishkin, M** (2001). Dissociations in cognitive memory: the syndrome of developmental amnesia. *Philosophical Transactions of the Royal Society of London B*, 356, 1435-1440.
- Visalberghi, E, Janson, CH & Agostini, I** (2003). Response towards novel foods and novel objects in wild *Cebus appella*. *International Journal of Primatology*, 24, 653-675.
- Visalberghi, E & Tomasello, M** (1998). Primate causal understanding in the physical and psychological domains. *Behavioural Processes*, 42, 189-203.
- Vries, Hd** (1993). The rowwise correlation between two proximity matrices and the partial rowwise correlation. *Psychometrika*, 58, 53-69.
- Vries, Hd, Netto, W & Hanegraaf, P** (1993). Matman: a program for the analysis of sociometric matrices and behavioural transition matrices. *Behaviour*, 125, 157-175.
- Watanabe, K, Urasopon, N & Malaivijitnond, S** (2007). Long-Tailed Macaques Use Human Hair as Dental Floss. *American Journal of Primatology* 69, 940-944.
- Wheeler, MA** (2000). Episodic memory and auto-noetic awareness. In: *The Oxford handbook of memory* (Ed. by Tulving, E & Craik, FIM), pp. 597-608. New York: Oxford University Press.
- Wheeler, MA, Stuss, DT & Tulving, E** (1997). Toward a Theory of Episodic memory: The Frontal Lobes and Auto-noetic Consciousness. *Psychological Bulletin*, 121, 331-354.
- Wimmer, H, Hogrefe, GJ & Perner, J** (1988). Children's Understanding of Informational Access as Source of Knowledge. *Child Development*, 59, 386-396.
- Wimmer, H & Perner, J** (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.
- Yonelinas, AP** (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, 46, 441-517.
- Zentall, TR** (2005). Animals may not be stuck in time. *Learning and Motivation*, 36, 208-225.
- Zentall, TR** (2006). Mental time travel in animals: A challenging question. *Behavioural Processes* 72, 173-183.
- Zentall, TR, Clement, TS, Bhatt, RS & Allen, J** (2001). Episodic-like memory in pigeons. *Psychonomic Bulletin & Review*, 8, 685-690.
- Zhou, W & Crystal, JD** (2009). Evidence for remembering when events occurred in a rodent model of episodic memory. *Proceedings of the National Academy of Sciences* 106, 9525-9529.
- Zinkivskay, A, Nazir, F & Smulders, TV** (2009). What-Where-When memory in magpies (*Pica pica*). *Animal Cognition*, 12, 119-125.

