

**Die Entwicklung der Sensitivität für das Wirken von Schwerkraft auf
die Bewegung unbelebter Objekte: Ein komparativer Zugang**

Abhandlung zur Erlangung der Doktorwürde der Philosophischen Fakultät der
Universität Zürich

vorgelegt von Trix Cacchione

von Luzern (LU)

Angenommen im Frühjahrssemester 2007 auf Antrag von

Herrn Prof. Dr. Friedrich Wilkening

Herrn Prof. Dr. Horst Krist

Zürich, 2008

Zusammenfassung

Sowohl Sensitivität für Schwerkraft als auch für Objektsolidität hat das Verhalten der nicht-menschlichen Primaten in den vorliegenden Studien beeinflusst, kam aber nicht in allen Kontexten konsistent zum Tragen. Der Ausdruck potentiellen Wissens war also von einer Vielzahl von Faktoren abhängig und damit – wie dies in der Literatur vielfach belegt ist – diskontinuierlich, inkonsistent und kontextabhängig. Menschenaffen zeigten weder bei der „Tisch-Aufgabe“ noch bei der „Röhrenaufgabe“ einen signifikanten Schwerkraftfehler. Trotzdem liess ihr Suchverhalten auf das Wirken naiver Schwerkraftkonzepte schliessen. Objektsolidität wurde nur bei der „Tisch-Aufgabe“ manifest, während es im Falle der „Röhrenaufgabe“ im Allgemeinen vernachlässigt wurde. Weiter konnte bei Weissbüscheläffchen eine Dissoziation zwischen Handlungs- und Wahrnehmungswissen im Rahmen der „Röhrenaufgabe“ nachgewiesen werden. Die Weissbüscheläffchen reagierten in ihrem Blickverhalten auf Verletzungen der Objektsolidität. Zudem neigten sie weniger zum Schwerkraftfehler als die bisher getesteten Neuweltaffenspezies. Insgesamt widersprechen diese Resultate der These, dass nur Menschen im Verlauf der Entwicklung die Fähigkeit zur Inhibition des Schwerkraftfehlers erwerben.

Summary

Sensitivity to gravity and solidity influenced performance of the non-human primates in the current studies, even though they did not apply their sensitivity in all situations. Thus, the use of knowledge depended on various factors and was, as often observed in developmental psychology, discontinuous and context-dependent. Neither in the “table task” nor in the “tubes tasks” great apes showed a reliable gravity bias, yet their search behaviour was clearly guided by naïve gravity concepts. Object solidity was only applied in the “table task”, whereas it was generally neglected in case of the “tubes tasks”. Further, a dissociation between action and perception based knowledge was identified in common marmosets with the “tubes tasks”. Marmosets looked longer at events involving a violation of solidity. Also, they were less gravity biased as prior tested New World monkey species. Overall, the present results challenge the thesis that only humans develop sufficient inhibitory control to suppress gravity errors.

| | |
|--|----|
| 1. Einleitung | 6 |
| 1.1 Die Entwicklung des Wissens über Schwerkraft bei Kindern | 7 |
| 1.2 Untersuchungen mit nicht-menschlichen Primaten..... | 10 |
| 1.3 Ziele der vorliegenden Studien | 12 |
| 1.4 Literaturangaben..... | 13 |
| 2. Gravity bias in perception and action | 19 |
| 2.1 Abstract | 19 |
| 2.2 Introduction | 20 |
| 2.2.1 Gravity-biased search..... | 20 |
| 2.2.2 Dissociation between perception and action | 21 |
| 2.3 Experiment 1 | 23 |
| 2.3.1 Method | 23 |
| 2.3.2 Results | 27 |
| 2.3.3 Discussion | 30 |
| 2.4 Experiment 2 | 31 |
| 2.4.1 Method | 32 |
| 2.4.2 Results | 34 |
| 2.4.3 Discussion | 35 |
| 2.5 General discussion..... | 35 |
| 2.6 References | 38 |
| 3. Gravity and solidity in great apes: The table task | 43 |
| 3.1 Abstract | 43 |
| 3.2 Introduction | 44 |
| 3.3 General methods..... | 52 |
| <i>Participants</i> | 52 |
| <i>Materials</i> | 54 |
| <i>Design</i> | 54 |
| <i>General Procedure</i> | 56 |
| <i>General Analyses</i> | 57 |
| 3.4 Experiment 1 | 58 |
| 3.4.1 Method | 58 |
| 3.4.2 Results | 60 |
| 3.4.3 Discussion | 63 |
| 3.5 Experiment 2 | 65 |
| 3.5.1 Method | 66 |
| 3.5.2 Results | 68 |
| 3.5.3 Discussion | 70 |
| 3.6 Experiment 3 | 70 |
| 3.6.1 Method | 71 |
| 3.6.2 Results | 72 |
| 3.6.3 Discussion | 75 |
| 3.7 General discussion..... | 75 |
| 3.8 References | 79 |

4. Do great apes refer to a tube as a causal device? Intuitions about gravity and solidity

| | |
|---------------------------------------|-----------|
| in four great ape species..... | 83 |
| 4.1 Abstract | 83 |
| 4.2 Introduction | 84 |
| 4.3 Experiment 1 | 88 |
| 4.3.1 Method | 88 |
| 4.3.2 Results | 91 |
| 4.3.3 Discussion | 92 |
| 4.4 Experiment 2 | 94 |
| 4.4.1 Method | 94 |
| 4.4.2 Results | 96 |
| 4.4.3 Discussion | 103 |
| 4.5 Experiment 3a | 105 |
| 4.5.1 Method | 105 |
| 4.5.2 Results | 106 |
| 4.5.3 Discussion | 106 |
| 4.6 Experiment 3b | 107 |
| 4.6.1 Method | 107 |
| 4.6.2 Results | 107 |
| 4.6.3 Discussion | 108 |
| 4.7 Experiment 4 | 108 |
| 4.7.1 Method | 109 |
| 4.7.2 Results | 109 |
| 4.7.3 Discussion | 110 |
| 4.8 Experiment 5 | 110 |
| 4.8.1 Method | 111 |
| 4.8.2 Results | 111 |
| 4.8.3 Discussion | 111 |
| 4.9 General discussion..... | 112 |
| 5. References | 114 |

1. Einleitung

Schwerkraft ist ein in unserem Alltag allgegenwärtiges physikalisches Prinzip, welches das Verhalten unbelebter als auch lebender Körper fundamental beeinflusst. Um das eigene Verhalten in einer von physikalischen Gesetzmässigkeiten regulierten Umwelt adaptativ zu organisieren, haben Menschen wie Tiere im Laufe der Evolution spezifische kognitive Mechanismen entwickelt. Beispielsweise enthalten perzeptive und motorische Systeme angeborene Strukturen, die als phylogenetisch erworbene Anpassungen an die Wirkung der Schwerkraft aufzufassen sind.

Auf der Ebene von Wissenssystemen hingegen zeichnen die heute vorliegenden Befunde ein ambivalentes Bild. Obwohl die Effekte der Schwerkraft konstant über den perzeptiven Input erfahrbar sind, erweist sich das menschliche Denken in dieser Hinsicht als erstaunlich inkonsistent und häufig bis ins Erwachsenenalter fehlerhaft. Es stellt sich in diesem Zusammenhang die Frage, wie Wissen über Schwerkraft von lebenden Organismen überhaupt erworben werden kann und gemäss welchen Kriterien es mental repräsentiert wird, um adaptatives Verhalten zu ermöglichen. Der momentane Forschungsstand zeigt, dass ein multikausales Modell aus angeborenen und über Erfahrung erworbenen Faktoren am wahrscheinlichsten ist. Wie aber greifen biologische, perzeptive und kognitive Mechanismen ineinander, um Wissensstrukturen zu bilden? Ist die traditionell zumindest im Erwachsenenalter unterstellte, weitgehende Kongruenz von physikalischer Gesetzmässigkeit und mentalem Alltagskonzept überhaupt angemessen? Diese Fragen bilden heute den Gegenstand einer intensiven Forschungstätigkeit, die sich hauptsächlich auf die ontogenetische Entwicklung besagter Wissenskonzepte beim Menschen konzentriert. Gerade im Hinblick auf den angenommenen Einfluss angeborener und erworbener Mechanismen drängt sich heute eine interdisziplinär abgestützte Ausdehnung des Forschungsfokus auf die phylogenetische Entwicklung auf: Naive Wissenskonzepte verschiedener lebender Organismen müssen geprüft und verglichen werden, um den unterschiedlichen Einfluss biologischer und psychologischer Faktoren isoliert betrachten zu können.

Die vorliegenden Studien adressieren die noch praktisch unerforschte phylogenetische Entwicklung intuitiver Objektkonzepte, namentlich das intuitive Wissen über das Verhalten bewegter physikalischer Objekte unter der Wirkung der Schwerkraft. Aus einer vergleichenden Perspektive wird das intuitive Denken über Schwerkraft bei vier verschiedenen Menschenspezies und einer Neuweltaffenspezies erhoben und mit bereits vorhandenen Daten aus dem Humanbereich, sowie den wenigen vorhandenen Daten von Neu-

und Altweltaffen, verglichen. Bei allen untersuchten Spezies kommen dieselben Methoden zum Einsatz, die auch zur Untersuchung von Menschenkindern sowie Neu- und Altweltaffen eingesetzt wurden, was einen direkten Vergleich ermöglicht. Die Studien zielen darauf ab, Gemeinsamkeiten und Unterschiede in den Wissensstrukturen besagter Spezies aufzudecken und zu beschreiben. Können Menschenaffen - wie Menschen und einzelne Neu- und Altweltaffen - die Effekte der Schwerkraft auf unbelebte Objekte repräsentieren, antizipieren und motorisch darauf reagieren? Bestimmen Intuitionen über das Wirken der Schwerkraft die Reaktionen dominant, so dass anderes Wissen über Merkmale und das Verhalten von Objekten (z. B. Objektsolidität) unterdrückt wird? Unterscheiden sich Menschenaffen diesbezüglich untereinander sowie von Menschen und Neu- und Altweltaffen? Wie können die gefundenen Gemeinsamkeiten und Unterschiede im Hinblick auf den Wissenserwerb unter dem Einfluss biologischer (angeborener) sowie perzeptiver und kognitiver (erworbener) Faktoren interpretiert werden?

1.1 Die Entwicklung des Wissens über Schwerkraft bei Kindern

Piaget (1975) entwickelte zur Untersuchung naiver physikalischer Konzepte bei jungen Kindern sog. „invisible displacement tasks“¹, die bis heute von der Mehrzahl der Studien verwendet werden, um physikalisches Wissen in der frühen Kindheit nachzuweisen. Die traditionelle Sichtweise geht davon aus, dass konzeptuelles Wissen erst nach Vollendung des ersten Lebensjahres vorhanden ist. Eine neue Forschergeneration kritisierte Piagets Verhaltensmass und machte geltend, dass mit sensibleren Methoden, wie Habituations- und Blickzeitenmassen, bereits im Alter ab 3 Monaten physikalisches Wissen nachgewiesen werden könne (z.B. Baillargeon, Spelke & Wasserman, 1985; Baillargeon, et al., 1995a; Spelke et al., 1992). Diese Befunde machen deutlich, dass Piagets Auffassung der kognitiven Entwicklung - als von einem Zustand frei von konzeptuellem Wissen bei Geburt, zu einem Zustand mit vollständigem konzeptuellem Wissen im Erwachsenenzustand führend - zu idealtypisch gezeichnet war. Diese Sicht der Dinge muss zugunsten eines eher inkohärenten, diskontinuierlichen Entwicklungsverlaufs, der sowohl frühe Kompetenzen als auch späte Inkompetenzen umfasst, korrigiert werden. Dies gilt ganz besonders für das Denken über Schwerkraft, das einerseits sehr robust und auf einer frühen Entwicklungsstufe erscheint, jedoch selbst für die meisten Erwachsenen nicht widerspruchsfrei anwendbar ist. Dies hat die

¹ Die Fähigkeit eines Organismus auf nicht direkt wahrnehmbare Objekte zu reagieren, gilt als Piagets Hauptkriterium zur Annahme mentaler Repräsentationen.

Diskussion um die Art der Konzeptorganisation angetrieben: Sind naive Konzepte überhaupt in analoger Weise wie wissenschaftliche Theorien organisiert (Carey, 1985; Howe, 1998; Mandler, 1998; Munakata, 2001; Thelen & Smith, 1994)?

Studien zu naivem Wissen über Schwerkraft bei Menschen befassten sich bisher mit zwei Problemtypen: Der Beurteilung der statisch relevanten Kontaktrelation (*engl.* support relation) zwischen zwei unbelebten Objekten (Baillargeon, Needham, & DeVos, 1992; Baillargeon, Kotovsky, & Needham, 1995a; Baillargeon, Raschke, & Needham, 1995b; Dan, Omori, & Tomiyasu, 2000; Huettel & Needham, 2000; Needham & Baillargeon, 1993a, 1993b; Spelke, Breinlinger, Macomber, & Jacobson, 1992) und der Beurteilung und der Antizipation der Bewegung unbelebter Objekte im Raum (Friedman, 2002; Hood, 1995, 98; Kaiser & Profitt, 1984; Kaiser, Profitt, & Anderson, 1985a; Kaiser, Profitt, & McCloskey, 1985b; Kaiser, McCloskey & Profitt, 1986a; Kim & Spelke, 1992, 99; Krist, 2000; McCloskey, 1983).

Nach Piaget reflektiert das Suchverhalten von 6-10 Monate alten Kindern zum ersten Mal das Wirken der Schwerkraft: Beobachteten sie wie ein Objekt fallen gelassen wurde, richteten sie ihre Aufmerksamkeit nicht der Hand - wo sie das Objekt zu letzt gesehen haben – sondern dem Boden zu, sie antizipierten also, dass sich das Objekt nach unten bewegt. Blickzeitenmasse belegen bereits bei 3 Monate alten Babys eine Sensitivität für Schwerkraft, wenn diese mit Abstützungs-Verhältnissen konfrontiert werden (Needham et al. 1993b; für eine Übersicht siehe Baillargeon, et al., 1995a; Baillargeon, 2002). Spelke et al. (1992) präsentierten Ereignisse, bei denen Objekte unsichtbar entlang der Vertikalen und Horizontalen verlagert wurden, dabei entweder ohne Kontakt zu anderen Objekten mitten im freien Fall zum Stillstand kamen oder sich scheinbar durch solide Hindernisse hindurch bewegten (diese Ereignisse wurden später von Autoren wie Hood und Hauser aufgegriffen und unter der Bezeichnung „table task“ als Handlungsaufgaben konzipiert). Spelke und Kollegen fanden keine Evidenz, dass 3 bis 4 Monate alte Babys erwarten, dass nicht abgestützte Objekte fallen, obwohl ihre Reaktionen in diesem Alter bereits auf eine Sensitivität für die Kontinuität und Solidität von Objekten schliessen liessen, da sie mit einer Verlängerung der Blickzeiten reagierten, wenn sich die Objekte durch solide Hindernisse bewegten. Die Reaktionen der Kinder änderten sich signifikant zwischen 3 und 6 Monaten, aber selbst im Alter von 6 Monaten waren die von Spelke und Kollegen gefundenen Schwerkraftreaktionen noch gering. Die nativistische Theoretikerin Spelke wertet diese Resultate als eine Bestätigung der von ihr postulierten „Core knowledge Hypothese“ und der Annahme, dass Schwerkraft im Gegensatz zu Solidität und Kontinuität nicht zu den

angeborenen Wissensbeständen gehört. Im Gegensatz zu Objektsolidität und –kontinuität sei Schwerkraft – obwohl in vielen Fällen salient beobachtbar - kein reliables Prinzip und deshalb nicht Teil der angeborenen Kernwissensbestände, die sowohl Wahrnehmung als auch Denken zu Grunde liegen. Tatsächlich belegen eine Vielzahl von Misskonzepten, dass Schwerkraft auch im erwachsenen Denken nicht zentral ist (z. B. Halloun & Hestenes, 1985; Howe, 1998; Kaiser, Jonides, & Alexander, 1986b; Krist, 2000; McCloskey, 1983; Shanon, 1976; White, 1988).

Welche Art von Wissen den mit Blickzeitenmessung erhobenen Reaktionen zugrunde liegt, ist heute heftig umstritten. Viele bezweifeln, dass es sich dabei um explizites konzeptuelles Wissen im Sinne Piagets handelt (z. B. Bogartz, Shinsky, & Shilling 2000; Cashon & Cohen, 2000; Haith, 1998). Tatsächlich können ältere Kinder viele von den bereits bei Babys diagnostizierten Wissensbeständen im Kontext von Handlungsaufgaben nicht anwenden. Es besteht also eine klare Dissoziation zwischen dem Wissen, welches sich mit Blickzeitenmassen diagnostizieren lässt und dem Wissen, welches sich mit Handlungsmassen nachweisen lässt. Interessanterweise werden Intuitionen über Schwerkraft in einem Handlungskontext früher manifest, als etwa Wissen über Objektsolidität, obwohl dies (wie oben ausgeführt) bei Blickzeitenmassen gerade in umgekehrter Reihenfolge feststellbar ist. Hood (1995) fand mit einer von ihm entwickelten „invisible displacement task“ (tubes task) nicht nur eine Dissoziation zwischen Wahrnehmen und Handeln, er identifizierte auch einen perseverativen Handlungsfehler, den sog. „gravity bias“, den er als Beleg für Wissen um Schwerkraft deutet. Er konfrontierte 2- und 3-Jährige Kinder mit einem Apparat, bei dem drei Einfüllstutzen mittels undurchsichtiger Röhren mit drei Zielbehältern verbunden waren. Die Kinder mussten angeben, in welchem Zielbehälter sich ein Objekt befindet, welches in eine der drei Röhren fallen gelassen wurde. Trotz massiver Gegenevidenz suchten Kinder im Alter von 2 bis 2.5 Jahren unsichtbar fallende Objekte überzufällig oft in Richtung der Schwerkraft, das heisst im Zielbehälter direkt unterhalb des entsprechenden Einfüllstutzens. Die Tendenz vertikal verlagerte Objekte in Schwerkraftrichtung zu suchen, scheint so stark, dass Wissen um Solidität - obwohl dies mit Blickzeitenmessung bereits ab 3 Monaten nachgewiesen wird – beim Lokalisieren der Objekte vollständig vernachlässigt wird. Hood geht davon aus, dass eine intuitive Theorie über das Wirken der Schwerkraft, wonach alle nicht abgestützten Objekte senkrecht nach unten fallen, für den Fehler verantwortlich ist. Tatsächlich fanden verschiedene Autoren auch bei Schulkindern und je nach Kontext sogar bei Erwachsenen Evidenz für einen „straight down belief“ (Bliss, Ogborn, & Whitelock, 1989; Eckstein & Shemesh, 1989; Kaiser, Proffitt, & McCloskey, 1985; Krist 2000; Marioni, 1989). Bliss,

Ogborn und Whitelock (1989) sprechen gar von einer "golden rule of everyday physics": "If an object is not supported, it falls until it is once more supported" (Bliss et al. 1989: 262). Hood nimmt an, dass der prepotenten Schwerkraftantwort ein Inhibitionsproblem zugrunde liegt: Diese intuitive Annahme ist so vorherrschend, dass sie von jüngeren Kindern selbst in inadäquaten Kontexten nicht unterdrückt werden kann. Für Hoods Postulat einer intuitiven Schwerkrafttheorie spricht, dass der Fehler nicht auftritt, wenn die Objektverlagerung von unten nach oben oder entlang der Horizontalen präsentiert wird (Hood, 1998; Hood, Santos, & Fieselman, 2000). Neuere Studien unterstützen auch die These, dass der Fehler durch die zunehmende Fähigkeit, unangebrachte Schwerkraftantworten zu unterdrücken, überwunden wird: Wird die Komplexität der Aufgabe angehoben und damit mehr kognitive Kapazität zur Lösung notwendig (wodurch Inhibitionsmechanismen behindert werden), zeigen auch 4 Jahre alte Kinder erneut den „gravity bias“ (Hood, Wilson, & Dyson, 2006). Allerdings kritisierte Hood selber, dass seine „tubes task“ nicht direkt mit den von Spelke im Rahmen der Habituationsstudie verwendeten Ereignissen vergleichbar ist. Er konstruierte daher in Anlehnung an die Habituationsstudie von Spelke et al. (1992) eine Handlungsaufgabe, die sog. „table task“ (Hood, Carey, & Prasadada; 2000). Dabei werden Kindern mit einem tischartigen Apparat konfrontiert, der hinter einem Schirm verborgen wird. Dann werden Objekte auf die Tischplatte fallen gelassen und geprüft, ob das Kind diese auf (im Einklang mit Wissen um Objektsolidität) oder unterhalb der Tischplatte (als Ausdruck des „gravity bias“) sucht. Bis zum Alter von 2;6 Jahren liess das Verhalten der Kinder weder Wissen um Objektsolidität noch Einsicht in das Abstützungsverhältnis zweier Objekte („support relation“) vermuten. Ebenso liess sich mit dieser Problemstellung kein „gravity bias“ nachweisen, vielmehr suchten die Kinder das Objekt dort, wo sie es während der Familiarisierung gesehen hatten. Es besteht also nicht nur eine Dissoziation zwischen dem Wissen, welches sich mit Handlungs- und Blickzeitenmassen nachweisen lässt. Auch im Kontext verschiedener Handlungsaufgaben lassen sich Kinder von unterschiedlichen Wissensbeständen leiten.

1.2 Untersuchungen mit nicht-menschlichen Primaten

In jüngster Zeit begannen einige Entwicklungspsychologen damit, kognitive Phänomene mit einem vergleichenden Ansatz zu untersuchen. Dabei kamen sowohl Blickzeitenmasse als auch Verhaltens- und Urteilsmasse zum Einsatz.

Povinelli (2000) beruft sich unter anderem auf die Pionierarbeit Köhlers (1957), der bereits vor hundert Jahren die These aufstellte, dass das statische Denken von Schimpansen - anders als beim Menschen - nicht in ein stabiles Koordinatensystem mit horizontaler und vertikaler Dimension eingebettet sei und dass sich Schimpansen ausschliesslich an perzeptiven Merkmalen (wie sichtbarem Kontakt zwischen Objekten) orientierten. Aufgrund von Studien mit Verhaltens- und Urteilsaufgaben geht Povinelli (2000) noch einen Schritt weiter und erklärt, dass sich Schimpansenkognition ausschliesslich auf wahrnehmbare Entitäten abstütze und frei von theoretisch-abstrakten Annahmen sei. Cacchione und Krist (2004) untersuchten, ob Menschenaffen wie menschliche Babys auf die Präsentation von gegen die Gesetze der Schwerkraft verstossenden Ereignissen mit einer Verlängerung der Blickdauer reagieren. Die Resultate zeigen, dass das Denken von Schimpansen und menschlichen Kindern im Alter von 6-7 Monaten Ähnlichkeiten aufweist. So reagieren auch Schimpansen deutlich auf die Existenz und die Menge von sichtbarem Kontakt zwischen zwei Objekten. Sie reagieren beispielsweise mit einer Verlängerung der Blickdauer bei der Präsentation frei „schwebender“ Objekte. Anders als Kinder scheinen sie dabei die Relevanz der räumlichen Orientierung nicht zu berücksichtigen. Sie behandeln beispielsweise vertikalen und horizontalen Kontakt zwischen Objekten als konzeptuell gleich. Dies wirft einmal mehr die Frage auf, ob mit der unterschiedlichen Sensibilität für die Bedeutung der Raumorientierung ein elementarer Unterschied zwischen menschlichem und nicht-menschlichem Denken formuliert worden ist.

Diese These erweist sich als problematisch, sobald man neuere Studien mit Verhaltens- und Urteilsaufgaben hinzu zieht. Verschiedene Studien fanden einen robusten „straight down belief“ bei nicht-menschlichen Primaten, was bedeutet, dass vertikale und horizontale Dimension im Kontext von Verhaltensaufgaben deutlich unterschieden werden (Hauser, 2001; Hauser, Williams, Kralik & Moskovitz, 2001; Hood, Hauser, Anderson & Santos, 1999). Hood, Hauser, Anderson und Santos (1999) testeten mit der für Kinder entwickelten „tubes task“ erwachsene Krallenaffen (*Saguinus oedipus oedipus*): Die Krallenaffen zeigten sogar eine noch hartnäckigere Tendenz, unsichtbar fallende Objekte in der Schwerkraftlinie zu suchen, als Kinder. In einer weiteren Studie prüfte Hauser, Williams, Kralik und Moskovitz, 2001 die Abhängigkeit des „gravity bias“ von der räumlichen Orientierung. Tatsächlich verschwindet der Fehler bei horizontaler Objektverlagerung völlig. Hauser (2001) fand, dass Rhesusaffen (*Macaca mulatta*) im Gegensatz zu Kindern den „gravity bias“ auch zeigen, wenn sie mit der „table task“ getestet werden. Wiederum verschwindet der Fehler, wenn schwerkraftrelevante Faktoren verändert werden: Wenn

potentielle Ziellokalitäten von der Vertikalen abweichen sowie wenn die Objektverlagerung entlang der Horizontalen geschieht. Wie Hood für Kinder postuliert, nimmt Hauser auch für nicht-menschliche Primaten an, dem Bias liege eine naive Schwerkrafttheorie zugrunde, die aufgrund von Inhibitionsproblemen auf zu viele Fälle angewendet wird (Hauser, 2003). Er geht davon aus, dass Rhesusaffen eigentlich Wissen über Objektsolidität haben, dieses aber im Handlungskontext nicht anwenden können, weil es durch die prepotente Schwerkraftantwort verdeckt wird. Tatsächlich gelang der Nachweis, dass auch bei Rhesusaffen eine Dissoziation zwischen Blickzeiten und Handlungsmassen besteht (Santos & Hauser, 2002; Santos, Seelig, & Hauser, 2006). Obwohl das Suchverhalten von Rhesusaffen keine Hinweise auf Soliditätswissen liefert, reagieren auch sie mit einer Verlängerung der Blickzeiten auf Soliditätsverletzung, wenn sie mit analogen Ereignissen konfrontiert werden, wie Spelke und Kollegen sie mit Kindern verwendet haben. Dies wertet Hauser als Beleg für die Inhibitionsthese und rückt damit den Schwerkraftfehler in einen ähnlichen Erklärungskontext wie den A-nicht-B-Fehler. Nach Diamond (1991) muss kognitive Entwicklung nicht nur als kontinuierlicher Wissenserwerb aufgefasst werden, sondern ebenso als die kontinuierliche Zunahme der Fähigkeit unerwünschte Reaktionen zu unterdrücken. Obwohl Kinder, die den A-nicht-B-Fehler machen, eigentlich wüssten, unter welcher Verdeckung sich das Objekt sich befindet, könnten sie die Greifbewegung an die falsche Position in einem frühen Entwicklungsstadium nicht unterdrücken. In analoger Weise stellt sich das Hauser für den „gravity bias“ vor. Im Unterschied zu Menschen jedoch, die den Bias im Verlauf der Entwicklung durch zunehmende Inhibitionskontrolle überwinden, bleibe der Schwerkraftfehler bei Affen über die gesamte Lebensspanne hinweg als prepotente Reaktion erhalten.

1.3 Ziele der vorliegenden Studien

Wie bei Menschen deuten die wenigen Befunde auch bei nicht-menschlichen Primaten die Existenz einer Dissoziation an, zwischen Wissen um Solidität und Schwerkraft welches mit Blickzeiten-, resp. mit Handlungsmassen gefunden werden kann. Beide reagieren mit Blickzeitenverlängerung bei Soliditätsverletzung, können dieses Wissen aber in einem Handlungskontext nicht anwenden, während Wissen um Schwerkraft vergleichsweise früh im Handlungskontext nachweisbar ist. Im Unterschied zu Kindern scheinen nicht-menschliche Primaten in ihrem Handeln noch ausgeprägter dem Schwerkraftfehler unterworfen. Er zeigt sich nicht nur robuster über die gesamte Lebensspanne hinweg, sondern auch in verschiedenen Handlungskontexten (table und tubes task). Bis heute gibt es keine einzige

Studie, die die Existenz eines „straight down beliefs“ bei Menschenaffen untersucht und damit diese (taxonomisch) wichtige Lücke zwischen Menschen und Affen schliesst. Weiter fehlen Untersuchungen zur Frage, ob die Dissoziation bezüglich Wissen um Solidität nur bei der „table task“ oder auch im Kontext der „tubes task“ gefunden werden kann.

Die im Manuskript I „Gravity bias in perception and action“ von Cacchione und Burkart zusammengefassten Untersuchungen, befassen sich mit der Frage, ob auch mit der „tubes task“ eine Dissoziation zwischen Wahrnehmungs- und Handlungswissen gefunden werden kann. Dabei wurden Weissbüscheläffchen (*Callithrix jacchus*) – eine in diesem Zusammenhang bisher noch nicht untersuchte Neuweltaffenart - sowohl mit einer Handlungs- als auch mit einer Erwartungsverletzungsversion (Blickzeitenmass) der „tubes task“ konfrontiert.

Die im Manuskript II „Gravity and solidity in great apes: the table task“ von Cacchione, Call und Zingg beschriebenen Experimente untersuchen die Existenz des Schwerkraftfehlers bei Menschenaffen. In einigen Varianten der „table task“ wird der Einfluss verschiedener schwerkraftrelevanter Variablen auf das Suchverhalten geprüft. Das Ziel ist es herauszufinden, ob die Schwerkraftthese oder einfachere Alternativhypothesen das Verhalten am Besten beschreiben.

Die im Manuskript III „Do great apes refer to a tube as a causal device? Intuitions about gravity and solidity in four great ape species“ von Cacchione und Call dargestellten Studien adressieren erneut die Frage, ob Menschenaffen in ihrem Handeln dem Schwerkraftfehler unterworfen sind, diesmal im Kontext der „tubes task“. Ziel ist es herauszufinden, welche perzeptiven Merkmale von Menschenaffen beim Lokalisieren unsichtbar verlagerter Objekte herangezogen werden, und ob sie Einsicht in die kausale Bedeutung der Röhre im Zusammenhang mit Objektverlagerungen haben.

1.4 Literaturangaben

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 47-83). Oxford: Blackwell.

Baillargeon, R., Spelke, E., S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, 20, 191-208.

Baillargeon, R., Kotovsky, L., & Needham, A. (1995a). The acquisition of physical knowledge in infancy. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: a multidisciplinary debate* (pp. 79-116). Oxford: Clarendon Press.

Baillargeon, R., Needham, A., & DeVos, J. (1992). The development of young infants' intuitions about support. *Early Development and Parenting, 1*, 69-78.

Baillargeon, R., Kotovsky, L., & Needham, A. (1995a). The acquisition of physical knowledge in infancy. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: a multidisciplinary debate* (pp. 79-116). Oxford: Clarendon Press.

Baillargeon, R., Raschke, H., & Needham, A. (1995b). Should objects fall when placed against or under other objects? The development of young infants' reasoning about support. *Unpublished manuscript*.

Bliss, J., Ogborn, J., & Whitelock, D. (1989). Secondary school pupils' common sense theories of motion. *International Journal of Science Education, II*, 261-272.

Bogartz, R. S., Shinskey, J. L., & Schilling, T. H. (2000). Object permanence in five-and-a-half-month-old infants? *Infancy, 1*, 403-428.

Bremner, G. J. (1997). From perception to cognition. In G. Bremner, A. Slater, & G. Butterworth (Eds.), *Infant development: recent advances* (pp. 55-74). East Sussex, UK: Psychological Press.

Cacchione, T., & Krist, H. (2004). Recognizing impossible object relations: intuitions about support in chimpanzees (*Pan troglodytes*), *Journal of Comparative Psychology, 118*, 140-148.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge: MIT Press.

Cashon, C. H. & Cohen, L. B. (2000). Eight month-old infants' perception of possible and impossible events. *Infancy, 1*, 429-446.

Dan, N., Omori, T., & Tomiyasu, Y. (2000). Development of infants' intuitions about support relations: Sensitivity to stability. *Developmental Science, 3*, 171-180.

Diamond, A. (1991). Neuropsychological insights into the meaning of object concept development. In S. G. Carey (Ed.), *The epigenesis of mind* (pp. 67-110). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Eckstein, S. G., & Shemesh, M. (1989). Development of children's ideas on motion: intuition vs. logical thinking. *International Journal of Science Education, 11*, 327- 336.

Friedman, W. J. (2002). Arrows of time in infancy: The representation of temporal-causal invariances. *Cognitive Psychology, 44*, 252-296.

Haith, M. M., (1998). Who put the cog in infant cognition: Is rich interpretation to costly? *Infant Behavior and Development*, 21, 167-179.

Halloun, I., A., & Hestenes, D. (1985). Common sense concepts about motion. *American Journal of Physics*, 53, 1056-1065.

Hauser, M. D. (2001). Searching for food in the wild. A nonhuman primate's expectations about invisible displacement. *Developmental Science*, 4, 84-93.

Hauser, M. D. (2003). Knowing about knowing. Associations between perception and action systems over evolution and during development. *Annual New York Academy of Sciences*, 1, 1-25.

Hauser, M. D., Williams, T., Kralik, J. D., & Moskovitz, D. (2001). What guides a search for food that has disappeared? Experiments on cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology*, 115, 140-151.

Hood, B. M. (1995). Gravity rules for 2-4-year-olds? *Cognitive Development*, 10, 577-598.

Hood, B. M. (1998). Gravity does rule for falling events. *Developmental Science*, 1, 59-64.

Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. (1999). Gravity biases in a nonhuman primate? *Developmental Science*, 2, 35-41.

Hood, B. M., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events. *Child Development*, 71, 1540-1554.

Hood, B. M., Santos, L., & Fieselman, S. (2000). Two year-old's naïve predictions for horizontal trajectories. *Developmental Science*, 3, 328-332.

Hood, B. M., Wilson, A., & Dyson, S. (2006). The effect of divided attention on inhibiting the gravity error. *Developmental Science*, 9, 303-308.

Howe, Christine J. (1998). *Conceptual structure in childhood and adolescence. The case of everyday physics*. London: Routledge.

Huettel, S. A., & Needham, A. (2000). Effects of balance relations between objects on infants' object segregation. *Developmental Science*, 3, 415-427.

Kaiser, M. K., & Profitt, D. (1984). The development of sensitivity to causally-relevant dynamic information. *Child Development*, 55, 1614-1624.

Kaiser, M. K., Profitt, D., & Anderson, K. (1985a). Judgements of natural and anomalous trajectories in the presence and absence of motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 795-803.

Kaiser, M. K., Profitt, D. R. & McCloskey, M. (1985b). The development of beliefs about falling objects. *Perception & Psychophysics*, 38(6), 533-539.

Kaiser, M. K., McCloskey, M., & Profitt, D. R. (1986a). Development of intuitive theories of motion in the absence of external forces. *Developmental Psychology*, 22, 67-71.

Kaiser, M., K., Jonides, J., & Alexander, J. (1986b). Intuitive reasoning about abstract and familiar physics problems. *Memory & Cognition*, 14, 308-312.

Kim, I. K., & Spelke, E. S. (1999). Perception and understanding of effects of gravity and inertia on object motion. *Developmental Science*, 2, 339-362.

Kim, I. K., & Spelke, E. S. (1992). Infants' sensitivity to effects of gravity on visible object motion. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 385-393.

Köhler, W. (1957). *The mentality of apes*. London: Pelican Books.

Krist, H. (2000). Development of naive beliefs about moving objects: The straight-down belief in action. *Cognitive Development*, 15, 281-308.

Mandler, J. M. (1998). Representation. In D. Kuhn & R. Siegler (Eds.), *Cognition, perception, and language*. Vol. 2, W. Damon (Series Ed.), *Handbook of child psychology* (pp. 255-308). New York: Wiley.

Marioni, C. (1989). Aspects of students' understanding in classroom settings (age 10-17): case study on motion and inertia. *Physics Education*, 24, 273-277.

McCloskey, M. (1983). Naïve theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental Models*, Hillsdale, N. J.: Lawrence Erlbaum Associates.

Munakata, Y. (2001). Task-dependency in infant behavior: Toward an understanding of the processes underlying cognitive development. Lacerda, von Hofsten, & Heimann (Eds.), *Emerging Cognitive Abilities in Early Infancy* (pp. 29-52). Mahwah, NJ: Lawrence Erlbaum Associates.

Needham, A., & Baillargeon, R. (1993a). Intuitions about support in 4.5-month-old infants. *Cognition*, 47, 121-148.

Needham, A., & Baillargeon, R. (1993b). Reasoning about support in 3-month-old infants. *Unpublished manuscript*.

Piaget, J. (1975). *Der Aufbau der Wirklichkeit beim Kinde*. Stuttgart: Klett.

Povinelli, D. J. (2000). *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford: University Press.

Santos, L. R. & Hauser, M. D. (2002). A non-human primates understanding of solidity: Dissociation between seeing and acting, *Developmental Science*, 5, F1-F7.

Santos, L. R., Seelig, D. & Hauser, M. D. (2006). Cotton-top tamarins' (*Saguinus oedipus*) expectations about occluded objects: a dissociation between looking and reaching tasks, *Infancy*, 9, 147-171.

Shanon, B. (1976). Aristotelianism, Newtonianism, and the physics of the layman. *Perception*, 5, 241-243.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, 99(4), 605-632.

Spelke, E.S. (1994). Initial knowledge: Six suggestions. *Cognition*, 50, 431-445.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: a dynamic field theory of infant perseverative reaching errors. *Behavioral and Brain Sciences*, 24, 1-86.

White, R., T. (1988). *Learning science*. Oxford: Basil Blackwell.

Running head: GRAVITY BIAS IN PERCEPTION AND ACTION

Gravity Bias in Perception and Action

Trix Cacchione¹, Judith M. Burkart²

¹Department of Psychology, Cognitive and Developmental Psychology, University of Zürich,
Switzerland

²Department of Anthropology, University of Zürich, Switzerland

Address correspondence to:

Trix Cacchione

Department of Psychology

University of Zürich

Binzmühlestrasse 14/21

8050 Zürich

Phone: +41 44 635 7498

Fax: +41 44 635 7489

E-Mail: t.cacchione@psychologie.unizh.ch

2. Gravity bias in perception and action

2.1 Abstract

While searching for an invisibly displaced object human and non-human primates as well as dogs are subject to a gravity bias: they erroneously search at the location underneath the release point, even when an obstacle deviates the object's path. Recent studies report knowledge dissociations between action and perception in a variety of tasks. The present studies investigate if gravity bias is observable in action only, or if it can be identified in perception too.

We tested common marmosets (*Callithrix jacchus*) with two versions of the so-called tube task where food is dropped through an opaque tube that is diagonally connected to one of three hiding places. Experiment 1 replicated the original search task ($n = 5$); in Experiment 2 a violation of expectancy version of the same task was used. Participants showed gravity biased search in the initial trials of Experiment 1. Results of Experiment 2 suggest a dissociation between action and perception: Naïve participants ($n = 7$) look longer when the object reappears in the hiding place specified by gravity.

2.2 Introduction

Gravity as an operating force on physical objects is omnipresent. Humans and non-human animals perpetually experience gravity and adaptations to gravity constraints got deeply ingrained in their perceptual and motor systems (e.g. Prechtel, 1989; Schone, 1984). Despite this, knowledge about gravity develops in a very slow and piecemeal fashion and remains rather fragile in humans over the whole live span (Baillargeon, 1995, 2002; Baillargeon, Raschke, & Needham, 1995; Baillargeon, Kotovsky, & Needham, 1995; Baillargeon, Needham, & DeVos, 1992; Dan, Omori, & Tomiyasu, 2000; Huettel & Needham, 2000; Kim & Spelke, 1992, 1999; Needham & Baillargeon, 1993; Spelke, Breinlinger and Macomber, & Jacobson, 1992). Other than knowledge about the physical principles of solidity or continuity, gravity is generally not supposed to be part of a potentially innate cognitive core (Spelke, 1994; Spelke, Breinlinger and Macomber, & Jacobson, 1992).

2.2.1 Gravity-biased search

Knowledge about gravity was found to have deep and reliable impact on search behavior in different mammal species: When searching for an invisibly displaced object, human children, non-human primates, and dogs expect that unsupported objects fall in a straight vertical line (Hauser, 2001; Hauser, Williams, Kralik, & Moskovitz, 2001; Hood, 1995, 1998; Hood, Hauser, Anderson, & Santos, 1999; Osthaus, Slater, & Stephen, 2003). The expectation that objects fall straight down appears to be so fundamental that even when a solid obstacle impedes a straight vertical fall, search is persistently directed at the location specified by the gravitational line.

Hood (1995) confronted 2- to 4- year-old children with an invisible displacement task where the goal was to locate a ball that was dropped down one of three intertwined opaque tubes connected to three hiding places. Search errors of 2- to 3- year-olds were not at random but occurred significantly more often at the hiding place directly underneath the release place. Strikingly, even though this box was never connected to the tube and the child was always shown the actual position of the object, children seemed unable to correct their performance. Hood interpreted this behavior as stemming from an underlying naïve theory of gravity: The belief that objects always fall straight down.

Subsequent studies revealed that gravity-biased search is not restricted to human children. Hood, Hauser, Anderson and Santos (1999) report a highly robust gravity bias in Cotton-top tamarins (Saguinus oedipus), when tested with the tube task outlined above. In contrast to children perseveration was even more pronounced. Also, in many cases the tamarins favored the middle box, a strategy that was not observed in children.

Recently Osthaus, Slater and Stephen (2003) run the tube task on dogs (Canis lupus familiaris). Like in toddlers and tamarins, the search of dogs was biased towards the box in gravity direction. A preference for the middle box was observed in dogs as well. Albeit they did not understand the mechanism of the tube, dogs were in contrast to children and tamarins able to overcome the bias: over trials they learned the specific location of the object. This led to the interpretation that strength of gravity bias might differ across species.

Gravity bias is not only traceable in different species; it is also observable in different tasks involving invisible displacement. Using another search task (table task) Hood, Carey, and Prasada (2000) confirm biased search in 2-year-old children. In this case an object is dropped behind a screen on top of a table. Again the children expected the object to fall in a vertical line to the lowest point: They searched under the table, neglecting the fact that a solid object cannot move through another solid object. With a similar set up of the table task, Hauser (2001) also found this gravity biased search in rhesus macaques (Macaca mulatta).

To sum up, the present state of research suggests that first, different mammals including human children rely on knowledge about gravity when searching for invisibly displaced objects; and that second, this knowledge is generally attributed to the operation of an underlying naïve theory.

2.2.2 Dissociation between perception and action

The conclusion that - in a search context - knowledge about gravity seems to override knowledge about solidity is in sharp contrast to findings obtained using looking time measures. Results based on looking times indicate that knowledge about solidity generally precedes knowledge about gravity: Sensitivity to solidity constraints seems to be robustly present in infants as young as 2.5 month (e.g. Hespos & Baillargeon, 2001; Spelke, Breinlinger and Macomber, & Jacobson, 1992), whereas measurable reactions to the violation of aspects of gravity can be traced earliest around 4 months of age (Baillargeon, 2002; Baillargeon, Kotovsky, & Needham, 1995; Kim & Spelke, 1992; Spelke, Breinlinger and Macomber, & Jacobson, 1992).

Even within a given task itself, in recent years dissociations between knowledge revealed by action and perception were frequently observed: perception-based tasks tend to reveal much earlier competences than tasks based on action. This contrast was initially attributed to the higher demand on executive functions in tasks that involve direct action (Diamond, 1991; Baillargeon, Graber, DeVos, & Black 1990; Hood, & Willats, 1986; Keen, 2003; Munakata, McClelland, Johnson, & Siegler, 1997). However, even older children, at an age where the executive problems associated with direct action control are no longer decisive, were found to be unable to act on knowledge properties whose violation they already perceptively detected as infants (Ahmed & Ruffman, 1998; Berthier, DeBlois, Poirier, Novak, & Clifton, 2000; Hood, 1995; Hood, Carey, & Prasada, 2000; Hood, Cole-Davies, & Dias, 2003). As a consequence the use of perception-based measures (such as habituation-dishabituation and violation of expectancy tasks) has been increasingly criticized (Baillargeon, 2004; Bogartz, Shinsky, & Schilling, 2000; Cashon & Cohen, 2000; Haith, 1998; Hood, 2001; Munakata, 2000; Rivera, Wakeley, & Langer, 1999; Schilling, 2000). Particularly controversial is the issue of what kind of knowledge systems are tapped by perception based measures, and of whether explicit conceptual knowledge can actually be demonstrated by perception based measures.

Dissociations between action and perception are not exclusive to human cognition but can equally be observed in non-human primates (Cacchione, & Krist, 2004; Santos, & Hauser, 2002; Leslie, 1994). To clarify why action and perception tasks often give contrasting evidence future research should focus on knowledge dissociations themselves. From a comparative perspective the most promising avenue is to compare responses of different species when tested with action- and perception-based versions of the same task type. Of special interest in the present context is the first experiment of Santos and Hauser (2002). With a looking version of a search task employed by Hauser (2001) they tested whether rhesus macaques (Macaca mulatta) would detect the violation of solidity when looking at invisible displacements without being engaged in active search. As in the search task, participants were presented with a solid table-like apparatus that consisted of two sidewalls through which a solid shelf could be inserted. An occluder was set up and an object dropped behind it. The removal of the occluder revealed two different outcomes: Either (consistent with the principle of solidity) the object was positioned on top of the shelf, or below the shelf (inconsistent with the principle of solidity). Rhesus macaques looked reliably longer when the object appeared to travel through the solid shelf, suggesting that they have some sensitivity for solidity constraints. These findings are in sharp contrast to responses in the action version

of the task where rhesus macaques tended to neglect the principle of solidity and searched below the shelf. So, at least in this looking context, naïve reasoning about gravity does not necessarily override knowledge about solidity: The expectation that all objects travel in a straight line did not influence the looking times of the rhesus macaques.

The current study set out to determine if a dissociation between perception and action can also be observed using the original tube task (Hood, 1995), in which different species were found to display a highly persistent gravity bias as revealed by direct search action. Would the naïve belief that all unsupported objects fall in a straight vertical line also guide looking behavior? Would therefore knowledge about gravity suppress knowledge about solidity both when acting and when looking at tubes? Or, does knowledge about solidity generally prevail in perception, irrespective of the specific task context? To test for these possibilities we run a looking and an action version of the tube task on common marmosets (*Callithrix jacchus*). To make sure that prior experience with one of the two task types does not influence reactions two sub samples were tested in a between-subject design.

2.3 Experiment 1

Experiment 1 was a replication of the original action task (Hood, 1995; Hood et al., 1999). Hood et al. (1999) report a marked gravity bias in cotton top tamarins, a New World monkey species closely related to common marmosets. To directly compare action and perception responses, we must provide evidence that gravity bias is present in common marmosets. In a first step, we therefore investigated whether common marmosets like cotton top tamarins are subject to the perseverative search in gravity direction when observing invisible object displacements.

2.3.1 Method

Participants

Participants were five adult common marmosets (*Callithrix jacchus*), a small New World monkey species. All individuals were born in captivity and mother-reared and are housed at the Primate Station of the Anthropological Institute of the University of Zürich. During experimentation, they were pairwise housed in outdoor cages that were equipped with branches, ropes, and different sitting places and covered with natural soil. There were

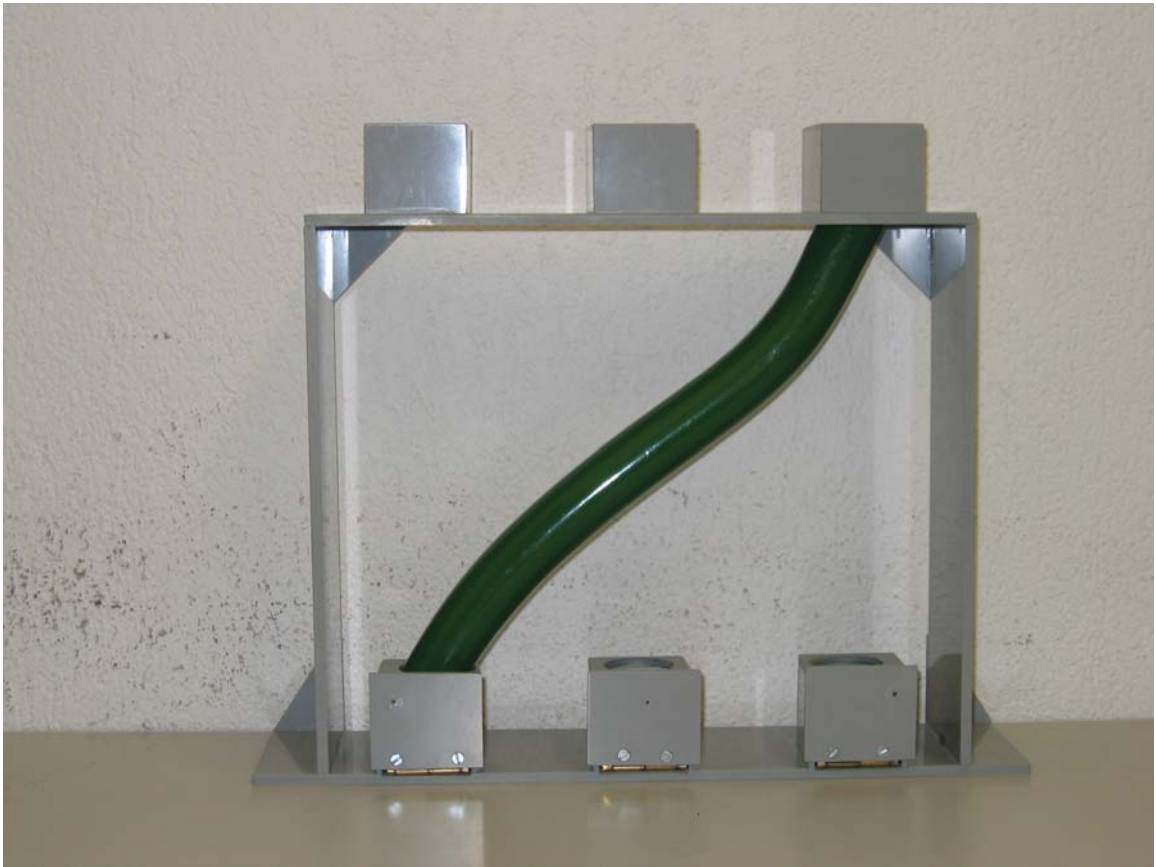
two adult males and three adult females, aged between 2 years 11 months and 7 years 1 month (average = 4 years 11 months, SD = 27.2 months). The animals formerly participated in experiments exploring social dynamics in the context of reproduction but had never been subjected to any cognitive task or invasive experimentation.

Apparatus

The apparatus (Figure 1) used in the action task was a scaled down remake of the original apparatus of Hood (1995). On the upper section there were three chimneys of 3.5 cm diameter separated by 6.5 cm. On the lower section (20 cm below) were three goal boxes (4.5 x 4.5 x 4.5 cm³). All goal boxes had a hole on the top and a front door. A single plastic opaque tube connected the upper and the lower section.

Subjects were lured into a transparent Plexiglas box (30 x 30 x 45 cm³) that was placed at the front door of the home cage. The front of the box was made of wire mesh and could be covered by an opaque screen, all other sides were opaque. This setting allowed us testing the animals individually without separating them from their mates acoustically or by long distances, which reduced the stress of being isolated almost completely.

Figure 1



Apparatus used in the action task.

Procedure

Training

Participants were first familiarized with the tube and the apparatus in their home cages. During the initial phase marmosets were allowed to explore the tube. The experimenter placed the tube in front of the animals and made them look through it. To ensure that they were aware that objects can travel through the tube, the experimenter dropped three non-food objects (piece of bark) down the tube that was in a vertical position. She made sure the marmosets tracked the complete process of dropping. After this, the apparatus without tube was placed in the home cage for 15 min to give the participants the possibility to explore it in detail.

The next day, marmosets were run on a training condition, to control for box preferences. One after another was lured in the Plexiglas box and the apparatus placed in front of it. The experimenter then opened all three boxes to demonstrate that they were empty and

closed them again. The participants were allowed to open the box doors and to explore the compartments. Then the apparatus was removed out of reach of the animals (10 cm away from the wire mesh) and a food item (half a raisin) visibly placed through the top whole in a random goal box. The food was first waved back and forth over the opening of the box. Only when the marmoset visually tracked the food it was placed inside the box. The apparatus was pushed back in reaching distance of the participants again and they were allowed to search for the food until found. For each animal 10 trials were run. The search behaviour was recorded on a check sheet. If the search was directed to different boxes in sequence, the search order was recorded. Participants that successfully found the food on the first attempt in at least 80% of all trials moved on to test condition. Participants that failed to do so were presented with additional 10 baseline trials the next day.

Test

Each participant was presented with two sessions on two consecutive days, immediately following the training. Two sessions were scheduled to assess the strength of a potential bias, and to determine whether it would decline over trials after the repeated localisation of the food in the box connected to the tube, the non-gravity box.

A session consisted of a test condition and a generalisation condition. Participants were again placed in the Plexiglas box and the apparatus positioned in front of them. To prevent that the marmosets located the food relying on acoustic cues, white noise was played throughout both sessions.

In the test condition the experimenter attached an opaque tube from the upper left chimney to the lower right goal box. Half of the participants were presented with the opposite configuration (upper right chimney– lower left box). During the second session, the configuration was changed, so that all animals were tested with both start configurations. A food item was moved back and forth over the chimney with the tube attached. If the participant tracked the food it was dropped. Then the apparatus was moved in reaching distance and the marmosets were allowed to search for the food until found. The order of the boxes searched was recorded. If the food was correctly located on the first attempt in 4 out of 5 trials, the participant moved on to the generalisation condition.

The generalisation was administered to investigate if successful participants truly understood the mechanism of the tube or if they merely searched at previously reinforced locations. Each participant reaching the criterion of 4 out of 5 trials correct was now presented with the opposite tube configuration (e.g. test: upper left – lower right;

generalisation: upper right – lower left). The same procedure as in the test condition was used. Again, if participants found the food in the new configuration on their first attempt in 4 out of 5 trials, the tube was switched back to the prior configuration (it was never attached to the middle locations). This was repeated up to a maximum of 16 trials per session. If search was not successful tube configuration was not changed until the 16 trials of a session were completed.

2.3.2 Results

Training

Four out of five participants reached the criterion of at least 80% correct during the first training session. One participant needed a second training session. Detailed succession of performance during all trials is shown in Figure 2.

Figure 2

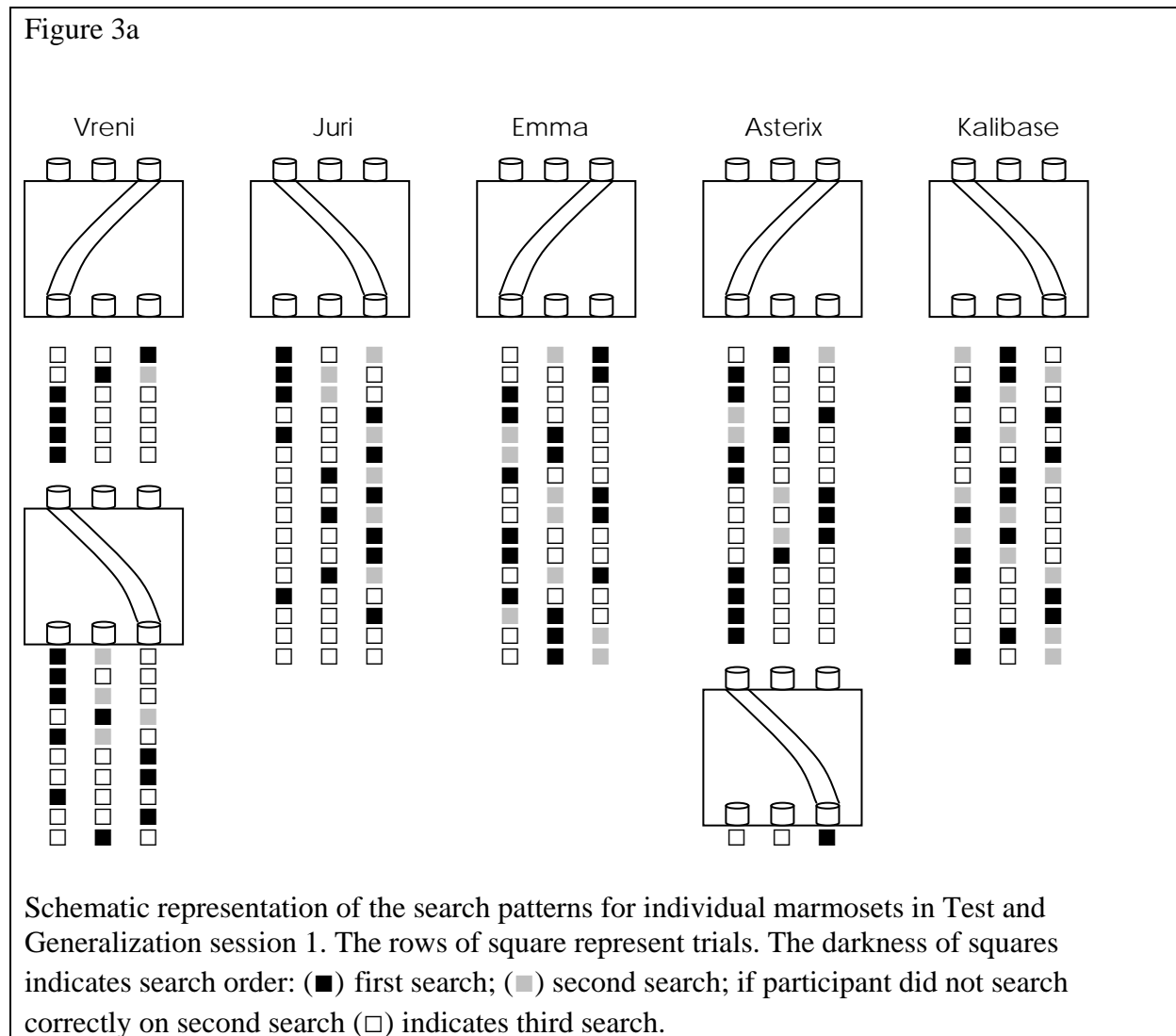
| Training session 1 | | | | | | | | | | | |
|--------------------|-----------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|
| Animal | Position of food item | | | | | | | | | | % correct |
| | R | L | R | R | M | R | L | M | L | M | |
| Vreni | R | L | R | R | M | R | L | M | L | M | 100 % |
| Juri | R | L | L | L | L | R | L | R | L | L | 50 % |
| Emma | R | L | R | R | M | R | L | M | M | M | 90 % |
| Asterix | R | L | R | R | M | R | L | M | L | M | 100 % |
| Kalibase | R | L | R | R | M | R | L | M | L | M | 100 % |

| Training session 2 | | | | | | | | | | | |
|--------------------|-----------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-------|
| | Position of food item | | | | | | | | | | |
| | L | R | L | R | L | M | M | L | M | R | |
| Juri | L | R | L | R | L | M | M | L | M | R | 100 % |

Succession of performance during all trials in both training sessions for individual marmosets.

Test and generalisation 1

Gravity biased search during the first confrontation with the task could be observed in 3 out of 5 participants (3-choice binomial distribution: $p < .05$). The remaining two animals first chose the middle box, and as second choice the gravity box. One animal passed the task criterion after six trials but showed gravity biased search again after the configuration was switched: it perseverated the previous response and chose the box that was the correct one prior to the configuration switch. One additional animal passed the task after 15 trials and searched in the correct box after the configuration switch. The remaining three animals did not pass the task within the 16 trials of test session 1. Their searches were equally distributed between the three boxes (gravity box: 34.75%, middle box: 30.5%, box connected to tube: 34.75%, Figure 3a).

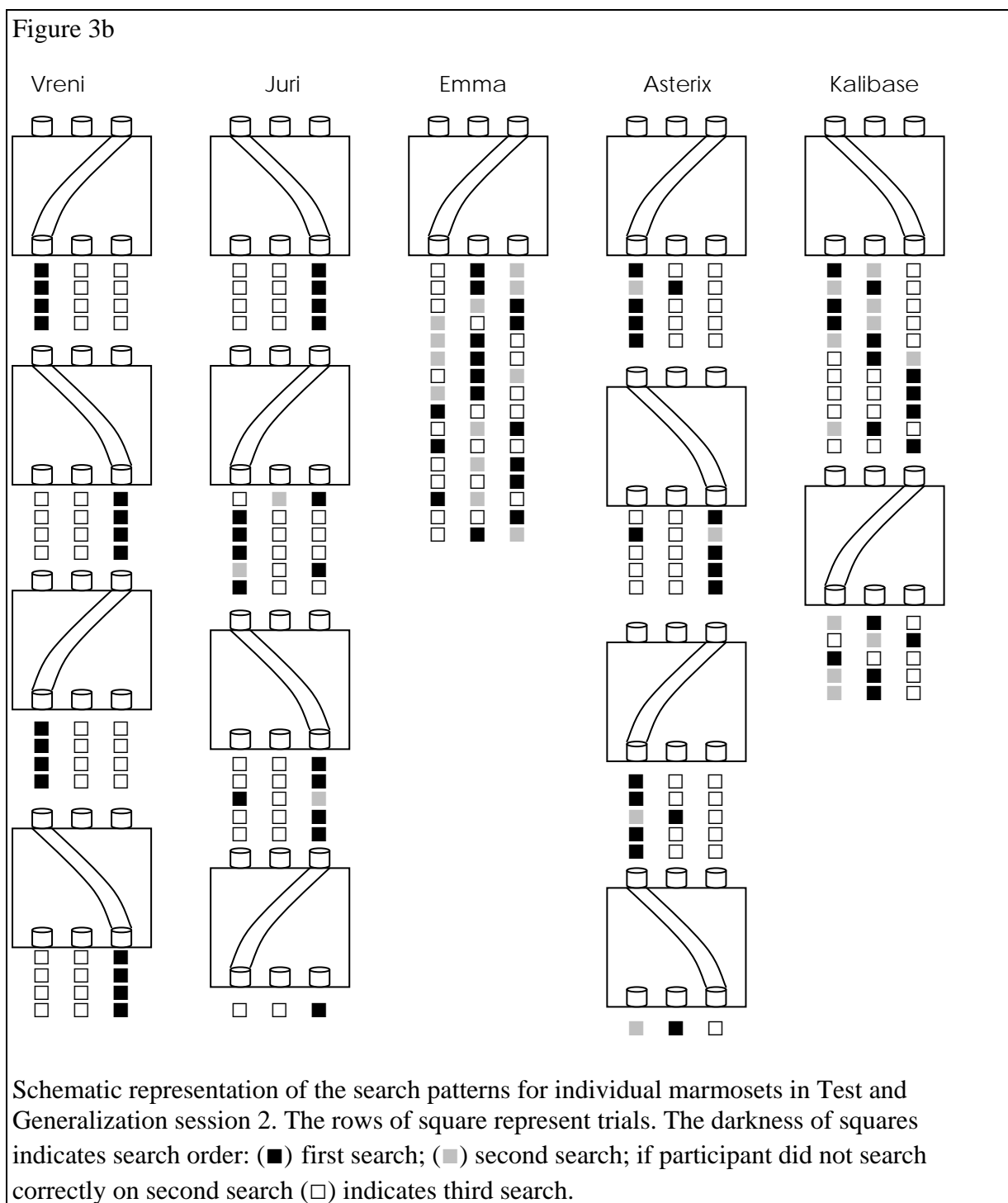


Test and generalisation 2

During test session two, the overall performance of the animals was improved and the gravity bias not traceable any more during the first choice: Three participants immediately searched the correct target box (3-choice binomial distribution: $p < .05$), one searched the middle container and only one still searched the container indicated by the gravity line.

One animal searched the correct box in all trials which resulted in three configurational swifts. This animal never perseverated to reach for the previously rewarded box but adjusted its choice consistently to the actual configuration of the apparatus. Two additional animals reached the criterion three times as well but gravity biased search eventually reappeared. One animal reached the criterion once after 11 trials but searched the middle container after configuration switch. One animal finally never reached the criterion for

the task. Her search was directed to the correct box in 18.75%, to the middle box in 43.75 %, and to the gravity box in 37.5 % (Figure 3b).



2.3.3 Discussion

As expected, the gravity bias is traceable in common marmosets too, as it is in human children and non-human animals tested so far. Unlike in tamarins, however, no pronounced

tendency for searching in the middle target box was observed. Four out of five marmosets overcame the gravity bias after a number of trials and started looking for the food in the correct target box, the one connected to the tube. Thus, at least four individuals were able to overcome the bias, two of them were quite easily able to do so. The transition to looking for the food item in the correct target box revealed that some deeper understanding of the situation might be involved in the animals: They first climbed up the wire mesh and tried to look and grasp into the upper chimney where the food disappeared, then followed the tube along the mesh and hence, finally arrived at the correct container. Due to the (expectable) scarcity of such behaviour however, the possibility of overcoming the gravity bias by simple place learning could not be ruled out.

To account for the possibility that the subjects learned to look for the food item in a specific container using a simple heuristic rule, we changed the configuration of the tube after reaching the criterion of four out of five correct choices, and tested whether the subjects would generalize. Would they continue to look in the old correct target box that now had become the gravity box, or would they change their search pattern according to the new configuration? At least two animals managed to deal with repeated reversals of the tube position, thereby showing some generalisation for locating the correct container.

To sum up, results demonstrate a gravity bias in marmosets during the search for an invisibly displaced object, as well as the possibility that they might overcome this misleading bias. While observations during the task strengthens the idea that the subjects who overcame the bias actually understand the tube mechanism, a formal demonstration of this understanding is still lacking (the same pattern of results could have emerged as a consequence of learning the heuristic rule of “look for the reward in the box attached to the tube”). However, like in human children and in macaques, in naïve marmosets knowledge about gravity overrides knowledge about solidity in this particular search task.

2.4 Experiment 2

Experiment 2 - a looking time version of Experiment 1 - sets out to determine whether a dissociation between action and perception can be found in this particular task context. Will the reappearance of an invisibly displaced object in the gravity container evoke longer looking times in common marmosets? Or, on contrary, will they look longer at an object in the

container connected to the tube, because it conflicts with the expectations that apparently guided their search behaviour during the first trials in Experiment 1?

2.4.1 Method

Participants

Participants were 7 naïve adult common marmosets (*Callithrix jacchus*), born in captivity and mother-reared and housed at the Primate Station of the Anthropological Institute of the University of Zürich. Their age ranged from 11 months to 15 years (average = 4 years 1 month, SD = 68 months). The animals formerly participated in experiments exploring social dynamics in the context of reproduction but had never been subjected to any cognitive task or invasive experimentation. Most importantly, none of the participants had prior experience with the search task.

Apparatus

The apparatus used in the looking task was identical to the one used in the action version, apart the following exceptions: Instead of three goal boxes only two goal boxes were on the lower section (the middle box was removed). The upper section of the tube was tamped to stop the dropped object. The food was always placed in the respective goal box beforehand and never travelled down the tube. A mechanism allowed opening the doors simultaneously. Subjects were individually lured into the testing box used in experiment 1. However, the wire mesh front of the box was replaced by a transparent Plexiglas which allowed to film the looking behaviour of the animals in detail. An additional screen allowed to interrupting visual access to the scene in front of the box. All other sides of the box were opaque.

Procedure

In the looking time version of the tube task we employed the so-called violation of expectancy methodology. It is based on the assumption that infants (and non-human primates) look longer at events that violate their expectations, this expectations being generated by knowledge representations. Therefore, longer looking times are thought to indicate violations of knowledge contents. The basic principle of violation of expectancy is to present each participant with two test events similar in structure, of which only one is congruent whereas the other is in conflict with the knowledge to be tested. Traditionally several familiarization

trials precede the test trials. However, it is often criticised that familiarisation itself may set up the expectations later tested (e.g. Thelen & Smith, 1994). To rule out this possibility, we omitted familiarisation trials completely.

After a pretraining that was identical to experiment 1 and that gave the animals the opportunity to explore the tube and apparatus, all participants saw two test events, a congruent and an incongruent one. In the congruent event, the food item dropped down the tube, the front door of the boxes opened and the food item was revealed in the box attached to the tube. In the incongruent event, the food item was revealed in the box directly underneath the release point. The order of the presentation was counterbalanced across participants.

Participants were lured in the Plexiglas box and the screen lowered. The apparatus was placed in front of them. An opaque tube was connected from one chimney to the opposite target container, for half of the animals from the upper right chimney to the left target container, for the other half in the other direction. A camera, positioned directly behind the apparatus, recorded the experimental session. It allowed the observation of the animals' looking behaviour but not of the content of the box once the door was opened. This allowed us to code the videotapes blindly with regard to the experimental condition (congruent vs. incongruent test event). In the congruent test event the experimenter placed a food item in the goal box attached to the tube and closed the door. Then the screen was raised. The experimenter took another food item and waved it above the chimney connected to the tube. If the participant tracked it, the food was released. If a participant did not track the entire releasing process, the dropping was repeated. Then both doors were opened simultaneously and the looking behaviour was recorded for the following 10 seconds.

The procedure in the incongruent test event was identical with the exception that the food was placed in the unconnected box specified by the gravitational line. The presentation order of the congruent and the incongruent test event was counterbalanced and both events presented during a single test session. At the end of the test all subject got a food reward, irrespective of their performance.

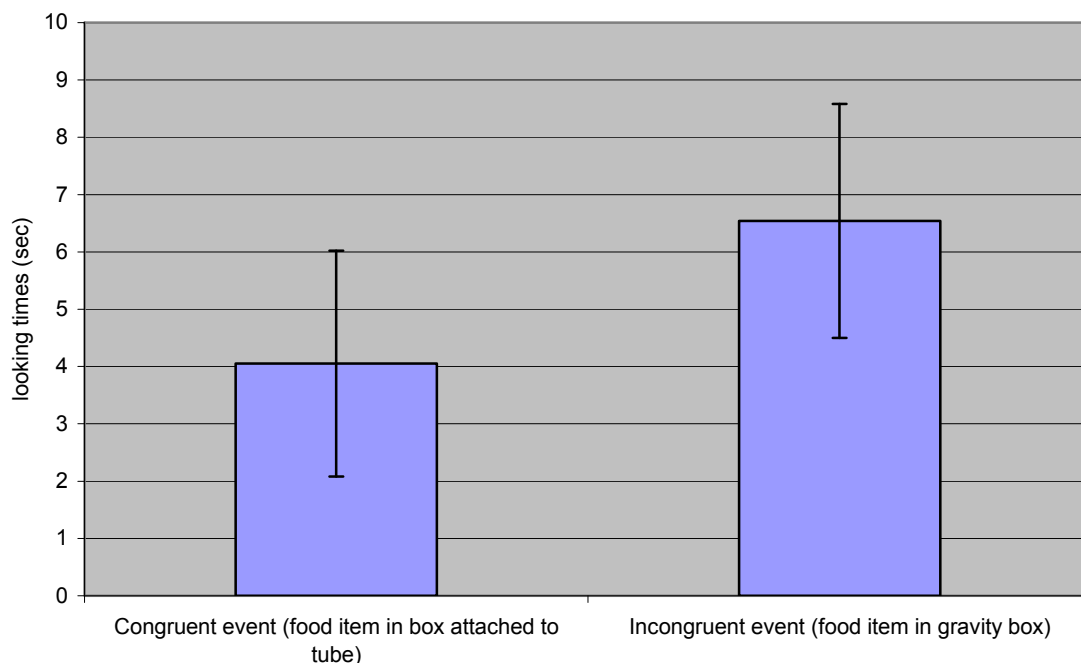
Analyses

Two raters analysed looking times frame by frame (25frames/ second) blindly with regard to the experimental condition. Looking times were assessed for a duration of ten seconds following the opening of the doors. Interrater concordance was 94.6%, interrater reliability $K = 0.892$ (Cohen's Kappa). Statistical analysis was performed on the raw data from the first rater.

2.4.2 Results

Subjects looked at the congruent event only during 4.05 sec (SD = 1.97 sec), while the incongruent test event did attract their attention for 6.54 sec (SD = 2.04, Figure 4). A 2 x 2 ANOVA with repeated measure was performed, for “congruence of test event” (within) and “trial order” (between). The analysis revealed a main effect for “congruence of test event” ($F(1) = 11,68$; $p < .05$), indicating that participants looked significantly longer at the incongruent test event where the food appeared in the gravity box than in the congruent test event where the food appeared in the box that was connected to the tube. No effect of trial order ($F(1) = 0.706$; n.s.) and no interactions ($F(1) = 0.373$; n.s.) were revealed by the statistical analysis.

Figure 4



Mean looking-time in Experiment 2 (n = 7).

2.4.3 Discussion

In the violation of expectancy version of the tube task, marmosets looked longer at the impossible outcome where the piece of food was revealed in the target box vertically below the releasing point, in the gravity box. When the food was revealed in the box connected to the tube, participants displayed shorter looking times.

The effect was not influenced by the trial order as no interaction between trial order and experimental condition was detected. Also, the effect cannot be attributed to expectations built during habituation because no habituation trials were administered at all. Therefore, the correct expectation about the result of the invisible displacement of the piece of food, correctly integrating knowledge about solidity as well as knowledge about gravity, already was pre-existing prior to experimentation.

This result is in contrast with the results from the action version of the tube task and suggests that a dissociation between perception based and action based measures is traceable in adult marmosets.

2.5 General discussion

Common marmosets are subject to gravity biased search: When attempting to localise an object dropped down an opaque tube, their immediate response is to search the gravity container directly below the releasing point. In this respect they react the same way as human infants, cotton top tamarins, rhesus macaques, and dogs do. However, if marmosets simply observe the dropping of an object down a tube they do not appear to expect the objects' reappearance in the gravity container. On contrary, they look longer, if the object is revealed in the gravity container as when it is revealed in the container actually connected to the tube. This suggests that a dissociation between knowledge revealed by action and by perception has been identified in this particular task-context: When tested with two versions of the tube task, common marmosets seem sensitive to violations of solidity only when perceiving objects, but not when acting on them. Or put differently: Only when marmosets are engaged in active search for an object, knowledge about gravity overrides knowledge about solidity. This "double" dissociation meets up to previous studies that report early evidence of gravity in

action and judgement tasks, whereas solidity was found to prevail in perception (e.g. Kim & Spelke, 1999).

How can the discordant results gathered from action- and perception-based measures be explained? In human children, corresponding dissociations are often discussed as developmental transitions, with knowledge being detectable early using the more sensitive perception tasks, but much later in tasks requiring action. Action tasks are more demanding on executive functions: planning and executing a successful search requires cognitive and sensorimotor capacities not available early in development (e.g. Diamond, 1991; Baillargeon, et al., 1990; Keen, 2003). However, the participants in both tasks being adult common marmosets, a developmental explanation is not appropriate here.

We favour the view that the occurrence of double dissociation bears evidence of different cognitive processes being involved in either task context. A current explanation is that action and perception response modes are based on different knowledge types. Mandler (2004) proposes to make a clear distinction between conceptual and perceptual knowledge, with perception-based measures only capturing the latter. In her framework, perceptual knowledge is conceived as principally procedural, implicit and unconscious - in contrast to conceptual knowledge which is supposed to be potentially accessible to conscious reasoning. In the present study marmosets perceptively detect a violation of solidity. However, knowledge about solidity did apparently not guide their search behaviour systematically: In many instances marmosets were subject to gravity biased search. Recall Hood's suggestion that gravity bias stems from an underlying naive theory. The bias might be observable only in a search context, because conceptual knowledge is revealed only if participants are engaged in overt action where an explicit prediction of the objects current localization is required. When merely witnessing the invisible displacement and revelation of the object, marmosets might implicitly detect a physical anomaly in the events presented without having any explicit understanding of the physical principles involved.

However, the results of Experiment 1 somewhat weakens the theory account, as the bias was found to be less persistent in non-human primates as previously thought. In contrast to cotton top tamarins, common marmosets were not found to persevere over scores of trials. Instead, search patterns reveal that throughout the second session gravity biases decreased. This finding is unexpected, given the performance of tamarins that searched the gravity location persistently even after repeated trials without direct reinforcement for choosing the gravity box (Hood et al., 1999). In fact the marmosets' responses appear to resemble that of dogs (Osthaus et al., 2003). Even though dogs did not understand the

functioning of the tube, they could overcome the gravity bias if they retrieved the food item at a constant location. Over trials they learned the specific location where the food was found. They did not, however, use the connecting tube as an indicator for the position of the food item. Eventually marmosets are more apt to use the tube as a cue to localize the food. It is at least possible that the prior exposure to the tube added to the larger flexibility of the marmosets in this respect. However, the current study does not permit (and was not aimed at) drawing any conclusion of why gravity biased search arises and why it appears to be less persistent in marmosets. The observed perseveration might just as well be explained by inhibitory failures – and has potentially nothing to do with theory-based explicit reasoning. Further research is needed to decide why gravity bias manifests itself in action; if the strength of the error varies across species as suggested by Osthaus et al. (2003) or if there are specific testing conditions under which gravity bias may be more or less pronounced (e.g. distance between releasing point and chimney where the object can be observed falling vertically before disappearing in the chimney, dimension of the apparatus with respect to the body size, etc.).

To sum up the present study leads to the conclusion that - depending on the response mode - different knowledge properties are traceable in common marmosets. Performance differences in perception and action might be attributed to the operation of functionally different knowledge systems. However, it would be wrong to jump to the conclusion that action measures generally offer a pristine approach to test conceptual knowledge. Even if the requirements to plan and exert action are complied, action systems often fail to directly transmit conceptual knowledge into appropriate action. Likewise, it would be wrong to underestimate the methodological relevance of perception-based measures. Future research should focus on a comparison of different methodological approaches to offer a more comprehensive investigation of the many ways in which knowledge concretises itself in observable responses.

2.6 References

Ahmed, A., & Ruffman, T. (1998). Why do infants make A not B errors in a search task, yet show memory for the location of hidden objects in a non-search task? Developmental Psychology, *34*, 441-445.

Baillargeon, R., Graber, M., DeVos, J., & Black, J. (1990). Why do young infants fail to search for hidden objects? Cognition, *36*, 255-284.

Baillargeon, R., Needham, A., & DeVos, J. (1992). The development of young infants' intuitions about support. Early Development and Parenting, *1*, 69-78.

Baillargeon, R. (1995). A model of physical reasoning in infancy. In C. Rovee-Collier & L. P. Lipsitt (Eds.), Advances in infancy research (Vol. 9, pp. 305-371). Norwood, NJ: Ablex.

Baillargeon, R., Kotovsky, L., & Needham, A. (1995). The acquisition of physical knowledge in infancy. In D. Sperber, D. Premack, & A. J. Premack (Eds.), Causal cognition: a multidisciplinary debate (pp. 79-116). Oxford: Clarendon Press.

Baillargeon, R., Raschke, H., & Needham, A. (1995). Should objects fall when placed against or under other objects? The development of young infants' reasoning about support. Unpublished manuscript.

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Ed.), Handbook of childhood cognitive development (pp. 47-83). Oxford: Blackwell.

Baillargeon, R. (2004). Infants' reasoning about hidden objects: Evidence for event-general and event-specific expectations. Developmental Science, *7*, 391-424.

Berthier, N., Deblois, S., Poirier, C. R., Novak, M. A., & Clifton, R. K. (2000). Where's a ball? Two and three year olds about unseen events. Developmental Psychology, *36*, 394-401.

Bogartz, R. S., Shinksey, J. L., & Schilling, T. H. (2000). Object permanence in five-and-a-half-month-old infants? Infancy, *1*, 403-428.

Cacchione, T., & Krist, H. (2004). Recognizing impossible object relations: Intuitions about support in chimpanzees (*Pan troglodytes*). Journal of Comparative Psychology, *118*, 140-148.

Cashon, C. H., & Cohen, L. B. (2000). Eight-month-old infants' perception of possible and impossible events. Infancy, *1*, 429-446.

Dan, N., Omori, T., & Tomiyasu, Y. (2000). Development of infants' intuitions about support relations: Sensitivity to stability. Developmental Science, *3*, 171-180.

Diamond, A. (1991). Neuropsychological insights into the meaning of object concept development. In S. Carey & R. Gelman (Eds.), The epigenesis of mind: Essays on biology and cognition (pp. 67-110). New Jersey: Erlbaum.

Haith, M. M. (1998). Who put the cog in infant cognition. Is rich interpretation too costly? Infant Behavior and Development, *21*, 167-179.

Hauser, M. D. (2001). Searching for food in the wild. A nonhuman primate's expectations about invisible displacement. Developmental Science, *4*, 84-93.

Hauser, M. D., Williams, T., Kralik, J. D., & Moskovitz, D. (2001). What guides a search for food that has disappeared? Experiments on cotton-top tamarins (Saguinus oedipus). Journal of Comparative Psychology, *115*, 140-151.

Hespos, S. J., & Baillargeon, R. (2001). Knowledge about containment events in very young infants. Cognition, *78*, 204-245.

Hood, B. M. (1995). Gravity rules for 2-4-year-olds? Cognitive Development, *10*, 577-598.

Hood, B. M. (1998). Gravity does rule for falling events. Developmental Science, *1*, 59-64.

Hood, B. (2001). When do infants know about objects, Perception, *30*, 1281-1284

Hood, B. M., & Willats, P. (1986). Reaching in the dark to an object's remembered position: Evidence for object permanence in 5-month old infants. British Journal of Developmental Psychology, *4*, 57-65.

Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. (1999). Gravity biases in a nonhuman primate? Developmental Science, *2*, 35-41.

Hood, B. M., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events. Child Development, *71*, 1540-1554.

Hood, B.M., Cole-Davies, V. & Dias, M. (2003) Looking and search measures of object knowledge in pre-school children. Developmental Psychology, *39*, 61-70.

Howard, I.P. (1982). Human visual orientation. New York: Wiley.

Huettel, S. A., & Needham, A. (2000). Effects of balance relations between objects on infants' object segregation. Developmental Science, *3*, 415-427.

Keen, R. (2003). Representations of objects and events: Why do infants look so smart and toddlers look so dumb? Current Directions in Psychological Science, *12*, 79-83.

Kim, I. K., & Spelke, E. S. (1992). Infants' sensitivity to effects of gravity on visible object motion. Journal of Experimental Psychology: Human Perception and Performance, *18*, 385-393.

Kim, I. K., & Spelke, E. S. (1999). Perception and understanding of effects of gravity and inertia on object motion. Developmental Science, *2*, 339-362.

Leslie, A. M. (1994). ToMM, ToBy, and agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), Mapping the mind: Domain specificity in cognition and culture (pp. 119-148). New York, NY: Cambridge University Press.

Mandler, J. M. (2004). A synopsis of "The foundations of Mind: Origins of conceptual thought". Developmental Science, *7*, 499-505.

Munakata, Y. (2000). Challenges to the violation-of-expectancy paradigm: Throwing the conceptual Baby out with the perceptual processing bathwater? Infancy, *1*, 471-477.

Munakata, Y., McClelland, J. L., Johnson, M. H., & Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. Psychological Review, *104*, 686-713.

Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. Cognition, *47*, 121-148.

Osthaus, B., Slater, A. M., & Stephen, L. E. G. (2003). Can dogs defy gravity? A comparison with the human infant and a non-human primate. Developmental Science, *6*, 489-497.

Prechtl, H., F., R. (1989). The development of postural control in infancy. In C. von Euler, T. Forsberg, & H. Lagercrantz (Eds.), Neurobiology of early infant behavior (pp. 59-68). London: Macmillan.

Rivera, S. M., Wakeley, A., & Langer, J. (1999). The drawbridge phenomenon: Representational reasoning or perceptual preference? Developmental Psychology, *35*, 427-435.

Santos, L. R. & Hauser, M. D. (2002). A non-human primates understanding of solidity: Dissociation between seeing and acting, Developmental Science, *5*, F1-F7.

Schilling, T. H. (2000). Infants' looking at possible and impossible screen rotations: The role of familiarization. Infancy, *1*, 389-402.

Schone, H. (1984). Spatial orientation. Princeton, NJ: University Press.

Spelke, E. S. (1994). Initial knowledge: six suggestions. Cognition, *50*, 431-445.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. Psychological Review, *99*, 605-632.

Thelen, E., & Smith, L. B. (1994). A dynamic systems approach to the development of cognition and action. Cambridge, MA: MIT Press.

Running Head: GRAVITY AND SOLIDITY IN GREAT APES

Gravity and solidity in great apes: The table task

Trix Cacchione¹, Josep Call², Robert Zingg³

¹Department of Psychology, Cognitive and Developmental Psychology, University of Zürich,
Switzerland

²Max-Planck-Institute for Evolutionary Anthropology, Developmental and Comparative
Psychology, Leipzig, Germany

³Zoo Zürich

Address correspondence to:

Trix Cacchione

Department of Psychology

University of Zürich

Binzmühlestrasse 14/21

8050 Zürich

Phone: +41 44 635 7498

Fax: +41 44 635 7489

E-Mail: t.cacchione@psychologie.unizh.ch

3. Gravity and solidity in great apes: The table task

3.1 Abstract

Three Experiments modeled after infant studies were run on four great ape species to investigate their reasoning about solidity and gravity constraints. The aims were: (a) to find out if great apes are subject to gravity biased search or display sensitivity for object solidity, (b) to check for species differences and (c) to assess if a gravity hypothesis or more parsimonious explanations best account for failures observed. Results indicate that overall great apes tend to perform in accord with solidity principles, that ape species differ in terms of their performance and that the errors made are best explained by a gravity account.

Key words: object permanence, gravity error, naïve physics

3.2 Introduction

For Piaget, the major criterion to decide if representations are present in the developing mind of young infants is their capability to refer to an object which is not directly perceivable in the current situation. If an organism represents knowledge about a given entity in the world, it must be able to make inferences about it, even if it is not perceptually accessible. For this reason, Piaget's tests for object concepts involved invisible displacement as a foundational feature: The future location of an object must be inferred by mentally reconstructing the movement it undergoes after disappearing from view.

Various constraints such as gravity, inertia or solidity act upon a moving object, shaping its pathway through time and space. To locate an invisibly displaced object an infant must have some knowledge of how these constraints interact to determine the future trajectory of the object moving out of sight (Baillargeon, 2002, 2004; Berthier, DeBlois, Poirer, Novak, & Clifton, 2000; Butler, Berthier, & Clifton, 2002; Kim & Spelke, 1992, 99; Hood, 1995, 98; Hood, Carey, & Prasada, 2000; Hood, Santos & Fieselman, 2000; Hood, Cole-Davies, & Dias, 2003; Mash, Keen, & Berthier, 2003; Spelke, Breinlinger, Macomber, & Jacobson, 1992).

In a series of studies, Spelke, Breinlinger, Macomber, and Jacobson (1992) set out to investigate at what age infants understand how solidity and gravity constrain the path of an invisibly moving object. For this purpose they invented a task that is generally referred to as the table task and was later adopted by various researchers (e.g. Berthier, DeBlois, Poirer, Novak, & Clifton, 2000; Hauser, 2001; Hood, Carey, & Prasada, 2000; Hood, Santos & Fieselman, 2000; Hood, Cole-Davies, & Dias, 2003) In a solidity version of the task, infants were presented with a solid shelf placed above a stage floor. A screen was set in place and an object dropped behind it. The removal of the screen revealed the object either resting on top of the shelf (possible test event) or under the shelf on the stage floor (impossible test event because object solidity would prevent the object from travelling through the shelf). The idea is that a preference to look at the impossible test event suggests that infants detect the physical anomaly of the event and have therefore some knowledge of solidity constraints. To test for gravity, Spelke and colleagues presented the same task, but this time no shelf was in place: The impossible test event showed the object floating in mid-air, whilst it rested on the stage floor in the possible test event. Again, if infants have some knowledge of gravity constraints they should look longer at the unsupported yet stationary object. Using these tasks, Spelke et al. (1992) found no evidence that 3 to 4 month old infants expected the unsupported object to fall. At the same age, however, infants' reasoning seemed to be in accord with the core

principles of continuity and solidity, as a horizontal version of the task suggested. Spelke et al. (1992) claimed that the negative results support the “core knowledge thesis” and the assumption that, while knowledge about object solidity is innate, knowledge about gravity develops slowly with increasing experience.

A completely different picture presents itself when young children are tested with tasks involving overt action. Despite the precocious hints of knowledge reported by Spelke and colleagues, Hood (1995) did not find older children’s search behavior to be in accord with solidity. He presented 2 to 3-year-olds with a tubes task in which the goal was to find a ball dropped down one of a set of three interwoven tubes landing in one of three hiding places. Hood identified an important phenomenon labelled “gravity bias”: Search errors occurred significantly more often at the hiding place directly underneath the release point than at other locations, suggesting some knowledge of gravity to be present in 2-year-olds. However, there was no evidence of knowledge about solidity at the same age, as children neglected the role of the solid tube shaping the path of the ball. Hood interprets the bias as stemming from a naïve folk theory of gravity which says that all unsupported objects fall in a straight line. In fact many researchers report similar (mis-) conceptions in preschoolers, pupils and in certain situations even in adults (Bliss, Ogborn, & Whitelock, 1989; Kaiser, Profitt, & McCloskey, 1985; Krist 2000). However, there is no consensus as to why the straight down belief occurs and explanations range from perceptual illusion to inhibition failure.

The table task adopted by Spelke et al. (1992) differs in many ways from the tubes task used by Hood (1995). A possible problem of the tubes task may be that it confounds multiple principles such as solidity, gravity and straight trajectories. To directly compare performance, Hood, Carey and Prasada (2000) tested knowledge of solidity and support in 4 search tasks, adapted from the looking time studies of Spelke et al. 1992. In one of the tasks children saw a toy being dropped behind a screen and were then asked to search for it. In a solidity condition a shelf was inserted above the stage floor, in a support condition the shelf was removed. In both conditions search behavior was recorded (search in the upper or in the lower location). Results showed that until 2;6 years of age, neither solidity nor support guides search behavior. No systematic search error was found in this task context, instead children tended to search for the object where they saw it during familiarization. Therefore, the results are in conflict not only between looking and search tasks but also within different types of search tasks. In the table task no gravity error appears, even though search performance is not correct.

In recent years the development of knowledge about physical object properties has also been addressed using a comparative perspective. Various studies have investigated reasoning about gravity and solidity constraints in different non-human primate species (Cacchione, & Krist, 2004; Hauser, 2001; Hauser, Williams, Kralik, & Moskovitz, 2001; Hood, Hauser, Anderson, & Santos, 1999) as well as in other non-human animals (Osthaus, Slater, & Lea, 2003). To address the question of the development of object knowledge from an evolutionary standpoint adds an important new perspective, as both dimensions of organism and environment are included. For example, if knowledge about solidity (in contrast to knowledge about gravity) is innate, as Spelke (1994) claims; non-human primates should also be very likely demonstrate such knowledge and aspects of ancestry may be important (i.e., cross species comparisons). On the other hand, if specific experiences lead to learning about physical object properties, then environmental factors might better account for differences found.

Various studies have found a robust gravity bias in non-human primates. (Hauser, 2001; Hauser et al., 2001; Hood et al., 1999). Hood et al. (1999) presented adult cotton-top tamarins (*Saguinus oedipus oedipus*) with a version of the original tube-task. Most tamarins failed to solve the task, but as in children, search was not at random: they too seem predisposed to search erroneously in the gravitational direction, that is, in the box directly underneath the release point. Osthaus, Slater and Stephen (2003) ran the tubes task on dogs (*Canis lupus familiaris*). Dogs, like human infants and tamarins were biased to search the box in gravity direction. There was some evidence that the strength of the gravity bias might differ across species: although dogs did not understand the mechanism of the tube, they were able to overcome the bias and to learn the location of the object.

In series of studies Hauser (2001), also tested for gravity biased search in the context of the table task. He presented rhesus monkeys (*Macaca mulatta*) with a table and two containers, one on top of the table shelf and one on the ground. Then the table was covered by an occluder and a food reward was dropped behind it. After the removal of the occluder the monkey was allowed to search for the reward. In contrast to human children, a strong gravity bias was identified in this context: the monkeys consistently searched in the box below the table. Santos and colleagues report dissociations between looking and action responses in rhesus macaques (*Macaca mulatta*), similar to those found in human infants (Santos & Hauser, 2002; Santos, Seelig, & Hauser, 2006).

In sum, the present state of research allows no consistent account for the development of knowledge of gravity and solidity constraints. First of all, evidence from perceptual and

action measures contradict one other and this seems to be true for both human and non-human primates. This dissociation of findings between action and looking studies is not just observable for solidity, but is manifest in various fields of developmental psychology and has led to heated methodological debate (e.g. Bogartz, Shinskey, & Shilling 2000; Cashon & Cohen, 2000; Haith, 1998). Additionally, in the case of solidity and gravity, findings between different action tasks differ, at least in human children. Why do intuitions about solidity measurably guide expectations only in looking times but not in action? Why does a gravity bias appear robustly in one search context but not in another? What is the reason for species differences in the existence and strengths of this bias? There appear to be multiple factors at work; task-specific, age-specific and species-specific - causing different response patterns.

In the present study we want to address the question of how these different factors come together to shape a behavioral response in one of the above presented gravity and solidity tasks: the table task. We use the table task because it is less complex than the tubes task and directly modelled after the looking tasks used in infancy (Spelke et al. 1992). We want to address this question from a comparative perspective testing four great ape species: gorilla, orangutan, chimpanzee, and bonobo. By doing this we add to the existing body of comparative data evidence from another non-human primate species that has not previously been tested in this context. This will fill the gap between monkey and infant data and offer new important insights into the development of gravity and solidity structures. First, we want to find out if great apes show a gravity bias at all. Human children, old and new world monkeys and even dogs have been found to be (at least in some task contexts) subject to some degree of gravity biased search. While all species tested with the tubes task displayed a bias, this is not true for the table task: Only rhesus monkeys were gravity biased when tested with this task, children, although not performing correctly, did not display a preference for the lower location (they searched the location where the object was during familiarization). Will the behavior of apes on this task mirror that of old world monkeys, thus exhibiting a search preference for the lower location, or will it be more similar to the search behaviour of 2;5 year old human children? Or will the apes do as older human children and adults do, and take object solidity into account, correctly searching at the upper location?

Further we are interested to see if aspects of organism and environment interact with task specific features which can lead to gravity or solidity based responses in the context of the table task. Do great apes differ concerning their sensitivity to gravity and solidity constraints? Differences could be connected to ancestry or environmental factors. If ancestry is an important factor modelling physical reasoning, African apes might differ from Asian

apes. Between species differences might also arise if aspects of environment play a major role in the development of physical cognition. Povinelli and Cant (1995) for example, argue that cognitive structures (in their model termed aspects of self-conception), may evolve in large-bodied apes as mechanisms to cope with problems posed by the need to prevent injuries when locomoting through a highly arboreal habitat which is fragile compared to the agents' body size. Living in an arboreal habitat may influence sensitivity for gravity events because a highly arboreal animal might experience gravity as more salient (e.g. by experiencing specific proprioceptive feedback when moving their own body along the vertical plane). Therefore it is feasible that the degree of arboreality of a species could have an impact on the search behavior in a gravity situation. There is some evidence that this might be the case: The highly arboreal tamarins evidence a very strong perseverating bias to the gravity location in the tubes task (Hood et al., 1999; Hauser et al., 2001). Whether the more terrestrial rhesus monkey would have perseverated on their gravity bias over many trials was not tested (Hauser, 2001). However Osthaus et al. (2003), found that dogs who did not learn to solve the tubes task were able to overcome the bias when tested with a similar amount of trials.

Finally, we are also interested to see if learning effects can be found over trials. Hauser (2003) argues that tamarins and rhesus monkeys (in contrast to human infants where the bias is only a transitory developmental phenomenon) can never acquire the necessary inhibitory skills to suppress gravity responses. He claims that they will cling to their straight down belief even in the face of massive contrary evidence, over myriads of trials. In fact, at least within the tubes task, monkeys were found to repeat this error despite never receiving positive feedback, i.e. never getting the reward. However, in the table task, only one trial was administered, so no conclusion is at hand. The studies of Osthaus et al. (2003) make clear that the bias is probably not as strong as often stressed, at least not in all species. So, we want to test if a potential bias is made across many trials or if search behavior improves with trials.

In the present study we test for different gravitationally relevant task factors and aim to gain a better insight into whether performance biases are best explained by a gravity account or if more parsimonious explanations will suffice. Hood (1995, 1998) mentioned two alternative explanations for the biased search behaviour observed: the infants may simply have searched in the closest location from the release point (proximity strategy) or searched in the aligned location (alignment account). Both strategies could account for the findings without referring to knowledge about gravity. In the case of the table task (where biased search appears to be under a solid shelf), Southgate (2004) mentions another alternative account in which the monkey might prefer to search a food reward in a sheltered location

(Karin-D'Arcy & Povinelli, 2002). In the following section these four conflicting hypotheses are briefly discussed.

The *Gravity hypothesis* claims that the bias is caused by a naïve theory that all unsupported objects fall in a straight line. Both Hood and Hauser favour this account. The idea is that in everyday interaction with the physical environment there is a vast exposure to gravity events. In the course of development sensitivity for statistical regularities in context of gravity develops. This sensitivity leads to specific expectations about the future behavior of (unseen) objects and culminates in a response bias: a hardwired prepotent response that has to be inhibited when inappropriate. As a consequence of inhibition problems, the bias leads to incorrect gravity answers in specific situations. In support of this account is the fact that the gravity bias gets weaker if aspects of spatial dimension are varied: Hood found 2-year-olds responses were significantly less biased when he presented the falling motion in reverse so that the object travelled through the tube in upward direction (Hood, 1998). Additionally, Hood, Santos and Fieselman (2000) did not find a similar bias on the horizontal plane, even though children's overall search performance was not better. However, it is doubtful as to whether the findings from upward trials really can be interpreted as supporting the theory account. Hood (1998) presented both the upward trials and downward trials as films on TV monitor and found that children were less biased than in the original search task on the downward trials, although still to a greater extent than in the upward version. This weakening of the bias can therefore be explained by the different mode of presentation. In fact Southgate (2004) found the error also in upward direction, if real events were presented.

Hauser (2001) also offers some support for a gravity account. He ran two additional versions of the table task where he manipulated gravitationally relevant factors to see if this influenced search: Firstly, he shifted the goal box away from the vertical trajectory and secondly, he presented the task along the horizontal plane. He found that a great deviation from the vertical falling line reliably improved the search performance, whereas a slight deviation did not alter performance. Also in case of horizontal displacement monkeys reliably performed the task correctly. These results offer strong evidence that erroneous search is tied to the vertical dimension, which supports a gravity explanation. These findings are further supported by results from a horizontal version of the tubes task (Hauser et al, 2001), where performance improvement in case of horizontal displacement was also found.

However, there is no unified evidence that the modification of dimensional aspects alter search performance in this paradigm. Similar tasks with human infants where object displacement was along the horizontal trajectory found no evidence of performance

improvements (Berthier, DeBlois, Poirer, Novak, & Clifton, 2000; Hood, Santos, & Fieselman, 2000; Mash, Keen, & Berthier, 2003; Frye, Zelazo, & Palfai, 1995). In one case similar perseveration in horizontal plane displacements was even reported (Frye et al., 1995). Also, Santos, Seelig and Hauser reported a proximity bias in tamarins tested with horizontal displacement. However, most of these studies are not directly comparable to the table task or used a more complex version of the task where the object had to be found in one of four possible locations, or where the object rolled down an inclined ramp. It is possible that due to the more complex setting performance deteriorated. These tasks are therefore not directly comparable to the Hauser (2001) study.

The *Alignment hypothesis* claims that a general bias to search in the aligned (or straight line) location is responsible for the search patterns observed. The visible portion of an objects trajectory before its disappearance (in this task typically a straight line) is extrapolated to infer the objects' current location. That a similar bias was found in upward (Southgate, 2004) and in horizontal directions (Frye et al. 1995) supports this account, although findings on the horizontal plane contradict it Hauser (2001). In the present context the *Alignment hypothesis* cannot help to explain search behaviour because in the standard table task, both top and down location are in an aligned position.

The *Proximity hypothesis* claims that the biased search occurs because participants search in the closest location to the position where the object was last seen. In case of the tubes task this is the box directly underneath the release point. However, in context of the table task, this hypothesis must only be considered in the case of correct search: In both, the vertical and the horizontal versions of the table task, the correct search location is the nearer to the release point. So, instead of knowledge about solidity, a proximity bias might also lead to correct performance in this task. The study of Santos, Seelig, and Hauser (2006) offers some evidence that this might be the case. In a horizontal displacement task they found tamarins' search biased to the compartment closest to the location where the object was last seen.

The *Shelter hypothesis* claims that the biased search occurs because monkeys prefer to search for food in a sheltered location in order to avoid food competition from conspecifics or because of predation evolutionary pressures might have let to a preference to forage in sheltered locations. This explanation is mainly relevant when it comes to non human primates and predicts location preferences only in the case of the table task: In contrast to the tubes task, only the lower location is "sheltered" by the shelf above it. Correct performance on a

horizontal version and with deviation from the vertical (Hauser 2001) are a problem for this hypothesis.

In a series of three Experiments we test for different gravitationally relevant factors trying to entangle which of the above listed hypotheses best matches the findings obtained. First we will run the original vertical table task, to see if great apes have a preference for the down location (Experiment 1). Authors such as Hood predict that a preference for the down location emerges as a procedural response to seeing an object fall, therefore in two conditions we vary the length of the falling trajectory before the object disappears behind the screen to see if this influences the search performance. Both gravity and shelter hypotheses predict a down location bias, whilst the *Proximity hypothesis* anticipates a preference for the upper location. Also, if participants understand that solidity constraints hinder the object from falling through the table they should choose the correct upper cup location. The variation of the length of falling trajectory helps to decide which of the explanations best account for findings: Only the gravity account predicts differences in performance after witnessing a long or a short falling trajectory respectively.

Next we investigate whether findings obtained in Experiment 1 are restricted to vertical object displacement. In Experiment 2 object displacement is shifted to the horizontal plane. In two conditions participants witness either a long or a short trajectory before the object moves behind the screen. If the performance in Experiment 1 is linked to sensitivity for gravity, performance should be generally better in case of horizontal displacement. Also, on the horizontal plane, variation of the length of perceived object trajectory should not influence search performance. Whilst the *Shelter hypothesis* would predict a back cup search preference, both gravity and proximity account expect search in the front cup. However, only in case of the gravity account there should be performance differences between vertical and horizontal plane displacements.

In Experiment 3 we test whether the degree of deviation from the vertical trajectory influences search behavior. If search performance is due to sensitivity to gravity, then a deviation of the down location from the vertical should positively influence search behavior leading to more searches in the upper location. The Shelter hypothesis predicts search in the down location, whilst the proximity and gravity accounts anticipate search in the upper location. However, only a gravity account could explain why differences in *degree* of deviation lead to performance differences.

3.3 General methods

Participants

We tested four species of great ape: Ten gorillas, eleven orangutans, six chimpanzees, and five bonobos (Table 1). Chimpanzees and bonobos are treated as a single subgroup. Five gorillas and three orangutans are housed in the Zurich Zoo, Switzerland. All other apes live in the Leipzig Zoo, Germany. The Leipzig group had a history of experience with various experiments concerning physical cognition. The Zurich group was completely naïve to experimental testing. Three additional orangutans and two additional gorillas from Zurich had to be excluded from the sample due to non-attendance.

Table 1. Species, age, sex, birthplace and rearing history of the apes participating in the in the Experiments 1-3.

| Participant | Species | Age (years) | Sex | Birthplace | Rearing history |
|-------------|------------|-------------|-----|-----------------------|---------------------------------|
| Azizi | Gorilla | 5 | m | Zurich | Mother |
| Bebe | Gorilla | 25 | f | Cameroon | Mother (wild born) |
| Binga | Gorilla | 4 | m | Zurich | Mother |
| Bonsenga | Gorilla | 4 | m | Zurich | Mother |
| Gorgo | Gorilla | 24 | m | Krefeld | Hand reared |
| N'Diki | Gorilla | 27 | f | Cameroon | Mother (wild born) |
| N'Gola | Gorilla | 28 | m | Jersey | Mother |
| N'Yokumi | Gorilla | 4 | f | Arnheim | Foster mother/Hand reared |
| Ruby | Gorilla | 7 | f | Arnheim | Hand reared |
| Viringika | Gorilla | 9 | f | Zurich | Mother |
| Bimbo | Orangutan | 25 | m | Duisburg | Hand reared |
| Dokana | Orangutan | 16 | f | Dresden | Mother |
| Dunja | Orangutan | 32 | f | Berlin | Mother |
| Oceh | Orangutan | 17 | f | Zurich | Mother |
| Padana | Orangutan | 8 | f | Leipzig | Mother |
| Pini | Orangutan | 17 | f | Leipzig | Mother |
| Salih | Orangutan | 13 | f | Zurich | Mother |
| Toba | Orangutan | 11 | f | Leipzig | Mother |
| Walter | Orangutan | 16 | m | Frankfurt a. M. | Mother |
| Xira | Orangutan | 8 | f | Zurich | Mother |
| Zora | Orangutan | 15 | f | München | Hand reared |
| Joey | Bonobo | 23 | m | Antwerpen | Hand reared |
| Kuno | Bonobo | 9 | m | Stuttgart | Hand reared |
| Limbuko | Bonobo | 10 | m | Stuttgart | Hand reared |
| Ulindi | Bonobo | 12 | f | Frankfurt a. M. | Mother |
| Yasa | Bonobo | 8 | f | Warwickshire | Mother |
| Alex | Chimpanzee | 4 | m | Plaisance du Touch | Hand reared |
| Alexandra | Chimpanzee | 6 | f | Rijswijk | Hand reared |
| Annette | Chimpanzee | 6 | f | Rijswijk | Hand reared |
| Fifi | Chimpanzee | 12 | f | Rijswijk | Mother |
| Jahaga | Chimpanzee | 12 | f | Rijswijk | Mother |
| Trudi | Chimpanzee | 12 | f | Rijswijk | Mother |

Materials

The apparatus was placed on a table like plastic testing surface. It consisted of a grey rectangular four-legged plastic table (height 23.8 cm/ length 42.9 cm/ depth 33 cm) and two oblong blue plastic cups (height 13.5 cm/ diameter 8.5 cm). One cup was placed on the tabletop, the other beneath it.

Grapes served as the object in the displacement events. Some apes could not be tested with grapes due to dietary reasons; therefore slices of banana were used in these cases.

Design

We ran a total of 3 Experiments. In Experiments 1 and 3, objects were invisibly displaced along the vertical plane, in Experiment 2 object displacement occurred along the horizontal plane. Each participant started with one of the vertical Experiments, moved on to the horizontal Experiment, and was finally run on the remaining vertical Experiment.

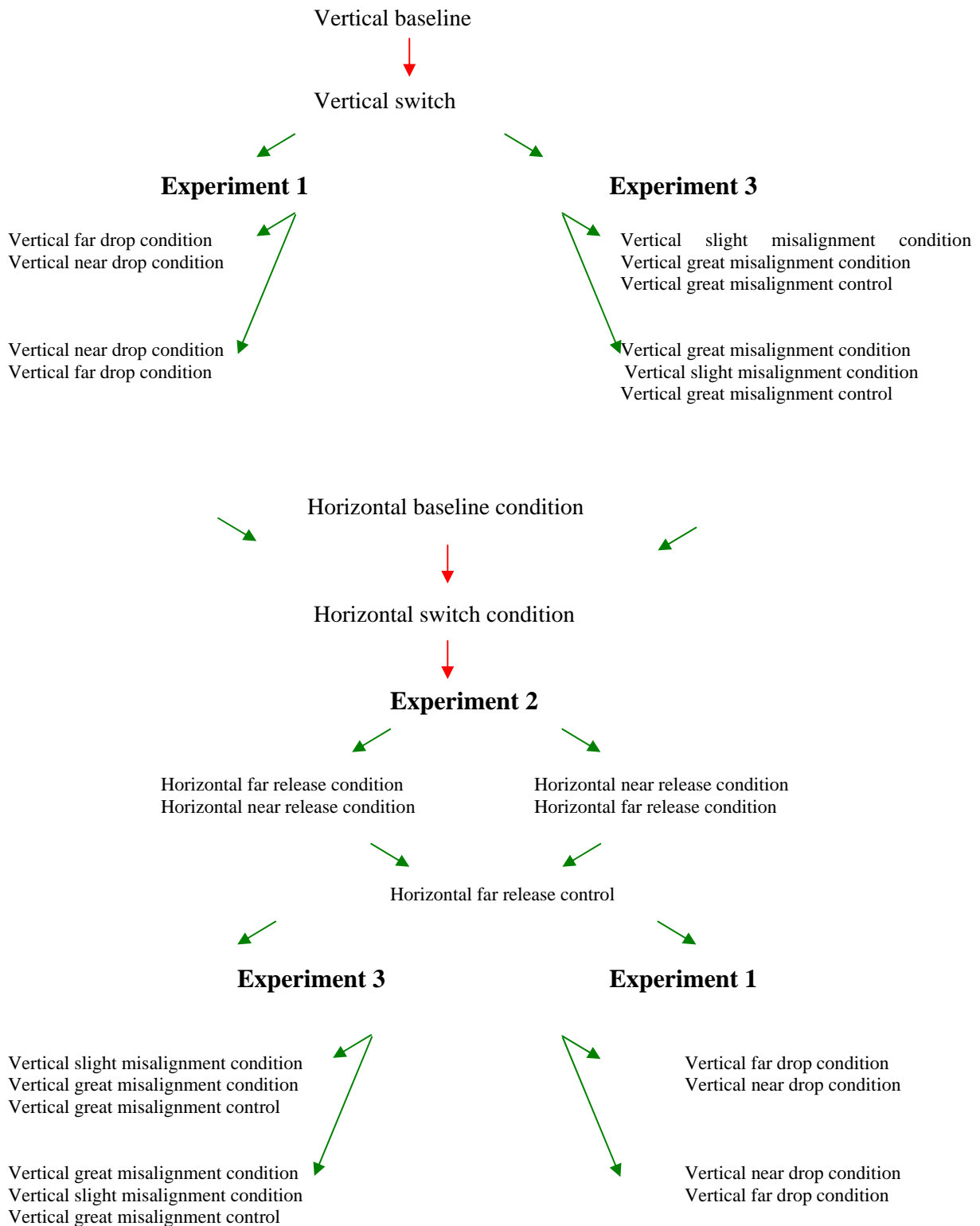
First, two baseline conditions involving vertical displacements were administered. The *vertical baseline condition* aimed at excluding potential preferences for one of the two cup positions. The *vertical switch condition* tested whether the participants are able to switch to a new location after finding the object repeatedly in the same location. After completing baseline conditions, half the participants were assigned to Experiment 1, and half to Experiment 3, respectively. The reason to counterbalance between the two vertical Experiments was to control for learning effects: As the apes encounter a similar situation in diverse test conditions and continuously get feedback about the actual location of the object, it is very likely that they will learn in the course of the study. Both, Experiments 1 and 3 consisted of two test conditions and Experiment 3 additionally included a control condition. Participants were counterbalanced across test conditions: In Experiment 1 the apes either witnessed a long object trajectory before the object moved behind a screen (*vertical far drop condition*) or a short one (*vertical near drop condition*). In Experiment 3 the cups were shifted from the original vertically aligned arrangement, resulting either in a minor deviation (*vertical slight misalignment condition*) or in a substantial deviation (*vertical great misalignment condition*). The *misalignment control condition* checked for side biases.

Second, all apes were run on two horizontal baseline conditions. Analogous to the vertical baselines, the apes now had to pass a *horizontal baseline condition* and a *horizontal switch condition* before moving on to Experiment 2. Experiment 2 consisted of two test

conditions and a control condition. Again, participants were counterbalanced across test conditions analogous to those of Experiment 1 but now transferred to a horizontal setting: In the *horizontal far release condition* the object travels along a long trajectory before disappearing behind a screen; in the *horizontal near release condition* it travels along a short one.

Finally, those participants who were first run on Experiment 1 now passed on to Experiment 3 and vice versa. Figure 1 illustrates the study design. Below, Experiments 1 to 3 are presented in a sequence. To simplify matters, procedure and results of the vertical baseline conditions are outlined together with Experiment 1.

Figure 1



General Procedure

The apes were tested alone either in an indoor observation room or in their sleeping room. Mothers with children younger than three years of age were tested in company of their

offspring. Some of the Zurich gorillas could not be separated from the group; they were tested in groups of two to four individuals, each was in turn lured to the test place while a keeper kept the others occupied.

All conditions were run by an experimenter who presented the stimuli. Both baseline conditions mounted up to ten trials. In the test conditions the apes were tested until they reached a criterion: If they were correct on the first choice in four out of five trials, the session was terminated. Those who did not reach this criterion were tested up to a maximum of ten trials.

The ape sat behind a mesh panel. A testing surface (slide table) was fixed by a metal frame directly underneath the mesh panel and the apparatus placed on top of it. An initial series of experiments in Leipzig was done using a wooden surface. Later in the course of the study the old wooden testing surfaces were replaced by the plastic slide tables described above. Facilities in Zurich did not enable the same sliding surfaces to be mounted so a metal trolley was used instead. The experimenter sat in front of the slide table which was constructed such that its surface could be shifted back and forth. The experimenter pulled the table back to present the displacement events (dropping/rolling of object), then she pushed it to the mesh to allow the ape to search for the food item until found by pointing out the cup location where it expected the food to be hidden. The search history was directly recorded on a sheet and all trials were videotaped.

General Analyses

All Analyses are made on a group level for apes in general. In further analyses chimpanzees and bonobos are treated as a single group and compared to gorilla and orangutan groups. As we are interested in ancestry factors, we also analyze subgroups of African apes (gorilla, bonobos and chimpanzees) versus Asian apes (orangutans). To test for arboreality we analyze the following subgroups: gorillas (low arboreality), chimpanzees and bonobos (medium arboreality) and orangutans (high arboreality). To test for learning effects we compare performance in the 1st and the 4th trials, as this enables us to include all apes in the sample (given that correct performance in 4 out of 5 trials is our criterion to end a test session).

We now outline the procedure and results of Experiments 1 to 3 in sequence; an in-depth examination of test order effects follows Experiment 3.

3.4 Experiment 1

Experiment 1 sets out to investigate the following questions: First, we are interested to find out if a gravity bias can also be observed in great apes, and if so, how persistently this influences search behavior across trials. Second, we would like to address whether the observation of the falling process intensifies this tendency. Does the perceived trajectory of an object falling straight down trigger gravity biased search or increase it? Does search performance improve, if the vertical fall is not or only partially observable? Finally, we are interested if all great apes perform similarly or if species differences can be observed.

To test this, we administer two test conditions where participants either witness a long or a short vertical fall before they can search the object. Both the gravity and shelter hypotheses predict a preference for the down location, whilst the Proximity hypothesis anticipates a preference for the upper location. Independent of the length of falling line the down location is always the sheltered one and the upper location always the one, next to the position where the object was last seen. However, the Gravity hypothesis expects a differential influence of perceived falling line: The perception of a long salient falling line might trigger gravity responses; in this case performance should deteriorate more after witnessing a long than a short object fall.

3.4.1 Method

Procedure

All apes first had to pass two baseline conditions: The *vertical baseline condition* and the *vertical switch condition*. They served to familiarize participants with the apparatus, to exclude baseline preferences for specific cup locations, and to exclude inhibitory problems influencing the search behavior.

Vertical baseline condition

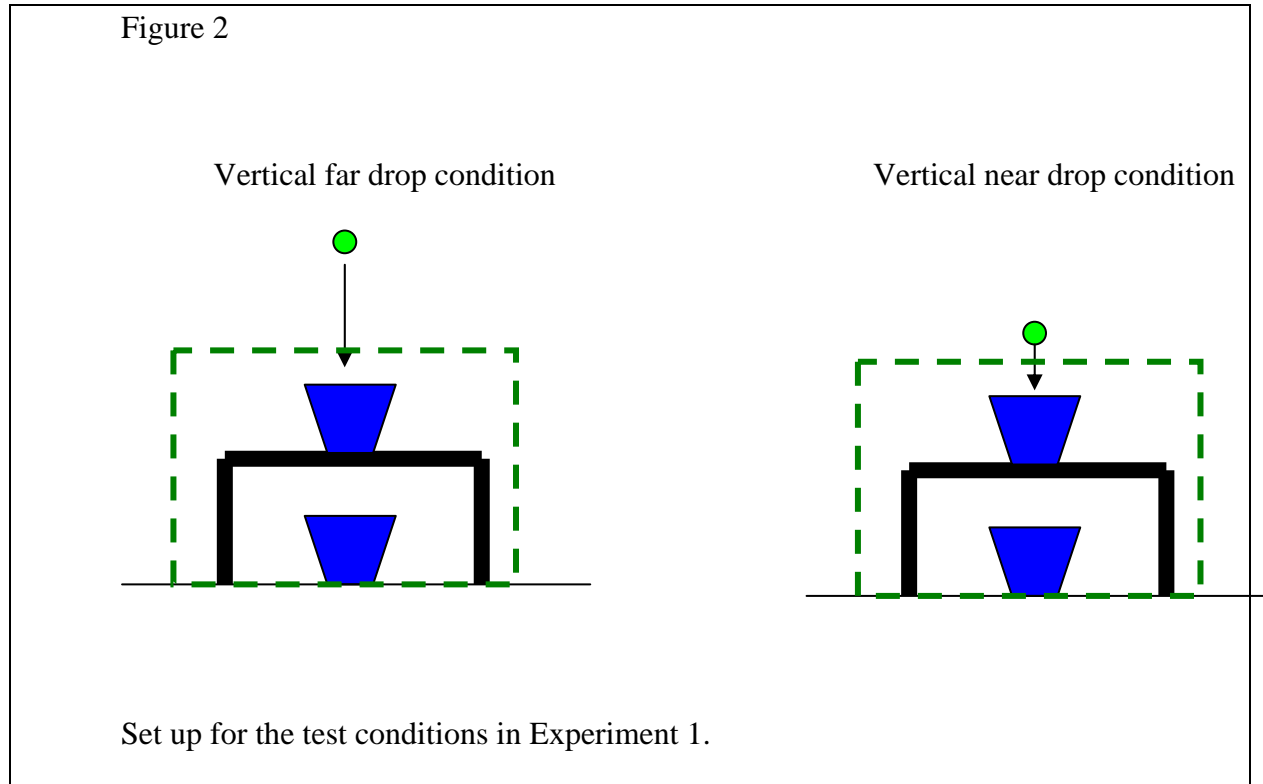
The apparatus was set up in front of the apes. The table was placed on the testing surface: One cup was positioned on top of the table, the other beneath it. The experimenter held out a grape and in full view of the ape placed it in one of the cups. Then the apparatus was pushed to the mesh. In order to get the food reward, the ape had to correctly point out the cup containing it. In ten trials the food item was placed randomly in both cups. If the search was correct on nine out of ten trials the participant moved on to the *vertical switch condition*. If not, the *vertical baseline condition* was repeated.

Vertical switch condition

The ape was presented with the same situation. Now, during the first five trials, the grape was placed in the same cup. In the last five trials it was again placed randomly in either location. If the ape choose correctly in the sixth trial (and could therefore suppress the tendency to select the location previously rewarded) it moved on to the vertical test conditions of either Experiment 1 or 3. If it failed the *vertical switch condition* was repeated.

Vertical far drop condition

Again, the table was placed on the testing surface, one cup on top, one beneath it (see Figure 2). Now a screen was raised and the experimenter waved the grape 25 cm above the screen. If the ape tracked it, the experimenter released the food. The screen was removed and the apparatus pushed to the mesh. The apes were allowed to indicate their choice by pointing to the cup in which they expected the food to be. Once the choice was made, the experimenter pulled back the apparatus. If they selected correctly, the experimenter took out the food reward and gave it to the ape. If they were wrong, she picked up the selected cup and bent it to the ape, so it could see it was empty. Then the ape was allowed to choose again. This procedure was repeated until the apes located the food item or abandoned search.



Vertical near drop condition

The same procedure was used as in the *vertical far drop condition* except that the grape was dropped directly above the top of the screen.

3.4.2 Results

First encounter

Within the first 5 trials of first encounter with the vertical tasks of Experiment 1 56.3 % of the apes reached criterion (4 out of five correct, n. s.), and 84.4 % within 10 trials ($\chi^2 = 15.125, p < .001$).

Figure 5

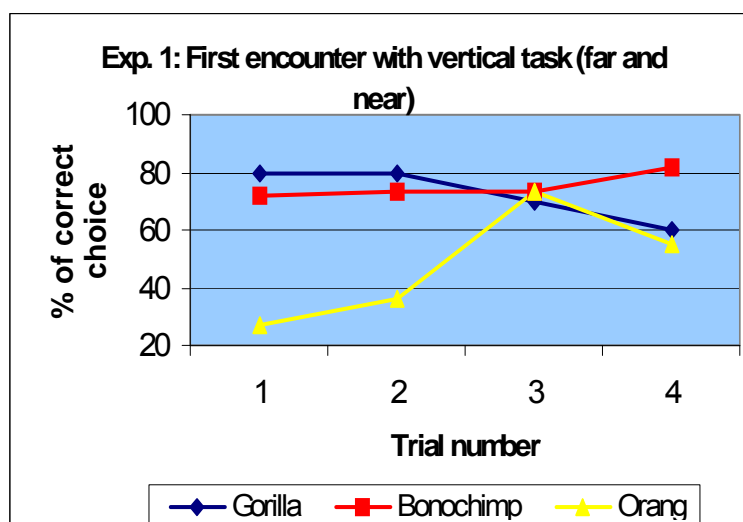


Figure 5 illustrates performance in the first encounter (including both far and near condition trials) with the vertical task in Experiment 1 for separate species. To check for gravity biased search, performance on the first trial of the first encounter with Experiment 1 was examined across all apes. A chi square analysis revealed no significant preference for a location on the very first test trial ($\chi^2 = 1.125, p = .289$) with 19 of 32 apes choosing the correct upper location. Performance improved on the fourth trial (21 of 32 choose correctly) but this tendency did not reach statistical significance ($\chi^2 = 3.125, p = .077$). This improvement was due to orangutans' performance: In the first trial they choose the correct upper location less often than other species, but across multiple trials their performance became more similar to that of gorillas and bonobo-chimpanzees.

A chi square analysis confirms this species difference: There was a significant ancestry group difference with African apes (gorillas and bonobo-chimpanzees) choosing the correct upper location more often in the first trial ($\chi^2 = 7.161, p < .05$), but by the fourth trial this group difference had all but disappeared ($\chi^2 = .912, p = .284$). This was not a mere aggregation effect, as other species combinations involving orangutans revealed no significant differences (e.g. gorilla and orangutan versus bonobo-chimpanzee, $\chi^2 = 1.239, p = .233$; bonobo-chimpanzee and orangutan versus gorillas, $\chi^2 = 2.565, p = .111$). Single chi square analyses were run for each subgroup alone and confirmed the above findings (see Table 2): African apes performed correctly on both the first ($\chi^2 = 5.762, p < .05$) and fourth trials ($\chi^2 = 3.857, p < .05$) of their first encounter with Experiment 1, whilst Asian apes' performance was at chance level on both first and fourth trials. Again, this was not due to an aggregation effect as the combination of gorilla and orangutan as well as the combination of bonobo-chimpanzee and orangutan failed to reveal above chance performance.

Vertical far condition

Within the first 5 trials of the vertical far drop condition 65.6 % of the apes reached criterion (4 out of five correct; marginally significant: $\chi^2 = 3.125, p = .077$), and 84.4 % within 10 trials ($\chi^2 = 15.125, p < .001$).

In the vertical far condition apes had no preference for either location in trial 1 ($\chi^2 = 1.125, p = .289$) with only 19 of 32 apes choosing the correct location, however by the 4th trial, 23 of 32 apes choose the correct search location ($\chi^2 = 6.125, p < .05$), indicating a learning effect.

There was no significant performance difference in arboreality or ancestry subgroups. However, if single chi square analyses were run for each subgroup alone differences were observable (see Table 2). Whilst performance of all arboreality subgroups in both trials 1 and 4 did not reach significance (with the exception that gorillas performed almost at significance on trial 4: $\chi^2 = 3.600, p = .058$), the performance of ancestry subgroups did differ: Performance of African apes mirrored that of apes in general (n. s. in trial 1 ($\chi^2 = 2.333, p = .127$) and correct in trial 4 ($\chi^2 = 5.762, p < .05$)), however Asian apes' search in both the first and the fourth trials was at chance level (t1: $\chi^2 = .091, p = .763$; t4: $\chi^2 = .818, p = .366$). This was not a mere aggregation effect as other species combinations did not lead to significance, with the exception of the combination of gorillas and orangutans in trial 4. However this finding can be attributed to the outstanding performance of gorillas in trial 4 of this task.

Vertical near condition

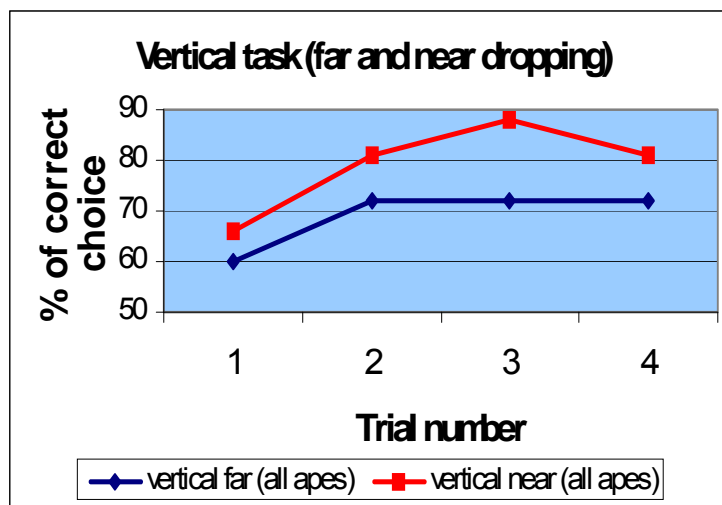
Within the first 5 trials of the vertical near drop condition 75 % of the apes reached criterion (4 out of five correct; $\chi^2 = 3.125, p < .01$), and 96.9 % within 10 trials ($\chi^2 = 28.125, p < .001$).

In the vertical near condition apes exhibited a non-significant trend to choose the correct location in trial 1 ($\chi^2 = 3.125, p = .077$) with 21 of 32 apes choosing correctly. Twenty-six of 32 apes searched the correct location in trial 4 ($\chi^2 = 12.500, p < .05$), indicating that, despite the performance improvement from trial 1 to 4, there were generally more correct searches in the vertical near condition than in the vertical far condition.

Crosstabs revealed no significant performance differences in arboreality or ancestry subgroups, but again, single chi square analysis did (see Table 2): Whilst the performance of Asian apes mirrors that of all apes (n. s. in trial 1 and correct in trial 4 ($\chi^2 = 4.445, p < .05$)), African apes' search is correct in the first ($\chi^2 = 3.857, p < .05$) and the fourth trials ($\chi^2 = 8.048, p < .01$). This was not a mere aggregation effect as other species combination did not lead to significance in both trials 1 and 4.

Figure 6 depicts performance in the far and near conditions. A non parametric Wilcoxon ranked test revealed no performance differences between far and near conditions for the percentage frequency of correct choices in the first four trials ($Z = -1.327, p = .184$) at the level of all apes. Nor were there differences between the two conditions if Wilcoxon tests were calculated for ancestry or arboreality subspecies alone. However, the results do indicate some differences between far and near conditions: more apes reached criterion in the near condition and in general they performed better in the near condition than in the far condition. Table 2 summarizes the findings (for trials 1 and 4): Apes in general tended to choose the correct location more often in the near condition than in the far condition. This was true for both the Asian and the African apes.

Figure 6



3.4.3 Discussion

Great apes performed quite well in this task, most of them reached criterion within 10 trials. They demonstrated no above chance preference for the lower location: in neither test condition, nor even in the first trial of the first encounter with the vertical task, was there a significant tendency to choose the lower location. Instead, the apes exhibited no significant preference for either of the two cups in the first trial in the first encounter or in the far test condition, and in fact showed a trend towards selecting the correct upper location in the near test condition. All apes choose the correct upper cup more often after witnessing a short falling line, whilst observing a long falling trajectory lead to more selections of the incorrect lower location. Additionally, apes reached criterion more often in the near test condition.

The responses of different ape species differed remarkably in the first trial of first encounter in Experiment 1: The African subspecies reliably choose the correct upper cup more often than did the Asian subspecies. However, even Asian apes who tended to select the upper cup location less often, demonstrated no significant preference for the down location in first encounter tests. Differences between Asian and African subspecies are also observable in both the far and near test conditions: In both cases, from the first trial onwards, African apes choose the correct upper location more often than Asian apes.

In both test conditions, and tendentially also in the first encounter with Experiment 1, the general performance improves from trial 1 to trial 4, with the apes reliably choosing the

correct location in the 4th trial; indicating that learning takes place over a few trials. This again suggests that even if a weak tendency to choose the lower location should be present in some species or individuals, this tendency does not reflect a resistant prepotent bias that inhibits learning.

It is not clear how the non-preference for neither cup should be interpreted. It could be that some apes simply have no idea of where the object might be and therefore initially choose both positions equally often. However, this possibility is somewhat at odds with the findings: There is no reason why the far and near drop conditions should lead to different responses if this were the case. Alternatively, it could also be that a lower position choice is due to sensitivity to gravity: Perhaps in principle the apes can infer that the object is in the upper cup (and are therefore to some extent aware of solidity constraints), but a prepotent gravity-response hinders them from directly choosing the upper location. Such a bias, however, is unlikely to be very robust and as a consequence both response possibilities initially occur. Only after some trials are the apes able to inhibit the gravity response in favour of the correct answer.

Thus, at a first glance none of the hypotheses outlined previously receive outright support from the present results. The gravity and shelter hypotheses are only partially supported, since there was no significant down location preference. Similarly, the *Proximity hypothesis* is not clearly supported as the apes did not consistently choose the upper location above chance level on the first trials. Although, one could argue that across trials the apes did not realize that solidity prevents the object from falling to the lowest point, but merely adopted a proximity strategy; a comparison of performance in the far and near conditions shows that this cannot be the case. A proximity strategy would lead to similar behavior in both conditions, as the object disappears in both cases at the same point from view. However the finding that the apes made more mistakes in the far condition goes against this possibility. One might still argue that the dropping hand is closer to the upper cup in the near condition, suggesting that it enhances the idea of the object being put directly into that cup. However even in the far drop condition the experimenter's hand is nearer to the upper cup than the lower - thus there is no reason why the ape should search more often in the lower now even more distant cup (which they do in the far drop condition). A more plausible explanation is that some sensitivity to gravity is present in apes and therefore looking at a longer falling trajectory triggers stronger gravity responses.

In sum, the results of Experiment 1 could be interpreted as showing that apes have some sensitivity to gravity, but they (at least the African apes) can in most cases inhibit this

tendency to search the lower location. At a group level, they are equally likely to respond with the correct or with the biased response during the first trials, but the capacity to inhibit this tendency increases when the event is repeated across several trials, resulting in correct search on the 4th trial. For reasons that remain unclear, orangutans seem to have more problems inhibiting this gravity response, and this may be associated with aspects of ancestry. However, both African and Asian apes seem to have more difficulties inhibiting the prepotent gravity response when watching a long fall prior to search. The observation that in many cases the apes erroneously directed their hand towards the lower location first, and then stopped to correct themselves before reaching to upper location, also supports the gravity account.

3.5 Experiment 2

Although most of the apes – at least in trial 4 – choose the correct upper cup; the findings of Experiment 1 suggest that in addition to sensitivity for solidity constraints, sensitivity to gravity may also influence the ape's behavior in this search task: apes searched more often at the lower location after witnessing a long vertical trajectory. What happens, if the events of Experiment 1 are presented in a horizontal version? Is the tendency to search at the down location after witnessing a long falling trajectory observable only in events where objects fall vertically? If this is found to be the case, then the results obtained in Experiment 1 can be interpreted as representing sensitivity to a gravitational force operating along the vertical plane. Alternatively, the results of Experiment 1 could be replicated in Experiment 2 in a horizontal setting.

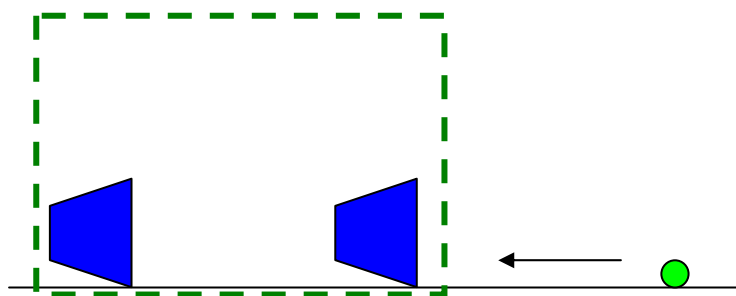
In two conditions participants witness either a long or a short horizontal trajectory before the object moves behind the screen. Does their performance generally improve? Does the perception of an object's trajectory have the same impact as in the vertical version even if the object is now moving along the horizontal plane? Or does this have no influence upon search behavior? The *Shelter hypothesis* predicts a back cup preference whilst both the gravity and proximity accounts anticipate search in the front cup. However, only in case of the gravity account are performance differences between vertical and horizontal plane displacements expected.

3.5.1 Method

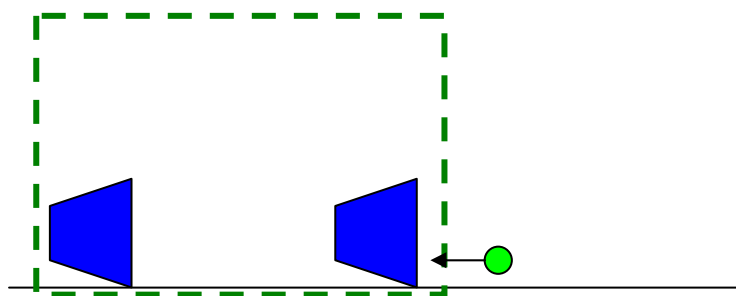
Procedure

In the horizontal events no table was used, instead the two cups were arranged successively along the horizontal dimension of the testing surface (see Figure 3). One cup was placed at the far end of the surface, with the other cup 19 cm away from it. The cups' openings faced the opposite end of the surface where the object was released. All horizontal events (test and baselines) were presented in two spatial arrangements: Half of the participants had the cup openings facing to the right side of the surface (hence the experimenter rolled the grape from the right side), and half to the left side. All participants were run on two baseline and two test conditions, followed by a control condition. To confirm that no side bias was liable for the observed behavior even on an individual level, participants were retested with the opposite cup arrangement.

Figure 3



Horizontal far release condition



Horizontal near release condition

Set up for the test conditions in Experiment 2.

As in the vertical plane, participants first had to pass two baseline conditions before moving on to the horizontal tests. Again, the rationale of the baselines was to introduce the horizontal setting and to control for baseline cup preferences as well as inhibitory limitations.

Horizontal baseline condition

To make sure that they were free from cup preferences, a grape was placed randomly in either cup. If the ape located it correctly in nine out of ten trials it moved on to the *horizontal switch condition*.

Horizontal switch condition

Similar to the vertical baselines, this condition controlled for perseveration after repeated enforcement in one cup location. The grape was first placed in the same cup on five consecutive trials, followed by five trials where it was randomly placed. If the ape correctly switched to the new location in the sixth trial it moved on to the test conditions.

Horizontal far release condition

The test conditions are depicted in Figure 3. The cups were arranged along the horizontal plane as in the baseline condition. A screen was raised and a grape held out to the ape. If he tracked the object, the experimenter brought the grape to the surface end opposite the cups and rolled it behind the screen (x cm from the first cup opening, x cm from the beginning of the screen). Then the screen was removed and the test surface pushed towards the mesh. After the ape made his choice by pointing to a cup location the experimenter pulled the table back. If the ape chose the correct cup, he got the reward. If he chose the wrong location, the experimenter lifted the cup and bent it towards the ape so he could ascertain himself that it was empty. Then the cup was placed in its original position and the ape was allowed to make a new choice. This was repeated until the ape found the grape or stopped searching.

Horizontal near release condition

The same procedure was used as in the horizontal far release condition, except that the grape was now released (start point of rolling) directly in front of the screen.

Horizontal far release control

To control for side bias within an individual, the far release test was repeated with the opposite spatial arrangement (cups right end, release left end versus cups left end, release right end).

3.5.2 Results

First encounter

Within the first 5 trials of first encounter with the horizontal tasks of Experiment 2 93.8 % of the apes reached criterion (4 out of five correct; $\chi^2 = 24.500$, $p < .001$), and 100 % within 10 trials.

The majority of apes searched the correct front location ($\chi^2 = 21.125$, $p < .001$) in the first trial of first encounter with Experiment 2 (including both near and far condition trials) with 29 of 32 apes choosing correctly. The same is true for the 4th trial ($\chi^2 = 18.000$, $p < .001$) with 28 of 32 choosing correctly. This was also true for both arboreality (gorilla, bonobo-chimpanzee and orangutan) and ancestry (African and Asian apes) subgroups (see Table 2), with the sole exception that gorillas performance on the 1st trial was only tendentially above chance ($\chi^2 = 3.600$, $p = .058$).

Table 2

| | VERTICAL DISPLACEMENT | | | | | | HORIZONTAL DISPLACEMENT | | | | | | |
|----------|-----------------------|---------------|----------|---------------|---------------|-------|-------------------------|------|-------------|------|--------------|------|------|
| | first encounter | | far drop | | near drop | | first encounter | | far release | | near release | | |
| | 1 | 4 | 1 | 4 | 1 | 4 | 1 | 4 | 1 | 4 | 1 | 4 | |
| ALL APES | n. s. | Trend .077 | n. s. | .05 | Trend .077 | .001 | .001 | .001 | .001 | .001 | .001 | .001 | .001 |
| AFRICAN | .05 | .05 | n. s. | .05 | .05 | .05 | .001 | .001 | .001 | .001 | .001 | .001 | .001 |
| ASIAN | n. s. | n. s. | n. s. | n. s. | n. s. | .05 | .01 | .01 | .01 | .05 | .01 | .01 | .01 |
| GORILLA | Trend .058 | n. s. | n. s. | Trend .058 | Trend .058 | n. s. | Trend .058 | .05 | .001 | .05 | n. s. | .05 | .05 |
| BONOCH | n. s. | .05 | n. s. | n. s. | n. s. | .001 | .001 | .05 | .001 | .05 | .001 | .001 | .001 |

Horizontal far and near release conditions

Within the first 5 trials of the horizontal far release condition 93.8 % of the apes reached criterion (4 out of five correct; $\chi^2 = 24.500$, $p < .001$), and 100 % within 10 trials.

In the near release condition, 96.9 % of the apes reached criterion (4 out of five correct) within the first 5 trials ($\chi^2 = 28.125, p < .001$) and 100 % within 10 trials.

Apes (31 of 32) search correctly in trials 1 ($\chi^2 = 28.125, p < .001, 31$) and 4 (27 of 32 are searching correct; $\chi^2 = 15.125, p < .001$) of the horizontal far condition. The same is true for the horizontal near condition in both trials 1 ($\chi^2 = 18.000, p < .001$) and 4 ($\chi^2 = 24.500, p < .001$) with 28 and 30 of 32 apes searching the correct location. Again, no differences were observed at the level of subgroups analyses (see Table 2), except for the gorillas who again did not reach a significant preference for the correct location in the first trial of the horizontal near test ($\chi^2 = 1.600, p = .206$). In accordance with the results above, non parametric Wilcoxon ranked tests revealed no differences in the frequency of correct choices (trials 1-4) between far and near release conditions ($z = -.758, p = .448$). Finally, there were no side preferences: apes performed correctly in the first trial of opposite side control condition ($\chi^2 = 18.000, p < .001$).

Further analyses were run to compare performances on vertical and horizontal test conditions: non parametric Wilcoxon ranked tests revealed significant differences in the frequency of correct choices (trials 1-4) between: (a) first encounter with horizontal conditions of Experiment 2 and vertical conditions of Experiment 1 ($z = -3.215, p < .01$), (b) horizontal far and vertical far conditions ($z = -2.898, p < .01$), and (c) horizontal near and vertical near conditions ($z = -2.621, p < .01$). In all cases, apes choose the upper location more often in the horizontal tests. All apes, in addition to arboreal and ancestry subgroups, performed better on horizontal tests, but differences between subgroups were observable nonetheless. Table 2 shows the performance on all vertical and horizontal conditions for trials 1 and 4. As African apes performed below chance only in the vertical far drop condition, their improvement between vertical and horizontal experiments was therefore not as pronounced as it was for apes in general, and particularly for the Asian subspecies who profited dramatically from the horizontal displacement experiment. Except for trial 4 of vertical near condition, the Asian apes never performed above chance on any of the vertical tests, but did perform reliably correctly on all horizontal tests. Comparisons between arboreality subgroups did not give such a clear picture, which was mainly due to large inconsistencies in the behavior of the gorillas. For example, given their general performance, it is unclear why they performed below chance on the 4th trials of both the vertical first encounter and near drop conditions, which indicates that their performance got worse across trials. It is also unclear how their below chance performance on the 1st trial of the horizontal near release condition should be interpreted.

3.5.3 Discussion

Performance is reliably better with horizontal displacement: In the first encounter with horizontal tasks of Experiment 2, as well as in both test conditions, apes generally performed better than in the vertical tests, with all apes reliably choosing the correct front cup in the first trial. All subspecies profited from the horizontal displacement, but African subspecies profited less than average whilst Asian species profited most. This difference is not surprising, given the performance of each subgroup in the vertical tasks of Experiment 1. Most importantly, and in contrast to the vertical tests, the length of the objects visible trajectory before it disappeared behind the screen had no impact on search behavior in horizontal displacements. This offers more evidence that search performance is somehow linked to the spatial dimension of object displacement and that search errors may reflect some sensitivity to the gravitational force operating along the vertical plane. The findings clearly do not support the *Shelter hypothesis*, which predicted a back cup search preference. One might argue that a proximity bias is responsible for the search in the front cup; however, a proximity account would anticipate the same performance on both vertical and horizontal trials. Only a gravity account can explain why performance reliably differs between vertical and horizontal plane displacements.

3.6 Experiment 3

The results of Experiment 2 suggest that search performance is linked to the spatial dimension of object displacement, i.e. horizontal or vertical. Experiment 3 investigates another gravitationally relevant factor by questioning the importance of the vertical alignment of the cups. What happens if the down-location cup is not placed straight under the dropping point, but shifted to the side? If search performance is due to sensitivity to gravity, then a deviation of the down location from the vertical should improve search performance. In two conditions the lower cup is shifted either slightly or greatly away from the vertical falling line. The *Shelter hypothesis* would predict search in the down location in both cases, and the *Proximity hypothesis* would contrarily expect search in the upper location in both conditions. However, only a gravity account, which anticipates greater performance improvements with greater deviations, could explain why differences in *degree* of deviation might lead to performance differences.

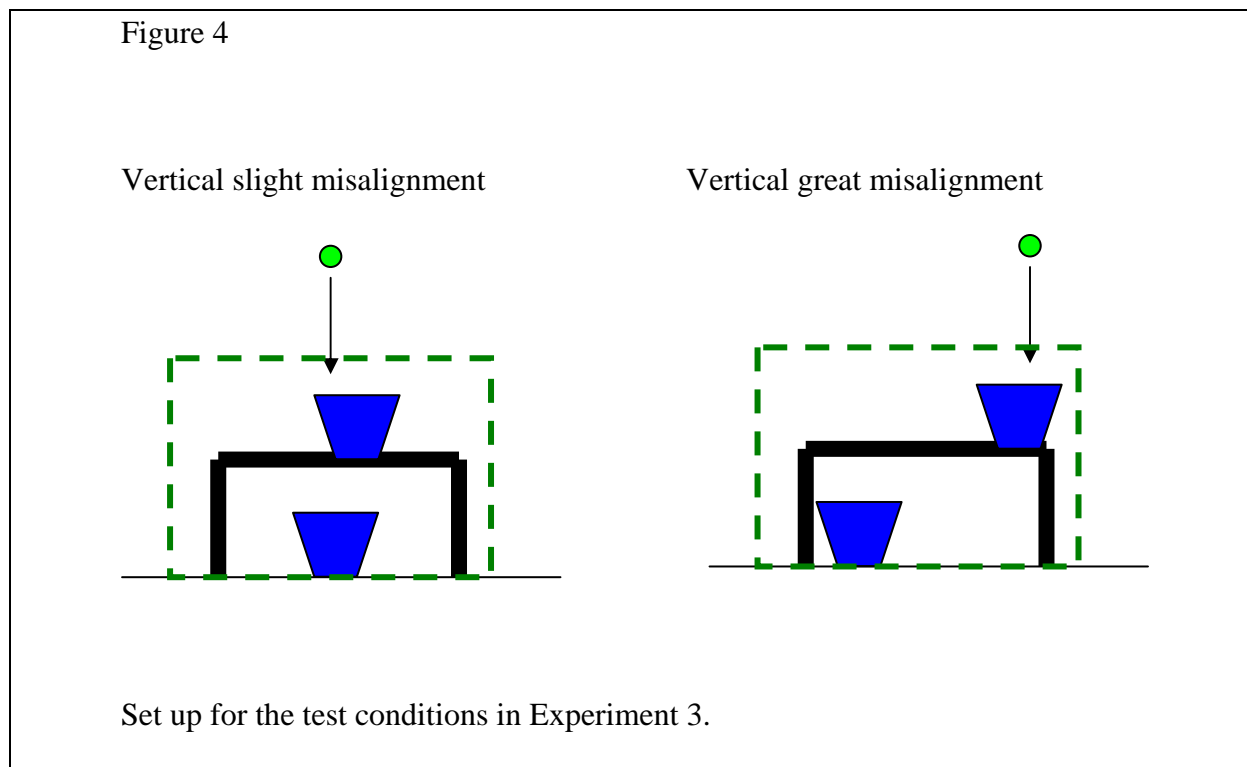
3.6.1 Method

Procedure

In Experiment 3 the apes were confronted with two test conditions and one control condition. Again, the control condition aimed at excluding side biases.

Vertical slight misalignment condition

As in Experiment 1, the table was placed on the testing surface, one cup on top, one beneath it (Figure 4). The setup and procedure was identical to the far drop condition of Experiment 1 except that the bottom cup was misaligned by 3 cm.



Vertical great misalignment condition

This condition was identical to the slight misalignment condition except that both cups were moved away from the table center and misaligned by 19 cm from each other. Half of the participants were presented with a top-left, down-right configuration, half of them with the opposite top-right, down-left configuration.

Vertical great misalignment control

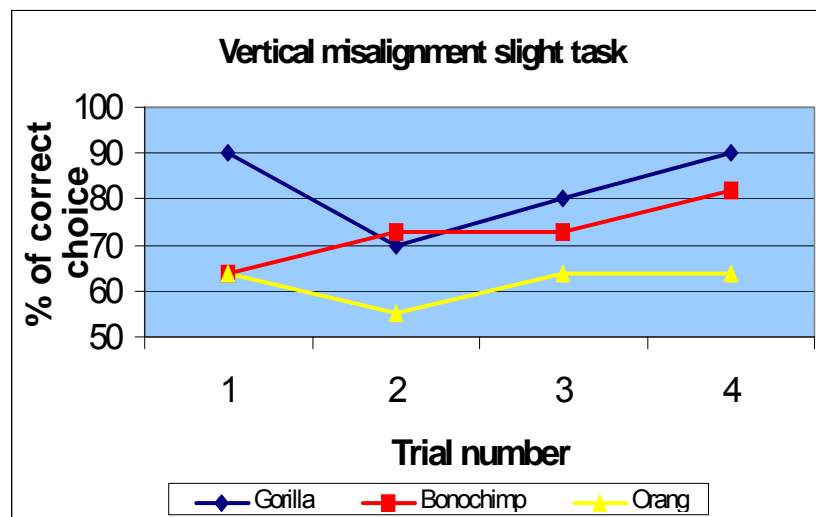
We retested all apes on the large misalignment condition but with the opposite cup configuration to check for side biases within participants.

3.6.2 Results

Within the first 5 trials of the misalignment slight condition 75 % of the apes reached criterion (4 out of five correct; $\chi^2 = 8.000$, $p < .01$), and 87.5 % within 10 trials ($\chi^2 = 18.000$, $p < .001$). In the misalignment great condition, 84.4 % of the apes reached criterion (4 out of five correct) within the first 5 trials ($\chi^2 = 15.125$, $p < .001$), and 93.8 % within 10 trials ($\chi^2 = 24.500$, $p < .001$).

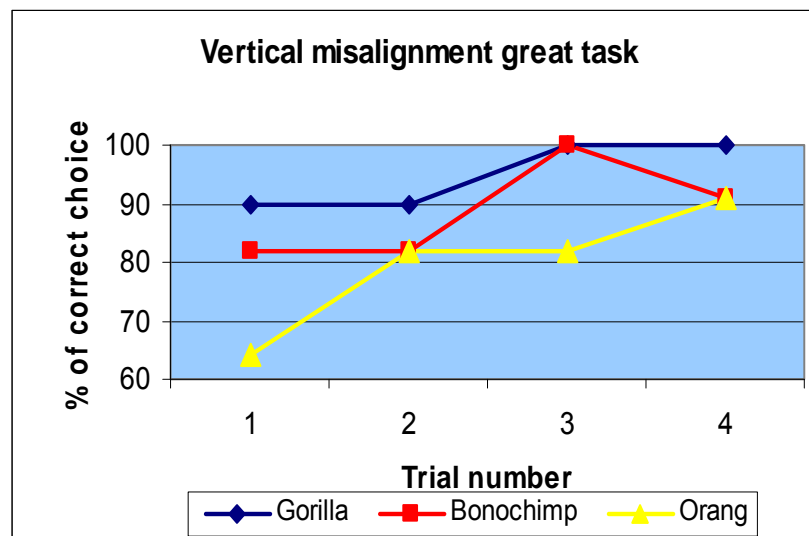
Apes (23 of 32) chose the correct upper cup on the first ($\chi^2 = 6.125$, $p < .05$) and the fourth trials (25 of 32, $\chi^2 = 10.125$, $p < .001$) of the misalignment slight condition (see Figure 7). Crosstabs revealed no differences between single species or between ancestry or arboreality subgroups.

Figure 7



The same pattern was found in the misalignment great condition (see Figure 8). Again all apes chose the correct upper cup in the first (25 of 32, $\chi^2 = 10.125$, $p < .001$) and the fourth trials (30 of 32, $\chi^2 = 24.500$, $p < .001$). Crosstabs revealed no differences between single species or ancestry and arboreality subgroups.

Figure 8



A non parametric Wilcoxon ranked tests revealed significant differences in the frequency of correct choices (trials 1-4) between slight and great deviation conditions ($z = -2.128, p < .05$). Wilcoxon tests run for ancestry and arboreality subgroups separately revealed a significant difference between performance in slight and great deviation conditions only for African apes who performed better in the great deviation condition ($z = -2.000, p < .05$) but not for Asian apes. Of the arboreality subgroups, only gorillas performed reliably in the great deviation condition ($z = -2.308, p < .05$).

Again we ran further single chi square analyses to check for species differences in both test conditions (see Table 3). As in the other Experiments, we found ancestry group differences with African apes performing above chance on both the first ($\chi^2 = 5.762, p < .05$) and the fourth trial ($\chi^2 = 10.714, p < .001$) of the misalignment slight condition and likewise in both the first ($\chi^2 = 10.714, p < .001$) and the fourth trials ($\chi^2 = 17.714, p < .001$) of the misalignment great condition. In contrast, Asian apes only performed above chance on the fourth trial of misalignment great condition. This does not seem to be a mere aggregation effect: Combinations of bonobo-chimpanzees and orangutans, or gorilla and orangutans, respectively, just mirror the performance of bonobo-chimpanzees or gorillas alone.

In case of the alignment slight condition, arboreality subgroups also differed from one another (see Table 3). Gorillas choose the correct location in both the first ($\chi^2 = 6.400, p < .05$) and fourth trials (all choose correct). Whereas bonobo-chimpanzees had no significant

preference in the first trial ($\chi^2 = .818, p = .366$), but performed correctly on the fourth trial ($\chi^2 = 4.455, p < .05$) of the slight deviation test. Finally, orangutans demonstrated no significant location preference in both trials 1 and 4. This indicates that there are not only differences between arboreality subgroups in the slight deviation condition, but also confirms the general difference between slight and great test deviation manipulations.

Table 3

| | VERTICAL DISPLACEMENT | | | | | |
|----------|-----------------------|-----------------------|---------------------|-------|--------------------|------|
| | vertical far drop | | misalignment slight | | misalignment great | |
| | 1 | 4 | 1 | 4 | 1 | 4 |
| ALL APES | n. s. | .05 | n. s. | .05 | .001 | .001 |
| AFRICAN | n. s. | .05 | .05 | .001 | .001 | .001 |
| ASIAN | n. s. | n. s. | n. s. | n. s. | n. s. | .01 |
| GORILLA | n. s. | Trend .058n. s. | .05 | .05 | .05 | .001 |
| BONUCH | n. s. | n. s. | n. s. | .05 | .05 | .01 |

Influence of test order in Experiments 1-3

A chi square analysis investigated test order effects². A reliable influence of test order was found in only one direction: Apes who participated first in Experiment 3 were found to have reliable effects on their behaviour in Experiment 1, whilst prior exposure to Experiment 1 did not lead to effects in performance on Experiment 3. Apes who were first run on Experiment 3 demonstrated a non-significant trend to perform better on the first trial of the first encounter with Experiment 1 ($\chi^2 = 3.802, p = .055$) and performed significantly better on trial 4 ($\chi^2 = 4.453, p < .05$) of Experiment 1. This effect was less pronounced in trial 1. Additionally, apes who were first run on Experiment 3 made significantly less errors on trial 4 of vertical near task ($\chi^2 = 5.744, p < .05$), however in the vertical far task there was no influence of test order. Therefore, prior experience with the vertical tasks of Experiment 3 helped to improve performance in Experiment 1, an effect only observable in first encounter and in the near test condition on trial 4.

² The gorilla data must be interpreted with special caution regarding test order effects because there was a selective drop out in the gorilla subgroup. Seven out of ten gorillas who were included in the final sample were run first on Experiment 1. Thus a comparably lower performance could be due to a lack of experience in comparison to the other groups tested. However, as gorillas performed rather better than other species we can rule out this possibility.

3.6.3 Discussion

Apes search correctly from the first trial on in both conditions, no matter whether there is a slight or a great deviation from the vertical falling line. In contrast to Experiment 1, where apes had problems with the far condition (both test conditions in Experiment 3 had a long drop), performance improves when the lower cup is at position that deviates from the vertical fall. This is the case even when the deviation is slight, as with the exception of the orangutans, all subspecies perform better in the slight misalignment condition than in the vertical far condition (see Table 3). The degree of deviation also appears to have had an influence on the search performance itself, with apes performing reliably better in great deviation trials. Again these findings speak against the *Shelter hypothesis* which predicts search in the down location in both experimental conditions. Although a proximity account would expect apes to search in the top location, which they correctly do, it fails to explain why findings differ between the vertical far task of Experiment 1 and the slight deviation condition of Experiment 3, and between the slight and great deviation conditions within Experiment 3. Only a gravity account, which anticipates performance improvements with deviation changes, can explain why a deviation from the vertical and also differences in *degree* of deviation might lead to performance differences.

Again we find some species differences, with African species generally performing better and apparently profiting more from a salient deviation from the vertical falling line. This is most pronounced in gorillas, followed by bonobo-chimpanzees, whilst orangutans tend to search at the down location more often, even in case of a great vertical deviation. If sensitivity to vertical gravity causes the tendency to search at the down location, one might conclude that African apes have fewer problems inhibiting gravity responses than Asian apes do, and that the more arboreal a species is, the greater its' difficulties are in suppressing gravity answers.

3.7 General discussion

Great apes do not show a gravity bias in the table task. In neither test condition were they found to show a reliable preference for the down location, not even on the first trial of their first encounter with the displacement of a vertical object in Experiment 1. This is true for

all four ape species tested. In this respect the great apes clearly differ from rhesus monkeys who were biased in an analogous task (Hauser, 2001). The successful performance of African apes can also be contrasted with that of 2:5 year old children who failed in this task even though they were not particularly biased to search in the incorrect down location: African apes correctly searched the upper cup already in the first trial of first encounter with Experiment 1. Although not all apes performed correctly on the first trial after witnessing a vertical drop, the overwhelming majority of them quickly learned to inhibit down location responses and to direct their search to the upper location. On the basis of this finding we conclude that, at least in this task context, a persistent prepotent gravity bias is not present in great apes. On the contrary, they are at least to some extent aware that solidity constraints guide the future motion of the object once it disappears from sight. Furthermore, they seem to infer that an object cannot move through a solid table and therefore search for the object in the upper cup.

This does not necessarily imply that sensitivity to the impact of gravity has no influence on the apes search behavior in this task. The frequency of errors observed was directly connected to factors associated with gravity: if various aspects of gravity were manipulated the frequency of errors decreased. For example, apes made fewer errors if the vertical fall was only partially visible. Similarly, a deviation of potential search locations from the vertically aligned falling trajectory helped to reduce search errors. Finally, search performance was reliably better with horizontal displacements than with vertical displacements. These findings strongly suggest that gravity has an impact on search behaviour, even though a gravity response can be inhibited in most cases.

Our findings also provide strong support for the *Gravity hypothesis* as an explanation for biased search behaviour. Both shelter and proximity hypotheses cannot account for the results as a whole, because they predict the same behaviour throughout all test conditions and fail to explain why error frequency differs between conditions. Only the gravity account can explain why error frequency differs after witnessing a short or a long vertical fall, when the degree of deviation from the vertical alignment is varied, or when vertical or horizontal displacement occurs.

In sum, it can be stated that apes appear to take into account both solidity and gravity constraints when searching for an invisibly displaced object. It is likely that the apes can principally infer that solidity prevents an object from moving through the table and that, as a consequence, the object can be found in the upper cup. Despite this, sensitivity to gravity constraints also influences search behaviour and under certain circumstances gravity

responses can prevail. Gravity manipulations such as the length of falling trajectory, degree of deviation from the vertical falling line and dimension of object displacement, may enhance the possibility of gravity responses. It is also important to note that gravity responses occurred most often in the initial trials, and that experience with the task across trials increased the apes' capacity to inhibit gravity answers. Finally, species differences also seem to be an important factor influencing the occurrence of gravity responses. All these factors interact in a test situation and may or may not turn the balance in favour of a gravity based response.

If our interpretation is correct, errors occur because gravity responses can not be suppressed. Why then do different ape species vary in their capacity to inhibit gravity responses? One possibility is that they differ in the extent to which they live in an arboreal habitat and therefore in the mode in which they experience gravity events. Species with a more arboreal habitat might experience gravity events more saliently and therefore develop a greater sensitivity for gravity constraints which could lead to defaults in the reported test situations. However, apart from weak effects found in Experiment 3, the degree of arboreality of the different subspecies tested here did not measurably influence their search behavior. Instead, the differences observed were more associated with ancestry. Orangutans were the species that was most frequently subject to gravity answers. They have a special position in the family of great apes and are commonly classified as belonging to the subfamily of Ponginae. In contrast gorillas, chimpanzees and bonobos are generally classified together with humans in the subfamily of Homininae. Orangutans diverged in an earlier period of evolutionary time than gorillas, chimpanzees and also humans which diverged from each other within in a very short time period (Glazko & Masatoshi, 2003; Stauffer, Walker, Ryder, Lyons-Weiler, & Blair Hedges, 2001). Thus it is possible that in the course of their divergent evolution great apes developed varying degrees of sensitivity for gravity events. The reason for this might still be associated with the specific environments they occupied: different habitats and living conditions call for specific adaptations which can be anatomical, behavioural and cognitive in nature.

In the face of our findings, is it valid to conclude that great apes have concepts of solidity (because they perform well on average) and gravity (because errors are associated with gravity)? We believe not. We find it hard to explain why an organism that is endowed with an *explicit concept of solidity* should be in any way context dependent in its reactions. If one has such a concept, there is no reason why witnessing a longer or shorter exposure to the trajectory of a falling object should influence one's judgement of events. In this respect, our findings mirror the present state of research which does not enable a consistent picture of

development to be drawn in this domain. It is clear that the traditional view of knowledge representations, in which they develop from a state of deficiency to a state of adequacy and that, once internalised, they operate in every situation; just does not fit with the evidence. However, the endeavour to identify the moment (developmental or evolutionary), from which a representation of physical laws guides behavior and provides true judgements, is probably equally doomed to failure.

What does make sense is to investigate *when and under what circumstances* an organism is sensitive to physical variables and *when and under what circumstances* this sensitivity manifests itself in overt action. This requires that the multiple factors present in any situation be taken into account, and furthermore that they should not be treated as single units but regarded in relation to one another. To account such 'multideterminedness' of behaviour and reasoning, Bremner (1997) proposed a dynamic systems approach as a metaphoric tool to resolve some of the problems connected with the various dissociations found throughout development. Factors related to the acting organism as well as factors specific to its environment and their complex systemic interaction must be taken into account to explain why they specific behavioural outcomes occur in specific situations.

Adopting this systemic view we are able to interpret our findings in a coherent framework: sensitivity to both solidity and gravity constraints influences the search behavior of the apes. But whether gravity or solidity has the stronger influence on behaviour in a given situation is dependent upon multiple interacting factors: species dependent factors, factors of experience and exposure to a specific task context, as well as task-specific factors that highlight the operating force of gravity.

3.8 References

- Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 47-83). Oxford: Blackwell.
- Baillargeon, R. (2004). Infants reasoning about hidden objects: Evidence for event-general and event-specific expectations. *Developmental Science*, 7, 391-424.
- Berthier, N., DeBlois, S. Poirer, C.R., Novak, M.A., & Clifton, R.K. (2000). Where's the ball? Two-and three-year-olds reason about unseen events. *Developmental Psychology*, 36, 394-401.
- Bliss, J., Ogborn, J., & Whitelock, D. (1989). Secondary school pupils' common sense theories of motion. *International Journal of Science Education*, II, 261-272.
- Bogartz, R. S., Shinskey, J. L., & Schilling, T. H. (2000). Object permanence in five-and-a-half-month-old infants? *Infancy*, 1, 403-428.
- Bremner, G. J. (1997). From perception to cognition. In G. Bremner, A. Slater, & G. Butterworth (Eds.), *Infant development: recent advances* (pp. 55-74). East Sussex, UK: Psychological Press.
- Butler, S. C., Berthier, N., & Clifton, R. K. (2002). Two-year-old's search strategies and visual tracking in a hidden displacement task. *Developmental Psychology*, 38, 581-590.
- Cacchione, T., & Krist, H. (2004). Recognizing impossible object relations: intuitions about support in chimpanzees (*Pan troglodytes*), *Journal of Comparative Psychology*, 118, 140-148.
- Cashon, C. H. & Cohen, L. B. (2000). Eight month-old infants' perception of possible and impossible events. *Infancy*, 1, 429-446.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, 10, 483-527.
- Glazko, G. V., & Masatoshi, N. (2003). Estimation of divergence times for major lineages of primate species, *Molecular Biology and Evolution*, 20, 424-434.
- Haith, M. M., (1998). Who put the cog in infant cognition: Is rich interpretation too costly? *Infant Behavior and Development*, 21, 167-179.
- Hauser, M. D. (2001). Searching for food in the wild. A nonhuman primate's expectations about invisible displacement. *Developmental Science*, 4, 84-93.

Hauser, M. D. (2003). Knowing about knowing. Associations between perception and action systems over evolution and during development. *Annual New York Academy of Sciences, 1*, 1-25.

Hauser, M. D., Williams, T., Kralik, J. D., & Moskovitz, D. (2001). What guides a search for food that has disappeared? Experiments on cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology, 115*, 140-151.

Hood, B. M. (1995). Gravity rules for 2-4-year-olds? *Cognitive Development, 10*, 577-598.

Hood, B. M. (1998). Gravity does rule for falling events. *Developmental Science, 1*, 59-64.

Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. (1999). Gravity biases in a nonhuman primate? *Developmental Science, 2*, 35-41.

Hood, B. M., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events. *Child Development, 71*, 1540-1554.

Hood, B. M., Santos, L., & Fieselman, S. (2000). Two year-old's naïve predictions for horizontal trajectories. *Developmental Science, 3*, 328-332.

Hood, B.M., Cole-Davies, V. & Dias, M. (2003) Looking and search measures of object knowledge in pre-school children. *Developmental Psychology, 39*, 61-70.

Kaiser, M. K., Proffitt, D. R. & McCloskey, M. (1985b). The development of beliefs about falling objects. *Perception & Psychophysics, 38*, 533-539.

Karin-D'Arcy, M. R., & Povinelli, D. (2002). Do chimpanzees know what each other see? A closer look. *International Journal of Comparative Psychology, 15*, 21-54.

Kim, I. K., & Spelke, E. S. (1999). Perception and understanding of effects of gravity and inertia on object motion. *Developmental Science, 2*, 339-362.

Kim, I. K., & Spelke, E. S. (1992). Infants' sensitivity to effects of gravity on visible object motion. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 385-393.

Krist, H. (2000). Development of naïve beliefs about moving objects: The straight-down belief in action. *Cognitive Development, 15*, 281-308.

Mash, C., Keen, R., & Berthier, N. E. (2003). Visual access and attention in two-year-old's event reasoning and object search. *Infancy, 4*, 371-388.

Osthaus, B., Slater, A. M., & Lea, S. E. G. (2003). Can dogs defy gravity? A comparison with the human infant and a non-human primate. *Developmental Science, 6*, 489-497.

Povinelli, D. J. (2000). *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford: University Press.

Povinelli, D. J., & Cant, G. H. (1995). Arboreal clambering and the evolution of self-conception, *The Quarterly Review of Biology*, 70, 393-421.

Santos, L. R. & Hauser, M. D. (2002). A non-human primates understanding of solidity: Dissociation between seeing and acting, *Developmental Science*, 5, F1-F7.

Santos, L. R., Seelig, D. & Hauser, M. D. (2006). Cotton-top tamarins' (*Saguinus oedipus*) expectations about occluded objects: a dissociation between looking and reaching tasks, *Infancy*, 9, 147-171.

Southgate, V. H. (2004). A comparative study of the understanding of invisible object displacement in macaque monkeys (*Macaca mulatta* and *arctoides*) and children (*Homo sapiens*). Doctoral thesis, University of St. Andrews.

Spelke, E. S. (1994). Initial knowledge: six suggestions. *Cognition*, 50, 431-445.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, 99(4), 605-632.

Stauffer, R. L., Walker, O. A., Ryde, M., Lyons-Weiler, M., & Blair Hedges, S. (2001). Humans and ape molecular clocks and constraints on paleontological hypotheses, *The American Genetic Association*, 92, 469-474.

Running Head: THE TUBE AS A CAUSAL DEVICE

Do great apes refer to a tube as a causal device? Intuitions about gravity and solidity in four great ape species

Trix Cacchione¹, Josep Call²

¹ Department of Psychology, Cognitive and Developmental Psychology, University of Zürich,
Switzerland

² Max-Planck-Institute for Evolutionary Anthropology, Developmental and Comparative
Psychology, Leipzig, Germany

Address correspondence to:

Trix Cacchione

Department of Psychology

University of Zürich

Binzmühlestrasse 14/21

8050 Zürich

Phone: +41 44 635 7498

Fax: +41 44 635 7489

E-Mail: t.cacchione@psychologie.unizh.ch

4. Do great apes refer to a tube as a causal device? Intuitions about gravity and solidity in four great ape species

4.1 Abstract

In a series of Experiments we investigated if great apes as human infants, monkeys and dogs are subject to a gravity bias when tested with the tubes task, or - in case of mastery – what is the source of competence on the tubes task. Results indicate that apes do neither have a reliable gravity bias when tested with the tubes, nor understand the causal function of the tube. Even though there is evidence that they can integrate tube related causal information to localize the object, they seem to depend mainly on non-causal inferences when searching for an invisibly displaced object.

Key words: invisible displacement, gravity bias, causal reasoning, naïve physics

4.2 Introduction

A vertically falling object represents a special case of object displacement for humans as well as for non-human primates. In recent years several studies investigated the understanding of invisible displacements involving gravity under a comparative perspective (Hauser, 2001; Hauser, Williams, Kralik, & Moskovitz, 2001; Hood, Hauser, Anderson, & Santos, 1999; Osthaus, Slater, & and Lea, 2003; Santos, & Hauser, 2002). The adoption of paradigms that are used in human infant research offers a direct comparability of the data collected with different species. Human infants, cotton-top tamarins (*Saguinus oedipus oedipus*), rhesus monkeys (*Macaca mulatta*) and dogs (*Canis lupus familiaris*) are all reported to make a “gravity error”: They seem to expect unsupported objects to fall vertically, even if solid obstacles impede their trajectory.

Hood (1995) presented 2- to 4-year old children with a task where an object is dropped down an opaque tube into one of three potential goal boxes. 2-year-olds repeatedly searched the object in the box directly beneath the dropping point, even though this box was not connected to the tube. Hood (1995, 1998, 2006) proposed that the error reflects the inability to suppress the naïve theory that all unsupported objects fall in a straight vertical line. The idea is that throughout the first years of life children have vast experience with falling events and eventually acquire a naïve gravity concept (straight down believe) that is later over generalized to cases where it can not be applied appropriately. Hood and colleagues did not find the gravity bias when upward or horizontal motion was presented (Hood, 1998; Hood, Santos, & Fieselman, 2000) what offers further evidence that the search error is associated with naïve gravity concepts. In fact different researchers report straight down believes in older children and – depending on the task context – even in adults (Bliss, Ogborn, & Whitelock, 1989; Kaiser, Profitt, & McCloskey, 1985; Krist 2000).

When tested with the same task also non-human primates and dogs show gravity biased search. Hood, Hauser, Anderson, and Santos (1999) presented adult cotton-top tamarins (*Saguinus oedipus oedipus*) with a version of the original tube-task. Tamarins like human infants had a preference to search directly underneath the dropping place, in the goal box specified by the vertical falling line. Performance of tamarins was generally lower than that of human children and in contrast to children, some tamarins developed a strategy to select the middle box. The tendency to persistently select the gravity box was more marked in tamarins than in human infants. Despite extensive training and cost incentives for correct choices (only one choice was allowed) they went on to select the wrong gravity box even

without receiving the food reward. The immunity to counter evidence is a crucial aspect suggesting again that naïve concepts cause the search errors: To achieve sufficient coherence naïve (and scientific) theories must possess a certain degree of resistance and should not be overturned easily by counter evidence. That the error reflects naïve gravity concepts (and not for example a mere alignment strategy) is further corroborated by the finding that also monkeys are not biased when tested with a horizontal version (Hauser, Williams, Kralik, & Moskovitz, 2001).

Using the same task Osthaus, Slater and Stephen (2003) found a gravity error in dogs (*Canis lupus familiaris*). As human infants and tamarins also dogs had a significant tendency to search the gravity box, but the error was less pronounced in dogs than in monkeys. Further, in contrast to the tamarins, dogs were found to be able to eventually overcome the bias and learn the location where the object can be found. However, the capacity to inhibit the error did not result from an understanding of the tubes' mechanism, since the dogs merely adopted a position preference. But, at any rate this suggests that the strength of gravity error may vary across species.

Not very much is known about how great apes reason reasoning about falling objects (Cacchione, & Krist, 2004; Cacchione, Call & Zingg, in prep.). Cacchione and Krist (2004) found some evidence that apes are sensitive to the effects of gravity on physical objects, namely, they seem (at least in some situations) to expect that unsupported objects fall. Recent studies (Cacchione, Call, & Zingg, in preparation) suggest that - even though apes search failures were clearly found to be related to gravity - apes do not show a reliable gravity error. At least not when tested with the so called table task, a paradigm developed by Spelke, Breinlinger, Macomber, and Jacobson (1992). This paradigm was originally used with a looking measure methodology, later adopted for action tests and used with human infants and non-human primates (e.g. Hauser, 2001; Hood, Carey, & Prasada, 2000). Participants are confronted with a table like apparatus and two cups, one placed on the table plate, the other below it. Then an object is dropped behind a screen, the screen is raised and the participant allowed to search the cup in which he expects the object to be. If participants understand object solidity and expect that an object can not move through a solid table they should select the upper cup. That great apes did not show a gravity bias but performed well when tested with the table task is especially interesting in the light of findings with other species: Hauser (2001) tested rhesus monkeys with this task and found them to be markedly biased towards the bottom cup. 2 years old human infants, however, did (while failing on this task) not show a general down location bias, but a preference for the location where the object was seen

during familiarization (Hood, Carey, & Prasada, 2000). So, at least in human infant's gravity errors only appear in the tubes task. Even though both task types test sensitivity to the effects of gravity and solidity on invisible object displacement, table and tubes task are not directly comparable. The tubes task is more demanding, placing greater cognitive load to the participant. This may be a reason why the gravity error becomes more manifest in the tubes task. So, even though great apes as human children did not show a reliable bottom preference in the table task, but, in contrast to children tended to solve the task, we do not know if the apes generally are endowed with a better capacity to inhibit unwarranted gravity reactions than monkeys. It seems therefore timely to relate great apes performance to that of monkeys and human infants also in context of the tubes task.

The present studies investigate how great apes perform when tested with versions of the original tubes task. In view of the performance of younger children, monkeys and dogs it seems possible that also great apes fail on this rather complex task, even though they showed no reliable bias in the table task. If so, error analyses may help to define possible causes of failure. It might be that apes simply do not know where the object is; in this case apes' selection of locations would be purely at random. On the other hand they might prefer specific goal locations, for example, as children, monkeys and dogs, the goal box specified by gravity, suggesting that errors might be due to a naïve folk theory of gravity. If apes fail in the tubes task it is interesting to see if they, as monkeys, are immune to evidence about the true location of the object, that is, if they stereotypically go on searching it in the gravity box. Or if they, as reported in dogs, eventually overcome the bias and learn to localize the object. As was demonstrated in dogs (who simply adopted a position strategy) learning to correctly localize the object does not imply an understanding of the task. It is therefore important to single out on what kind of information the apes base their box selection. Do they over trials eventually realize the causal relation of tube configuration, object displacement and goal location, or do they merely adopt strategies as location learning or rule learning (e.g. the object is always in the box connected to the tube)? This said it becomes apparent that, even if great apes should master the tubes task from the first trial on, one should be very careful to find out on what information they rely to do so. Is the search behavior really based on the conceptualization of the causal effects of physical forces on a moving objects path? Do they appreciate that a solid tube will determine the future trajectory of an object dropped into it and do they therefore expect that the object will come to rest in the goal box connected to the tube? For example, an abstract conception of the tubes' casual functioning should be independent of experience

made with a specific task type and should be independent of perceptive cues present in a given situation.

In a series of experiments we confront great apes with four versions of the tubes task to find out if they appreciate that a causal connection exists between the tube and the future location of an object dropped into it. The first three versions correspond to three different levels of abstraction in which the degree of mental reconstruction that is needed to infer the future location of the invisible displaced object is varied. On a low abstraction level a transparent tube connects the dropping and goal location. The observer must simply visually track the object moving through the tube and infer that the object rests in the connected opaque goal box. The trajectory of the object can be directly observed and must not be mentally extrapolated from the point of disappearance. On an intermediate level an opaque tube connects dropping and goal location but the object travelling through the tube is accompanied by a moving sound. The acoustic cue signals that the object after its disappearance from view is still travelling inside the tube. The observer can track the objects movement acoustically or, if acoustic cues do not support tracking, infer that an object moving inside the tube comes to rest in the goal box connected to it. (In the later case the acoustic cue may serve as a reminder that the objects is still moving and that the movement takes place inside the tube; this may reduce the cognitive load to infer the correct location.) Again, the observer must not understand the causal relation of tube and goal location of the object, at least not, if he manages to acoustically track the object. On a high abstraction level participants are confronted with a “silent” tube, that is with a situation where only an opaque tube connecting dropping and goal location are present as the only information perceptually accessible. The participants must mentally reconstruct the trajectory and infer that the tube determines the objects future motion. Finally, in a fourth version we confront apes with a painted two-dimensional “tube”. The three-dimensionality of the tube is an important visual cue that signals that the tube can *contain* other objects and is therefore basic to assume a causal relation of tube and object movement. If participants really appreciate the causal function of the tube, they should be aware of the fundamental importance of three-dimensionality as prerequisite to causal functioning and react accordingly.

In Experiment 1 apes are confronted with task versions that supply the apes with perceptive information about the function of the tube: The “acoustic” tube, where acoustic cues are provided and the transparent tube, where the object’s movement is (in addition to the acoustic cue) visually fully accessible. In Experiment 2 we a) withdraw perceptive cues and test if performance changes in the “silent” tube version of the task and, b) present the two-

dimensional “tube”, a non-causal version of the task, to see, if apes appreciate the causal relation of tube, object movement and goal location. If they understand the tubes’ causal function, they should a) localize the object from the first trial on, b) localize the object irrespective if a transparent, an “acoustic” or a “silent” tube is presented, c) differentiate between a causal, three-dimensional and a non-causal two-dimensional tube. Experiments 3-5 were run as controls on apes that passed Experiment 1 and 2 to analyze the nature of successful performance and to rule out the use of simpler strategies.

4.3 Experiment 1

Experiment 1 investigates the apes’ reactions to a scaled down version of the tubes task. In an initial task an acoustic cue is provided to facilitate the notion that, firstly, the object (albeit invisible) is actually in motion, and, secondly, the movement occurs inside the tube. Can apes successfully localize the invisibly displaced object using the acoustic information? If not, we are interested to see: a) if more gravity than non-gravity errors occur, and b) if performance eventually improves. If performance remains low, apes are run on a second version of the task where a transparent tube is presented. Does performance improve if (in addition to the acoustic cue) object movement is fully visible? Finally, we compare if all ape species perform in a similar manner, or if systematic species differences occur.

4.3.1 Method

Participants

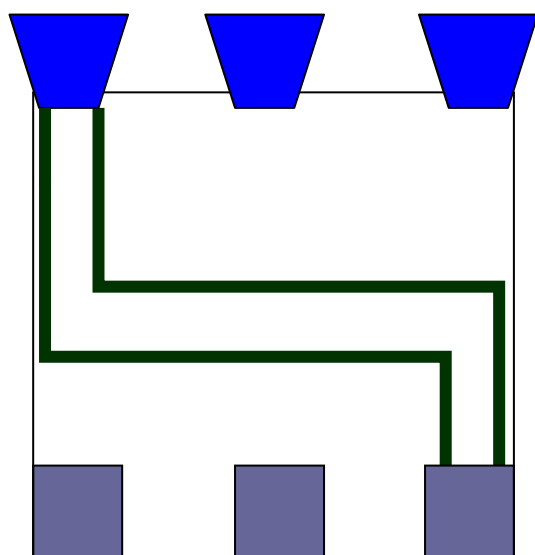
6 gorillas (*Gorilla gorilla*), 5 orangutans (*Pongo pygmaeus*), and 5 bonobos (*Pan paniscus*) participated in Experiment 1. All apes were housed at the Wolfgang Köhler Research Center (Zoo Leipzig), in Germany. Except 1 gorilla they all had prior experience with various experiments investigating physical cognition. Apes were tested alone either in an indoor observation room or in their sleeping room, respectively. Mothers with children younger than three years of age were tested in company of their offspring.

Apparatus

The apparatus was modeled after Hood’s (1995) original apparatus (see Figure 1). It consisted of a grey plastic frame and a transparent rear panel made of plexiglas (height 42.5

cm/ length 68.2 cm/ depth 11.3 cm). On the upper section three opaque plastic chimneys were attached. On the lower section three opaque goal boxes (height 9.8 cm/ length 9.1 cm/ depth 9.4 cm) were placed, directly below the chimneys. Chimneys and goal boxes could be connected by opaque tubes (length 74 cm/ diameter 3 cm). The tubes were made of pieces of a flexible hose. Hazelnuts were used as food reward to drop down the tubes. The traveling of the hazelnut through the tubes was accompanied by a moving sound. A deformable cotton mat served as absorbing device to absorb the sound of arrival in the goal box.

Figure 1



Set up for the test condition in Experiment 1

Design and Procedure

There were two experimental phases: A baseline phase and a testing phase (consisting of two conditions each). All conditions were run by an experimenter who presented the stimuli. Both baseline and test conditions were presented in sessions mounting up to nine trials. In both baseline conditions the sessions were repeated until the ape was correct in eight out of nine trials. Then it proceeded to the testing conditions: An acoustic tube condition and a transparent tube condition. The acoustic test condition included three sessions consisting of nine trials each. If the ape was correct in seven out of nine trials in one of the three sessions it was judged to have passed the acoustic test condition. Only apes that did not pass the acoustic

test condition moved on to the transparent tube condition. In the transparent tube condition they were presented with one session consisting of nine trials. If they were correct in seven out of nine trials they were judged to have passed the transparent test. All apes were presented with only one baseline or test session per day.

The baseline conditions served to familiarize the apes with the apparatus, to exclude baseline preferences for a specific goal box, and to make sure that apes can direct search to a new goal box after finding the food reward repeatedly in the same location. In the first baseline condition the hazelnut was placed alternately in each goal box in a random fashion. In the second baseline condition the hazelnut was placed in the same goal box for three consecutive trials, then three times in a second goal box and finally three times in the remaining goal box. Again the order of the boxes was at random. The procedure in first baseline condition was as follows.

The ape sat behind a plexiglas panel. A testing surface (slide table) was fixed by a metal frame directly underneath the panel and the apparatus placed on top of it. The Plexiglas panel had three holes through which the ape could point at each of the goal boxes. The experimenter sat in front of the slide table. The slide table was constructed such that its surface could be shifted back and forth. The experimenter pulled the table back and placed a hazelnut in full view of the ape in one of the goal boxes and closed it. Then she pushed it to the panel. The ape could now point to the goal box in which expected the food to be hidden. The experimenter draw the table back again. If the ape had chosen correctly she opened the box and gave the hazelnut to the ape. If he had pointed to the wrong goal box the experimenter pushed the table back to the panel allowing for another choice. This was repeated until the ape localized the nut. The search history was directly recorded on a sheet. If the ape was correct in eight out of nine trials he proceeded to the second baseline trial. If not the first baseline condition was repeated. All trials were videotaped.

The procedure of the second baseline was identical to the first except that the hazelnut was now placed on three consecutive times in the same goal box. If the ape was correct in eight out of nine trials he proceeded to the acoustic tube test. If not the second baseline condition was repeated.

In the acoustic tube condition half of participants were presented with a single tube connecting the upper left chimney with the lower right goal box, the second half was tested with the opposite tube configuration. The experimenter pulled the table back and waved the nut above the chimney. If the ape tracked it, the experimenter released the nut. Then the apparatus was pushed to the panel and the apes were allowed to indicate their choice by

pointing to a goal box. The experimenter now pulled back the apparatus. If the ape selected correctly, the experimenter took out the nut gave it to the ape. If the wrong box was selected the experimenter pushed the apparatus back to the panel and the ape was allowed to choose again. This procedure was repeated until the ape located the nut or abandoned search. As in the baseline sessions search history was directly recorded on a sheet and all trials were videotaped. The procedure in the transparent tube condition was the same except that instead of an opaque a transparent tube connected chimney and goal box, so that the nut moving through the tube was now - in addition to the acoustic cue due to movement sound - fully visible.

4.3.2 Results

Only 18.8 % of the apes reached criterion (7 out of 9 correct) within the first session ($\chi^2 = 6.250, p < .05$) of Experiment 1 (none of the gorillas, 20 % of the bonobos and 40 % of the orangutans), but 87.5 % of the apes reached it in one of the three test sessions ($\chi^2 = 9.0, p < .01$).

Five of 16 apes chose the correct box on first attempt in the first trial. This is not different from the amount of correct responses expected by chance (5.3 correct responses in 16 apes). All orangutans chose erroneously in the first trial, 3 of 6 gorillas and 3 of 5 bonobos did so. Of the 11 apes that chose the incorrect box, 8 selected the gravity box and 3 the middle box (n. s.). In the first trial of the second test session 6 of 16 apes chose correct, 9 of 10 erroneous searches were directed to the gravity box ($\chi^2 = 6.400, p < .01$). In the first trial of the third session 9 of 16 apes chose the correct box, this trend was only marginally significant (binomial test, $p = .056$). Five of 7 incorrect searches were directed to the gravity box.

On average only 48 % of the choices were correct during test session one. There was a marginally significant tendency of improvement within the first test session (39.5 % were correct in trial 1-3 and 62.5 % in trial 7-9; Wilcoxon, $z = -1.701, p = .089$). In test session two 66.9 % and in session three 72.19 % of choices were correct. There were no species differences except that, on average, gorillas made reliably more errors in session two (Kruskal-Wallis, $\chi^2 = 6.117, p < .05$).

Figure 2

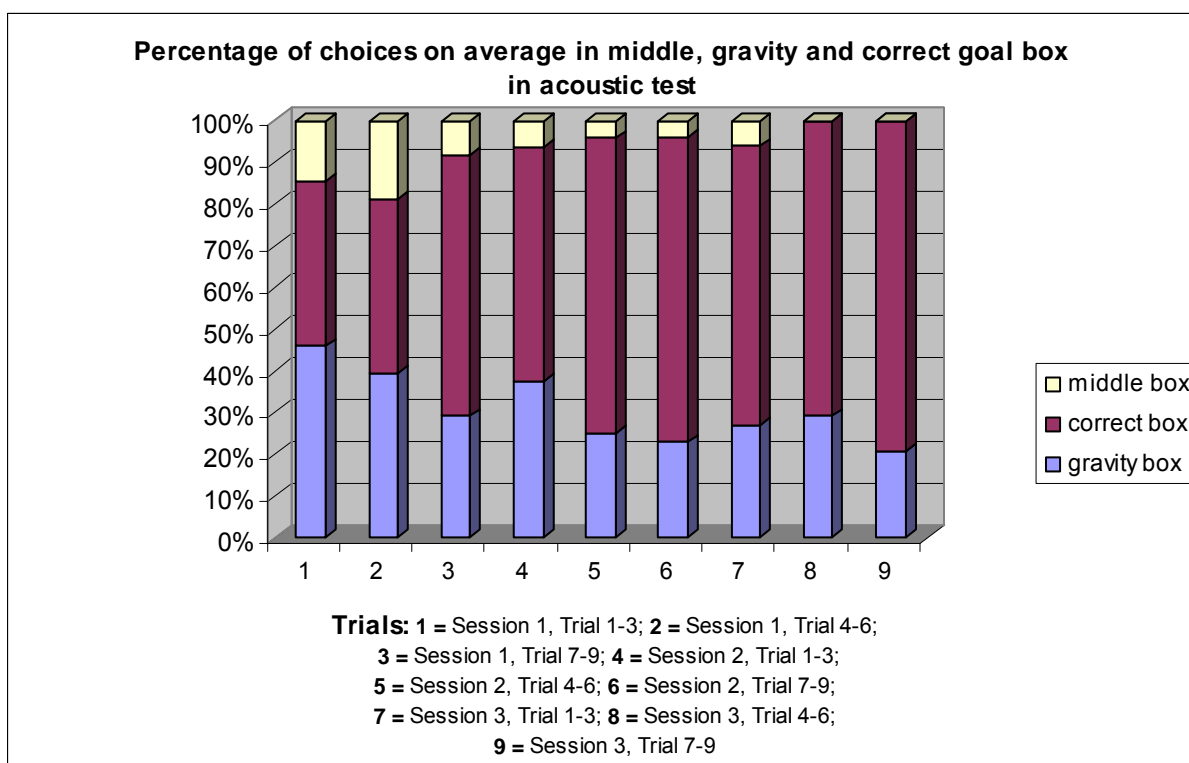


Figure 2 shows the percentage of choices throughout all three test sessions. The percentage of correct choices increases over trials, the percentage of gravity answers declines. Few apes select the middle box, in the last 6 trials of session 3 the middle box is never selected. Performance did significantly improve between test sessions: Apes chose more often correct in session two than one (Wilcoxon, $z = -2.210, p < .05$) and more often in session three than one (Wilcoxon, $z = -2.019, p < .05$). An ANOVA with order (frequency of correct choice in test session 1, 2 and 3) as within subject factor revealed a significant test order effect $F(2, 30) = 4.153, p < .05$.

Two gorillas did not reach criterion during the three test sessions of acoustic test and were run on a transparent test session. Both selected the correct box already in the first trial. Also, both reached criterion in the transparent test. On average 94.5 % of the choices were correct.

4.3.3 Discussion

Apes are not able to locate a nut dropped down an opaque tube, not even when acoustic cues accompany its movement through the tube. Only a small proportion of the

sample reached criterion in the first test session, the first trial performance was at random and, on average less than 50 % of the choices were correct during test session one. This suggests that first, apes do not understand the function of the tube, and, second, the provision of acoustic cues does not help substantially to localize the object. Despite the fact that a) the gravity box was more often selected in the first trial than both other goal boxes and b) more gravity than non-gravity errors occurred, there was no significant tendency to select the gravity box in the first test session. So, in contrast to children, monkeys and dogs, apes do not show a marked preference for the gravity box in this task.

Over trials the performance reliably improved and finally the major part of the apes chose the correct goal box. Likewise, the amount of erroneous choices declined, but much more pronounced in case of middle box choices, so finally gravity errors prevailed. Only two apes did not reach criterion in Experiment 1. However, both of them passed the transparent test condition smoothly, suggesting that visual information is much more valuable to enhance performance than acoustic information. Only visual perception of the movement supports successful tracking, while acoustic perception does apparently not suffice for this purpose.

In sum the results show that first, apes initially fail in an acoustic version of the tube, that second, they do not have a reliable gravity box preference and that third, they eventually learn to select the correct goal box. The question remains, however, what exactly they learned in terms of content during Experiment 1. Of particular importance is the question if the acoustic information provided in Experiment 1 is related to the learning. Even though acoustic cues do not seem to be helpful in initial trials, their value to localize the object might still have been detected later in course of the task. Moreover, it can not be ruled out that acoustic information already in the first test session influenced the performance. Even though the apes performed low, their performance in absence of acoustic cues might have been even lower. For example, it is possible that the apes that chose correctly in the first session were completely relying on acoustic information. On the other hand it is possible that acoustic information did not enhance performance in any respect, but that the improvement of performance throughout Experiment 1 just reflects the adoption of simple strategies (e.g. position learning). The question if acoustic information was crucial to explain performance in Experiment 1 is further investigated in Experiment 2.

4.4 Experiment 2

Successful localization of the object in Experiment 1 could have been due to: a) the use of acoustic information, b) position learning, c) learning to search the location with the tube attached or d) an understanding of the tubes' causal function. In Experiment 2 apes are confronted with two conditions to investigate the possible causes of correct performance: First, a "silent" tube condition where object movement is both invisible and non-audible, and, second, a two-dimensional gloss paint tube condition where no causal relation between tube shape and goal location exists. If correct performance is simply associated with the use of acoustic information apes should (at least initially) fail to localize the object in the silent test (as well as in the painted tube test). If mere position learning was the reason for correct performance apes should equally localize the object in the real and in the painted tube condition. If they simply learned to search the location with the tube attached they should find the object in the silent test condition, but not (at least not initially) in the painted tube condition. Finally, if correct performance was associated with an understanding of the causal function of the tube apes should react differentially to the silent and the painted tube, displaying greater difficulties to localize the object in the painted tube test.

4.4.1 Method

Participants

The same 6 gorillas (*Gorilla gorilla*), 5 orangutans (*Pongo pygmaeus*), and 5 bonobos (*Pan paniscus*) as participated in Experiment 1 also were run on Experiment 2. Additionally, a group of 6 chimpanzees (*Pan troglodytes*) was tested. Also chimpanzees were housed at the Wolfgang Köhler Research Center (Zoo Leipzig) in Germany and had prior experience with various experiments investigating physical cognition. The same testing facilities were used as in Experiment 1.

Apparatus

In the silent tube condition, the same apparatus as in Experiment 1 was used with one alteration: The upper part of the tube was blocked by a barrier so that the nut inserted in the chimney rested invisibly inside the tube and did not travel through it. In the gloss paint pattern

condition, again, the same apparatus was used, but instead of a tube a plexiglas panel was inserted in flush with the rear side panel. On the panel a two-dimensional tube was painted with black gloss paint. As the real tube the panel could be inserted in way that it appeared to “connect” the upper left chimney with the lower right goal box or vice versa. The chimney was blocked by a barrier, so that the hazelnut inserted rested invisibly inside the chimney. An oblong screen could be mounted in front of the apparatus so that the hazelnut could be removed between trials.

Procedure

Before testing the chimpanzee group was first presented with the same two baseline conditions as described in Experiment 1. The procedure in the baseline conditions was the same as in Experiment 1. All other ape species were directly confronted with the test conditions.

Experiment 2 consisted of two test conditions: A silent tube condition and a gloss paint pattern condition. One half of the apes was tested with the silent tube condition first, the other half was first run on the gloss paint pattern condition. Both test conditions added up to three sessions, each session consisted of nine trials. If the apes were correct in 7 out of 9 trials in one of the three sessions the test condition was judged as being passed. All apes were presented with only one test (or baseline) session per day.

In both test conditions the apes were tested in the same setting as in Experiment except for the above described alterations on the tube apparatus. One half of the apes was presented with an upper-left-lower-right configuration, the other half with the opposite configuration. The apes that were already tested in Experiment 1 were presented with the same configuration as in Experiment 1.

In the silent tube condition the experimenter presented the tube and inserted it in full view of the ape inside the apparatus, then she pulled the table back and waved a nut above the chimney connected to the tube. If the ape tracked the nut the experimenter released it. Then she pushed the table to the panel and allowed the ape to choose one goal box by pointing through the panel. The table was pulled back again. If the ape pointed to the correct box, the experimenter reached into the box, produced a nut and gave it to the ape (the box was not baited before hand but the nut surreptitiously placed there while reaching into the box). If the ape did not select the correct goal location the experimenter pushed the table back to the panel allowing for another choice. This was repeated until the ape correctly located the nut or abandoned search. Before moving on to the next trial, the experimenter raised a screen in

front of the apparatus and the nut resting in the in the tube was removed. The procedure in the gloss paint pattern condition was identical with the following exception: Before testing the ape with the gloss paint pattern, the plexiglas with the pattern painted on it was held on front of the ape and rotated to make sure that he appreciated that it was a two-dimensional pattern. Then it was inserted in full view of the ape into the apparatus and the test began. The search behavior was recorded on a sheet and all sessions videotaped.

4.4.2 Results

Silent test condition

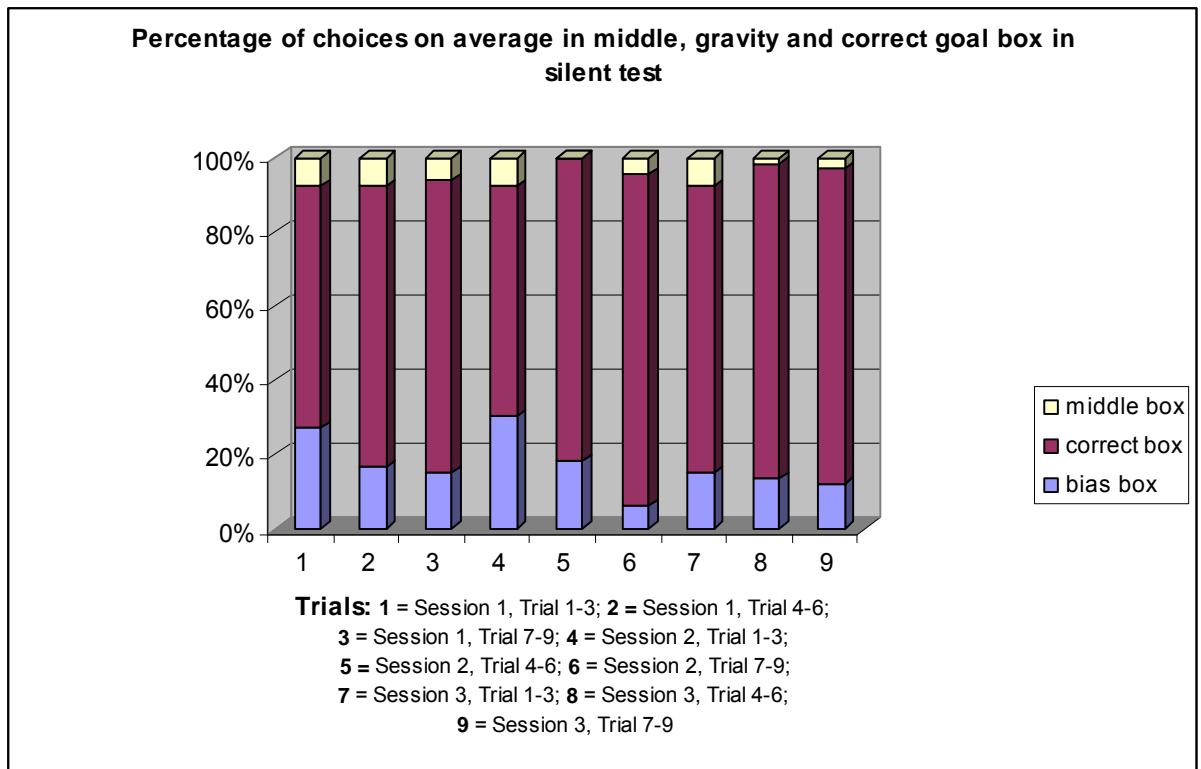
Within the first session of the silent test condition 59.8 % of the apes reached criterion (7 out of 9 correct; n. s.), 81.8 % reached it in one of the three test sessions of silent test condition ($\chi^2 = 8.909, p < .01$). Compared to orangutans and bonobos (in both groups 80 % reached criterion) chimpanzees (33.3 % reached criterion) and gorillas (50 % reached criterion) had more difficulties in the first session of silent test. The performance of chimpanzees quickly improved and 83.3 % of them reached criterion in one of the three test sessions of silent condition, gorillas performance continued to be low (50 % reached criterion in one of the three test sessions).

Fourteen of 22 apes chose the correct box on first attempt in the first trial of silent test. This is reliably more than the 7.3 correct responses expected by chance (binomial test, $p < .01$). Of the 8 apes that responded wrong all chose the gravity box. Species differed in their first trial performance of silent test: Gorillas and chimpanzees selected equally often the correct and the gravity box. Orangutans and bonobos in contrast chose reliably correct in the first trial (binomial test, $p < .05$). Also in session two and three of silent test apes chose reliably correct in the first trial (binomial test, $p < .05$ and $p < .001$). Again there are species differences: In the first trial of session two gorillas have a preference for the gravity box (binomial, test, $p < .05$), chimpanzees do not perform above chance, and orangutans and bonobos prefer the correct box (binomial, test, $p < .01$ and $p < .05$). In the first trial of session three only bonobos select the correct box more often (binomial, test, $p < .05$). Of the 10 erroneous first trial choices in session two, 7 were directed at the gravity box (n. s.). In session three 4 of 6 wrong first trial answers were gravity responses (n. s.).

On average 73.27 % of the choices were correct during test session one, 77.77 % in session two and 82.32 % in session three. There was a significant improvement within the second test session (Wilcoxon, $z = -3.269, p < .001$). There were non-significant trends for

species differences: On average gorilla and chimpanzees tended to make more erroneous choices than orangutans and bonobos in session one ($\chi^2 = 7.442, p = .059$), and, gorillas were also more often mistaken in session two ($\chi^2 = 6.587, p = .086$) and three ($\chi^2 = 6.387, p = .094$). In Figure 3 the percentage of choices throughout all three sessions of silent test are depicted. As a tendency, the percentage of correct choices increases over trials (at least within sessions) and the percentage of gravity answers declines. However, there are no significant differences in performance between test sessions. Again few apes select the middle box a tendency that rather declines over sessions.

Figure 3



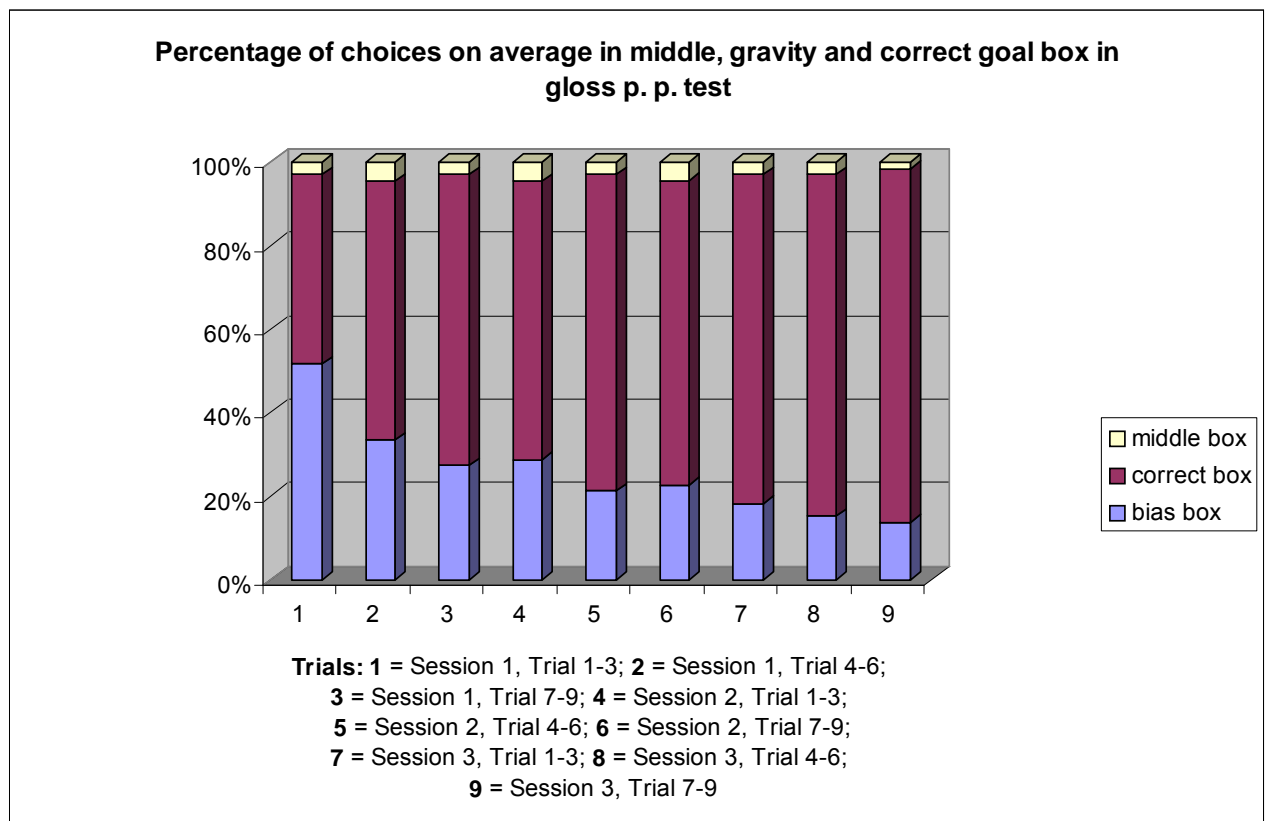
Gloss paint pattern condition

In the paper test condition 40.9 % of the apes reached criterion in the first session (n. s.) and 68.2 % within one of the three test sessions (n. s.). Again gorillas' and chimpanzees' performance was lower in the first session (in both groups 33.3 % reached criterion) and throughout session one to three (50 % of the chimpanzees and 66.7 % of the gorillas reached criterion) than that of orangutans and bonobos (60 % resp. 40 % in the first session, 80% in session one to three).

Nine of 22 apes chose the correct box on first attempt in the first trial of gloss paint pattern test (n. s.) and 12 the gravity box (binomial test, $p < .05$). Of 13 erroneous choices 12 were directed at the gravity box ($\chi^2 = 9.308$, $p < .01$). There were species differences in the first trial performance of gloss paint pattern test: Only orangutans performed correctly (binomial test, $p < .05$), gorillas reliably often chose the gravity box (binomial test, $p < .05$), chimpanzees showed a very weak trend to do so (binomial test, $p = .097$) and bonobos did not have an above chance box preference. In session two and three of gloss paint pattern test apes chose reliably correct in the first trial (binomial test, $p < .01$ and $p < .001$). Again there are species differences: In session two again orangutans perform correctly (binomial test, $p < .05$) and chimpanzees at least show a very weak trend to do so (binomial test, $p = .097$), while gorillas have a weak tendency to choose the gravity box (binomial test, $p = .097$), and bonobos have no above chance preference. In session three all except gorillas perform correctly (binomial test, orangutans: $p < .01$; bonobos: $p < .05$; chimpanzees: $p < .01$). 9 chose erroneously in session two, 8 of them selected the gravity box ($\chi^2 = 5.444$, $p < .05$). In session three finally all five incorrect choices were directed to the gravity box.

During test session one 59.05 % of the choices were correct, 71.68 % in session two and 81.91 % in session three. There was a significant improvement within the first test session (Wilcoxon, $z = -3.002$, $p < .01$). No species difference was identified in session one to three. In Figure 4 the percentage of choices throughout all three sessions of silent test are depicted. The percentage of correct choices increases over trials and the percentage of gravity answers declines. Again few apes select the middle box, a tendency that rather declines over sessions. Performance did significantly improve between test sessions: Apes chose more often correct in session two than one (Wilcoxon, $z = -2.287$, $p < .05$), more often in session three than one (Wilcoxon, $z = -3.155$, $p < .01$) and more often in session two than three (Wilcoxon, $z = -2.586$, $p < .01$).

Figure 4



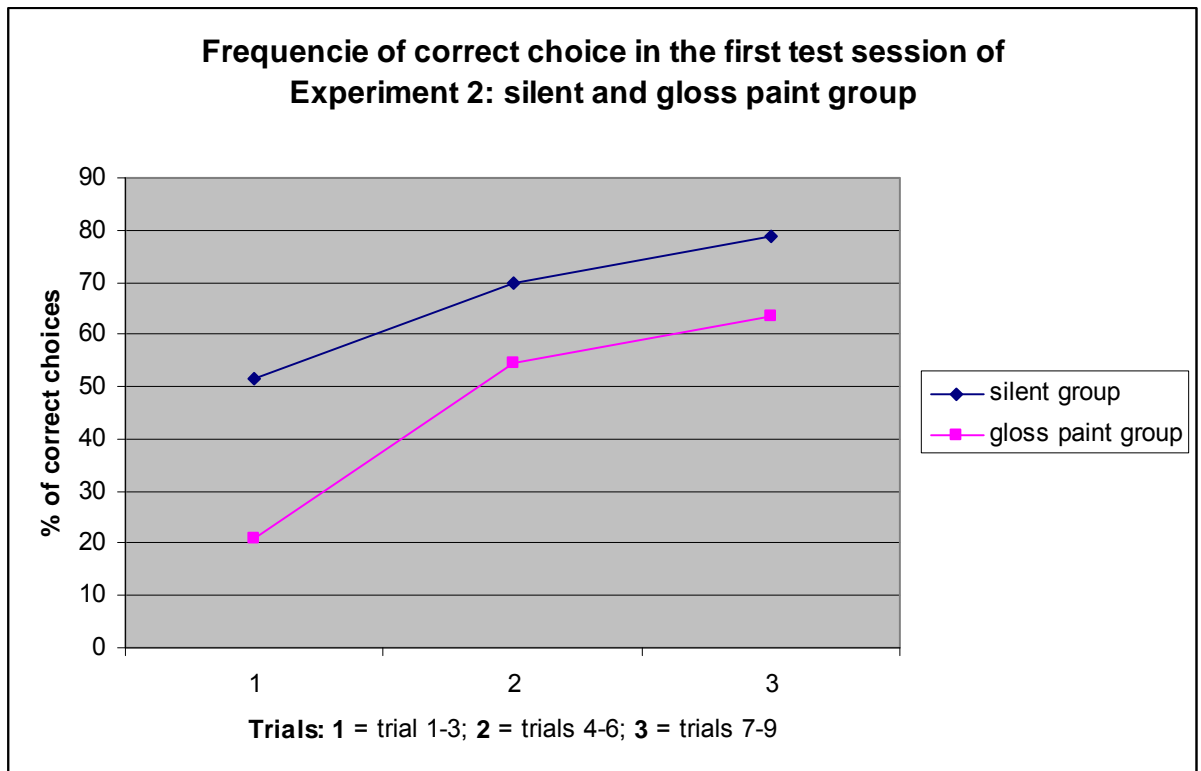
Differences between conditions and test order effects

For the sub sample that participated in both Experiment 1 and 2 we calculated Wilcoxon tests to compare between performance in the last session of acoustic test and first session of silent test to see if performance declined, when the acoustic cue was no longer available. A comparison of the percentage of correct choices of the third session of acoustic test and the first session of silent test revealed no significant differences, neither in the first three trials nor throughout all nine trials. The same was true if Wilcoxon tests were calculated for the group that was tested with the silent condition first in Experiment 2. There were no species differences in this respect.

Next, analyses were run to compare performance between silent and gloss paint pattern condition. Wilcoxon tests were computed to check for performance difference between the first session of silent and gloss paint pattern test. In the gloss paint condition the gravity box was reliably more chosen ($z = -2.098, p < .05$). However, there were only marginally more correct responses in the first three trials of silent session ($z = -1.791, p = .079$). This was due to the low performance of chimpanzees: If the data was analyzed for gorillas, bonobos and orangutans this trend reached significance ($z = -2.116, p < .05$).

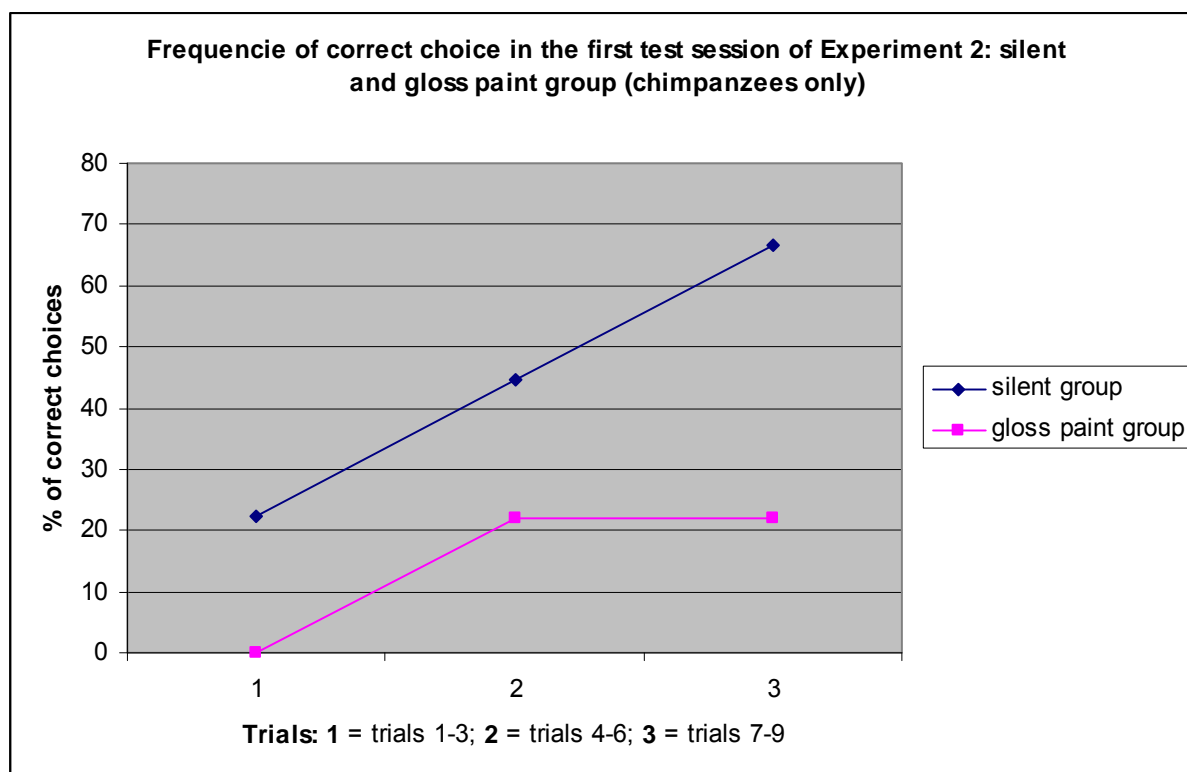
Because half of the apes were tested with silent condition first, and half of them with gloss paint first, experience might diminish between-condition effects. We therefore analyzed the data for the first test condition run only, comparing performance of the silent and the gloss paint sub sample. In the first trial of first test session 7 apes of the silent group chose the correct box and 4 the gravity box. In the gloss paint group 3 apes chose the correct, 7 the gravity and 1 the middle box. Binomial tests were run to analyze the first trial choices. They revealed that the silent test sub sample chose reliably correct in the first trial of first test session ($p < .05$), the number of apes choosing the gravity box was not above chance. In the gloss paint test sub sample the gravity box was chosen significantly more than expected by chance ($p < .05$), while the number of correct choices did not differ from chance. In the first test session 66.73 % of the choices made by the silent group were correct and 24.18 % directed to the gravity box. The gloss paint group chose in 46.45 % of the cases the correct box and in 48.45 % the gravity box. Figure 5 shows the performance in the two conditions. Over all nine trials of the first test session the performance difference did not reach significance, but performance in the first three trials reliably differed: Apes of the silent group chose more often correct ($z = -1.977, p < .05$), apes of the gloss paint group chose more often the gravity box ($z = -2.561, p < .01$).

Figure 5



The chimpanzee sub sample is the only one directly confronted with Experiment 2, that is, they have no prior experience with the acoustic tube. To see, if prior experience with the acoustic tube is the cause for the observed performance differences between silent and gloss paint conditions, we qualitatively analyzed the chimpanzees' data for the first test session run (again comparing the silent and the gloss paint subgroup) alone, to see, if we find the same performance patterns as in the whole sample. Throughout the first test session 44.67 % of the choices of the silent group were correct and 37 % were directed to the gravity box; compared to only 14.67 % correct choices but 81.67 % gravity choices in the gloss paint group. Figure 6 shows performance of chimpanzees in the first test session for both sub groups. It mirrors the pattern produced by the whole sample. (The chimpanzee sub sample is too small to carry out statistical analyses.)

Figure 6



Additionally, we analyzed if prior experience with Experiment 1 had a different impact on performance on subsequent silent or gloss paint condition tests of Experiment 2. We compared performance differences between acoustic and silent test and, between acoustic and gloss paint pattern test, analyzing only the data from gorillas, orangutans and bonobos. Wilcoxon tests revealed significant more correct responses in the first session of silent test than in the first session of acoustic test ($z = -2.790, p < .01$), while performance between first sessions of acoustic and gloss paint tests did not differ. Also in the first three trials of gloss paint test the percentage of correct responses was reliably lower than in last session of acoustic test (Wilcoxon, $z = -2.167, p < .05$), while performance between third session of acoustic test and the first session of silent test did not differ.

Finally, we checked for test order effects. An ANOVA with order (frequency of correct choice in test sessions 1-6 of Experiment 2) as within subject factor revealed a significant test order effect $F(5, 105) = 6.901, p < .001$. We also tested for order effects over all test sessions in Experiment 1 and 2: Also over all test sessions of Experiment 1 and 2 an ANOVA revealed a significant order effect (frequency of correct choices in test session 1-9) with $F(8, 120) = 3.346, p < .01$.

4.4.3 Discussion

Apes performance in the last acoustic and the first silent test session did not differ, no decline in performance occurred when the acoustic cue was withdrawn. So, whatever apes learned during Experiment 1 was not linked to the presence of acoustic cues.

When confronted with the silent tube, two thirds of the apes reached criterion in the first test session, the great majority of them within three sessions. From the first trial they reliably chose the correct box and, in case of erroneous responses, much more gravity than non-gravity errors occurred. There were some species differences with chimpanzees and gorillas performing lower. Chimpanzee's lower performance was expected and can be explained by their lack of prior experience with Experiment 1. In general, their performance mirrored that of the rest of the sample in Experiment 1: They initially had no above chance preference, made more gravity than non-gravity errors and learned quite quickly to localize the object. The cause of the low performance of gorillas, however, is less clear. Apparently they have generally more difficulty with the tube task (already in Experiment 1 gorillas' performance was lower) and profit less from experience with the task.

Apes obviously differentiated between a causal three-dimensional real tube and a non-causal two-dimensional painted tube. Less apes passed criterion in the gloss paint test than in the silent test. In contrast to the silent test (where apes performed correctly) they were significantly biased in the first trial of gloss paint test. This is remarkable, because neither in the acoustic test of Experiment 1 nor in the silent test in Experiment 2 apes reliably preferred the gravity box. Maybe the gravity response occurs as a default when no valuable information about the location of the object is available (this issue is further addressed in the general discussion). In the first test session apes made reliably more correct choices in the silent condition and reliably more gravity errors in the gloss paint condition. Also, there were no reliable performance differences between silent test sessions while performance improved significantly between all three gloss paint test sessions. Finally, only within the silent test species differences were manifest, with both chimpanzees and gorillas initially performing lower than orangutans and bonobos, while in the gloss paint test there were no reliable species differences. Together this suggests that the apes experienced a qualitative difference between silent and painted tube.

Apes not only reacted differentially to the silent and the painted tube, their reactions also suggest that they experienced a difference between acoustic tube and painted tube, but not between acoustic and silent tube. Performance deteriorated only between last session of

acoustic test and first session of gloss paint test, while there was no performance difference between last session of acoustic and first session of silent test. Also, only in the painted tube test a reliable gravity bias occurred in the first trial. Apparently, only the painted tube represented something qualitatively different than both acoustic and silent tube. Also, experience made in Experiment 1 could not be transferred equally to both test conditions of Experiment 2: While reliably more apes selected the correct box in the first session of silent test than acoustic test, first session performance did not differ between acoustic and gloss paint test. This indicates that what was learned in the acoustic test could be applied in the silent test, but not in the gloss paint test. However, apes' different reactions to the real three-dimensional and the painted two-dimensional tube can not solely be due to the experience with the acoustic tube in Experiment 1 (e.g. because they experienced the silent tube as perceptually more similar with the already familiar acoustic than the painted tube). Chimpanzees did not partake in Experiment 1, but showed the same response difference as the other three ape species: They chose more often correct in the silent test and made more gravity errors in the gloss paint test.

Why do apes treat a real three-dimensional and a painted two-dimensional tube as two qualitatively different phenomena? The two tube versions differ in causal and perceptual properties, but only the difference in causal properties can explain differential success in retrieving the object. Only the real tube represents a causal mechanism and provides the observer with information about the future movement of the disappearing object. In case of a painted tube no causal connection exists between the locations where the object disappears and is retrieved. The object could not possibly have continued to move after its disappearance in a chimney without being perceived and can therefore not have changed its position to one of the three goal boxes. So, the painted tube test confronts the ape with a task that has no solution, with a situation where it is impossible to infer in which goal box the object might be found. At best apes can guess in the initial trials and eventually learn to find the object by establishing a position strategy or a rule "the object is always in the location specified by the line". If the apes only experienced a perceptual difference between real and painted tube, then they would in neither situation have valuable information about the objects' current location. The different performance can only be explained by assuming that they were to some degree sensitive for the causal difference: Only in the real tube condition they were provided with information about the objects current position (while in the painted tube condition they could only guess), therefore the object was more often found in the real tube condition.

As in Experiment 1 also in Experiment 2 learning effects were observable. There was reliable performance improvement not only within Experiment 2, but also over all sessions of Experiment 1 and 2. The question remains what exactly the apes learn and on what information they rely when they select the correct box. Obviously, they did not just learn to track the object acoustically in Experiment 1, as the object is equally localized when no acoustic cues are given. This rules out the possibility that correct search is solely dependent on the use of acoustic information. The results further indicate that something more than a position strategy is at the root of correct choices. Since they clearly differentiate a real from a painted tube they seem to integrate at least to some extent information regarding the functionality of the tube and are not reacting solely to position properties. However, that correct choices are based on a complete causal understanding of the tubes function must also be ruled out, because they do not find the object initially. Experiment 3, 4, and 5 further investigates if apes prefer to a) select the prior enforced position, b) select the position with the tube attached, or c) select the box specified by the tubes' causal function.

4.5 Experiment 3a

Even though apes initially failed to find the hazelnut they eventually learned to localize it over trials. Moreover, they clearly found it easier to localize the object when presented with a causal tube than with a non-causal drawing. It remains unclear however, on what information apes rely to select a goal box. In Experiment 3a apes are presented with the opposite tube configuration to that of Experiment 1 and 2. If location learning is the cause for correct choices apes should fail to localize the object. Successful localization on the other hand indicates either the use of a more elaborate strategy (e.g. "the object is at the location with the tube attached") or a true understanding of the tube.

4.5.1 Method

Participants

Only those apes that passed the silent tube test condition of Experiment 2 were run on Experiment 3. This was true of 19 of the 22 apes tested on Experiment 2: 3 gorillas did not pass the silent test; therefore only 3 gorillas remained in the sample. All other apes were the same as in Experiment 2.

Apparatus

The apparatus was identical to that in Experiment 1 and 2. As in the silent test condition of Experiment 2 the blocked tube was used.

Procedure

All apes were directly run on the single test condition of Experiment 3a. Only one session consisting of nine trials was administered. If the ape was correct in 7 out of 9 trials it was judged to have passed the test. All apes were now tested with the opposite configuration: Those who were run on an upper-left-lower-right configuration in Experiment 1 and 2 were now presented with an upper-right-lower-left configuration and vice versa. The procedure was identical to that of Experiment 2.

4.5.2 Results

Only 21.1 % of the apes reached criterion ($\chi^2 = 6.368, p < .01$). None of the chimpanzees and gorillas reached criterion, but 40 % of the orangutans and bonobos did. Only 1 out of 19 apes chose the correct box (binomial test, $p < .01$) on first attempt in the first trial of Experiment 3 and 17 the prior rewarded (now gravity) box (binomial test, $p < .001$). Of 18 erroneous first trial choices 17 were directed to the gravity box ($\chi^2 = 14.222, p < .001$) that is, the box were actually containing the hazelnut in the prior tests in Experiment 1 and 2. On average only 33.26 % were correct choices throughout trial one to nine. But most errors occurred during the initial trials, and then performance quickly improved. A Wilcoxon test revealed a significant difference in the frequencies of correct choices between trial 1-3 and trial 7-9 (15.68 %, 49.05 %; $z = -2.758, p < .01$). But even in the last three trials only 49.05 % of the choices were correct, while in 42.11 % of the cases still the box rewarded on prior test was chosen and in 8.84% of the cases the middle box.

4.5.3 Discussion

Practically all apes selected the prior rewarded position in the first trial. This strongly suggests that apes used position as the strongest cue to localize the object. A substantial part of the apes did not stick to the position strategy, but learned to switch to the new location. This is interesting compared to the first session performance of Experiment 1 where apes did

not appear to learn so substantially: By the time the apes were run on Experiment 3a, they have been rewarded over many sessions to select a specific goal box. So, one would expect them to be even more reluctant to switch to a new location than in Experiment 1 where they lacked prior experience. This might suggest that the apes learned something additional than just to remember a specific location, even though the position seems to be by far the most important cue for them.

4.6 Experiment 3b

It seems highly likely that apes mainly used position information to localize the object. However, in Experiment 3a the gravity box was the same box as the box rewarded in Experiment 1 and 2. This leaves a minor possibility, that apes did not have a position preference, but reacted with a default gravity answer when confronted with the new tube configuration. Experiment 3b decides if position learning or gravity bias is at the root of findings of Experiment 3a. Further, it investigates the strength of the association of food and specific location built up during Experiment 1 and 2.

4.6.1 Method

Participants

The same 19 apes as in Experiment 3a also participated in Experiment 3b.

Procedure

Apparatus and procedure was identical to that of Experiment 3a with the exception that all apes now were presented with an upper-middle-box-lower-outside-box configuration. Those who were run on Experiment 1 and 2 with an upper-left-lower-right configuration were now presented with an upper-middle-lower-left configuration and vice versa.

4.6.2 Results

In Experiment 3b 57.9 % of the apes reached criterion (n. s.).

6 out of 19 apes chose the correct box (not different from the 6.3 correct responses expected by chance) on first attempt in the first trial and 12 selected the prior rewarded position (binomial test, $p < .01$). In only one case the middle (gravity) box was selected but 12 of the 13 erroneous first trial choices were directed to the box containing the hazelnut in the prior tests of Experiment 1 and 2 ($\chi^2 = 9.308$, $p < .01$). On average 68.53 % of the choices were correct throughout trial one to nine with most errors occurring during the initial trials. A Wilcoxon test revealed a significant difference in the frequencies of correct choices between trial 1-3 and trial 7-9 ($z = -3.355$, $p < .001$). In the last three trials 87.74 % of the choices were correct, in 10.52 % of the cases the prior rewarded box was chosen and 1.74 % of the cases the gravity box.

4.6.3 Discussion

Experiment 3b confirms that apes select the box where the object was found in Experiment b1 and 2 and not the box specified by the gravity line. Also it corroborates the impression that apes more quickly come to localize the object than in Experiment 1. Taken together this suggests that great apes a) do not appear to use information about the tubes function, b) do not rely on a naïve theory of gravity when selecting a box, c) learn to localize the correct goal location and d) mainly use a position strategy to do so. However, that the strategy to select the prior enforced location is the only factor that has an impact on the apes' behavior in the tubes task is doubtful given the results of Experiment 2. If apes would merely adopt a position strategy they would not differ in their reactions to the real and the painted tube. Experiment 4 further investigates this issue by and confronts the apes with a situation where both a real and a painted tube are at choice.

4.7 Experiment 4

The results of Experiment 2 and Experiments 3a and 3b seem somewhat contradictory. Findings of Experiment 2 can only be explained by assuming that apes have at least to a certain degree an intuitive sensitivity to the functional properties of the tubes' mechanism and are able to integrate this information to make inferences about the future location of an invisibly displaced object. This can not be a very elaborate understanding and has only minor influence on box selection since apes are initially unable to find the object. Findings of

Experiment 3a and 3b suggest that apes mainly rely on position information to localize the object. However, in Experiment 1 and 2 the object could actually be found repeatedly at the same location, it could be that under these circumstances position advanced to the major cue to localize the object. Eventually most apes were able to abandon their position preference and to select the new location. In Experiment 4 only apes that passed Experiment 3a or 3b are confronted with a conflict situation where both a real and painted tube are presented simultaneously. To successfully localize the object in Experiment 4 apes must a) differentiate between a causal and a non-causal tube and can not b) refer to position information.

4.7.1 Method

Participants

Only the 11 apes that passed either Experiment 3a or 3b were tested on Experiment 4. 2 gorillas, 3 orangutans, 3 bonobos and 3 chimpanzees were remaining in the sample.

Apparatus

The apparatus was the same as used in the other tests with one alteration. In addition to a silent tube with a blocked neck also a gloss paint pattern tube painted on a plexiglas was inserted into the apparatus. Instead of a single tube, now both a tube *and* a gloss paint pattern “connected” the upper middle chimney to the lower-left, or the lower-right goal box, respectively.

Procedure

Only one test session consisting of 9 trials was run. Apes that searched correctly in 7 out of 9 trials were judged to have passed the conflict test. The procedure was identical to that of Experiment 2 with the following exceptions. All apes started with an arrangement where the silent tube was inserted as in Experiment 3b (e.g. upper middle-lower left) and the gloss paint pattern inserted the opposite way (e.g. upper-middle-lower-right). After each trial the configuration was switched (e.g. tube now upper-middle-lower-right and gloss paint pattern now upper middle-lower left).

4.7.2 Results

In Experiment 4 27.3 % of the apes reached criterion (n. s.).

Three of 11 apes chose the correct box already in the first trial of Experiment 5 (this does not differ from the expected 3.6 correct responses by chance) and 7 chose the box that was rewarded in Experiment 1 and 2 (binominal test, $p < .05$). Only 1 ape chose the middle (gravity) box and 7 of the 8 erroneous choices were directed to the box where the hazelnut could be found during prior tests of Experiment 1 and 2 ($\chi^2 = 4.500, p < .05$).

On average 48.36 % of the choices were correct throughout trial 1 to 9. Performance improved over trials: A Wilcoxon test revealed a significant difference in the frequencies of correct choices between trial 1-3 and trial 7-9 ($z = -2.013, p < .05$). In trial 7-9 60.64 % of the choices were directed to the box connected to the tube and 39.36 % of the choices were directed to the gloss paint pattern box. The middle (gravity) box was never chosen in the last three trials of the test.

4.7.3 Discussion

The major part of the apes did not pay attention if a real tube or a painted tube was “connected” to a goal box but clearly preferred the box where the object could be found in Experiment 1 and 2: They reliably often selected the position box in trial 1; an average of about 50 % of the choices were correct from trial 1-9. This is exactly what can be expected if apes continue to select the position box. The middle (gravity) box was only scarcely selected. However, the number of correct choices significantly increased, indicating that some apes changed their search strategy throughout the test session. Also, 3 apes passed criterion of Experiment 4: They chose the box that was connected to real tube. To decide whether they understood the tubes mechanism or if they just adopted a more elaborate strategy (e.g. to select the box with the tube attached) is investigated in Experiment 5.

4.8 Experiment 5

In Experiment 5 the apes that passed Experiment 4 are confronted with two real tubes simultaneously. If they depend on the strategy to select the box with the tube attached they can no longer localize the object.

4.8.1 Method

Participants

Only apes that passed the conflict test in Experiment 4 were tested on Experiment 5. This was true for only 3 apes: 1 gorilla and 2 bonobos.

Apparatus

The apparatus was identical to that of the other test. Now two silent tubes with blocked necks were inserted into it. One tube connected the upper-left chimney with the lower right goal box and the other connected the upper-right chimney with the lower left goal box.

Procedure

Only one session with 9 trials was administered. If apes pointed to the correct goal box on first attempt in 7 out of 9 trials they were judged to have passed the test. The procedure was identical to that of Experiment 2. After each trial the dropping chimney was changed (e.g. first trial left, second trial right chimney).

4.8.2 Results

Only 1 of the 3 apes (the gorilla) reached criterion and passed this test condition. In the first trial 1 ape chose the correct box, the 2 others pointed at the gravity box. On average 33.33 % of the choices were correct. The middle box was never chosen. There was no improvement in the last trials, on contrary, the performance declined: While 44.33 % of the choices were correct on trial 1-3 only 22 % were correct on trial 7-9.

4.8.3 Discussion

2 of the 3 apes passed Experiment 4 were using the strategy to select the box with the tube attached. Only the gorilla passed this condition suggesting that he can infer the objects location using information regarding the tubes function.

4.9 General discussion

Great apes do not have a reliable tendency to select the gravity box when tested with the tubes task. In this respect they contrast to 2-year-old infants, old- and new world monkeys and dogs who were all found to be biased in this task context. Even though apes selected the gravity box more often than the correct and the middle box in the first trial, this preference did not reach statistical significance and was abandoned in a relatively short time. So, as dogs, but unlike human infants and cotton-top tamarins, apes seemed to be quite flexible in this task: The tendency to select the gravity box - if appearing at all – was not persistent over trials. However, this does not mean that the gravity error is non-existent in apes. When apes were confronted with a situation where no logical solution could be inferred they reverted to gravity responses as a default answer. Obviously also apes hold naïve beliefs about gravity function, but it is dormant and reappears only under specific circumstances. This corroborates the idea that when development progresses the gravity bias is not abandoned, but suppressed because the capacity to inhibit prepotent answers is increasing. Recently Hood, Wilson and Dyson (2006) showed that the gravity bias also reappears in 4 years old children if the recruitment of inhibitory mechanisms is hindered by increasing the cognitive load the task poses on infants.

The present study reveals that apes have some intuitions about the causal function of the tube, but that it is not substantial enough to support a true understanding of the task. Apes localize the object more often when they are presented with a three-dimensional tube in contrast to a two-dimensional painted line. This suggests first, that they are to some extent aware of the causal function that the tube exerts on the objects trajectory and second, that they are to some extent aware that the spatial extension is the fundamental property that supports the causal functioning. However, these intuitions seem to be of a very weak nature because they do not lead to consistent correct solutions in this task. This suggests that both, naïve gravity concepts and intuitions about the causal function of the tube have only weak impact on apes' behavior in this task context, leaving room for the adoption of various non-causal strategies.

The main strategy employed by the apes in this study was a position strategy, the same strategy also adopted by dogs in the Osthau et al. (2003) study. This might be an artifact of the repeated reinforcement: Apes (as the dogs) were tested over many trials with same tube configuration and therefore repeatedly rewarded to choose a specific goal box. This in itself is interesting because other non-human primates did not build up a position strategy under the

same circumstances but perseverated to choose the gravity location. Overall apes' use of strategies was indicative of a remarkable flexibility: Not all apes developed a position preference, and if so, the position strategy was quite easily abandoned when no more successful (e.g. in Experiment 3a, 3b and 4) and replaced by other more expedient strategies (e.g. the strategy to select the box with the tube attached).

An additional finding of the present study is that apes did not make substantial use of acoustic cues to track an invisible object in this task. The provision of visual information, however, had deep impact on the apes' ability to localize the object. While the apes search success was not affected if acoustic information was available or not, the object was found practically always when its motion through the tube was fully visible.

Finally, we identified between species differences as well as a substantial within species variance in this task. Gorillas obviously had more difficulties to retrieve the object than the three other species, they made more errors and they more often failed to reach criterion. But this species difference was outweighed by the variance on an individual level: Some apes were initially gravity biased, others chose correct already in the first trial. Not all apes adopted the same strategy to localize the object and not all were able to change strategy in a similarly flexible way. Most of the apes did not pass all test conditions, a minor part passed the majority of test conditions and only one ape passed each test of the present study. Unexpectedly, it was a gorilla what further emphasizes the great individual variability in this task.

In sum, the tube task seems to be much more demanding for great apes than the table task that they generally solved well. But despite their lack of understanding the tubes great apes were not found to be reliably biased towards the goal location specified by gravity. Moreover, apes seem to exhibit some sensitivity to the causal relation of tube, object motion and goal location. However, they were unable to adopt their intuitions to localize an object dropped down the tube. Further research should address the question why apes causal sensitivity remains subliminal and under which circumstances it could be utilized more yielding.

5. References

- Bliss, J., Ogborn, J., & Whitelock, D. (1989). Secondary school pupils' common sense theories of motion. *International Journal of Science Education, II*, 261-272.
- Cacchione, T., & Krist, H. (2004). Recognizing impossible object relations: Intuitions about support in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 118*, 140-148.
- Cacchione, T., Call, J., & Zingg, R. Gravity and solidity in great apes: The table task. *In preparation.*
- Hauser, M. D., & Santos, L. R. (2002). A non-human primate's understanding of solidity: Dissociations between seeing and acting. *Developmental Science, 5*, F1-F5.
- Hauser, M. D., Williams, T., Kralik, J. D., & Moskovitz, D. (2001). What guides a search for food that has disappeared? Experiments on cotton-top tamarins (*Saguinus oedipus oedipus*). *Journal of Comparative Psychology, 115*, 140-151.
- Hood, B. M. (1995). Gravity rules for 2-4-year-olds? *Cognitive Development, 10*, 577-598.
- Hood, B. M. (1998). Gravity does rule for falling events. *Developmental Science, 1*, 59-64.
- Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. (1999). Gravity biases in a nonhuman primate? *Developmental Science, 2*, 35-41.
- Hood, B. M., Carey, S., & Prasada, S. (2000). Predicting the outcomes of physical events. *Child Development, 71*, 1540-1554.
- Hood, B. M., Santos, L., & Fieselman, S. (2000). Two year-old's naïve predictions for horizontal trajectories. *Developmental Science, 3*, 328-332.
- Hood, B. M., Wilson, A., & Dyson, S. (2006). The effect of divided attention on inhibiting the gravity error. *Developmental Science, 9*, 303-308.
- Kaiser, M. K., Proffitt, D. R. & McCloskey, M. (1985b). The development of beliefs about falling objects. *Perception & Psychophysics, 38(6)*, 533-539.
- Krist, H. (2000). Development of naïve beliefs about moving objects: The straight-down belief in action. *Cognitive Development, 15*, 281-308.
- Osthaus, B., Slater, A. M., & Lea, S. E. G. (2003). Can dogs defy gravity? A comparison with the human infant and a non-human primate. *Developmental Science, 6*, 489-497.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605-632.

Curriculum Vitae

Trix Cacchione

Luzernerstrasse 79

6030 Ebikon

041/ 450 07 14

t.cacchione@psychologie.uzh.ch

Geburtsdatum 10.07.1970
Zivilstand ledig
Heimatorte Luzern, Rom

Ausbildung

1977-1983 Primarschule, Maihofschulhaus, Luzern

1983-1986 Sekundarschule, Mariahilfsschulhaus, Luzern

1986-1991 Lehrerseminar der Stadt Luzern

1991-1992 Deutschlehrerin und Betreuerin im Durchgangszentrum für AsylbewerberInnen, Morschach (SZ)

1993-1995 Deutschlehrerin in den Gemeindeschulen Kriens, LU (Deutsch für Fremdsprachige, Deutsch für Asylbewerberkinder) sowie diverse Stellvertretungen an der 1. bis 6. Primarstufe (Schulen der Stadt Luzern, Gemeindeschulen Kriens, LU)

1995-1996 Je ein Semester Grundstudium der Biologie und der Kunstgeschichte; zwei Semester Latein, Universität Zürich

1996-2003 Studium der Allgemeinen und Entwicklungspsychologie im Hauptfach, der Soziologie im 1. Nebenfach und der Ethnologie im 2. Nebenfach, Universität Zürich

- 2003 Lizentiatsarbeit: „The development of relational object concepts:
Representation of support in chimpanzees“, Universität Zürich
- 2003- Assistentin / Doktorandin am Psychologischen Institut, Abteilung Allgemeine
und Entwicklungspsychologie, Universität Zürich
- 2004-2006 Forschungsaufenthalt im Rahmen der Dissertation am Max Planck Institut für
Evolutionäre Anthropologie