Tool use in great apes and human children : the impact of prior experience and visual feedback

Sonja Jördis Ebel

A thesis submitted for the degree of PhD
at the
University of St Andrews

2019

Identifier to use to cite or link to this thesis:

DOI: https://doi.org/10.17630/10023-18968

Candidate's declaration

I, Sonja Jördis Ebel, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 52,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in June 2015.

I, Sonja Jördis Ebel, received assistance in the writing of this thesis in respect of Y, which was provided by Christoph Völter; Daniel Hanus; Esther Herrmann; Martin Schmelz; Andreas Domberg; Sarah Peoples.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Sonja Jördis Ebel, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

Embargo on all of print copy for a period of 5 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for printed embargo request

Studies still have to be published

Electronic copy

Embargo on all of electronic copy for a period of 5 years on the following ground(s):

- Publication would preclude future publication

Supporting statement for electronic embargo request

Studies still have to be published.

Title and Abstract

- I agree to the title and abstract being published.

Date Signature of candidate

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Sonja Jördis Ebel, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date Signature of candidate

Acknowledgements

I would like to start off by thanking Josep Call for giving me the opportunity to study the cognition of great apes, which has been a wonderful journey. Planning and conducting studies with him, Christoph Völter, Esther Herrmann, Daniel Hanus, and Martin Schmelz was a delightful experience and a pleasure that I do not want to miss. Many thanks to all of them. I would also like to thank Michael Tomasello for giving me the opportunity to join his department at MPI-EVA in Leipzig, including trips to international conferences. Many thanks to all research assistants at the Wolfgang Köhler Primate Research Center (WKPRC) at Leipzig Zoo and at the child lab at MPI-EVA – but especially to Hanna Petschauer and Susanne Mauritz – for their help with organization of the studies. Further thanks to Ilona Schappert for giving me the opportunity to collect data with the orang-utans at Dortmund Zoo. Moreover, I would like to thank all the caregivers of the apes at WKPRC, Sweetwaters Chimpanzee Sanctuary and Dortmund Zoo involved with data collection for their great support, good conversations and a delightful time.

Special thanks to Esther Herrmann for all her work and great support around the research trip to Sweetwaters Chimpanzee Sanctuary at the Ol Pejeta Conservancy in Kenya. Many thanks to Henriette Zeidler, who helped with organization during my stay, Jan Engelmann for great conversations, and Josep Call and Martin Schmelz for giving me the opportunity to conduct this trip in the first place. Many thanks to Richard Vigne, Stephen Ngulu, the board members, and the staff of Sweetwaters Chimpanzee Sanctuary for their great support. I would also like to thank Ol Pejeta Conservancy, Kenya Wildlife Service (KWS), and the National Council for Science and Technology (NCST) for approving the research.

Moreover, I would like to thank Roger Mundry and Colleen Stephens for their excellent statistical advice that lent a great deal of sophistication to the analyses. Many thanks to Josep Call, Christoph Völter, Daniel Hanus, Esther Herrmann, Martin Schmelz, Andreas Domberg, and Sarah Peoples for commenting on various parts of the draft for this thesis. I would also like to thank Sylvio Tüpke for creating the figures of the experimental setups in Chapter 3, 4, and 5B-C and Laura Därr for the one in Chapter 5D. Many thanks to Raik Pieszek for building the apparatus used in the experiments reported in Chapter 3. I would also like to thank the many student assistants and interns for their help with data collection with human children, and Sarah Peoples, Elisabeth Beier and Lisa Wartig for their help with reliability coding.

Furthermore, I would like to thank my colleagues and dear friends who joined me on the journey during my pleasant time at MPI-EVA in Leipzig. First of all, I would like to mention the coffee club, at which the “zoo people” came together as a daily routine. Joint trips to conferences completed my experience with them. Many thanks to Johanna Eckert, Christoph Völter, Alejandro Sanchez-Amaro, Martin Schmelz, Shona Duguid, Manuel Bohn, Matthias Allritz, Katharina Helming, Marina Proft, Andreas Domberg, Sarah Peoples, Marie Nitzschner, Suska Nolte, Amy Lewis and Christopher Krupenye for discussions, laughter, and making the MPI such a nice working environment. Additionally, I would like to thank Katja Liebal for giving me the stepladder to study ape cognition for my master’s thesis and being a great mentor.

Finally – and most importantly –, I would like to thank my family (Romy, Christian, Jonas, Fridtjof, Britta, and Marianne Ebel) for their constant support and patience throughout my studies and the Ph.D. Moreover, I would like to thank Christina Freihorst, Konstantin Richter, Kathrin Kopp, Christian Kietzmann and Jana Heidenfelder for their support and encouragement – and I ask indulgence from those I omitted here. My special thanks to all the children and apes who participated in the studies and I hope that they enjoyed being part of the scientific endeavour as well.

Funding

This work was supported by the Max Planck Society [Department for Developmental and Comparative Psychology at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany].

Contents

List of Tables

List of Figures

List of Appendices

Abstract

Human and primate tool use has been the focus of intensive research for many decades. Studies with non-human great apes are of special interest for the question when certain cognitive abilities evolved. This thesis investigates the role of prior experience and visual feedback in great apes' and human children's tool use. Prior experience with tools is normally regarded as beneficial, helping individuals to find successful strategies. Also, visual feedback and additional information about the solution of a problem can deliver crucial insight into task components. Following an introductory and a methodological chapter, Chapter 3 explores the role of visual feedback and additional information in great ape problem-solving using the Floating Peanut Task (FPT), which requires pouring water into a tube to extract an object. Findings suggest that visual feedback was necessary for success at first, but later became redundant, and end-state information (seeing a water-filled tube) helped some individuals independently.

As a downside of experience, familiar strategies may restrict the analysis of novel problems. Most interestingly, prior use of a tool can discourage using it with a novel function (functional fixedness effect). Chapter 4 investigates functional fixedness in 6- to 8-year-old children using the FPT, focusing on how prior tool use and task presentation predict success. Findings suggest low success rates overall and no effect of experience; however, greater tool salience increased success. Chapter 5 investigates functional fixedness in great apes, varying their experience with three tools to be used each with a novel function. Prior experience lowered success and increased latency on

novel problems, and prior use as a food item kept apes from using a bread stick as a raking tool. Chapter 6 discusses the overall findings in terms of the evolutionary origins of the negative impact of prior experience with tools, object representations, and learning mechanisms.

Chapter 1: General introduction[1]

Summary

In this thesis, I focus on the role of prior experience in great apes' and human children's tool use. This first chapter consists of three parts in which I introduce the notion of tool use, the Floating Peanut Task (FPT), and the functional fixedness effect. First, I define tool use and differentiate it from other terms, such as problem solving and object manipulation. I give an overview about tool use in our closest living relatives, the non-human primates, which is of special interest to better understand the evolution of human physical cognition in general and of tool use in particular. I further provide an overview about the cognitive underpinnings of tool use.

Second, I introduce the FPT, which requires subjects to pour water into a vertical tube to retrieve a floating object. I define tool innovation and tool knowledge and relate the FPT to two additional innovation problems, namely the hook task and Aesop's Fable task. I then propose a shared structure in all three problems. The hook task requires bending a wire into a hook to retrieve from a vertical tube a small bucket containing a reward. Aesop's Fable task entails dropping stones into a vertical tube that contains some water to make the water level rise until a floating object can be

---

[1] Parts of the material from this chapter formed the basis for the introductions of the following papers. Under review: Ebel, S. J.; Schmelz, M.; Herrmann, E.; Call, J.: Innovative problem solving in great apes is fostered by visual feedback and the end-state of the solution. In preparation for submission: Ebel, S. J.; Völter, C. J.; Call., J.: Functional fixation in the tool use of captive great apes (*Pan paniscus, Pan troglodytes, Pongo abelii*); Ebel, S. J.; Völter, C. J.; Call., J.: Functional fixedness in great apes invoked by a food item.

reached. I then elaborate on tool knowledge and discuss studies with the FPT in human children and great apes.

Third, I review the literature on the role of prior experience in problem solving in general and tool use in particular. Prior experience has a positive effect on problem-solving skills in many cases, however, it can also lead to a decrease in performance. This detrimental effect of experience is the one I focus on in this thesis. More specifically, I investigate the concept of functional fixedness, which refers to the negative impact of experience with tools by blocking the employment of novel functions. I further give an overview about studies that have been conducted with human children and great apes.

A. Tool use

An orang-utan craving for a peanut inside a tube eagerly tried to bite open the Plexiglas that the tube was made from. She found some water inside the tube on which the peanut was floating, but it was too far away from the top of the tube to reach it with her fingers. After minutes of relentless trying, she turned her head around and peered at the upper corner of the room where a water dispenser was installed. Trying one more time to extract the peanut with her fingers, she climbed up to the water dispenser and returned with a mouthful of water that she added to the tube. Briefly assessing the effect of her actions by peering into the tube, she collected another mouthful of water, thus raising the water level to a point where she could reach the peanut, satisfying her craving (personal communication D. Hanus; Mendes, Hanus, & Call, 2007).

This example of the orang-utan Toba solving the Floating Peanut Task (FPT) is an illustrious example of non-human great apes' impressive tool-use abilities (henceforth: great apes or apes): The orang-utan used water as a tool to access a food reward. While about 150 years ago, people were stunned when Darwin came to the conclusion that humans descended from an ape-like creature (Darwin, 1859, 1871), nowadays no scientifically educated person will deny our evolutionary heritage (Mayr, 2002). Since the first establishment of comparative psychology – the study of non-human animal cognition – as a discipline by C. Lloyd Morgan (1882, 1890–1891, 1894), researchers have explored the impressive minds of monkeys and apes and have created a more and more complete picture of our shared evolutionary history (Tomasello, 2014;

Tomasello & Call, 1997). Wolfgang Köhler (1925) and Robert Yerkes (1916) were two pioneers in the field of primate cognition research who further set the field in motion. Their books “The mentality of apes“ and “The mental life of monkeys and apes: a study of ideational behaviour” that were first published 1917 and 1916 respectively have influenced generations of researchers and still continue doing so. However, before we delve into the multifaceted world of primate cognition and tool use in particular: What is tool use anyway?

1. What is tool use

Tool use has been studied from many different perspectives and is the topic of ongoing interdisciplinary research (e.g., McCormack, Hoerl, & Butterfill, 2011; van Elk, van Schie, & Bekkering, 2014) and of animal research in particular (e.g., Sanz, Call, & Boesch, 2013; Shumaker, Walkup, & Beck, 2011). Tools are generally described as objects that are used as means to an end by altering the position of other objects (Tomasello & Call, 1997). Shumaker and colleagues (2011) define tool use in the following way (see also Beck, 1980):

> *“The external employment of an unattached or manipulable attached environmental object to alter more efficiently the form, position, or condition of another object, another organism, or the user itself, when the user holds and directly manipulates the tool during or prior to use and is responsible for the proper and effective orientation of the tool.”*

This definition entails many different types of behaviours such as using objects for protection (e.g., a hermit crab using a snail shell), for modifying one's senses (e.g., the usage of glasses in humans), or to alter the state of another individual (e.g., a chimpanzee poking a conspecific with a stick). Since I concentrate on tool use in physical problems in this thesis, mostly in the foraging context, I define tool use more narrowly. This allows me to focus on the body of literature most relevant for my research. Tool use is defined as

> *"[t]he external employment of an unattached (…) environmental object to alter (…) the (…) position (…) of [other] object[s] (…), when the user holds and directly manipulates the tool during or prior to use and is responsible for the proper and effective orientation of the tool."* (adapted from Shumaker et al., 2011)

This definition describes tool use as altering the position of an object by another object, leading to the question if any object can be a tool or if tools are somewhat special.

Tools, objects, and artefacts

Objects are permanent entities that have certain properties and that can be manipulated (Santos & Hood, 2009). Some authors use the terms "object" and "tool" interchangeably (Greif & Needham, 2011; van Elk et al., 2014), whereas others have argued for a conceptual distinction (e.g., Rothi, Ochipa, & Heilman, 1991). The latter refer to the hierarchical structure between tool and object: The tool is used to change the position of the object, and not vice versa. For example, a stick is used to rake in a food item (see Figure 1.1B). However, the former argue that any object can be used as a tool so that the distinction may be true for a specific tool-use event, but that it is not relevant for classifying objects versus tools. For example, a food item can also be used to reach a stick, too (see Chapter 5). I follow this idea and use “tool” and “object” interchangeably throughout the thesis.

Research in humans further involves the notion of “artefacts”. Artefacts are objects that have been made by an agent for a specific purpose and are therefore mainly defined by their function (Elsner & Pauen, 2007; Nelson, 1973). Most of the objects in humans’ environment concern artefacts, whereas most objects in apes’ (natural) environment concern non-modified (natural) objects such as twigs or stones (although wild chimpanzees also manufacture tools such as termite-fishing tools that have a brush-like end; e.g., Sanz, Call, & Morgan, 2009). Humans typically reason about what an artefact has been made for and even young children take teleological-intentional information into account when using tools (design stance; e.g., Casler & Kelemen, 2005, 2007; Csibra & Gergely, 2007, 2009; Hernik & Csibra, 2009; Hernik & Csibra, 2015; Ruiz & Santos, 2013). It seems unlikely that apes engage in such an activity as well, given their natural ecology. Since I aim at comparing great apes and human children more closely, I refer to “objects” and “tools” in this thesis only. However, having so many artefacts in their environment may have shaped humans’ understanding of the physical world in general and that of tools in particular. Thus, I discuss the design stance and potential differences in object representations between human children and great apes further in Chapter 6.

Figure 1.1 Two examples of ape problem-solving involving tool use: The ape can access a food reward by dropping a stone into an apparatus that releases a platform located inside the apparatus (A) or by raking it in with a wooden stick. (The pictures are taken from Ebel and Call (2018) and a study presented in Chapter 5 respectively.)

Object manipulation and tool use

I have defined objects as permanent entities that have certain features and that can be manipulated (Santos & Hood, 2009). Object manipulation refers to touching, moving, or transforming an object with one's own body parts, usually with the hands and mouths in the case of primates (see also Burghardt, 2006; Glickman & Sroges, 1966). Object manipulation and tool use are often regarded as a continuum with increasing complexity from the former to the latter (Greif & Needham, 2011; Parker & Gibson, 1977; Piaget, 1977): Since an individual has to relate her body to the tool and the tool to another object during tool use, this activity seems more complex than object manipulations in which she only refers the object towards her own body (Parker & Gibson, 1977; Tomasello & Call, 1997). I discuss the continuum between object manipulation and tool use further in the following section about the evolutionary origins of tool use.

Objects are entities that are acted upon (Santos & Hood, 2009). Object properties therefore influence the type of manipulations that can be performed with an object; these manipulations may vary with an object's shape, size, material etc. (Lockman, 2000). For example, a wooden stick may be used to reach items that are located out of reach, but it could not be used for sponging water due to its solidity. Eleanor and James Gibson have coined the term "object affordances", which describes the possibility of actions that can be done with an object (e.g., Gibson, 1982). This approach focuses on the object itself; it mainly refers to processes of perception and action (see also Lockman, 2000). Remarkably, when Worgotter et al. (2013)

investigated manipulation actions based on hand-object relations in humans, they found less than 30 types of manipulations. This suggests that there is a limited number of manual manipulations that can be performed with an object such as put on top or put together, push on top or push together, put over, push over etc. (these four examples belong to the category “release determined actions” that serve the goal to hide or to construct; see Worgotter et al., 2013).

Tool use and problem solving

Tools are often employed when an individual cannot reach her goal directly. For example, when a food item lies out of her reach, a chimpanzee might use a wooden stick to rake in the food. Thus, she has to find a way to overcome an obstacle to reach her goal, that is, she has to solve a problem. Figure 1.1 presents two examples for problem solving involving tool use in captive great apes: In Figure 1.1A, an orang-utan male is confronted with an apparatus that requires him to drop a stone inside a tube. The stone causes the collapsing of a platform inside the apparatus, which releases a food reward. In Figure 1.1B a bonobo female is presented with an out of reach reward on a platform. She uses a wooden stick to rake in the food. While the first example involves a tool that has to be released from the ape’s hand, the second one requires the contact with the tool until the food is reached. Seed and Mayer (2017) characterize the relationship of problem solving and tool use as follows: although tool use can be part of the broad domain of problem solving, not all cases of tool use classify as such.

Problem solving is defined by

> *"overcoming some obstacle to achieve a goal when the entire solution is neither in the species-typical repertoire nor socially learned."* (Seed & Mayer, 2017)

For example, many cases of hard-wired, rigid, and socially learned tool use would not fit Seed and Mayer's definition of problem solving. On the contrary, problem solving comprises such diverse problem-solving situations as taking detours and short cuts, puzzle boxes, executive function tasks etc., which may rely on cognitive processes such as perception, representation, learning, memory, planning, or decision-making (DeLoache, Miller, & Pierroutsakos, 1998; Seed & Mayer, 2017). DeLoache et al. (1998) define problem solving more broadly as a "goal-directed cognitive activity" that usually involves inference. Inference is further defined as "going beyond the information given to reach a new conclusion, form of generalization, find a solution" (DeLoache et al., 1998). Since I focus on problem solving involving tool use in this thesis, I define problem solving as:

> *"overcoming some [physical] obstacle to achieve a goal [that involves the usage of a tool] when the entire solution is [not] in the species-typical repertoire (…)."*
> (adapted from Seed & Mayer, 2017)

I investigate problem solving in the context of individual and social learning in this thesis. Thus, I do not restrict the definition to problem solving that is solely based on the physical information that an individual gathers in the problem-solving situation. However, an important aspect of the definition is the fact that it excludes species-typical behaviour such as more hard-wired tool use (although there may be a continuum from species-typical to more flexible behaviours, see Seed & Mayer, 2017). In addition, I use the term “problem” and “task” interchangeably throughout the thesis. The problems that I used for my studies with the great apes were foraging tasks, which required extracting food from an apparatus with a tool. In case of the children, a token was used instead of food that could be exchanged for some stickers later.

Flexible tool use

Not all occurrences of tool use are intelligent, but many are hard-wired and inflexible (Shumaker et al., 2011). Hard-wired tool use represents an adaptation to a specific niche with little variation in its execution; the ability to somewhat flexibly use tools (as any other cognitive adaptation) is shown by variation and requires the control by the individual (Tomasello & Call, 1997). Thus, *flexibility* and *selection* are hallmarks of intelligent tool use which may be based on some form of causal understanding and planning (Byrne, Sanz, & Morgan, 2013; Seed & Byrne, 2010; Tomasello & Call, 1997). The term “flexible tool use” often refers to the ability of a species or an individual to use multiple tools for one purpose and one tool for multiple purposes (Greif & Needham, 2011). Call (2013) suggests that flexible (or creative) tool use is based on three components: the ability to accumulate knowledge, the ability to recombine pieces of information in novel ways, and a disposition for object manipulation (Call, 2013).

Focus of this thesis: tool use in great apes

Finally, although tool use is extremely rare among the animal kingdom, there is still a variety of species using tools such as New Caledonian crows or capuchin monkeys (Shumaker et al., 2011). Yet, in this thesis I mainly concentrate on humans and great apes because I am interested in the evolution of human physical cognition. Therefore, it is most relevant to study our closest living relatives (i.e., bonobos, chimpanzees, gorillas, and orang-utans) and to compare them to human children (Nielsen & Haun, 2016; Tomasello, 2014). For example, tool use in great apes and New Caledonian crows may resemble an interesting case of convergent evolution (Mayr, 2002), yet, this is not the focus of this thesis. Moreover, I concentrate on studies with great apes more narrowly. The topics that I discuss throughout this thesis like the mechanisms underlying the FPT (Chapter 3 and 4) or the functional fixedness effect (Chapter 4 and 5) seem to be above the capacities of most monkey species who may lack the required causal understanding (Visalberghi, Fragaszy, & Savage-Rumbaugh, 1995; Visalberghi & Trinca, 1989). Additionally, I mainly refer to studies that have been conducted with captive great apes. Studies in captivity allow for using controlled experiments to investigate cognitive mechanisms more specifically (Tomasello, 2014). Moreover, only chimpanzees and Sumatran orang-utans habitually use tools in the wild, whereas all great ape species skilfully use tools in captivity (Köhler, 1925; Lethmate, 1977; Tomasello & Call, 1997). This shows that all great ape species have the capacity to use tools, but not all of them develop this behaviour in the wild. I discuss studies with wild great apes throughout the thesis whenever they yield insights on apes' cognitive capacities.

As demonstrated, tool use is a fascinating research topic. However, why does tool use matter from an evolutionary perspective? To better understand human and great ape tool-use, I give an overview about its evolutionary origins.

2. Evolutionary origins

As demonstrated, tool use is necessary when faced with a problem that requires moving an object (e.g., a food reward) so that it comes into reach. In order to better understand the evolutionary origins of human tool use, it is important to look at primates' general manipulation behaviour of objects. Even though also few primate species habitually use tools in the wild, quite a few of them have evolved a specific morphological adaptation that generally facilitates tool use: Their hands allow them to grasp and move objects (Shumaker et al., 2011; Tomasello & Call, 1997). Thus, it may be possible to further answer the following questions: Why has tool use evolved in some primate species, but not in others? Which environmental factors bring about tool use?

In order to answer these questions I first outline the ancestral tree of the Old World monkeys and apes to give a better understanding of the evolutionary timeline. I then present Piaget's six stages of sensorimotor intelligence in human children, which serve as a foundation to discuss object manipulation and tool use in primates. I present diverse hypotheses about the occurrence of tool use in primate species such as the feeding adaptation hypothesis and discuss the influence of factors such as terrestriality and food abundance.

The ancestral tree of great apes

Primates living today can be classified into two major subgroups (Fleagle, 1988): prosimians (lemurs, lorises, and tarsiers) and simians (monkeys and apes, including humans). In the past, prosimians were not considered to be “real” primates (reflected by the German word “Halbaffen” which means “half-ape”), but by now a close relationship to monkeys and apes has been demonstrated (Fleagle, 1988; Silvertown, 2008). Tarsiers, monkeys, and apes further belong to a suborder of primates referred to as “haplorhini” (i.e., "dry-nosed" primates) and lemurs and lorises as “strepsirrhini” (i.e., “wet-nosed” primates; Fleagle, 1988). Haplorhini can be further differentiated into New World monkeys, Old World monkeys, and apes. In this thesis, I focus on apes who diverged from Old World monkeys about 23 million years ago (Silvertown, 2008; Figure 1.2).

Humans’ most distant ape relatives are the lesser apes (gibbons and siamangs) that diverged from the apes about 17 million years ago. Great apes further diverged into orang-utans (about 14 million years ago), gorillas (about 8 million years ago), and chimpanzees and humans (about 6 million years ago; Silvertown, 2008; Figure 1.2). The great ape genera consist of eight species that are living today (and further subspecies that are not listed here): The *Homo* species(*Homo sapiens*), two *Pan* species, chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*), two *Gorilla* species, Western gorillas (*Gorilla gorilla*) and Eastern gorillas (*Gorilla beringei*), and three *Pongo* species, Bornean orang-utans (*Pongo pygmaeus*), Sumatran orang-utans (*Pongo abelii*), and Tapanuli orang-utans (*Pongo tapanuliensis*; not depicted in Figure 1.2). The latter has only recently been discovered (Nater et al., 2017). Figure 1.2 illustrates the ancestral tree of the Old World monkeys and apes.

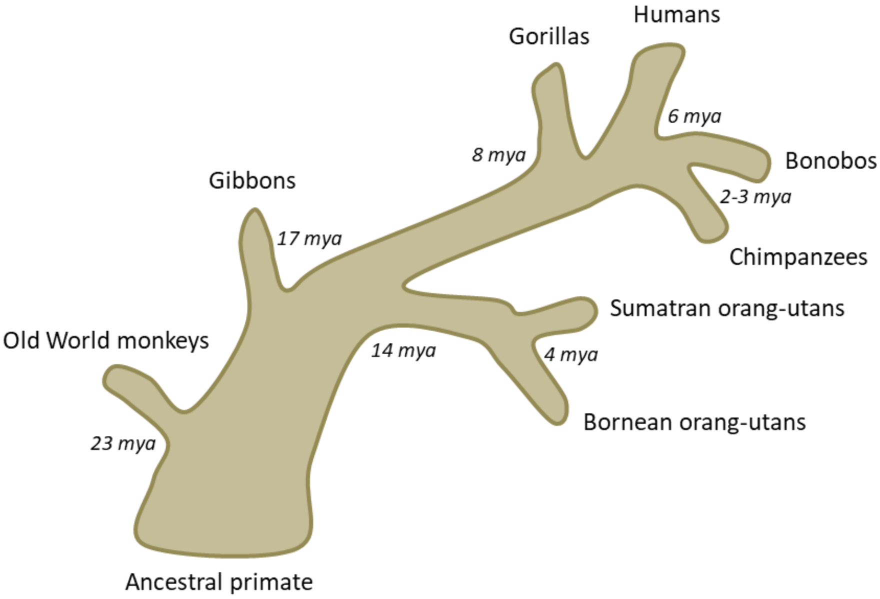


Figure 1.2 The ancestral tree of the Old World monkeys and apes.

The comparative approach

Wild chimpanzees have been shown to use multiple tools on a regular basis, the same tool for multiple purposes, and tool sets (Boesch, Head, & Robbins, 2009; Brewer & McGrew, 1990; Sanz et al., 2009; Tomasello & Call, 1997; Whiten & Boesch, 2001).

Sumatran orang-utans also habitually use tools in the wild (van Schaik et al., 2003). However, all great apes are adept tool users in captivity (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007; Herrmann, Wobber, & Call, 2008; Köhler, 1925; Lethmate, 1977; Manrique, Gross, & Call, 2010; Martin-Ordas, Schumacher, & Call, 2012), so what can be learned from this finding?

Extant great apes species shared a last common ancestor about 14 million years ago (Fleagle, 1988; Silvertown, 2008; Figure 1.2). Thus, finding the same cognitive adaptations in orang-utans, gorillas, chimpanzees and humans would indicate that these cognitive abilities have evolved in their last common ancestor (Tomasello, 2014; Tomasello & Call, 1997). The occurrence of the same cognitive ability in these genera could theoretically also be taken as a case of convergent evolution. However, it is rather unlikely that the same ability has evolved multiple times among such close relatives. A comparison of living great ape species with human children can therefore be seen as a fruitful approach to a better understanding of the evolution of human cognition (Moll & Tomasello, 2007; Nielsen & Haun, 2016; Tomasello, 2014; Tomasello & Call, 1997; Tomasello, Carpenter, Call, Behne, & Moll, 2005).

The evolution of tool use in primate species may have its precursor in the manipulation of objects. Piaget (1977) has defined six stages of sensorimotor intelligence in the first two years of a child’s life. Parker and Gibson (1977) used this classification system to study object manipulations in non-human primates. Piaget’s stages and Parker & Gibson’s findings are presented in the subsequent sections.

Table 1.1 Piaget's six stages of sensorimotor intelligence.

Stage	Learned schemes	Behavioural manifestation	Age
1	*Reflexes*	Actions under little voluntary control	0-1 month
2	*Primary circular reactions and first acquired adaptations*	Actions directed to one's own body	1-4 months
3	*Secondary circular reactions*	Actions aimed at reproducing interesting effects in the environment	4-8 months
4	*Coordination of secondary schemata and their applications to new situations*	Hierarchical embedding of secondary schemes (differentiation between means and ends; intentionality)	8-12 months
5	*Tertiary circular reactions and discovery of new means through active experimentation*	Actions aimed at relating external entities to one another	12-18 months
6[1]	*The invention of new means through mental recombination*	Insightful problem solving and imitation, internalized trial and error; variation, differentiation and recombination of schemes	18-24 months

Adapted from Parker and Gibson (1977); Piaget (1977); Tomasello and Call (1997)

[1] Tomasello and Call (1997) note that stage 6 is not strictly about sensorimotor skills as it explicitly involves mental representations.

Piaget's six stages of sensorimotor intelligence

Object manipulation on the one side (i.e., the handling of enduring physical entities that have certain properties; Santos & Hood, 2009) and tool use on the other side (i.e., the usage of an objects to alter other objects in the environment; Greif & Needham, 2011) can be seen as a ontogenetic and phylogenetic continuum (Greif & Needham, 2011; Parker & Gibson, 1977; Piaget, 1977; Vauclair, 1984). Parker and Gibson (1977) differentiate between simple prehension, simple object manipulation, object-substrate manipulation, complex object manipulation, and social-object manipulation (which can involve all other forms of object manipulation).

The developmental psychologist Jean Piaget (1977) proposed six distinct stages of sensorimotor intelligence in the development of human children (Table 1.1). The first four stages describe object manipulation in the first year of life. In the first stage, babies show reflexes with little voluntary control of their actions. In the second and third stage, they start to direct actions towards their own bodies and repeat actions on objects that caused interesting effects. In the fourth stage, they show hierarchical organisation of actions such as transferring an object from one hand to the other. They further differentiate between means and ends and understand intentionality. In the fifth stage, 1-year-old babies engage in purposeful trial-and-error learning and manipulate objects to discover new means to old and new ends. In the sixth stage, 1- to 1.5-year-olds engage in insightful problem solving and exhibit a mental recombination of learned behaviour sequences such as using a stick to access an out of reach reward without overt trial-and-error behaviour (Parker & Gibson, 1977; Piaget, 1977; Tomasello & Call, 1997; Table 1.1). What are the implications of these developmental steps during human ontogeny for the evolution of tool use in primate species?

The feeding adaptation hypothesis

Primate species differ vastly in their interest towards objects and the complexity of their object manipulations (e.g., Burghardt, 2006; Glickman & Sroges, 1966; Jolly, 1964a, 1964b; Jordan, 1982; Parker & Gibson, 1977; Poti & Spinozzi, 1994; Takeshita & Walraven, 1996; Tomasello & Call, 1997; Torigoe, 1985; Vauclair, 1984; Welker, 1956). For example, Torigoe (1985) explored object manipulations with a nylon rope and a wooden cube in 74 primate species living in captivity. He found that lemurs, marmosets, spider monkeys, and leaf-eaters had the smallest manipulation repertoire among the species tested. Old World monkeys (except leaf-eaters) exhibited a medium repertoire and capuchin monkeys and apes displayed the largest repertoire (see also Glickman & Sroges, 1966). Additionally, when Glickman and Sroges (1966) compared two subfamilies of the Old World monkeys, they found a greater manipulation repertoire in cercopithecine monkeys (to which such species as baboons, macaques, or vervet monkeys belong) compared to colobine monkeys (to which such species as the black-and-white or the red colobus monkey belong). Among the cercopithecine monkeys, baboons and macaques showed an elaborated repertoire of object manipulations. An interesting difference between cercopithecine and colobus monkeys is that the former have an omnivorous diet comprising fruits, insects, and small vertebrates, whereas the latter feed on leaves (Glickman & Sroges, 1966).

Parker and Gibson (1977) more specifically proposed that object manipulations in primates resembled feeding adaptations. They used Piaget’s six stages of sensorimotor intelligence to analyse the complexity of object manipulation in capuchins and great apes. Their analyses taken together with the previously reported studies suggest that first, leaf-eating species such as prosimians, New World monkeys (except capuchins) and some Old World monkeys exhibit little diversity in their object manipulations. They do not have to handle objects during their daily feeding routine. This is also reflected by the missing physiological adaptations of their hands for object manipulation (Tomasello & Call, 1997). Second, frugivorous species such as Old World monkeys (except for leaf-eaters) show more diverse object manipulations since they regularly engage in object manipulations when feeding on fruits. Third, omnivorous species such as capuchin monkeys and apes exhibit the greatest diversity in their object manipulations, which their feeding routines may require. Sometimes this also involves the usage of tools to extract embedded foods (Glickman & Sroges, 1966; Jolly, 1964b; Parker & Gibson, 1977; Tomasello & Call, 1997; Torigoe, 1985).

This explains why frugivorous and omnivorous primate species living in captivity are more interested in novel objects: They have evolved a disposition to explore and handle objects which could potentially resemble a novel food source (please note that primates are often much more neophobic in the wild than they are in captivity though; Forss, Schuppli, Haiden, Zweifel, & van Schaik, 2015). Parker and Gibson (1977) further proposed that the most complex type of object manipulation, tool use, has evolved in capuchin monkeys and apes as an adaptation to only seasonally available embedded foods.

Nevertheless, not all differences between species in their object manipulation repertoire can be explained by their diet. For example, only looking at prosimians reveals a different picture: The group of prosimians interested most in objects consists of omnivorous, frugivorous, and herbivorous species, whereas the second most interested group consists of omnivorous species only, and the group interested least comprises insectivorous species (Jolly, 1964b). Torigoe (1985) also suggests that some type of object manipulations may have their origin in feeding contexts, while others may have developed in different contexts such as throwing in an agonistic social setting.

Moreover, the correlation between object manipulations with novel (inedible) objects and diet does not reveal why a certain species has evolved this specific diet. Parker and Gibson (1977) have proposed that tool use has evolved in capuchin monkeys and apes due to an environment with seasonally limited embedded foods. Two famous hypotheses concern the environmental and social conditions: the necessity and the opportunity hypothesis. The necessity hypothesis predicts that tool use develops as niche construction due to selection pressure; the opportunity hypothesis predicts that tool use emerges in species or populations that live terrestrially, in social groups, and that have high food abundance, i.e., species or populations which have the opportunity for tool use (e.g., Koops, McGrew, & Matsuzawa, 2013; Meulman, Sanz, Visalberghi, & van Schaik, 2012; Spagnoletti et al., 2012; Visalberghi, Fragaszy, Izar, & Ottoni, 2005). However, these two competing hypotheses are not discussed further at this point as ecological questions are not the focus of this thesis and are better investigated with wild animals than with captive ones.

3. Cognitive underpinnings

Tool use requires the coordination of one's own body with the tool and the tool with other objects. Thus, two important cognitive underpinnings of tool use are perception and action processes (i.e., sensorimotor skills; Greif & Needham, 2011; Roche, Blumenschine, & Shea, 2009; Seed & Byrne, 2010; Stout & Chaminade, 2007; Vaesen, 2012; Völter & Call, 2014a). Intelligent tool use may further involve goal-directedness, physical reasoning, and planning (Seed & Byrne, 2010). Furthermore, executive functions such inhibition, working memory, and attentional shifting are likely to play an important role in tool use (e.g., Meulman, Seed, & Mann, 2013; Seed & Byrne, 2010; Seed, Call, Emery, & Clayton, 2009; Vaesen, 2012), yet these are not the focus of this thesis. In this section, I provide an overview of cognitive underpinnings of intelligent tool use (Seed & Byrne, 2010).

Intelligent tool use

Intelligent tool use is goal-directed: The individual acts upon a goal and expects certain outcomes of her actions; the tool-use behaviour is *selective* and *flexible* (Seed & Byrne, 2010). For example, great apes choose tools with certain functional properties and manufacture tools (selective tool use); they also use one tool for several purposes and several tools for one purpose (flexible tool use; Shumaker et al., 2011).

Moreover, tool use is based on the comprehension of the causal relationships between objects (Seed & Byrne, 2010). However, other forms of physical reasoning can also lead to success in some tool use task, such as those based on certain perceptual strategies (Seed & Byrne, 2010). With regard to great apes, studies have shown that they understand causal relationships between objects (Call, 2004, 2007; Hanus & Call, 2008, 2011; Völter, Sentís, & Call, 2016). Moreover, when gathering or making a tool, some kind of planning is involved (Byrne et al., 2013; Seed & Byrne, 2010). The individual selects or manufactures the tool in relation to her higher order goal, sometimes this involves hierarchical planning. Great apes have been shown to use tools sequentially (i.e., use a tool to get a tool) and to use tool sets (i.e., several functionally distinct tools in sequence; Boesch, 2013; Boesch et al., 2009; Martin-Ordas et al., 2012).

Finally, the intelligent tool user may engage in insight by producing a novel solution, which is not part of her past behaviour (Seed & Byrne, 2010). I do not consider insight as a pre-requisite of intelligent tool use here. Firstly, insight is hard to assess in non-human animals because they cannot report on how they came up with the solution (Shettleworth, 2012). Secondly, the combination of selectivity and flexibility, planning, and physical reasoning already provides an adequate framework for intelligent tool use. However, what is insight anyway? Problem solving via insight is usually characterized by an initial period of unsuccessful attempts at finding the solution, followed by an impasse that suddenly dissolves and culminates in the solution to the problem (Bowden, Jung-Beeman, Fleck, & Kounios, 2005; Sternberg & Davidson, 1995). The term was coined by Gestalt psychologists in the early 20th century and was also used to describe the behaviour of chimpanzees (Köhler, 1925). It is still under debate if it is applicable in non-human animals (Shettleworth, 2012). One important aspect of studying insight is to investigate the role of visual feedback: For example, New Caledonian crows were presented with a string-pulling task that required them to pull up a string to which a food reward was attached (Taylor, Medina, et al., 2010). The crows solved the task when they received visual feedback for their action, but most of them failed when a platform restricted their view. These findings suggest that they may not have solved the task via insight (see Taylor, Medina, et al., 2010).

Furthermore, studies that investigate when problem-solving performance breaks down in a species deepen our understanding about cognitive underpinnings of tool use. For example, chimpanzees perform better in a trap or a maze task in which they have to plan their moves ahead to avoid traps, when they use their fingers than when they use tools (Seed et al., 2009; Völter & Call, 2014a). Thus, performance breaks down when one relation is added to the planning (i.e., hand – object versus hand – tool – object; see also Girndt, Meier, & Call, 2008). One explanation could be that action planning and tool use are based on the same neural substrates (Seed et al., 2009; Völter & Call, 2014a).

In this section, I have presented the most important cognitive underpinnings of intelligent tool use. Since I also refer to social learning mechanisms in this thesis, which allow the acquisition of tool use by observation, I briefly summarize the two most relevant ones in the subsequent section.

Social learning

Many different social learning strategies have been suggested throughout the literature (e.g., Galef & Laland, 2005; Heyes, 1994; Nehaniv & Dautenhahn, 2007). Since social learning is not the main focus of this thesis, I only briefly introduce the two most relevant social learning strategies in human children and great apes for my thesis: imitation and emulation learning. In general, individuals may learn by copying actions, goals, and/or results (Call, Carpenter, & Tomasello, 2005; Horner & Whiten, 2005; McGuigan, Whiten, Flynn, & Horner, 2007; Tennie, Call, & Tomasello, 2006; Tennie, Call, & Tomasello, 2010; Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009). First, imitation learning refers to copying the precise action performed by the demonstrator (Call et al., 2005; Horner & Whiten, 2005). However, this imitation can be rational, that is, only intentional actions are copied while accidental ones are not. For example, children and enculturated chimpanzees may only copy an adult switching on the light with her head when she has her hands free. However, they do not copy her actions when she was not able to use her hands for some clearly visible reason (Buttelmann, Carpenter, Call, & Tomasello, 2007; Gergely, Bekkering, & Király, 2002). When individuals faithfully copy actions that are clearly redundant (and individuals know this), one refers to this phenomenon as “over-imitation” (Lyons, Young, & Keil, 2007; McGuigan, Makinson, & Whiten, 2011; Whiten et al., 2009). Second, emulation learning refers to copying goals, and/or results by one’s own means (McGuigan & Whiten, 2009; Tennie et al., 2006; Tennie, Call, et al., 2010). For example, chimpanzees may observe how a conspecific opens a puzzle box by pushing a door open, whereas they pull it open, i.e., they reproduce the solution by their own means (Tennie et al., 2006).

After introducing the concept of tool use and giving an overview of its evolutionary origin, I turn my eye on one specific problem in the following section. The Floating Peanut Task requires the usage of an unusual tool: water.

B. The Floating Peanut Task (FPT)

The problem that I have referred to at the beginning of the general introduction is the Floating Peanut Task (FPT): This task requires subjects to pour water into a vertical tube to retrieve a buoyant object (Hanus, Mendes, Tennie, & Call, 2011; Mendes, Hanus, & Call, 2007; Nielsen, 2013; Renner, Abramo, Karen Hambright, & Phillips, 2017; Tennie, Call, et al., 2010). It has been proposed being a candidate for an insight problem as discussed in Chapter 1A (Mendes et al., 2007; Shettleworth, 2012).

In the subsequent sections, I first define innovative problem solving, tool innovation, and tool knowledge, and then review studies on three innovation problems in human children and great apes: the FPT, the hook task, and Aesop’s Fable task.

1. Tool innovation

The term “innovation” is used with many meanings, one is invention and adoption by group members (Hochberg, Marquet, Boyd, & Wagner, 2017; Reader & Laland, 2003), another one is innovation as producing a novel solution to a problem (Beck et al., 2011; Beck et al., 2012; Beck et al., 2014; Beck et al., 2016; Chappell et al., 2013; Cutting et al., 2011; Cutting, Apperly, Chappell, & Beck, 2014; Griffin & Guez, 2014; Laumer, Bugnyar, Reber, & Auersperg, 2017; Nielsen, 2013; Nielsen, Tomaselli, Mushin, & Whiten, 2014). Importantly, while problem solving may comprise behavioural innovations, innovations do not necessarily entail problem solving (Seed &

Mayer, 2017). Innovation problems may be ill-structured: Their structure does not provide information about the precise actions required to transform the task from the start to the goal state (i.e., ill-structured problems; Cutting et al., 2011; Jonassen, 1997). In this thesis, I define innovative problem solving as

> *“overcoming some [physical] obstacle to achieve a goal [that involves the creative usage of a tool] when the entire solution is [not] in the species-typical repertoire (…) [and the structure of the task components does not provide information on how to transform them from the start to the goal state].”*
>
> (adapted from Seed & Mayer, 2017)

Innovation problems demand a creative approach to finding a solution, which may involve the use of novel tools, such as bending a hook out of a wire (Bateson, 2014; Beck et al., 2011; Beck et al., 2012; Beck et al., 2014; Beck et al., 2016; Chappell et al., 2013; Cutting et al., 2011; Cutting et al., 2014; Griffin & Guez, 2014; Hennessey & Amabile, 2010; Laumer et al., 2017; Nielsen, 2013; Nielsen et al., 2014). When innovation problems involve tool use, such as manufacturing a novel tool or using an unusual object such as water as tool, I refer to these events as “tool innovation”. Tool innovation seems to develop late during human ontogeny with children becoming proficient by the age of six to eight years (Beck, Apperly, Chappell, Guthrie, & Cutting, 2011; Beck, Chappell, Apperly, & Cutting, 2012; Beck, Williams, Cutting, Apperly, & Chappell, 2016; Chappell, Cutting, Apperly, & Beck, 2013; Cutting, Apperly, & Beck, 2011; Hanus et al., 2011). But why does it develop that late?

2. Function knowledge

One possibility is that tool innovation is partially based on function knowledge. Function knowledge refers to information about the intentions of the tool's designer, the current tool user, or any cultural practice that restricts the tool's function (Greif & Needham, 2011). Function knowledge is also referred to as "tool knowledge" or "conceptual knowledge" (e.g., Manrique et al., 2010). These terms are used interchangeably in this thesis. Tool knowledge develops late during human ontogeny. For example, Defeyter, Avons, and German (2007) asked 5- and 7-year-olds to come up with possible functions for a familiar or novel tool. The older children came up with more possible functions with the novel objects than younger ones. Yet, the younger children came up with more possible functions with the familiar objects than the older ones. This study indicates that 7-year-olds have a broader repertoire of possible tool functions, yet, they may be more restricted by design features than 5-year-olds. This suggests that their function knowledge is already more profound than the knowledge of younger children. German, Truxaw, and Defeyter (2007) emphasize that knowledge about artefacts involves the integration of multiple cognitive domains, such as object mechanics (i.e., perception and action) or reasoning abilities. Thus, function knowledge may require higher order cognition that only develops late in human children.

Object representations

Some authors argue that human object representations are unique among the animal kingdom and only humans represent objects as being *made for* a specific function by an agent (design stance or teleological-intentional stance; Casler & Kelemen, 2007; Csibra & Gergely, 2009; German & Johnson, 2002; Hernik & Csibra, 2009; Ruiz & Santos, 2013; Vaesen, 2012). Two-year-old children already assign functions to objects after single demonstrations and 2- to 4-year-olds show a dissociation of tool functions: Children avoid using a familiar tool for a novel function when a functionally equivalent, but novel tool is available (Casler & Kelemen, 2005, 2007). Additionally, young children already reason about what an object has been made for (Defeyter, Hearing, & German, 2009; German et al., 2007; Hernik & Csibra, 2009).

In contrast, many authors question if great apes have enduring functional representations of objects (e.g., Vaesen, 2012). In fact, wild apes disregard tools after they have used them and they do not carry them around (Boesch & Boesch, 1984; McGrew, 1992), even though they may occasionally re-use rare objects such as hammers and anvils (Boesch & Boesch, 1984; Carvalho, Biro, McGrew, & Matsuzawa, 2009). However, captive great apes have been shown to be able to save tools for future usage. There is also evidence for future planning abilities in apes from other behavioural domains such as planning travel routes and foraging (Boesch & Boesch, 1984; Janmaat, Polansky, Ban, & Boesch, 2014; Mulcahy & Call, 2006; Osvath & Osvath, 2008; van Schaik, Damerius, & Isler, 2013). Still, these findings do not require enduring functional representations of the tools. So apes' notion of *function* may not go beyond the current situation or the situation in mind (i.e., when planning to use a tool): They may perceive, remember, or imagine task affordances (of familiar problems) and select tools due to their functional properties without assigning an enduring function to the object. For example, when wild chimpanzees select tool sets and carry them to tool use sites (Boesch et al., 2009; Brewer & McGrew, 1990; Sanz et al., 2009), this potentially involves memory processes and foresight, but it could be explained by chimpanzees selecting tools by their functional properties only. They may know which functional properties are needed to fish termites or extract honey, but this does not necessarily show that these tools are considered to be exclusively *for* this specific purpose (see Csibra & Gergely, 2009; Hernik & Csibra, 2009; Vaesen, 2012).

In the subsequent section, I review studies on three innovation problems: the FPT, the hook task, and Aesop's Fable task from a developmental and comparative perspective.

3. The developmental and comparative perspective

The FPT has first been employed with great apes, but due to reasons of consistency I report studies with human children first. However, first I discuss some important differences between studies conducted with apes and humans. First, children use a tool to transport the water, which they pour into the tube (Hanus et al., 2011; Nielsen, 2013), while apes use their mouths (Hanus et al., 2011; Mendes et al., 2007; Tennie, Call, et al., 2010). On rare occasions, apes have also been reported to urinate into the tube or to transport water with the palm of their hands (see Chapter 3). One could argue that the task is more difficult for the children than for the apes because they have to employ meta-tool use (i.e., use a tool to use the water). Yet, studies showed that children struggled to come up with the solution strategy, but not to execute it once they had found it (Chapter 4; Hanus et al., 2011). This is also reflected by the observation that children continuously pour water into the tube until they can reach the floating object (Chapter 4). In opposition, some apes stop spitting water into the tube without retrieving the peanut (Hanus et al., 2011; Tennie, Call, et al., 2010). Also, apes generally have less fine-tuned motor skills. Thus, employing a container to transport the water and then, to pour it into the tube would increase the task's difficulty a lot for them while it does not for the children.

Second, children are presented with a visible water source (e.g., a transparent pitcher or bottle; Hanus et al., 2011; Nielsen, 2013), whereas apes are tested with a water dispenser that does not allow visible access to the water (Hanus et al., 2011; Mendes et al., 2007; Tennie, Call, et al., 2010). To reduce differences between methods, I use an open water source (a bucket and a water basin) in my studies with the FPT, both for great apes and human children (Chapter 3 and 4). To conclude, testing each species with an adapted paradigm seems adequate to evaluate the cognitive underpinnings that success in the FPT is based on. Yet, if possible, one could adjust some parameters such as visibility and salience of the tool. In the following, I summarize the studies that have been conducted with the FPT in human children and great apes (as well as one monkey species) so far.

Figure 1.3 The Floating Peanut Task in a human child (A) and an orang-utan (with a transparent and an opaque tube respectively). The pictures relate to studies reported in Chapter 3 and 4.

The FPT – spontaneous success

Human children

Only a few studies have employed the FPT to study problem solving in human children (see Figure 1.3A for an illustration). Hanus et al. (2011) investigated the performance of 4-, 6-, and 8-year-olds with a dry and wet (i.e., quarter-filled) tube. Children were presented with a vertical tube that contained a peanut and that was attached to the table; a transparent water pitcher was located close by. Findings suggest a two-way interaction with age and tube condition. Older children performed better than younger ones, especially in the wet tube condition (4-year-olds – dry: 0%, wet: 17%; 6-year-olds – dry: 33%, wet: 50%; 8-year-olds – dry: 42%, wet: 75%). Interestingly, all children used the pitcher to water plants prior to the test. (If their prior experience with the tool with a different function influenced the likelihood of success is discussed further in Chapter 4.)

Another study conducted by Nielsen (2013) examined the performance of 4-year-olds (and adults) in the FPT. The findings corroborate the ones from Hanus et al. (2011): Children were presented with a tube contained a toy monkey and a water bottle plus a small and a big cup close by. All adults spontaneously solved the task in the baseline, whereas only a minority of children did (14%). Successful children either directly poured water from the bottle into the tube (two children) or employed the two cups as well (three children). The social demonstrations that unsuccessful children received are described below (see *FPT – additional information*).

Great apes

Orang-utans – Mendes et al. (2007) presented five Sumatran orang-utans with the FPT in the wet condition (i.e., the tube was quarter-filled). All individuals solved the task spontaneously in the first session (mean delay to success: 540 seconds). Hanus et al. (2011) further tested ten Bornean orang-utans who received either eight sessions with the wet condition or four sessions with the dry condition and subsequently four sessions with the wet condition. Two of these comprised a quarter-filled tube and two a half-filled tube. None of the Bornean orang-utans solved the task. Two subjects who belonged to the dry-condition group added water to the tube, but not enough to access the peanut (please note that it is not specifically reported in which condition spitting occurred).

Chimpanzees – Hanus et al. (2011) presented 19 chimpanzees with the FPT in the wet condition for eight sessions. None of the chimpanzees solved the task or added water to the tube. The authors then tested 24 chimpanzees from another population. Subjects either received four sessions with the dry condition and four sessions with the wet condition (i.e., two sessions with a quarter-filled tube and two sessions with a half-filled tube) or eight sessions with the wet condition (i.e., a quarter-filled tube). Five chimpanzees solved the task spontaneously, two in the dry condition and three in the wet condition. Four of them solved the task in the first session and the other one did so in the second session (delay to first spit [median]: 232 seconds; range: 5-533 seconds; delay to success [median]: 578 seconds; range: 459-811 seconds). Four additional subjects added water to the tube, but failed to retrieve the peanut. One of them belonged to the dry-condition group and three to the wet-condition group (please note that it is not specifically reported in which condition apes added water to the tube).

Chimpanzees from the first population were tested with a familiar water dispenser, whereas the ones from the second population were given a novel one. Hanus et al. (2011) therefore presented the 19 chimpanzees from the first population with an additional novel water dispenser. Subjects either received two sessions with the dry condition and then, two sessions with the wet condition (i.e., a quarter-filled tube) or four sessions with the wet condition. Two subjects solved the task in the wet condition. Three additional subjects added watered to the tube, one in the wet condition and two in the dry condition. Hanus et al. (2011) conclude that the familiar function of the water dispenser prevented chimpanzees' success in the first place (functional fixedness effect; see Chapter 1C).

Gorillas – Hanus et al. (2011) tested five Western gorillas who received eight sessions with the wet condition (i.e., a quarter-filled tube). None of the gorillas solved the task or added water to the tube.

Monkeys – To my knowledge, Renner et al. (2017) conducted the only study with the FPT in a monkey species. Seven brown capuchin monkeys (*Cebus apella / Sapajus paella*) received four sessions with the dry tube condition and four sessions with the wet tube condition (i.e., the tube was quarter-filled). None of monkeys solved the task or added water to the tube.

The FPT – additional information

As demonstrated, some Sumatran orang-utans and chimpanzees have solved the FPT, whereas there is no evidence in Bornean orang-utans or gorillas so far (Hanus et al., 2011; Mendes et al., 2007). Some further studies have explored the impact of additional information or hints towards the solution in the FPT. One piece of information that may help finding the solution has already been reported in the previous section: In the wet condition, subjects could pick up on the information that first, the object inside the tube is buoyant and second, that the tool required is water. These pieces of information helped children with increasing age, whereas great apes did not benefit from them. In the following, I briefly summarize studies with the FPT that employ additional information or hints at the task's solution such as a social demonstration.

Human children

After a baseline reported in the previous section, Nielsen (2013) presented unsuccessful 4-year-old children with a social demonstration. They observed how a human experimenter poured water directly from the bottle into the tube (bottle condition) or from the bottle first into the small cup (small cup condition) or the large cup (large cup condition) and then into the tube until the tube was eighth-filled. Then, children were given a try. After receiving a demonstration, over half of the children succeeded (61%). Most of them copied the pouring method they had observed before (bottle and small cup condition: 60%; big cup condition: 64%). The findings indicate that children benefit from demonstrations which they copy confidentially, even if a simpler solution was available such as pouring water directly into the tube (for a discussion on overimitation, see Lyons et al., 2007; Nielsen, 2013; Taniguchi & Sanefuji, 2017; Whiten et al., 2009).

Great apes

Chimpanzees – Chimpanzees are the only great ape species that have been studied with regard to additional information towards the task's solution in the FPT. Tennie, Call, et al. (2010) presented 27 chimpanzees with a social demonstration that was either performed by a conspecific or a human demonstrator. While the conspecific used the precise actions required for solving the task (i.e., collecting water with one's mouth and spitting it into the tube), the human reproduced the solution by different means (i.e., collected water with a bottle and poured it into the tube from outside the test room). They received one session with four to six demonstrations and one with two demonstrations. In both sessions, subjects were presented with the original test thereafter. Eight chimpanzees solved the FPT in total, five after a demonstration by a conspecific and three after one by a human. Two additional subjects added water to the tube after a demonstration by a conspecific and three after one by a human, but they did not retrieve the reward. Tennie, Call, et al. (2010) conclude that chimpanzees benefitted from receiving pieces of information about the results and goals (i.e., the end-state), but not further from information about the precise actions required to produce these results. The authors propose emulation learning to explain the results best: Chimpanzees were able to reproduce the result by their own actions after a short exposure.

Monkeys – Renner et al. (2017) presented seven brown capuchin monkeys with two vertical tubes next to each other. Subjects then received four sessions with a full demonstration performed by a human experimenter. The human experimenter poured water from a cup into one of them until the marshmallow inside came into reach. Thereafter, the monkey was given a chance to extract the marshmallow from the second tube. None of the monkeys added water to the tube. One female retrieved the marshmallow in her last session by employing a wad of straw. Yet, this incidence does not count as success in the FPT since the task requires the absence of functional tools other than water.

In the next section, I discuss evidence from two further innovation problems: the hook task and Aesop's Fable task.

The hook task

In the hook task, individuals have to make a hook out of a wire or a pipe cleaner to retrieve a bucket with a reward inside from a vertical transparent tube (Beck, Apperly, Chappell, Guthrie, & Cutting, 2011; Laumer, Bugnyar, Reber, & Auersperg, 2017; Nielsen, Tomaselli, Mushin, & Whiten, 2014; Weir et al., 2002). Interestingly, human children only became proficient in this task by the age of six to eight years (Beck et al., 2011; Beck, Chappell, Apperly, & Cutting, 2012; Beck, Williams, Cutting, Apperly, & Chappell, 2016; Chappell, Cutting, Apperly, & Beck, 2013; Cutting, Apperly, & Beck, 2011). Their poor performance seems quite robust because providing them with experience about the tool materials, or verbally prompting them to try something else did not improve performance (Chappell et al., 2013). However, showing them a ready-made hook, especially if paired with giving them experience with the tool materials (i.e., their flexibility) substantially enhanced the performance in 6- to 7-year-olds, but not that of 4- to 5-year-olds. Moreover, children's success rates also improved after they had received a social demonstration by the experimenter (Chappell et al., 2013; Cutting, Apperly, Chappell, & Beck, 2014).

To my knowledge, no study has been published so far in which great apes are presented with the hook task. Yet, two other non-human animals have been shown to solve the task after receiving prior experience with a target tool: New Caledonian crows (Corvus moneduloides) and Goffin's cockatoos (Cacatua goffiniana) bent a hook out of a wire and a pipe cleaner respectively to retrieve the bucket from the tube (Laumer et al., 2017; Weir et al., 2002). In both cases, the birds had received prior experience with the target tool (i.e., a ready-made hook).

Aesop's Fable task

Aesop's Fable task consists of a vertical tube that contains some water and a floating object which can be retrieved by dropping multiple stones into the tube so that the water level rises (e.g., Cheke, Loissel, & Clayton, 2012; Jelbert, Taylor, Cheke, Clayton, & Gray, 2014; Jelbert, Taylor, & Gray, 2015). Cheke et al. (2012) presented 4- to 10-year-old children with Aesop's Fable task in multiple configurations. Before the test, children received training with the collapsible platform task, in which they also had to drop a stone into a tube (see Figure 1.1A). Thus, children's performance in the test might be dependent on this prior experience. In one condition of the test, they had to choose between two tubes of which one contained water and one contained sawdust. They had to drop marbles in the tube containing water to be able to extract the floating object. In another condition, they had to choose from two types of objects that were either buoyant or not (cork balls versus marbles). The findings from these and further tests indicated that children performed decently by the age of five to seven years in most conditions and nearly perfectly by the age of eight years (Cheke et al., 2012). The authors also report that 4- to 5-year-olds learned slower than older children. Since children had received prior experience with the collapsible platform task, they could transfer the solution to the novel problem. However, it is interesting that they perform well in diverse conditions assessing their causal understanding around the same age in which they solve the FPT and the hook task.

Comparing the FPT, the hook task, and Aesop's Fable task

Although further studies and more direct comparisons between the FPT, the hook task, and Aesop's Fable task are needed, I consider it possible that all three would reveal a comparable age pattern in human children and perhaps comparable success rates among great apes. All three tasks fit to the definition of innovation problems, in which the actions required to transform the problem from its starting to its solution state are unknown (ill-structured problem; Cutting et al., 2011; Jonassen, 2000). They require unusual tool use (i.e., using water as tool, manufacturing a novel tool, or using heavy objects for water displacement). Comparing these tasks more closely in future studies may help to learn more about the underlying mechanisms in tool innovation and the role of prior experience with the problems’ components.

As demonstrated, the FPT and other innovation problems are fascinating tasks to explore human children’s and apes’ problem solving. In the subsequent chapter, I turn my eye to a psychological effect that is more specifically relates to the tool’s identity: the functional fixedness effect.

C. The functional fixedness effect

When Hanus et al. (2011) presented chimpanzees with the FPT and a familiar water dispenser, none of the apes solved the task. The authors subsequently installed an additional, novel water dispenser. As a result, a number of individuals added water to the tube, some even enough to obtain the peanut. Was this behaviour evidence of a functional fixedness effect in great apes? Karl Duncker (1945) coined the term by proposing that

> *"'fixedness' may (…) be conditioned functionally as well as by such factors of visual organization. For instance, a stick that has just been used as a ruler is less likely to appear as a tool for other purposes than it would normally be."*

In other words, when a problem requires the usage of a familiar tool with a novel function, humans – and potentially great apes as suggested by the example – struggle to come up with the solution (e.g., Adamson, 1952; Birch & Rabinowitz, 1951; Defeyter & German, 2003; Dominowski & Dallob, 1995; Duncker, 1945; Flavell, Cooper, & Loiselle, 1958; German & Barrett, 2005; German & Defeyter, 2000; Glucksberg, 1964; Glucksberg & Weisberg, 1966; Maier, 1931; Yonge, 1966). One famous example is Duncker's box problem, which requires using tacks to attach three boxes to a wall, light three candles and fix them inside the boxes (function of box: container versus support; Duncker, 1945; Figure 1.4). The crucial manipulation consists of the location of the additional experimental materials (i.e., the tacks, the candles, and the matches): They are either presented inside the boxes or next to them (Figure 1.4A,B). Human adults are more likely to solve the box problem when no pre-utilization of the boxes is given, that is, the boxes are empty.

Before I give an overview about studies on functional fixedness, I discuss the puzzling negative effect of experience in problem-solving situations more closely: How is it that prior experience can have detrimental effects on one’s problem solving abilities, given that it is usually regarded as beneficial? Also, how does the functional fixedness effect relate to other fixation effects such as mental set? When does it develop in human children and primates?

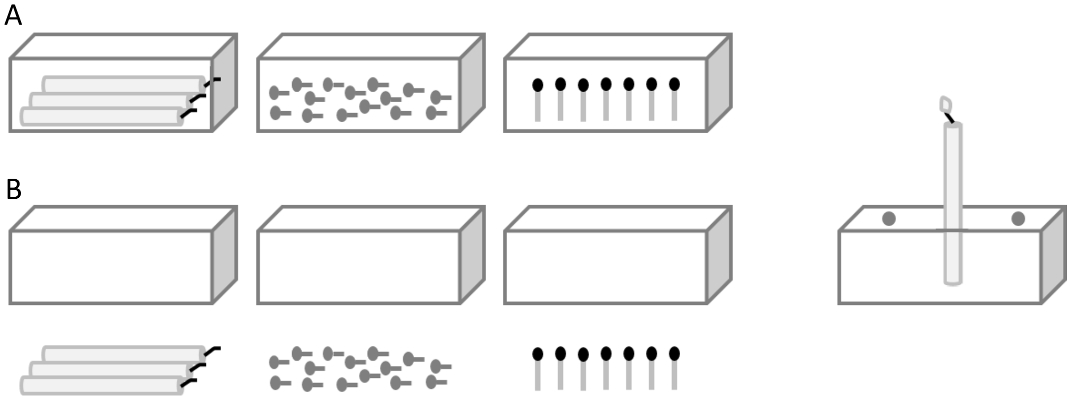


Figure 1.4 Duncker’s box problem: The experimental materials (candles, tacks, and matches) are either presented inside the boxes (A) or next to them (B).

1. The role of prior experience

It is widely accepted that experience can lead to increased problem-solving performance. Encountering the same or similar problems repeatedly allows for employing the same solution strategies (e.g., Ebel & Call, 2018; Jackson, 1942; Koton, 1988; Mendes et al., 2007; Reed, Ernst, & Banerji, 1974; Schultz, 1960; Sheridan & Reingold, 2014). For example, expert chess players solve chess problems faster than novices (Sheridan & Reingold, 2014). The notion of prior experience is defined as knowledge which an individual has gathered during her past. Importantly, this knowledge is available to the individual and could be taken into account when making decisions, although it is not directly perceivable in the current situation (although environmental stimuli can lead to a cued recall; see also Tomasello & Call, 1997 for a discussion of mental representations). The retrieval of past knowledge may be based on memory processed that could be voluntary or involuntary (i.e., cued recall; Lewis, Call, & Berntsen, 2017), a differentiation not important for our purposes.

However, prior experience can also hinder problem solving (for a review, see Chown, 1959; Luchins & Luchins, 1959; Schultz & Searleman, 2002). A novel problem may require abandoning a previously acquired solution strategy. Otherwise it may prevent the discovery of more efficient or effective solutions (Bilalić, McLeod, & Gobet, 2008b; Dominowski & Dallob, 1995; Schultz & Searleman, 1998; Schultz, Stone, & Christie, 1997; Sheridan & Reingold, 2013). Thus, experience can help and hinder the individual at the same time. Or to put it in other words, the individual has to decide when to rely on prior knowledge and when to abandon it to attain the maximal results.

For example, although fixation effects are defined as representing an obstacle to the individual's goals, routines can also make you productive in everyday life (see also Schultz & Searleman, 2002).

Novices in a certain domain show more variable responses than experts, making their problem solving less efficient, but also more flexible than the ones of experts in case of unusual problems (Sheridan & Reingold, 2014). While experts may be fixated on their familiar solution strategies, the more expert they are, the less they are fixated, suggesting an inverted U-shaped curve of fixation across expertise (Bilalić, McLeod, & Gobet, 2008a; Bilalić et al., 2008b; Sheridan & Reingold, 2013, 2014). In the subsequent section, I give an overview about fixation effects and their relation to functional fixedness.

2. Defining functional fixedness

The blocking effect of the familiar solution has been referred to as Einstellung effect, (mental) set, fixation, or, more generally, as behavioural rigidity (for a review, see Chown, 1959; Luchins & Luchins, 1959; Schultz & Searleman, 2002). Behavioural ”rigidity is the [general] tendency of an individual *not* to change”, as Schultz and Searleman (2002) put it. The concept has not only been referred to mental sets, but also to extreme attitude, stereotypy, few flexibility, perseveration, authoritarianism, or difficulties to change habits (Schultz & Searleman, 2002). Rigidity involves the acquisition and perseveration of set; and set refers to a mental or behavioural pattern, which is established via experience and that enables the individual to make predictions about the future (Schultz & Searleman, 2002). Einstellung and set may be used interchangeably (Luchins & Luchins, 1959; Schultz & Searleman, 2002). The functional fixedness effect could be considered to be a subdomain of the Einstellung effect (Duncker, 1945; see Figure 1.5): It involves the learning of a tool function and its perseveration, which ultimately prevents success in problem-solving tasks or makes problem solving much less efficient.

Functional fixedness has been tested with various paradigms in human adults. For example, the function of the tool is either induced by usage or by presentation (Duncker, 1945). An example of function by presentation (also referred to as “pre-utilization condition”) is the box problem mentioned above (Defeyter & German, 2003; Duncker, 1945; German & Barrett, 2005). Some studies also employ a different method to induce (or re-fresh) a tool’s function: Subjects actively use a familiar or novel tool by which a function is assigned to the tool or a previously known function is re-activated (e.g., when a pair of pliars is used to pull out nails and then, has to be used as support as in the "pliers problem"; Duncker, 1945). Duncker (1945) has found no difference between these two paradigms in human adults (i.e., function by presentation versus function by usage). Another interesting aspect concerns the type of experience or *how* the information is learned. Individuals can either learn about the tool’s function by their own actions or by observing others (e.g., Sommerville, Hildebrand, & Crane, 2008). Finally, individuals may generalize from familiar tools to novel tools, which are similar in their appearance (Casler & Kelemen, 2005, 2007). They may subsequently exhibit a functional fixedness effect with these novel tools, too (see Figure 1.5). In the following section, I review studies on functional fixedness in human children and great apes.

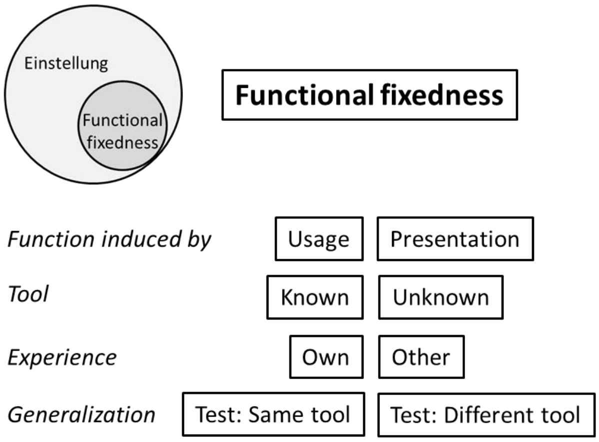


Figure 1.5 Experimental procedures to study the functional fixedness effect.

3. The developmental and comparative perspective

Growing research deals with the question of how functional fixedness develops in human children (e.g., Clegg & Legare, 2016; Defeyter & German, 2003; German & Defeyter, 2000). Yet, the majority of research on tool functions has focused on perseveration with tools in babies and young children (e.g., Elsner & Schellhas, 2012; Smitsman & Cox, 2008). Few studies have been conducted with great apes on this issue (e.g., Gruber, Muller, Reynolds, Wrangham, & Zuberbuhler, 2011; Gruber, Muller, Strimling, Wrangham, & Zuberbuehler, 2009; Marshall-Pescini & Whiten, 2008). In this section, I review studies on functional fixedness in human children and great apes.

Human children

Developmental studies suggest that the functional fixedness effect emerges around the age of six years in human children (Defeyter & German, 2003; German & Defeyter, 2000). For example, German and Defeyter (2000) presented 5-, 6-, and 7-year-old children with a problem that required selecting and stapling a box and several building blocks. The crucial manipulation was that the box served as container for the additional experimental materials for half of the children (pre-utilization condition), whereas the other half was presented with all materials lying next to each other (no pre-utilization condition). Findings suggested that 6- and 7- year-olds took longer in the pre-utilization condition to solve the task, whereas the performance of 5-year-olds was not affected by the manipulation.

Some studies suggest that children younger than six years may exhibit a functional fixedness effect because they show perseveration behaviour with tools that have a familiar function (Elsner & Schellhas, 2012; Smitsman & Cox, 2008). For example, 2- to 4-year-olds either learned stabbing polystyrene chips through the top opening of an apparatus or pushing them out of the box via lower holes with a rod (Elsner & Schellhas, 2012). When the previously used holes were blocked and the respective other holes were opened, children spent more time with the previously successful strategy (i.e., stabbing or pushing) which was now less efficient than the respective other strategy (i.e., pushing or stabbing) than naïve children during the first acquisition phase (Elsner & Schellhas, 2012). This study shows that prior experience influenced young children’s tool use; yet, it is unclear if children were fixated on the tool’s function or the task’s solution since both the tool and the apparatus were the same in both phases (the latter would be a more general fixation effect, also referred to as "(mental) set", or "rigidity"; Chown, 1959; Duncker, 1945; Luchins & Luchins, 1959; Schultz & Searleman, 2002; Sternberg & Davidson, 1995). However, if children exhibit a functional fixedness effect earlier with a different task design than those previously used will be a matter of future studies. Remarkably, a recent study tested adolescent children and young adults and found that functional fixedness is also apparent in a culture that incorporates a limited number of artefacts in peoples’ everyday lives (German & Barrett, 2005).

Great apes

Behavioural rigidity has been observed in wild as well as captive great apes: The apes stayed with their first solution strategy even if a more efficient one became available (Gruber et al., 2011; Gruber et al., 2009; Harrison & Whiten, 2018; Hrubesch, Preuschoft, & van Schaik, 2009; Marshall-Pescini & Whiten, 2008; Price, Lambeth, Schapiro, & Whiten, 2009). In this section, I give an overview on fixation effects in wild and captive great apes.

Two groups of wild chimpanzees were tested with a honey trap task and a tool with two functional ends (i.e., a twig with a stick-end and a leaf-end; Gruber et al., 2011). One of the groups used sticks regularly in their everyday lives, whereas the other one did not. The authors found that chimpanzees of the stick-using population used the stick-end of the tool to extract honey from the log, whereas the population without regular stick tool use either tried to access the honey directly with their fingers or produced a sponge out of the leaf-end of the tool. Even when the stick-end was pre-inserted into a hole containing honey, apes from the latter population did not use it to extract the honey. The authors conclude that both populations found different parts of the tool salient (Gruber et al., 2011; see also Gruber et al., 2009) which potentially resulted in a fixation of one functionally aspect of the tool (for a potential dissociation of tool end functions see also Sanz et al., 2009).

In Hrubesch et al. (2009), captive chimpanzees learned to retrieve food rewards from a platform using a stick. Some individuals discovered that the platform could be shaken so that the food items could be retrieved more easily and rapidly. Yet, many of the subjects continued using the stick technique to retrieve the food although the more efficient solution strategy was clearly visible. In another study, chimpanzees were presented with an apparatus with two feeding options (Marshall-Pescini & Whiten, 2008). First, they learned from a demonstrator to dip honey from the top area of an apparatus. Then, they were shown a more efficient strategy that required them to insert the same stick at the frontal area of the apparatus. Instead of learning the novel, more efficient technique, chimpanzees stayed with the one, which they had learned first. Another study showed that chimpanzees were more likely to solve the FPT when they were presented with a novel water dispenser (Hanus et al., 2011). Hanus et al. (2011) suggested that chimpanzees may have been fixated on the familiar function of the water dispenser (i.e., being for drinking and spitting at conspecifics and people). These studies are consistent with the existence of a functional fixation in the tool use of great apes. However, to my knowledge, there has been no systematic investigation of this effect so far. The studies presented investigate if apes stay with a familiar solution strategy or if they switch to a more efficient one when using a tool. Yet, they do not force the apes to use the tools with a novel function to be still able to solve the task (i.e., block the old solution) while controlling for novelty of the tool (e.g., in case of the study with the FPT; Hanus et al., 2011). Thus, we lack studies investigating the functional fixedness effect with experimental designs comparable to the ones used with humans so far.

Studies on functional fixedness with apes are important for our understanding of the evolution of human object representations. First, functional fixedness may imply enduring functional representations of tools because otherwise there would be no fixation effect. Second, overcoming this fixation seems to involve a capacity to have two different representations of one object between which the individual can switch. Thus, studies on functional fixedness in great apes and human children may deepen our knowledge about the development and evolution of object representations.

D. Aim of thesis

In this thesis, I investigate the role of prior experience in the tool use of great apes and human children. First, I use the Floating Peanut Task (FPT) to study insightful problem solving; more precisely, the role of visual feedback and additional information about task solution in great apes (Chapter 2). There has been an ongoing debate over decades whether non-human animals engage in insightful problem solving. Removing perceptual feedback allows investigating if apes anticipate the outcomes of their actions in the FPT. To explore their task understanding further, they receive several hints towards the task solution (involving the end state and/or the actions required to solve the task).

Second, I use the FPT to investigate the functional fixedness effect and how it is influenced by learning, namely, from own experience or observing another individual, the amount of experience, and task presentation in 6- to 8-year-old human children (Chapter 3). Third, I explore the functional fixedness effect in great apes from three different angles (Chapter 4): The first experiment involves a tool with two subparts offering different functionality (i.e., a brush with a brush-like and a pointed end); the second one a tool with two different functions that are not separated in space (i.e., a hose); and the third one a food item that can be used as a tool (i.e., a bread stick). Whereas there is some evidence for behavioural rigidity in apes, only few studies investigate the functional fixedness effect, so that this thesis helps to assess if great apes are vulnerable to this effect as humans are.

In the general discussion of the thesis (Chapter 5) I draw conclusions about the double-edged sword of experience. It is characterised by improvements from tool-use experience, but also a decline in performance due to the functional fixedness effect. From an ontogenetic and phylogenetic perspective, I discuss the mechanisms that may be involved. I conclude that there are probably two types of functional fixedness of which one is based on perception–action learning and one on conceptual tool knowledge. Moreover, I discuss functional fixedness designs as a more general methodology to investigate various aspects of tool knowledge.

Finally, I propose a theoretical account for why tool innovation arises late during human ontogeny and phylogeny, and I define two steps on the way to become a flexible tool user. These are related to the two types of the functional fixedness effect. Firstly, individuals have to learn about perception–action contingencies from object manipulations, and later from repeated tool use. Individuals master simple tool use tasks, but perseverate on previous tool actions and do not show great flexibility with tools. Secondly, individuals have to combine this procedural knowledge with conceptual tool knowledge that finally allows them using the same tool in different contexts and other tools as a replacement of the familiar one. I conclude that this integration of the two systems is crucial for tool innovation and that it may be only partially present in great apes, whereas it fully develops in humans over time.

Chapter 2: General methods

Summary

In this chapter, I provide an overview about the general methods of the studies that I report throughout the thesis. First, I give details about the great apes and human children who participated in the experiments. Second, I describe the general testing procedure for both groups of participants. Third, I explain the procedure of data collection and analyses.

A. Participants

In this thesis, I report on ten experiments conducted with non-human great apes and human children. All studies have been approved by an ethics committee of the University of St Andrews. The studies were purely behavioural and did not involve any invasive methods. All applicable international, national, and institutional guidelines for behavioural research with non-human animals were followed. In the subsequent sections, I give an overview of the great apes and human children who participated in the experiments.

1. Great apes

The studies with great apes have been conducted with three populations of great apes at three different holding facilities. Two populations were housed at German zoos and the third one at a Kenyan sanctuary. I briefly present each holding facility in the following.

Wolfgang Köhler Primate Research Center / Leipzig Zoo

The majority of experiments were conducted at the Wolfgang Köhler Primate Research Center (WKPRC) which is maintained by the Max Planck Institute for Evolutionary Anthropology (MPI-EVA) in cooperation with the Leipzig Zoo (Leipzig, Germany). The WKPRC houses four species of non-human great apes: bonobos (*Pan* *paniscus*), chimpanzees (*Pan troglodytes*), Western gorillas (*Gorilla gorilla*), and Sumatran orang-utans (*Pongo abelii*). Fifty great apes of various ages were living at the WKPRC in 2017. I report apes' sex, age, and rearing history in the relevant chapters.

The great apes lived in social groups of various sizes with access to indoor and outdoor enclosures during the daytime. The chimpanzees lived in two separate groups. The enclosures comprised various enrichment devices such as trees and ropes to climb on, shaking boxes and poking bins. Additionally, the apes were provided with enrichment materials daily. They spent the night in their indoor sleeping rooms and were released into the enclosures in the morning. The apes received their usual diet throughout the study period. Water was available ad libitum during testing. Any food that apes gained during the experiments was an additional reward to their usual diet.

Apes participated on studies on cognition at WKPRC on a daily basis, including studies on tool use. For each of the experiments, I assessed their prior knowledge with certain tools (e.g., a brush tool or hose) by looking at the study archive and asking the caregivers of the apes. This was an important consideration because prior experience could influence my experiments.

Sweetwaters Chimpanzee Sanctuary

One of the experiments was conducted at the Sweetwaters Chimpanzee Sanctuary located in the Ol Pejeta Conservancy (Laikipia, Kenya). Thirty-nine chimpanzees of various ages lived in the sanctuary in 2015, when the studies were conducted. The majority of chimpanzees were orphans, were born in the wild and came to the sanctuary after being confiscated from the illegal bushmeat and pet trade. They were all raised by humans in a highly comparable way, living together with peers after arriving at the sanctuary.

Chimpanzees lived in two social groups in extensive outside enclosures that comprised bushland, trees, and open areas. Apes stayed in their indoor sleeping rooms during the night. They were fed multiple times per day and tests were conducted in the indoor sleeping rooms on a voluntary basis. Chimpanzees participate on studies on cognition somewhat regularly, including some studies on tool use.

Dortmund Zoo

I conducted one experiment with two Sumatran orang-utans (*Pongo abelii*) at the Dortmund Zoo (Dortmund, Germany). The subjects lived in a social group and had access to indoor and outdoor areas comprising multiple enrichment devices. All tests were conducted on a voluntary basis in the sleeping room. The orang-utans received their usual diet during the period of the study and water was available ad libitum. Both orang-utans used to live at WKPRC where apes participate on cognitive studies on a daily basis, including studies on tool use.

Sampling

I tested all apes available at the time of data collection so that most of the experiments comprise several species. Only gorillas were excluded from most of my studies because of the small sample size of three individuals (WKPRC), which does not allow for controlling the factor “species” in the statistical analyses. Most of the experiments involved within- and between-subjects designs. Within- and between-subject-variables were counterbalanced as well as possible. For example, in Chapter 3 in Experiment 2, subjects received three conditions of which a third of the subjects started with each of the conditions. Likewise, in Chapter 5 in Experiment 1 and 2, subjects were distributed randomly into two groups after matching pairs as close as possible with regard to species, sex, and age.

2. Human children

Children were recruited from a database of children in kindergartens in Leipzig, a mid-sized German city. Some of them had already participated in studies on cognitive development. The socioeconomic background of children was considered diverse, but I did not collect any data on children’s socioeconomic status. The children’s parents gave their informed consent before the children participated in the study.

I tested children from various kindergartens in each of the experiments and went to each kindergarten only once. The experiments required that children did not tell each other the solution to the problem. They involved a between-subject design with multiple variables. All of the conditions were tested in each kindergarten if possible to control for differences between kindergartens. In the following, I describe the general testing procedures for apes and children.

B. General testing procedures

Testing great apes and human children followed certain general testing procedures of which I give an overview in the following, starting with the apes. I further report the specific methods of each experiment in the relevant chapters

1. Great apes

The testing of the apes took place in their indoor sleeping rooms. Some of the tests at WKPRC were also conducted in the so called “observation rooms”. Observation rooms were comparable to apes’ sleeping rooms with regard to their configuration and comfort (see Figure 2.1 for two examples). In the following, all rooms in which testing took place will be referred to as “testing rooms”. Testing took place in the morning and apes entered the testing room from their indoor enclosures. Dependent on the study, the apparatus was either already installed when apes entered the testing room or it was prepared when the subjects were waiting in an adjacent room. Apes were released back into their natural group after the test. Testing was stopped if apes exhibited any signs of stress, such as displays, screaming, or scratching. Infants were sometimes tested with the hydraulic door half-open so that they could go back to their mothers at any time during the test. The amount of trials per session was adjusted to the amount of food that apes were allowed to consume per test to prevent over-feeding. Apes were regularly fed in the morning and any food gained in the experiments was additional. They also had access to water throughout the study.

The caregivers of the apes were responsible for calling them inside, separating them, and moving them from one room to the other. They also decided if it was possible to test a specific individual on a given day. For example, sometimes apes were sick or changes in the social dynamics of the group required them to stay with the group to keep the group stable and/or calm. Apes were separated from the group for a fixed amount of time and returned to their group thereafter. In some cases, another individual stayed in an adjacent room during the test. For example, mothers stayed inside when infants were tested. Also, some apes were more relaxed when a conspecific was close to them with whom they shared an affiliative relationship. In these cases, I tested both individuals, one after the other. When a second individual was inside, I made sure that they could not observe the experimental setup from their position. This was an important consideration since they also participated in the test as subjects. I report the precise testing procedures in the relevant chapters.

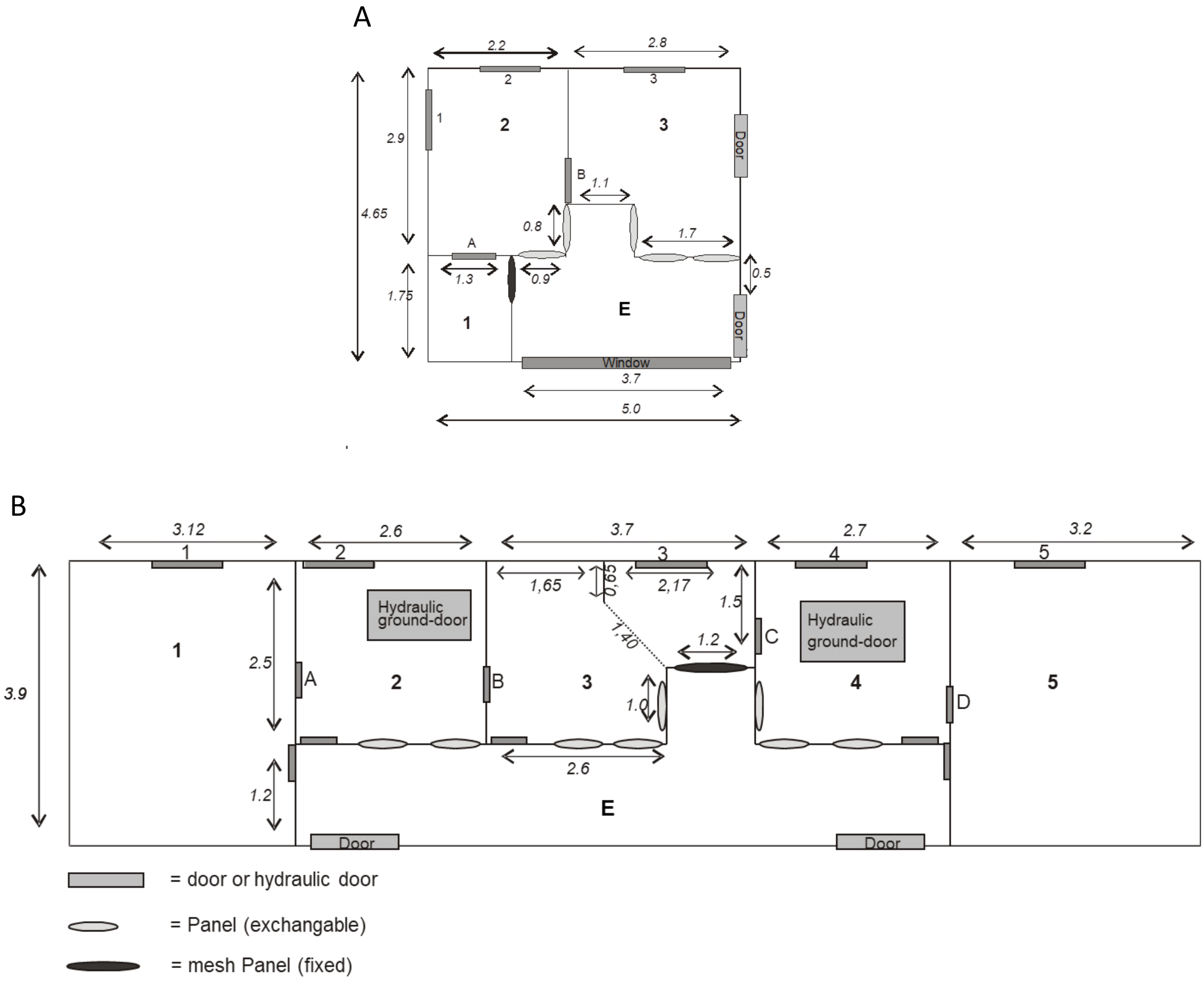


Figure 2.1 Two examples of testing rooms: The observation (A) and sleeping rooms (B) of the larger group of chimpanzees at WKPRC. The numbers indicate separate rooms (sleeping room: 1-4, observation room: 1-3) and the letter “E” the experimenter area. All measurements are provided in meters. (The figure is reproduced with permission from MPI-EVA.)

2. Human children

Children were tested in quiet rooms of the kindergartens. I set up the apparatus and then collected one to four children at a time from the same class. After approval of the kindergarten teachers, children were asked if they would like to play a game with me. If they refused, I did not try to convince them to participate. If they expressed any concerns throughout the experiment or seemed anxious, I asked them if they were feeling well and offered them to bring them back to their classes. Since the experiments investigated individual problem-solving skills, I aimed at reducing the likelihood of communication about the solution between participants. Thus, a second experimenter played a game with the small group of children in front of the test room. One child participated in the study at a time. All children received a reward (i.e., three stickers) regardless of their performance. After the last child had finished the test, the children were brought back to their class-room. Additional to the stickers, they received a postcard for their parents that contained information on details of the study.

C. Data collection and analyses

Great apes and human children were video-taped during the experiments. All videos were stored on file servers belonging to the MPI-EVA. The data was either live-coded during the experiments or coded from videos. I used the program Solomon Coder for behavioural coding (Péter, 2011). A second coder, usually a student assistant or intern, coded 20% of the data in each of the experiments. Reliability between both coders was assessed by performing a Pearson's correlation or calculating Cohen's Kappa. The data was analysed in R and the alpha level for the analyses was set to 0.05 (R Core Team, 2013). If possible, I performed Generalized Linear Mixed Models (GLMMs) to analyse the data; GLMMs are based on a fixed and random effects structure, which controls for the fact that individuals contributed to more than one data point (e.g., Barr, Levy, Scheepers, & Tily, 2013). Thus, GLMMs provide a powerful tool to analyse data in comparative research with multiple measurements of the same individuals. I report the specific statistical analyses in the respective chapters.

Chapter 3: Great apes – Visual feedback in the FPT[2]

Summary

In this chapter, I report three experiments with great apes that test the role of visual feedback and its timing in the Floating Peanut Task (FPT). The FPT requires apes to collect water from a water dispenser or basin and spit it into a vertical tube to retrieve a peanut. Previous studies have established that some individuals solve the FPT with a transparent tube. However, these studies left open the question if apes anticipated the outcome of their actions and acted upon an insight into the task's solution, or if they were spitting into the tube for some reason and then, they were reinforced by the peanut moving closer to the opening of the tube. This question can be tackled by confronting apes with an opaque tube that deprives them of any visual feedback for their actions.

In Experiment 1, apes were presented with an opaque tube without any hints given at the solution in the baseline condition. Then, they received an end-state demonstration (i.e., they encountered a water-filled tube with a peanut floating atop) and thereafter, a bottle demonstration (i.e., they observed how a human experimenter solved the task by pouring water from a bottle into the tube) to investigate the impact of additional information on task performance. Unfortunately, the conditions could

---

[2] Material from Chapter 3 formed the basis of the following paper (under review): Ebel, S. J.; Schmelz, M.; Herrmann, E.; Call, J.: Innovative problem solving in great apes is fostered by visual feedback and the end-state of the solution.

not be counterbalanced for order across subjects because this experiment was part of a test battery requiring the same order for all subjects.

In Experiment 2, I further tested for the effect of additional information in the FPT. Apes were presented with a transparent tube in the baseline. Then, they received an end-state demonstration (i.e., they encountered a water-filled tube with a peanut floating atop) and two types of water tap demonstrations (i.e., a water tap located above the tube was turned on by the ape or the human experimenter). The three additional information conditions were counterbalanced for order across subjects.

In Experiment 3, I investigated if performance in the FPT is influenced by the timing with which visual feedback is blocked. More precisely, while subjects in Experiment 1 first encountered an opaque tube, subjects in Experiment 3 had already solved the transparent tube and were subsequently given an opaque one. The sample comprised successful subjects from Experiment 2 and from previous studies. First, subjects were given the transparent tube as a reminder and then, they encountered the opaque one. I further assessed the role of visual feedback on successful apes' behaviour. More precisely, subjects who were successful with the opaque tube were presented with two additional tubes that could not be solved. One tube had a hole at its front and one at its back so that water was flowing out. The two tubes were presented counterbalanced for order across subjects. I assessed if apes changed their spitting behaviour dependent on the availability of visual feedback about apparatus malfunctioning.

The findings of the three experiments indicate that apes who had identified the solution with a transparent tube also solved an opaque version (Experiment 3). However, apes starting with the opaque one failed to solve the task (Experiment 1). An end-state demonstration promoted success independent on the availability of visual feedback (Experiment 1 and 2). Experiencing how water was poured into the tube either by a human demonstrator or by a water tap that had been opened either by the ape or a human did not seem to be of further assistance (Experiment 1 and 2). The findings suggest that great apes require visual feedback for solving the FPT, which is no longer required after initial acquisition. Moreover, some subjects benefit from encountering the end-state of the solution, a finding that corroborates previous studies.

A. Introduction

The FPT has been used by various authors to study different aspects of problem solving in human children and great apes (Hanus et al., 2011; Mendes et al., 2007; Nielsen, 2013; Tennie, Call, et al., 2010). In the FPT, individuals have to pour water into a vertical transparent tube which contains a peanut (or any other buoyant object) that floats upwards until it comes into reach. Recent studies left unclear whether apes solving this task really anticipated the effect that spitting water into the tube would have on the peanut’s position or if they had added water to the tube for some other reason and upon seeing its positive effects repeated the action until they managed to extract the peanut from the tube. Mendes et al. (2007) argued that apes‘ behaviour in the FPT may be regarded as insightful since the apes suddenly started adding water to the tube and continued their goal-directed behaviour until they had reached the peanut (see also Shettleworth, 2012). The authors suggested using an opaque tube to address this question. This is an important consideration because if subjects were to succeed in such a task, it would rule out that visual feedback is essential and would suggest that individuals can anticipate the effect of pouring water on the position (and therefore accessibility) of the peanut.

Visual feedback plays an important role in problem solving (e.g., Köhler, 1925; Taylor, Elliffe, Hunt, & Gray, 2010; Völter & Call, 2012) and refers to any visual stimuli that serve as positive or negative feedback for an individual’s actions. This feedback helps to assess if the actions are likely to obtain the desired goal, for example, when a chimpanzee is raking in a food reward with a stick, she can assess her progress by observing the food coming closer. Visual feedback generated by an individual's own actions can facilitate or impede the appearance of an efficient solution to a problem. For example, pushing an object away from the subject to overcome a barrier that is preventing its direct retrieval is difficult because subjects cannot resist bringing the object closer and consequently, after pushing it away, they repeatedly bring it back to the starting position (e.g., Guillaume & Meyerson, 1930; Köhler, 1925).

The timing of feedback in relation to the solution, and not just its nature, is also important. For example, Taylor, Elliffe, et al. (2010) presented New Caledonian crows with a vertical string pulling task in which they could pull up a string to which a piece of food was attached. All crows succeeded when they had full visual access to the string. Thereafter, the crows also solved a visually restricted version of the task in which a platform limited visual access. However, blocking visual feedback before first acquisition substantially hindered the solution and only one crow succeeded spontaneously (Taylor, Elliffe, et al., 2010). Völter and Call (2012) presented non-human great apes with an analogous task in which apes could crank up a piece of food that was attached to a string inside an either transparent or opaque apparatus (Völter & Call, 2012). Some subjects spontaneously solved the task with the transparent version, but all subjects failed with the opaque one. However, after apes had acquired the solution with the transparent apparatus, they transferred it to the opaque one (Völter & Call, 2012). Both studies suggest that while feedback was required for acquisition of the solution, it was no longer needed to maintain performance.

An important question is if the impact of visual feedback would also be modulated by task difficulty. When Völter and Call (2012) presented apes with two less complex problems (i.e., pushing out a food item from a horizontal tube or removing sticks from a tower to release a food reward), the visual feedback was not necessary to succeed. However, apes were faster when visual feedback was available than when it was not, suggesting that visual feedback supports more efficient problem solving (Völter & Call, 2012). These results highlight that the effect of visual feedback is modulated by task complexity and its timing (i.e., before or after the solution).

Another finding regarding the effect of visual feedback comes from a study in which apes witnessed parts of the solution in the FPT. Visual feedback may not only be present when an individual receives information about the consequences of her own actions, but also when additional pieces of information about task affordances are presented such as an end-state demonstration (i.e., showing the task in its solution state without indicating the actions required to reach it) or a full demonstration (i.e., showing an agent performing all actions required to reach the solution state). For instance, apes who witnessed water being poured into the tube so that the peanut floated up performed better than subjects who did not receive such information (Tennie, Call, et al., 2010). Interestingly, Tennie, Call, et al. (2010) found no evidence that observing another ape solving the task was more useful than just observing the changes that occurred to the peanut's location when a human experimenter poured water into the tube from a bottle (38% vs. 21%). Tennie, Call, et al. (2010) concluded that emulation learning explains the finding, that is, chimpanzees reproduced the end-state using their own actions, since there was no difference between a demonstration showing the precise actions required and a bottle demonstration. Interestingly, chimpanzees did not benefit more from encountering a partial solution, namely a tube that already contained some water compared to a dry one (Hanus et al., 2011).

Völter and Call (2012) have shown that great apes' problem-solving performance depended on visual feedback and task difficulty. When simple manipulations led to success in a task, visual feedback was not required. Yet, when an action had to be performed repeatedly to solve the task apes needed visual access to the effect of their actions to succeed. In the current chapter, I investigated great apes' problem-solving performance with another task which required a creative idea to solve it: the FPT. Firstly, the mechanism of the FPT could perhaps be considered more intuitive to apes than the one of the crank task because apes interacted with water during their everyday lives and they had seen objects floating on water before. However, they probably had not used a crank before and thus, did not understand its mechanism. Thus, it is perhaps easier for them to anticipate the outcome of their actions in the opaque version of the FPT than in the opaque version of the crank task. Secondly, adding water to the tube in the FPT is not a straightforward manipulation as manipulations with one's hands and mouth as in the crank task and the two other tasks in Völter and Call (2012) that could theoretically lead to a success by chance. Thus, the opaque version of the FPT also allowed me to investigate if apes solve the task via insight, i.e., after getting stuck, suddenly coming up with the solution strategy and acting on it in a goal-directed manner (Mendes et al., 2007).

In Experiment 1, I therefore presented naïve chimpanzees with an opaque version of the FPT that prevented them from receiving visual feedback about the effect that spitting water into the tube had on the peanut's position. After a baseline with a dry tube, chimpanzees received some hints about how to solve the task. I hypothesized that a minority of the individuals would be able to solve the task with an opaque tube, which would indicate that they potentially solved it via insightful learning. Yet, success rates in the FPT had been low with the transparent tube in previous studies (e.g., Hanus et al., 2011) so that I complemented the baseline with two additional conditions in which pieces of information were given on how to solve the task. In the end-state condition, apes encountered a water-filled tube with the peanut floating atop and in the human demonstration condition, they watched how the experimenter poured water into the tube until the peanut emerged at the opening. I hypothesized that some individuals would solve the task after receiving the end-state condition because a previous study showed emulation learning in the FPT with a transparent tube (Tennie et al., 2010). Yet, I did not expect that chimpanzees who did not solve the task in the end-state condition would solve it in the human demonstration condition. If chimpanzees were emulating the results (or goals) in the FPT (Tennie et al., 2010), receiving pieces of information about the actions required to establish the end-state should not be of further help to them. However, if some chimpanzees solved the task after receiving the human demonstration (and maybe none after the end-state demonstration), this would indicate that actions resembled a valuable source of information for apes in the FPT. Since water was added to the tube in both conditions (chimp and human demonstration) in Tennie et al. (2010) and an end-state condition was not directly compared to a demonstration condition, this would be an interesting finding. In the current experiment, I measured success and latency to success to assess apes' problem-solving performance. Additionally, I measured the number of spits and the mean inter-spit-interval as behavioural indicators to better understand apes effort in this task (number of spits) and goal-directedness (inter-spit-interval). These measurements also allow the reader to compare the current study to previous ones (Hanus et al., 2011; Mendes et al., 2007; Tennie et al., 2010). I also measured the same behaviours in the following two experiments.

As the classical FPT is typically solved by a minority of apes only, it seemed likely that depriving subjects of any visual feedback may make the task extremely difficult. I therefore conducted two additional experiments to complement the first one. The second experiment aimed at firstly, increase the sample size of subjects who solved the FPT with a transparent tube for Experiment 3, secondly, validate the findings from Experiment 1 and Tennie et al. (2010) about the impact of emulation learning in the FPT, and thirdly, test bonobos for the first time with the FPT. In Experiment 2, I therefore presented bonobos and chimpanzees with a baseline with a transparent tube followed by three conditions that provided apes with additional information about task affordances (end-state demonstration, water tap turned on by the ape, water tap turned on by the experimenter). I hypothesized that apes would solve the task in the baseline with a somewhat smaller success rate than in previous studies because some of them had been unsuccessful with the FPT before. Moreover, I expected that some individuals would solve the task after receiving additional information, although the impact of the different conditions was exploratory (i.e., if causing the action of water pouring into the tube was more informative on the task's solution than observing a demonstrator causing the action). If results showed instead a distinct pattern in relation to the three conditions, this would indicate that apes payed attention to the end-state, actions and potentially who caused them (dependent on the precise results).

Finally, I aimed at investigating if experienced apes would solve the FPT with an opaque tube. In Experiment 3, I therefore presented apes who had previously solved the transparent tube with an opaque tube to find out whether disrupting visual feedback after acquisition affected their performance. Therefore, I presented subjects with a transparent tube first and then with an opaque one. I hypothesized that apes would be able to solve the opaque version of the FPT because experienced apes were also able to solve the opaque version of the crank task in Völter and Call (2012). If they would not solve the opaque version of the FPT, this would indicate that they were dependent on visual feedback in the FPT despite their experience. Furthermore, I investigated on an exploratory basis whether successful subjects confronted with an ineffective opaque tube differentially perseverated in pouring water depending on whether the cause of failure was visible or not. In the visual cause condition, subjects could see that the water escaped through a hole at the front of the tube while in the no visual cause condition, they did not receive such feedback as the water escaped at the back of the tube. I hypothesized that apes would notice the cause of their failure in the visible cause condition and would stop earlier with spitting water into the tube than in the no visual cause condition. If they did not, this would suggest that they do not pay attention to visual feedback in the FPT that indicates failure and perhaps that they do not understand the FPT's mechanism fully.

B. Experiment 1

1. Methods

Subjects

Twenty-four chimpanzees living at Sweetwaters Chimpanzee Sanctuary (Ol Pejeta Conservancy, Kenya) participated in the study ($N_{females}$ = 14; age range: 8-28 years; mean: 18 years; Table 3.1).

Materials

I used an opaque Plexiglas tube (26 cm x 5 cm) that was closed at both ends. A hole (about 3 cm x 3.5 cm) was drilled on its top front (Figure 3.1). The size of the tube and the position of the hole were such that they blocked visual access to a peanut located at the bottom of the tube. In fact, the peanut became visible only as it neared the hole. I attached a water dispenser to a grey PVC plate about a meter away from the tube.

Table 3.1 Subjects participating in Experiment 1.

Subject	**Sex**	**Age**
Jane	Female	11
Joy	Female	11
Victoria	Female	11
Saidia	Female	12
Eva	Female	13
Julia	Female	15
Mwanzo	Female	19
Bahati	Female	22
Tess	Female	23
Chipie	Female	24
Dufatanya	Female	24
Amizero	Female	26
Akela	Female	27
Cheetah	Female	28
Roy	Male	8
Ali Kaka	Male	13
Cumbo	Male	15
Edvard	Male	15
William	Male	15
Zee	Male	15
Amihirwe	Male	17
Niyonkuru	Male	25
Uruhara	Male	26
Ndaronse	Male	27

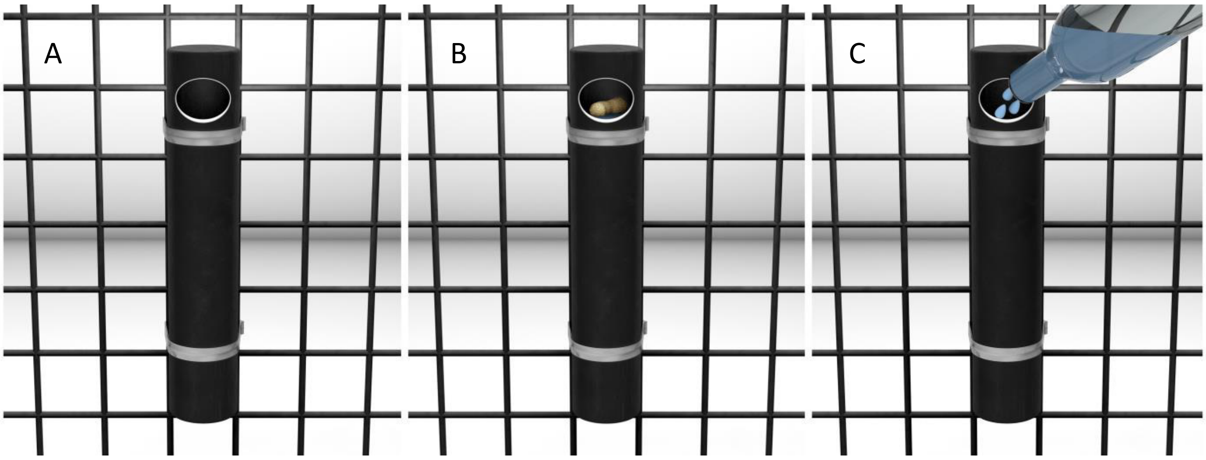


Figure 3.1 Setup of Experiment 1: Baseline condition (A), end-state condition (B) and human demonstration condition (C).

Procedure

Chimpanzees received a maximum of six sessions with one session per day. After they had solved the task once, they did not receive any further sessions. First, they received two sessions with the baseline condition followed by two sessions with the end-state and two with the human demonstration condition (Figure 3.1). Since the study was part of a larger study which required the same order for all individuals, I did not counterbalance the order of conditions across individuals. Besides, I did not expect a great improvement in the human demonstration condition as a previous study suggested that chimpanzees mainly benefit from encountering the end-state (Tennie, Call, et al., 2010). I employed a human demonstration using a bottle (i.e., showing how to reach the outcome by slightly different methods) instead of one performed by a conspecific (i.e., showing the exact actions needed) because a recent study had shown that there was no difference between these two types of demonstrations (Tennie, Call, et al., 2010).

In the baseline condition, chimpanzees watched from an adjacent room how the experimenter dropped a peanut into the tube. After entering, they had ten minutes to retrieve the peanut. In the end-state condition, they encountered the water-filled tube with the peanut floating atop which allowed them to retrieve the peanut from the water. After they had taken the peanut from the tube, the door to the adjacent room was opened and the subject left the room so that the experimenter could prepare the test trial. In the human demonstration condition, apes witnessed from an adjacent room how the experimenter filled a bottle (maximum capacity: 500 ml) with water from the water dispenser and poured it into the tube. The experimenter repeated these actions three times until the tube was filled and the peanut came into reach (duration: about 60-90 seconds). Then, the chimpanzee entered the room and could retrieve the peanut from the tube. After the chimpanzees had obtained the peanut in the end-state and the human demonstration conditions, they waited in an adjacent room until the experimenter had exchanged the wet tube for a dry one. Then, the sessions continued as in the baseline condition and subjects had ten minutes to solve the task. Subjects were tested individually in all conditions. The experimenter only performed actions (dropping the peanut into the tube in the baseline or pouring water into the tube in the human demonstration) when subjects were sitting at the mesh with their heads facing the tube. When subjects moved away, the experimenter called them and demonstrations continued as soon as they returned to their position.

Coding and analyses

I coded success (i.e., retrieval of the peanut), latency to success, latency to first spit, number of spits, and the mean inter-spit-interval using Solomon Coder (Péter, 2011). When subjects spat several times with one mouthful of water this was still counted as one spit.

2. Results

None of the chimpanzees acquired the solution or added water to the tube during the baseline. After receiving an end-state demonstration, one chimpanzee solved the task without ever having seen someone adding water to the tube before. Thus, she solved the task without attaining visual feedback for her actions. Remarkably, she continuously added water to the tube without pausing once (Jane, 11 years – session 4; 7 spits, first spit: 500 seconds, success: 589 seconds, mean inter-spit-interval: 11 seconds). Two additional females added water to the tube after an end-state demonstration, but not enough to obtain the peanut (Cheetah, 28 years – session 3; 2 spits, first spit: 263 seconds, inter-spit-interval: 25 seconds; session 4; 1 spit, first spit: 556 seconds; Julia, 15 years – session 4; first spit: 550 seconds). After her first and second spit in session 3, Cheetah found a vegetable stalk and inserted it into the tube repeatedly for about a minute. Thereafter, she quit the task. It is possible that the stick-like object distracted her from adding more water to the tube. Interestingly, all spitting behaviour except for one event (Cheetah, session 3) occurred very late in the session, after more than eight minutes. None of the chimpanzees added water to the tube in the human demonstration condition.

3. Discussion

One chimpanzee solved the FPT when visual feedback was blocked after receiving information about the end-state of the task, i.e., encountering a water-filled tube with a peanut floating atop. Thus, at least one individual solved the task without receiving any immediate visual feedback for her spitting actions and without ever having seen water being poured into the tube before. This perseveration constitutes the first evidence, albeit weak, that an individual may have anticipated the consequences of her actions in the FPT. No other subjects showed this behaviour and two chimpanzees added water to the tube once or twice and quit, perhaps because they obtained no feedback.

The two additional chimpanzees added water to the tube after receiving information about the end-state, but failed to obtain the peanut. However, none of the chimpanzees acquired the solution during the baseline, suggesting that apes require visual feedback to solve the FPT spontaneously since at least some apes solved the task spontaneously in previous studies (Hanus et al., 2011; Mendes et al., 2007). Moreover, chimpanzees were not further benefitting from a human demonstration, that is, none of the subjects added water to the tube in this condition. Thus, these findings are consistent with the idea that end-state information facilitated spitting behaviour in some individuals in the FPT when visual feedback was blocked. Although emulation and imitation learning were not directly compared in the current study, findings corroborate the idea that emulation learning was enough to explain chimpanzees' success in a previous study (see also Call et al., 2005; Tennie, Call, et al., 2010).

The conditions differed in regard to their reinforcement structure. While subjects' actions were reinforced during the end-state and the human demonstration condition when they retrieved the peanut from the tube, they were not reinforced during the baseline. Such reinforcement might have led to a higher motivation to engage with the task thus providing a better explanation than emulation learning for the spitting behaviour of the three individuals. However, a recent study showed a positive effect of social demonstrations in the FPT, even though apes were not reinforced during these demonstrations since a dominant conspecific was in the same room with the subject and always took the peanut (Tennie, Call, et al., 2010). Additionally, chimpanzees readily manipulated the tube and the water dispenser during baseline sessions, something that indicates that the chimpanzees were motivated to engage with the task. It is still possible that observing a conspecific accessing the peanut might have a similar effect as eating the peanut oneself, an issue that requires further investigation. Note, however, that reinforcement may also play a role in emulation learning in natural settings where food leftovers, which can be considered in some cases "end-states", may act as reinforcers. Future studies could directly compare the relationship between end-state and social demonstrations on the one hand and the reinforcement that subjects experience themselves or observe in others on the other hand.

One reason why the pouring demonstrations in the current and previous studies might not have been more effective is that the change in the position of the peanut was not caused by the subjects' own actions. In other words, if the subject had caused the change and not been a mere observer of both the cause and the effect, this would have been more effective in producing a solution. Therefore, in the next experiment I devised a task in which the subject caused the change during the demonstration phase, but using different means of what she would be required to do during the test phase (i.e., pour water from her mouth into the tube). More specifically, I presented chimpanzees and bonobos with the transparent version of the FPT followed by three conditions: an end-state demonstration condition, a condition in which they themselves activated a tap that filled the tube with water and brought the peanut within reach and a condition in which the experimenter activated the tap.

C. Experiment 2

1. Methods

Subjects

Six bonobos and 18 chimpanzees housed at WKPRC participated in the study ($N_{females}$ = 19; age range: 6-48 years; mean age: 23 years; Table 3.2). While all six bonobos and five chimpanzees were naïve to the FPT, 13 chimpanzees had been presented with this task in previous studies, but had failed to solve it (Hanus et al., 2011; Tennie, Call, et al., 2010; Table 3.2). Two chimpanzees (Annett, Swela) were excluded from the study because one of them refused to eat wet peanuts and one did not pull the string to activate the water tap (N_{final} = 22).

Materials

I used a dry transparent Plexiglas tube (26 cm x 5 cm) that was closed at the bottom with a peanut inside (Figure 3.2). For the bonobos, I could not use peanuts due to a peanut allergy of one individual. Thus, I assessed bonobos' preference for dried pieces of apple and banana and used their preferred food item for the test. A black steel container (50 cm x 20 cm x 30 cm) filled with water was attached to the mesh (same as in Allritz, Tennie, & Call, 2013). I used an open water source instead of a water dispenser to potentially increase the likelihood of success due to the salience of the water. The distance of the water to the tube varied across the ape groups due to the conditions of the respective sleeping rooms (bonobos, chimpanzees group 1: about 2.25 m, chimpanzees group 2: about 1.40 m). In the “water tap by ape” and “water tap by human” condition, a water tap was attached above the tube (see Figure 3.2C,D; Appendix B). The tap could be opened by pulling a string so that water would flow into the tube. By pulling the string, a metal ring at the tap was moved towards the ape which produced a banging sound when the movement stopped. The string-pulling was actually non-functional and the water was turned on by the experimenter operating a valve out of sight.

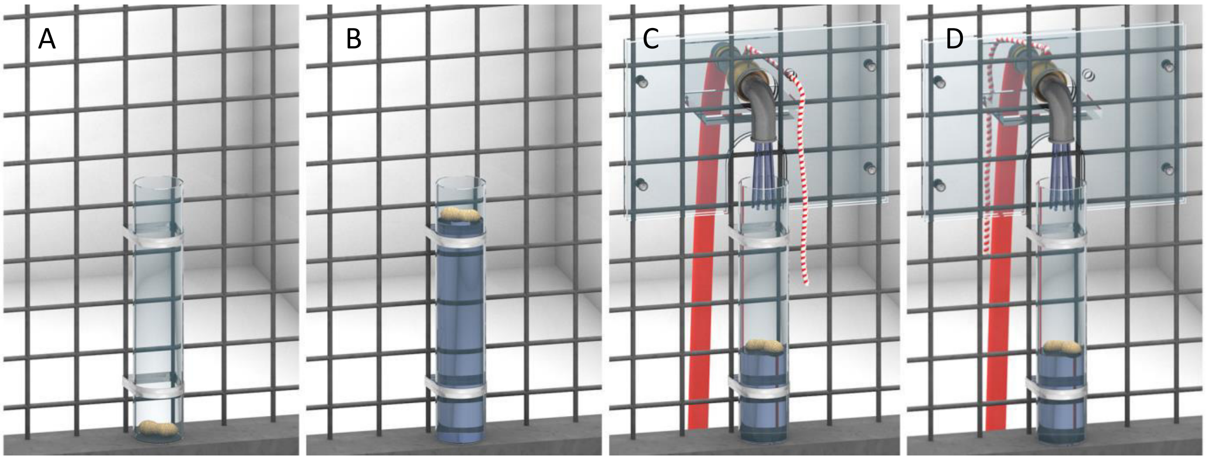


Figure 3.2 Setup of Experiment 2: baseline condition (A), end-state condition (B), water tap by ape condition (C) and water tap by human condition (D).

Table 3.2 Subjects participating in Experiment 2. Some apes had participated in previous studies using the FPT.

Subject	Species	Sex	Age	Rearing	Participated in previous studies	Spat water
Fimi	Bonobo	Female	6	Mother	No	No
Gemena	Bonobo	Female	9	Mother	No	No
Luiza	Bonobo	Female	9	Mother	No	No
Lexi	Bonobo	Female	15	Nursery	No	No
Yasa	Bonobo	Female	17	Mother	No	No
Kuno	Bonobo	Male	18	Nursery	No	No
Kara	Chimp	Female	9	Mother	No (with her mother, see Fraukje)	No[2]
Alexandra	Chimp	Female	15	Nursery	Hanus et al. 2011, Exp. 1 & Exp. 3	No
Annett[3]	Chimp	Female	15	Nursery	Hanus et al. 2011, Exp. 1 & Exp. 3 / Tennie et al. 2010	Maybe[1]
Swela[3]	Chimp	Female	19	Mother	Hanus et al. 2011, Exp. 3	Maybe[1]
Sandra	Chimp	Female	21	Mother	Hanus et al. 2011, Exp. 1 & Exp. 3 / Tennie et al. 2010	Maybe[1]
Jahaga	Chimp	Female	21	Mother	Hanus et al. 2011, Exp. 1 & Exp. 3	Yes
Daza	Chimp	Female	28	Unknown	No	No
Natascha	Chimp	Female	34	Nursery	Hanus et al. 2011, Exp. 1	No
Riet	Chimp	Female	36	Nursery	Hanus et al. 2011, Exp. 1 & Exp. 3	No
Corrie	Chimp	Female	37	Nursery	Hanus et al. 2011, Exp. 1 & Exp. 3	No
Ulla	Chimp	Female	37	Nursery	Hanus et al. 2011, Exp. 3 / Tennie et al. 2010	Yes
Fraukje	Chimp	Female	38	Nursery	Hanus et al. 2011, Exp. 1 & Exp. 3 / Tennie et al. 2010	Maybe[1]
Frederike	Chimp	Female	40	Unknown	No	No

Jeudi	Chimp	Female	48	Unknown	No	No
Kofi	Chimp	Male	9	Mother	No (with his mother, see Ulla)	No[2]
Lobo	Chimp	Male	10	Mother	Hanus et al. 2011, Exp. 3	No
Alex	Chimp	Male	13	Nursery	Hanus et al. 2011, Exp. 1 & Exp. 3 / Tennie et al. 2010	Maybe[1]
Robert	chimp	Male	38	Nursery	Hanus et al. 2011, Exp. 1	No

---

[1] Maybe in Tennie et al. 2010 or Hanus et al. 2011 (not reported)

[2] Kofi has seen his mother adding water to the tube once. Kara has experienced a chimp demonstration, but probably no spitting by her mother. Both were two to three years old at that time.

[3] Excluded from the study

Procedure

Apes received a maximum of eight sessions with one session per day. When apes solved the task once, they were not given further sessions. First, they received two sessions with the baseline condition followed by six sessions in which they received additional information (end-state, water tap by ape and water tap by human), counterbalanced for order across individuals. Each of the additional information conditions was given on two consecutive sessions. In the baseline condition, apes were presented with a dry tube containing a peanut (or a piece of dried apple or banana in case of the bonobos) for ten minutes.

In the end-state condition, apes encountered a tube filled with water and the peanut floating atop (Figure 3.2B). In the water tap by ape condition, apes faced a tube with a peanut located inside. They could pull a string which moved a metal ring to “turn on” the water, while the experimenter actually operated a valve (Figure 3.2C). When the tube was filled with water, they could retrieve the peanut from the tube. In the water tap by human condition, the water was turned on by the experimenter by moving the metal ring herself in view of the ape (Figure 3.2D). After the ape had obtained the peanut the experimenter “switched off” the water tap by pushing the metal ring back into its original position.

Each session in the additional information phase consisted of three demonstrations, followed by the original test with a dry tube. Between the demonstrations the experimenter emptied the tube and placed a new peanut inside. After the demonstrations the water tap was removed and the wet tube was exchanged for a dry one. While the setup was prepared, subjects waited in an adjacent room.

Coding and analyses

I coded success (i.e., retrieval of the peanut), latency to success, latency to first spit, number of spits and mean inter-spit-interval using Solomon Coder (Péter, 2011). As before, in case subjects spat several times with one mouthful of water this was still counted as one spit. A second coder coded all videos with spitting behaviour from Experiment 2 and 3 and reliability was excellent (Pearson’s correlation coefficient: latency to success, $r = 0.996$, $df = 20$, $p < 0.001$; latency to first spit, $r = 0.993$, $df = 50$, $p < 0.001$; number of spits, $r = 0.995$, $df = 50$, $p < 0.001$).

2. Results

Two chimpanzees solved the FPT spontaneously in the baseline (Kofi – session 1; 6 spits, first spit: 40 seconds, success: 577 seconds, mean inter-spit-interval: 102 seconds; Sandra – session 2, 9 spits, first spit: 454 seconds, success: 562 seconds, mean inter-spit-interval: 12 seconds). Three additional chimpanzees acquired the solution in the additional information phase, with two of them succeeding in the first session. More specifically, one subject solved the task after an end-state demonstration (Lobo – session 3; 12 spits, first spit: 150 seconds, success: 380 seconds, mean inter-spit-interval: 21 seconds) and one after activating the water tap herself (Alexandra – session 3; 7 spits, first spit: 101 seconds, success: 152 seconds, mean inter-spit-interval: 7 seconds). Moreover, one chimpanzee solved the task after an end-state demonstration (Kara – session 5; 5 spits, first spit: 91 seconds, success: 250 seconds, mean inter-spit-interval: 39 seconds), after she had already passed two unsuccessful sessions with the water-tap by human condition. One additional chimpanzee and one bonobo added water to the tube, but not enough to obtain the peanut (all unsuccessful spitting behaviour is summarized in Table 3.3). One subject employed two additional techniques to add water to the tube next to spitting: Kofi urinated into the tube and used his hand to transport the water when solving the task.

Table 3.3 Unsuccessful spitting behaviour in Experiment 2.

Subject	**Species**	**Age**	**Session**	**Condition**	**Number of spits**
Fimi	Bonobo	6	1, 5	Baseline, water tap by ape	1 (+4)[1], 2
Kara	Chimp	9	2, 3, 4	Baseline, water tap by human (2x)	1, 1, 2
Lobo	Chimp	10	2	Baseline	1
Riet	Chimp	36	1	Baseline	1

[1] The last four spits occurred after ten minutes had past (in the end the tube was quarter-filled).

3. Discussion

Five chimpanzees solved the FPT by repeatedly spitting water into the transparent tube until they could reach the peanut. Two of them did so spontaneously during the baseline while the other three solved the task after receiving additional information about the solution that always comprised the end-state (i.e., the peanut floating on a water-filled tube). Another chimpanzee and one bonobo added water to the tube, but not enough to extract the peanut. This finding corroborates the results from Experiment 1 with an opaque tube and the ones by Tennie, Call, et al. (2010) with a transparent tube which showed that chimpanzees benefited from encountering the end-state in the FPT (see also Vale, Davis, Lambeth, Schapiro, & Whiten, 2017 for the positive effect of model demonstrations on the occurrence of solutions). I did not find evidence for a difference between conditions in which the ape or the human controlled the water tap to fill the tube, but the low success rate make the interpretation of this result difficult. Future research could investigate whether the end-state produced by the individual's own action versus the action of someone else affect the likelihood of learning by emulation.

While the findings showed low success rates in the FPT with a transparent tube, a word of caution in interpreting these results is necessary. Some of the subjects had participated in previous studies using the FPT, but had failed the task while previously successful individuals who might have had a greater potential to solve the task were not included (see Table 3.2; Hanus et al., 2011; Tennie, Call, et al., 2010). Furthermore, it is difficult to compare the baseline to the conditions in the additional information phase given that their order of presentation was not counterbalanced across subjects. However, this was not the goal here. A previous study had already established that additional baseline sessions did not improve performance (Hanus et al., 2011). More specifically, solutions typically occurred in the first or second baseline sessions or not at all (Hanus et al., 2011). The fact that three additional individuals apparently benefited from end-state conditions is therefore entirely consistent with previous studies.

As far as I know, I tested bonobos for the first time with the FPT. Although none of the six individuals solved the task, one subject added water to the tube in two sessions, but not enough to obtain the dried piece of fruit. However, why no bonobo in comparison to chimpanzees solved the task remains an open question. Future studies should investigate factors like the difference in food reward, or differences in persistence and food motivation across species.

The findings further support the idea that certain forms of visual feedback facilitated the solution in the FPT. As a next step, I assessed whether apes would solve the FPT with an opaque tube after having solved the transparent version before. In Experiment 1, one ape potentially anticipated the effect of the water on the peanut's position, but all other subjects provided no evidence. Here I investigated whether apes would continue solving the task (by repeatedly pouring water in the tube) despite not being able to see the peanut moving upwards. This is equivalent to the manipulation by Taylor, Elliffe, et al. (2010) in the string pulling task and Völter and Call (2012) in the crank task.

To do so, I presented chimpanzees and orang-utans that had already acquired the solution in the FPT with a transparent tube in Experiment 2 or in previous studies (Mendes et al., 2007; Tennie, Call, et al., 2010) with the opaque tube that I used in Experiment 1. In a final manipulation I confronted successful apes with the opaque tube with a hole drilled near the bottom so that any water that was poured into the tube escaped via this hole, thus, preventing the peanut from moving upwards. In the "visual cause" condition, the water escaped from a hole at the front, thus providing information about the cause for the peanut's lack of upward movement. In the "no visual cause" condition, the water escaped from a hole at the back of the tube, out of sight of the apes. I examined if apes would change their behaviour (i.e., stop adding water to the tube) because of the visual feedback that they received and that contained information about the tube's malfunctioning.

D. Experiment 3

1. Methods

Subjects

Eight chimpanzees and five Sumatran orang-utans participated in the study ($N_{females}$ = 9; age range: 9-25 years; mean age: 15 years; Table 3.4). Eleven of them were housed at WKPRC and two at Dortmund Zoo to which they had been transferred since the first study (Mendes et al., 2007). The subjects had previously solved the FPT with a transparent tube except for one orang-utan (Tao) who had witnessed her mother solving the task as an infant. I tested this subject to see if she remembered observing her mother solving the task (transparent tube) and to increase my sample size (opaque tube). The time delay from first success to the re-test and the number of successful trials they had completed differed for the individuals.

While the five successful chimpanzees from Experiment 2 had solved the FPT only once about one month ago (mean: 29 days, minimum: 24 days, maximum: 33 days), the other subjects solved the task more than once, but several years ago. More precisely, Dokana, Padana, Pini (with her 2-year-old daughter Raja), and Toba (with her 6-month-old daughter Tao) were tested nine years prior to the study (Mendes et al., 2007). Padana and Dokana were re-tested five years ago and they plus Pini and Raja were re-tested again about two years ago. Dokana additionally solved the FPT multiple times for diverse TV documentaries, but also did so for the last time about two years ago. Frodo was trained to function as a demonstrator about seven years ago (Tennie, Call, et al., 2010) and Lome and Tai also solved the task about seven years ago (Hanus et al., 2011). Tai additionally participated at least in one TV documentary one-and-a-half years ago.

Table 3.4 Subjects of Experiment 3.

Subject	Species	Sex	Age	Rearing	Facility	First condition
Kara	Chimp	Female	9	Mother	Leipzig	No visual cause
Tai	Chimp	Female	12	Mother	Leipzig	Visual cause
Alexandra	Chimp	Female	15	Nursery	Leipzig	Visual cause
Sandra	Chimp	Female	21	Mother	Leipzig	-
Kofi	Chimp	Male	9	Mother	Leipzig	Visual cause
Lobo	Chimp	Male	10	Mother	Leipzig	No visual cause
Lome	Chimp	Male	13	Mother	Leipzig	-
Frodo	Chimp	Male	20	Mother	Leipzig	-
Tao	Orang	Female	9	Mother	Dortmund	-
Raja	Orang	Female	10	Mother	Leipzig	-
Padana	Orang	Female	16	Mother	Leipzig	Visual cause
Toba	Orang	Female	20	Mother	Dortmund	No visual cause
Dokana	Orang	Female	25	Mother	Leipzig	-

Materials

The same transparent tube was used as in Experiment 2. Additionally, I employed a modified version of the opaque tube from Experiment 1 that included the following changes (see Figure 3.3): The tube was glued to a Plexiglas plate and a hose was attached to its back. The hose was connected to a valve that could be switched on and off. It was closed throughout the opaque condition to prevent the water from escaping the tube (Figure 3.3A). The mesh surrounding the tube was covered so that apes could not see behind the tube. Apes were tested with the water source that was used when they first acquired the solution. Thus, I either used the black steel container from Experiment 2, a novel water dispenser or the familiar water dispenser. The distance of the water to the tube varied across the ape groups due to the different water sources and conditions of the sleeping rooms (chimpanzees: about 2.25 m or 1.40 m, orang-utans: about 2.25 m + 2 m in height). Due to experimenter error, I tested one chimpanzee (Frodo) with the familiar water dispenser although this individual had been tested with a novel one before.

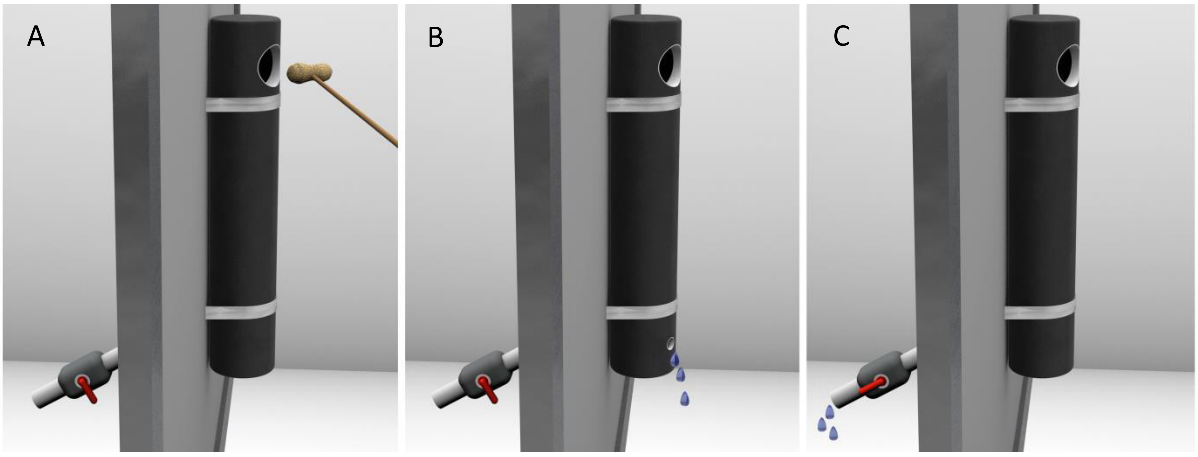


Figure 3.3 Setup of Experiment 3: Opaque condition (A), visual cause condition (B) and no visual cause condition (C).

Procedure

Apes received a maximum of four sessions with one session per day. Each session lasted a maximum of ten minutes. First, I presented apes with the transparent tube followed by the opaque tube, each for two sessions. As soon as they solved the task in a given condition, they did not get a second session with that same condition. Only apes who solved the transparent tube received the opaque tube. There were about 24 hours between success with the transparent tube and presentation of the opaque tube. I therefore administered one additional session with the transparent tube with one orang-utan (Padana) to ensure the same timing between conditions (data of first successful session in results section).

Another orang-utan (Toba) broke off the bottom of the transparent tube in her first session because the rectangular plate at the bottom of the tube protruded due to circumstances of the mesh and I repeated the session on the next day. The caregivers had reported beforehand that the orang-utans would probably know about the water dispensers in their sleeping rooms, although they had rarely seen them drinking from these dispensers. Especially with Tao they were unsure if she was actually aware of this drinking option. Thus, in her first session, I provided Tao with a bowl of water which she spilled after a while so that I do not know if she had a chance to solve the task. Before Toba broke off the bottom of the tube in her first session, she spat saliva several times into the tube so that I was also not sure if she was aware of the water dispenser. To give them a fair chance, I provided both with a large bucket of water in the second session that I placed on the floor of the test room.

Three orang-utans (Dokana, Padana, Raja) received two additional sessions with the opaque tube prior to the transparent tube, resulting in a maximum of six sessions for these subjects (i.e., opaque – transparent – opaque). I did so because these individuals had already been re-tested with the transparent tube in a recent study in which they all solved the task (unpublished data; for reasons of consistency I should have treated Tai in the same way because she had been re-tested for a TV documentary, but I was not aware of this at the time of conducting the study). Thus, they were directly confronted with the opaque tube. I additionally gave them a reminder with the transparent tube thereafter because the delay from last success with the transparent tube to encountering the opaque tube was about 24 hours for the apes, while it was much longer for these three individuals.

In the transparent condition, upon entering the room the ape encountered the peanut located inside a dry and transparent tube (Figure 3.2A). In the opaque condition, the experimenter placed a peanut near the opening of the tube with the aid of a stick, while a caregiver attracted the ape to one side of the room (see Figure 3.3A). When the ape started moving towards the tube, the caregiver gave a signal and the experimenter dropped the peanut into the tube and retracted the stick.

Upon completion of the main phase of the experiment, apes who had been successful with the opaque tube received a follow-up test composed of two conditions presented in separate sessions with the order of presentation counterbalanced across individuals. In the "visual cause" condition, the opaque tube had a hole located at its bottom front so that any water poured into the tube escaped through it (Figure 3.3B). In the "no visual cause" condition, the hole was located at the back of the tube hidden from the subject’s view (Figure 3.3C). Both conditions were impossible to solve. Sessions lasted ten minutes each.

Coding and analyses

I coded success (i.e., retrieval of the peanut), latency to success, latency to first spit, number of spits, latency to first spit and the mean inter-spit-interval using Solomon Coder (Péter, 2011). As before, in case subjects spat several times with one mouthful of water this was still scored as one spit. I compared conditions for successful subjects performing Exact Wilcoxon signed rank tests in R (Hothorn & Hornik, 2015; R Core Team, 2013).

2. Results

Transparent tube

Ten out of 13 apes solved the FPT with a transparent tube in the re-test (77%), i.e., they added enough water to obtain the peanut (except Frodo, Lome, Raja; Table 3.5). More specifically, seven individuals solved the task in the first session and three (Kara, Tao, Toba) in the second one. Two of these individuals (Toba, Tao) had been transferred to a new holding facility and it could be that they were not aware of the water dispenser during the first session. Additionally, Toba broke off the bottom of the tube in the first session after she had repeatedly spat saliva into the dry tube. When both individuals were provided with a water bucket in the second session, they solved the task.

Three subjects remained unsuccessful in the re-test with the transparent tube, yet, two of them added water to it. Raja’s failure was caused by wood wool that she stuffed into the tube. Although she subsequently filled it with water to the top, the peanut got stuck by the wood wool and therefore, she failed to retrieve it. On the second session, she quit after three spits, thus, failing the task. Frodo added five mouthful of water in his second session, but failed to obtain the peanut while Lome did not add any water to the tube.

Opaque tube

Seven out of ten apes solved the opaque tube (70%), i.e., they added enough water to the tube to obtain the peanut without receiving visual feedback for their actions (all except Dokana, Sandra, Tao; Table 3.5). Six of them did so in the first session and one (Alexandra) in the second one. All unsuccessful subjects except one (Tao) added water to the tube eventually (Sandra, session 2, 6 spits; Dokana, session 1, 1 spit). All three orang-utans (Raja, Dokana, Padana) that received two additional sessions with the opaque tube *before* they encountered the transparent tube (and then the opaque tube again, see methods) added water to the opaque tube in the first two sessions, but none of them solved the task.

There was no difference between the transparent and the opaque condition with regard to latency to success (Wilcoxon test: T = 22, $p = 0.219$, $N = 7$; increase from transparent to opaque in 5/7;transparent: 137±75 seconds; opaque: 215±76 seconds; mean±sd.), latency to first spit (Wilcoxon test: T = 20, $p = 0.\ 343$, $N = 7$; increase from transparent to opaque in 5/7; transparent: 57±69 seconds; opaque: 71±40 seconds), mean inter-spit-interval (Wilcoxon test: T = 18, $p = 0.156$, $N = 6$, one tie; increase from transparent to opaque in 4/7; transparent: 18±11 seconds; opaque: 57±51 seconds) or number of spits ($N = 5$, two ties; decrease from transparent to opaque in 5/7; transparent: 5±2 spits; opaque: 4±2 spits).

Table 3.5 Results of Experiment 1-3.

Experiment 1 – Tube: Opaque | Subjects: Naïve				
	Baseline[2]	End-state[2]	Human dem.[2]	
Success	0/24	1/24	0/23	
Spitting[1]	0/24	3/24	0/23	
Experiment 2 – Tube: Transparent | Subjects: Naïve or previously unsuccessful				
	Baseline[2]	End-state[3]	Water tap by ape[3]	Water tap by human[3]
Success	2/24	2[4]/21	1[5]/20	0/20
Spitting[1]	6/24	2/21	2/20	1/20
Experiment 3 – Tube: Transparent and opaque | Subjects: Previously successful				
	Transparent[2]	Opaque[2]		
Success	10/13	7/10		
Spitting[1]	12/13	9/10		

[1] Overall spitting behaviour (incl. successful subjects).

[2] Fixed order of the conditions.

[3] Counterbalanced order of the conditions. (Sample sizes differ between conditions because successful subjects did not receive any further sessions.)

[4] First and second condition that the successful subjects received respectively.

[5] First condition that the successful subject received.

Follow-up test

There were significant differences in spitting frequency between the baseline and experimental conditions (Friedman test: $\chi^2 = 9.0$, *df* = 2, *p* = 0.011, *N* = 7). Subjects spat significantly less often in the baseline compared to the front and back conditions

(Wilcoxon test: T = 21, *p* = 0.031, *N* = 6 in both cases), but there were no significant differences between the front and back conditions (Wilcoxon test: T = 22, *p* = 0.234, *N* = 7; baseline: 4±2, front: 17±13, back: 9±6; number of spits, mean±sd). However, subjects spat more often in the first compared to the second experimental condition that they received (Wilcoxon test: T = 26.5, *p* = 0.047, *N* = 7; first: 19±12; second: 7±4). There were also significant differences between conditions with regard to the mean spitting frequency (Friedman test: χ^2 = 6.0, *df* = 2, *p* = 0.050, *N* = 7). However, pairwise comparisons failed to confirm the differences between conditions (Wilcoxon tests: T = 24, *p* = 0.109, *N* = 7 in all cases; baseline: 57±51, front: 55±42, back: 104±90; inter-spit-interval in seconds, mean±sd.). Similarly, there was no significant difference in the latency to spit between the first and the second condition that subjects received (Wilcoxon test: T = 21, *p* = 0.300, *N* = 7; first: 62±76, second: 98±69).

3. Discussion

Most of the chimpanzees and orang-utans that had first acquired the solution in the FPT ranging from one month up to nine years ago solved the task again in Experiment 3. Most of the successful subjects also transferred the solution to an opaque tube that deprived them of any visual feedback, i.e., they could not perceive the effect that their spitting actions had on the peanut's position. These results suggest that apes became independent of visual feedback after first acquisition of the solution in the FPT.

Apes solved the opaque tube although it required repeated actions over the course of about two-and-a-half minutes (from first spit to the retrieval of the peanut) without being able to assess if their manipulation was successful. While most apes showed this high level of persistence, three apes failed to transfer the solution to the opaque tube. Since two of them still added water to the tube, visual feedback might have been essential for them to solve the task. These results, taken together with those of Experiment 1, are consistent with the findings of two recent studies employing a cranking task and a vertical string pulling task with great apes and New Caledonian crows respectively (Taylor, Elliffe, et al., 2010; Völter & Call, 2012). In these studies, some individuals acquired the solution when visual feedback was available and also transferred it to an apparatus that restricted or completely blocked visual feedback. However, none of them (except for one crow) acquired the solution when visual feedback was restricted or blocked, like in Experiment 1 (Taylor, Elliffe, et al., 2010; Völter & Call, 2012).

Apes solved the FPT although they had not faced the task for a period that ranged from one month to nine years. Although this may be an indication of good memory performance, it may also be a sign of problem-solving consistency. That is, those individuals who solved it originally, also solved it (independently) a few years later and without necessarily recalling that solution. Without comparing the initial latencies to solve the task with the latencies in the current study it is unclear whether their success represents a case of good memory or re-innovation. Although this would have been a desirable comparison, I was unable to carry it out because only few successful subjects had also participated in previous studies and had solved the task spontaneously. A recent study with a tool-use task in chimpanzees in which the authors were able to carry out such a comparison yielded positive results (Vale et al., 2016) and also studies with other paradigms showed impressive long-term memory performance in great apes (Janmaat, Ban, & Boesch, 2013; Lewis et al., 2017; Martin-Ordas, Atance, & Call, 2014; Martin-Ordas, Berntsen, & Call, 2013).

When subjects faced failure in the follow-up test, they perseverated in adding water to the tube although their attempts substantially decreased in the second session. Indeed, the order of presentation of the conditions rather than the conditions themselves (i.e., seeing the cause of failure or not) seemed to be the factor that best explained subjects' reduction in spitting frequency. In contrast, mean spitting frequency did not differ between conditions. These findings suggest that apes did not take into account visual feedback about the cause of their failure because they did not decrease their spitting behaviour when they could see the water flowing out of the tube. However, a larger sample would be needed to analyse this in greater detail.

E. General discussion

Visual feedback plays a pivotal role in the initial acquisition of the solution in the FPT, but decreases in its importance thereafter. While great apes were able to solve the task again after a period of time without the benefit provided by visual feedback (Experiment 3), they failed to solve the task when visual feedback was absent when they first confronted the task (Experiment 1). Additionally, the type of feedback about failure (i.e., seen or unseen cause) did not alter apes' behaviour. That is, apes did not adapt their spitting behaviour when the water escaped through a hole at the front or back of the tube (Experiment 3). Intriguingly, observing the solution led to success in some individuals: Apes who experienced a water-filled (transparent or opaque) tube solved the task subsequently while experiencing how the water was added to the tube by a human demonstrator or a water tap did not seem to be of further assistance (Experiment 1 and 2).

Contrary to my predictions for Experiment 1, none of the apes solved the FPT spontaneously with an opaque tube. Thus, I did not find evidence that they solved the FPT via insightful learning, but potentially added water to the tube for some reasons in previous studies (e.g., to move the peanut or to make contact with it) and then, perceived the effect that their action had on the peanut. As predicted, the end-state of the solution helped some individuals to solve the task in Experiment 1 and 2. Yet, I expected more individuals to solve the task after the hint. One would need a larger sample of apes to address the significance of an end-state demonstration compared to a baseline group statistically. Since our conditions were presented in a fixed order, I was not able to conduct such an analysis. The results matched my predictions in Experiment 3 as the majority of experienced apes were able to transfer the solution from a transparent to an opaque tube. Overall, the findings reported in this chapter validate previous ones about the role of visual feedback in ape problem-solving and tentatively, their emulation learning abilities (Tennie et al., 2010; Völter & Call, 2012)

One individual solved the opaque tube after experiencing an end-state demonstration, that is, she solved the task without experiencing the effect that adding water to the tube had on the peanut's position. This provides some (albeit weak) evidence that one subject may have anticipated the outcome of her actions in the FPT. However, the FPT may have been too difficult for apes to show their anticipatory abilities (see also Redshaw & Suddendorf, 2016; Völter & Call, 2012). Even in its easier version (transparent tube), only a minority of apes solved the task (Hanus et al., 2011; Tennie, Call, et al., 2010).

Furthermore, the task required the ability to delay gratification as well as the necessary motivation to continue spitting despite not obtaining anything. Although apes generally perform well in delay of gratification tasks and can wait for 60-180 seconds to get a higher valued reward (Beran, 2002; Rosati, Stevens, Hare, & Hauser, 2007), not seeing any change in the peanut's position may have discouraged them. Recall that apes needed on average about 150 seconds from their first spit to retrieve the peanut from the opaque tube in Experiment 3. Therefore, I must interpret my results with caution. Apes may be able to anticipate the outcome of their actions with easier tasks since studies have shown that great apes possess some future planning abilities (Janmaat et al., 2014; Mulcahy & Call, 2006; Osvath & Osvath, 2008; van Schaik et al., 2013; Völter & Call, 2014b).

The findings are consistent with previous studies showing that great apes and New Caledonian crows were dependent on visual feedback for acquisition, but not for maintenance of the solution in a string pulling task (Taylor, Elliffe, et al., 2010; Völter & Call, 2012). Acquisition of the solution in the FPT may be based on at least three different processes. First, apes might have solved the FPT by anticipating the outcome of their actions and some form of causal understanding (see also Köhler, 1925; Redshaw & Suddendorf, 2016). There was little evidence for this. Second, apes may have added water to the tube to move the peanut intentionally (i.e., acted on a creative idea) and then, were differentially reinforced by visual feedback (i.e., the peanut nearing the tube's opening; see also Bateson, 2014). Third, apes may have solved the task by trial-and-error learning and differential visual reinforcement. In this case, they would have added water to the tube by chance repeatedly resulting in differential reinforcement.

The second alternative may be considered more likely than the third one because spitting into the tube is a novel and unusual response (recall that only few apes spat into the tube, see also Hanus et al., 2011; Tennie, Call, et al., 2010). Moreover, apes sometimes solve tasks apparently with little visual feedback about task affordances (Boesch, 2013) or without any evidence of learning by inferring the task's causal structure (Boesch, 2013; Hanus & Call, 2008, 2011; Völter et al., 2016).

The apes solved the task again after months and even years of the original solution. This may be an indicator of intra-individual consistency in problem solving and apes may have re-innovated when presented with the task for a second time. However, it is an open question which characteristics would classify an innovator in the FPT (Uher, Asendorpf, & Call, 2008). A recent study showed that success in another innovation problem was not predicted by divergent thinking or executive functions such as inhibition, working memory, or attentional flexibility in human children (Beck, Williams, Cutting, Apperly, & Chappell, 2016). However, this study found that a measurement that is potentially associated with general intelligence predicted success.

Another interpretation could be that apes' success may be an indicator of long-term memory. They may have remembered the solution, potentially via involuntary memory processes that led to a cued recall when facing the tube (Lewis et al., 2017; Martin-Ordas et al., 2014; Martin-Ordas et al., 2013). Generally, it seems parsimonious to me that apes remembered the solution with both tubes because recent studies have shown a decent long-term memory performance in wild and captive apes, spanning months and even years (Janmaat et al., 2013; Janmaat et al., 2016; Janmaat et al., 2014; Kano & Hirata, 2015; Lewis et al., 2017; Martin-Ordas et al., 2014; Martin-Ordas et al., 2013; Martin-Ordas, Haun, Colmenares, & Call, 2010; Mendes & Call, 2014). However, to clearly disentangle these two possibilities one would need to compare latencies, an analysis I was not able to carry out in this study. However, a recent study has found that chimpanzees who had learned to manufacture an elongated tool three years and seven months ago used the same solution strategy and did so faster when presented with the task again than during first acquisition and also transferred the solution to an opaque apparatus (Vale et al., 2016).

One possible explanation for why subjects solved the opaque versions of these tasks after having solved the transparent ones is that they were able to recall the effect of their actions despite not seeing it. Alternatively, after solution of the task motor programs alone were capable of sustaining the solution despite the lack of visual feedback. Although this could explain the results, it does not seem enough to explain the results of two simpler tasks employed by Völter and Call (2012) that subjects solved even without the benefit of visual feedback. In these tasks, apes had to either poke out a food reward from a transparent or opaque tube or to remove sticks from a transparent or opaque tower so that a food reward was released. Although one could argue that the two less complex tasks provide this evidence given that solutions occurred even in the opaque versions that required multiple steps (Völter & Call, 2012), the actions required to solve these tasks were relatively simple (insert a stick in a tube or remove sticks from a tower), thus, raising the possibility that subjects may have arrived to them by chance. Furthermore, even those simple actions caused some visible change in the state of the world (e.g., sticks off the box) that the other tasks (crank task, FPT) did not provide.

Some of the apes acquired the solution in the FPT after experiencing a water-filled tube with the peanut floating atop. More specifically, in case of the transparent tube two apes benefited from an end-state demonstration and one individual from perceiving how water from a water tap filled the tube until the peanut could be reached (Experiment 2). In case of the opaque tube one subject solved the task and two further individuals added water to the tube after receiving an end-state demonstration once or twice (Experiment 1). Admittedly, this is not a major improvement, but one has to consider that in case of the transparent tube half of my sample comprised previously unsuccessful apes so that chances of them being successful were reduced (Hanus et al., 2011; Tennie, Call, et al., 2010). In fact, previous studies have established that subjects who had failed to solve the FPT in the first two sessions were unable to improve if they were simply given additional sessions (Hanus et al., 2011; Tennie, Call, et al., 2010).

The relative low success attained by subjects even after witnessing the peanut floating upwards suggests that visual feedback (about the water causing the peanut's movement) per se is not a clue that any subject would use to solve the task (contrary to children who would imitate the precise actions required; Nielsen, 2013). One possible explanation is that witnessing an effect is less memorable than causing an effect, but when subjects also had the chance to make the tap drop water in the tube this did not increase success rates. Obviously, here the means that they experienced or used themselves (pulling a string to release water from a tap) during the exposure phase and those that they would have to use during the test (pouring water from the mouth) were different, and consequently subjects may have not transferred the solution using different means. Thus, the findings corroborate the ones of Tennie, Call, et al. (2010), although I did not explicitly test for imitation versus emulation learning since apes were not given the chance to imitate the precise actions (i.e., using a bottle to transport water; for a direct comparison see Call et al., 2005; Horner & Whiten, 2005; Tennie et al., 2006; Tennie, Greve, Gretscher, & Call, 2010; Whiten et al., 2009).

Interestingly, Hanus et al. (2011) found that encountering a partial solution (i.e., a quarter-filled tube with the peanut floating atop) did not facilitate the solution for apes, while it did for children (Hanus et al., 2011), perhaps because in the end-condition apes could access the peanut and touch the water (i.e., were reinforced), whereas they could not do so in the case of the quarter-filled tube. However, future studies are needed to investigate the difference between learning from a partial and a full solution in problem solving situations more closely.

Great apes apply a different technique than children to solve the FPT, that is, they have to pour water from their mouths compared to children who use a pitcher, bottle, cup, or any other hollow object (Hanus et al., 2011; Mendes et al., 2007; Nielsen, 2013; Tennie, Call, et al., 2010). In a recent study, children used a pitcher to water plants prior to the FPT to become familiar with the water (Hanus et al., 2011). However, this prior experience with the water could have resulted in children perceiving the water pitcher as *being for* watering plants which subsequently might have prevented them using it in the FPT (functional fixedness). A similar argument has been made by Hanus et al. (2011) for chimpanzees who did not solve the FPT with a familiar water dispenser which *was for* drinking etc., but overall spitting behaviour towards the tube increased when apes were presented with a novel water dispenser. I therefore investigate in the following chapter if children exhibit a functional fixedness effect when watering plants prior to the FPT.

Chapter 4: Human children – The functional fixedness effect in the FPT[3]

Summary

In this chapter, I report four experiments with 6- to 8- year-old children that investigate the role of prior experience and task presentation in the Floating Peanut Task (FPT). Prior experience with a tool is usually regarded as beneficial for problem solving. Yet, it can also have a negative impact on problem-solving performance by preventing the use of the tool for a novel purpose (functional fixedness). Children were presented with a game that involved collecting three balls of which one was located inside the tube. This ball could only be retrieved by gathering and pouring water from a bucket with a small cup into the tube. Half of the children received prior experience with the water and learned that it was used for watering plants while the other half did not.

In Experiment 1, I tested the functional fixedness effect in 6-year-olds. Children either received prior experience with the tool or not and either performed the actions with the tool themselves or observed them in the experimenter. More specifically, I

---

[3] Material from Chapter 4 formed the basis of the following paper (under review): Ebel, S. J.; Hanus, D.; Call, J.: How prior experience and task presentation modulate innovative problem solving in six-year-old-children.

varied the amount of plants that were watered prior to the test (zero, one, or five) and who watered the plants (own-experience, other-experience).

In Experiment 2, 6-year-olds were presented with the same setup as in the previous experiment. This time I varied the distance of the water bucket to the tube (close, far) and the tube condition, i.e., for half of the children the tube was quarter-filled with water. I used a transparent water bucket for this experiment, while the bucket was opaque in the previous experiment.

In Experiment 3, I tested 6-year-olds in the same conditions reported for Experiment 1 with one crucial change: I used the transparent water bucket from Experiment 2, located close to the tube. Moreover, unsuccessful children received a hint towards the water by the experimenter.

In Experiment 4, 7- to 8-year-olds were presented with the same setup as in the previous experiments. Half of them watered five plants prior to the test, whereas the other half did not water any plants. The tube already contained some water (wet condition of Experiment 2) and the transparent water bucket was placed far from the tube on the floor for all subjects.

The findings of the four experiments suggest robustly low success rates in the FPT for 6-year-olds (Experiment 1-3). However, children’s performance improved to some extent with increased salience of the water bucket (transparent and located close by) as well as a hint given by the Experimenter (Experiment 2 and 3). Due to the low success rates in this age group, I was not able to assess the effect of prior experience with watering plants on performance. However, I did not find evidence for such an effect in 7- to 8-year-olds (Experiment 4). Overall, the findings suggest that 6-year-olds struggle to solve the FPT, but that they are more likely to do so if crucial aspects of the task are made more salient. Thus, although 6-year-olds can find innovative solutions, they require more physical and social scaffolding than older children.

A. Introduction

As suggested in the previous chapters, prior experience with parts of a problem can influence task performance (e.g., Birch & Rabinowitz, 1951; Flavell et al., 1958; Yonge, 1966). While some prior experience may lead to a fixation effect, too much experience can cause a reversed pattern. For example, experts in a given field might flexibly choose from different solution strategies because of their diverse experience (e.g., Bilalić et al., 2008a; Flavell et al., 1958; Star & Seifert, 2006). Previous studies suggest that the functional fixedness effect in humans (Duncker, 1945), which entails a fixation on the function of an object, seems to develop around six years of age (e.g., Defeyter & German, 2003; German & Defeyter, 2000). Interestingly, at this age children also start solving the FPT and comparable innovation tasks such as the hook task (Beck, Apperly, Chappell, Guthrie, & Cutting, 2011; Hanus et al., 2011), whereas they perform poorly at younger ages (see Chapter 1 for a discussion).

One aspect that has received little attention regarding functional fixedness is the role of own- versus other-experience. In other words, is it necessary for an individual to experience the function herself or is it enough to observe the function with another person? From teleological-intentional perspective one would expect that observing the function is enough to establish the idea what the object is *for* and indeed some findings suggest that this is the case in children as young as two-and-a-half years (e.g., Casler & Kelemen, 2005; Defeyter et al., 2009; German & Johnson, 2002; Hernik & Csibra, 2009). While previous studies explored whether children assign functions to objects after observing another individual using them, I focused on whether observing the function would also induce functional fixedness. The FPT seemed a good task to study this effect because it had the right level of difficulty which allowed for a two-sided hypothesis.

In the current chapter, I combined the question of the role of own- versus other-experience in relation to functional fixedness in human children with a paradigm used by Hanus et al. (2011): Children watered plants prior to the FPT to get familiar with the water. It remained an open question if this prior experience with the water potentially helped (facilitating effect) or hindered (functional fixedness) children to come up with the solution in the FPT thereafter. In Experiment 1, I therefore investigated the effect of watering plants (five, one, or zero plants) prior to being confronted with the FPT in 6-year-old children on their success in this task and whether it mattered how children experienced this, namely if they watered the plants themselves or they watched an experimenter doing so (own versus other experience). I chose 6-year-olds because they performed at an intermediate level in the FPT in a previous study, allowing us to entertain a two-sided hypothesis (Hanus et al., 2011). Moreover, the functional fixedness effect seems to develop around the age of six years (Defeyter & German, 2003; German & Defeyter, 2000). I implemented the FPT in a game that required children to collect three balls of which one was located inside the tube. The game was implemented to induce a positive mood in the children and to decrease social pressure since positive affect has been shown to facilitate solutions in creative problems (e.g., Lin, Tsai, Lin, & Chen, 2014). I hypothesized that watering more plants would either have a positive (i.e., facilitating) or a negative (i.e., functionally fixating) effect on success rates with more plants being watered leading to a stronger effect. If results indicated an inverted U-shaped curve instead, this would show that a fixedness effect is only established with a medium amount of experience with the tool. I did not expect own- versus other-experience to modulate the effect since children learn tool use well socially, yet, it would still be an interesting result from a comparative perspective.

Since success rates were extremely low in Experiment 1, I conducted another experiment to find an experimental setup in which children's performance was increased. In Experiment 2, I therefore focused on the salience of the tool and varied the distance of the water to the tube (close versus far) and the condition of the tube (dry versus wet). I hypothesized that success rates would increase with water being close to the tube and that this effect would be even more pronounced when the tube already contained some water. For Experiment 3, I used the most successful condition of Experiment 2 and repeated Experiment 1 with a modified experimental setup. The salience of the water source was increased (i.e., the bucket was placed close to the tube and it was transparent instead of opaque). I had the same hypotheses as in Experiment 1. Since results from Experiment 3 did not allow me to answer my main question if watering plants prior to the FPT had an impact on children's performance due to overall low success rates, in Experiment 4, I focused on 7-to-8-year-olds who either watered five plants or none at all. I hypothesized that watering plants would again have an effect on their problem-solving performance.

B. Experiment 1

1. Methods

Subjects

Participants were 96 6-year-old children (48 girls; age range: 6.0-6.5 years, mean: 6.2 years). For each of the six conditions, I tested 16 children including the same amount of girls and boys. I tested nine additional children that were excluded from the analyses because they either reported to have encountered the task before, e.g., in a teaching context (*N* = 3), because another child had told them the solution (*N* = 2), or because they did not touch the setup (*N* = 4).

Materials

Two tables (L 59 cm x W 30 cm x H 50) were placed next to each other. On one table, there was a Plexiglas tube (L 26 cm x W 5 cm) attached to a piece of wood, a grey tube (about L 8 cm x W 6 cm, diameter of 4 cm), a preserving jar (about H 7 cm, diameter 7 cm) and a wooden pirate ship (L 19.5 cm x W 5.5 cm x H 22.5 cm; see Figure 4.1). A blue ball made of foam (diameter: 2.5 cm) was put inside the vertical tube, a corresponding one in red inside the grey horizontal tube and a yellow one inside the jar (see Figure 4.1). The table was covered with a white sheet before the children entered. On the other table, five, one or no plants at all were placed in a row

(*Spathiphyllum*, about 22 cm high; see Figure 4.1). A round yellow mat was positioned next to the table on the floor (about 89 cm distance to tube).

Depending on the condition, a yellow five-litre bucket (H 22.5, diameter 22 cm) was already standing on the yellow mat (one and five plants condition) or placed at the entrance of the room (zero plants condition). The bucket was filled with water (H 4 cm) onto which a blue cup (H 5.5 cm, diameter 6.2 cm) was floating. To make my study more comparable to studies with non-human great apes, I used a bucket and a cup as water source to investigate if children would always pour several times to fill the tube as apes sometimes stop after a few spits (Hanus et al., 2011; Tennie, Call, et al., 2010). The water source was out of sight when children faced the tube as it was for the apes as well. When children failed the task, they were presented with an additional task that consisted of a wooden box from which they could easily retrieve another blue ball so that all children succeeded to collect the three balls and gained three stickers as a reward.

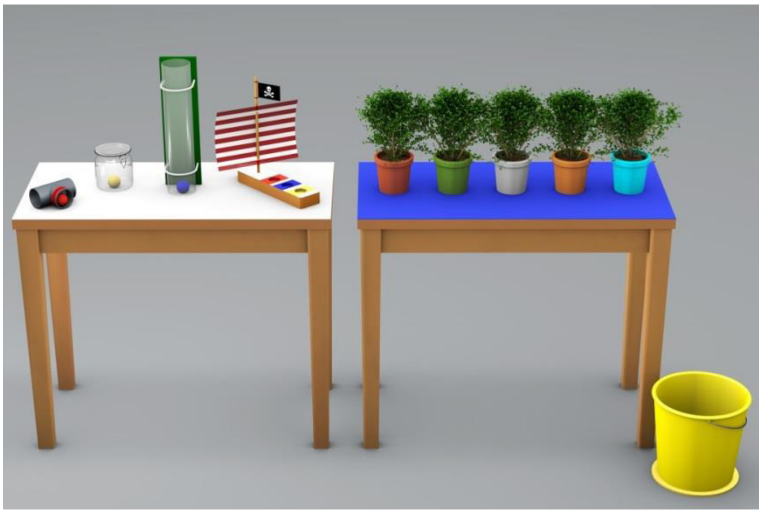


Figure 4.1 Setup of Experiment 1 (the five plants condition is shown).

Procedure

Two factors were manipulated in a between-subjects design: how many plants were watered (zero, one, five) and who did so (own-experience: child, other-experience: experimenter). In the prior experience phase, the experimenter asked the children to water one or five plants with the cup from the water bucket dependent on the condition. In the condition without any plants present, children were asked to carry the water bucket inside upon entering the room. Then, they placed it onto a yellow mat next to the table. This action was performed so that children become aware of the bucket. In the own-experience condition, children performed the watering of the plants (one or five plants condition) or carried in the bucket (zero plants condition) while in the other-experience condition, the experimenter accomplished these actions while the children were watching.

In the test phase, the experimenter retrieved a pirate ship from underneath the white sheet that covered the setup and told the children that they would get a surprise if they managed to collect three balls and to place them into the ship. While children could retrieve two balls easily from a jar and a horizontal tube, one ball was at the bottom of a long vertical dry tube that required children to pour water into the tube to obtain the ball. After explaining the game, the experimenter revealed the setup by removing the white sheet from the table. She told the children that they could try out whatever came to their minds and sat down at the corner of the room. The experimenter stated a motivating sentence every minute (“Just try out another thing! Maybe you have another idea?” or “You can try out whatever comes to your mind.”).

Children had five minutes time to solve the task. In case they did not solve it within this time period, the experimenter would go over and ask them if they had any further ideas what one could try. Children were then allowed to act on the idea if they stated the correct solution. When children did not state the correct solution, they received another (easier) task to obtain a blue ball so that in the end, all children completed the game and won a prize, namely three stickers (for the full text of the procedure, see Appendix A).

Coding and analyses

Children's performance was videotaped. I measured success defined as extracting the ball from the tube. I conducted a generalized linear mixed model (GLMM) with a binomial error structure but it failed to converge due to a floor effect (for details on model formulation, please see Experiment 2 and 3).

2. Results

Figure 4.2 presents the number of children who solved the task as a function of the number of plants watered and the ID of the person who watered them. The extremely low innovation rates (8%) prevented us from assessing differences between conditions.

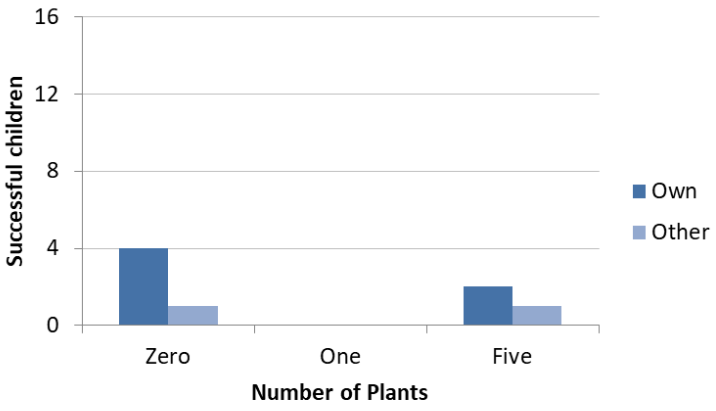


Figure 4.2 Results of Experiment 1.

3. Discussion

This result was quite unexpected as a previous study found that 42% of the 6-year-olds tested solved the FPT (wet and dry condition pooled together; Hanus et al., 2011). Yet, there are some differences between Hanus et al. (2011) and the current study. Most importantly, the water was presented in a much more salient way in the previous study as the transparent water-filled pitcher was placed onto the table in close proximity of the tube (Hanus et al., 2011). Proximity has been shown to determine which parts of the environment subjects see as the problem space (e.g., Simon & Newell, 1971). In Experiment 2, I therefore manipulated the distance of the water (close or far) and the condition of the tube (dry or wet) to increase water salience of the water as a "tool" and boost innovation rates. Besides, I increased the salience of the water by using a transparent bucket. I hypothesized that especially the water being close to the tube would help children to solve the task.

C. Experiment 2

1. Methods

Subjects

Participants were 64 6-year-old children (32 girls; age range: 6.0-6.5 years, mean: 6.2 years). For each of the four conditions, I tested 16 children including the same amount of girls and boys. I tested twelve additional children that were excluded from the analyses because they either reported to have encountered the task before, e.g., in a teaching context (*N* = 6), because another child had told them the solution (*N* = 5) or because they did not touch the setup (*N* = 1).

Materials

The same materials as in Experiment 1 were used. The opaque bucket was replaced by a transparent rectangular one (L 22 cm x W 17 cm x H 16 cm; water H 5.5.cm). No plants were used in Experiment 2.

Figure 4.3 Setup of Experiment 2 (the dry tube and close water condition is shown).

Procedure

Two factors were manipulated in a between-subjects design: the distance of the bucket to the tube (close, far) and if there was already water inside the tube (dry, wet). I placed the bucket on the table about 30 cm to the tube in the close condition while I placed it on the floor next to the table about 89 cm to the tube in the far condition. The tube was completely dry in the dry condition whereas it was quarter-filled with water in the wet condition. Additionally, all children were asked to carry the bucket with water to its predetermined location to reduce their fear of using it. Otherwise, the procedure was the same as in Experiment 1 (see also Figure 4.3).

Coding and analyses

All trials were videotaped. I noted down if children solved the task as in Experiment 1. To analyse the data, I conducted a GLMM with a binomial error structure with solution (yes / no) as a response (R-package lme4, Bates, Maechler, Bolker, & Walker, 2015; R Core Team, 2013). The model included distance of water (close / far), tube condition (dry / wet), sex, and age (z-transformed) as predictors, as well as the interaction between distance of water and tube condition. I included kindergarten as random effect into the model. I assessed model stability by comparing the estimates derived by a model based on all data with those obtained from models with levels of the random effect excluded one at a time. Model stability was acceptable. Variance Inflation Factors (VIF, Field, 2005) were derived using the function vif of the R-package car (Fox & Weisberg, 2011) applied to a standard linear model excluding random effects and interactions, and did not indicate collinearity to be a concern. The significance of the full model in comparison to the null model (comprising only the random effects) was assessed using a likelihood ratio test (R function anova with argument test set to "Chisq"). As a next step, I excluded non-significant interactions from the model and established p-values for the individual effects with likelihood ratio tests comparing the full with respective reduced models (Barr et al., 2013; R function drop1).

2. Results

Figure 4.4 presents the number of children who solved the task as a function of the distance of the water to the tube and the tube condition. The full model did not differ significantly from the null model (GLMM; likelihood ratio test: $\chi^2 = 6.05$, $df = 5$, $p = 0.301$) so that I did not investigate the effects of single predictors further. Apparently, there was no significant difference between conditions (close dry: 50%, close wet: 50%, far dry: 19%, far wet: 25%).

After inspecting the data visually, I decided to run another, exploratory analysis in which I added the interaction of distance of water and sex to the model. The full-null-model-comparison revealed significance (GLMM; likelihood ratio test: $\chi^2 = 12.98$, $df = 6$, $p = 0.043$). Analysing the predictors further indicated that significantly more boys solved the task when the water was close than any other sex and condition combination (Distance of water x Sex, $p = 0.009$; boys close: 63%, boys far: 6%, girls close: 38%, girls far: 38%).

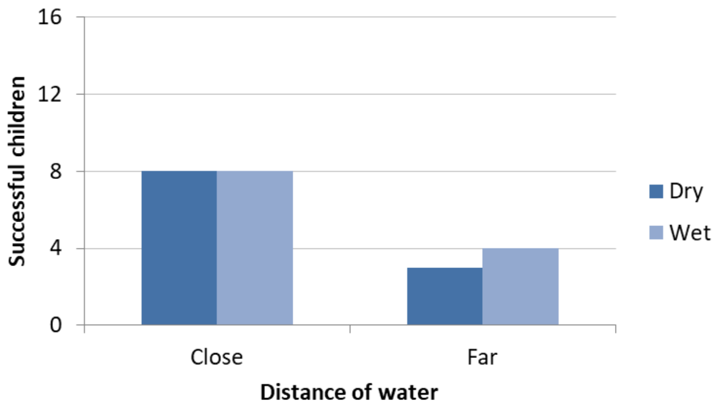


Figure 4.4 Results of Experiment 2.

3. Discussion

Increasing the salience of the water by using a transparent water bucket and placing the bucket closer to and on the same level as the tube helped some children (i.e., mainly the boys) to come up with the solution. However, my conclusion is only tentative because of the post-hoc exploratory nature of this analysis.

In an attempt to confirm this result, in Experiment 3, I used the most successful condition from Experiment 2 (close water) and investigated the same variables as in Experiment 1 (number of watering events and type of experience). The chosen condition allowed us to investigate the direction of the effect of watering plants to go into both directions, either increasing or decreasing innovation rates.

D. Experiment 3

1. Methods

Subjects

Participants were 96 6-year-old children (48 girls; age range: 6.0-6.5 years, mean: 6.1 years). For each of the six conditions, I tested 16 children including the same amount of girls and boys. I tested 15 additional children that were excluded from the analyses because they either reported to have encountered the task before, e.g., in a teaching context (*N* = 4), because another child had told them the solution (*N* = 1), because they did not touch the setup (*N* = 3), because of experimenter error (*N* = 3) or because of other reasons (*N* = 4).

Materials

The same materials as in Experiment 1 were used. The opaque bucket was exchanged by the transparent one from Experiment 2 (Figure 4.5).

Figure 4.5 Setup of Experiment 3 (the five plants condition is shown).

Procedure

I investigated two factors in a between-subjects design: how many plants were watered (five, one, zero) and who watered the plants (child: own-experience, experimenter: other-experience). The procedure was the same as in Experiment 1 except for the following changes: The bucket was transparent and it was picked up at the door and placed onto the yellow mat close to the tube in all conditions (distance: 30 cm). Besides, when children had not solved the task after five minutes, I gave them a hint: The experimenter took the cup from the bucket and poured water with it once inside the bucket mumbling “hmm”. No eye contact was made during this action to keep it as unintentional as possible. She then stated that the child may perhaps have another idea and that she would sit down again for a moment. Children had one additional minute to solve the task.

Coding and analyses

The same recording, scoring, and analytical procedure were followed as in Experiment 1 and 2. The model included number of plants watered, type of experience, sex, and age as predictors, as well as the interaction between number of plants and type of experience and kindergarten as random effect. Model stability and VIFs looked acceptable.

2. Results

Figure 4.6 presents the number of children who solved the task as a function of the number of plants watered and the ID of the person who did so. The full-null-model-comparison did not reach significance (GLMM; likelihood ratio test: $\chi^2 = 4.24$, $df = 7$, $p = 0.752$). Overall, 22% of the children solved the FPT revealing again unexpectedly low innovation rates as in Experiment 1 (Figure 4.6A). When adding the children who solved the task after receiving a hint, 53% of all children innovated. This resembles 40% of the children (30 out of 75) that had failed to solve the task spontaneously (Figure 4.6B).

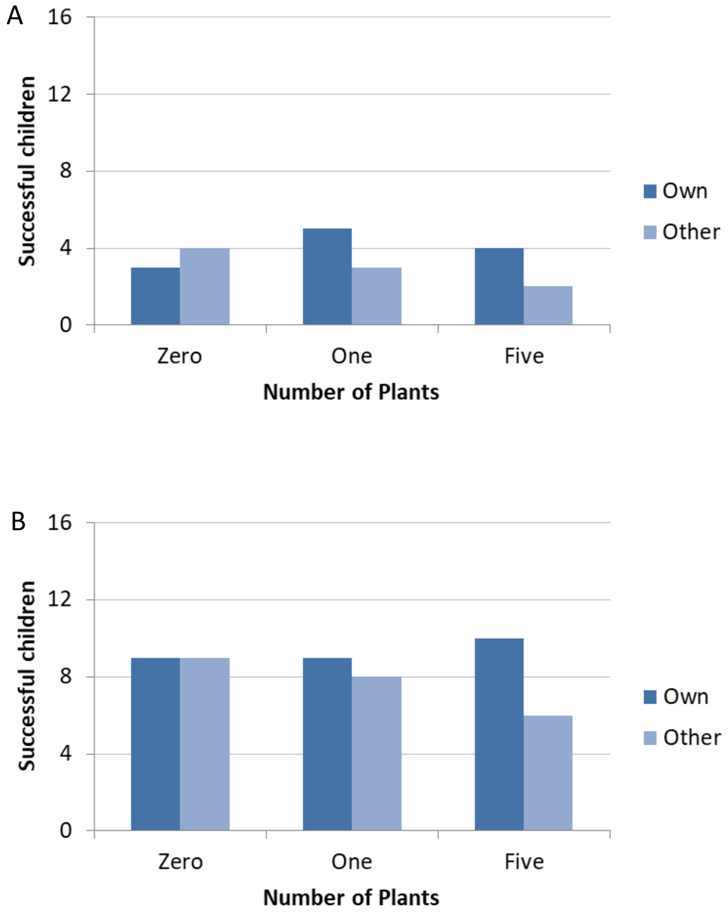


Figure 4.6 Results of Experiment 3: spontaneous solutions (A) and spontaneous solutions and solutions occurring after an experimenter-given hint summarized (B).

3. Discussion

It is remarkable that relatively few children succeeded after their attention was drawn to the water. To investigate the impact of the salience of the water, I directly compared Experiment 1 and 3. I found that significantly more children innovated in Experiment 3 (χ^2-test: $\chi^2 = 5.85$, $df = 1$, $p = 0.016$). Thus, children were more successful when the water bucket was made more salient (i.e., it was transparent and close to the tube plus it was placed on the table either by the children or the experimenter). As only a few 6-year-olds solved the FPT in Experiment 1 and 3, I decided to test 7-to-8-year-olds to tackle my initial question if watering plants (five or zero plants) prior to the FPT had an influence on innovation rates.

E. Experiment 4

1. Methods

Subjects

Participants were 33 7-to-8-year-old children (17 girls; age range: 7.5-8.0 years; mean age: 7.7). For the five and zero plants condition, I tested 16 and 17 children respectively. Children were recruited from a database of children in after-school care centres in a mid-sized German city and some of them had already participated in studies on cognitive development. The socioeconomic background of children was diverse and the parents of the participants had given their informed consent for the study. The study was conducted in a quiet room provided by the after-school care centres. I tested two additional children that were excluded from the analyses because they reported to have encountered the task before, e.g., in a teaching context.

Materials

I used the same materials as in the previous experiments, including the transparent bucket from Experiment 2 and 3. I placed the setup on tables provided by the after-school care centres dependent on their sizes since the previously used tables were too small for the older children. As usual, one ball was inside the transparent vertical tube. The two additional balls were inside a jar and a piece of tube which were slightly harder to open compared to the previously used ones to adjusted to children's age.

Procedure

One factor was investigated in a between-subjects design, namely how many plants were watered (five, zero). The bucket was placed on the floor next to the table (as in Experiment 1) before children entered the room and the tube was always wet (i.e., quarter-filled with water as one of the conditions in Experiment 2). The procedure was the same as in the previous experiments (for an overview of experimental conditions in Experiment 1-4 see Table 4.1).

Coding and analyses

The same type of binomial model was used to analyse the data as in Experiment 3 but only included the number of plants watered, sex, and age as fixed effects and kindergarten as random effect. Model stability and VIFs looked adequate.

Table 4.1 Experimental conditions in Experiment 1-4.

Exp.	**Age**	**Number of plants**	**Type of experience**	**Distance of water**	**Condition of tube**	**Bucket**
1	6	*Five, one, or zero*	*Own or other*	Far	Dry	Opaque
2	6	Zero	Own	*Close or far*	*Dry or wet*	Transparent
3	6	*Five, one, or zero*	*Own or other*	Close	Dry	Transparent
4	7-8	*Five or zero*	Own	Far	Wet	Transparent

2. Results

The full-null-model-comparison did not reach significance (GLMM; likelihood ratio test: $\chi^2 = 0.88$, $df = 3$, $p = 0.831$). Half of the children solved the task in both conditions. Figure 4.7 presents the number of children who solved the task as a function of the number of plants watered.

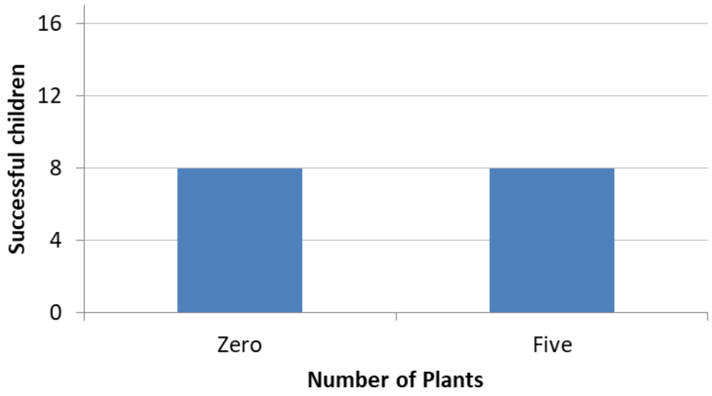


Figure 4.7 Results of Experiment 4.

3. Discussion

There was no evidence of functional fixedness or a facilitating effect of watering five plants prior to the FPT in 7- to 8-year-old children. Half of the children solved the task in both conditions. Children performed somewhat worse than in a previous study (75% success in the wet condition; Hanus et al., 2011), but studies differed with regard to multiple variables (for an overview, see Table 4.2). First, Hanus et al. (2011) used a pitcher, whereas I made use of a water bucket. Second, the setup of the other authors comprised a tube and a water pitcher only, which were both located on a table. I embedded the task in a game to give children some experience of success (i.e., when collecting two additional balls that were easy to retrieve from two containers). Thus, there were multiple objects on the table: the tube, two containers with two foam balls inside and a pirate ship (to collect the balls). Additionally, the water bucket with a small cup floating on the water and one or five plants were located on the adjacent table. Third, a session in Hanus et al. (2011) lasted ten minutes, whereas it lasted only five minutes in mine. Fourth, the other authors used a peanut as floating object, whereas I employed a foam ball.

Especially the number of objects on the table as well as the tool (i.e., pitcher or bucket) might have influenced children’s performance. The objects might have distracted children from the task so that they spent time manipulating them. They also tried to use them to solve the task, e.g., by relating them to the tube’s opening (although the objects were obviously too large to fit inside the tube). Moreover, a pitcher is more likely to be associated with pouring water so that object affordances might have facilitated solutions in Hanus et al. (2011). The other two aspects are less likely to have influenced children's behaviour. I suppose that first, foam balls are more likely to be associated with floating on water and second, most children solved the task within the first few minutes in the previous study.

Table 4.2 Comparison of developmental studies with the FPT.

Study	**Age**	**Tool**	**Distance water**	**Condition of tube**	**Floating object**	**Additional objects**
Hanus et al., 2011	4, 6, 8 years	Transparent pitcher	50-80 cm; on table	Dry; wet	Peanut	No
Nielsen, 2013	4, 20 years	Transparent bottle; 2 cups	Close; on table	Dry	Plastic monkey	No
Experiment 1	6 years	Opaque bucket; 1 cup	89 cm; on floor	Dry	Foam ball	Yes
Experiment 2	6 years	Transparent bucket; 1 cup	30 cm; on table; 89 cm; on floor	Dry; wet	Foam ball	Yes
Experiment 3	6 years	Transparent bucket; 1 cup	30 cm; on table	Dry	Foam ball	Yes
Experiment 4	7-8 years	Transparent bucket; 1 cup	89 cm; on floor	Wet	Foam ball	Yes

F. General discussion

No evidence for a functional fixedness effect with regard to prior experience (i.e., watering plants) in the FPT was found in 6-year-old children despite repeated attempts. Overall, success rates in 6-year-olds remained very low (20%, pooled data from Experiment 1-3, 52/256 children). Performance in this age group increased when children were presented with a non-social cue (increased salience of the water bucket) or a social cue (experimenter given hint towards the water). An additional non-social cue (tube quarter-filled with water) did not have the same effect. The ID of the person watering the plants (child or experimenter) did not influence task performance.

The findings of this chapter do not reflect most of my predictions. Overall, I found much lower success rates in human children than a previous study (Hanus et al., 2011; for a discussion, see below) and none of the tested conditions had a significant impact on children's performance. Contrary to my predictions, the study did not reveal that experience with the tool influenced children's problem-solving performance in the FPT. I could not assess the impact of own versus other experience and the amount of experience statistically due to low success rates in Experiment 1 and 3. Yet, I would not expect an effect of both variables with this specific experimental setup because both variables are dependent on a significant effect of prior experience on task performance, which I did not find in Experiment 4. In the following, I would like to discuss potential reasons for the unexpected findings.

No functional fixedness effect was found with regard to watering plants prior to the FPT. Perhaps using a pitcher or a bottle (Hanus et al., 2011; Nielsen, 2013) instead of a water bucket would have changed the results. Buckets are commonly used for multiple purposes but they are associated with pouring water less often than pitchers and bottles (i.e., they have different tool affordances). It would be interesting to use a watering can instead, which is *made for* watering plants (see also Defeyter & German, 2003; Defeyter et al., 2009; Hernik & Csibra, 2009; Ruiz & Santos, 2013).

Another possibility why I did not find an effect could be that five pouring actions were not enough. Future studies could explore more closely how much exposure is needed to induce a functional fixedness effect in children (see also Flavell et al., 1958; Yonge, 1966). As I did not find a functional fixedness effect, I could not investigate the impact of own- versus other-experience. Some authors have suggested that humans are unique in their representation of objects, that is, they see artefacts as made by other individuals for specific purposes (e.g., Defeyter & German, 2003; Defeyter et al., 2009; German et al., 2007; Hernik & Csibra, 2009; Ruiz & Santos, 2013). Thus, one may propose that human children show the same degree of functional fixation independent on if they make the experience themselves or if they observe someone. This is another aspect that remains under studied and which deserves further research attention. Moreover, it would be important to investigate factors modulating the functional fixedness effect in slightly older children since this phenomenon only emerges around the age of 6 years (e.g., Defeyter & German, 2003; German & Defeyter, 2000) and the children in the current study had just turned six.

Another reason for the low success rates may be that children hesitated to use water indoors for fear of spilling it on the floor. Many children indeed asked if they could use the water before doing so, even in the wet condition in which there was already some water located inside the tube. To reduce fear of using the water, I told children spilling water was no problem when they watered the plants. I also encouraged them to try out any idea they had. After the test, I asked them for further ideas to give them a chance to state the solution to rule out that they did not dare to act on their correct idea. It would still be interesting to present children with the FPT on an outdoor playground to lower the hesitation to employ water as well as to remove the constraints of a test situation (see Bonawitz et al., 2009). Besides, there is no evidence that low innovation rates in the hook task can be explained by children's hesitation to manipulate the target object, namely bending the pipe cleaner (Cutting, Apperly, & Beck, 2011). In sum, children may hesitate to employ the water in the FPT but it is unlikely that this is the main reason why they struggled with this problem.

Perhaps the late emergence of innovative problem solving in children is not that surprising. Since adults take care of children, there is no need for them to innovate, e.g., to find novel food sources. While there is evidence that children in some hunter-gatherer groups already contribute substantially to sustain themselves, the amount of food gained depends on the children's age with older ones contributing more than younger ones (e.g., Hawkes, O'Connell, & Blurton Jones, 1995). Besides, human children show a strong bias towards social learning (e.g., Behne et al., 2005; Csibra & Gergely, 2009; Wood, Kendal, & Flynn, 2013) which sometimes even leads to imitation of clearly irrelevant actions ("overimitation", e.g., Lyons et al., 2007). To sum up, there may be no need for younger children to have innovative abilities because they are taken care of by older group members and their main learning focus is to copy others rather than to learn individually (e.g., Csibra & Gergely, 2009).

Success rates in Experiment 3 increased when children received a non-social cue about the water bucket by increasing its salience. When the bucket was transparent and placed close to the tube, children were more likely to succeed. Maybe enhanced proximity and visibility allowed children to perceive the bucket as part of the problem space and therefore, as a potential "tool". Interestingly, water that was already located inside the tube did not have the same effect in 6-year-olds. When Hanus et al. (2011) presented 4-, 6- and 8-year-olds with a dry or a wet tube, they found increased innovation rates with age and tube condition. However, when only looking at 6-year-olds, only two additional children solved the FPT when there was already water located inside the tube (dry: 33%, wet: 50%), indicating no major difference within this age group. Taken together, these two studies suggest that 6-year-olds did not understand that the water inside the tube was a hint to the solution, perhaps because it did not draw their attention to towards the "tool" itself (i.e., the water bucket).

Some children found the solution after they had obtained a social cue about the water bucket (i.e., the tool). They benefitted from observing the experimenter pouring water with the cup inside the water bucket once. After receiving a hint, 40% of the beforehand unsuccessful children (*N* = 30) came up with the correct solution. One possibility is that the hint drew their attention towards the water bucket which then became part of the problem space, leaving open the question if the experimenter given hint was taken as an ostensive cue by the children. Most of the children (86%)

solved the task immediately after the hint, either stating the solution while the experimenter was still touching the water or straight away when the experimenter sat down at the corner of the room. Yet, when children were asked how they had come up with the solution, only few children (27%) referred to the action of the experimenter and then, they did so in a descriptive way. Only one of these children reported that the experimenter had actually given her a hint. Most children (70%) instead reported other ways how they had come up with the solution (and one child did not give a sensible answer).

Children’s behaviour may therefore indicate a facilitated recombination by drawing attention to the water while not necessarily being caused by ostensive communication. However, even if the majority of the successful children did not report that they were given a hint, I cannot be sure if they nevertheless interpreted the actions of the experimenter as an ostensive cue. It is possible that the question was quite difficult to answer for 6-year-olds and they sometimes seemed confused about us asking. It is possible that a clearly ostensive cue like the experimenter pointing to the water, then looking at the children, smiling and raising her eyebrows would have resulted in much higher innovation rates as children are known to be sensitive to ostensive communication (e.g., Behne, Carpenter, & Tomasello, 2005; Csibra & Gergely, 2009).

Finally, children showed a clear pattern when it comes to pouring water into the tube. Once they had the idea, they continued pouring the water until they could reach the ball. Recent studies showed a slightly different pattern in non-human great apes, with some of them acting the same as the children while others stopped adding water without obtaining the peanut (e.g., Hanus et al., 2011). Children often stated the solution before employing it, probably to make sure that they were allowed to use the water. Thus, they clearly anticipated the outcome of their actions. Encountering a quarter-filled tube neither helped 6-year-olds, nor apes (Hanus et al., 2011). This is surprising as a quarter filled tube constitutes a partial solution and I know that very young children and non-human great apes benefit from encountering the full solution (the "end-state", e.g., Bellagamba & Tomasello, 1999; Huang, Heyes, & Charman, 2002; Tennie, Call, et al., 2010). Only by the age of eight years, children seem to benefit from encountering a partial solution in the FPT (Hanus et al., 2011).

In conclusion, I did not find a functional fixedness effect with regard to prior experience in the floating peanut task in 6-year-olds. Yet, I found robust low innovation rates in 6-year-olds. A non-social cue (proximity and visibility of the water) and a social cue (an experimenter given hint) increased performance though overall innovation rates still stayed modest. Nonetheless a minority of children found the innovative solution suggesting that some 6-year-olds have the capacity to innovate but that they may be more dependent on greater physical and social scaffolding than older children and adults. Although I did not find a functional fixedness effect in my experiments with human children, perhaps due to the experimental design, previous studies have already established the effect in children (Defeyter & German, 2003; German & Defeyter, 2000). In the following chapter, I therefore return to great apes and tackle the question if our closest living relatives are vulnerable to the functional fixedness effect as well.

Chapter 5: Great apes – The functional fixedness effect revisited[4]

Summary

In this chapter, I discuss three experiments on the role of prior experience with tools in great ape problem-solving. More precisely, I investigate if apes have difficulties to overcome the familiar function of a tool and struggle to use it for a novel function (functional fixedness). Apes receive experience with three different types of tools: one with two functional ends, one in which the two functions are not separated in space, and a food item that can be used as a tool.

In Experiment 1, apes either received prior experience with the brush end of a tool (i.e., they dipped juice from a container) or not. In the test, they were presented with the same container, but this time the dipping option was blocked. Instead, they had to use the pointed end of the same tool to punch a hole into the container to access the juice.

In Experiment 2, apes either received prior experience with a hose (i.e., they drank juice from a container) or not. In the test, they were presented with a horizontal

---

[4] Material from Chapter 5 formed the basis of the following papers (in preparation for submission): Ebel, S. J.; Völter, C. J.; Call., J.: Functional fixation in the tool use of captive great apes (*Pan paniscus, Pan troglodytes, Pongo abelii*). Ebel, S. J.; Völter, C. J.; Call., J.: Functional fixedness in great apes invoked by a food item.

tube with blockages close to both openings. Subjects had to select the flexible hose to poke out the food reward from the tube.

In Experiment 3, apes were first fed with a novel food item (grissini / bread stick) as a whole, broken into pieces or they did not receive any experience with the food. In the test, they were presented with an out of reach reward on a platform. To access the food, they had to use the bread stick as a raking tool.

Results indicated that prior experience shaped apes' manipulation pattern with the apparatus (Experiment 1) and led to a decrease in problem-solving abilities (Experiment 2 and 3). More specifically, apes who had prior experience with a tool were less likely to use the tool with a novel function than naïve ones. Finally, apes who received experience with a novel food item were less likely to employ it as a tool in the first trial than naïve ones (Experiment 3). The findings suggest that great apes, like humans, are vulnerable to the functional fixedness effect.

A. Introduction

Some authors have argued that human object representations are unique among the animal kingdom and only humans represent objects as being *made for* a specific function by an agent (design stance or teleological-intentional stance; Casler & Kelemen, 2007; Csibra & Gergely, 2009; German & Johnson, 2002; Hernik & Csibra, 2009; Ruiz & Santos, 2013; Vaesen, 2012). According to this account, great apes would not exhibit a functional fixedness effect possibly due to a lack of (enduring) artefacts in their environment and a different social setting than that of humans (e.g., Csibra & Gergely, 2009; Hernik & Csibra, 2009; Vaesen, 2012). This position, however, neglects comparative evidence that is consistent with functional fixation in non-human primates. As I have already elaborated on fixation effects in primate tool-use in the general introduction, I summarize the evidence only briefly in the following.

Recent studies have shown that great apes exhibit rigid or conservative behaviour in problem-solving situations: They remain with the solution strategy that they have acquired first, even if a more effective one becomes available to them (Gruber et al., 2011; Gruber et al., 2009; Harrison & Whiten, 2018; Hrubesch et al., 2009; Marshall-Pescini & Whiten, 2008; Price et al., 2009). Some of these studies involve tool use; for example, Marshall-Pescini and Whiten (2008) presented captive chimpanzees with an array of several honey boxes. The chimpanzees received two types of demonstrations by a human demonstrator: Firstly, they observed how the demonstrator dipped honey with a rod on the top opening of the honey boxes. Secondly, they witnessed how the same rod was used to poke the side of the box to release a bolt which allowed for the lid of the box to be opened. Chimpanzees who had learned to dip honey with the rod from the top opening of the boxes sustained their solution strategy, even if they were demonstrated a more effective strategy.

In another study, two groups of wild chimpanzees were presented with a honey trap task and a tool with two functional ends (i.e., a twig with a stick-end and a leaf-end; Gruber et al., 2011). One group of chimpanzees regularly used the sticks as a tool in their everyday lives, whereas the other group did not. Both groups made sponges out of leaves that they then employed as tools. The authors found that chimpanzees of the stick-using population used the stick-end of the tool to extract the honey, whereas the population without regularly occurring stick tool usage in their everyday lives either tried to access the honey with their fingers or produced a sponge out of the leaf-end of the tool in order to extract the honey more efficiently. Pre-inserting the stick-end into the honey trap did not alter their behaviour. The authors conclude that both populations found different parts of the tool functionally salient (Gruber et al., 2011; see also Gruber et al., 2009). These two examples of perseverative behaviour in apes show that prior experience with tools made their problem solving less effective: They required longer time to retrieve the food rewards from the apparatus than if they had switched to the more effective solution strategy. I aim to further extend these findings by investigating functional fixedness in great apes in more detail in three different experiments.

Functional fixedness had not been studied systematically with an experimental design in captive great apes yet. In Experiment 1, I therefore let great apes either dip juice with the brush end of a tool from the top opening of an apparatus or let them explore the tool without a task present during the prior experience phase. In the test, the dipping option was blocked and subjects were required to use the pointed end of the same tool to puncture a hole in the bottom area of the apparatus to access the juice. I measured success and time to success to assess problem-solving performance in the test. Moreover, I coded manipulations with and without the tool to better understand what a potential difference between groups could be based on. I hypothesized that apes with prior functional experience with the tool would show a worse problem-solving performance than those from the control group. Moreover, I expected them to manipulate the apparatus more with the brush end of the tool at the top opening (as in the prior experience phase), whereas I expected the control group to explore the apparatus more in general than the experimental group.

Since the task and the tool were the same in the prior experience and the test phase, this setup did not differentiate between a functional fixedness (which is a fixation on the function of a tool) and the Einstellung effect (which is more broadly the fixation on a solution strategy, potentially also comprising the tool, but also the apparatus). I therefore conducted Experiment 2 to specifically test for a functional fixedness effect in great apes. Here, great apes either drank juice from a container using a hose, or explored the hose with the empty container present during the prior experience phase. In the test, subjects had to select the hose among three tools (a hose, a stick, or a string) to poke out a food reward from a horizontal tube that required using a flexible tool to get it out. I measured again success and time to success to assess problem-solving performance. I additionally measured time until target tool selection as three tools were involved because some researchers consider this measurement to be more precise to assess functional fixedness than time until success because individuals vary in their time to execute the task (Defeyter & German, 2003; German & Barrett, 2005). Moreover, I coded time until touching the hose, sucking attempts and which tool apes selected first to use it at the tube to better understand a potential difference between both groups. I again hypothesized that apes with prior experience with one function of the tool would solve the tasks less often, take longer to do so and also to select the target tool than apes who had not assigned a function to the tool yet.

The first two experiments comprised non-edible tools, which functions had to be learned by the apes. In Experiment 3, I aimed at using a more intuitive function of a tool, namely being *being for* eating or nourishing oneself. I investigated whether apes from the four great ape species would use a food item (i.e., a grissini/ bread stick) as a tool depending on their experience with it. Apes were pre-fed with either whole grissini, grissini pieces, or they were not pre-fed at all. In the test, I presented subjects with a reward located out of their reach, which they could rake in with a grissini. I measured tool use with the food item and hesitation behaviour to use it as a tool. I hypothesized that the apes would generally be able to use a food item as a tool and based on the functional fixedness effect, that naïve apes would be more likely to do so. If they were not able to use the food item as a tool, this would potentially suggest that they were either not able to use a food item with a different function or that they lacked the required inhibitory control. If I found that some apes were able to use the grissini as a tool, but their performance was not dependent on their experience with the food, this would indicate that there was no evidence of functional fixedness with a food item from the current study. I conducted some control tests to rule out that apes' performance was dependent on their general raking abilities or on individual food preferences.

B. Experiment 1

1. Methods

Subjects

Twenty-seven great apes (bonobos, *Pan paniscus*, $N = 6$; chimpanzees, *Pan troglodytes*, $N = 15$; orang-utans, *Pongo abelii*, $N = 6$; $N_{females} = 18$; age range: 7-40 years; mean age: 21 years; Table 5.1) participated in the study. Four apes (three chimpanzees and one bonobo) were excluded from the study because they did not reach the dipping criterion in the prior experience phase ($N_{final} = 23$). Five of the chimpanzees had acquired a preference for a wooden stick with a broadened (brush like) end over a non-modified one for dipping a liquid food in a previous study. Some of these individuals had previously also manufactured a brush-like tool themselves by chewing one end of the stick (unpublished data). I distributed these individuals equally among the two groups to account for these individual differences in prior experience.

Table 5.1 Subjects participating in Experiment 1.

Subject	**Species**	**Sex**	**Age**	**Rearing**
Fimi	Bonobo	Female	8	Mother
Gemena	Bonobo	Female	10	Mother
Luiza	Bonobo	Female	11	Mother
Lexi	Bonobo	Female	17	Nursery
Yasa[1]	Bonobo	Female	19	Mother
Kuno	Bonobo	Male	19	Nursery
Kisha[1]	Chimp	Female	12	Mother
Tai	Chimp	Female	14	Mother
Swela[1]	Chimp	Female	20	Mother
Sandra	Chimp	Female	23	Mother
Dorien	Chimp	Female	35	Nursery
Natascha	Chimp	Female	36	Nursery
Riet	Chimp	Female	38	Nursery
Corrie[1]	Chimp	Female	39	Nursery
Fraukje	Chimp	Female	40	Nursery
Bangolo	Chimp	Male	7	Mother
Kofi	Chimp	Male	11	Mother
Lobo	Chimp	Male	12	Mother
Lome	Chimp	Male	15	Mother
Frodo	Chimp	Male	22	Mother
Robert	Chimp	Male	40	Nursery
Raja	Orang	Female	13	Mother
Padana	Orang	Female	18	Mother
Dokana	Orang	Female	27	Mother

Pini	Orang	Female	28	Mother
Suaq	Orang	Male	7	Mother
Bimbo	Orang	Male	36	Nursery

[1] were excluded from the study

Materials

I used a Plexiglas tube (H 17.5 cm; outer diameter: 9 cm) with an opening at the top that was either left open (during the prior experience phase) or covered with a metal mesh (during the test phase); it also contained a hole at the bottom on the front side (diameter: 0.9 cm) that either was connected to the inner part of the tube (test phase) or was not (prior experience phase; Figure 5.1). The bottom hole was covered with a piece of tape from the inside in both phases. The tube was filled with grape juice (chimpanzees, orang-utans: 120 ml; bonobos: 100 ml) and was attached to the mesh of apes' sleeping rooms (hereafter: test rooms) from the outside. In case of the orang-utans, I attached a thin tube (length 9.5 cm) at the back of the apparatus in the test phase in case orang-utans would spit water into the tube (Mendes et al., 2007), which they did not. I therefore had to slightly reduce the grape juice for the orang-utans slightly due to this adjustment (100 ml; test phase only). The wooden brush tool (about L 30 cm x W 2 cm x H 0.7 cm) comprised two functional ends: a brush end (about L 5 cm) and a pointed end (about L 2.5 cm).

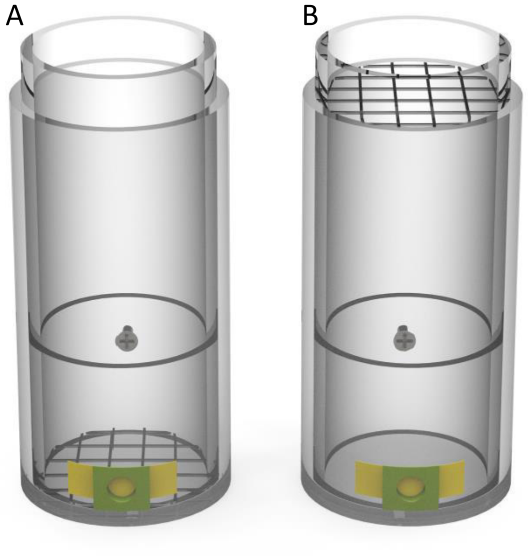


Figure 5.1 Setup of Experiment 1: In the prior experience phase, apes could dip juice from the top opening of the apparatus with the brush end of the tool (A); in the test phase, the top opening was blocked by a mesh so that apes had to poke a whole into the apparatus at the bottom opening with the pointed end of the tool (B).

Procedure

Apes were distributed pseudo-randomly into two groups with restrictions that I counterbalanced the two groups as much as possible for prior experience, species, age, and sex. Both groups entered a prior experience phase and a test phase. In the prior experience phase, the prior experience group (N = 11) learned that the brush end of the tool functioned to extract the grape juice in an efficient way (4-13 sessions, mean: 7 sessions), whereas the no experience group (N = 12) was presented with the brush tool lying on the metal frame of the mesh panel in the test room, with no apparatus present to control for the novelty of the tool (two sessions). All sessions lasted five minutes each and one session was conducted per day. When subjects from the prior experience group were still dipping after five minutes, they were permitted to finish the juice.

Subjects from the prior experience group were presented with the tool lying on the metal frame (counterbalanced for direction of the brush end) in front of the tube (baseline). However, in the first session, the brush end of the tool was pre-inserted into the tube to facilitate finding the dipping solution. Additionally, I applied three steps if apes were not dipping on their own, each for two sessions: 1) subjects were presented with a second session with the brush tool pre-inserted into the tube (as mentioned); 2) the experimenter showed them how to use the tool (i.e., she offered them juice from the tube and then, inserted the brush end of the tool into the tube to give subjects a try; this was sometimes performed multiple times within one session); 3) subjects were given a different brush in case that their failure was dependent on the specific brush that was being used (this condition was only applied with three chimpanzees that were excluded from the study eventually). When subjects dipped the grape juice on their own, they received baseline sessions thereafter.

The prior experience phase was completed when subjects had reached the criterion of 200 dipping events with the brush end of the tool (brush end: 200-261 events, mean: 222 events; pointed end: 0-29 events, mean: 7 events). I added one session with one bonobo (Luiza) because she dipped the juice 17 times with the pointed end in her last session and I aimed at strengthening the function of the brush end of the tool. If the apes did not dip the brush end of the tool into the juice on their own accord within the six shaping sessions, they were excluded from the sample; this applied to three chimpanzees. Additionally, one subject stopped dipping after 98 dipping events and was excluded as well (Yasa).

In the test phase, subjects were presented with the same vertical tube with a top opening blocked by a mesh material. Here the subjects had to puncture the tape covering the hole at the bottom front part of the tube with the pointed end of the tool to access the juice. Subjects received two sessions with five minutes each. When apes were only to solve the task in their second session, they received a third one.

Coding and analyses

I coded overall if subjects solved the task as well as the survival time until first success. The survival time is a compound of time passed and success. Moreover, I measured duration of manipulation by type (brush end, pointed end, touch with hand or mouth) at three different areas of the tube (top opening, bottom hole, other areas). I calculated the relative manipulation time for each combination of manipulation type and tube area for all subjects and sessions. A second person coded 20% of the videos and both coders were in good agreement (Cohen’s Kappa: success: $K = 1$, $N = 10$, $p = 0.002$; Pearson’s correlation: survival time: $r = 1$, $df = 8$, $p < 0.001$; manipulations – all combinations of tube area and manipulation type combined: $r = 0.99$, $df = 88$, $p < 0.001$; dipping – brush end and pointed end combined: $r = 1$, $df = 24$, $p < 0.001$).

All analyses were performed with R 3.0.2 (R Core Team, 2013). Success was analysed with a generalized linear model (GLM) with a binomial error structure (*N* = 23; R package "lme4", Bates et al., 2015). The model comprised group and age as fixed effects. Age was first log-transformed and then z-transformed to a mean of zero and a standard deviation of one. I derived Variance Inflation Factors (VIFs) to check for collinearity, which indicated collinearity to be no issue (R function “vif” of the package “car”; Field, 2005; Fox & Weisberg, 2011). I performed a full-null-model comparison by comparing the full with a reduced model that included the intercept only with a likelihood ratio test (LRT) to establish the overall effect of the predictors (R function "anova" with argument "test" set to "Chisq"; Schielzeth & Forstmeier, 2009). In case of a significant difference between both models, I established p-values for the individual predictors with LRTs comparing the full with respective reduced models (R function "drop1"; Barr et al., 2013). Survival time was analysed with a survival model and involved the same predictors as the model with success as the response (*N* = 23; R package “survival”; Therneau, 2015; Therneau & Grambsch, 2000). Moreover, the treatment of the model was the same as in the previous one.

Manipulation was analysed in a two-step approach, given that 51% of a given combination of manipulation type and tube area equalled zero. First, I analysed the occurrence of manipulation as binary variable (yes/no) for all combinations of manipulation type and area using a generalized linear mixed model (GLMM) with a binomial error structure (*N* = 441). Thereafter, I analysed the subset of the data in which manipulation occurred with a GLMM with a Gaussian error structure and relative duration of manipulation as the response (*N* = 214). Both models comprised the following fixed and random effect structure. I included the four-way-interaction between group, manipulation type, tube area and age, but the model did not converge. Thus, I included the three-way interaction between group, manipulation type and tube area as well as age, species, sex, and session as fixed effects and the random slopes of area, manipulation type, and session within subjects (Barr et al., 2013; Schielzeth & Forstmeier, 2009). The binomial model additionally included an offset term (log-transformed duration of session divided by 60) to take care of varying session lengths. Age was log-transformed and age and session were z-transformed to a mean of zero and a standard deviation of one.

VIFs were assessed for both models and did not find collinearity to be an issue. Moreover, I evaluated normal distribution and homogeneity of the residuals of the Gaussian model by plotting the residuals which looked acceptable. I also assessed model stability by looking at the impact of excluding one level of the random effect at a time for the estimates of the fixed effects. As a result of this, model stability seemed adequate. The Gaussian model was highly unstable when excluding one specific chimpanzee (Lobo) being that he was the only representative in one of the combinations of the three-way interaction (the Gaussian model was based on a subset of the data). I compared the full models with respective reduced models that only comprised the random effect terms using a LRT; I then, excluded non-significant interactions and established p-values using LRTs (see model with success as the response). In order to further investigate significant interactions, I re-leveled the respective factors (binomial model) or subsetted the data (Gaussian model).

2. Results

Success and survival time

The GLM with success as dependent variable was not significant when compared to the null model (LRT: $\chi^2 = 4.59$, $df = 2$, $p = 0.101$). Fifty percent of the subjects from the no experience and 18% from the prior experience group solved the task (Figure 5.2A). More specifically, four subjects from the no experience and one from the prior experience group solved the task in the first session and two from the no experience and one from the prior experience group solved it in the second session.

The model with survival time as dependent variable indicated a trend when compared to the null model (LRT: $\chi^2 = 5.91$, $df = 2$, $p = 0.052$). Subjects from the no experience group tended to solve the task faster than those from the prior experience group (LRT: $\chi^2 = 3.06$, $df = 1$, $p = 0.080$; Figure 5.2B). Moreover, younger subjects tended to solve the task faster than older ones (LRT: $\chi^2 = 3.14$, $df = 1$, $p = 0.076$; age of successful subjects – mean: 17 years, minimum: 7 years, maximum: 40 years; mean age of sample: 21 years).

One orang-utan (Padana) from the prior experience group managed to break the mesh in her second session after 211 seconds and dipped the juice with the brush end of the tool. I gave her two additional sessions in which I replaced the mesh material with a round plate made of polyvinyl chloride (PVC) with small holes inside. The subject did not solve the task during these additional sessions.

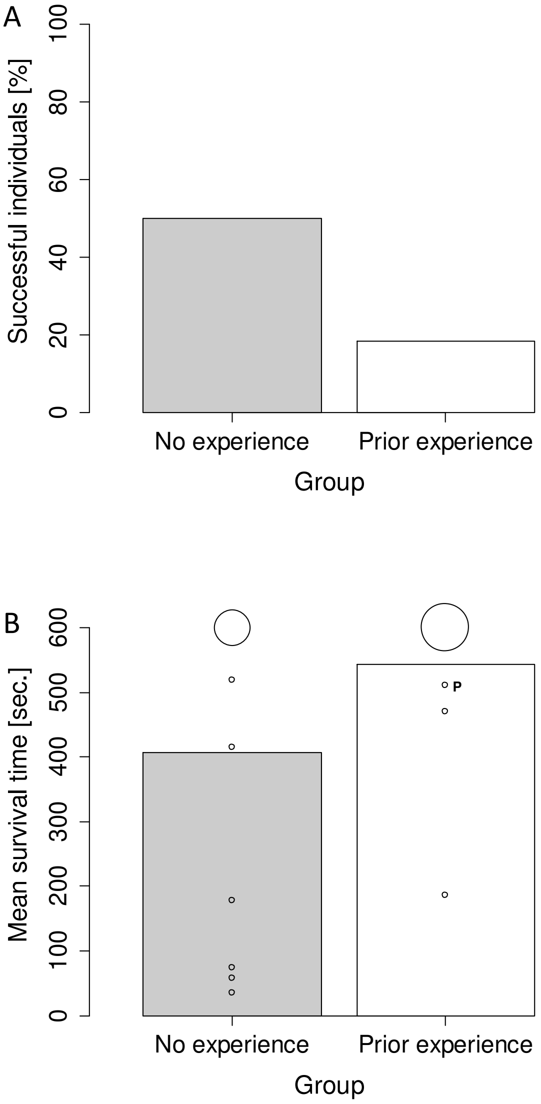


Figure 5.2 Results of Experiment 1: Success (A) and survival time (B) as a function of group. Circles indicate individual performance with larger circles referring to more individuals. ^{P} One subject (Padana) broke the apparatus before the end of the session.

Manipulation: binomial model

I analysed the manipulation of the tube in the test phase in a two-step approach. First, I looked at the occurrence of manipulation for all combinations of manipulation type (brush end, pointed end, touch with hand or mouth) and tube area (top, bottom, other). The full model (GLMM) fitted the data significantly better than the null model (LRT: $\chi^2 = 102.18$, $df = 22$, $p < 0.001$; Appendix C, Table C.1). I first removed the non-significant three-way interaction between group, manipulation type, and tube area (LRT: $\chi^2 = 1.19$, $df = 4$, $p = 0.880$) and the non-significant two-way interaction between group and manipulation type (LRT: $\chi^2 = 1.01$, $df = 2$, $p = 0.603$) from the model. I obtained a significant interaction between group and tube area (LRT: $\chi^2 = 12.63$, $df = 2$, $p = 0.002$).

Examining the interaction further exposed three significant main findings (see Appendix C, Table C.2): Firstly, subjects from the prior experience group manipulated the top area more frequently than the ones from the no experience group and even within the group they also manipulated it more often than the additional two areas. Secondly, I found the reverse to be true for the bottom area of the tube, that is, subjects from the no experience group manipulated this area more frequently than the ones from the prior experience group. Thirdly, overall subjects from both groups manipulated the bottom area less than the other two areas.

Moreover, the interaction between manipulation type and tube area was significant (LRT: $\chi^2 = 40.66$, $df = 4$, $p < 0.001$). Inspecting the interaction further revealed three significant main findings: Firstly, overall subjects manipulated the top area more frequently than the bottom one and the other area with all three manipulation types. Secondly, they manipulated the bottom area most often with the hand or mouth, followed by the pointed end, and least with the brush end of the tool. Third, they manipulated the bottom area less often in comparison to the other two areas with the brush end of the tool. Furthermore, older subjects manipulated the tube less than younger ones (LRT: $\chi^2 = 16.69$, $df = 1$, $p < 0.001$), but manipulation time did not differ in regard to sex (LRT: $\chi^2 = 0.15$, $df = 1$, $p = 0.703$). Overall, subjects manipulated the tube less frequently over sessions (LRT: $\chi^2 = 6.80$, $df = 1$, $p = 0.009$).

There was a significant effect of species, with chimpanzees manipulating the tube more often than bonobos and orang-utans (LRT: $\chi^2 = 7.08$, $df = 2$, $p = 0.029$). However, the factor species was estimated with the age effect given as I did not include an interaction between species and age in the model. Additionally, the duration of the session (included as an offset term) was overall shorter for chimpanzees than for the other two species. Thus, an occurrence (i.e., a “1”) weighed more and a non-occurrence (i.e., a “0”) weighed less for chimpanzees compared to bonobos and orang-utans. Moreover, bonobos’ age range was smaller than that of chimpanzees and orangs-utans.

Manipulation: Gaussian model

As a second step, I looked at the duration of manipulation for all combinations of tube area and manipulation type. The full-null-model comparison was significant (GLMM; LRT: $\chi^2 = 100.31$, $df = 22$, $p < 0.001$; Appendix C, Table C.1; Figure 5.3). I found a significant three-way-interaction of group, tube area and manipulation type (LRT: $\chi^2 = 11.73$, $df = 4$, $p = 0.019$), whereas none of the other predictors did (age, species, session, sex). Subjects from both groups manipulated the top area for a longer time than the additional two areas and longer with the brush end and the pointed end than by hand or mouth (Appendix C, Table C.3). However, the prior experience group still manipulated the top area longer than the no experience group. The no experience group manipulated the bottom area for a longer time than the prior experience group. When looking at the manipulation types, they especially manipulated the bottom area and the other area longer by hand or mouth than the prior experience group.

To compare both models, the binomial model revealed a significant two-way interaction between group and area, whereas the Gaussian model revealed a significant three-way interaction between group, area, and manipulation type. The findings of both models somewhat corroborate each other by showing a manipulatory focus on the top area of the tube with both ends of the tool by the prior experience group. Moreover, the no experience group explored the additional two areas of the tube more than the prior experience group (Figure 5.3).

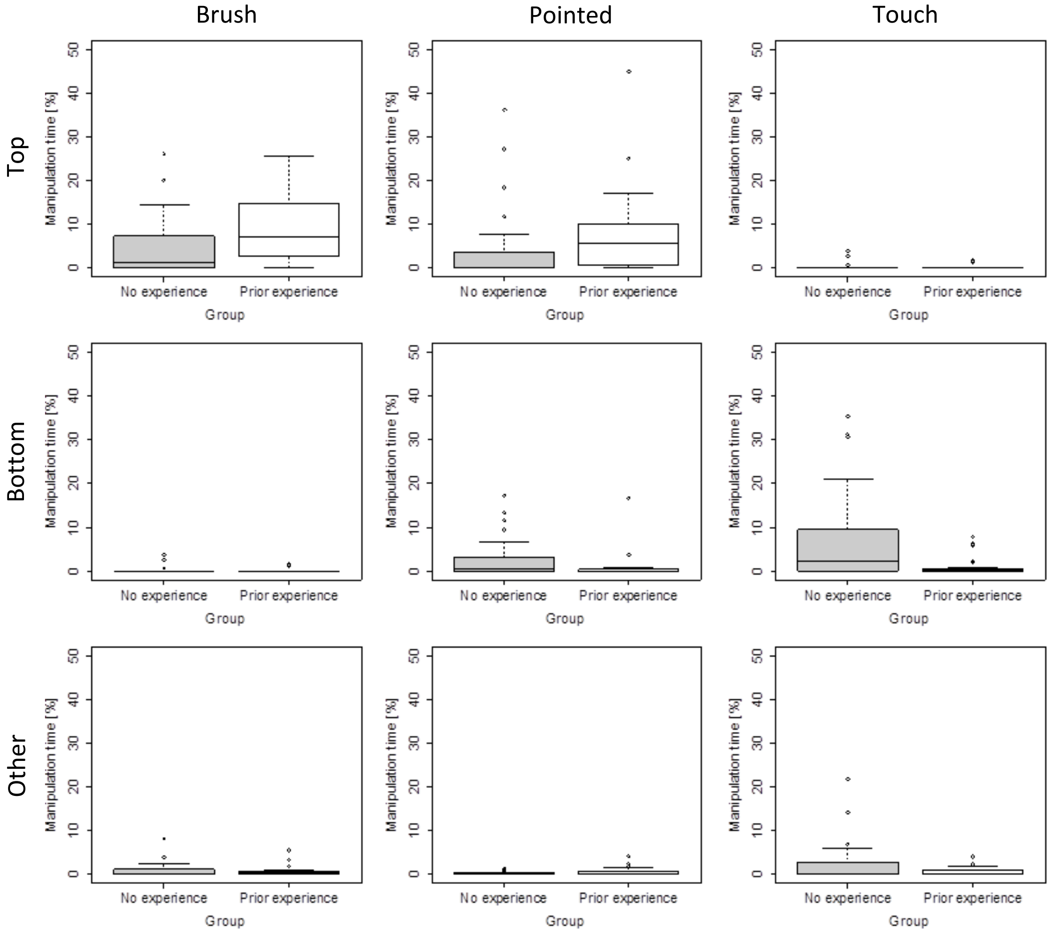


Figure 5.3 Results of Experiment 1: Manipulation time is shown as a function of group, tube area (top, bottom, other), and manipulation type (brush, pointed, touch; durations, based on all data).

3. Discussion

Great apes with prior experience manipulated the area of the apparatus that was relevant for the previous solution more than naïve individuals. The naïve individuals, in contrast, explored the whole apparatus more, especially with their hands and mouths. The task comprised of dipping juice with the brush end of a tool from the top area of the apparatus during the prior experience phase; during the test phase, they were required to use the pointed end of the same tool to puncture a tape covering a hole in the bottom area of the same apparatus. However, this difference in manipulation pattern between the two groups did not result in a difference in success, potentially due to an overall low success rate of 35%. I still found a trend of naïve subjects solving the task faster than experienced ones.

Experienced individuals did not direct significantly more manipulations with the brush end of the tool to the (blocked) top area of the apparatus in comparison to naïve individuals as I would have predicted based on a functional fixedness effect. However, the prior experience group manipulated the top area overall to a greater extent than the no experience group. This finding is consistent with an Einstellung effect or a mental set (Chown, 1959; Luchins & Luchins, 1959; Schultz & Searleman, 2002), that is, subjects perseverated on the former solution strategy in their tool use (Elsner & Schellhas, 2012; Smitsman & Cox, 2008). However, it does not indicate a functional fixedness effect (Adamson, 1952; Duncker, 1945; Flavell et al., 1958; German & Barrett, 2005; German & Defeyter, 2000; Legare & Nielsen, 2015; Maier, 1931).

It is conceivable, however, that subjects from the prior experience group were indeed fixated on the tool's function (i.e., being *for* dipping juice), but not specifically on the brush end of the tool: Firstly, they manipulated the top area significantly more with the pointed end than subjects from the no experience group. Secondly, eight out of eleven subjects dipped juice with the pointed end during the prior experience phase at least once so that they whole tool might have been assigned the function of dipping juice. Moreover, it is possible that subjects indeed became fixated on the brush end of the tool, but that they were flexible enough to integrate causal information in their judgements and understood that the brush end would not fit through the mesh, whereas the pointed end would be useful in breaking through the mesh.

Although I believe that it is generally possible that subjects became fixated on the tool's function (i.e., being *for* dipping) regardless of the functional end, I consider it more parsimonious that they perseverated on their former solution strategy, commonly referred to as the Einstellung effect. To disentangle both effects, one would have to employ two different tasks and/or give subjects the choice between at least two tools during the test. In addition, it is important to keep in mind that I was only able to find a trend for subjects from the no experience group performing better in the task. Therefore, the differences in manipulation style did not lead to a significant effect in the current experiment, potentially because the task was too difficult (only a 50% success rate in the control group).

Findings indicated a tentative age effect. Younger subjects tended to solve the task faster in addition to manipulating the apparatus significantly more frequently (but not longer) than older subjects. These findings suggest that younger subjects showed an increased exploration rate as indicated by their larger diversity of manipulation actions directed at the different parts of the apparatus. The greater diversity of manipulations might also explain why younger individuals tended to solve the task more effectively than older ones. A broad range of manipulations therefore seemed to be beneficial in this specific task. Irrespective of individual's age, apes had a clear preference on how to manipulate the top opening of the tube. How the age differences play into behavioural rigidity and conservatism will be a question for future research; large samples covering a wide age range will be needed to answer this question. Intriguingly, some recent studies suggest that younger apes outperform older ones in tasks that have a strong self-control component (Manrique & Call, 2015; Tennie, Call, et al., 2010; Völter & Call, 2014b).

Although naïve subjects spent a lot of time exploring all areas of the apparatus, half of them did not come up with the solution, indicating that this was a rather challenging task for these apes. Furthermore, my design potentially involved not only a fixation on the tool's function (i.e., functional fixedness effect; Duncker, 1945), but also involved a fixation on the task and its solution (i.e., Einstellung effect or mental set; Luchins & Luchins, 1959). Thus, even if a significant difference between groups had been found, a disentanglement of both types of fixation effects would not have been possible. In Experiment 2, I therefore administered two different tasks for the prior experience and the test phase. Moreover, I used a tool whose functional parts were not separated in space (Gruber et al., 2011; Gruber et al., 2009; Sanz et al., 2009) to avoid subjects from perceiving two distinct tools (corresponding to distinct functional parts). This tool consisted of a hose that could be used for both drinking (i.e., similar to a straw) and poking. These two functions relied on different physical properties of the tool (i.e., the hollowness and pliability of the tool).

C. Experiment 2

1. Methods

Subjects

Twenty-eight great apes (bonobos, *Pan paniscus*, $N = 4$; chimpanzees, *Pan troglodytes*, $N = 17$; orang-utans, *Pongo abelii*, $N = 7$; $N_{females} = 17$; age range: 7-42 years; mean age: 21 years; Table 5.2) participated in the study and of these, 22 had already participated in Experiment 1. Three apes (two bonobos and one chimpanzee) were excluded from the study because they did not fulfil the criterion for successful drinking during the prior experience phase. One additional orang-utan was excluded from the analysis because her 7-year-old offspring who was not separable during the test solved the task ($N_{final} = 24$).

Four apes (one chimpanzee and three orang-utans) already had experience using a straw to drink juice (Manrique & Call, 2011). I distributed these individuals equally among the two groups as I did in Experiment 1. It is likely that most of the apes have drunk juice through plastic tubes before during eye-tracking studies (e.g., Krupenye, Kano, Hirata, Call, & Tomasello, 2016). However, these tubes were attached to a Plexiglas panel and different in appearance from the hose used in the current study.

Table 5.2 Subjects participating in Experiment 2.

Subject	**Species**	**Sex**	**Age**	**Rearing**
Fimi	Bonobo	Female	8	Mother
Gemena[1]	Bonobo	Female	10	Mother
Luiza	Bonobo	Female	11	Mother
Kuno[1]	Bonobo	Male	19	Nursery
Tai	Chimp	Female	14	Mother
Bambari	Chimp	Female	16	Mother
Swela	Chimp	Female	20	Mother
Sandra	Chimp	Female	23	Mother
Hope	Chimp	Female	26	Mother
Daza	Chimp	Female	30	Unknown
Dorien[1]	Chimp	Female	35	Nursery
Riet	Chimp	Female	38	Nursery
Fraukje	Chimp	Female	40	Nursery
Frederike	Chimp	Female	42	Unknown
Bangolo	Chimp	Male	7	Mother
Kofi	Chimp	Male	11	Mother
Lobo	Chimp	Male	12	Mother
Alex	Chimp	Male	15	Nursery
Lome	Chimp	Male	15	Mother
Frodo	Chimp	Male	22	Mother
Robert	Chimp	Male	40	Nursery
Raja	Orang	Female	13	Mother
Padana	Orang	Female	18	Mother
Dokana[1]	Orang	Female	27	Mother

Pini	Orang	Female	28	Mother
Batak	Orang	Male	7	Mother
Suaq	Orang	Male	7	Mother
Bimbo	Orang	Male	36	Nursery

---

[1] were excluded from the study

Materials

For the pre-experience phase, I used a Plexiglas container (L 10 cm x W 10 cm x H 25 cm) with two holes: one at the front for inserting the hose (diameter: 4.5 cm) and one at the top for re-filling (diameter: 3 cm) that was filled with diluted grape juice (100 ml juice + 300 ml water). For the test phase, I used a horizontal Plexiglas tube (L 40 cm, outer diameter: 4 cm) with two blockages spread five cm apart to its openings (the tube was the same as in Experiment 1 in Völter & Call, 2012). In addition, I used a grey plastic hose (L 40 cm, diameter: 1.6 cm) that is usually used for cable protection as the tool and a wooden stick (L 40 cm, diameter: 1 cm) and a string (L 40 cm x W 0.7 cm x H 0.2 cm) as distractor objects (Figure 5.4).

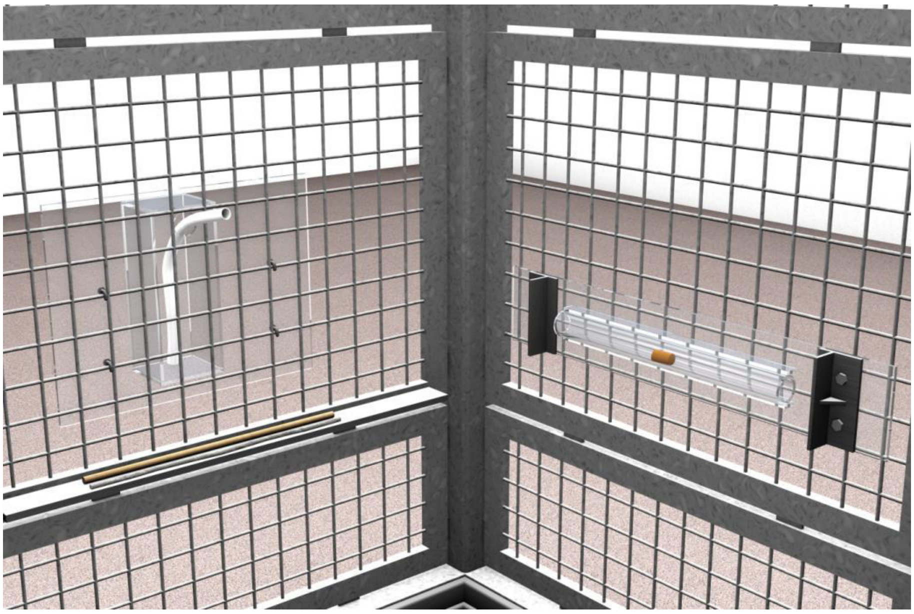


Figure 5.4 The setup of Experiment 2: In the prior experience phase, apes could drink juice from a drinking container with a hose; in the test phase, apes had to extract the hose from the drinking container to poke out a food reward from a horizontal tube which required a flexible tool.

Procedure

Apes were pseudo-randomly distributed into two groups with the restriction that I counterbalanced both groups for species, age, sex, and straw-tool experience as much as possible. Both groups were presented with a prior experience phase and a test phase as in Experiment 1. In the prior experience phase, the prior experience group (*N* = 10) learned to use the hose for drinking juice (5-12 sessions, mean: 7 sessions), whereas the no experience group (*N* = 14) was presented with the hose lying on the metal frame of a panel in the test room with the empty drinking container present (two sessions). The no experience group served to control for novelty of the tool (and of the drinking container). All sessions lasted five minutes each and one session was conducted per day; when subjects from the prior experience group were still drinking after five minutes, they were allowed to finish drinking the juice. In general, subjects from the prior experience group were presented with the tool lying on the metal frame in front of the drinking container (baseline). However, in the first session, the hose was pre-inserted into the drinking container to facilitate the task.

Additionally, I applied a three-step scaffolding procedure if apes were not drinking on their own, each for two sessions: 1) Subjects were presented with the hose pre-inserted into the container (as aforementioned); 2) subjects were presented with the hose attached to the container and a narrow plastic tube, which subjects knew from previous eye-tracking studies, was attached to the hose; 3) subjects received help by the experimenter, i.e., they were fed from the hose with juice, then, the hose was inserted into the container and they were given a chance (both was repeated multiple times). When subjects drank through the hose in the second or third condition, they received one session with a pre-inserted condition and then, baseline sessions thereafter (i.e., they had to insert the tool themselves in order to drink). The prior experience phase was completed when subjects had reached the criterion of five successful drinking sessions (pre-inserted and baseline condition). If apes did not drink from the hose on their own within the six shaping sessions, they were excluded from the sample (one chimpanzee and two bonobos). Additionally, one orangutan (Dokana) was excluded from the analyses because her offspring solved the task during the test.

In the test phase, subjects were presented with the empty drinking container with the hose pre-inserted and a wooden stick and a string of the same length lying in front of the container on the metal frame. In case of one chimpanzee (Fraukje), the hose was pre-inserted in the top hole of the container (that was used for re-filling the juice) because this subject had inserted and drunk juice from this hole before instead of using the hole at the front. A horizontal tube with two blockages close to its openings was attached to the mesh in a 90° angle to the container (Figure 5.4). Subjects were required to select the hose to use in poking out the banana pellet from the tube. Subjects received two test sessions lasting five minutes each. When they solved the task for the first time in the second session, they received a third one.

Coding and analyses

Overall success across all sessions, survival time until first success, latency until the first touch of the hose, and survival time until the first selection of the hose (i.e., extracting it from the container) was coded from the videos. I also scored the first tool use attempt at the tube (either with the hose, stick, or string) and whether any sucking attempts with the hose occurred across all sessions (similar to the drinking actions during the pre-experience phase). In regards to survival time until success, I considered the moment the hose came into contact with the pellet as "success" due to the fact that the pellet once got stuck due to the apparatus malfunctioning. A second coder scored 20% of all videos and both coders were in good agreement (Cohen's Kappa; success: $K = 1$, $N = 10$, $p = 0.002$; first tool use attempt: $K = 1$, $N = 10$, $p = 0.002$;

drinking attempt: $K = 1$, $N = 10$, $p = 0.002$; Pearson's correlation; survival time to success: $r = 1$, $df = 8$, $p < 0.001$; latency to touch hose: $r = 0.989$, $df = 8$, $p < 0.001$; survival time to select hose: $r = 1$, $df = 8$, $p < 0.001$).

The two groups were not balanced for species due to several drop-outs (bonobos: $N_{NoExp} = 2$, $N_{PriorExp} = 0$; orang-utans: $N_{NoExp} = 4$, $N_{PriorExp} = 2$). I therefore conducted all analyses once based on all data and once based on a balanced subset comprising chimpanzees only ($N_{NoExp} = 8$, $N_{PriorExp} = 8$). Analyses of overall success and survival time until first success were performed as in the previous experiment (see methods of Experiment 1). VIFs did not indicate collinearity to be a problem. Additionally, I analysed survival time until first selection of the hose with a survival model that was built and treated the same way as with survival time until first success as the response. I aimed to analyse latency until first touching the hose with a linear model but these model assumptions were not fulfilled. Instead, I performed a Mann-Whitney-U test to compare the latencies between the two groups. Moreover, I used Fisher's exact test to analyse drinking attempts (yes/no) and first tool use attempts (hose – yes/no).

2. Results

Success and survival time

The GLM with success as the response was significant when compared to the null model (LRT: $\chi^2 = 8.73$, $df = 2$, $p = 0.013$). Members of the no experience group solved the task significantly more often than the ones of the prior experience group (LRT: $\chi^2 = 8.48$, $df = 1$, $p = 0.004$; Figure 5.5A), whereas age did not influence performance (LRT: $\chi^2 = 0.57$, $df = 1$, $p = 0.451$). In fact, 13 out of 14 subjects from the no experience but only four out of ten subjects from the prior experience group solved the task in the first session. Since I found large standard errors due to overall few subjects not solving the task, I additionally performed a Fisher’s exact test for success which corroborated my finding (Fisher's exact test: $p = 0.020$).

The full model with survival time as the response fitted the data significantly better when compared to the null model (LRT: $\chi^2 = 8.18$, $df = 2$, $p = 0.017$). Members of the no experience group solved the task significantly faster than the ones of the prior experience group (LRT: $\chi^2 = 8.06$, $df = 1$, $p = 0.005$; Figure 5.5B), whereas age did not significantly affect performance (LRT: $\chi^2 = 0.50$, $df = 1$, $p = 0.481$). I found the same pattern when looking at the subset of chimpanzees (survival model; LRT comparing full and null model: $\chi^2 = 6.63$, $df = 2$, $p = 0.036$; LRT for group: $\chi^2 = 5.88$, $df = 1$, $p = 0.015$; LRT for age: $\chi^2 = 1.45$, $df = 1$, $p = 0.228$).

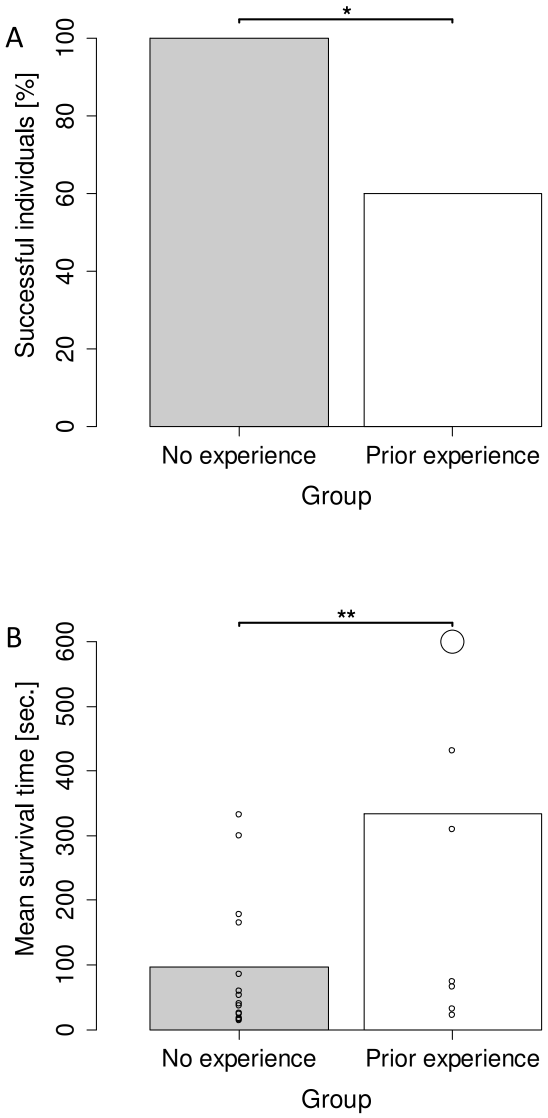


Figure 5.5 Results of Experiment 2: Success (A) and survival time (B) as a function of group. Circles indicate individual performance with larger circles referring to more individuals. Significance level: ** $p < 0.01$, * $p < 0.05$

Behavioural measurements

Subjects from both groups did not differ in their latency until their first contact with the hose in the test (Mann-Whitney U test: $U = 45$, $N_{NoExp} = 14$, $N_{PriorExp} = 10$, $p = 0.15$; Figure 5.6A). I found a slightly different pattern when looking at the subset of chimpanzees: Subjects from the prior experience group tended touch the hose faster than those from the no experience group, potentially because they tried to drink from it (Mann-Whitney U test: $U = 14$, $N_{NoExp} = 8$, $N_{PriorExp} = 8$, $p = 0.060$).

Moreover, the model with survival time until hose selection (i.e., extracting the hose from the container) as the response was significant when compared to the null model (LRT: $\chi^2 = 8.53$, $df = 2$, $p = 0.014$). Subjects from the no experience group selected the hose significantly faster than those from the prior experience group (LRT: $\chi^2 = 4.56$, $df = 1$, $p = 0.033$; Figure 5.6B) and younger subjects selected the hose faster than older subjects (LRT: $\chi^2 = 4.26$, $df = 1$, $p = 0.039$). Interestingly, four subjects (including three chimpanzees) from the prior experience group did not extract the hose at all from the container. I found the same pattern to be true when looking at the subset of chimpanzees (survival model; LRT comparing full and null model: $\chi^2 = 12.58$, $df = 2$, $p = 0.002$; LRT for group: $\chi^2 = 7.00$, $df = 1$, $p = 0.008$; LRT for age: $\chi^2 = 8.15$, $df = 1$, $p = 0.004$).

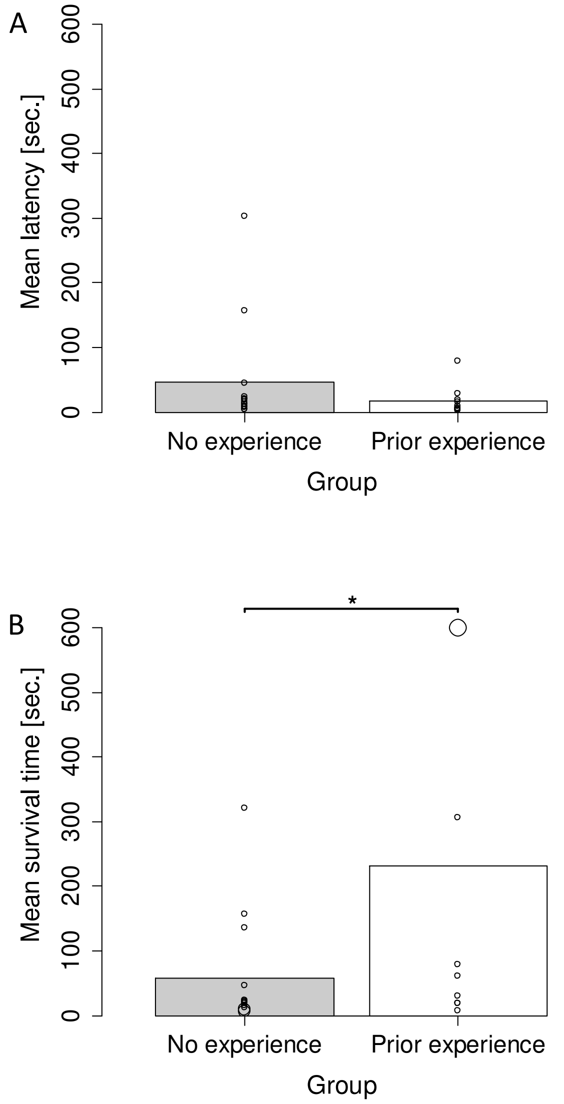


Figure 5.6 Results of Experiment 2: Latency to touch hose (A) and survival time to hose selection (B) as a function of group in Experiment 2. Circles indicate individual performance with larger circles referring to more individuals. Significance level: * $p < 0.05$

Moreover, I investigated whether subjects from both groups differed in regard to how often they chose the hose as the first tool to manipulate the tube. In doing so, I did not find a difference, neither between the two groups (Fisher's exact test: *p* = 0.358; NoExp – 4:10, PriorExp – 1:9; hose as first tool – yes:no), nor for the chimpanzees (Fisher's exact test: *p* = 0.569; NoExp – 3:5, PriorExp – 1:7). Subjects from both groups seemed to have a preference for sticks as their choice of first tool use attempt (NoExp – hose: 29%, stick: 64%, string: 7%, none: 0%; PriorExp – hose: 10%, stick: 70%, string: 10%, none: 10%).

In addition, I analysed drinking attempts and found a significant difference between both groups, that is, the prior experience group performed significantly more sucking actions with the hose inserted in the empty container (Fisher's exact test: *p* = 0.010; NoExp – 2:12, PriorExp – 7:3; drinking attempt – yes:no). Further evidence that these sucking actions reflected drinking attempts was evident, in most cases, with the subjects moving their cheeks and throats while sucking at the hose, which was similar to the actions while drinking and swallowing juice. Some subjects sucked at the hose several times throughout one session and one subject even re-inserted the hose into the container and tried to drink thereafter.

3. Discussion

In line with my predictions for a functional fixedness effect, after drinking juice from a container with a hose, apes were less likely to use the hose for poking out a food reward from a tube thereafter compared to a control group. Moreover, these individuals with prior with the straw function of the hose often tried to drink from the empty hose and took longer to extract it from the container than the control group. Contrary to this, at least chimpanzees with prior experience tended to touch the hose faster than the control group which indicates that the tool was more salient to them than for the control group. Together these findings suggest that subjects remembered the previous function (i.e., drinking) in the novel problem solving context and that this knowledge hindered them in finding the novel solution.

This second experiment corroborates the findings of the first experiment and shows that a change of manipulation style due to prior experience can also lead to a difference in success, that is, a decrease in problem-solving abilities. This study supports the notion that prior experience is a two-sided coin: It can facilitate future problem solving when encountering similar problems, but it can also have detrimental effects by blocking novel solution strategies. In the third experiment, I aimed at extending this finding by creating a situation in which apes are required to switch flexibly between different tool functions to solve a problem. I therefore employed a food item as a tool with an intrinsic function that is easy to discover, available at all times and, thus, is hard to disregard. More specifically, I investigated if prior experience with a novel food would modulate the apes’ tool-use behaviour with it.

D. Experiment 3

1. Methods

Subjects

Thirty-seven great apes participated in the study (bonobos, *Pan paniscus*, $N = 7$; chimpanzees, *Pan troglodytes*, $N = 19$; gorillas, *Gorilla gorilla*, $N = 3$; Sumatran orang-utans, *Pongo abelii*, $N = 8$; $N_{females} = 24$; age range: 7-50; mean age: 21 years; Table 5.3). One orang-utan (Raja) was excluded from the study because she stopped eating the grissini during the pre-feeding phase. One gorilla (Abeeku) was further excluded from the analyses because he did not manage to rake in the grapes with a wooden stick during the control test ($N_{final} = 35$).

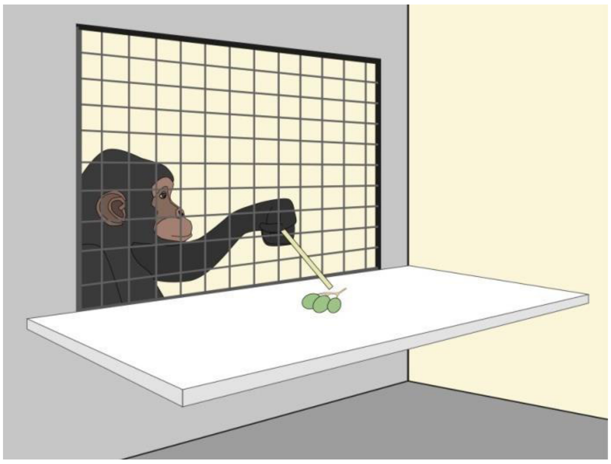


Figure 5.7 The setup of Experiment 3: Apes could rake in grapes with a grissini.

Table 5.3 Subjects participating in Experiment 3.

Subject	Species	Sex	Age	Raising
Fimi	Bonobo	Female	8	Mother
Gemena	Bonobo	Female	11	Mother
Luiza	Bonobo	Female	11	Mother
Lexi	Bonobo	Female	17	Nursery
Yasa	Bonobo	Female	19	Mother
Kuno	Bonobo	Male	19	Nursery
Jasongo	Bonobo	Male	26	Mother
Tai	Chimp	Female	14	Mother
Bambari	Chimp	Female	15	Mother
Swela	Chimp	Female	21	Mother
Sandra	Chimp	Female	23	Mother
Hope	Chimp	Female	25	Mother
Daza	Chimp	Female	30	Unknown
Dorien	Chimp	Female	36	Nursery
Natascha	Chimp	Female	36	Nursery
Riet	Chimp	Female	39	Nursery
Fraukje	Chimp	Female	40	Nursery
Frederike	Chimp	Female	42	Unknown
Jeudi	Chimp	Female	50	Unknown
Bangolo	Chimp	Male	7	Mother
Kofi	Chimp	Male	11	Mother
Lobo	Chimp	Male	12	Mother
Lome	Chimp	Male	15	Mother
Alex	Chimp	Male	15	Nursery
Frodo	Chimp	Male	23	Mother

Robert	Chimp	Male	41	Nursery
Kumili	Gorilla	Female	13	Mother
Kibara	Gorilla	Female	13	Mother
Abeeku[1]	Gorilla	Male	17	Mother
Tanah	Orang	Female	7	Mother
Raja[1]	Orang	Female	13	Mother
Padana	Orang	Female	19	Mother
Dokana	Orang	Female	28	Mother
Pini	Orang	Female	28	Mother
Batak	Orang	Male	7	Mother
Suaq	Orang	Male	7	Mother
Bimbo	Orang	Male	36	Nursery

[1] were excluded from the study

Materials

For the study, a table (L 80 cm x W 40 cm x H 55 cm), grissini (i.e., long, hard bread sticks; length: about 20 cm, diameter: about 0.6 cm), wooden sticks (length: 20 cm, diameter: 0.5 cm), a cluster of three grapes (i.e., three grapes that were connected), and wooden blocks (5 cm x 2.4 cm x 1.8 cm) were used. To my knowledge, the grissini were a novel food item for bonobos, gorillas, and orang-utans, whereas many of the chimpanzees had been tested with grissini about five years before (Karg, Schmelz, Call, & Tomasello, 2016). I consider the impact of this prior experience on my findings in the discussion. Additionally, most apes had previously received pretzel sticks, which may be regarded as somewhat similar in appearance and taste.

Therefore, I made certain that there was a gap of at least two months between the last test with pretzel sticks and the current study with the grissini (which was only relevant for gorillas and orang-utans).

Procedure

Apes were distributed into three groups. One group was pre-fed with whole grissini, one pre-fed with pieces of grissini and one group was not fed before the test. The two pre-feeding groups received a total of nine grissini with three grissini per day. The grissini were broken into three pieces of equal length out of sight of the subjects. After the pre-feeding, apes received a total of six sessions with one session being conducted per day. First, apes received two test sessions with three trials each in which they could use the grissini as a rake tool (Figure 5.7). After measuring apes' finger length by placing a small piece of food (such as a raisin, or a piece of pellet) out of their reach, the experimenter placed clusters of grapes on a table about two cm out of the apes' reach. After five seconds, she gave a grissini to the subject, which could be used to rake in the grapes into reach. After the subject had eaten the grissini, the experimenter removed the grapes (if needed) and started the next trial. The third session conducted comprised of two control tests with three trials each. Apes could either rake in the cluster of grapes with a wooden stick or a non-edible object (a wooden block) with a grissini, with the procedure remaining the same as in the original test. The trials were presented in a randomized order with a maximum of two trials of the same type presented consecutively.

The last three sessions consisted of a food preference test consisting of six trials each (plus one additional trial) in which subjects could choose between two options: three, four or five grapes versus one grissini or one wooden stick. Apes received each of the six combinations once per session, resulting in a total of 18 trials. To minimize carry over effects (i.e., apes assuming that they could access both options by choosing the tool first and using it to rake in the grapes) I used a Plexiglas panel (test: mesh panel) and the food was placed onto two white plates (test: no plates). The experimenter baited both plates and pushed the sliding table toward the subjects who could choose a plate by touching one of them. If a choice was unclear (e.g., both plates were touched at the same time), the trial was repeated. At the end of each food preference test session, I conducted an additional trial in which I placed a grissini and a cluster of grapes on the floor close to each other (counterbalanced for side) in an adjacent room. The subjects were then allowed to enter the room and eat the food items so that the order of consumption in the absence of a task could be assessed. All sessions were videotaped.

Coding and analyses

The following behaviours were measured in the test trials: success (accessing the grapes), attempt (touching the grapes with the grissini / wooden stick, but not accessing them), rake hesitation (sticking the grissini / stick through the mesh without touching the grapes), exploration with mouth (putting the grissini / wooden stick) in contact with one's mouth *before* the first tool use event in a trial) and food item eaten first in case of success (grissini or grapes). For the main analyses I collapsed success and attempts into a "tool use" (yes/no) variable as a result of the grissini sometimes breaking while trying to rake in the grapes or due to an individual's clumsiness. For the two control tests, I scored tool use in the same way. Moreover, I scored the total number of grapes across trials that apes had sacrificed for a grissini in the food preference test, which indicated their relative preference for grissini over grapes, with a higher score showing a higher preference for grissini (score: 0-36). A second coder coded 20% of the videos and both coders were in good agreement (Cohen's Kappa; tool use: $K = 1$, $N = 43$, $p < 0.001$; tool taken by hand or mouth: $K = 0.93$, $N = 43$, $p < 0.001$; exploration with mouth before first rake attempt: $K = 1$, $N = 22$, $p < 0.001$; food item eaten first: $K = 1$, $N = 18$, $p < 0.001$).

Two analyses were conducted, one for the first trial and one comprising of all six trials. I expected larger differences in performance between groups in the first trial seeing as all subjects (including the ones assigned to the no grissini pre-feeding group) experienced that the grissini was edible during the first trial. I performed a Generalized Linear Model (GLM) with a binomial error structure with tool use (yes/no) in the first trial as the response ($N = 35$). The model included group, species and grissini preference as fixed effects. The grissini preference was z-transformed to a mean of zero and a standard deviation of one. I derived Variance Inflation Factors (VIFs) to check for collinearity using the function vif of the R-package car (Field, 2005; Fox & Weisberg, 2011) which were adequate. I established the overall effect of the predictors by comparing the full with a null model including only the intercept using a likelihood ratio test (LRT; R function anova with argument test set to "Chisq";

Schielzeth & Forstmeier, 2009). I also performed LRTs comparing the full with respective reduced models to get p-values for the individual predictors (R function drop1; Barr et al., 2013). I explored significant predictors further by re-levelling the respective factors.

Additionally, I performed a Generalized Linear Mixed Model (GLMM) with a binomial error structure with tool use (yes/no) as the response across all trials (N = 210). The model included group, species, age, grissini preference, and trial as fixed effects as well as the random slope of subject within trial (Barr et al., 2013). Age was log-transformed and age, trial and grissini preference were z-transformed to a mean of zero and a standard deviation of one. VIFs were established as before and did not indicate collinearity to be a problem. I evaluated model stability by comparing the estimates from a model based on all data versus several models based on data with one level of the random effect taken out at a time. Model stability was acceptable except for the level “gorilla” of the predictor “species”, indicating some instability when one of the two gorillas was taken out. This was not surprising due to the small sample size (N = 2). I first performed a full-null model comparison and then established p-values for the individual predictors the same way as in the previous model. In order to better understand the apes’ behaviour, I also conducted several models to analyse additional behavioural measurements.

2. Results

First trial analysis

First, tool use in the first trial was analysed in order to have a clean measurement of my manipulation of the variable group. Overall, the full model (GLM) was significant when compared to a null model (LRT: $\chi^2 = 19.37$, $df = 6$, $p = 0.004$; Appendix C, Table C.4). More specifically, I found a significant effect of group (LRT: $\chi^2 = 9.57$, $df = 2$, $p = 0.008$), showing that apes who were naïve to the grissini were more likely to use it as a tool than apes who had been pre-fed with either whole grissini or grissini pieces (Figure 5.8). Furthermore, I found a significant effect of species ($\chi^2 = 9.66$, $df = 3$, $p = 0.022$) with bonobos outperforming chimpanzees. Orang-utans also tended to perform better than chimpanzees (bonobos: 86%; chimpanzees: 32%; gorillas: 50%; orang-utans: 71%; trials with tool use). Individual preferences for grissini and grapes did not influence the response significantly ($\chi^2 = 1.66$, $df = 1$, $p = 0.197$).

The confidence intervals of the model indicated some uncertainty. Therefore, I conducted non-parametric analyses to confirm my findings for the main predictors group and species. The tests revealed a significant difference between groups (χ^2- homogeneity test: $\chi^2 = 7.92$, $df = 2$, $p = 0.019$) as well as between species (Fisher's exact test: $p = 0.042$), corroborating the results of the model.

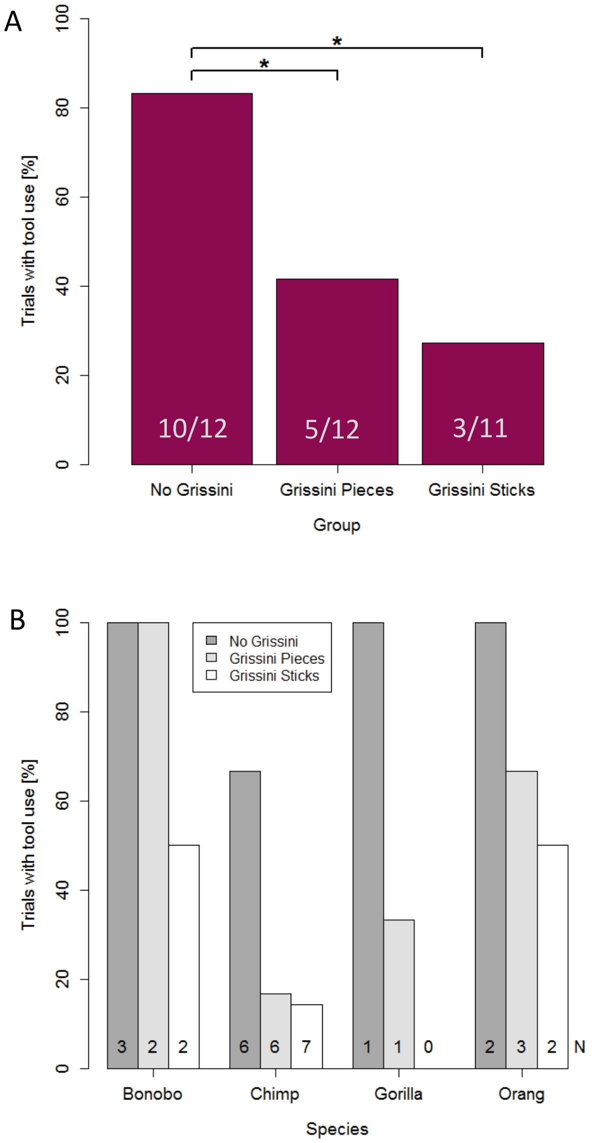
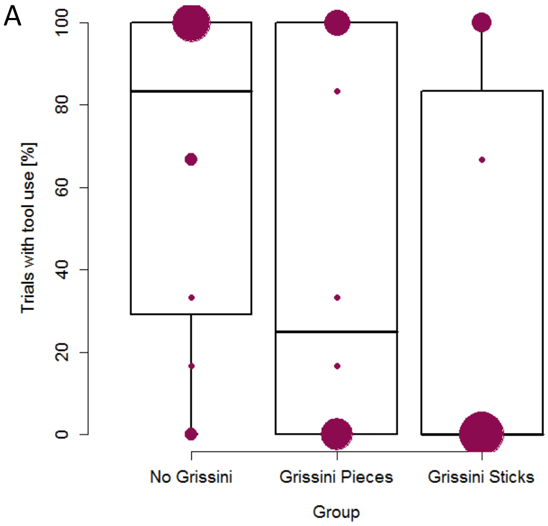


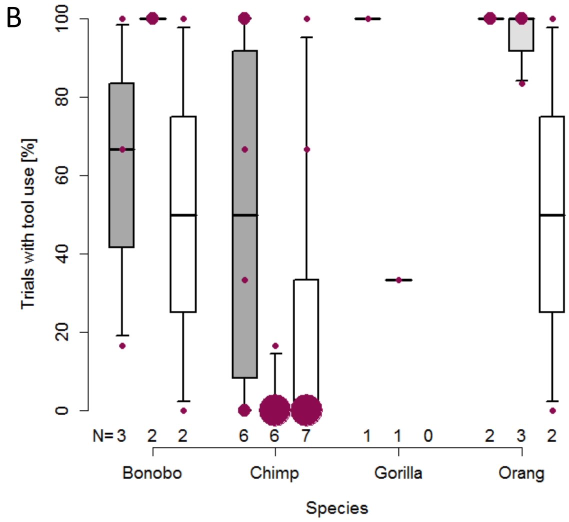
Figure 5.8 Results of Experiment 3: Tool use in the first trial as a function of group (A) and as a function of group and species (B).

All trials

As a next step, I analysed tool use for all trials and found the model (GLMM) to be significant when compared to the null model (LRT: $\chi^2 = 22.31$, *df* = 7, *p* = 0.002; Appendix C, Table C.5). While group did not have a significant effect on the response (LRT: $\chi^2 = 0.74$, *df* = 2, *p* = 0.690; Figure 5.9A), species did ($\chi^2 = 14.66$, *df* = 3, *p* = 0.002; Figure 5.9B). More specifically, bonobos and orang-utans outperformed chimpanzees (bonobos: 70±43%; chimpanzees: 25±39%; gorillas: 67±47%; orang-utans: 83±37%; trials with tool use, mean±SD; Figure 4). There was not a significant influence of individuals' preference for grissini ($\chi^2 = 0.27$, *df* = 1, *p* = 0.603) or of trial ($\chi^2 = 0.38$, *df* = 1, *p* = 0.540) on the response. Apes also did not seem to differ in their preference for grissini on the species level (bonobos: 12±13; chimpanzees: 12±7; gorillas: 16±22; orang-utans: 12±10; preference for grissini, score 0:36, mean±SD; Figure 5.9C).

Confidence intervals for this model could not be established as a result of several models not converging during the boot-strapping process due to a problem of complete separation (i.e., many subjects either always used the grissini as a tool or never did). I therefore conducted non-parametric tests again to confirm my findings for my main predictors group and species. While the test revealed no significant difference between the groups (Kruskal-Wallis H-test: $\chi^2 = 3.40$, *df* = 2, *p* = 0.183), I found a significant difference between species (Kruskal-Wallis H-test: $\chi^2 = 10.68$, *df* = 3, *p* = 0.013), thus confirming my previous findings.


A
Trials with tool use [%]
0
20
40
60
80
100
No Grissini
Grissini Pieces
Grissini Sticks
Group


B
Trials with tool use [%]
0
20
40
60
80
100
N= 3
2
2
6
6
7
1
1
0
2
3
2
Bonobo
Chimp
Gorilla
Orang
Species

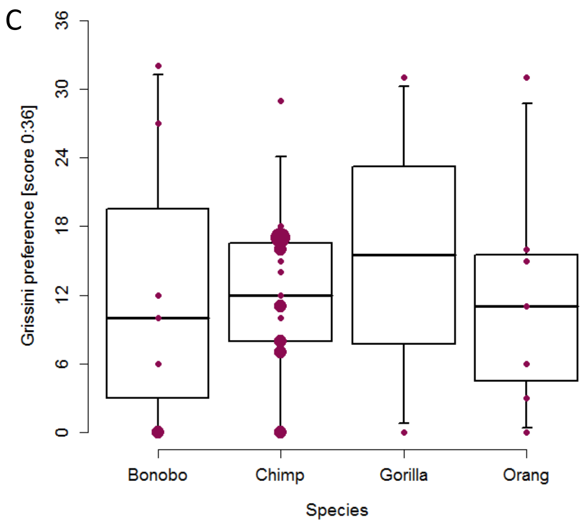


Figure 5.9 Results of Experiment 3: Tool use across all trials as a function of group (A) and as a function of group and species (B); and grissini preference as a function of species (C). Circles indicate individual data points with larger circles referring to more individuals. Groups in (B): No grissini (dark grey), grissini pieces (light grey), grissini sticks (white).

Furthermore, I found that naïve apes explored the grissini more often with their mouths – sometimes biting off a piece – before the first tool use event in a trial than apes that had been pre-fed whole grissini (GLMM; LRT comparing full and null model: $\chi^2 = 22.46$, *df* = 7, *p* = 0.002; LRT for group, $\chi^2 = 8.30$, *df* = 2, *p* = 0.016; no grissini: 74%, grissini pieces: 44%, whole grissini: 36%; trials with tool use only). Additionally, bonobos did so more often than gorillas and chimpanzees in turn did so more often than gorillas and orang-utans (LRT for species; $\chi^2 = 11.21$, *df* = 3, *p* = 0.011; bonobo: 72%, chimpanzee: 76%, gorilla: 13%, orang-utan: 37%). I found a similar pattern in trial

1, even though it was not significant (Fisher's Exact Test: $p = 0.118$; no grissini: 90%, grissini pieces: 60%, whole grissini: 33%; however, the two pre-feeding groups were summarized because of the few tool use occurrences). There was also no significant difference in trial 1 with regard to species (Fisher's Exact Test: $p = 0.621$; bonobo: 83%, chimpanzee: 83%, gorilla: 0%, orang-utan: 60%; the one successful gorilla was excluded from this analysis). On some occasions apes already took the grissini from the experimenter with their mouths instead of using their hands (tool taken by mouth in trials with tool use: 20%; in trials without tool use: 32%; all trials – bonobos: 29%, chimpanzees: 30%, gorillas: 0%, orang-utans: 21%).

After raking in the grapes, apes often did not consume the grissini immediately. Instead they held it with their hands or feet, or left it between the thigh and body while first consuming the grapes (88% of trials). However, apes also consumed the grapes first (90% of trials) when there was no raking task, potentially because grapes were easier to consume than the grissini. Yet, grissini were readily eaten in all trials with one exception: One gorilla (Kumili) did not eat the grissini in her first four trials, potentially because she did not realize that it was food (although her offspring sometimes picked it up and ate it). I also coded hesitation behaviour (i.e., sticking the grissini through the mesh, but not far enough to touch the grapes either because retracting it before or because it was already too short), which occurred ten times (no grissini: 6%, grissini pieces: 6%, grissini sticks: 3%; trials with hesitation). Of these, three events occurred in the first trial (grissini pieces: 2x, grissini sticks: 1x).

Follow-up tests: Raking abilities and food preference

I investigated apes’ general raking abilities in two control tests to make sure that they had the prerequisites skills necessary to solve the task. Firstly, I found that all apes were able to use a wooden stick as a tool to rake in grapes (98% of trials) except for one gorilla (Abeeku) who was excluded from the analyses. Secondly, I also presented apes with trials in which they could rake in a wooden block with a grissini. Apes used the grissini less often as a tool with this non-edible wooden object than they did with the grapes (wooden block: 21%, grapes: 48%; percentage of trials) even though there might have been a carry-over from the main experiment (potentially explaining the high raking rates with the non-edible object). Interestingly, one chimpanzee (Frederike) who had not used the grissini as a tool with the grapes did so with the wooden block. This potentially again indicates a carry-over effect seeing that the trials of the two control tests (i.e., raking in grapes with a wooden stick and raking in a piece of wood with a grissini) were intermixed.

In the food preference test, most individuals did not show an absolute preference for grapes or grissini and sacrificed at least some grapes for a grissini, but they rarely did so for wooden sticks (grissini: 12±9, wooden sticks: 2±3; score 0-36, mean±sd). In the rare cases in which they chose the wooden stick over the grapes, they often tried to rake in the grapes while the experimenter was pulling the sliding table away, suggesting a carry-over effect from the test and not a preference for wooden sticks over grapes.

3. Discussion

Members of all tested great ape species successfully used a food item as a tool. More precisely, apes used a grissini stick to rake in grapes that they could not have reached otherwise. Moreover, they were more likely to use the grissini as a tool in the first trial when they had not experienced it as food before, providing it to be consistent with a functional fixedness effect. Intriguingly, the majority of naïve apes (90%) explored the food item with their mouths before employing it as a tool, suggesting that they were aware that the tool was edible. The findings suggest that great apes are capable of forming two representations of one object and can become fixated on one of these representations by experience.

The pre-feeding experience only affected apes' first trial experience. Two factors contributed to the transient nature of this manipulation: Across trials the initially naïve apes also gained experience with the object as food, and some of the apes from the pre-feeding groups started using the grissini as a tool. Individual preferences for grapes over grissini did not predict success. Bonobos and orang-utans showed more tool use with the food item than chimpanzees, either indicating a difference between species or one in experience. Many of the chimpanzees had received grissini as a reward in a study five years before.

Naïve apes were more likely to use a food item as a tool in the first trial than apes who had been pre-fed with it, a finding consistent with a functional fixation of the object (e.g., Defeyter & German, 2003; Duncker, 1945; Flavell et al., 1958; German & Defeyter, 2000; Hanus et al., 2011). One interpretation may be that experienced apes had already formed a representation of the food item as being *for* eating that hindered them employing it as a tool. The object comprised two functions: The disposition to nourish oneself (i.e., it was *for* eating) and the disposition to extend one's reach (i.e., it was *for* raking) that were available at the same time. I discuss the implications of tool affordances furher in the general discussion.

The four great ape species were not equally likely to use the food item as a tool. More precisely, bonobos and orang-utans outperformed chimpanzees. I cannot draw conclusions about gorillas because only two subjects were involved, but at minimum both of them used the food item as a tool eventually. One could interpret the species difference in three possible ways. First, chimpanzees potentially preferred eating the grissini over the grapes than bonobos and orang-utans; thus, the task was more difficult for them being as they needed more inhibitory control to not eat their preferred food item immediately. However, individual food preferences for grissini over grapes did not predict success and the mean preference for grissini was about the same for all species. Secondly, the finding could possibly reflect differences in temperament that modulated performance (such as general activity level, emotional arousal or inhibitory control). I did indeed observe that bonobos and orang-utans often used the tool in a calm manner taking their time whereas many chimpanzees produced food grunts and/or acted in a hectic way. Differences in inhibitory control could thus play a role, although a recent study suggests comparable inhibitory skills in bonobos, chimpanzees, and orang-utans (Amici, Aureli, & Call, 2008; but see Vlamings, Hare, & Call, 2010). Thirdly, chimpanzees might have formed a stronger representation of grissini being food because it was not completely novel as food for most of them.

Chimpanzees had gained experience with grissini as rewards (but not as tools) in a previous study that was conducted about five years before (Karg et al., 2016). However, to my knowledge, the grissini was a novel type of food for the other species. Intriguingly, “naïve” chimpanzees in the current study still outperformed chimpanzees that were pre-fed (trial 1 – no grissini: 67%, grissini pieces: 17%, whole grissini: 14%; all trials – no grissini: 50%, grissini pieces: 3%, whole grissini: 24%; percent of trials with tool use). This finding suggests that chimpanzees’ experience with the food item in a previous study might have reduced the overall occurrence of tool use, but the pre-feeding still influenced their behaviour.

Therefore, I consider it possible that prior experience with the food item some years before has potentially caused a decreased performance in chimpanzees compared to bonobos and orang-utans, but to be certain one would have to test a naïve population of chimpanzees.

E. General discussion

The impact of prior experience on great ape tool use was investigated in the course of three experiments. In Experiment 1, prior experience with a tool altered subjects' subsequent manipulation pattern of the apparatus. More precisely, subjects who had dipped juice with the brush end of a tool from the top area of the apparatus, focused on this solution area even when it was blocked, whereas their naïve counterparts explored the whole apparatus, including the area relevant for the new solution. Although prior experience modulated where apes targeted the apparatus, they did not specifically use the former functional tool end (i.e., the brush end) for manipulation, potentially reflecting an Einstellung effect. In Experiment 2, prior experience with a tool had a negative impact on apes' problem-solving performance. More specifically, subjects who had used a hose for drinking juice before solved a task that required using the hose to poke out a food reward from a tube less often than naïve individuals. Additionally, they were slower in finding the solution, took longer to extract the hose from the drinking container, and tried to drink from the hose more frequently than the naïve control group. In Experiment 3, apes used a grissini to rake in grapes dependent on their experience with it. Naïve apes were more likely to use the food item as a tool in the first trial than experienced ones who had been pre-fed with grissini. Together these findings indicate that both a previous experience with a tool function and with features of the problem situation (i.e., the apparatus) can induce fixation effects in great apes.

My predictions with regard to a decreased problem-solving performance in the experimental group compared to the control group in Experiment 1 were not met. However, only half of the individuals from the control group solved the task which means that it had been too difficult for the apes. I expected that most of them would solve the task since they did not receive any prior experience with the tool that potentially influenced their performance. I further expected the apes from the experimental group in Experiment 1 to preferentially manipulate the top opening of the apparatus with the brush end of the tool. Yet, they showed a preference for this apparatus area, but not for the tool end. As expected, apes exhibited a functional fixedness effect in Experiment 2 when the task was easier to solve than the previous one (i.e., all individuals from the control group solved the task). We found a difference between both groups in all three measurements that are commonly used in the literature (success, time until success and time until target tool selection). Functional fixedness is often broadly described as a decrease in one's problem-solving performance due to the fixation on a previous function of a tool. Yet, some authors consider tool selection as the best measurement because it focuses on the moment when an individual considers the tool for the novel function and less on the execution of the task which may vary in time between individuals (Defeyter & German, 2003; German & Barrett, 2005). Finally, I expected that some apes would be able use a food item as a tool and that they would be more likely to do so if they had not experienced it as food before. Results validated my predictions, however, the effect vanished over trials, suggesting that the functional fixedness effect disappears quickly with food items in great apes.

Recent studies have suggested that great apes are conservative tool users in the sense that they tend to stick with a problem-solving strategy as long as it is working (Gruber et al., 2011; Gruber et al., 2009; Hanus et al., 2011; Hrubesch et al., 2009; Manrique, Völter, & Call, 2013; Marshall-Pescini & Whiten, 2008). The current study corroborates these findings by showing that captive great apes become fixated on the previous function of a tool which either hindered success or delayed the solution to a novel problem. Apes also performed the familiar action with the tool although it was non-functional in the novel context (i.e., they tried to drink from the hose although there was clearly no juice available).

Chimpanzees tended to touch the hose faster when they had prior experience with it in a functional context compared to naïve ones. Nevertheless, subjects with prior experience took significantly longer to extract the hose from the drinking container which suggests that prior experience changed their perception of the tool: Although they tended to notice it faster than the control group, they did not consider it for the novel task and this is exactly what the functional fixedness effect is about (Adamson, 1952; Birch & Rabinowitz, 1951; Defeyter & German, 2003; Dominowski & Dallob, 1995; Duncker, 1945; Flavell et al., 1958; German & Barrett, 2005; German & Defeyter, 2000; Glucksberg, 1964; Glucksberg & Weisberg, 1966; Hanus et al., 2011; Maier, 1931; Yonge, 1966).

However, this evidence does not prove that apes understood that the hose *was for* drinking (i.e., assigning the function to the tool itself that constitutes its disposition). Instead, they may have learned that the hose *is used for* drinking (i.e., exhibiting a fixation on the *action* performed with the tool). It would be a matter of debate if it was necessary to differentiate between the notion “tool X is for drinking” and “tool X is (or can be) used for drinking” which potentially reflects a difference semantic knowledge (or concepts) and procedural knowledge. As Hernik and Csibra (2009) point out humans consider tools having a specific functions even if they are currently not in use. This conceptualization of tools is likely to be amplified by norms in humans which may lead to a different type of object representation in humans (i.e., all artefacts are designed for specific purposes) compared to non-human animals and to a stronger or even a different type of fixation effect (Csibra & Gergely, 2009; Hernik & Csibra, 2009; Ruiz & Santos, 2013; Vaesen, 2012). I discuss the social embeddedness of objects in humans and great apes further in Chapter 6.

It also remained unclear from Experiment 2 if apes were aware of the two functions of the tool simultaneously or sequentially. It appears unlikely that apes had forgotten the previous function, given the evidence for apes’ excellent long term memory for distinct past events (e.g., Janmaat et al., 2013; Kano & Hirata, 2015; Lewis et al., 2017; Martin-Ordas et al., 2013). For Experiment 3 I was seeking a tool that could be seen with two functions at the same time (or at least very fast sequentially as it remains a general question if one could have two representations simultaneously). Thus, I used an object that had an intuitive function (i.e., being *for* eating) that could be easily discovered, that is available when facing the raking problem and thus, could be hardly overlooked in Experiment 3. Moreover, I used a task with clear task affordances. All apes were familiar with raking tasks and using a stick as a rake tool seems to be a widely spread tool use found in captive great apes (Herrmann et al.,

2007) and for the apes in the current study in particular (Manrique et al., 2010; Martin-Ordas et al., 2012). Therefore, apes did not have to be trained with any aspect of the task and experience with the object could be established as part of their daily routine. The findings of Experiment 3 suggest that great apes may be able to form two representations of one object (i.e., one as a food item and one as a tool), and that they can be fixated on one of these representations because without an enduring representation of the first function, there could be no fixation effect.

In the following chapter, I summarize my findings, refer to some limitations and implications of the results, present my thoughts for future research, and give some final remarks.

Chapter 6: General discussion[5]

Summary

The aim of this thesis has been to investigate the role of prior experience and visual feedback in human and great ape tool use. It is likely that shared abilities have evolved in the last common ancestor of humans and great apes (Nielsen & Haun, 2016; Tomasello, 2014). Therefore, the comparative approach of studying our closest living relatives helps us better understand the cognitive abilities underpinning tool use in humans. Previous research has focused on the positive impact of experience with tools (e.g., Hirata, Morimura, & Houki, 2009) and the importance of visual feedback in problem solving (e.g., Völter & Call, 2012). The present findings show how experience with tools can negatively impact problem solving in great apes based on functional fixedness, and shed light on the processes that underlie success in the FPT. They additionally indicate great apes' reliance on visual feedback in complex tasks such as the FPT. These findings also show that human children's performance in the FPT is dependent on task presentation: Children are more likely to solve the FPT when the necessary tool is within their field of view.

---

[5] Material from Chapter 6 contributed to the following manuscripts (in preparation for submission): Ebel, S. J.; Völter, C. J.; Call., J. Functional fixation in the tool use of captive great apes (*Pan paniscus, Pan troglodytes, Pongo abelii*). Ebel, S. J.; Völter, C. J.; Call, J. Functional fixedness in great apes invoked by a food item.

Chapter 3 focused on the role of visual feedback and its timing in the FPT, also investigating whether additional information about task affordances facilitates identifying the solution in great apes. Apes were presented with several conditions comprising transparent and opaque tubes from which they could retrieve a peanut by adding water to the tube. Results indicate that subjects relied on visual feedback to initially identify the solution, but then kept up their performance independently of visual feedback. Moreover, encountering the end state of the solution (i.e., a water-filled tube with the peanut floating atop) increased success rates. These findings suggest that apes require visual feedback to solve the FPT, and that some solve the FPT by emulation learning.

Chapter 4 employed the FPT to study the functional fixedness effect in 6-year-old human children. On the one hand, the focus was on the modulating effect of own and other experience, and on the other hand, on the amount of experience with the tool. Children were presented with a transparent tube from which they could extract a ball by pouring water into it. The difference between groups was in whether prior to the task, children gathered experience watering plants. The results indicate that watering plants prior to the task did not lead to a functional fixedness effect. Rather, children benefited from an increased salience of the tool (i.e., the water bucket being transparent and close to the tube).

Chapter 5 examined the functional fixedness effect in great apes with three experiments. In the first experiment, the tool had two ends that each were functional in a different way, namely, a brush and a pointed end. In the second experiment, the tool also had two functions, but these were not spatially separated: a hose for drinking and poking. In the third experiment, the object in question, a grissini, combined the role of a food item and a raking tool. The findings suggest that apes were influenced by their prior experience with the tools: They narrowed down their manipulatory focus to those areas of the apparatus from which they had previously received rewards (Experiment 1). They took longer to solve the task or did not solve it at all (Experiment 2). Moreover, experience with the food item decreased the likelihood that they would employ it as a tool (Experiment 3).

The findings reported in Chapter 3 suggest that apes are dependent on visual feedback to solve the FPT. These results do not support the hypothesis that apes solve the task via insight learning (Köhler, 1925; Mendes et al., 2007; Shettleworth, 2012). Instead, adding water to the tube may have been reinforced by the perception of the peanut nearing the opening. Future studies may reveal if apes anticipate the outcome of their actions in less complex tasks than the FPT.

The findings of Chapter 4 further deepen our understanding of children's ability to solve the FPT. The study revealed unexpectedly low success rates in 6-year-olds (e.g., 8% in Experiment 1) compared to a previous study (42%; Hanus et al., 2011). This could be due to the arrangement of the task components and in fact, findings indicate a somewhat better performance when the tool is made more salient. However, no functional fixedness effect resulted when children used the water for plants prior to the test. Therefore, the impact of own and other experience, and of the amount of experience children had with the tool could not be investigated. These two topics should be addressed by future studies. Keeping the focus on the functional fixedness effect, Chapter 5 reports results clearly indicating that prior experience with a tool decreased apes' problem-solving abilities. However, it remains an open question if the fixation effect in this task was based on sensorimotor processes or conceptual representations. I discuss this topic further in Chapter 6B.

A. Limitations

The studies that have been reported in this thesis have several limitations, three of which I discuss here. The first limitation concerns the lack of a direct comparison between great apes and human children. In addition, only one population of apes was tested for each study and the sample size was rather small. The second limitation deals with the procedures employed to study specific concepts (i.e., “insight” and “functional fixedness”). The third limitation points out the ignorance of individual difference throughout the studies.

1. Direct comparison and sampling

In this thesis, the cognitive abilities of great apes (Chapter 3 and 5) and human children (Chapter 4) are studied separately. A direct comparison would be fruitful to better compare across species. Having the exact same procedures and experimenter would allow us to draw more specific conclusions about the evolution of human cognition as the species are studied under comparable conditions.

Moreover, the ape sample often comprised only one population of apes (the one from the WKPRC). Therefore, my conclusions from the reported studies may not be generalizable to other populations of apes. Most experiments comprised 20-25 apes which is a decent sample size for research in comparative psychology; however, it is not satisfactory from a statistical perspective. These individuals were also from different species, which reduces the actual sample size further and neglects species differences that cannot be detected with such small species-specific samples. Additionally, prior to my studies, the apes had participated in various experiments studying the cognitive abilities of apes, also involving tool use. Thus, individual apes may have differed in their prior experience with tools. Also, we could not counter-balance age perfectly across experimental groups due to the limited number of subjects available (yet, we did so as good as possible). This potentially leads to a neglect of developmental trajectories.

Sampling of the children entailed diverse populations (i.e., kindergartens). Every condition was run in each population tested (whenever possible) to control for differences between populations. However, these populations could also be grouped together as being Western, Educated, Industrialized, Rich, and Democratic (WEIRD; Nielsen & Haun, 2016), or more specifically, from a middle-sized German city. Comparative research with children from additional cultures, comprising also a larger age range, may yield further insights about the dependency of study results on specific populations and about developmental trajectories (Nielsen & Haun, 2016).

2. Concepts and procedures

In Chapter 1 and 3, I have discussed the notion of insight in ape problem-solving (e.g., Köhler, 1925; Sternberg & Davidson, 1995). Yet, it is still under discussion if this concept leads to a fruitful debate when studying humans and non-human animals (Bowden et al., 2005; Shettleworth, 2012). Insightful problem solving in humans is considered to be accompanied by an “aha”-moment that occurs during the restructuring of the representation of a problem (e.g., Dominowski & Dallob, 1995; Shen, Yuan, Liu, & Luo, 2016). Yet, non-linguistic animals cannot report their experience and it is hard to judge their subtle emotions from their behaviour, such as positive affect when coming up with the solution. Future studies may employ innovative techniques such as eye-tracking, measuring body temperature with heat cameras, or tracking skin conductance to explore apes’ emotional reactions during problem solving.

Duncker (1945) refers to the concept of “functional fixedness” by describing it as a functional representation of an object blocking other potential representations. This is an important consideration, yet, it leaves open two major questions. First, it does not refer to the type of experience required with an object to establish a fixation. In many studies familiar objects are used that subjects had gained experience with before the study took place (and then, the function of the object was re-activated in the study). Thus, the role of one’s actions with the object is unclear and how much experience is needed to get a fixation effect. Second, since the experience with the object is not exactly known, the processes the effect is based on are open for discussion. While experience with using a tool may establish both, manipulation and function knowledge (e.g., Buxbaum & Saffran, 2002), encountering novel tools with an ascribed function may only bring about function knowledge. I discuss this aspect further in Chapter 6B.

3. Individual differences

Recent studies on differential psychology in great apes have proposed stable individual differences between apes (e.g., Uher & Asendorpf, 2008; Uher et al., 2008). This was not the focus of my thesis, but it is an intriguing idea that apes consistently differ in their approach towards problems. For example, persistence has been proposed to play a key role in problem solving (e.g., Huebner & Fichtel, 2015). Also, humans have been shown to differ in their level of rigidity in their behaviour (Schultz & Searleman, 2002), suggesting that individual apes may be more vulnerable to functional fixedness than others. Finally, Beck et al. (2016) explored which factors influence success in the hook task in human children. Findings indicated that neither executive functions, nor divergent thinking abilities play a major role in solving the task. Thus, what characterizes individuals who solve innovation problems remains an open, but exciting question for future research.

B. Implications

1. Prior experience: A double-edged sword

As suggested in Chapter 5, prior experience with tools represents a double-edged sword: On the one hand, it can facilitate problem solving by reducing latency to success (Ebel & Call, 2018; Mendes et al., 2007; Vale et al., 2016) and by helping to transfer solutions to similar problems (Chapter 2; Manrique et al., 2010; Vale et al., 2016). On the other hand, it can complicate problem solving or even prevent solutions when problems require using a tool for a novel function (Chapter 5; Gruber et al., 2011; Gruber et al., 2009; Marshall-Pescini & Whiten, 2008). This finding is strikingly counterintuitive: We usually assume that experience with a tool should increase an individual's performance. For example, chimpanzees crack open nuts more efficiently with increased experience, that is, they learn how to select the best hammers and anvils, how to place a nut on an anvil, how to orient and move a hammer in relation to the nut and so on (Biro, Sousa, & Matsuzawa, 2006; Hirata, Morimura, & Houki, 2009).

While experience with a tool in one specific function may lead to functional fixedness, experience with the same tool in diverse functions may increase a user's flexibility with it (e.g., Flavell et al., 1958). Similarly, while novices in a given field are slower in finding solution strategies to typical problems from that field, they may be more flexible in their responses than experts. Experts exhibit biases towards familiar strategies and when unusual solutions are needed, only an even higher level of expertise helps to overcome these biases and to quickly find novel strategies (e.g., in chess; Bilalić et al., 2008a; Bilalić et al., 2008b; Sheridan & Reingold, 2013, 2014). So which processes underlie the diverging effects of prior experience in tool use that make experience a double-edged sword?

Two integration steps in human ontogeny

Before turning to apes, I explore general processes underlying tool use in human children. I argue that there are two major integration steps that occur at different times during human ontogeny. The first one concerns the integration of perceptual and motor processes (Lockman, 2000) and the second one relates to the integration of sensorimotor processes and tool knowledge (Greif & Needham, 2011).

There is evidence for a dissociation between perceptual and motor processes in toddlers: In a study by DeLoache, Uttal, and Rosengren (2004), children aged 2.5 years who had previously acted upon a toy in a specific way, subsequently encountered the same item in a different size. They tried to perform the same actions with it as before even if these were obviously not possible any more (for a study about the scale error with tools, see Hunley & Hahn, 2016). For example, children tried to get into a toy car that looked like a previously experienced large exemplar, but was so small that even their feet would not fit in (DeLoache et al., 2004). Babies have to learn how to grasp a tool and how to orient and move it in relation to other objects. They do this by repeating learned actions (Greif & Needham, 2011; see also Thelen, Schöner, Scheier, & Smith, 2001 for a dynamical system approach on children's learning with objects). Perseveration with certain types of grasping tools occurs early during ontogeny and leads to decreased problem-solving success in young children when the solution requires grasping a familiar tool in a novel way. For example, children struggle to grasp and hold a spoon at its bowl instead of its handle when this is required to be effective in a specific task (e.g., Barrett, Davis, & Needham, 2007; McCarty, Clifton, & Collard, 1999, 2001). Thus, young children seem to focus more on where to hold a tool than on its function during this first learning phase (Barrett et al., 2007).

As expected, children become more flexible with grasping tools with increasing age. Yet, 2- to 4-year-olds still continue to perform familiar actions with tools that have become inefficient or even prevent solutions (Elsner & Schellhas, 2012; Smitsman & Cox, 2008). In addition to studying these types of perseveration behaviour with tools in children, researchers have investigated children’s understanding of tool functions (see also *6.B.2. Object representations*). Once 2-year-olds are able to use a novel tool as demonstrated by an adult, they also prefer to use this familiar tool instead of a novel one when presented with a novel task, i.e., they use the familiar tool for a novel function (Casler & Kelemen, 2007). However, by the age of 2.5 years, they choose the novel tool for the novel task, suggesting a dissociation of tool functions (Casler & Kelemen, 2005, 2007).

By the age of six years, children also exhibit a functional fixedness effect: Their knowledge about the function of a tool prevents them from using it in a novel context (Defeyter & German, 2003; Duncker, 1945; German & Defeyter, 2000). In these studies, the function is often established by presenting the tool in a specific and familiar function, e.g., a spoon is located inside a cup with rice (German & Barrett, 2005). When asked to suggest functions for a tool, 7-year-olds take information about its design into consideration (i.e., what it may have been made for). In contrast, 5-year-olds are more flexible in suggesting tool functions than the older children (Defeyter et al., 2007). Yet, the older children produce more diverse functions for tools when information on design is not given. These studies suggest that younger children respond more to tool affordances (i.e., how tools can be grasped and what can be done with them) than older children. The latter, in contrast, integrate information on design (either inferred from the tool or provided by an agent) better than younger children.

Interestingly, children also start to become proficient tool innovators around the age of six to eight years. For example, they may solve the FPT and the hook task, which require pouring water into a tube to retrieve a floating object and bending a pipe cleaner to extract a small bucket from a tube respectively (Chapter 3 and 4; Beck et al., 2011; Beck et al., 2016; Chappell, Cutting, Apperly, & Beck, 2013; Cutting et al., 2011; Cutting, Apperly, Chappell, & Beck, 2014; Hanus et al., 2011). Beck and colleagues (2011) have proposed that children struggle with tool innovation tasks because these type of problems are ill-structured (Jonassen, 1997): The actions required to transform the starting state into the solution state are unknown (e.g., Chappell et al., 2013; Cutting et al., 2011; Cutting et al., 2014). As proposed, perceptual and motor processes are integrated first during human ontogeny and tool knowledge is integrated with sensorimotor processes second (Greif & Needham, 2011). It seems an intriguing possibility that the second integration step leads to the ability to solve ill-structured problems since these may require the integration of all three domains. So what can be learned from these findings about the underlying processes of functional fixedness in human children and great apes?

Two types of fixation effects

Assuming that during the development of tool use perceptual and motor processes are integrated first and later sensorimotor processes are integrated with tool knowledge (Greif & Needham, 2011), two possible types of functional fixedness can be defined. One is based on sensorimotor processes (Elsner & Schellhas, 2012; Smitsman & Cox, 2008) and the other one on tool knowledge (Defeyter & German, 2003; German & Defeyter, 2000). Authors have referred to these two types of knowledge on the one hand as “practical”, “procedural”, or “manipulation” knowledge, and on the other as “conceptual” or “functional” knowledge (please note that the definitions may also vary across authors; e.g., Buxbaum & Saffran, 2002; Manrique et al., 2010). More precisely, young children perseverate on their *actions* with tools, causing inefficiency or barring success (Elsner & Schellhas, 2012; Smitsman & Cox, 2008). In contrast, by the age of six years they perseverate on the *concept* or *function* of tools, preventing their use with a novel function (Defeyter & German, 2003; German & Barrett, 2005; German & Defeyter, 2000). The developmental literature refers to the former type of fixation effect in tool use as “perseveration” (Elsner & Schellhas, 2012; Smitsman & Cox, 2008), and to the latter as “functional fixedness” (Defeyter & German, 2003; German & Defeyter, 2000).

However, to my knowledge these terms are usually not differentiated from each other explicitly in this way. Indeed, this may explain the puzzling variation across studies with regard to the age when fixation effects occur in children's tool use (Barrett et al., 2007; Defeyter & German, 2003; Elsner & Schellhas, 2012; German & Barrett, 2005; German & Defeyter, 2000; McCarty et al., 1999, 2001; Smitsman & Cox, 2008). Based on this working hypothesis, which type of fixation effect did great apes exhibit throughout the experiments discussed in Chapter 5?

Recall the findings from Experiment 2: Apes who previously used a hose for drinking were less likely to use the same hose to poke a banana pellet out of a tube than apes without such functional tool experience. This finding indicates that apes exhibited a fixation effect, but leaves open what processes the effect was based on. Interestingly, apes without prior experience extracted the hose from the empty drinking container significantly sooner than apes with prior experience. However, these latter chimpanzees with prior experience tended to touch the hose sooner within a session. They entered the testing room and briefly touched the hose or sucked it, but then left it in the drinking container and manipulated the apparatus with another tool instead. Thus, they apparently did not consider the hose as a potential tool for the task, even though they had noticed it early in the session. On the one hand, sucking the hose may be interpreted as perseveration behaviour: Apes responded to the previous tool affordances, and this did not bring them closer to their goal of extracting the pellet from the tube. On the other hand, by itself, sucking the hose did not keep them from using it to poke out the pellet from the tube. Rather, they failed to come up with the hypothesis that the hose could be used in a different function – this may suggest a fixation due to tool knowledge. So perhaps what is needed is an experimental design that allows to differentiate between both types of functional fixedness (see also Chapter 6.C.1.).

Functional fixedness in adult humans has been tested with two types of paradigms: Participants either used a tool prior to testing (often using familiar tools such as gimlets or pliers) or the presentation of the tool established its function (again familiar tools were used such as boxes as containers). Duncker (1945) did not find a difference between both paradigms when testing human adults. Children have been mainly tested with the latter paradigm (Clegg & Legare, 2016; Defeyter & German, 2003; German & Barrett, 2005; German & Defeyter, 2000). As I have argued above, tool knowledge may develop later in children than perception and manipulation of tools (Greif & Needham, 2011). This leaves open the possibility that different paradigms for testing functional fixedness lead to different results in children, whereas they may not do so in adults. Therefore, a promising approach would be to test children at various ages and great apes with both paradigms. Tool functions would be either established through prior action or by presentation, employing both familiar and novel tools. This would allow us to explore the cognitive underpinnings of functional fixedness and the role of prior experience with tools.

To conclude, further research is needed to determine, if there are two distinct types of fixation effects in children's tool use, and if so, how they are related to each other, what role experience with the tool in one specific or multiple functions plays, and to which of these effects great apes are susceptible.

Alternatively, one could also question the conclusion that there are two distinct types of functional fixedness. Some authors do argue that conceptual tool knowledge always comprises sensorimotor processes as well and that this knowledge only accumulates with age (Thelen et al., 2001). Further theoretical work is needed to shed light on the cognitive underpinnings of human and non-human primate tool use to settle this question. In the next chapter, I discuss great apes' tool knowledge and understanding of tool functions, and how they may represent objects as compared to humans.

2. Object representations

Humans surround themselves with a multitude of artefacts that have been produced for specific purposes (Csibra & Gergely, 2009; Hernik & Csibra, 2009; Ruiz & Santos, 2013; Vaesen, 2012). Young children infer a tool's function by observation, but they do not exclusively use the tool with one specific function yet. When growing older they start using tools exclusively for specific functions, which later leads to a functional fixedness effect. At the same time during development, children seem to start to reason about what tools have been *made for* (design stance or teleological-intentional stance; Casler & Kelemen, 2005, 2007; Csibra & Gergely, 2007; Hernik & Csibra, 2009; Kelemen, 1999; Kemler Nelson, Frankenfield, Morris, & Blair, 2000; Nielsen, 2006; Ruiz & Santos, 2013; Siegel & Callanan, 2007).

Great apes, in contrast, usually do not save tools for future usage, that is, they do not collect tools and carry them with them if they do not need them in the current situation (Boesch & Boesch, 1984; McGrew, 1992; although they are generally able to do so, see Mulcahy & Call, 2006; Osvath & Osvath, 2008). So why should they have enduring functional representations of tools? However, there is some evidence that tools which endure over time such as hammer and anvils for nut-cracking or abandoned sticks for ant dipping may facilitate future tool use (Fragaszy et al., 2013). Especially with more complex tool use this may be advantageous for young individuals to acquire these skills (e.g., when tool use involves tool manufacturing; Fragaszy et al., 2013; Sanz et al., 2009).

While children experience objects as embedded in a social world (i.e., they are told what certain objects are for or reason about their function from their design or others’ actions), great apes do not seem to teach learners actively (i.e., they do not show them intentionally how to use a tool; Lonsdorf, 2006). Some authors consider apes to teach infants, but they mostly refer to instances which comprise the provisioning of tools and social tolerance with youngsters than active (intentional) teaching (Biro et al., 2003; Boesch, 1991; Musgrave, Morgan, Lonsdorf, Mundry, & Sanz, 2016). Moreover, while human children mainly learn by imitation and faithfully copy an adult’s action even if it was not necessary to achieve a certain goal ("over-imitation"; e.g., Lyons et al., 2007; Taniguchi & Sanefuji, 2017), chimpanzees often seem to emulate the result (i.e., reproduce the task’s solution by different means than the ones that they have observed) and only imitate actions when the causal structure of the task is opaque (Call et al., 2005; Hernik & Csibra, 2009; Horner & Whiten, 2005;

McGuigan et al., 2007; Tennie et al., 2006; Tennie, Call, et al., 2010; Whiten, Custance, Gomez, Teixidor, & Bard, 1996; Whiten et al., 2009). Thus, apes seem to be less interested in the precise actions of others than human children. When they can approach a tool-use problem with their own solution strategy, they will. On the contrary, children pay attention to what a tool *is for* (function) or how it *should* be used (norm). The interplay between these two aspects of human tool use (i.e., function knowledge and knowledge about conventions or norms) may be a fruitful topic for future research.

Functional fixedness to study object representations

Are humans unique with regard to their object representations? Even though I give an outlook on possible future research questions in Chapter 6C, I would like to propose a possible procedure to tackle this question at this point: One could use the functional fixedness effect to evaluate if apes possess function knowledge of tools and if this knowledge is dependent on the type of experience (i.e., if they use the tool themselves or observe another individual using it). Consider the experimental setup in Chapter 5 (Experiment 2): Some apes learned that a hose could be used to drink grape juice. In the test, apes were then presented with the (dry) hose inserted into the empty drinking box and two additional tools were lying at the side. Apes had to select the hose from the box to poke out a food item from a novel apparatus.

What if this experiment was repeated, but this time the hose was transferred into a novel context? For example, the apes could select the hose among three tools from a platform to solve the novel task. One could also change perceptual features of the tool in the test to see if apes generalize their tool knowledge. This paradigm would prevent apes from perseverating on the previous tool function and test for their representations of the tools more closely: Would they avoid using a familiar tool for a novel function even if this function is not apparent in the current situation?

This paradigm could also be transferred into the social domain: Subjects learn about the tool's function by observing a conspecific. If apes then showed a functional fixedness effect, this would be an indicator that observing how a conspecific uses the tool is processed in a similar way as when the individual uses the tool herself. Or to put it in other words, it would show that objects can also be socially embedded in great apes, which has been questioned by many authors (Csibra & Gergely, 2009; Hernik & Csibra, 2009; Ruiz & Santos, 2013; Vaesen, 2012). As suggested, the functional fixedness effect has a great potential to study function knowledge and object representations in great apes. The presentation of the tool will be crucial in these studies – as will be the presentation of other task components, as discussed in the following section.

3. Insight into task components

As demonstrated in Chapter 4, children's problem-solving performance in the FPT was dependent on the presentation of the task components. Success rates increased when the tool (i.e., the water) was made more salient by presenting it closer to the tube. Moreover, success rates were much lower than in a previous study (Hanus et al., 2011). I presented the FPT embedded into a game that comprised multiple objects, whereas the previous study only used the tube and a water pitcher. This again indicates that the specific presentation of the task's components matter for children's problem-solving performance (see also Jonassen, 2000). One possible underlying process might be children's attentional focus. It has been shown for insight problems that subjects suddenly switch their attentional focus to a task component they have disregarded before (e.g., Sternberg & Davidson, 1995). In the condition in which the water bucket was located on the floor next to the table (Chapter 4, Experiment 1), children sometimes wandered around the testing room. They suddenly stopped in front of the water bucket and stared at it, before happily stating and executing the solution. Sometimes they also alternated their gaze between the water and the tube multiple times, before they smiled and reported their idea. This behavioural evidence – although anecdotal – may suggest that children suddenly had an insight into the task's solution: They became aware of the water as a potential component of the task and this happened when the water came into their (visual) focus.

The notion of insight is controversial in animal behavioural research, but some researchers have tried to find concrete behavioural indicators of insight (Bowden et al., 2005; Knoblich, Ohlsson, & Raney, 2001; Köhler, 1925; Shettleworth, 2012; Sternberg & Davidson, 1995). The FPT has been suggested to be a problem which is usually solved via insight (Mendes et al., 2007; Shettleworth, 2012). To explore this possibility one could code children's behaviour and interviewing them after the test. Additionally, it would be interesting to compare children and apes directly with regard to some measurements such as the delay from first pouring action to success. Insightful problem solving would suggest a fast solution after the first pouring action (i.e., goal directed). Great apes did not solve the task spontaneously with an opaque tube (Chapter 3) so that I did not find evidence that they anticipated the outcome of their actions in the FPT. However, this task was extremely difficult to solve. It would be interesting to present children of various ages with the opaque tube and investigate first, their overall success and second, which hints at the solution help them with solving the task.

C. Future directions

1. Functional fixedness

Functional fixedness paradigms seem to yield a great opportunity to study function knowledge in great apes. In the following, I propose a diverse application of these paradigms to study enduring functional representations in great apes (and human children). For example, would apes dissociate tool functions? When presented with two functional tools of which one is familiar while the other is novel would they select the novel tool for the novel task (Casler & Kelemen, 2005, 2007)? How much experience is needed with tools to exhibit a functional fixedness effect? What are the time constraints, i.e., when does the effect decay (Flavell et al., 1958; Yonge, 1966)? Moreover, would apes exhibit a functional fixedness effect if they were presented with the tool in its function instead of using the tool? For example, would they solve a task more slowly that required using a box with a novel function (e.g., box as support for climbing) when presented with the experimental materials lying inside the box (box as container) than when the material is presented next to the box (“pre-utilization condition”; Defeyter & German, 2003; Duncker, 1945; German & Defeyter, 2000)?

To further evaluate if apes represent function knowledge, one could give apes prior experience with a tool and conduct the test in a completely novel task environment that prevents perseveration behaviour of the previous function. Alternatively, one could exchange the tool in the test by a similar tool that differs with regard to some perceptual features, or one from the same class or category of tools to assess apes’ generalization ability (Casler & Kelemen, 2005, 2007).

Furthermore, would apes exhibit a functional fixedness effect after observing a conspecific? If so, this would show that objects have a social dimension for apes as well (Hernik & Csibra, 2009; Ruiz & Santos, 2013). Moreover, would an apparatus with an opaque causal structure strengthen or weaken the effect in humans and great apes (in an individual and social context)? Functions of human artefacts are often opaque and children show a strong bias toward imitation learning (with tools) irrespective of the causal structure of the task. Chimpanzees only imitate actions when the causal structure of a task remains opaque.

Finally, although I focus on functional fixedness with tools in this thesis, many apes also managed to use the familiar tool with a novel function throughout the studies reported in Chapter 5. Repeated innovation, that is, finding a novel solution to a problem after the previously known one is blocked, is not well studied in great apes yet (e.g., Manrique et al., 2013). It would be interesting to investigate the relationship of functional fixedness to functional flexibility and to examine their cognitive underpinnings. Studying individual differences with regard to flexibility and fixation would be a great endeavour as well. Moreover, investigating the relationship of the Einstellung effect and functional fixedness in great apes and human children more closely would reveal if they rely on the same cognitive processes.

2. Own- versus other-experience

In Chapter 3 and 4, I report on two experiments that involve the comparison between the impact of performing an action oneself or observing the action in another agent. Unfortunately, I could not assess this variable due to low success rates in both experiments. Recent studies have shown that human children and non-human primates take own- and other-actions into account (Kuroshima, Kaiser, & Fragaszy, 2014; Sommerville et al., 2008; Tomasello, 2014); however, both types of learning are understudied and need further consideration in future research, especially in the case of great apes.

Finally, studying own- and other-experience more closely in great apes and human children could reveal species specific learning biases. For example, do apes and children process information about objects the same way if they use the objects themselves compared to when they observe others using the objects? And is the former type of experience more memorable than the latter one? It is also an interesting question how these two types of experience relate to the functional fixedness effect and distinct learning strategies in different species. Especially in the case of children it would be interesting to also investigate the role of norms. Many studies involve a demonstration by a human experimenter who tells the child to watch while she uses a tool (e.g., Casler & Kelemen, 2005; Casler & Kelemen, 2007). Then, children's understanding of the tool's function is investigated. It would be stimulating to compare children's function understanding of tools whose function they had learned through individual learning with one of tools whose function they have learned by others.

It would further be interesting to study the role of reinforcement in social learning more intensely. What do individuals exactly learn when they copy actions, goals and/or results? For example, social demonstrations in which subjects perceive the whole action sequence required to solve the task may differ in their reinforcement structure: In some studies, subjects receive the reward from the apparatus after the demonstration, whereas in other studies the demonstrator gets it. Moreover, the visibility of the reward may play a major role, i.e., if a transparent or opaque apparatus is used as we know that apes narrow down their attentional focus onto the food when it is visible (Birch, 1945; Ebel & Call, 2018; Vlamings, Uher, & Call, 2006).

3. Tool innovation

Beck and colleagues discovered only recently that children struggle with relatively simple tool innovations, e.g., when they have to bend a pipe cleaner to extract a small bucket from a tube (Beck et al., 2011; Beck et al., 2014; Chappell et al., 2013; Cutting et al., 2011; Cutting et al., 2014). This phenomenon seems quite robust and most hints at the solution do not help children to come up with the solution. It would be interesting to directly compare human children and great apes with the exact same setup to investigate apes’ performance in relation to children’s performance at different ages.

A related area of future research may be the direct comparison of different innovation tasks such as the Floating Peanut Task, the hook task, and the Aesop’s Fable task. I suggest that they may share the same underlying structure since no hint is given how to transform the problem from the starting state to the goal state (ill-structured problem; Cutting et al., 2011; Jonassen, 2000). It would be stimulating to compare all three tasks with comparable samples of children and perhaps great apes. Would children start solving the tasks around the same age? Recent studies suggest this. So what make this age special? One could relate individuals’ success in these tasks to other tasks requiring function knowledge. This would reveal if children that have generally a more elaborated knowledge about tools are more likely to solve innovation problems. Since all three tasks involve transparent vertical tubes and two tasks involve water, they provide the opportunity to study fixation on the task’s solution

(Einstellung) and on the tool (functional fixedness) as well as repeated innovation, i.e., the overcoming of such fixations.

Finally, it would be interesting to explore further modulating variables that influence task performance. For example, which combinations of hints at the solution facilitate success in both human children and great apes? Robust low success rates have been found in children (and apes), but a combination of manipulation experience with the tool and seeing the target tool helped older children. To my knowledge, there is no published study about the hook task with apes, however, the New Caledonian crow Betty that was first tested with the task and Goffin's cuckatoos are dependent on experience with the target tool to solve the hook task (Laumer, Bugnyar, Reber, & Auersperg, 2017; Weir, Chappell, & Kacelnik, 2002). So how would great apes from various species perform in this task? It would generally be fascinating to study innovation problems with regard to the role of prior experience with parts of the problems, especially tools further.

D. Conclusion

Since Wolfgang Köhler’s pioneering work on primate cognition, tool use has been the focus of intensive research (Köhler, 1925). Studies comparing non-human great apes with human children have improved our understanding of the evolution of cognitive abilities (Tomasello, 2014). Many authors have suggested that human physical cognition is special, namely, that object representations are embedded in a social infrastructure (Ruiz & Santos, 2013). Human children already reason about what an object is made for (Hernik & Csibra, 2009). One question is whether functional knowledge of objects as *being for* a certain purpose is really unique to humans. The notion of *being made for* something more likely is, because apes do not encounter many artefacts in their daily lives. But humans’ intense reasoning about object functions also has a drawback. It narrows down their attentional focus to certain functions of objects, which obstructs creative problem solving (functional fixedness; Duncker, 1945).

The aim of this thesis has been to investigate the role of prior experience and visual feedback in great apes’ and human children’s tool use. I found remarkable parallels between apes and human children in the negative impact of prior experience with tools. These findings also call for more elaborated studies on fixation effects and their cognitive underpinnings. Potential candidates are sensorimotor processes and function knowledge (Greif & Needham, 2011). Perseveration with tools that is possibly based on sensorimotor processes is already found in young infants, whereas fixation on tool functions (i.e., tool knowledge) occurs later in human ontogeny (Elsner &

Schellhas, 2012; German & Defeyter, 2000). The studies reported in this thesis leave open the question which type of fixation great apes exhibited, an important issue for future studies. Investigating functional fixedness in great apes is a powerful method for studying function knowledge: If apes showed a functional fixedness effect, this would indicate that they represent a tool as having a certain function. Otherwise, there would be no fixation effect. However, to study a fixation on the function of the tool (i.e., on the tool's representation), fixation based on motor processes has to be excluded as an alternative explanation.

Visual feedback can play a pivotal role in identifying problem solutions in both children and apes (e.g., Völter & Call, 2012). This can occur either by revealing the effect of one's actions or by giving hints about the task solution. Apes focus on emulation learning (i.e., reproducing the end-state of the solution) and only copy actions if required (e.g, if the task is opaque or imitation rational; Buttelmann et al., 2007; Horner & Whiten, 2005). Human children confidently imitate demonstrated actions, even if these are redundant (overimitation; Horner & Whiten, 2005; Lyons et al., 2007). How these species-specific learning biases relate to the learning of tool functions and to functional fixedness more specifically will be a matter of continued research. Future studies, elaborating on tool functions, object representations, and learning mechanisms in great apes will contribute to our knowledge about the evolution of physical cognition in humans.

References

Adamson, R. E. (1952). Functional fixedness as related to problem solving: a repetition of three experiments. *Journal of Experimental Psychology, 44*(4), 288-291. doi:10.1037/h0062487

Allritz, M., Tennie, C., & Call, J. (2013). Food washing and placer mining in captive great apes. *Primates, 54*(4), 361-370. doi:10.1007/s10329-013-0355-5

Amici, F., Aureli, F., & Call, J. (2008). Fission-fusion dynamics, behavioral flexibility, and inhibitory control in primates. *Current Biology, 18*(18), 1415-1419. doi:10.1016/j.cub.2008.08.020

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278. doi:10.1016/j.jml.2012.11.001

Barrett, T. M., Davis, E. F., & Needham, A. (2007). Learning about tools in infancy. *Developmental Psychology, 43*(2), 352.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Bateson, P. (2014). Play, playfulness, creativity and innovation. *Animal Behavior and Cognition, 1*(2), 99-112.

Beck, B. B. (1980). *Animal tool behavior. The use and manufacture of tools by animals*. New York & London: Garland STPM Press.

Beck, S. R., Apperly, I. A., Chappell, J., Guthrie, C., & Cutting, N. (2011). Making tools isn't child's play. *Cognition, 119*(2), 301-306. doi:10.1016/j.cognition.2011.01.003

Beck, S. R., Cutting, N., Apperly, I. A., Demery, Z., Iliffe, L., Rishi, S., & Chappell, J. (2014). Is tool-making knowledge robust over time and across problems? *Frontiers in psychology, 5*, 1395.

Beck, S. R., Williams, C., Cutting, N., Apperly, I. A., & Chappell, J. (2016). Individual differences in children's innovative problem-solving are not predicted by divergent thinking or executive functions. *Philosophical Transactions of the Royal Society B-Biological Sciences, 371*(1690). doi:10.1098/rstb.2015.0190

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science, 8*(6), 492-499. doi:10.1111/j.1467-7687.2005.00440.x

Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: comparing 12- and 18-month-olds. *Infant Behavior & Development, 22*(2), 277-282.

Beran, M. J. (2002). Maintenance of self-imposed delay of gratification by four chimpanzees (*Pan troglodytes*) and an Orangutan (*Pongo pygmaeus*). *Journal of General Psychology, 129*(1), 49-66. doi:10.1080/00221300209602032

Bilalić, M., McLeod, P., & Gobet, F. (2008a). Inflexibility of experts—Reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology, 56*(2), 73-102. doi:http://dx.doi.org/10.1016/j.cogpsych.2007.02.001

Bilalić, M., McLeod, P., & Gobet, F. (2008b). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition, 108*(3), 652-661. doi:10.1016/j.cognition.2008.05.005

Birch, H. G. (1945). The role of motivational factors in insightful problem-solving. *Journal of Comparative Psychology, 38*(5), 295-317. doi:10.1037/h0059937

Birch, H. G., & Rabinowitz, H. S. (1951). The negative effect of previous experience on productive thinking. *Journal of Experimental Psychology, 41*(2), 121-125. doi:10.1037/h0062635

Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C., & Matsuzawa, T. (2003). Cultural innovation and transmission of tool use in wild chimpanzees: evidence from field experiments. *Animal Cognition, 6*(4), 213-223. doi:10.1007/s10071-003-0183-x

Biro, D., Sousa, C., & Matsuzawa, T. (2006). Ontogeny and cultural propagation of tool use by wild chimpanzees at Bossou, Guinea: Case studies in nut cracking and leaf folding. In T. Matsuzawa, M. Tomonaga, & M. Tanaka (Eds.), *Cognitive Development in Chimpanzees* (pp. 476-508). Tokyo: Springer Tokyo.

Boesch, C. (1991). Teaching among wild chimpanzees. *Animal Behaviour, 41*, 530-532. doi:10.1016/s0003-3472(05)80857-7

Boesch, C. (2013). Ecology and cognition of tool use in chimpanzees. In C. M. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals. Cognition and ecology*. Cambridge: Cambridge University Press.

Boesch, C., & Boesch, H. (1984). Mental map in wild chimpanzees: An analysis of hammer transports for nut cracking. *Primates, 25*(2), 160-170. doi:10.1007/bf02382388

Boesch, C., Head, J., & Robbins, M. M. (2009). Complex tool sets for honey extraction among chimpanzees in Loango National Park, Gabon. *Journal of Human Evolution, 56*(6), 560-569. doi:http://dx.doi.org/10.1016/j.jhevol.2009.04.001

Bonawitz, E., Shafto, P., Gweon, H., Chang, I., Katz, S., & Schulz, L. (2009, July29th-Aug.1st). *The double-edged sword of pedagogy: modeling the effect of pedagogical contexts on preschoolers' exploratory play.* Paper presented at the 31st Annual Meeting of the Cognitive Science Society, CogSci, VU University Amsterdam, Netherlands.

Bowden, E. M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences, 9*(7), 322-328. doi:10.1016/j.tics.2005.05.012

Brewer, S. M., & McGrew, W. C. (1990). Chimpanzee use of a tool-set to get honey. *Folia primatologica, 54*(1-2), 100-104. doi:10.1159/000156429

Burghardt, G. M. (2006). *The genesis of animal play - testing the limits* (First MIT paperback ed.). Cambridge, Massachusetts; London, England: The MIT Press.

Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated chimpanzees imitate rationally. *Developmental Science, 10*(4), F31-38. doi:10.1111/j.1467-7687.2007.00630.x

Buxbaum, L. J., & Saffran, E. M. (2002). Knowledge of object manipulation and object function: dissociations in apraxic and nonapraxic subjects. *Brain and language, 82*(2), 179-199.

Byrne, R. W., Sanz, C. M., & Morgan, D. B. (2013). Chimpanzees plan their tool use. In C. M. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals. Cognition and ecology*. Cambridge: Cambridge University Press.

Call, J. (2004). Inferences about the location of food in the great apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla*, and *Pongo pygmaeus*). *Journal of Comparative Psychology, 118*(2), 232-241. doi:10.1037/0735-7036.118.2.232

Call, J. (2007). Apes know that hidden objects can affect the orientation of other objects. *Cognition, 105*(1), 1-25. doi:10.1016/j.cognition.2006.08.004

Call, J. (2013). Three ingredients for becoming a creative tool user. In C. M. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals. Cognition and ecology* (pp. 3-20). Cambridge: Cambridge University Press.

Call, J., Carpenter, M., & Tomasello, M. (2005). Copying results and copying actions in the process of social learning: chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Animal Cognition, 8*(3), 151-163. doi:10.1007/s10071-004-0237-8

Carvalho, S., Biro, D., McGrew, W. C., & Matsuzawa, T. (2009). Tool-composite reuse in wild chimpanzees (*Pan troglodytes*): archaeologically invisible steps in the technological evolution of early hominins? *Animal Cognition, 12 Suppl 1*, S103-114. doi:10.1007/s10071-009-0271-7

Casler, K., & Kelemen, D. (2005). Young children's rapid learning about artifacts. *Developmental Science, 8*(6), 472-480. doi:10.1111/j.1467-7687.2005.00438.x

Casler, K., & Kelemen, D. (2007). Reasoning about artifacts at 24 months: The developing teleo-functional stance. *Cognition, 103*(1), 120-130. doi:10.1016/j.cognition.2006.02.006

Chappell, J., Cutting, N., Apperly, I. A., & Beck, S. R. (2013). The development of tool manufacture in humans: what helps young children make innovative tools? *Philosophical Transactions of the Royal Society B-Biological Sciences, 368*(1630). doi:10.1098/rstb.2012.0409

Chown, S. M. (1959). Rigidity: A flexible concept. *Psychological Bulletin, 56*(3), 195.

Clegg, J. M., & Legare, C. H. (2016). Instrumental and conventional interpretations of behavior are associated with distinct outcomes in early childhood. *Child Development, 87*(2), 527-542. doi:10.1111/cdev.12472

Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica, 124*(1), 60-78. doi:10.1016/j.actpsy.2006.09.007

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Science, 13*(4), 148-153. doi:10.1016/j.tics.2009.01.005

Cutting, N., Apperly, I. A., & Beck, S. R. (2011). Why do children lack the flexibility to innovate tools? *Journal of Experimental Child Psychology, 109*(4), 497-511. doi:10.1016/j.jecp.2011.02.012

Cutting, N., Apperly, I. A., Chappell, J., & Beck, S. R. (2014). The puzzling difficulty of tool innovation: Why can't children piece their knowledge together? *Journal of Experimental Child Psychology, 125*, 110-117. doi:10.1016/j.jecp.2013.11.010

Defeyter, M. A., Avons, S. E., & German, T. C. (2007). Developmental changes in information central to artifact representation: evidence from 'functional fluency' tasks. *Developmental Science, 10*(5), 538-546. doi:10.1111/j.1467-7687.2007.00617.x

Defeyter, M. A., & German, T. P. (2003). Acquiring an understanding of design: evidence from children's insight problem solving. *Cognition, 89*(2), 133-155.

Defeyter, M. A., Hearing, J., & German, T. C. (2009). A developmental dissociation between category and function judgments about novel artifacts. *Cognition, 110*(2), 260-264. doi:10.1016/j.cognition.2008.10.014

DeLoache, J. S., Miller, K. F., & Pierroutsakos, S. L. (1998). Reasoning and problem solving. In W. Damon (Ed.), *Handbook of child psychology* (5th ed., Vol. 2: Cognition, perception, and language, pp. 801-850). New York: John Wiley & Sons, Inc.

DeLoache, J. S., Uttal, D., & Rosengren, K. S. (2004). Scale errors offer evidence for a perception-action dissociation early in life. *Science, 304*(5673), 1027-1029. doi:10.1126/science.1093567

Dominowski, R. L., & Dallob, P. (1995). Insight and problem solving. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight*. Cambridge, Massachusetts; London, England: The MIT Press.

Duncker, K. (1945). On problem-solving. *Psychological Monographs, 58*(5), 1-112.

Ebel, S. J., & Call, J. (2018). The interplay of prior experience and motivation in great ape problem-solving (*Gorilla gorilla, Pan paniscus, Pan troglodytes*, and *Pongo abelii*). *Journal of Comparative Psychology. Advance online publication*. doi:10.1037/com0000117

Elsner, B., & Pauen, S. (2007). Social learning of artefact function in 12-and 15-month-olds. *European Journal of Developmental Psychology, 4*(1), 80-99.

Elsner, B., & Schellhas, B. (2012). The acquisition of flexible tool use in preschoolers: the impact of prior experience. *Zeitschrift Fur Psychologie-Journal of Psychology, 220*(1), 44-49. doi:10.1027/2151-2604/a000090

Field, A. P. (2005). Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychological Methods, 10*(4), 444-467. doi:10.1037/1082-989x.10.4.444

Flavell, J. H., Cooper, A., & Loiselle, R. H. (1958). Effect of the number of pre-utilization functions on functional fixedness in problem-solving. *Psychological Reports, 4*(2), 343-350.

Fleagle, J. G. (1988). *Primate adaptation & evolution*. San Diego: Academic Press.

Forss, S. I., Schuppli, C., Haiden, D., Zweifel, N., & van Schaik, C. P. (2015). Contrasting responses to novelty by wild and captive orangutans. *American Journal of Primatology, 77*(10), 1109-1121. doi:10.1002/ajp.22445

Fox, J., & Weisberg, S. (2011). An {R} Companion to Applied Regression. Thousand Oaks CA: Sage. Retrieved from http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Fragaszy, D. M., Biro, D., Eshchar, Y., Humle, T., Izar, P., Resende, B., & Visalberghi, E. (2013). The fourth dimension of tool use: temporally enduring artefacts aid primates learning to use tools. *Philosophical Transactions of the Royal Society B-Biological Sciences, 368*(1630). doi:10.1098/rstb.2012.0410

Galef, B. G., & Laland, K. N. (2005). Social learning in animals: empirical studies and theoretical models. *American Institute of Biological Sciences bulletin, 55*(6), 489-499.

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature, 415*(6873), 755.

German, T. P., & Barrett, H. C. (2005). Functional fixedness in a technologically sparse culture. *Psychological Science, 16*(1), 1-5. doi:10.1111/j.0956-7976.2005.00771.x

German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review, 7*(4), 707-712. doi:10.3758/bf03213010

German, T. P., & Johnson, S. C. (2002). Function and the origins of the design stance. *Journal of Cognition and Development, 3*(3), 279-300. doi:10.1207/s15327647jcd0303_2

German, T. P., Truxaw, D., & Defeyter, M. A. (2007). The role of information about "convention," "design," and "goal" in representing artificial kinds. *New Directions for Child and Adolescent Development*, (115), 69-81.

Gibson, E. L. (1982). The concept of affordances in development: the renascence of functionalism. In W. A. Collins (Ed.), *The concept of development: the Minnesota symposia on child psychology* (Vol. 15).

Girndt, A., Meier, T., & Call, J. (2008). Task constraints mask great apes' ability to solve the trap-table task. *Journal of Experimental Psychology-Animal Behavior Processes, 34*(1), 54-62. doi:10.1037/0097-7403.34.1.54

Glickman, S. E., & Sroges, R. W. (1966). Curiosity in zoo animals. *Behaviour, 26*, 151-&. doi:10.1163/156853966x00074

Glucksberg, S. (1964). Problem solving - response competition and the influence of drive. *Psychological Reports, 15*(3), 939-942.

Glucksberg, S., & Weisberg, R. W. (1966). Verbal behavior and problem solving: some effects of labeling in a functional fixedness problem. *Journal of Experimental Psychology, 71*(5), 659-664.

Greif, M. L., & Needham, A. (2011). The development of human tool use in early life. In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool use and causal cognition* (pp. 51-68). Oxford: Oxford University Press.

Gruber, T., Muller, M. N., Reynolds, V., Wrangham, R., & Zuberbuhler, K. (2011). Community-specific evaluation of tool affordances in wild chimpanzees. *Scientific Reports, 1*, 128. doi:10.1038/srep00128

Gruber, T., Muller, M. N., Strimling, P., Wrangham, R., & Zuberbuehler, K. (2009). Wild chimpanzees rely on cultural knowledge to solve an experimental honey acquisition task. *Current Biology, 19*(21), 1806-1810. doi:10.1016/j.cub.2009.08.060

Guillaume, P., & Meyerson, I. (1930). Recherches sur l'usage de l'instrument chez les singes. I. Le probleme du detour. *Journal de Psychologie Normale et Pathologique, 27*, 177–236.

Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology, 18*(9), R370-R372. doi:10.1016/j.cub.2008.02.039

Hanus, D., & Call, J. (2011). Chimpanzee problem-solving: contrasting the use of causal and arbitrary cues. *Animal Cognition, 14*(6), 871-878. doi:10.1007/s10071-011-0421-6

Hanus, D., Mendes, N., Tennie, C., & Call, J. (2011). Comparing the performances of apes (*Gorilla gorilla, Pan troglodytes, Pongo pygmaeus*) and human children (*Homo sapiens*) in the floating peanut task. *PLoS One, 6*(6), e19555. doi:10.1371/journal.pone.0019555

Harrison, R. A., & Whiten, A. (2018). Chimpanzees (*Pan troglodytes*) display limited behavioural flexibility when faced with a changing foraging task requiring tool use. *PeerJ, 6*. doi:10.7717/peerj.4366

Hawkes, K., O'Connell, F., & Blurton Jones, N. G. (1995). Hadza children's foraging: juvenile dependency, social arrangements, and mobility among hunter-gatherers. *Current Anthropology, 36*(4), 688-700.

Hernik, M., & Csibra, G. (2009). Functional understanding facilitates learning about tools in human children. *Current Opinion in Neurobiology, 19*(1), 34-38. doi:10.1016/j.conb.2009.05.003

Hernik, M., & Csibra, G. (2015). Infants learn enduring functions of novel tools from action demonstrations. *Journal of Experimental Child Psychology, 130*, 176-192.

Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science, 317*(5843), 1360-1366. doi:10.1126/science.1146282

Herrmann, E., Wobber, V., & Call, J. (2008). Great apes' (*Pan troglodytes, Pan paniscus, Gorilla gorilla, Pongo pygmaeus*) understanding of tool functional properties after limited experience. *Journal of Comparative Psychology, 122*(2), 220-230. doi:10.1037/0735-7036.122.2.220

Heyes, C. M. (1994). Social learning in animals: categories and mechanisms. *Biological Reviews, 69*(2), 207-231.

Hirata, S., Morimura, N., & Houki, C. (2009). How to crack nuts: Acquisition process in captive chimpanzees (*Pan troglodytes*) observing a model. *Animal Cognition, 12*, S87-S101. doi:10.1007/s10071-009-0275-3

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan trogiodytes*) and children (*Homo sapiens*). *Animal Cognition, 8*(3), 164-181. doi:10.1007/s10071-004-0239-6

Hothorn, T., & Hornik, K. (2015). Exact rank tests: exact distributions for rank and permutation tests. R package version 0.8-28. Retrieved from http://CRAN.R-project.org/package=exactRankTests

Hrubesch, C., Preuschoft, S., & van Schaik, C. (2009). Skill mastery inhibits adoption of observed alternative solutions among chimpanzees (*Pan troglodytes*). *Animal Cognition, 12*(2), 209-216. doi:10.1007/s10071-008-0183-y

Huang, C.-T., Heyes, C., & Charman, T. (2002). Infants' behavioral reenactment of "failed attempts": Exploring the roles of emulation learning, stimulus enhancement, and understanding of intentions. *Developmental Psychology, 38*(5), 840-855. doi:10.1037/0012-1649.38.5.840

Huebner, F., & Fichtel, C. (2015). Innovation and behavioral flexibility in wild redfronted lemurs (*Eulemur rufifrons*). *Animal Cognition, 18*(3), 777-787. doi:10.1007/s10071-015-0844-6

Hunley, S. B., & Hahn, E. R. (2016). Labels affect preschoolers' tool-based scale errors. *Journal of Experimental Child Psychology, 151*, 40-50. doi:10.1016/j.jecp.2016.01.007

Jackson, T. A. (1942). Use of the stick as a tool by young chimpanzees. *Journal of Comparative Psychology, 34*(2), 223.

Janmaat, K. R. L., Ban, S. D., & Boesch, C. (2013). Chimpanzees use long-term spatial memory to monitor large fruit trees and remember feeding experiences across seasons. *Animal Behaviour, 86*(6), 1183-1205. doi:10.1016/j.anbehav.2013.09.021

Janmaat, K. R. L., Boesch, C., Byrne, R., Chapman, C. A., Bi, Z. B. G., Head, J. S., . . . Polansky, L. (2016). Spatio-temporal complexity of chimpanzee food: how cognitive adaptations can counteract the ephemeral nature of ripe fruit. *American Journal of Primatology, 78*(6), 626-645. doi:10.1002/ajp.22527

Janmaat, K. R. L., Polansky, L., Ban, S. D., & Boesch, C. (2014). Wild chimpanzees plan their breakfast time, type, and location. *Proceedings of the National Academy of Sciences of the United States of America, 111*(46), 16343-16348. doi:10.1073/pnas.1407524111

Jolly, A. (1964a). Choice of cue in prosimian learning. *Animal Behaviour, 12*(4), 571-577.

Jolly, A. (1964b). Prosimians' manipulation of simple object problems. *Animal Behaviour, 12*(4), 560-570.

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development, 45*(1), 65-94. doi:10.1007/bf02299613

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*(4), 63-85. doi:https://doi.org/10.1007/BF02300500

Jordan, C. (1982). Object manipulation and tool-use in captive pygmy chimpanzees (*Pan paniscus*). *Journal of Human Evolution, 11*(1), 35-39.

Kano, F., & Hirata, S. (2015). Great apes make anticipatory looks based on long-term memory of single events. *Current Biology, 25*(19), 2513-2517. doi:10.1016/j.cub.2015.08.004

Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2016). Differing views: Can chimpanzees do Level 2 perspective-taking? *Animal Cognition, 19*(3), 555-564. doi:10.1007/s10071-016-0956-7

Kelemen, D. (1999). Function, goals and intention: children's teleological reasoning about objects. *Trends in Cognitive Sciences, 3*(12), 461-468. doi:Doi 10.1016/S1364-6613(99)01402-3

Kemler Nelson, D. G., Frankenfield, A., Morris, C., & Blair, E. (2000). Young children's use of functional information to categorize artifacts: three factors that matter. *Cognition, 77*(2), 133-168.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition, 29*(7), 1000-1009. doi:10.3758/bf03195762

Köhler, W. (1925). *The mentality of apes*. London: Kegan Paul, Trench, Trubner [Original work published: Intelligenzprüfungen an Anthropoiden (1917). Königliche Akademie der Wissenschaften: Berlin].

Koops, K., McGrew, W. C., & Matsuzawa, T. (2013). Ecology of culture: do environmental factors influence foraging tool use in wild chimpanzees, *Pan troglodytes verus*? *Animal Behaviour, 85*(1), 175-185. doi:https://doi.org/10.1016/j.anbehav.2012.10.022

Koton, P. A. (1988). *Using experience in learning and problem solving.* (Ph.D. Thesis.), Massachusetts Institute of Technology.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science, 354*(6308), 110-114.

Kuroshima, H., Kaiser, I., & Fragaszy, D. M. (2014). Does own experience affect perception of others' actions in capuchin monkeys (*Cebus apella*)? *Animal Cognition, 17*(6), 1269-1279. doi:10.1007/s10071-014-0760-1

Laumer, I. B., Bugnyar, T., Reber, S. A., & Auersperg, A. M. I. (2017). Can hook-bending be let off the hook? Bending/unbending of pliant tools by cockatoos. *Proceedings of the Royal Society B: Biological Sciences, 284*(1862). doi:10.1098/rspb.2017.1026

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in Cognitive Sciences, 19*(11), 688-699.

Lethmate, J. (1977). *Problemlöseverhalten von Orang-Utans (Pongo pygmaeus)* (Vol. 19). Berlin; Hamburg: Paul Parey.

Lewis, A., Call, J., & Berntsen, D. (2017). Non-goal-directed recall of specific events in apes after long delays. *Proceedings of the Royal Society B: Biological Sciences, 284*(1858). doi:10.1098/rspb.2017.0518

Lin, W. L., Tsai, P. H., Lin, H. Y., & Chen, H. C. (2014). How does emotion influence different creative performances? The mediating role of cognitive flexibility. *Cognition & Emotion, 28*(5), 834-844. doi:10.1080/02699931.2013.854195

Lockman, J. J. (2000). A perception–action perspective on tool use development. *Child Development, 71*(1), 137-144.

Lonsdorf, E. V. (2006). What is the role of mothers in the acquisition of termite-fishing behaviors in wild chimpanzees (*Pan troglodytes schweinfurthii*)? *Animal Cognition, 9*(1), 36-46. doi:10.1007/s10071-005-0002-7

Luchins, A. S., & Luchins, E. H. (1959). Rigidity of behavior. *Eugene, OR: University of Oregon Books*.

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences of the United States of America, 104*(50), 19751-19756. doi:10.1073/pnas.0704452104

Maier, N. R. F. (1931). Reasoning in humans. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology, 12*(2), 181-194.

Manrique, H. M., & Call, J. (2011). Spontaneous use of tools as straws in great apes. *Animal Cognition, 14*(2), 213-226. doi:10.1007/s10071-010-0355-4

Manrique, H. M., & Call, J. (2015). Age-dependent cognitive inflexibility in great apes. *Animal Behaviour, 102*, 1-6. doi:10.1016/j.anbehav.2015.01.002

Manrique, H. M., Gross, A. N.-M., & Call, J. (2010). Great apes select tools on the basis of their rigidity. *Journal of Experimental Psychology-Animal Behavior Processes, 36*(4), 409-422. doi:10.1037/a0019296

Manrique, H. M., Völter, C. J., & Call, J. (2013). Repeated innovation in great apes. *Animal Behaviour, 85*(1), 195-202. doi:10.1016/j.anbehav.2012.10.026

Marshall-Pescini, S., & Whiten, A. (2008). Chimpanzees (*Pan troglodytes*) and the question of cumulative culture: An experimental approach. *Animal Cognition, 11*(3), 449-456. doi:10.1007/s10071-007-0135-y

Martin-Ordas, G., Atance, C. M., & Call, J. (2014). Remembering in tool-use tasks in children and apes: The role of the information at encoding. *Memory, 22*(1), 129-144. doi:10.1080/09658211.2013.806553

Martin-Ordas, G., Berntsen, D., & Call, J. (2013). Memory for distant past events in chimpanzees and orangutans. *Current Biology, 23*(15), 1438-1441. doi:10.1016/j.cub.2013.06.017

Martin-Ordas, G., Haun, D. B. M., Colmenares, F., & Call, J. (2010). Keeping track of time: Evidence for episodic-like memory in great apes. *Animal Cognition, 13*(2), 331-340. doi:10.1007/s10071-009-0282-4

Martin-Ordas, G., Schumacher, L., & Call, J. (2012). Sequential Tool Use in Great Apes. *PLoS One, 7*(12). doi:10.1371/journal.pone.0052074

Mayr, E. (2002). *What evolution is*. London: Phoenix.

McCarty, M. E., Clifton, R. K., & Collard, R. R. (1999). Problem solving in infancy: The emergence of an action plan. *Developmental Psychology, 35*(4), 1091-1101. doi:10.1037/0012-1649.35.4.1091

McCarty, M. E., Clifton, R. K., & Collard, R. R. (2001). The beginnings of tool use by infants and toddlers. *Infancy, 2*(2), 233-256. doi:doi:10.1207/S15327078IN0202_8

McCormack, T., Hoerl, C., & Butterfill, S. (Eds.). (2011). *Tool use and causal cognition*. Oxford: Oxford University Press.

McGrew, W. C. (1992). *Chimpanzee material culture: implications for human evolution*. Cambridge: Cambridge University Press.

McGuigan, N., Makinson, J., & Whiten, A. (2011). From over-imitation to super-copying: Adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology, 102*(1), 1-18. doi:doi:10.1348/000712610X493115

McGuigan, N., & Whiten, A. (2009). Emulation and “overemulation” in the social learning of causally opaque versus causally transparent tool use by 23- and 30-month-olds. *Journal of Experimental Child Psychology, 104*(4), 367-381. doi:https://doi.org/10.1016/j.jecp.2009.07.001

McGuigan, N., Whiten, A., Flynn, E., & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3-and 5-year-old children. *Cognitive Development, 22*(3), 353-364. doi:10.1016/j.cogdev.2007.01.001

Mendes, N., & Call, J. (2014). Chimpanzees form long-term memories for food locations after limited exposure. *American Journal of Primatology, 76*(5), 485-495. doi:10.1002/ajp.22248

Mendes, N., Hanus, D., & Call, J. (2007). Raising the level: orangutans use water as a tool. *Biology Letters, 3*(5), 453-455. doi:10.1098/rsbl.2007.0198

Meulman, E. J. M., Sanz, C. M., Visalberghi, E., & van Schaik, C. P. (2012). The role of terrestriality in promoting primate technology. *Evolutionary Anthropology: Issues, News, and Reviews, 21*(2), 58-68.

Meulman, E. J. M., Seed, A. M., & Mann, J. (2013). If at first you don't succeed… Studies of ontogeny shed light on the cognitive demands of habitual tool use. *Philosophical Transactions of the Royal Society B-Biological Sciences, 368*(1630), 20130050.

Moll, H., & Tomasello, M. (2007). Cooperation and human cognition: the Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society B-Biological Sciences, 362*(1480), 639-648. doi:10.1098/rstb.2006.2000

Mulcahy, N. J., & Call, J. (2006). Apes save tools for future use. *Science, 312*(5776), 1038-1040. doi:10.1126/science.1125456

Musgrave, S., Morgan, D., Lonsdorf, E., Mundry, R., & Sanz, C. M. (2016). Tool transfers are a form of teaching among chimpanzees. *Scientific Reports, 6*, 34783. doi:10.1038/srep34783

Nater, A., Mattle-Greminger, M. P., Nurcahyo, A., Nowak, M. G., de Manuel, M., Desai, T., . . . Roos, C. (2017). Morphometric, behavioral, and genomic evidence for a new Orangutan species. *Current Biology, 27*(22), 3487-3498. e3410.

Nehaniv, C. L., & Dautenhahn, K. E. (2007). *Imitation and social learning in robots, humans and animals: behavioural, social and communicative dimensions*: Cambridge University Press.

Nelson, K. (1973). Some evidence for the cognitive primacy of categorization and its functional basis. *Merrill-Palmer Quarterly of Behavior and Development, 19*(1), 21-39.

Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology, 42*(3), 555-565. doi:Doi 10.1037/0012-1649.42.3.555

Nielsen, M. (2013). Young children's imitative and innovative behaviour on the floating object task. *Infant and Child Development*(22), 44-52. doi:10.1002/icd

Nielsen, M., & Haun, D. (2016). Why developmental psychology is incomplete without comparative and cross-cultural perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1686). doi:10.1098/rstb.2015.0071

Osvath, M., & Osvath, H. (2008). Chimpanzee (*Pan troglodytes*) and orangutan (*Pongo abelii*) forethought: self-control and pre-experience in the face of future tool use. *Animal Cognition, 11*(4), 661-674. doi:10.1007/s10071-008-0157-0

Parker, S. T., & Gibson, K. R. (1977). Object manipulation, tool use and sensorimotor intelligence as feeding adaptations in cebus monkeys and great apes. *Journal of Human Evolution, 6*(7), 623-641. doi:10.1016/s0047-2484(77)80135-8

Péter, A. (2011). Solomon Coder (version beta 12.09.04): A simple solution for behavior coding. Retrieved from http://solomoncoder.com/

Piaget, J. (1977). *La naissance de l'intelligence chez l'enfant* (9th ed. ed.). Neuchâtel; Paris: Delachaux & Niestlé S.A. [First published: 1936].

Poti, P., & Spinozzi, G. (1994). Early sensorimotor development in chimpanzees (*Pan Troglodytes*). *Journal of Comparative Psychology, 108*(1), 93-103. doi:Doi 10.1037/0735-7036.108.1.93

Price, E. E., Lambeth, S. P., Schapiro, S. J., & Whiten, A. (2009). A potent effect of observational learning on chimpanzee tool construction. *Proceedings of the Royal Society B: Biological Sciences, 276*(1671), 3377-3383. doi:10.1098/rspb.2009.0640

R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Redshaw, J., & Suddendorf, T. (2016). Children's and apes' preparatory responses to two mutually exclusive possibilities. *Current Biology, 26*(13), 1758-1762. doi:10.1016/j.cub.2016.04.062

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology, 6*(3), 436-450. doi:https://doi.org/10.1016/0010-0285(74)90020-6

Renner, E., Abramo, A. M., Karen Hambright, M., & Phillips, K. A. (2017). Insightful problem solving and emulation in brown capuchin monkeys. *Animal Cognition*, 1-6. doi:10.1007/s10071-017-1080-z

Roche, H., Blumenschine, R. J., & Shea, J. J. (2009). Origins and adaptations of early Homo: what archeology tells us. In F. E. Grine, J. G. Fleagle, & R. E. Leakey (Eds.), *The First Humans – Origin and Early Evolution of the Genus Homo: Contributions from the Third Stony Brook Human Evolution Symposium and Workshop October 3 – October 7, 2006* (pp. 135-147). Dordrecht: Springer Netherlands.

Rosati, A. G., Stevens, J. R., Hare, B., & Hauser, M. D. (2007). The evolutionary origins of human patience: temporal preferences in chimpanzees, bonobos, and human adults. *Current Biology, 17*(19), 1663-1668. doi:10.1016/j.cub.2007.08.033

Rothi, L. J. G., Ochipa, C., & Heilman, K. M. (1991). A Cognitive Neuropsychological Model of Limb Praxis. *Cognitive Neuropsychology, 8*(6), 443-458. doi:10.1080/02643299108253382

Ruiz, A. M., & Santos, L. R. (2013). Understanding differences in the way human and non-human primates represent tools: The role of teleological-intentional information. In C.

M. Sanz, J. Call, & C. Boesch (Eds.), *Tool use in animals. Cognition and ecology*. Cambridge: Cambridge University Press.

Santos, L. R., & Hood, B. M. (2009). Object representation as a central issue in cognitive science. In L. R. Santos & B. M. Hood (Eds.), *The origin of object knowledge*. Oxford: Oxford University Press.

Sanz, C. M., Call, J., & Boesch, C. (Eds.). (2013). *Tool use in animals. Cognition and ecology*. Cambridge: Cambridge University Press.

Sanz, C. M., Call, J., & Morgan, D. (2009). Design complexity in termite-fishing tools of chimpanzees (*Pan troglodytes*). *Biology Letters*. doi:10.1098/rsbl.2008.0786

Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. *Behavioral Ecology, 20*(2), 416-420. doi:10.1093/beheco/arn145

Schultz, P. W., & Searleman, A. (1998). Personal need for structure, the Einstellung task, and the effects of stress. *Personality and Individual Differences, 24*(3), 305-310.

Schultz, P. W., & Searleman, A. (2002). Rigidity of thought and behavior: 100 years of research. *Genetic, social, and general psychology monographs, 128*(2), 165.

Schultz, P. W., Stone, W. F., & Christie, R. (1997). Authoritarianism and mental rigidity: the Einstellung problem revisited. *Personality and Social Psychology Bulletin, 23*(1), 3-9. doi:10.1177/0146167297231001

Schultz, R. W. (1960). Problem solving behavior and transfer. *Harvard Educational Review, 30*, 61-77.

Seed, A., & Byrne, R. (2010). Animal tool-use. *Current Biology, 20*(23), R1032-R1039.

Seed, A., Call, J., Emery, N. J., & Clayton, N. S. (2009). Chimpanzees solve the trap problem when the confound of tool-use is removed. *Journal of Experimental Psychology-Animal Behavior Processes, 35*(1), 23-34. doi:10.1037/a0012925

Seed, A., & Mayer, C. (2017). Problem solving. In J. Call, G. M. Burghardt, I. M. Pepperberg, C. T. Snowdon, & T. Zentall (Eds.), *APA Handbook of comparative psychology* (Vol. 2, pp. 601-625). Washington: American Psychological Association.

Shen, W. B., Yuan, Y., Liu, C., & Luo, J. (2016). In search of the "Aha!' experience: Elucidating the emotionality of insight problem-solving. *British Journal of Psychology, 107*(2), 281-298. doi:10.1111/bjop.12142

Sheridan, H., & Reingold, E. M. (2013). The mechanisms and boundary conditions of the Einstellung effect in chess: Evidence from eye movements. *PLoS One, 8*(10), e75796. doi:10.1371/journal.pone.0075796

Sheridan, H., & Reingold, E. M. (2014). Expert vs. novice differences in the detection of relevant information during a chess game: evidence from eye movements. *Frontiers in psychology, 5*. doi:10.3389/fpsyg.2014.00941

Shettleworth, S. J. (2012). Do animals have insight, and what is insight anyway? *Canadian Journal of Experimental Psychology, 66*(4), 217-226. doi:10.1037/a0030674

Shumaker, R. W., Walkup, K. R., & Beck, B. B. (2011). *Animal tool behavior. The use and manufacture of tools by animals* (Revised and updated ed.). Baltimore: The Johns Hopkins University Press.

Siegel, D. R., & Callanan, M. A. (2007). Artifacts as conventional objects. *Journal of Cognition and Development, 8*(2), 183-203. doi:Doi 10.1080/15248370701202414

Silvertown, J. (2008). *99% ape: how evolution adds up*. Chicago: The University of Chicago Press.

Simon, H. A., & Newell, A. (1971). Human problem solving: the state of the theory in 1970. *American Psychologist, 26*(2), 145-159. doi:10.1037/h0030806

Smitsman, A. W., & Cox, R. F. A. (2008). Perseveration in tool use: A window for understanding the dynamics of the action-selection process. *Infancy, 13*(3), 249-269. doi:doi:10.1080/15250000802004379

Sommerville, J. A., Hildebrand, E. A., & Crane, C. C. (2008). Experience matters: The impact of doing versus watching on infants' subsequent perception of tool-use events. *Developmental Psychology, 44*(5), 1249-1256. doi:10.1037/a0012296

Spagnoletti, N., Visalberghi, E., Verderane, M. P., Ottoni, E., Izar, P., & Fragaszy, D. (2012). Stone tool use in wild bearded capuchin monkeys, *Cebus libidinosus*. Is it a strategy to overcome food scarcity? *Animal Behaviour, 83*(5), 1285-1294. doi:https://doi.org/10.1016/j.anbehav.2012.03.002

Star, J. R., & Seifert, C. (2006). The development of flexibility in equation solving. *Contemporary Educational Psychology, 31*(3), 280-300. doi:http://dx.doi.org/10.1016/j.cedpsych.2005.08.001

Sternberg, R. J., & Davidson, J. E. (Eds.). (1995). *The nature of insight*. Cambridge, Massachusetts; London, England: MIT Press.

Stout, D., & Chaminade, T. (2007). The evolutionary neuroscience of tool making. *Neuropsychologia, 45*(5), 1091-1100. doi:https://doi.org/10.1016/j.neuropsychologia.2006.09.014

Takeshita, H., & Walraven, V. (1996). A comparative study of the variety and complexity of object manipulation in captive chimpanzees (*Pan troglodytes*) and Bonobos (*Pan paniscus*). *Primates, 37*(4), 423-441. doi:10.1007/bf02381377

Taniguchi, Y., & Sanefuji, W. (2017). The boundaries of overimitation in preschool children: Effects of target and tool use on imitation of irrelevant actions. *Journal of Experimental Child Psychology, 159*, 83-95. doi:10.1016/j.jecp.2017.01.014

Taylor, A. H., Elliffe, D., Hunt, G. R., & Gray, R. D. (2010). Complex cognition and behavioural innovation in New Caledonian crows. *Proceedings of the Royal Society B: Biological Sciences, 277*(1694), 2637-2643. doi:10.1098/rspb.2010.0285

Taylor, A. H., Medina, F. S., Holzhaider, J. C., Hearne, L. J., Hunt, G. R., & Gray, R. D. (2010). An investigation into the cognition behind spontaneous string pulling in New Caledonian crows. *PLoS One, 5*(2), e9345. doi:10.1371/journal.pone.0009345

Tennie, C., Call, J., & Tomasello, M. (2006). Push or pull: Imitation vs. Emulation in great apes and human children. *Ethology, 112*(12), 1159-1169. doi:10.1111/j.1439-0310.2006.01269.x

Tennie, C., Call, J., & Tomasello, M. (2010). Evidence for emulation in chimpanzees in social settings using the floating peanut task. *PLoS One, 5*(5), e10544. doi:10.1371/journal.pone.0010544

Tennie, C., Greve, K., Gretscher, H., & Call, J. (2010). Two-year-old children copy more reliably and more often than nonhuman great apes in multiple observational learning tasks. *Primates, 51*(4), 337-351. doi:10.1007/s10329-010-0208-4

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences, 24*(1), 1-86. doi:10.1017/S0140525X01003910

Therneau, T. M. (2015). A package for survival analysis in S. R package version 2.38. Retrieved from http://CRAN.R-project.org/package=survival

Therneau, T. M., & Grambsch, P. M. (2000). *Modeling survival data: extending the cox model*. New York: Springer.

Tomasello, M. (2014). *A natural history of human thinking*: Harvard University Press.

Tomasello, M., & Call, J. (1997). *Primate cognition*. New York; Oxford: Oxford University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675+. doi:10.1017/s0140525x05000129

Torigoe, T. (1985). Comparison of object manipulation among 74 species of non-human primates. *Primates, 26*(2), 182-194. doi:10.1007/bf02382017

Uher, J., & Asendorpf, J. B. (2008). Personality assessment in the Great Apes: Comparing ecologically valid behavior measures, behavior ratings, and adjective ratings. *Journal of Research in Personality, 42*(4), 821-838. doi:10.1016/j.jrp.2007.10.004

Uher, J., Asendorpf, J. B., & Call, J. (2008). Personality in the behaviour of great apes: temporal stability, cross-situational consistency and coherence in response. *Animal Behaviour, 75*, 99-112. doi:10.1016/j.anbehav.2007.04.018

Vaesen, K. (2012). The cognitive bases of human tool use. *Behavioral and Brain Sciences, 35*(4), 203-+. doi:10.1017/s0140525x11001452

Vale, G. L., Davis, S. J., Lambeth, S. P., Schapiro, S. J., & Whiten, A. (2017). Acquisition of a socially learned tool use sequence in chimpanzees: Implications for cumulative culture. *Evolution and Human Behavior, 38*(5), 635-644. doi:10.1016/j.evolhumbehav.2017.04.007

Vale, G. L., Flynn, E. G., Pender, L., Price, E., Whiten, A., Lambeth, S. P., . . . Kendal, R. L. (2016). Robust retention and transfer of tool construction techniques in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 130*(1), 24-35. doi:10.1037/a0040000

van Elk, M., van Schie, H., & Bekkering, H. (2014). Action semantics: A unifying conceptual framework for the selective use of multimodal and modality-specific object knowledge. *Physics of Life Reviews, 11*(2), 220-250. doi:10.1016/j.plrev.2013.11.005

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., . . . Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science, 299*(5603), 102-105. doi:10.1126/science.1078004

van Schaik, C. P., Damerius, L., & Isler, K. (2013). Wild Orangutan Males Plan and Communicate Their Travel Direction One Day in Advance. *PLoS One, 8*(9). doi:10.1371/journal.pone.0074896

Vauclair, J. (1984). Phylogenetic approach to object manipulation in human and ape infants. *Human Development, 27*(5-6), 321-328.

Visalberghi, E., Fragaszy, D. M., Izar, P., & Ottoni, E. B. (2005). Terrestriality and tool use. *Science, 308*(5724), 951-952. doi:10.1126/science.308.5724.951c

Visalberghi, E., Fragaszy, D. M., & Savage-Rumbaugh, S. (1995). Performance in a tool-using task by common chimpanzees (*Pan troglodytes*), bonobos (*Pan paniscus*), an orangutan (*Pongo pygmaeus*), and capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology, 109*(1), 52.

Visalberghi, E., & Trinca, L. (1989). Tool use in capuchin monkeys - distinguishing between performing and understanding. *Primates, 30*(4), 511-521. doi:10.1007/bf02380877

Vlamings, P. H., Hare, B., & Call, J. (2010). Reaching around barriers: the performance of the great apes and 3-5-year-old children. *Animal Cognition, 13*(2), 273-285. doi:10.1007/s10071-009-0265-5

Vlamings, P. H., Uher, J., & Call, J. (2006). How the great apes (*Pan troglodytes, Pongo pygmaeus, Pan paniscus*, and *Gorilla gorilla*) perform on the reversed contingency task: the effects of food quantity and food visibility. *Journal of Experimental Psychology-Animal Behavior Processes, 32*(1), 60-70. doi:10.1037/0097-7403.32.1.60

Völter, C. J., & Call, J. (2012). Problem solving in great apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla*, and *Pongo abelii*): the effect of visual feedback. *Animal Cognition, 15*(5), 923-936. doi:DOI 10.1007/s10071-012-0519-5

Völter, C. J., & Call, J. (2014a). The cognitive underpinnings of flexible tool use in great apes. *Journal of Experimental Psychology-Animal Behavior Processes, 40*(3), 287-302. doi:10.1037/xan0000025

Völter, C. J., & Call, J. (2014b). Younger apes and human children plan their moves in a maze task. *Cognition, 130*(2), 186-203. doi:10.1016/j.cognition.2013.10.007

Völter, C. J., Sentís, I., & Call, J. (2016). Great apes and children infer causal relations from patterns of variation and covariation. *Cognition, 155*, 30-43. doi:https://doi.org/10.1016/j.cognition.2016.06.009

Weir, A. A. S., Chappell, J., & Kacelnik, A. (2002). Shaping of hooks in new Caledonian crows. *Science, 297*(5583), 981-981. doi:10.1126/science.1073433

Welker, W. I. (1956). Some determinants of play and exploration in chimpanzees. *Journal of Comparative and Physiological Psychology, 49*(1), 84-89. doi:10.1037/h0044463

Whiten, A., & Boesch, C. (2001). The cultures of chimpanzees. *Scientific American, 284*(1), 60-67.

Whiten, A., Custance, D. M., Gomez, J. C., Teixidor, P., & Bard, K. A. (1996). Imitative learning of artificial fruit processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 110*(1), 3-14. doi:10.1037//0735-7036.110.1.3

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1528), 2417-2428. doi:10.1098/rstb.2009.0069

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Developmental Review, 33*(4), 341-356. doi:10.1016/j.dr.2013.08.002

Worgotter, F., Aksoy, E. E., Kruger, N., Piater, J., Ude, A., & Tamosiunaite, M. (2013). A simple ontology of manipulation actions based on hand-object relations. *IEEE Transactions on Autonomous Mental Development, 5*(2), 117-134. doi:10.1109/tamd.2012.2232291

Yonge, G. D. (1966). Structure of experience and functional fixedness. *Journal of Educational Psychology, 57*(2), 115-120.

Appendices

Appendix A. The detailed procedure of the studies reported in Chapter 4.

1.a The experimenter enters the test room with the child. In the five plants and the one plant condition, she asks her to water the plant(s).

> E: "Oh yes, I have forgotten to water the plants. Could you maybe do that? (…) See, with this cup here you can water the plant(s) a little. (…) Right, each plant just a bit. (…) It doesn't matter if you spill some water. (…) Great! Now we can play the game!"

1.b In the zero plants condition, the experimenter asks the child to carry in the bucket with water and to place it onto the yellow mat.

> E: "Oh yes, the bucket. Perhaps you could carry it inside? (…) You can place the bucket onto the yellow mat over there. (…) Great! Now we can play the game!"

2. The experimenter retrieves the pirate ship from the white sheet and explains the game to the child which consists of placing three balls into the ship.

> E: "Look, I have brought this ship with me. Do you know what type of ship this is? [answer] Yes, exactly, this is a pirate ship! And the game works the following way: you can collect and place three balls inside the ship. The blue ball goes to the blue position. And the red ball goes – do you know where the red ball goes? [answer] Exactly! And which one is missing? [answer] Right! So… here's the ship. [E hand over ship to the child] And if you succeed in placing the three balls

inside the ship, you will get a surprise. Do you have any further questions? [answer] You are allowed to try out whatever comes to your mind. Then we can get started – and go!"

The experimenter removes the white sheet from the table and sits down at the corner of the room.

E: "I have some work to do. I will sit down over there."

The child has five minutes time to collect the three balls. The experimenter states one of two motivating sentences (M1, M2) about every minute. Sometimes a motivating sentence is a little postponed, e.g., if such a sentence was just stated (see below).

M1, E: "Just try out another thing! Maybe you have another idea?"
(minute 1 & 3)

M2, E: "You can try out whatever comes to your mind." (minute 2 & 4)

If a child asks whether she is allowed to use the water or states the idea of pouring water into the tube, the experimenter will repeat that the child can try out whatever comes to her mind.

3.a <u>If the child solves the task</u> by pouring water into the tube until she reaches the blue ball, the experimenter will go over, praise her and ask her how she has come up with the idea.

E: "Super! (…) Great! (…) Wow… (…) And how did you come up with this idea?"

3.b If the child does not solve the task during the five minute session, the experimenter will go over and ask her if she has any further ideas. If she has not, the experimenter will tell her that this was a really difficult puzzle and that she has brought yet another puzzle from which the child can retrieve a blue ball.

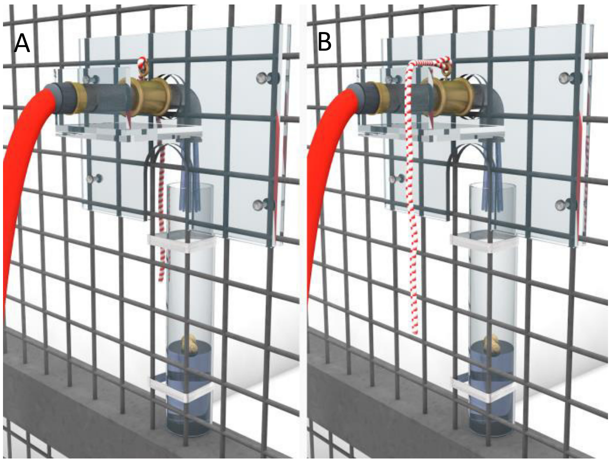
> E: "So (…) do you have any further ideas? [answer] But do you know what? This is a really difficult puzzle… Yet, I have brought another puzzle from which you can also get a blue ball. Do you want to try it?"

3.c If the child states the solution within the five minutes, but does not report it at the end when asked for further ideas (instead stays silent), the experimenter will refer to her previous statement. [This is done because children are generally hesitant to use the water and some children are additionally shy.] If the child does not refer to their previous idea after this question, she will receive the alternative puzzle and is coded as "unsuccessful".

> E: "Well, but you had another idea, right?"

4. Experiment 3 only: After five minutes have passed and children have been asked for further ideas they are given a hint. The experimenter pours water once with the cup inside the bucket stating "hmm". No eye contact is made and the action is done as incidentally as possible. Then, the experimenter tells the children to try again something out and sits down at the corner of the room. Children have one more minute to solve the task.

E: “Hmm… But do you know what, maybe you have another idea – I’ll just sit down over there for a short moment.”



Appendix B. The back view of the experimental setup from Chapter 3 (Experiment 2): The activated water tap in the “water tap by ape” (A) and “water tap by human” (B) condition is shown from the perspective of the experimenter.

Appendix C. Details about the models in Chapter 5.

Table C.1 Results of the models with manipulation time as the response (binomial and Gaussian) in Experiment 1.

Binomial model

	df	χ^2	*p*
Age(1)	1	16.69	**<0.001**
Sex (Female)	1	0.15	0.703
Species (Bonobo)	2	7.08	0.029
Session(1)	1	6.80	**0.009**
Group (PriorExp) x Area (Lower)	2	12.63	**0.002**
Manipulation type (Brush) x Area (Lower)	4	40.66	**<0.001**
Gaussian model			
Age(1)	1	0.55	0.459
Sex (Female)	1	0.73	0.393
Species (Bonobo)	2	2.14	0.344
Session(1)	1	0.88	0.347
Group (PriorExp) x Area (Lower) x Manipulation type (Brush)	4	11.73	**0.019**

(1) Log-transformed (age) and standardized to their respective means (age, session)

Table C.2 Results of the pairwise comparisons for the significant two-way interactions of the binomial model in Experiment 1, obtained by re-levelling the respective factors.

Posthoc: Manipulation type x Area(1)

		Manipulation type					
		Brush		Pointed		Touch	
		Est+/-SE	*p*	Est+/-SE	*p*	Est+/-SE	*p*
Area	Lower vs. Upper	5.977+/- 0.895	**<0.001**	2.388+/- 0.703	**0.001**	1.723+/- 0.733	**0.019**
	Lower vs. Other	3.513+/- 0.774	**<0.001**	-0.229+/- 0.620	0.712	-0.751+/- 0.605	0.215
	Upper vs. Other	-2.469+/- 0.715	**0.001**	-2.621+/- 0.704	**<0.001**	-2.474+/- 0.723	**0.001**

		Area					
		Lower		Upper		Other	
		Est+/-SE	*p*	Est+/-SE	*p*	Est+/-SE	*p*
Manipul type	Brush vs. Pointed	2.617+/- 0.790	**0.001**	-0.973+/- 0.681	0.153	-1.125+/- 0.650	***0.083***
	Brush vs. Touch	4.299+/- 0.775	**<0.001**	0.041+/- 0.657	0.950	0.036+/- 0.584	0.951
	Pointed vs. Touch	1.682+/- 0.735	**0.022**	1.015+/- 0.756	0.180	1.161+/- 0.737	0.115

Posthoc: Group x Area(2)

		Group			
		Prior exp.		No exp.	
		Est+/-SE	*p*	Est+/-SE	*p*
Area	Lower vs. Upper	5.977+/- 0.895	**<0.001**	3.107+/- 0.778	**<0.001**
	Lower vs. Other	3.513 +/- 0.774	**<0.001**	2.263+/- 0.707	**0.001**
	Upper vs. Other	-2.469+/- 0.715	**0.001**	-0.844+/- 0.639	0.187

		Area					
		Lower		Upper		Other	
		Est+/-SE	*p*	Est+/-SE	*p*	Est+/-SE	*p*
Group	Prior exp. vs. No exp.	1.284+/- 0.567	**0.023**	-1.595+/- 0.607	**0.009**	0.030+/- 0.510	0.953

(1) Estimates are given for the reference category of group set to “Prior experience”

(2) Estimates are given for the reference category of manipulation type set to “Brush”

Table C.3 Results of the pairwise comparisons for the significant three-way interaction of the Gaussian model in Experiment 1, obtained by subsetting the data by factor and examining the respective two-way interactions. When the two-way interaction was non-significant, it was excluded from the model (shown in brackets).

Post-hoc: Group x Area x Manipulation type

Subset by Manipulation type			
Brush	*df*	χ^2	*p*
(Group x Area	2	1.70	0.427)
Group	1	0.01	0.926
Area	2	22.42	**<0.001**
Wood			
Group x Area	2	9.68	**0.008**
Touch			
Group x Area	2	18.85	**<0.001**
Subset by Area			
Upper	*df*	χ^2	*p*
(Group x Manipulation type	2	0.07	0.966)
Group	1	1.80	0.180
Manipulation type	2	17.24	**<0.001**
Lower			
(Group x Manipulation type	2	1.53	0.465)
Group	1	18.78	**<0.001**
Manipulation	2	1.73	0.421
Other			
Group x Manipulation type	2	7.94	**0.019**

Table C.4 Results of the GLM with first trial data (Experiment 3).

Term	**Estimate**	**SE**	**lowerCL**	**upperCL**	χ^2	**Df**	**P**
Intercept	4.629	1.843	1.647	9.173	(4)	(4)	(4)
Group (pieces)(2)	-2.330	1.152	-4.893	-0.242	9.57(3)	2(3)	0.008(3)
Group (sticks)(2)	-3.453	1.375	-6.648	-1.076	(4)	(4)	(4)
Species (chimp)(2)	-3.769	1.702	-7.994	-0.999	9.66(3)	3(3)	0.021(3)
Species (gorilla)(2)	-3.232	2.924	-10.033	1.813	(4)	(4)	(4)
Species (orang)(2)	-1.622	1.738	-5.621	1.528	(4)	(4)	(4)
Grissini preference(1)	-0.689	0.565	-1.969	0.344	1.66	1	0.197

(1) Standardized to its mean

(2) Reference category: group (none), species (bonobo)

(3) Overall effect of the predictor (group, species)

(4) Not shown because of having a very limited interpretation

Table C.5 Results of the GLMM with all trials (Experiment 3).

Term	**Estimate**	**SE**	χ^2	**Df**	**P**
Intercept	10.589	4.588	(4)	(4)	(4)
Trial(1)	0.543	0.888	0.38	1	0.540
Age(1)	-0.491	1.887	0.08	1	0.774
Group (pieces)(2)	-2.324	3.973	0.74(3)	2(3)	0.690(3)
Group (sticks)(2)	-3.362	4.628	(4)	(4)	(4)
Species (chimp)(2)	-17.437	6.717	14.65(3)	3(3)	0.002(3)
Species (gorilla)(2)	-2.529	6.729	(4)	(4)	(4)
Species (orang)(2)	1.067	4.462	(4)	(4)	(4)
Grissini preference(1)	0.543	0.888	0.26	1	0.603

(1) Log-transformed (age), standardized to its mean (age, Grissini preference, trial)

(2) Reference category: group (none), species (bonobo)

(3) Overall effect of the predictor (group, species)

(4) Not shown because of having a very limited interpretation

Appendix D. Ethical approval letters.

University of St Andrews

School of Psychology & Neuroscience Ethics Committee

11 September 2015

Project Title:	The role of visual feedback in the floating peanut task
Researcher's Name:	Sonja Ebel
Supervisor:	Professor Josep Call

Thank you for submitting your application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 30th June 2015. The following documents were reviewed:

1. Animal Ethics Form	04/09/2015
2. Extended Details	04/09/2015
3. External Permissions	04/09/2015

The School of Psychology & Neuroscience Ethics Committee approves this study from an ethical point of view.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to the School Ethics Committee.

You must inform the School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to the School Ethics Committee to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an application for ethical amendment Form submitted where appropriate.

Approval is given on the understanding that the ASAB Guidelines for the Treatment of Animals in Behavioural Research and Teaching (published in Animal Behaviour, 2003, 65, 249-255, http://www.sciencedirect.com/) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof Josep Call (Supervisor)
School Ethics Committee
Dr Tamara Lawson (Home Office Liaison Officer)

University of St Andrews

School of Psychology & Neuroscience Ethics Committee

2 February 2016

Project Title:	Nonhuman great apes' performance in the floating peanut task using an opaque tube
Researcher's Name:	Sonja Ebel
Supervisor:	Professor Josep Call

Thank you for submitting your application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 26th January 2016. The following documents have been reviewed:

1.	Animal Ethics Form	02/02/2016
2.	Leipzig Protocol	02/02/2016

The School of Psychology & Neuroscience Ethics Committee approves this study from an ethical point of view.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to the School Ethics Committee.

You must inform the School Ethics Committee when the research has been completed. If you are unable to complete your research within the three year validation period, you will be required to write to the School Ethics Committee to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an application for ethical amendment Form submitted where appropriate.

Approval is given on the understanding that the ASAB Guidelines for the Treatment of Animals in Behavioural Research and Teaching (published in Animal Behaviour, 2003, 65, 249-255, http://www.sciencedirect.com/) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Professor Josep Call (Supervisor)
School Ethics Committee
Dr Tamara Lawson (Home Office Liaison Officer)

SEC Convenor, St Mary's Quad, St Andrews, Fife KY16 9JP, Scotland
Email: psyethics@st-andrews.ac.uk Tel: 01334 462071
The University of St Andrews is a charity registered in Scotland: No SC013532

University of St Andrews | FOUNDED 1413

University Teaching and Research Ethics Committee

14 March 2016

Dear Sonja

Thank you for submitting your ethical application which was considered at the School of Psychology & Neuroscience Ethics Committee meeting on 5th February 2016; the following documents have been reviewed:

1. Ethical Application Form
2. Study Design
3. Information and consent procedure at the MPI-EVA
4. MPI-EVA Flyer
5. Letter to Parents
6. Consent Form
7. Police Check
8. Data Management Plan

The School of Psychology & Neuroscience Ethics Committee has been delegated to act on behalf of the University Teaching and Research Ethics Committee (UTREC) and has granted this application ethical approval. The particulars relating to the approved project are as follows -

Approval Code:	PS11986	**Approved on:**	10/03/2016	**Approval Expiry:**	10/03/2021
Project Title:	Do children show a functional fixedness effect in the floating peanut task?				
Researcher:	Sonja Ebel				
Supervisor:	Professor Josep Call				

Approval is awarded for five years. Projects which have not commenced within two years of approval must be re-submitted for review by your School Ethics Committee. If you are unable to complete your research within the five year approval period, you are required to write to your School Ethics Committee Convener to request a discretionary extension of no greater than 6 months or to re-apply if directed to do so, and you should inform your School Ethics Committee when your project reaches completion.

If you make any changes to the project outlined in your approved ethical application form, you should inform your supervisor and seek advice on the ethical implications of those changes from the School Ethics Convener who may advise you to complete and submit an ethical amendment form for review.

Any adverse incident which occurs during the course of conducting your research must be reported immediately to the School Ethics Committee who will advise you on the appropriate action to be taken.

Approval is given on the understanding that you conduct your research as outlined in your application and in compliance with UTREC Guidelines and Policies (http://www.st-andrews.ac.uk/utrec/guidelinespolicies/). You are also advised to ensure that you procure and handle your research data within the provisions of the Data Provision Act 1998 and in accordance with any conditions of funding incumbent upon you.

Yours sincerely

Convener of the School Ethics Committee

cc Professor Josep Call (Supervisor)

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife KY16 9JP
Email: psyethics@st-andrews.ac.uk Tel: 01334 462071

University of St Andrews | FOUNDED 1413

University Teaching and Research Ethics Committee

24 March 2016

Dear Sonja

Thank you for submitting your amendment application which comprised the following documents:

1. Ethical Amendment Application Form

The School of Psychology & Neuroscience Ethics Committee is delegated to act on behalf of the University Teaching and Research Ethics Committee (UTREC) and has approved this ethical amendment application. The particulars of this approval are as follows –

Original Approval Code:	PS11986	**Approved on:**	10/03/16
Amendment Approval Date:	23/03/16	**Approval Expiry Date:**	10/03/2021
Project Title:	Do children show a functional fixedness effect in the floating peanut task?		
Researchers:	Sonja Ebel, Nadine Kanate	**Supervisor:**	Josep Call

Ethical amendment approval does not extend the originally granted approval period of three years, rather it validate the changes you have made to the originally approved ethical application. If you are unable to complete your research within the original five year validation period, you are required to write to your School Ethics Committee Convener to request a discretionary extension of no greater than 6 months or to re-apply if directed to do so, and you should inform your School Ethics Committee when your project reaches completion.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that you adhere to the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf).

Yours sincerely

Convener of the School Ethics Committee

cc Josep Call (Supervisor)

School of Psychology & Neuroscience, St Mary's Quad, South Street, St Andrews, Fife, KY16 9JP
Email: psyethics@st-andrews.ac.uk Tel: 01334 462071
The University of St Andrews is a charity registered in Scotland: No SC013532

University of St Andrews

School of Psychology & Neuroscience Ethics Committee

14 March 2016

Dear Sonja

Thank you for submitting your application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 9th February 2016. The following documents have been reviewed:

1. Animal Ethics Form
2. Leipzig: Information for ethics committees

Project Title:	Functional fixedness in nonhuman great apes		
Researcher's Name:	Sonja Ebel		
Supervisor:	Professor Josep Call		
Approved on:	11/03/2016	**Approval Expiry:**	11/03/2019

The School of Psychology & Neuroscience Ethics Committee approves this study from an ethical point of view.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to the School Ethics Committee.

You must inform the School Ethics Committee when the research has been completed. If you are unable to complete your research within the three year validation period, you will be required to write to the School Ethics Committee to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an application for ethical amendment Form submitted where appropriate.

Approval is given on the understanding that the ASAB Guidelines for the Treatment of Animals in Behavioural Research and Teaching (published in Animal Behaviour, 2003, 65, 249-255, http://www.sciencedirect.com/) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Professor Josep Call (Supervisor)
School Ethics Committee
Dr Tamara Lawson (Home Office Liaison Officer)

University of St Andrews

School of Psychology & Neuroscience Ethics Committee

22 November 2016

Dear Sonja

Thank you for submitting your application for amendment which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 17th November 2016. The following documents have been reviewed:

1. Application for change(s) to a School Ethics Committee Form
2. Amended Animal Ethics Form

Project Title:	Functional fixedness in nonhuman great apes		
Researcher:	Sonja Ebel		
Supervisor:	Professor Josep Call		
Original Approval:	11/03/2016	**Amendment Approval:**	17/11/2016
Approval Expiry:	11/03/2021		

The School of Psychology & Neuroscience Ethics Committee approves the amendment to this study from an ethical point of view.

Approval is given for five years from the date of approval of the original application. Projects, which have not commenced within two years of original approval, must be re-submitted to the School Ethics Committee.

You must inform the School Ethics Committee when the research has been completed. If you are unable to complete your research within the five-year validation period, you will be required to write to the School Ethics Committee to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and a further application for amendment submitted where appropriate.

Approval is given on the understanding that the ***ASAB Guidelines for the Treatment of Animals in Behavioural Research & Teaching (ANIMAL BEHAVIOUR, 2012, 83, 301-309)*** are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Professor Josep Call (Supervisor)
School Ethics Committee
Dr Tamara Lawson (Home Office Liaison Officer)

University of St Andrews

School of Psychology & Neuroscience Ethics Committee

22 November 2016

Dear Sonja

Thank you for submitting your application which was considered at the Psychology & Neuroscience School Ethics Committee meeting on the 17th November 2016. The following documents have been reviewed:

1. Animal Ethics Form
2. Leipzig Protocol

Project Title:	Functional fixedness in nonhuman great apes 2		
Researchers' Names:	Sonja Ebel and Christoph Völter		
Supervisor:	Professor Josep Call		
Approved on:	17/11/2016	**Approval Expiry:**	17/11/2021

The School of Psychology & Neuroscience Ethics Committee approves this study from an ethical point of view.

Approval is given for five years. Projects, which have not commenced within two years of original approval, must be re-submitted to the School Ethics Committee.

You must inform the School Ethics Committee when the research has been completed. If you are unable to complete your research within the five-year validation period, you will be required to write to the School Ethics Committee to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an application for ethical amendment Form submitted where appropriate.

Approval is given on the understanding that the ***ASAB Guidelines for the Treatment of Animals in Behavioural Research & Teaching (ANIMAL BEHAVIOUR, 2012, 83, 301-309)*** are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Professor Josep Call (Supervisor)
School Ethics Committee
Dr Tamara Lawson (Home Office Liaison Officer)