CHILDREN'S TOOL MAKING:

FROM INNOVATION TO MANUFACTURE

by

NICOLA CUTTING

A thesis submitted to the

University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

University of Birmingham

September 2012

Abstract

Through eight experiments this thesis investigated the divergence in children's abilities in the domain of tool making. Despite being excellent tool manufacturers following full instruction, children displayed great difficulty in innovating novel tools to solve problems. Experiments 1 to 3 found four-to-seven-year-olds' tool-innovation difficulty to be a robust phenomenon that extended to new tasks requiring different tools made by a variety of methods and materials. Experiments 3 and 4 aimed to discover whether some tool-innovation tasks are harder for children than others. Together these experiments suggested that the difficulty of tool innovation is due to the type of transformation required. Experiments 5 to 8 investigated why children find tool innovation so difficult. Experiments 5 and 6 ruled out singular executive functions as limiting factors on children's performance. Experiments 7 and 8 found that young children have great difficulty in generating and coordinating the components of a problem even if aspects of the task are highlighted for them. Overall this thesis led to the conclusion that tool-innovation difficulty is due the ill-structured nature of the task. Additionally this thesis provides new definitions and frameworks with which to study tool-related behaviour that will benefit both the developmental and comparative literatures.

Acknowledgements

First, I would like to thank Sarah Beck and Ian Apperly for their support and encouragement over the past three years. It has been great to work with people who are as excited by pipecleaner hooks as I am. I hope that a completed thesis on hooks will make up for the disappointment that you never managed to get me to drink coffee or beer. Perhaps you could accomplish this over the next three years!

The next thank you has to go to the teachers and children at the schools that participated in this research. Without you this thesis wouldn't have been possible. David Grist and Bev Archer deserve a special mention for allowing me to visit their schools on multiple occasions, and always being so enthusiastic about my research.

Thank you to members of the Cognitive Development Group for providing helpful comments and suggestions following a number of talks in department. Thanks also to Emma Tecwyn and Zoe Demery for offering a new perspective on my work, and for initiating the fabulous conferences we have organised.

Thank you to all my friends in office 517 for providing a fantastic working environment full of fun, laughter, and occasionally, work. Special thanks goes to Hayley Mace, without whom this thesis would have been finished a long time ago, but wouldn't have been half as much fun!

Thanks to my dad for his support and encouragement, and always believing in me. Thanks also to Anne and Graham. Anne for always being so optimistic and encouraging, and Graham for his help with making new apparatus.

Last but definitely not least, I would like to thank Ivan. Without the support, both from yourself and from your wallet I wouldn't have made it to this point. Thanks for putting up with the endless talk about pipecleaner hooks, and the vast amount of pipecleaners that seem to appear all around the house. Unfortunately for you this isn't the end of pipecleaner hooks, this is just the beginning….!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.0 Overview

Human life is dominated by the use of tools. From 'simple' tools such as hammers and pencils to more complex ones such as computers and mobile phones, life as we know it would not exist if we had not evolved such a tool-rich culture. Children build a vast knowledge and experience of the tools that surround them from a very young age. However, to get to the tool-rich culture we currently experience someone must have invented and made those tools in the first place. Surprisingly there has been very little investigation into the tool-making abilities of modern humans. This thesis presents the first series of studies to investigate tool making in young children. The findings of this thesis challenge the assumption that humans are experts in all aspects of tool-related behaviour. Drawing on definitions and evidence from the comparative literature this thesis introduces a new way of thinking about tool making and provides new definitions and ways of categorising tool behaviour that benefit both the developmental and comparative literatures.

In this introductory chapter I will first review human children's tool-using behaviours. I will then review evidence for tool making in the comparative literature before discussing the limited evidence for human children's tool-making difficulties. I will go on to review areas of the literature that may give some indication for the roots of children's difficulties. Finally I will outline the studies included in this thesis.

1.1 Tool Use

1.1.1 Children's Tool Use

Children have a vast aptitude for learning about tools. Tool use is observed in infants as young as 9 months old who have been shown to pull a cloth towards themselves to retrieve a toy placed on top (Willatts, 1984). The most commonly used tool in early life is the spoon (Barrett, Davis & Needham, 2007), and as such spoon use has been widely studied. Connolly and Dalgleish (1989) conducted a longitudinal study investigating the development of infant spoon use. They found that spoon use improved due to the emergence of handedness, convergence on a singular grip type and increased visual monitoring which helped the infant improve their control over the spoon.

It is proposed that our cognitive system facilitates tool learning due to the ability to represent tools in terms of functions (Hernik & Csibra, 2009). From a young age children think about tools in terms of the outcomes they can achieve. Children strive to learn as much as possible about the functions of tools and artefacts. Any information that they are unable to learn directly for themselves, children gain from observing others. In this section on children's tool use I will first review the literature on children's function learning before assessing what we know about children's ability to learn about tools from others.

1.1.2 Children's Learning about Functions

The majority of research reviewed in this section refers to artefacts in general rather than tools specifically. An artefact is any man-made object, whereas the definition of a tool is an '*unattached environmental object*…(used)…*to alter more efficiently the form, position or condition of another object'* (B. Beck, 1980, pg. 10). Tools are therefore artefacts that have specific functions.

Children understand that artefacts are for something from a very young age. This functional understanding is not simply linked to actions, but exists independently (Hernik & Csibra, 2009). Children understand that an artefact has a function even if it is not currently being used. They do not need to see someone using an artefact to know that it must have some use. All artefacts are expected to have functions and children are not satisfied until they have learnt what they are. This has been demonstrated in children as young as 2 years old, who after asking ‘what is it?’ were more satisfied with an answer about function than they were with merely a label (Kemler Nelson, Egan & Holt, 2004). In conditions where an artefact was labelled only with its name, children asked more follow-up questions to try and discover the artefact’s function. This drive to learn about artefacts and their functions is adaptive as it allows us to behave in efficient ways.

Adults are known to take a ‘design stance’ (Dennett, 1987) when considering the functions of artefacts, meaning they relate functions to the historical origins of artefacts, namely the intentions of the inventor/designer (Chaigneau, Castillo, & Martinez, 2008). Bloom (1996) suggested this was due to artefacts being the manifestation of intention. Artefacts are created to have specific functions intended by the designer. Other factors such as current use, context and appearance have been shown to be important, but adults’ categorisation of artefacts is predominantly determined by the original intended design (Kelemen, 1999; German & Johnson, 2002). Having a ‘design stance’ is adaptive and efficient as it ensures that adults know which object they will need to use to perform a specific task without having to think through their actions each time to see if something will work. However, this relationship with objects causes adults to have functional fixedness (Duncker, 1945), in that they assign a function to an object and then find it difficult to use the object for any other means.

There is some debate as to when children adopt the design stance approach towards objects, but it is argued to be around 6 years of age (Defeyter & German, 2003; German & Defeyter, 2000). Defeyter and German (2003) demonstrated functional fixedness in children from the age of 6. These children were less likely to choose the only object able to push a toy out of a tube, a pencil, when its function as a writing implement was demonstrated than when no demonstrations occurred. Below this age there was no difference between the two conditions. Defeyter and German concluded that below age 6 children’s views on object function are based on the current goals of the user rather than the designed for function. More recent work with 4- to 6-year olds (Defeyter, Hearing & German, 2009), i.e. before functional fixedness becomes established, has shown that judgements about function are based on current goals, but judgements about category are more likely to be based on the designers intentions, indicating the beginnings of ‘design stance’.

Casler and Kelemen (2005) have also investigated children’s approach towards object functions before functional fixedness occurs, and found that children need only have one experience to deem that an object is ‘for’ a particular function. They went on to test younger children and found that although 2-year-olds quickly learn that a tool has a particular function, children did not readily make exclusive object categories at this age (Casler & Kelemen, 2007). Although 2-year-olds will use a tool for a particular function, they will happily use that tool for a second function. Casler, Terziyan and Greene’s (2009) recent research has shown that children as young as 2 years of age show evidence that they strongly believe there to be right and wrong ways to act with objects. For both familiar and novel objects children protested when a puppet used the object in a non-demonstrated way. Jaswal (2006) provides further evidence that young children (3- to 4-years) have the beginnings of design stance as they are more likely to name an object with an obviously ‘wrong’ name if they are told the person made it rather than found it. This shows evidence of naming taking the inventor’s intentions into consideration.

In sum, humans have adapted to learn about the functions of artefacts from an extremely young age. This adaption allows children to navigate the world and its artefacts quickly and efficiently. This may however come at a cost if children have set ideas about what artefacts are for and become unable to use them in ways other than their intended use.

1.1.3 Children's Social Learning of Tools

When children are unable to learn how artefacts function for themselves they look to others for this information. Gardiner, Bjorklund, Greif and Gray (2012) have recently demonstrated that learning about tools from others is much more efficient than exploring objects and discovering functions for oneself. There has been a vast amount of research investigating how children learn to use tools from others. Children learn to use tools from an extremely young age (Connolly & Dalgleish, 1989; Meltzoff, 1995; McGuigan & Whiten, 2009) and in this section I will review the literature detailing the methods used by children to learn about tools.

Early research classified any type of behaviour replication as imitation, but more recently different types of copying behaviour have been more precisely defined. Want and Harris (2002) proposed five types of social learning: local enhancement, stimulus enhancement, emulation, mimicry and imitation. Local enhancement was first defined by Thorpe (1956) and refers to an increased interest in the location where an action has taken place. Stimulus enhancement (Spence, 1937) more specifically refers to an increased interest in the object on which an action was carried out. Local and stimulus enhancement only give the observer information that there is something of interest about the location or object. The observer must engage with the object to work out for themselves what the goal is and how they can achieve it. Emulation (Tomasello, 1990) refers to when an observer learns the properties of objects and the causal relationships between them. Using this information the observer is able to adopt his or her own strategy to achieve the goal. Whiten and Ham (1992) defined goal emulation, and used it to refer to instances when an observer learns about a goal but then uses their own means to achieve it. Mimicry (Tomasello, Kruger & Ratner, 1993) refers to instances when an observer replicates a model's actions without any insight in to why they might be effective or even what goal those actions might serve. Lastly, imitation (Tomasello, 1990) refers to the recognition and reproduction of a goal and the specific actions required to bring about that goal. Two types of imitation have been suggested; blind imitation refers to the replication of actions and goals without any understanding of the affordances or causal mechanisms involved, whereas insightful imitation refers to when one learns about the affordances and causal mechanisms via the imitative process. Most developmental tool research has focused on the imitation-emulation distinction, there is very little research regarding mimicry, stimulus and local enhancement.

There are differing views on children's motivations to copy others' behaviours, and children's motivations are thought to change through development. Young children are thought to be motivated to imitate for cognitive reasons, namely to learn about causal events (Uzgiris, 1981). As infants are interested in learning how the world works, under the age of 2 they often emulate end results when a model's actions appear to contain causally irrelevant behaviours (Nielsen, 2006). However, infants returned to imitating full actions if they could see that extra actions are justified (Gergely, Bekkering & Király, 2002).

As children get older their motivations for imitating others change. By the age of 2 as well as imitating for cognitive advancement children begin to imitate for social reasons. By 18 months infants are more likely to imitate a model if they were behaving socially rather than aloof (Nielsen, 2006) and if the actions to be copied are socially cued (Brugger, Lariviere, Mumme & Bushnell, 2007). Infants become more faithful at imitation with age and their desire to be sociable leads them to imitate all aspects of a model's behaviour. Killen and Uzgiris (1981) found that infants were willing to imitate nonsense actions in order to keep the game going and continue their interaction with the model. Children become such faithful imitators that their behaviours often become maladaptive regarding the endstate they are trying to achieve. Children have been shown to imitate all aspects of a model's behaviour even when they are obviously causally irrelevant; this phenomenon has been termed over-imitation (Horner & Whiten, 2005; Lyons, Young & Keil, 2007).

Horner and Whiten (2005) tested whether 3-to-4-year-olds would selectively use emulation or imitation depending on the causal information available to them. Children watched as an experimenter carried out relevant and irrelevant actions with a tool on opaque and transparent puzzle boxes to retrieve a reward. Horner and Whiten suggested that the children possessed the required causal knowledge to understand which actions were relevant. Based on this they suggested that children should imitate all actions in the opaque condition when the relevance of the actions could not be observed, but should emulate in the transparent condition when it should be obvious that the actions were not needed. This was the pattern of behaviour observed in a group of chimpanzees tested on the same task. However, the children in this study were shown to imitate in both conditions, a finding consistent with previous research (Nagell, Olguin & Tomasello, 1993; Whiten et al., 1996). Horner and Whiten suggested that children's tendency to imitate obviously irrelevant actions could be due to them focusing on actions rather than goals, and/or could be due to children interpreting the actions as intentional. The actions were presented three times and so would look purposeful to an observer. Horner and Whiten's (2005) findings were replicated by McGuigan, Whiten, Flynn and Horner (2007) who confirmed the presence of over-imitation in a larger sample of 3-year-olds and extended the findings to include 5-year-olds. Surprisingly the older children were found to be higher fidelity over-imitators than the 3-year-olds suggesting that imitation might be an adaptive strategy which is more frequently adopted with age.

There is some debate about the reasons underlying over-imitation in children. Lyons, Young and Keil (2007) proposed three potential explanations for over-imitation. First, children may over-imitate to satisfy social motivations, meaning they are more interested in the imitative interaction than the actual action being copied. Second, children may over-imitate because they view the actions of the model as being intentional. Third, imitation may occur due to it becoming habitual. Through a series of manipulations Lyons et al. investigated these opposing suggestions. Children continued to over-imitate despite attention being brought to 'silly' actions and despite the experimenter leaving the room. This suggests that children were not over-imitating to satisfy social motivations. Over-imitation persisted despite children being asked to quickly check the apparatus worked, a manipulation designed to rule out children's assumption that they are supposed to copy the experimenter and to overcome any reluctance to contradict the model's actions. Lyons et al. concluded that children over-imitate in order to learn causal relations; imitation becomes automatic in children leading to them mis-encoding actions as being causally relevant even if there is evidence to suggest otherwise. Lyons et al. concluded this on the basis of their final manipulation where children stopped imitating if the actions to be imitated occurred on one half of an object that was not connected to the other half containing the reward. When the model's actions clearly violated the contact principle children did not encode them as causally relevant and so did not copy them.

Simpson and Riggs (2011) argued that children's over-imitation cannot easily be explained as a desire to learn about causal mechanisms. They manipulated the order of presentation for relevant and irrelevant actions and tested whether children would imitate or emulate in a puzzle box task after either a short or long delay. They found that 3-and-4-year-olds continued to imitate irrelevant actions even if they occurred after the reward had been retrieved, suggesting that children may imitate for social reasons rather than to learn about causality. They provide further evidence against the causal argument for imitation as children stopped imitating irrelevant actions after a delay, suggesting that irrelevant actions were not encoded as well as relevant actions.

Whatever the mechanism underlying it, imitation has proven to be adaptive for humans. It is a useful method that can be used to learn new skills. Over evolution humans have learnt to be faithful imitators as this ensures that the fruitful products of other's knowledge are transmitted between individuals and generations. Imitation ensures that as well as the goals and end-products being transmitted the methods of how to achieve these are also shared. If only the end result was known it would take a lot more effort and be much more cognitively demanding to realise the means to achieve the end result for oneself. Social learning is clearly important for the spread of culture amongst a group; however, for a culture to evolve there must be innovations. In turn, these innovations must spread via transmission for progression to occur.

Research in to children's tool behaviour has predominantly investigated the ability to use tools. To my knowledge there have been only two studies investigating children's ability to make tools. This is a surprising gap in the literature as tool making is a vital component of human evolution, for humans to possess the tool-rich culture that we currently experience someone must have made the tools to begin with. The current interest in human cultural evolution is focused on the social learning of tool use, but to fully understand how human tool culture evolved we need to know about tool making. The evolution of tools requires the initial innovation of a tool and then social learning within a group to spread information about how to make the newly innovated tool. To understand about tool making we need to discover how innovations are achieved. What skills are required to invent new tools? Under what conditions do innovations occur? Second, once someone has discovered how to make a tool how does this knowledge spread amongst a group. It seems likely that the skills involved in making tools would spread through social learning in a similar way to knowledge about tool use, but we are currently lacking the evidence to support this suggestion.

1.2 Tool Making

As mentioned above there is very little research into the tool making abilities of modern humans. In the developmental literature there have been only two published studies investigating children's tool making, and so first I will look to research into tool making in the comparative literature, before discussing the limited research on human children's tool making.

1.2.1 Non-human Animal Tool Making

As with the tool-related research investigating human children's abilities research into animal tool behaviour has predominantly focused on abilities to use tools. Tool use has been reported in a wide range of animals from invertebrates to the great apes (Schumaker, Walker & Beck, 2011). Tool use is most commonly associated with obtaining food, as evidenced by the nut-cracking behaviours of capuchin monkeys (Ottoni & Manu, 2001), gorillas (Owen, 2005) and chimpanzees (Sugiyama, 1989), who used stones to pound open nuts. Other food related tool use includes the use of twigs, feathers or shells to probe into cracks by woodpecker finches (Merlen & Davis-Merlen, 2000) and observations of octopus dropping coral into open oysters to prevent them from closing (Lane, 1957). Animals have also been reported to use tools for other means such as protection or comfort. Hermit crabs have been reported to carry anemones to protect themselves against predators (Ross, 1971). When the crabs retreat into their shells the anemones block the shell entrance and prevent predators from gaining access. Elephants have been widely reported to brandish branches to rid themselves of flies, a practice commonly known as fly switching (Hart & Hart, 1994; Hart, Hart, McCoy & Sarath, 2001). As these observations clearly demonstrate, non-human animal tool use is a widespread phenomenon with reports of tool-using behaviours steadily increasing.

Recently there has been growing interest in non-human animals' abilities to make tools. In his book cataloguing animal tool behaviour Benjamin Beck (1980) defined four different modes of tool making. The most commonly observed mode of tool making is termed detach, and refers to the making of a tool by the severation of a fixed attachment, such that the severed material is used as a tool. A common example of detachment is when an animal may break a branch off a tree and use that branch as a tool. Chimpanzees (Sugiyama, 1989) and orang-utans (Kaplan & Rogers, 1994) have been observed to detach branches and use them to repel insects or to keep cool.

Subtract refers to the removal of parts of an object so that the remaining core can be used as a tool. The most common example of subtraction is the removal of twigs and leaves from a branch so that a long smooth core is left. Orang-utans have been observed to subtract leaves and twigs from a branch to leave a sharp tool used for peeling fruits (van Schaik & Fox, 1994). Bonobos subtract leaves and bark from branches to leave a core that can be used to fish for termites in termite mounds (Parish, 1994).

Add/combine refers to when two or more objects are connected together to form a tool. It is very rare to see this form of tool making in the wild, but it has been observed in the laboratory (see Price, Lambeth, Schapiro & Whiten, 2009; Bania, Harris, Kinsley & Boysen, 2009 for adding/combining in chimpanzees).

The final mode of tool making is reshaping; this refers to when a material is fundamentally restructured in to a new shape that can be used to serve as a tool. An example of reshaping is the crumpling of leaves so that they more usefully serve as a sponge tool. Sponge making has been observed in orang-utans (van Schaik et al., 2003) and bonobos (Walraven, van Elsacker & Verheyen, 1993) amongst many other species. Another form of reshaping occurs when chimpanzees (Sugiyama, 1985) chew the ends of sticks to make brush like tools which are used for termite fishing.

Much of the literature on non-human animal tool making is very descriptive and simply catalogues tool-making behaviour seen in the wild (see Schumaker et al., 2011 for a recent catalogue of animal tool making). More recent research has started to investigate causal understanding in animals and has set out to find the limits of non-human animal tool capacity. It is not surprising that the majority of evidence for non-human tool making comes from our nearest relatives – non-human primates. As such I will begin by discussing recent advances in the understanding of non-human primate tool making, before going on to discuss the remarkable tool-making abilities of animals less closely related to ourselves.

Bania et al. (2009) investigated the ability of chimpanzees to make tools by either assembling or disassembling materials. Chimpanzees were presented with a piece of dowel into the ends of which two shorter pieces of dowel could be inserted to make an H-shape. In a hook retrieval task chimpanzees were presented with the separated materials and were required to construct a hook tool to pull a reward towards them. In a probing task chimpanzees were presented with the H-shaped tool and were required to remove one of the cross-pieces to create a long stick capable of probing into a narrow tube to retrieve their reward. B. Beck (1980) would label these tool-making techniques adding and subtracting respectively. As noted above the adding method of tool making is rarely seen in the wild and as such this is one of very few studies to demonstrate non-human primate abilities for this tool-making mode. Bania et al.'s study was a replication of research carried out with chimpanzees by Povinelli et al. (2000). Povinelli et al. found low success rates for chimpanzees for both adding and subtracting tasks, these findings were inconsistent with the reported tool-making abilities observed in the wild. In contrast Bania et al.'s chimpanzees performed extremely well on both types of task, although the subtraction task was found to be easier for the chimpanzees to achieve than the adding task. This was not surprising as subtraction is the more common form of tool making seen in chimpanzees. Bania and colleagues suggest the difference between their results and those of Povinelli et al. to be due to the environment in which the chimpanzees had been reared. Povinelli et al.'s chimpanzees had not had opportunity to explore raw materials in their early development, and were also only given a short amount of time to interact with the materials prior to the start of the study. The chimpanzees were therefore not given the best chances possible to help them with the tasks. Together these studies indicate the effect that enculturation could have on non-human primate performance on tool-related tasks.

Visalberghi, Fragaszy and Savage-Rumbaugh (1995) tested the abilities of capuchin monkeys, chimpanzees, bonobos and orang-utans on tasks requiring the retrieval of a reward from a horizontal tube. After familiarisation with a simple condition where a single straight stick was needed to retrieve the reward participants were then given two complex conditions requiring the making of a tool. In one condition participants were presented with a bundle of sticks tied together with tape or an elastic band. The bundle as a whole was too wide to fit into the tube and so participants were required to detach a single stick from the bundle. In the second condition participants were presented with an H-shaped stick, a long stick with two shorter sticks at each end. These shorter sticks prevented the tool from entering the tube and as such participants needed to subtract them to make a functional tool. The great apes passed all of the tasks with ease. They never tried to insert the bundle as a whole and although they made more errors on the H-shaped stick task they quickly refined their behaviour and became quicker and more accurate over trials. The capuchins also managed to solve both of the tasks within the given timeframes, however they made many more errors than the great apes, and these errors did not reduce over trial blocks. Visalberghi and colleagues concluded that success alone does not mean causal comprehension. The fact that the errors the capuchins made did not decrease over time shows that they did not gain any understanding about the task. They did get quicker over time suggesting that they were making associations, i.e. if I put this stick in here I will get the reward, but they did not have any causal understanding about how this was achieved.

In the last decade corvid research has been at the forefront of investigations into non-human tool making. The remarkable abilities of New Caledonian crows, a member of the corvid family, have been studied in the wild and in the laboratory. In the wild New Caledonian crows have been reported to make hook tools to retrieve grubs from holes. The birds have demonstrated two methods for making hooks, both using a subtraction method of tool manufacture. The first method New Caledonian crows use is to use their beak to shape pandanus leaves into hooks (Hunt & Gray, 2003). The second method involves twigs and is achieved by breaking off parts of a twig in specific areas to leave behind a strong hooked tool (Hunt, 1996; Hunt & Gray, 2004).

Building on these natural behaviours New Caledonian crow hook use and manufacture was tested in the laboratory. Weir, Chappell & Kacelnik (2002) originally set out to test whether New Caledonian crows could choose between a hooked wire tool and a straight wire tool to retrieve a bucket containing a food reward from a tall vertical tube. Shortly into the experiment the male crow removed the hook tool and so the female subject was left with the straight wire. The female crow, Betty, spontaneously bent the wire in to a hook and used it to retrieve the bucket and her reward. The participants had had prior experience with the apparatus but the only opportunity they had to bend pliable materials was with pipecleaners a year before the current experiment, and this experience had only been for one hour. Based on this remarkable observation of spontaneous tool making, the experimenters continued to give Betty only the straight wire with which to solve the task. Betty continued to make hook tools using a variety of different techniques. For example on one trial she held the wire with her feet whilst pulling the other end upwards with her beak, whilst on another trial she pushed the wire into the sticky tape around the bottom of the apparatus and then pulled the other end away with her beak. Betty provided the first evidence in corvids of innovative tool-making methods with novel materials.

Bird and Emery (2009) replicated this study with four rooks and found that they were all successful at bending hooks to retrieve the reward. Three of the rooks successfully manufactured a hook from straight wire on their first attempt. Unlike Betty the rooks had not seen a functional wire hook in the context of the hook-making task. However, the rooks had used other hooks made of different material to retrieve buckets from the same apparatus. The rooks therefore had some indication of the type of tool that the task required.

Following on from evidence of successful hook making Weir and Kacelnik (2006) set about discovering the extent of Betty's abilities, and investigated how much Betty understood about the physical causality underlying her tool making. First, Weir and Kacelnik presented Betty with the same hook-making apparatus but a different material for making hooks. Betty was given an aluminium strip that could only be bent in one plane. After initially trying to use similar techniques that had previously been successful with the thin wire (Weir et al., 2002) Betty quickly adapted her technique and found new methods with which to solve the task. In experiment 3 Betty was required to perform a different action with the aluminium strip. This time the strip was bent into a U-shape and required unbending to allow it to fit into a horizontal tube and be long enough to reach the reward. In this study Weir and Kacelnik aimed to discover how Betty would perform on a task in which repeating previously successful actions would lead to failure. Betty successfully modified the tool by unbending on three of the four trials she was given. Despite Betty's success Weir and Kacelnik were cautious about the claims they made. It was not possible to state whether Betty had full causal understanding because she continued to make errors and enter unmodified materials throughout trials. Entering unmodified materials into the tubes on subsequent trials following successful ones where correct modifications were made appears to suggest that Betty did not have full understanding of what was required for the task. However, there may have been a logical reason for her actions. Until recently non-human primates were thought to lack understanding about non-functional upside-down traps on a trap-tube task, but recent evidence has shown that human adults did the same thing despite verbalising knowledge that the trap was non-functional. This demonstrated that even if actions are not relevant for the task these actions may not be maladaptive, especially if there is no cost to conducting them (Silva, Page & Silva, 2005). Weir and Kacelnik suggested that rather than causal knowledge being an all-or-nothing phenomenon, causal understanding is more likely to be a continuum.

Taken together these studies on animal tool making provide evidence for abilities to make tools via detaching, subtracting, adding and reshaping. The success rates from Bania et al.'s (2009) study with chimpanzees suggest that making tools by subtraction may be easier than making tools by adding pieces of material together. Results from Visalberghi et al. (1995) suggest that detaching may be an easier method of tool making than subtracting. Kacelnik, Chappell, Kenward and Weir (2006) proposed a tool making complexity hierarchy based on the degree of transformation that a tool-making episode requires. Adapting and extending the definitions of B. Beck (1980) Kacelnik et al. suggest there to be four levels of tool making complexity. The first level 'none' refers to the use of unmodified materials to act as tools. The second level consists of 'detach' and 'subtract' as defined by B. Beck. The third level consists of 'add/combine' and 'reshape', again as defined by B. Beck. Finally the fourth level consists of multi-step manufacture and fine crafting. These involve two or more manufacturing steps or sculpting of three-dimensional raw material respectively. Despite the suggestion of a tool making hierarchy and studies that appear to support aspects of it (Bania et al.) to my knowledge there have been no studies which directly investigate whether the different tool-making modes differ in complexity.

1.2.2 Human Children's Tool Making

After extensive literature searches there appear to be only two studies investigating human children's tool making. Mounoud (1996) investigated the construction of a tool from Lego blocks in children between 4 and 9 years old and Beck, Apperly, Chappell, Guthrie and Cutting (2011) used a simpler hook-bending task based on studies with corvids (Weir et al., 2002). I will discuss the findings from these studies here.

Mounoud (1996) devised a task which required children to construct a tool needed to push a cube located inside a box to another location. This task was made more complex by the addition of barriers which the tool was required to go around. Children were given Lego blocks that were easily connected together in order to construct their tool. Four-and-five-year-olds predominantly used a trial-and-error step-by-step approach. Six-to-nine-year-olds were more likely to make the whole tool before entering it in to the apparatus and then made corrections as required. The younger children in this latter age group conceived the tool in segments and did not consider the relationships between the different tool components. From the age of 7 children began to see the tool as a whole and were better able to reason about the changes that needed to be made.

Mounoud's (1996) study suggested that tool making might be very difficult for young children; however it is a very complex task which leaves us unsure as to what aspects children might find difficult. There is no way of knowing if tool making itself is the difficult part, or whether the difficulty lies in children's understanding of exactly what the task requires of them.

Beck et al. (2011) tested children's abilities on a different tool-making task and more closely defined different aspects of tool making in order to discover the extent of children's difficulties. Based on the hook-making study with New Caledonian crows Beck et al. (2011) used similar apparatus to test human children's tool-making abilities. Children were presented with a clear vertical tube containing a bucket with a wire handle that contained a sticker, with the instruction that if they could get the sticker out they could keep it. Children were unable to reach into the tube with their hands to retrieve the sticker and as such were required to make a tool from the materials they were given; a pipecleaner, a piece of string and 2 short coloured matchsticks. The easiest solution to the task would be to bend the pliable pipecleaner in to a hook and use it to fish out the bucket and retrieve the sticker.

Beck and colleagues (2011) divided tool making in to two distinct types: tool manufacture and tool innovation. Tool manufacture was defined as the physical transformation of materials in to a tool. Tool innovation required a prior step where the tool maker must first imagine what tool they needed to make before constructing it. Beck et al. first tested children's ability to imagine the tool they needed and construct it using the materials given (tool innovation). If unsuccessful at tool innovation the experimenter demonstrated how to bend a pipecleaner hook, and children were then tested on their ability to manufacture a hook tool following this demonstration. Children were found to be surprisingly poor at innovating the hook tool required. Under the age of 5 children very rarely innovated a hook. Success levels gradually improved between the ages of 5 and 11, but by age 8 only around half of children successfully innovated a hook. A mature comparison group of 16 and 17-year-olds demonstrated that 100% of adults could successfully complete the task and innovate the required tool. Despite difficulties in tool innovation children in all age groups showed a great aptitude for tool manufacture with 120 out of 124 children successfully manufacturing a hook tool following the demonstration.

In a preliminary study children's understanding of hooks was tested using a forced choice paradigm. Children were presented with the apparatus and told it was their job to retrieve the bucket to win the sticker; they were then given the choice between a hooked pipecleaner and a straight pipecleaner. Children as young as 4 years old were above chance at picking up the hooked pipecleaner first and entering it into the tube, demonstrating an understanding for the use of hook tools.

In experiment 3[1] of the same paper Beck and colleagues (2011) checked to ensure that failure at tool innovation was not due to the experimenter poorly communicating the task to the children. It is possible that the children may have been unaware they were allowed to transform the materials, alternatively children may not have known about the pliable function of pipecleaners. To test for these alternative explanations in the third experiment half of the children were given the opportunity to manipulate the materials prior to the main task. In this familiarisation phase children watched as the experimenter demonstration transformations with the materials which they then had the opportunity to copy. The pipecleaner was wound around a pen and the pulled off to show that it keeps its shape. The string was laid over an s-shaped pattern printed onto card. Finally, the matchsticks were laid on the table to make a square shape. As before success levels for the innovation phase were very low and the familiarisation phase did not aid children's performance suggesting that difficulty in tool-innovation tasks could not easily be explained by a lack of knowledge about the properties of materials or lack of awareness about permission to alter the materials.

1.3 Why might tool innovation be so difficult?

Beck et al.'s (2011) studies demonstrate a divergence in children's tool-making abilities. Despite being excellent tool manufacturers following instruction, children displayed great difficulty innovating simple novel tools, with most children not succeeding until they were 8 years old. Building on this new finding this thesis aimed to determine the extent of children's difficulty and discover why tool innovation is so complex. In the

[1] Experiment 3 in Beck et al. (2011) was the first study conducted as part of my PhD. As it was part of a published series of studies and more usefully provides background to the rest of this thesis the results of this study are only reported here.

sections below I review different areas of the literature that may help to shed some light on why tool innovation is so hard.

1.3.1 Functional Fixedness

Children appear to have a disposition to learn about object functions, this leads to them adopting a design stance as described in section 1.1.2. One possible explanation for children's tool-innovation difficulties could be that children are unable to use the materials they are given to create tools because they allocate functions to them which then become fixed. This has been demonstrated in numerous tasks where as children develop the design stance they become slower at recognising alternative uses for objects (Defeyter & German, 2003; German & Defeyter, 2000). There are two reasons why functional fixedness is unlikely to explain children's tool-innovation difficulties. First, the hook-innovation task requires children to make a hook from a pliant wire pipecleaner. Pipecleaners are a common material for children that are often used in crafts, their known function is that they can be bent into whatever shape one requires. As such they do not have one particular function that one can become fixed upon. Second, studies demonstrating functional fixedness have shown that it develops around age 6, meaning that younger children often perform better than older children. In Beck et al.'s extensive developmental study children's success levels for innovating a hook tool were shown to increase with age, if functional fixedness was the cause of children's difficulty we would expect there to be a dip in performance around age 6.

Based on these two inconsistencies I do not believe functional fixedness to be the cause of children's tool-innovation difficulty and as such will not actively pursue this line of research. The studies in this thesis will however ensure that children have full awareness of the pipecleaner properties and affordances, and it will be made clear that they have permission to modify the materials they are given.

1.3.2 Executive Function

Tool making may place cognitive demands on a child. During childhood vast development of executive control is observed. Executive control is an umbrella term used to describe the psychological processes involved in goal-directed actions. The executive system is made up from three dissociated but connected components: inhibitory control, working memory (needed for monitoring and updating knowledge) and switching (also known as shifting) (Diamond, 2006; Miyake et al., 2000).

Broadly speaking inhibitory control refers to the ability to stop what one is doing. Most research in to inhibitory control defines inhibition as the ability to overcome a prepotent response. Tasks to measure inhibitory control typically consist of a situation where the participant must choose between two responses. One of these responses is incorrect but proves to be prepotent; this situation arises because it is the more habitual or practiced response. An example of an inhibitory control task used with children is the day-night stroop task. This task consists of cards that either have a picture of a sun or a moon on them. Children are instructed to answer day to the picture of the moon and night to the picture of the sun. This requires children to engage their inhibitory control because they must suppress their prepotent response of answering with the word that corresponds with the picture on the card. Inhibitory control develops markedly between the ages of 3 and 5 (Davidson, Amso, Anderson & Diamond, 2006), and continues to develop through childhood into adulthood (Williams, Ponesse, Schachar, Logan & Tannock, 1999).

The ability to monitor and update knowledge is interchangeably known as ‘updating’ or ‘working memory’. In this thesis I will use the term working memory. Working memory is the ability to monitor incoming information and code it according to its relevance for the task at hand. This involves reviewing items held in memory with new incoming information and replacing old information with new where relevant (Morris & Jones, 1990). Working memory is not merely a passive store for information but plays an active role in the manipulation and revision of knowledge (Lehto, 1996; Morris & Jones). In fact neuroimaging studies have shown that different brain areas are responsible for the different components of memory. Storage and maintenance are associated with premotor areas of the prefrontal cortex and the parietal lobe, whereas working memory has been associated with dorsolateral prefrontal cortex (Jonides & Smith, 1997).

Task switching, also known as 'attention switching' or 'shifting' refers to the ability to switch between different tasks or operations (Monsell, 1996). In simple terms task switching requires disengagement with a current but now irrelevant task, and re-engagement with a new relevant task (Miyake et al., 2000). However, Allport and Wylie (2000) suggest there to be additional difficulties due to proactive interference. When switching to a new task interference from the previous task must be overcome in order to switch successfully. Children's task switching abilities are demonstrated in card-sorting tasks such as the Dimension Change Card Sort (DCCS) task (Espy, 1997; Frye, Zelazo, & Palfai, 1995). In this task children are presented with cards depicting images that differ in two dimensions, for example flowers and cars that are depicted in either blue or red. Children are first instructed to sort the cards by one dimension i.e. category, cars vs. flowers. They are then instructed to sort by the second dimension, i.e. colour. Independent of which order children are asked to sort the cards young children fail to switch to the new rule on their second go. That is they perseverate with the sorting strategy they used for their first turn. Children's task switching ability improves dramatically between the ages of three and five (Chevalier & Blaye, 2009; Diamond, 2006), with further advances between the ages of 5 and 11 where we see improvements in speed and accuracy (Meiran, 1996).

Innovation of novel tools is likely to place demands on children's executive functions. When given a tool-innovation problem children must use their working memory. They must hold in mind the rules of the task and the different components of information. As they engage with the task they must update their knowledge based on feedback and coordinate this knowledge in to a useful solution. Children must use their inhibitory control to suppress irrelevant actions. They must also be able to stop what they are doing if their current strategy is unsuccessful. Finally, children must be able to switch between different strategies. If their current strategy is unsuccessful children must disengage with the task and re-engage using a new more efficient strategy. Based on this it appears likely that children's tool-innovation difficulties could be explained by underdeveloped executive function.

Another possibility for the difficulty of tool innovation is that it may be a more complex form of executive task that uses executive functions in a different way to the tasks outlined above. In the section below I review evidence to suggest that tool innovation may be an ill-structured problem.

1.3.3 Ill-structured problem solving

The distinction between well-structured and ill-structured problems was first made by Reitman (1965). Reitman defined problems in terms of their start state, goal state and the transformation required to go between the two. If information regarding all three of these components were present problems were regarded as being well-structured. If information was missing from one or more of the components Reitman defined the problem as being ill-structured. An example of a well-structured task is the Dimension Change Card Sort (DCCS) (Frye et al., 1995). This task has a well-defined start state and stimuli, cards with coloured pictures; a clearly defined endstate, cards should be in correct boxes; and a clearly defined set of transformations, sort the cards into boxes based on the rules given, e.g. sort by colour or sort by category. In contrast ill-structured tasks such as preparing a meal for others have ill-defined start and end states (Goel & Grafman, 2000). A cook may not completely know the start state as they will not know how hungry their diners will be. They may not fully know the goal state if they are unaware of how many courses they should make or whether their guests like what they plan to cook. The transformations would also be ill-defined as there are many different options as to how to make the meal i.e. cook it yourself or order a take-away. Well-structured problem solving has been widely studied; however there has been much less research in to ill-structured problems. In this section I will first discuss explanations for the difficulty of ill-structured tasks from the problem solving literature. I will then go on to discuss more recent developments for the use of ill-structured problem solving in the domain of cognitive neuroscience.

One strand of research investigating ill-structured problems has suggested that these types of problem are particularly difficult because simply having domain knowledge may not be sufficient to solve them (Chen & Bradshaw, 2007). To be able to use the domain knowledge we have flexibly in the context of an ill-structured task that knowledge must be well-integrated into what researchers term structural knowledge. Only when people have structural knowledge are they able to flexibly access this knowledge and manipulate it in to a successful solution to solve a problem (Jonassen, Beissner & Yacci, 1993). This has been demonstrated in tasks comparing experts and novices (Voss, Blais, Means & Greene, 1986; Wineburg, 1998). Although novices may have the component pieces of information needed to solve a problem they find themselves unable to coordinate this knowledge into a useful solution. Conversely, the experts could flexibly consider all of the relevant pieces of information and coordinate them in the required way.

Ill-structured problems have been used to advance understanding of the executive deficits seen in clinical populations. In turn this research has led to a greater understanding of brain function. As early as the 1930s it has been reported that some brain damaged patients display a huge discrepancy between scores on clinical tests in the laboratory and ability to function in everyday life (Penfield & Evans, 1935). Patients have been observed to perform at normal levels in tests given to them by experimenters, yet these same patients had difficulty in carrying out everyday simple tasks such as cooking a meal or doing the shopping (Eslinger & Damasio, 1985; Mesulam, 1986; Goel & Grafman, 2000). Based on these peculiar findings Shallice and Burgess (1991) devised new ill-structured tasks that were more closely related to everyday scenarios and required multi-tasking and prospective memory. One task, the Multiple Errands Test, took place in a shopping centre and required patients to retrieve items and information listed for them whilst following simple rules such as only being able to enter each shop once, and only entering a shop if they purchased something. Another, laboratory based task, the Six Element Test, required patients to complete six tasks of three sub-types whilst following a set of rules such as not completing two parts of the same sub-type in a row. Shallice and Burgess found that their clinical patients performed comparatively worse on these ill-structured tasks than age and IQ matched controls, despite performing at similar levels on traditional executive measures. These findings have been extended to other clinical patients (Goel, Pullara & Grafman, 2001) and autistic children (White, Burgess & Hill, 2009).

In the last ten years research in to this area has progressed rapidly. The use of brain-imaging has enabled researchers to discover that patients who display difficulty on ill-structured tasks despite performing normally on traditional executive function tasks have damage to an area of the brain termed Brodmann Area 10, also known as the rostral prefrontal cortex. Brodmann Area 10 has been shown to have a protracted maturation throughout childhood and adolescence (Dumontheil, Burgess, & Blakemore, 2008), and has the highest rate of growth for any area of the brain between the ages of five and eleven (Sowell et al., 2004; Burgess, Dumontheil & Gilbert, 2007).

Ill-structured problem solving research could help to inform us about the difficulties children display in tool-innovation tasks. The defining feature of ill-structured problems is to generate solutions that are not directly supplied by the task (Goel, 2010). Tool innovation is therefore an ill-structured problem. Innovation is a complex task involving multiple components, much like the devised ill-structured tasks described above. There is little research in to children's performance on ill-structured problems. However evidence of the protracted maturation of Brodmann Area 10 between five and eleven years could explain the development seen in children's tool-innovation abilities between these ages (Beck et al., 2011).

1.4 This thesis

The ability to innovate and manufacture tools is a vital component of human cumulative culture. Current research has focused on children's capacity for social learning and has neglected to investigate children's innovative abilities. Using Beck et al.'s surprising finding that children have great difficulty in innovating novel tools as a starting point, this thesis aimed to discover the extent of children's difficulty and explored possible explanations. Chapter 2 replicated the findings from the original hook-innovation task and demonstrated children's difficulties to extend to a new task requiring the innovation of a different tool. Chapters 3 and 4 drew on definitions of tool making from the comparative literature and tested children's tool-innovation abilities for different methods of tool making. Chapters 5 and 6 investigated possible explanations for children's difficulties examining the roles of executive function and aspects of ill-structured problem solving respectively. Finally in chapter 7 I summarise the findings of this thesis and propose a new framework in which tool making should be studied in both developmental and comparative literatures.

Chapter 2

Why do children lack the flexibility to innovate tools?

This chapter, largely in its current form, is published as:

Cutting, N., Apperly, I. A., & Beck, S. R. (2011). Why do children lack the flexibility to innovate tools? *Journal of Experimental Child Psychology*, 109, 497–511.

2.0 Abstract

Despite being proficient tool users, young children have surprising difficulty in innovating tools (making novel tools to solve problems). Two experiments found that 4-to-7-year-olds had difficulty on two tool-innovation problems, and explored reasons for this inflexibility. Experiment 1 (N=51), showed that children's performance was unaffected by the need to switch away from previously correct strategies. Experiment 2 (N=92) suggested children's difficulty could not easily be explained by task pragmatics or permission issues. Both experiments found evidence that some children perseverated on a single incorrect strategy, but such perseveration was insufficient to explain children's tendency not to innovate tools. We suggest children's difficulty lies not with switching, task pragmatics or behavioural perseveration, but with solving the fundamentally "ill-structured" nature of tool-innovation problems.

2.1 Introduction

Human life revolves around the use of tools. It is almost impossible to consider life without them. How would we cook without pans and utensils? How would we even catch or dig up our food? Humans are believed to be experts in all tool-related behaviours (Defeyter & German, 2003; Herrmann, Call, Hernàndez-Lloreda, Hare, & Tomasello, 2007). However, despite extensive research indicating children's early competence for tool use, children's tool-making abilities have been neglected in the developmental literature. In this paper we distinguish between two types of tool-making: tool-manufacture – the ability to make tools after instruction or observation; and tool-innovation - independently making a novel tool to solve a problem. The present studies focused on children's tool innovation, and explored whether children's difficulty with tool innovation was due to mental inflexibility.

Early hominid tool use is thought to have propelled human evolution, making us the advanced social beings we are today (Csibra & Gergely, 2009; Kacelnik, 2009). Tool-related activities are implicated in the development of social behaviours such as imitation, teaching, and language (Csibra & Gergely, 2009; Gibson & Ingold, 1993). The Cultural Intelligence Hypothesis proposes that the advancement of these social capacities has allowed humans to develop cognitive skills not possessed by our nearest primate relatives (Herrmann et al., 2007). Our ability to collaborate and share knowledge permitted massive technological advances in our manufacture and use of an extensive range tools. Tools quite clearly have, and continue to play, an integral part in human life.

There is a substantial literature on the development of children’s tool-related behaviours. Competent tool use is evident from an early age, demonstrated, for example, by the skilful manipulation of spoons (Connolly & Dalgleish, 1989), hooks and rakes (Brown, 1990), and many more tools in the second year of life. A large literature on social learning shows that young children are also able to learn about novel tools by imitating others from 2 or 3 years of age (McGuigan & Whiten, 2009; Meltzoff, 1995; Want & Harris, 2002). Young children can not only use tools, but show early abilities to infer their intended use (Casler & Kelemen, 2005), design (Casler, Terziyan, & Greene, 2009) and how they should be categorized (Defeyter, Hearing, & German, 2009).

Furthermore, a recent study has shown children to be competent in tool manufacture (making tools after instruction) (Beck, Apperly, Chappell, Guthrie & Cutting, 2011). Children as young as 3 years old readily manufactured a simple hook tool when the hook-making action was demonstrated. This is in line with the findings of research investigating infant memory for actions, which shows that at around 30 months children readily imitated a model who constructed a non-tool object e.g. a rattle (Hayne, Herbert & Simcock, 2003;

Herbert & Hayne, 2000; Barr & Wyss, 2008). In this paper we explore a possible limit on children's excellent tool-related capabilities.

Tools were once thought to be a uniquely human phenomenon, but tool-related behaviour is now widely studied comparatively. Recent research has focused on the making of tools. Chimpanzees *(Pan troglodytes),* have demonstrated the ability to manufacture a wide range of tools both in the wild (Boesch & Boesch, 1990), and in captivity (Bania, Harris, Kinsley, & Boysen, 2009; Visalberghi, Fragaszy, & Savage-Rumbaugh, 1995; Povinelli, 2000). However, there is some debate as to whether such behaviour, especially when seen in the laboratory, is insightful, or merely results from a trial-and-error approach (Povinelli, 2000).

New Caledonian crows *(Corvus moneduloides)* are also well-known for their tool-manufacturing abilities. Specifically, they manufacture hooks from twigs to retrieve food in the wild (Hunt & Gray, 2002). More recently, impressive tool-manufacturing abilities have been seen in the laboratory. To retrieve a bucket from a tall narrow tube, one crow, Betty, bent a piece of wire into a hook, which she then used to solve the task. What was impressive was that Betty made a tool from a piece of wire, a material that crows would not encounter in the wild. Furthermore, on repeated trials she employed a variety of bending techniques, suggesting that her success was not the result of associative learning (Weir, Chappell, & Kacelnik, 2002). More recently, four rooks, a species that does not use tools in the wild, have also solved this tool-manufacture task (Bird & Emery, 2009).

Tool Innovation

Being able to make tools allows individuals to perform a much wider range of acts than they could without tools or with only found tools. However, we should remember that, in the corvid and child studies described above, when individuals make tools they have already seen an example of the required tool, and in the child study (Beck et al., 2011) when children made tools successfully it was when the experimenter had even demonstrated how to make the tool. We term the ability to make tools having been instructed or having seen an example, tool manufacture. This begs the question of where tools come from in the first place. Tool innovation - making a novel tool to solve a problem - has been largely neglected by the comparative and developmental literatures. This is surprising because tool innovation must be the foundation for all other tool-related behaviour: children's (and adults') evident capacity to make tools and use tools that they see used by others would be of little use if nobody innovated tools in the first place.

There has been only one study to date of tool innovation. Using an apparatus based on that used by Weir et al. (2002), Beck et al. (2011) investigated children's ability to innovate a simple hook tool needed to retrieve a bucket from a narrow vertical tube. Children were given a straight pipecleaner, a long piece of string, and some small matchsticks. The most obvious solution was to bend the pipecleaner into a hook. This is what most adults did when confronted by the task (a few individuals made a functional tool by attaching a matchstick to the pipecleaner to make an inverted "T"). The critical difference between this study and the studies with corvids (Bird & Emery, 2009; Weir et al., 2002) is that the child participants had not seen an example of the appropriate tool within the context of this task. Instead they had to imagine the solution themselves; that is, they had to innovate a novel tool. As mentioned above, when children in the same study saw the experimenter demonstrate making the appropriate tool, they had no difficulties repeating this tool manufacture.

Children performed remarkably poorly on the tool-innovation task. Children aged 3- to 5- rarely made a hook, or any other functional tool; fewer than half of 7-year-olds succeeded; and children did not perform at high levels until the age of 9 or 10. These findings are even more striking in the context of further evidence presented by Beck et al. (2011). Even 4-year-olds understood that a hook was the best tool for the job and chose it over a straight pipecleaner significantly more than chance (Experiment 1). Children's difficulties persisted even after receiving a warm-up exercise with the materials that ensured they knew manipulation of materials was allowed and the pipecleaner is pliable (Experiment 3). Finally, the fact that tool innovation was the limiting step for children was underscored by the finding that almost all children successfully completed the task when they received a demonstration of hook-bending after their initial failure. This is consistent with literature that shows children are very successful at social learning. The question that arises from this finding is why do children find tool innovation so difficult?

Since young children clearly have the competence to manufacture and use tools, it seems unlikely that any difficulty with tool innovation would be due to a lack of understanding of what tools are, or any difficulty with the practical business of shaping a tool and executing tool-using actions. Instead we look to a cognitive explanation. Difficulty with tool innovation might be a consequence of the mental inflexibility that is commonly observed in early childhood. One way to characterize this mental inflexibility is to think about children's developing executive control.

Executive control is an umbrella term for psychological processes involved in the conscious control of thought and action (Zelazo & Muller, 2002; Anderson, 1998). Executive control is needed for novel tasks, or situations that require concentration, planning, strategy development, coordination, or choosing between alternative options (Diamond, 2006; Anderson, 1998; Brocki & Bohlin, 2004). Imagining what kind of tool is needed to solve a problem and how to make it (tool innovation) is likely to tap many of these demands and to a greater extent than simply using or manufacturing tools, which rely mainly on imitating actions.

There are different ways in which we might construe the role of mental flexibility in tool innovation. One possibility is that children are able to generate potential tool-innovation solutions to the task, but find it difficult to move on from unsuccessful ideas, and so tend to become “stuck in set”. The ability to select and switch between multiple perspectives, tasks or strategies to determine the optimal option for the current situation is a well-known component of executive function that develops significantly between 3 and 5 years (e.g., Diamond, 2006; Chevalier & Blaye, 2009). This is demonstrated in simple card-sorting tasks (Frye, Zelazo & Palfai, 1995; Espy, 1997), where children begin to demonstrate the ability to shift flexibly between rules. Between the ages of 5 and 11 further improvements in cognitive flexibility occur, with children passing more complex tasks (Luna et al., 2001) and improving in speed and accuracy (Meiran, 1996). It seems plausible that difficulty with switching between alternatives might contribute to children’s difficulty with tool innovation.

Thus, in Experiment 1 we investigated the role of switching in tool innovation. We tested children on two tool-innovation tasks that required ‘opposite’ solutions (hook-making required a pipecleaner to be bent; the new task required a pipecleaner to be unbent). We speculated that if children’s difficulty was with switching between strategies then having succeeded (before or after a demonstration) using one strategy on one task, children might find it particularly difficult to adopt a different strategy on the second task. Furthermore, introducing a second tool-innovation task also allowed us to generalize Beck et al.’s claims, which were based only on a hook-making task.

A second possibility is that children have the capability to innovate the tools required for the tasks, but other features of our tool-innovation task created unintended difficulties, making it difficult for them to demonstrate this flexible behaviour. For example, in the hook-innovation task a straight pipecleaner is presented along with other distracter items as materials that can solve the task. Children may perseverate with the first material with which they attempt to solve the task and fail to switch to another material if the first proves unsuccessful. Alternatively, they may restrict themselves to using only unmodified materials rather than making them in to a new tool. We will discuss this further in the introduction to Experiment 2, where we adapt the task instructions so as to reduce the chances that children will perseverate with unmodified materials.

A third possibility is that, despite being able to make and use tools, young children lack the mental flexibility necessary to innovate tools because tool innovation is an “ill-structured” problem. Executive function researchers distinguish between “well-structured” and “ill-structured” problems (Burgess et al., 1996; Goel, 1995). Most commonly-used tests of executive function (including those used with children) are well-structured insofar as they have a clearly-defined set of stimuli (e.g., cards with coloured pictures) a clearly-defined set of responses (boxes in which to sort the cards) and a clearly-defined set of rules (sort according to the colour of the picture; then switch to sorting according to shape). In contrast, ill-structured tasks lack information either in their start or goal states or in the transformations needed to get from one to the other, and so part of the task requirement is for the participant to supply this for themselves. The difference between well- and ill-structured executive tasks is underscored by the observation that some brain-injured patients (Shallice & Burgess, 1991) and children with autism (White, Burgess & Hill, 2009) may pass traditional, well-structured, executive function tasks, yet show impairment on ill-structured tasks, and experience difficulties with mental flexibility in their everyday lives.

Tool innovation is an excellent example of an ill-structured task. The participant has information about the start and goal states, but lacks information about how to get from one to another. They must devise and hold in mind a solution to the problem, inhibit irrelevant actions and plan a sequence of actions to achieve their goal. We return to whether tool innovation might be thought of as an ill-structured task in the General Discussion, in the light of our tests of the role of cognitive flexibility.

2.2 Experiment 1

The first experiment replicated Beck et al.'s (2011) hook study, with the addition of a second tool-innovation task, unbending. In the new task a pipecleaner was presented bent in half and had to be unbent to make it long enough to push a ball from a tube. An unbending task was chosen as it requires the opposite action to the hooks (bending) task. Following Beck et al. children were given a piece of string as a distracter as well as a pipecleaner (although unlike Beck et al. we did not include small sticks, in order to prevent the making of other functional tools). The distracter material allowed us to see if the first material children selected was the functionally relevant pipecleaner. Also, all children received a warm-up exercise in which they manipulated materials (as in Experiment 3: Beck et al.) to ensure they had experience of the materials' properties. Although we did not explicitly check, all children were expected to have had previous experience working with pipecleaners in a craft context in school.

2.2.1 Method

2.2.1.1 Participants

The final sample consisted of 24 4- to 5- year olds (13 boys), mean age 4 years 10 months (4; 10), (range 4; 3 to 5; 3), and 27 6- to 7-year-olds (10 boys), mean age 6; 8 (range 6; 3 to 7; 2) from a Primary School in South Birmingham, UK. The ethnic composition of the sample was 91% Caucasian, 7% Black and 2% other/unknown. A further 5 children were tested but excluded from analysis, 3 children from the 6 to 7 age group retrieved the sticker without making a functional tool (e.g. by catching the bucket on the folded end of the wire pipecleaner) and 2 children from the 4 to 5 age group, one who retrieved the sticker without making a tool and one who had seen another child perform the task.

2.2.1.2 Materials

For the warm-up task a pipecleaner (length 29cm), pen (length 14cm), a piece of string (length 29cm) and a template of an S-shape printed onto A4 card (height 12cm, width 9cm) were used. For the hooks task the materials were a transparent plastic tube (height 22cm, width of opening 4cm) attached vertically to a cardboard base (length 35cm, width 21cm), a bucket with a wire handle, a pipecleaner (length 29cm), a piece of string (length 29cm), and a sticker (See Figure 2. 1). For the ‘Unbending’ task the materials were a transparent plastic tube (length 22cm, width of opening 4cm) attached horizontally to a cardboard base (length 33cm, width 15cm), a pipecleaner bent in half (unbent length 22cm), a piece of string (length 29cm), and a small spherical pompom (like those used in crafts; diameter 4cm) with sticker attached (See Figure 2. 2). We used a small clock to time the task.

Figure 2. 1 *Tall tube containing bucket (with sticker inside), pipecleaner and string.*

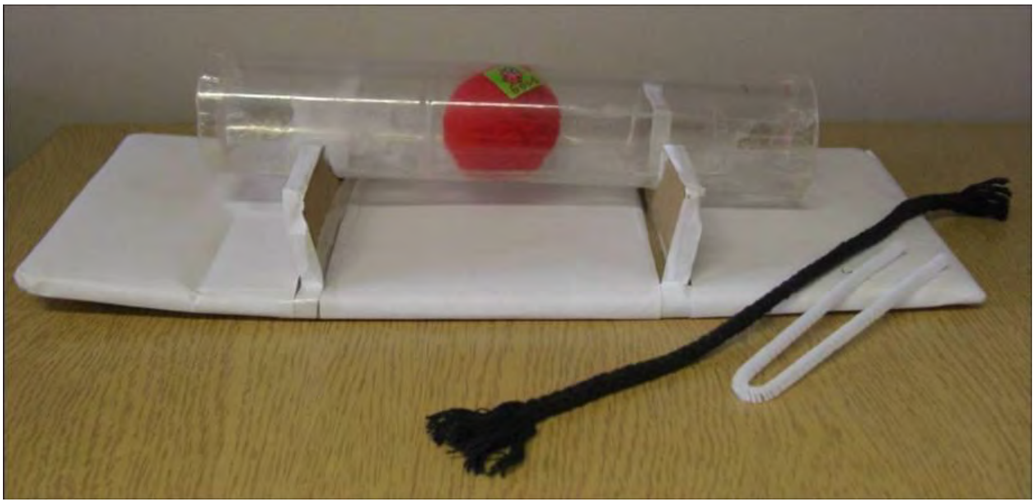

Figure 2. 2. *Horizontal tube containing pompom with sticker attached, bent pipecleaner and string.*

2.2.1.3 Procedure

Before testing began children were instructed by their class teacher not to tell other children how to play the games they would be playing with the experimenter in order for them to be a nice surprise for everyone. Participants were tested by a female experimenter in a quiet area just outside the main classroom. The child and experimenter sat facing each other across a table. First, children completed the warm-up exercise. After this, all children received both the hook and unbending tasks. The order was counterbalanced across participants.

2.2.1.3.1Warm-up task

Children watched as the experimenter demonstrated actions with the string and pipecleaner (order counterbalanced), which the child then copied. The pipecleaner was wound around a pen, and then removed to demonstrate that it kept its shape. The string was laid over the template to follow the S-shaped pattern.

2.2.1.3.2 Hooks task

Children were shown the vertical transparent tube with the bucket containing a sticker already in place in the bottom. They were told that if they could get the sticker out of the tube they were allowed to keep it. The experimenter then brought out the string and pipecleaner and told the child that these were things that 'may help' to get the sticker out. The children were then given one minute to try to retrieve the sticker. No feedback was given, but children were given neutral prompts if required. Examples of prompts used include 'Can you think how you might be able to get the sticker out?' and 'Maybe you could use these things to help you.' If, after one minute, the child had not retrieved the sticker, they were encouraged by the experimenter to put down the materials they were using. With the materials remaining on view in front of the participant, the experimenter then said 'watch this,' and using another pipecleaner held in the middle, bent one end to form a hook. The children were again encouraged to retrieve the sticker. They were not given the experimenter's hooked pipecleaner.

2.2.1.3.3 Unbending task

Children were shown the horizontal tube with the sticker attached to a pompom held in the middle. As in the hooks task, they were told that if they could get the sticker out they could keep it. The experimenter introduced the string and the bent pipecleaner as things that 'may help' to retrieve the sticker. If, after one minute had elapsed, the child had not retrieved the sticker, then they were encouraged to put down the material they were using. With the materials remaining on view in front of the participant, the experimenter then demonstrated 'unbending.' with another bent pipecleaner. The children were again encouraged to try to retrieve the sticker (using their own materials).

2.2.1.4 Measures

Children's behaviours were recorded online, using a coding system to differentiate their actions across time. The system coded materials selected (including whether they were touched, picked up, or entered into the tube), whether the material was manipulated and what shape was made, and whether the participant was successful before and after the experimenter's demonstration.

2.2.2 Results

There was no difference in success based on gender (Hooks – Fisher's Exact Test, p=.739, Unbending - χ^2 (1, N=51) =.123, p=.723), and so all data were combined for subsequent analyses.

Table 2. 1. *Children's Behaviours during Innovation Tasks for Experiment 1.*

	Touched first		Used first				**Success**	
Age Group	Pipe-cleaner	String	Pipe-cleaner as presented [a]	String	Pipe-cleaner adapted[b]	Combo [c]	**Before demo**	**Only after demo**
				Hooks Task				
4- to 5- *n=24*	21	3	20	3	1	0	**2 (**8%)	**19** (79%)
6-to 7- *n=27*	25	2	23	1	2	1	**8** (30%)	**18** (67%)
				Unbending Task				
4- to 5- *n=24*	17	7	17	5	1	1	**8 (**33%)	**14** (58%)
6-to 7- *n=27*	22	5	20	4	3	0	**15** (56%)	**11** (41%)

Note. [a]Hooks task: pipecleaner presented straight, unbending task: pipecleaner presented bent in half. [b]Hooks task: pipecleaner bent into hook, unbending task: pipecleaner unbent. [c]Subject combined string and pipecleaner, usually by tying them together

As can be seen in table 2.1 children did not assume all materials to be equally useful; there was a strong bias for children to both touch and use the pipecleaner first in each task. Furthermore, children tended to use the materials as they were presented and very rarely made any attempt to adapt them. However, as is clear in the final column of the table, once tool-manufacture was demonstrated, children easily succeeded in these tasks.

First we focused on the main variable of interest: successful tool innovation before demonstration. Children were coded as successful in the hooks task if they bent the pipecleaner into a hook, within the one minute time limit, and used this to retrieve the bucket from the tube. Children were coded as successful in the unbending task if they unbent the pipecleaner (within the time limit) making it long enough to push the pompom from the tube. It was occasionally unclear whether unbending had been an intentional act as exerting force on the bent pipecleaner sometimes allowed it to unbend. As insight is difficult to establish all cases of unbending were coded as successful.

The low success rates before demonstration for the hooks task are consistent with Beck et al. (2011), demonstrating a stable finding that children display difficulties in innovating a simple hook tool. The new unbending task also yielded low success rates, with only a third of 4- to 5-year-olds and half of 6- to 7-year-olds unbending the pipecleaner to make it long enough to push the pompom from the tube. Since these results may include a small number of children who unbent the pipecleaner unintentionally, the results for true insightful tool innovation may be lower still.

Comparison of success across age groups reveals a trend that older children successfully innovate more tools than younger children, but unlike Beck et al. (2011) we did not find a significant difference between age groups (Fisher's Exact Test, Hooks, p = .081; Unbending, p = .160). Therefore, data for the two age groups were combined for subsequent analyses.

Success before demonstration on the unbending task was better than on the hooks task (McNemar test, p = .011). We used Chi-square tests to investigate whether task order affected children's performance. Whether the hooks task was presented first or second had no effect on whether children succeeded in making a hook (Fisher's Exact Test, p >.999). Similarly order had no effect on success in the unbending task (χ^2 (1, N=51) = .167, p = .683). These results indicate that children's success on one task (whether spontaneous, or following demonstration) neither aided nor hindered their spontaneous success on the second task.

Having established there to be no relationship between behaviours between tasks, we decided to look more closely at both unsuccessful and successful (before demonstration) children's behaviours within each task. Although children were not perseverating on techniques across tasks, one possible reason for failure could be that children were perseverating on techniques within a task. We coded unsuccessful children as perseverators if they only ever entered one 'tool' into the tube and persisted in trying to retrieve the sticker with this 'tool' for the whole time period (1 minute). As can be seen in Table 2.2 perseveration was not a common occurrence for either the 4- to 5-, or 6- to 7-year-olds. Chi-square analyses show there to be no difference in perseveration between the two age groups (Hooks: χ^2(1, N=41) =.149, p=.699; Unbending: Fisher's Exact Test, p=.613). Although it is a potential stumbling block to overcome if you first approach the task with the wrong 'tool', perseveration cannot explain why children are not successful in innovating tools in this study.

Although unsuccessful children rarely perseverated with one material for the whole time period, few manipulated the materials in any way, i.e. bent the pipecleaner or combined materials. In the hooks task only 18% of 4- to 5-year-olds and 26% of 6- to 7-year-olds manipulated materials, and similarly in the tube task the figures were only 25% (4 to 5 years) and 17% (6 to 7 years).

Table 2. 2 *Frequencies of perseveration in unsuccessful children, and number of insertions into tube for successful children for Experiment 1.*

Age Group (years)	Unsuccessful			Successful			
		Perseveration			Insertion Into tube		
	N	No	Yes	N	Immediate Tool	1 unsuccessful then tool	2+ unsuccessful then tool
Hooks							
4 to 5	22	**15**	**7**	2	**1**	**1**	**0**
6 to 7	19	**14**	**5**	8	**2**	**4**	**2**
Unbending							
4 to 5	16	**13**	**3**	8	**2**	**4**	**2**
6 to 7	12	**11**	**1**	15	**3**	**7**	**5**

Next we examined the actions of successful tool-innovators within each task. Successful tool-innovators were coded as to the number of different items inserted into the tube before retrieving the sticker. Table 2.2 shows the majority of successful tool-innovators either entered a successful tool into the tube immediately (i.e. a hook or unbent pipecleaner), or entered one unsuccessful ‘tool’ (always an unmodified material) before making and entering a successful one. These results suggest tool innovation resulted more from insightful solving of the task, rather than trial and error learning.

2.3 Experiment 2

In Experiment 1 children's difficulties in tool innovation were shown to extend beyond hooks to another task: unbending. Investigation of children's success showed there was no effect of task order, which indicates that children's inflexible behaviour on one tool-innovation task was not modified by a prior experience of making a tool on another task. Importantly, children's inflexibility did not appear due to perseveration on one unsuccessful strategy. Unsuccessful children in both tasks rarely perseverated with the same material for the whole time period. However, it was also notable that unsuccessful children made few attempts to modify the materials they were given. In our second experiment we explored the possibility that children may fail to modify the materials because they think that they are not allowed to.

Although children in Experiment 1 experienced a warm-up task in which they manipulated string and pipecleaner materials, it remains possible that these children did not realize that they were allowed to alter the materials given in the context of the main task. Alternatively, children may have failed to modify materials due to the pragmatics of the task. Children were presented with the materials as things that *'may help'* to retrieve the sticker. This may have been interpreted by children as the experimenter proffering the materials as tools that could be used *as presented* as a solution to the task, thus preventing modification. In Experiment 2, we sought to minimize the likelihood of permission or pragmatics playing a role in children's poor performance on the tool-innovation task by telling children they needed to make something with the materials.

2.3.1 Method

There were two conditions. In the control condition, children received the same instructions as in Experiment 1. In the experimental condition, children received the new instruction to *make something* with the materials. This instruction was used to avoid any assumption children may have had that the materials must be used as they were to solve the task. Also, we tried to reduce any possibility that children thought the experimenter was giving them pre-made tools to solve the task by introducing the children to a puppet that happened to have some materials with him. The aim of the puppet was to draw attention away from the experimenter; making the task appear more general rather than one which the experimenter had created and had the answer to. Because of this we excluded the warm-up phase of the experiment in which children completed an unrelated task that involved manipulating the materials. Previous results (Beck et al., 2011: Experiment 3) indicated that the warm-up exercise had no effect on task success. The materials in the control condition were also presented by a puppet, and the wording changed to 'Here are some things that *can* help you.' We used the word "can" rather than "may" (as we had in Experiment 1) to match the certainty implied by the instructions in the experimental condition. Thus, the only difference between the experimental and control conditions was the instruction to make something.

2.3.1.1 Participants

The final sample consisted of 44 4- to 5-year-olds (17 boys), mean age 4; 10, (range 4; 5 to 5; 5), and 48 6- to 7-year-olds (25 boys), mean age 6; 10 (range 6; 5 to 7; 4) from a Primary School in South Birmingham. The ethnic composition of the sample was 48% Caucasian, 27% Black, 10% Asian, and 15% other/unknown.

2.3.1.2 Materials

The materials for Experiment 2 were the same as in Experiment 1 except for the addition of a short stick (5cm) presented as an additional distracter material (this matched the materials used by Beck et al., 2011), a puppet, and a box (20cm x 13 cm x 5cm) in which the puppet carried the materials.

2.3.1.3 Procedure

Participants were tested in a similar environment to that outlined in Experiment 1. All participants received both the hooks and unbending tasks, order counterbalanced. Children were alternately assigned to either the control (help) group or to the experimental (make) group based on the teacher's class list. Children were introduced to the puppet, 'Heinz', and told 'Heinz really likes to play games, so he might come back later to see what we are doing'. The procedure for both groups was identical apart from the instructions given by Heinz.

Both the hooks and unbending tasks followed the same procedure as in Experiment 1, but after showing the tube apparatus, the experimenter exclaimed, 'Oh look here's Heinz; let's see what he has to say'. The experimenter then listened as Heinz spoke in her ear and then told the children either, 'Heinz says he has some things here that can help you to get the sticker' (control group) or 'Heinz says he has some things here you can make something with to get the sticker' (experimental group). As before, if the children had not retrieved the sticker after one minute, bending or unbending was demonstrated by the Experimenter.

2.3.2 Results

Examination of success rates showed there to be no effect of gender (Hooks - χ^2 (1, N=92) =.058, p=.809, Unbending - χ^2 (1, N=92) =.097, p=.755), and so all data were combined for subsequent analyses.

Table 2. 3. *Tool-innovation Behaviours as a Function of Age and Condition for Experiment 2.*

			Touch First			Use first					**Success**	
Age group	Condition	N	Pipe-Cleaner	String	Match-stick	Pipe-cleaner as presented[a]	string	Match-stick	Pipe-cleaner adapted [b]	Combo [c]	**Before demo**	**After demo**
						Hooks						
4 to 5	Help	22	20	0	2	19	1	1	1	0	**1** (5%)	**19** (86%)
	Make	22	21	1	0	17	2	0	1	2	**2** (9%)	**17** (77%)
6 to 7	Help	23	21	2	0	19	3	0	1	0	**8** (35%)	**14** (61%)
	Make	25	23	1	1	17	1	0	4	3	**12** (48%)	**13** (52%)
						Unbending						
4 to 5	Help	22	14	7	1	11	7	1	3	0	**10** (45%)	**11** (50%)
	Make	22	16	2	4	11	2	4	3	2	**8** (36%)	**9** (41%)
6 to 7	Help	23	14	1	8	7	1	4	9	2	**16** (70%)	**7** (30%)
	Make	25	19	2	4	9	1	1	8	5	**18** (72%)	**4** (16%)

Note. [a]Hooks task pipecleaner presented straight, unbending task pipecleaner presented bent in half. [b]Hooks task pipecleaner bent into hook, unbending task pipecleaner unbent. [c]Subject combined string and pipecleaner, usually by tying them together. [d]1 participant in this group did not attempt to use a 'tool' and spent their time trying to make something.

Both the hooks and unbending tasks showed the same pattern of behaviour seen previously (see Table 2.3). Children had a strong bias to both touch and use the pipecleaner first, but very few then went on to manipulate the pipecleaner and innovate a tool.

This Experiment provides further evidence for the stability of the finding that young children do not readily innovate a hook tool to solve a task, with only 3 out of 44 4- to 5-year-olds and 20 out of 48 6- to 7-year-olds innovating a hook to solve the task. The results for the new unbending task are also found to be consistent with the previous success rates, with 18 out of 44 4- to 5-year-olds and 34 out of 48 6- to 7-year-olds unbending the pipecleaner to retrieve the sticker. As in Experiment 1, the unbending task was easier for children to achieve than the hooks task, (McNemar Test, $p < .001$). Chi-square tests were used to investigate whether task order affected children's performance. Whether each task was presented first or second had no effect on whether children succeeded in making a hook (χ^2 (1, N = 92) = .000, p >.999) or unbending (χ^2 (1, N = 92) = .003, p =.956), indicating the absence of transfer effects.

There was significant improvement in performance with age. Older children were more successful in innovating tools on both the hooks task (χ^2 (1, N =92) = 14.869, p < .001), and the unbending task (χ^2 (1, N = 92) = 8.365, p =.004). Although no age difference was found in Experiment 1, age effects were observed in this age range by Beck et al. (2011) and we conclude that the most likely reason for the difference between Experiment 1 and 2 is the larger sample size in Experiment 2.

The main aim of Experiment 2 was to test whether instructing children to make something with the materials helped them to be more flexible at innovating tools. Chi-square analyses revealed no difference between success rates for the Experimental and Control conditions for either Hooks (χ^2 (1, N = 92) =1.174, p = .278) or for Unbending (χ^2 (1, N = 92) =0.057, p = .812). This was also true for the two age groups independently (4-

to 5-year-olds – Hooks: Fisher's Exact Test, p >.999; Unbending: χ^2 (1, N = 44) = 0.376, p = .540: 6- to 7-year-olds – Hooks: χ^2 (1, N = 48) = 0.861, p = .353; Unbending: χ^2 (1, N = 48) = 0.034, p = .853). This suggests that it is unlikely that children's difficulty with innovating a tool is due to a mis-perception that they are not allowed to modify the tool-making materials.

Table 2. 4. *Frequencies of perseveration in unsuccessful children, and number of entries into tube for successful children for Experiment 2.*

	Unsuccessful			Successful			
		Perseveration			Entry Into tube		
Age Group (years)	N	No	Yes	N	Immediate Tool	1 unsuccessful then tool	2+ unsuccessful then tool
				Hooks			
4 to 5	41	**17**	**24**	3	**2**	**1**	**0**
6 to 7	28	**24**	**4**	20[a]	**7**	**8**	**3**
				Unbending			
4 to 5	26	**19**	**7**	18[b]	**10**	**1**	**6**
6 to 7	14	**13**	**1**	34	**20**	**8**	**6**

Note. [a]Two participants retrieved the sticker without making a hook. [b]One participant retrieved the sticker without unbending the pipecleaner.

As in Experiment 1 we then examined children's behaviours more closely (see Table 2.4). For unsuccessful participants we again coded whether they perseverated on one technique for the whole time period. Six- to 7 year olds' performance was consistent with Experiment 1. They did not perseverate with one object. In contrast, 4-to-5-year-olds displayed higher levels of perseverative behaviour. Chi-square analysis of the hooks task revealed 4-to-5-year-olds to be significantly more likely than 6-to-7-year-olds to perseverate on one unsuccessful technique for the whole time period ($\chi^2(1, N = 69) = 13.511$, $p < .001$). This same trend was seen for the unbending task, but did not reach significance (Fisher's Exact Test, p=.222), most likely due to the lower number of unsuccessful participants.

Examination of the behaviours of successful tool innovators paints a similar picture to Experiment 1. In both tasks the majority of successful participants succeeded immediately or after just one unsuccessful insertion, suggesting a role for insight rather than trial and error learning.

2.4 General Discussion

The present experiments suggested that young children show striking inflexibility on two tasks that require them to innovate a simple tool, and investigated alternative reasons for this inflexibility.

An important first objective was to test whether difficulties previously observed by Beck et al. (2011) when children were required to innovate a hook tool would also be apparent on another task. Our novel "unbending" task was easier to solve than the hooks task, yet, overall performance was still poor. Around two-thirds of 4- to 5-year-olds and a third of 6- to 7-year-olds spent their time probing with inadequate materials rather than performing the simple action of unbending the pipecleaner needed to solve the task. A reason why the unbending task may be easier for children to solve could be because the final shape of the required tool is much simpler to manufacture than the hook. The fact that unbending is easier is consistent with research in the comparative literature that has shown chimpanzees have more difficulty assembling tools than disassembling them (Bania et al., 2009). In this study chimpanzees were given a tool composed of a long stick with two short sticks that could be added to each end to make an H-shape. Chimpanzees either had to assemble a hook to retrieve an object, or they had to disassemble the H-shape to form the long stick needed to probe in a tube. As stated above chimpanzees found it easier to disassemble the tool, which fits with our finding that children found it easier to unbend and therefore disassemble what they had been given, than they did assemble a hook. Further developmental research is needed to investigate different types of tool manufacture that may have differing levels of complexity.

Having established that tool-innovation difficulties are robust across two different tasks, we next considered whether the findings could be explained by children having difficulty with switching between possible task solutions. Experiment 1 revealed that in their second task children did not perseverate on techniques that had been successful in the first task. For example, children who, before the demonstration, successfully bent the straight pipecleaner to make a hook on their first task were just as likely to switch to the correct strategy of unbending for their second task, compared with children who did not bend the pipecleaner on their first task. However, it is also noteworthy that children did not demonstrate any positive transfer effects, meaning that succeeding or being shown how to succeed in the first task did not allow children to gain insight and facilitate their tool-innovation ability, and so did not increase the likelihood of success on the second task. This suggests that tool innovation may not be an all-or-nothing insight that generalizes easily from one task to another.

Experiment 2 investigated whether children's inflexible behaviour was due to a misunderstanding that they should not alter the given materials. By telling children they could make something with the materials we aimed to overcome any tendency for children to believe that the materials were things that should be used without modification. In fact, children who were prompted to make something were no more likely to make a tool than children who were only told that the materials "could help" with retrieving the sticker.

Further evidence against the possibility that children thought they were not permitted to modify materials comes from the absence of transfer effects (in experiment 1) after the warm-up phase and (in both experiments) after their first task. Given that, in experiment 1, children modify a pipecleaner in the warm-up phase, and again when they either solve or are shown the solution to their first task, it seems even less likely that they still believe they are not permitted to modify the materials when they begin their second task. Yet, we observed no difference in children's levels of success between their first and second tasks. We believe that these considerations make it unlikely that task pragmatics or misunderstanding about permission to modify the materials are adequate explanations of children's tool-innovation difficulties. Nonetheless, it would be valuable for future work to include yet more explicit indications that the puppet or experimenter no longer needed the materials and that the child was allowed to change the materials.

To gain a better understanding of what children were doing within each task we analysed the behaviours of both unsuccessful and successful participants. For unsuccessful participants we focused on perseverative behaviour. Although perseveration was rare in Experiment 1, Experiment 2 yielded much higher perseveration rates for the younger, 4 to 5 year old children, with these children perseverating significantly more than the older children. As the levels of perseveration in each experiment are similar for the 6- to 7- year olds we suggest that the difference seen in the younger children is likely to be due to task differences between the two experiments, rather than differences between the two samples. In this regard it is notable that in Experiment 1 all children received a warm-up exercise in which they manipulated the task materials whereas this was excluded from Experiment 2 in order to make the materials appear more incidental to the overall task. It is possible that the warm-up exercise in Experiment 1 helped the younger children avoid perseverative behaviours, perhaps by priming them to manipulate the materials given in the main task. However, despite this finding it is clear from our results that children's tool-innovation difficulties are not merely due to an inability to overcome such perseverative behaviour. First, many children did not display perseverative behaviours yet were still not able to innovate tools. Second, for many of the children who succeeded, there was no apparent need to overcome perseveration on an initial unsuccessful solution because they immediately innovated successful tools. Nevertheless, although the current studies suggest that overcoming such perseveration is not the limiting step for tool-innovation success, the data do suggest that it may be a necessary condition for success. For if children initially use an unsuccessful 'tool' and then fail to stop using it, they can never go on to succeed in innovating a tool.

Together, then, Experiments 1 and 2 suggest that children's difficulty with tool innovation may not derive from difficulty with switching between alternative tool-innovation solutions nor from difficulty overcoming a bias to view the tool-making materials as having pre-established, fixed functions. We can also rule out the possibility that difficulty arises from the need to overcome a tendency to perseverate with incorrect solutions. This raises the question of what other factors might lead to children's apparent lack of flexibility on tool-innovation tasks. One possibility is that, unlike many tasks examining the development of mental flexibility and executive function in young children, which are "well-structured" problems, tool innovation is an intrinsically difficult, "ill-structured" problem (Shallice & Burgess, 1991).

To see why it might be appropriate to view tool innovation as an intrinsically ill-structured task, it is useful to compare the tool-innovation tasks to a well-structured tool task. In Experiment 1 in Beck et al. (2011) children were given the same goal of retrieving a bucket containing a sticker from a deep, narrow container, but with the choice between a straight or a hooked pipecleaner. This is a well-structured task that has clear initial and goal states, and clearly defined strategies for how to move between them, and on this task children performed very well from the age of 4. Together with the evidence of children's success after the experimenter's demonstration of tool making, this clearly demonstrates that children can recognize the solution to the problem when they see it, and can execute all of the relevant actions necessary to make and use the tool. What they find difficult is generating their own solution when it is not directly supplied.

A requirement to generate a solution that is not directly supplied by the task is the defining feature of "ill-structured" executive tasks. For example, in the Six Elements Test (Burgess et al., 1996) participants are presented with six tasks to complete and are asked to achieve as many points as possible by completing as many of the tasks as they can within a time limit, and whilst following rules, such as having to attempt every task. Thus, the task explicitly supplies the starting conditions (the games and the rules) and the objective (maximizing points scored on the games), but it is ill-structured because participants must devise their own strategy for tackling the problem. Such problems undoubtedly require multiple executive processes (including memory, inhibition, and switching), but as noted in the introduction, they do not seem to reduce simply to the sum of these components. It is possible to be impaired on ill-structured problems despite showing no impairment on standard, well-structured tests of executive function (e.g., Shallice & Burgess, 1991; White et al., 2009). We suggest that children's difficulty with tool innovation may stem from the ill-structured nature of such problems. Although there is little evidence on the development of children's performance on ill-structured executive tasks, it is noteworthy that the ability to solve ill-structured tasks has been specifically associated with regions of medial prefrontal cortex (Brodmann area 10) that show protracted maturation throughout childhood and adolescence (Dumontheil, Burgess, & Blakemore, 2008). If children's difficulty with tool innovation is derived from a broader, domain-general, difficulty with solving ill-structured problems, then it would be expected that individual differences in performance at tool innovation should be correlated with individual differences in performance on other ill-structured problems in non-tool contexts, and this relationship should be independent of general intelligence and other executive functions, such as inhibition and working memory.

Alternatively, it could be that children's difficulty with generating structured solutions for tool-innovation problems lies with a lack of domain-specific knowledge about the mechanical properties of tool-making materials, rather than with a domain-general problem with ill-structured tasks. If this were the explanation for children's difficulties with tool innovation then individual differences in successful innovation should correlate with other tasks that require knowledge of the mechanical properties of tools but do not require ill-structured problem solving. Moreover, such a correlation should persist even if children's performance on an ill-structured problem of another kind were partialled out. Future work would be necessary to distinguish between these possibilities.

Finally, whatever the detailed reason for children's difficulties, perhaps the most important conclusion from our studies is the simple and robust finding that tool innovation is a difficult and late-developing ability. Even when children are excellent tool users and tool manufacturers they fail to innovate simple tools. It is often noted that children are excellent social learners (e.g. Csibra & Gergely, 2009). Our findings highlight the importance of social learning in children's developing ability to use tools, since "reinventing the wheel" for themselves is comparatively difficult. We might speculate that two factors were critical in the historical evolution of humans' tool-rich cultures. The ability to innovate tools is clearly vital for technological advancement, but it is equally important that the valuable products of this effortful process are preserved and passed on through social learning. Either of these abilities has the potential to be the limiting step on the development of tool-rich cultures. However, we venture that the capacity for cognitively demanding tool innovation, rather than tool use, or tool manufacture, is what makes human tool culture stand out as uniquely complex.

Chapter 3

Is There a Complexity Hierarchy in Human Children's Tool Making?

This chapter, largely in its current form, is under submission as the paper:

Cutting, N., Beck, S.R., & Apperly, I.A. (under submission). Is There a Complexity Hierarchy in Human Children's Tool Making?

3.0 Abstract

The belief that humans are experts at all tool-related behaviour has been undermined by research showing children to have great difficulty in tool innovation (making novel tools to solve problems). The current paper investigated whether children's tool making follows trends for a hierarchy of difficulty seen in non-human animals. We tested 4-to7-year-olds (N=192) on tool-making problems requiring different levels of transformation complexity. Children showed poor innovation for all levels of complexity. In a second phase, children's tool-manufacturing ability was tested following two stages of demonstration. No hierarchy was observed, but many children manufactured tools successfully. Patterns of success suggest children's ability to recognize relationships between their raw materials and the target tool demonstration is critical to their performance.

3.1 Experiment 3: Introduction

Humans are thought to be the ultimate tool users and tool makers (Defeyter & German, 2003; Herrmann, Call, Hernàndez-Lloreda, Hare, & Tomasello, 2007). Comparative research uses our abilities as a benchmark with which to compare non-human animal behaviour. But are we really as good with tools as we think we are? There is no doubt that very young children show impressive abilities in tool use (Connolly & Dalgleish, 1989; Casler & Kelemen, 2005; McGuigan & Whiten, 2009), but evidence from tool making shows a divergence in children's abilities.

Tool making can be split into two distinct types – tool manufacture and tool innovation. Children demonstrate great aptitude for tool manufacture – making a tool following instruction (Beck, Apperly, Chappell, Guthrie & Cutting, 2011). Conversely, tool innovation – the ability to design and make a novel tool to solve a task – was surprisingly difficult for young children, even when the tools needed were very simple (Beck et al., 2011; Cutting, Apperly & Beck, 2011).

3.1.1 Children's Tool Innovation

Although the ability to both use tools and learn by observing others how to manufacture tools have no doubt been essential skills in the evolution of human culture, tool innovation could be the key skill that set humans apart from other animals (Beck, Chappell, Apperly, & Cutting, 2012). An aptitude for innovation seems a likely explanation for our complex tool-rich culture. The ability to learn tool manufacture (fashioning a tool having seen a model tool or demonstration c.f. tool innovation) from others is essential for transmission between individuals, but without innovation human culture would not have evolved to the extent that it has. However, as stated above, the ability to innovate simple tools is thought to have a long developmental trajectory, appearing much later than the ability to manufacture tools based on imitation. So, although tool-innovation ability may be the key skill in the advancement of human culture, it may also dictate some limits on the ontogeny of tool cognition in children.

Children's innovation difficulties were first demonstrated in a task requiring the retrieval of a bucket from a narrow vertical tube. Children failed to innovate a hook tool by the simple action of bending a pipecleaner until the surprisingly late age of 8 years (Beck et al., 2011; Cutting et al., 2011). Of course, it would be of limited interest if children only had difficulty with innovating hooks. It is therefore of critical importance to test whether children's difficulty generalizes to different materials, and also to different categories of tool problem. Cutting et al. (2011) found evidence that children's difficulty was not restricted to hook-making, but extended to another task requiring innovation of a long straight tool to push a ball out of a horizontal tube, by unbending a pipecleaner. However, this still required children to understand the physical properties of the same pliant material (a pipecleaner), and still required the material to be transformed by re-shaping. The latter point is particularly important because comparative researchers suggest that the difficulty of tool-making may be strongly influenced by the kind of transformation that is required (Kacelnik, Chappell, Weir, & Kenward, 2006). A key aim of the current study was to test whether young children's success at tool-making also varies as a function of the required physical transformation, and whether this follows the same pattern as proposed in non-human animals.

3.1.2 Non-human Animal Tool Manufacture

There are many ways in which to make tools, and it seems plausible that some methods of tool making will be easier and more common than others. As there is little research on the simple tool-making skills of modern humans, we look instead to the non-human animal literature. An important point to note is that the animal studies cited below demonstrate tool manufacture rather than tool innovation, as the animals had seen or were sometimes even instructed how to make the required tools. Although there is evidence of animals innovating new methods to make known tools (e.g. crows bending hooks, Weir, Chappell & Kacelnik, 2002) to our knowledge there are no experimental studies of *tool-innovation* in the comparative literature in which animals innovated novel tools.

In his influential work on animal tool use, Benjamin Beck (1980) defined four modes of tool making. The most commonly catalogued mode of tool making is termed *Detaching,* and is defined by the severation of a fixed attachment between two objects. *Subtracting* is the removal of parts of an object to leave behind a more functional tool. *Adding* is when two or more objects are combined to form a tool. And finally, *reshapin*g is defined as restructuring material into a functional tool (see table 3.1 for examples).

Table 3. 1 Beck's (1980) Tool making definitions.

Mode of tool making	Description	Example
Detach	Severing a fixed attachment between two objects, so that the removed object can be used as a tool.	Breaking a branch off a tree to use as a weapon
Subtract	Removing parts of an object to make it into a more useful tool.	Removing leaves from a twig.
Add/Combine	The connection of 2 or more objects to form a tool.	Connecting 2 short sticks to make a longer stick.
Reshape	Fundamentally restructuring an object's material to produce a tool.	Scrunching up leaves to make a sponge.

An example of an experimental task requiring detaching comes from Visalberghi, Fragaszy and Savage-Rumbaugh (1995) who tested primates (great apes and capuchin monkeys) on a task requiring them to push a food treat out of a narrow horizontal tube. In one condition participants were required to detach a suitable stick from a bundle of sticks held together with either an elastic band or tape. The bundle itself was too large to fit into the tube and so detaching a single stick was the only solution to the task. All participants from both species solved the task on every trial. Great apes found the task trivially easy and never attempted to insert the whole bundle. In contrast capuchin monkeys made many errors before solving the task on a given trial, and these errors did not significantly decrease over trial blocks. It should be noted that it is not perfectly clear that this task really requires detachment. The task does not require animals to sever any fixed attachments and as such does not operationalise detaching as defined by B. Beck (1980). Alternatively, one might construe it as 'just' a tool selection task, i.e. if the bundle of sticks is seen as a collection rather than a single object. We will return to this point in the discussion.

Tasks with chimpanzees illustrate the subtracting and adding modes of tool making (Bania, Harris, Kinsley, & Boysen, 2009). In this experiment participants were presented with pieces of dowel or PVC. There was one long piece into the ends of which two shorter pieces could be inserted to make an H-shape. Participants were presented with two tasks: in a hook retrieval task they were required to add a smaller piece in to the long piece of dowel to make a hook tool to retrieve a reward. In a second task they were required to subtract the smaller pieces of dowel to leave a straight stick tool that would fit inside a tube to push out a reward. Although levels of success were high for both tasks, chimpanzees performed better when they had to disassemble a tool (subtract) than they had to assemble (add). Thus, in addition to Beck's observation of a hierarchy in the frequency of different modes of tool making in non-human animals, there is experimental evidence that some modes of tool making may be harder than others.

The reshaping mode of tool manufacture is exemplified by the hook-making and unbending tasks with children, cited above (Beck et al., 2011; Cutting et al., 2011). These experiments were based on a study with New Caledonian crows in which one crow spontaneously bent a piece of wire into a hook to retrieve a bucket containing a reward from a narrow vertical tube (Weir et al., 2002. See Bird & Emery, 2009 for success on this task by rooks).

Although in his catalogue of natural tool behaviour B. Beck (1980) did not explicitly describe a hierarchy of difficulty for tool making, he did suggest that detaching was the simplest and most commonly seen mode. Kacelnik et al. (2006) went one step further and suggested four levels of tool-making complexity based on the amount of modification required. Level 0, ‘none’, is the simplest and refers to when unmodified materials are used as tools (n.b. selecting a tool would fall under this level). Level 1 comprises detach and subtract as defined by B. Beck. Level 2 comprises adding/combining and reshaping. Finally, level 3 consists of multi-step manufacture and fine crafting. The suggestion that detaching and subtracting are less complex modes of tool making requiring less transformation than adding and reshaping fits with the animal evidence cited above: Detaching was reported to be remarkably easy for great apes (Visalberghi et al., 1995); and chimpanzees found adding more difficult than subtracting (Bania et al., 2009). However, the full hierarchy has yet to be tested empirically in any species, including humans.

3.1.3 Testing for a hierarchy in children’s tool innovation.

In humans there is no systematic evidence on tool-making abilities in different modes. The aim of this current work was to test for the presence of a hierarchy based on transformation complexity in young children’s tool making, and whether this follows the same pattern as proposed in the comparative literature. A hierarchy was tested for in children’s ability to innovate tools via different methods, and children’s ability to manufacture tools given different levels of instruction. We tested children because previous work had shown that their tool-making abilities were under development and so we would be more likely to see a hierarchy emerging than if we had focused on adults whose tool-making is good (see Beck et al., 2011).

Using both the vertical tube and bucket apparatus (requiring a hook tool) (Beck et al., 2011; Cutting et al., 2011) and the horizontal tube apparatus (requiring a long, straight tool) (Cutting et al., 2011), the present work examined children’s tool-innovation abilities using B. Beck’s (1980) four modes of tool making. The previous reshaping tasks (Beck et al., 2011; Cutting et al., 2011) were compared to tasks requiring participants to innovate tools by detaching, subtracting and adding materials. Note that by using B. Beck's four tool-making modes this study only tests for a difference in complexity between Levels 1 and 2 as defined by Kacelnik et al. (2006).

For the vertical-tube task children were required either to *detach* the relevant hook tool from a bundle containing other non-functional tools, *subtract* pieces of dowel from a hooked tool to allow it to fit inside the vertical tube, *add* a small piece of dowel into a long stick to form a hooked tool, or *reshape* a pipecleaner into a hook as in the previous experiments (Beck et al., 2011; Cutting et al., 2011). For the horizontal-tube task children were required to either *detach* the relevant tool from a bundle, *subtract* parts of the tool that make it non-functional, *add* pieces of dowel together to make a longer tool, or *reshape* a pipecleaner in to a long straight tool as previously (see figure 3.1).

3.1.4 Testing for a hierarchy in children's tool manufacture following demonstrations.

If children were unsuccessful at spontaneous tool innovation within the given timeframe, they progressed to a tool manufacture phase of the experiment, in which we assessed children's ability to make different kinds of tool following two phases of demonstration. This is potentially important because much of children's abilities to use tools may derive from social learning rather than first-person innovation (Vaesen, 2012). Previous research has demonstrated high success levels in children after witnessing a demonstration of the action required to reshape a tool (Beck et al., 2011; Cutting et al., 2011). However, based on previous findings we do not know whether children benefitted from the provision of a solution to the problem (e.g. they saw an example of a functional hook tool) or whether they needed to see the action required to form the tool. In the present work we used a two-stage demonstration to explore this. In the initial demonstration children were shown a pre-made example of the tool they were trying to create with the simple instruction 'Look at this'. If this demonstration did not elicit tool manufacture, they then received the original manufacturing action demonstration with the instruction 'Watch this'. These two stages of demonstration allowed us to see how much information about the manufacturing process children needed to be shown before they were able to manufacture a tool.

As well as seeking the existence of a tool-making hierarchy, testing children on the different tool-making modes addressed two major limitations of the previous research into children's tool innovation. First, it allowed us to test whether children's innovation difficulties extend to other modes of tool making as well as reshaping. Second, we needed to check that children's difficulties were not limited to tasks involving pipecleaners – the only material that had been used in previous studies. Demonstrating such generalization is crucial for the conclusion that children's difficulty is with tool innovation *per se*, and not with understanding the physical properties or affordances of particular materials, such as pliant wire pipecleaners.

Children aged 4 to 7 years were tested on the tasks. This age group was chosen as previous work has shown that 4-year-olds rarely innovate tools, and so any improvement with the other modes of tool making would be easy to see. Children's success gradually increases, and around age 7 50-75% of children succeed at reshaping, and so if detaching and subtracting were easier modes we would expect to see near ceiling performance in this age group.

3.2 Method

3.2.1 Participants

Ninety-seven 4- to 5- year-olds (49 boys), mean age 5 years 2 months (5; 2), (range 4; 3 to 5; 11), and 95 6- to 7-year-olds (49 boys), mean age 6; 10 (range 6; 0 to 7; 9) participated. Mean ages and ranges were based on the data available from 65 younger and

69 older children. Children were recruited from and tested at urban schools serving a working and middle class population.

3.2.2 Materials

Table 3.1 depicts the task apparatus and the materials available to the children in each condition. For the vertical-tube task the main apparatus was a plastic tube (length = 22cm, width of opening = 4cm) attached vertically to a cardboard base (length = 35cm, width = 21cm) and a bucket with a wire handle containing a sticker. The materials for each of the conditions were as follows: Detach – a bundle held together with elastic made up from a wooden rectangular block (width = 4.5cm, length = 28cm, depth = 2cm), a piece of dowel (length = 12cm, diameter = 1.5cm), and a wooden hook composed of a piece of dowel (length = 28cm, diameter = 1.5cm) with three holes in, the end hole containing a smaller piece of dowel (length = 4cm, diameter = 0.5cm), and a separate piece of dowel (length = 10cm, diameter = 0.5cm). Subtract – A piece of dowel (length = 28cm, diameter = 1.5cm) with three holes into which were placed three pieces of dowel (end and middle pieces- length = 10cm, diameter = 0.5cm, other end piece length = 4cm, diameter = 0.5cm), and a separate piece of dowel (length = 10cm, diameter = 0.5cm). Add – A piece of dowel (length = 28cm, diameter = 1.5cm), with three holes and two shorter pieces of dowel (1 x length = 10cm, diameter = 0.5cm, 1 x length = 4cm, diameter = 0.5cm). Reshape – a pipecleaner (length = 29cm) and a piece of string (length = 29cm).

For the horizontal-tube task the main apparatus was a clear plastic tube (length = 22cm, width of opening = 4cm) attached horizontally to a cardboard base (length = 33cm, width = 15cm), and a pompom (diameter = 4cm) with sticker attached. The materials for each of the conditions were as follows: Detach - a bundle held together with elastic made up from a wooden rectangular block (width = 4.5cm, length = 28cm, depth = 2cm) and two pieces of dowel (1 x length = 24cm, diameter = 1.5cm, 1 x length = 7cm, diameter =

1.5cm), and a separate short piece of dowel (length = 4cm, diameter = 0.5cm). Subtract - a piece of dowel (length = 24cm, diameter = 1.5cm) with three holes, into the end and middle holes were two pieces of dowel (length = 10cm, diameter = 0.5cm), and a separate short piece of dowel (length = 4cm, diameter = 0.5cm). Add – Three pieces of dowel (length = 7cm, diameter = 1.5cm) with hook and loop squares attached to both ends and two pieces in the middle, and a short piece of dowel (length = 4cm, diameter = 0.5cm). Reshape – pipecleaner bent in half (unbent length = 22cm) and a piece of string (length = 29cm).

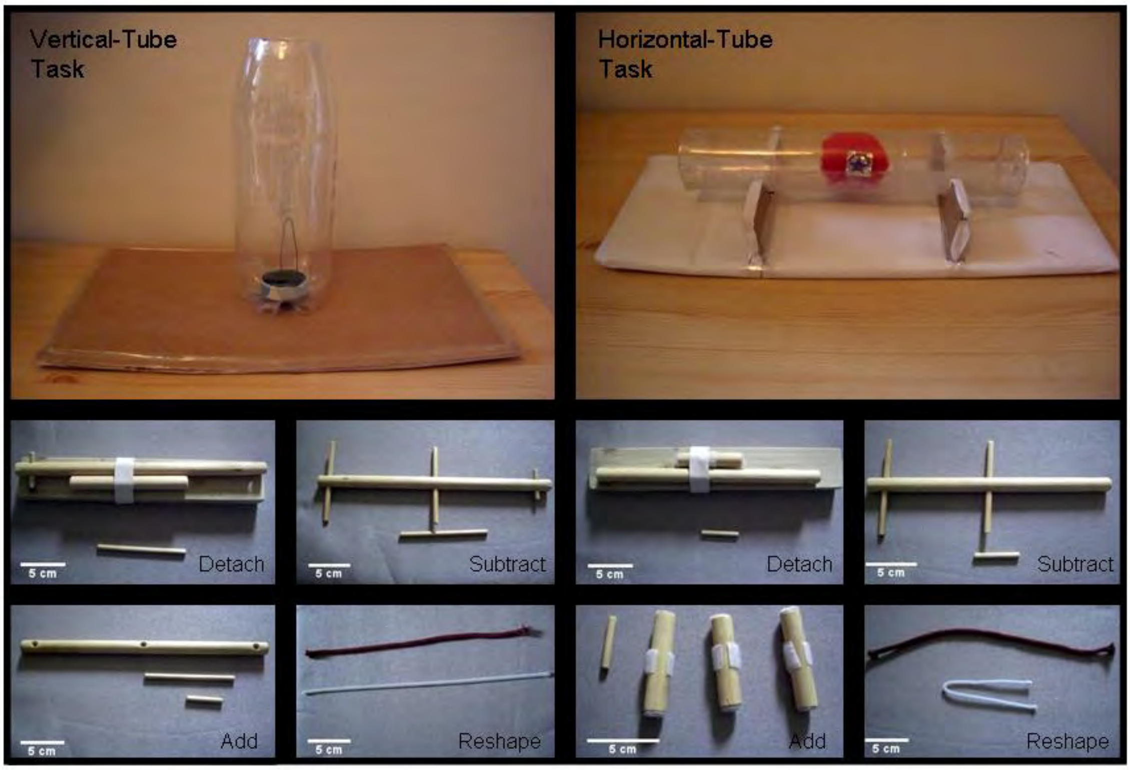

Figure 3. 1 *Apparatus for vertical- and horizontal-tube tasks.*

3.2.3 Procedure

Prior to testing children were instructed by their class teacher not to tell other children about the games they would be playing with the experimenter so they would be a nice surprise for everyone. Participants were tested individually in a quiet area just outside the main classroom. All children received both the vertical- and horizontal-tube tasks. For each task children received one of the four tool-making conditions, a different mode of making for each task (for example, one participant would complete the vertical-tube task detach version and the horizontal-tube task add version, another the vertical-tube task reshape version and the horizontal-tube task subtract version). The order of tasks and tool-making modes were counterbalanced across participants.

For each task children were shown the relevant transparent tube (vertical or horizontal) and their attention was drawn to the sticker (either in the bucket or attached to the pompom). Children were told that if they could get the sticker out of the tube they were allowed to keep it. The experimenter then brought out the relevant materials for the particular condition and told the child that these were things that ‘can help’ to get the sticker out. The children were given one minute to try to retrieve the sticker. No feedback was provided, only neutral prompts if required. Examples of prompts used include ‘Can you think how you might be able to get the sticker out?’ and ‘Maybe you could use these things to help you.’ If, after one minute, the child had not retrieved the sticker, they were encouraged by the experimenter to put down the materials they were using. With the materials remaining on view, the experimenter then said ‘look at this,’ and using their own materials held out a pre-made target tool for the child to view (target-tool demonstration). The children were again encouraged to retrieve the sticker. If, after 30 seconds, children were still unsuccessful they were again encouraged to put down the materials they were using. The experimenter then said ‘watch this’ and again using their own materials (target tool had been returned to original state), demonstrated the action required to make a functional tool (action demonstration).

3.2.4 Measures

Behaviours were recorded online by the experimenter. A coding system was used to differentiate children's actions across time. The system coded which materials were selected by the child, including whether the selection involved touching, picking up and/or entering the material into the tube; how the materials were manipulated, for example if pieces were added or removed or if the material was modified into a new shape; and finally whether success was achieved before or after each demonstration.

It should be noted that if children did the correct action but then failed to use the tool correctly they were coded as being unsuccessful pre-demonstration. However, as the demonstrations would have no effect due to the child already having performed the required action, these children received a verbal prompt as to how to use the tool they had created.

3.3 Results

The data were first analysed for any differences due to gender, age, and task order, including whether success on children's second task was affected by the mode of tool making required for their first task. Then for each of the tasks independently (vertical and horizontal) we ran chi-square analyses to compare the success rates for the different modes of tool making. Finally chi-square tests were again used to compare each mode directly against each other to determine where the significant differences could be found.

3.3.1 Tool Innovation

Following the coding of Cutting et al. (2011) and Beck et al. (2011), children were coded as successful on each of the tasks if they retrieved the sticker having made an appropriate tool within the one minute timeframe. For the horizontal tube-add task children were coded as successfully innovating a tool if they added the pieces of dowel together prior to entering them into the tube. Pushing extra pieces of dowel into the tube once a piece had already been inserted was not deemed to be true tool innovation (seen in 11 younger and 13 older children).

There were no effects of gender for either the vertical, χ^2 (1, N = 176) = 0.571, p =.45, or the horizontal tube tasks, χ^2 (1, N = 190) = 0.00, p =.988, on all modes combined, or for any mode independently (all *p*s ns). No difference in success was found based on whether tasks were presented first or second. Similarly, participants' success on the second task was unaffected by the mode required for the first task. As such, data for each mode of manufacture on each task were combined irrespective of whether participants received the task first or second.

For the vertical-tube task, the detach and reshape modes yielded ceiling and floor success rates respectively and so no difference between the two age groups was observed. Chi-square tests revealed an improvement with age for both the subtract, χ^2 (1, N = 45) = 4.874, p = .027, and add modes, Fisher's Exact Test, p = .048. For the horizontal-tube task no difference in age was observed for the detach (100% success) or add modes. Older children were more successful on the subtract, χ^2 (1, N = 51) = 11.004, p = .001 and reshape modes, χ^2 (1, N = 49) = 3.960, p = .047. Despite these differences, separate comparisons for the two age groups for the following analyses did not reveal a pattern that differed for the data collapsed over age. Therefore, subsequent analyses will report both age groups combined.

For the vertical-tube task, chi-square analyses found a significant difference in success between the different modes of manufacture χ^2 (3, N=176) = 64.574, $p < .001$. Children were significantly more successful on the detach task than on the subtract, add or reshape tasks when compared individually (all chi-squares, $p<.001$, see Table 3.2). There were trends for differences between the subtract, add and reshape tasks, but these were smaller (lowest p = .035) and were not statistically significant once a Bonferroni correction for the 6 multiple comparisons were made (the alpha level was set at $p < .008$).

Table 3. 2 *Frequencies of success before and after demonstrations for the vertical-tube task.*

Age group (years)	N	Success: Before Demo	Success: Only after Demo 1	Success: Only after Demo 2
Detach				
4-5	23	18 (78%)	5* (22%)	-
6-7	20	17(85%)	3* (15%)	-
Subtract				
4-5	22	3 (14%)	5 (23%)	10 (45%)
6-7	23	10 (43%)	5 (22%)	7 (30%)
Add				
4-5	21	0	2 (10%)	8 (38%)
6-7	22	5 (23%)	4 (18%)	9 (41%)
Reshape				
4-5	23	1 (4%)	7 (30%)	5 (22%)
6-7	22	4 (18%)	11 (50%)	6 (27%)

*Note: * Verbal prompt only, as had already carried out the required action.*

These results were mirrored on the horizontal-tube task (see table 3.3). Once again chi-square analyses found a significant difference in success between the different modes of manufacture, χ^2 (3, N = 190) = 69.677, p <.001. Children were significantly more successful on the detach task than on the subtract, add or reshape tasks when compared individually (all chi-squares, p <.001). Again there were trends for differences between the subtract, add and reshape modes (lowest p = .06) which did not reach significance.

Table 3. 3 *Frequencies of success before and after demonstrations for the horizontal-tube task.*

Age group (years)	N	Success		
		Before Demo	Only after Demo 1	Only after Demo 2
		Detach		
4-5	24	24 (100%)	-	-
6-7	21	21 (100%)	-	-
		Subtract		
4-5	28	2 (7%)	5 (18%)	5 (18%)
6-7	23	11 (48%)	4 (17%)	6 (26%)
		Add		
4-5	22	4 (18%)	4 (18%)	3 (14%)
6-7	23	7 (30%)	3 (13%)	1 (14%)
		Reshape		
4-5	22	6 (27%)	4 (18%)	5 (23%)
6-7	27	15 (56%)	6 (22%)	6 (22%)

3.3.2 Tool Manufacture following demonstration

First, we excluded children who had already succeeded in the spontaneous innovation of a tool, because they could not benefit from any demonstration. We also excluded results from the detach mode of manufacture as the small number of children who made errors in this condition were only coded as unsuccessful due to using the tool incorrectly and failing to retrieve the sticker. These children, therefore, did not receive the demonstrations as in the other conditions but merely received a verbal prompt as to how to us the tool correctly. As in previous research children were very likely to succeed at the tasks having seen a demonstration (see tables 3.2 and 3.3).

Using chi-square tests we analyzed whether there were differences with age in the success levels following the two types of demonstration for all remaining modes of manufacture combined. Following the first, 'target-tool', demonstration chi-square tests showed trends for older children to be more successful than younger children (Vertical Tube: $p = .03$, Horizontal Tube: $p = .057$). Following the second, 'action', demonstration older children were more successful than younger children (Vertical Tube: $p=.006$, Horizontal Tube: $p=.001$). These analyses suggest that older children benefitted more from both of the experimenter's demonstrations than younger children.

Next, we looked for differences in success for the modes of manufacture independently for both age groups combined. Following the first 'target tool' demonstration there were some significant differences in children's rate of success for different modes of manufacture, but they were inconsistent across the two tasks. The only difference on the vertical-tube task was that children were significantly more successful at reshaping following demonstration than they were at adding (Combined ages: χ^2 (1, N = 76) = 7.04, p =.008). For the horizontal-tube task, children were more likely to succeed following the first demonstration in the adding task rather than the subtracting task (Fisher's Exact Test, p =.01).

Following the second, 'action' demonstration there were no differences in success rates in either the vertical- or horizontal-tube tasks.

3.3.3 Comparing modes across tasks.

The following analyses will not be discussed in this current chapter, but will be referred to in Chapter 4.

We compared success levels for each of the tool-making modes across the two tasks (vertical-tube task and horizontal-tube task). We did not compare success across the two detaching tasks due to ceiling performance. There was no difference in the level of success for the two subtracting tasks, χ^2 (1, N = 96) = 1.40, p = .708. Nor was there any difference in the level of success between the two adding modes on each of the tasks, χ^2 (1, N = 88) = 2.428, p = .119. However, we found that the horizontal-tube reshaping task was significantly easier than the vertical-tube reshaping task, χ^2 (1, N = 91) = 9.456, p = .002.

3.4 Discussion

Children's tool-innovation difficulties were shown to extend to both new materials and new methods of tool making. We did not, however, find a hierarchy of complexity for the different modes as suggested in the comparative literature. In the unsuccessful innovators we observed differences with age as to the amount of instruction children needed to manufacture the required tool. Children's difficulties spanned different materials and different modes of tool making. Thus, the current study significantly extends previous findings that tool innovation is a difficult and late developing aspect of tool-related behaviour.

The first aim of this study was to test for the presence of a mode of tool-making hierarchy in tool innovation as suggested by Kacelnik et al. (2006). We did not find evidence for any such hierarchy based on transformation complexity. There were no differences in the levels of success between the subtract, and the add and reshape modes of tool making. Success on these modes in both age groups was relatively low. At best there was a 50% success rate in the older age group. However, there were three aspects of our findings and our data coding that warranted further consideration.

First, we considered the possibility that the low success rates in the new tasks could be explained by the crafted nature of the materials presented, which may have prevented children from modifying them. However, this explanation seems unlikely given that children perform at the same level on both their first and second tasks. If children were unaware they had permission to alter their materials we would expect children's success rates to dramatically improve on their second task after they had manipulated materials in their first task (see also Cutting et al., 2011).

Second, the only evidence that appeared to support a hierarchy was that the detaching task was substantially easier than the three other modes of tool making. There are two alternative explanations for the high success rates on these tasks. The first explanation is that the operation of detaching is truly easy. The high success rates for the detaching tasks are consistent with the comparative literature (both capuchins and great apes succeeded in the task when it was given in a manufacturing rather than innovation format (Visalberghi et al., 1995), and fit with the complexity hierarchy in the sense that it is easier than adding or reshaping. However, Kacelnik's hierarchy also predicts that subtracting should be as easy or difficult as detaching (Kacelnik et al., 2006). An alternative explanation is that these tasks did not truly represent detaching. Although we based our task directly on one previously used in the comparative literature (Visalberghi et al.), it is a concern that this does not in fact correspond to detaching as defined by B. Beck (1980). B. Beck defines detaching as the severation of a fixed attachment between objects, and examples given include breaking a branch off of a tree. Merely separating a tool from a bundle does not really capture this mode of tool manufacture, which appears to be more like a tool-selection task. If we take this second explanation, then the 'detaching' task in this current work more closely represents Kacelnik et. al.'s (2006) baseline Level 0, where unmodified materials are used as a tool. If this is the case then these results indicate significantly greater difficulty for tasks requiring tool modification (Kacelnik et al.'s Levels 1 and 2) compared with those that do not (Level 0). This provides support for the first stage of Kacelnik's complexity hierarchy but does not provide evidence for levels of complexity in tool manufacture or innovation.

Of course, if we interpret this "detach" task as measuring tool selection rather than tool modification this calls in to question claims that non-human primates were making tools in the task used by Visalberghi et al. (1995). Perhaps a better example of a detaching task would be one more closely modelled on behaviour seen in the wild, such as the detaching of branches from trees by elephants, which they then used as fly switches (Hart, Hart, McCoy &, Sarath, 2001). Future work in both children and non-human primates might take this as inspiration for developing better tests of detachment operations.

Third, one might question whether the horizontal-tube-add task results may not have given a true indication of children's abilities. Children were only coded as successful on this task if they added materials together outside of the tube, as we felt that adding extra material once a piece had been inserted did not represent true tool making. This poses a problem as children were not instructed to make the whole tool prior to entry, and so children that were coded as unsuccessful may have been successful if they were aware of this. However, children also performed poorly on the vertical-add task, and so our results appear to be representative of the difficulty of this mode.

In sum, we do not believe that any of these considerations affect our conclusion that children in our study failed to show a clear hierarchy of difficulty for tool innovation. However, the current findings extend knowledge by suggesting that innovation difficulties are not limited to a single mode of tool making, and are not due to the specific material used in the tasks.

Our second aim was to conduct the first study to investigate children's learning about tools from demonstrations, and to examine whether the ease of such learning followed the hierarchy described by Kacelnik et al. (2006).

We introduced a two-stage demonstration procedure. After failure on the innovation part of the task, the participant was first shown the target tool that they needed to produce, and then, if required, were shown the action needed to make that tool. Older children benefitted more from both demonstrations than younger children. Although there has been much research on children's observational learning of tool use (McGuigan & Whiten, 2009; Meltzoff, 1995; Want & Harris, 2002), these results are the first steps in understanding children's observational learning of how to manufacture tools. Only about half the children succeeded having seen just the target tool, whereas the rest also needed to see the required action.

Comparison of success rates following the first 'target-tool' demonstration reveals that children were more successful for certain modes of manufacture than others. However, the same pattern was not seen across the two tasks. For the vertical task children were more successful after the first demonstration for the reshaping mode than the adding mode, but no differences were seen between the reshaping and subtracting or the subtracting and adding modes. In contrast, for the horizontal task children were more successful following the first demonstration for the adding task than for the subtracting task, but no differences were observed between the subtracting and reshaping or adding and reshaping modes. Notably, these varied patterns add to the view that there is no systematic hierarchy in children's difficulty with different modes of tool manufacture. Instead, a possible reason for the differences between the two tasks could be that the relation between the target tool and the raw materials the participant possesses may be clearer in some cases than others. Success in tool manufacture may be dependent on children's ability to recognize the relationship between their raw materials and the target tool. For example, in the horizontal-tube experiment the demonstration for the add mode may be a clearer transformation from the materials the participant possesses (3 blocks combined together to make a long stick versus 3 small individual blocks) than the demonstrated tool for the subtract mode (straight stick versus straight stick with smaller pieces inserted). Despite these differences the high success levels following the demonstrations highlight how much easier it is to learn how to manufacture a tool from others than it is to innovate a tool for oneself.

Altogether this current work makes two important advances on our understanding of children's tool-related behaviour. First, children showed surprising difficulty in innovating novel tools via a variety of methods and with a number of materials, and this showed little evidence of following the hierarchy of tool difficulty described by (Kacelnik et al., 2006). Second, this work demonstrates that success following tool manufacturing instruction increases with age and may be affected by the transparency of the transformation required. These findings pave the way for future research into children's social learning of tool manufacture. Finally, this work, with other recent studies, confirms that innovation is a difficult and late-developing aspect of tool-related behaviour. Though late developing, our ability to overcome this challenge could explain how human tool behaviour has advanced so far beyond our nearest relatives.

Chapter 4

Is tool making complexity due to the transformation required?

4.1 Experiment 4: Introduction

Chapter 3 explored Kacelnik, Chappell, Weir and Kenward's (2006) suggestion that different methods of tool making differ in complexity. Three different modes of tool making were compared on two different tasks; a vertical-tube task requiring a hooked tool, and a horizontal-tube task requiring a straight tool. No differences in success levels between the different modes of tool making were found within each task (vertical or horizontal tube). It was concluded that all tool-making modes were equally difficult for young children. In this current chapter tool-making complexity was explored further. This chapter investigated the suggestion that comparing tool-making modes was too broad a categorisation of methods, and instead attention should focus at a lower level of tool making.

In the main analyses in chapter 3 the different tool-making modes were only compared within each task. Chapter 3 did not consider how difficulty of modes (i.e. subtracting/adding/reshaping) may vary across tasks (i.e. vertical-tube/horizontal-tube). In additional analyses, reported in chapter 3 but not discussed, (see chapter 3, section 3.3.3) the different modes of tool making were compared across the two task types. These additional analyses found no difference in success levels on the different tasks for the subtract or add modes of tool making, but did find a difference in the levels of success for the reshaping mode, with the horizontal-tube task that required unbending being significantly easier than the vertical-tube task that required the bending of a hook. This is a stable finding that has been demonstrated in previous experiments (Cutting et al., 2011, experiments 1 and 2). In this current chapter we aimed to explore why there might be differences in success rates within the different tool-making modes for some tasks but not for others.

We suggest the possibility that in chapter 3 we were focusing at the wrong level of tool making. Perhaps the level of tool-making mode was too broad and instead attention should be focused at a lower level, because a different level might be the determining factor for task difficulty. We propose three levels of tool making (see Figure 4. 1). The first and highest level being the general level of tool-making mode as defined by B. Beck (1980). B. Beck defined four types of tool-making; detaching, subtracting, adding and reshaping (see chapter 3 for full details of B. Beck's work). The second level proposed is the transformation of materials required to create each tool, for example for the reshaping mode this could be bending or unbending. Lastly, we suggest a third level termed action that refers to the specific action required to create each tool, for example two tasks may require bending, but one may require the specific action of bending a pipecleaner in to hook, whereas another may require bending of a different material in to a hook or bending a material in to another type of tool. Tool-making complexity may lie at the level of transformation or at the level of action.

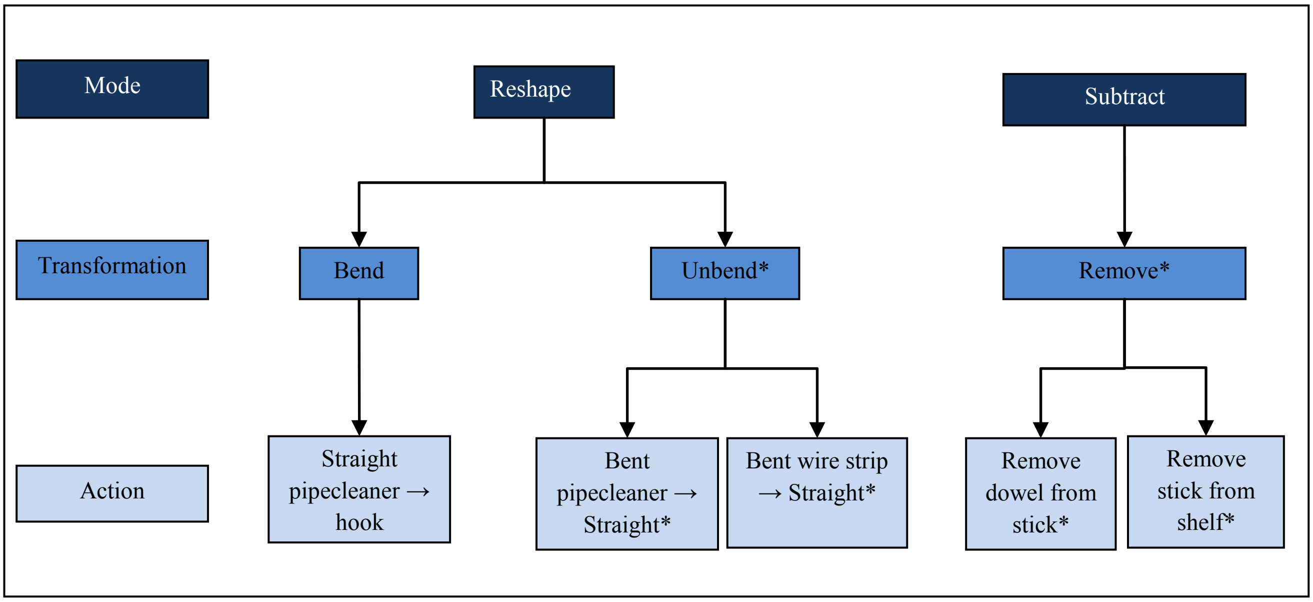

Figure 4. 1 *Diagram to show the different levels of tool making. * denotes the transformations and actions used in the current experiment.*

The tool-making modes tested in the two tasks in chapter 3 can be examined more closely in terms of the transformations and specific actions that were required to create each tool. Comparing the transformations and actions required for the same mode of tool making across the two tasks may provide some insight in to the differences in success levels that were seen. First, for the tools made via reshaping, the specific transformations required on each of the tasks were different. In the vertical-tube task, participants were required to take a straight pipecleaner and transform it by bending it in to a hook, conversely in the horizontal-tube task, participants were required to take a bent pipecleaner and transform it by unbending it in to long tool. The mode of tool making for both of these tasks was the same – reshaping, however, they clearly differ in the type of transformation required – bending vs. unbending. It is therefore possible that the reshaping tasks in chapter 3 differ in difficulty due to differences in the transformation required. Further evidence for the potential importance of transformation comes from the subtracting tasks in chapter 3. In the subtracting versions of the tasks children were required to alter the materials using the same transformation - removing. An example of a different transformation for the subtract mode of tool making would be breaking or snapping off parts of a tool to leave a more functional core behind. Children performed comparably on the two subtracting trials in chapter 3 and it is possible that this was because they were also the same at the level of transformation. These findings suggest the possibility that differences in the level of difficulty found in tool-making could lie at the level of transformation rather than at the level of tool-making mode.

From the chapter 3 findings however we cannot disentangle whether difficulty lies at the level of transformation (e.g. bending vs. unbending) or at a lower level we have termed action (see Figure 4. 1). The action level refers to the specific manipulation that is carried out to form a tool. In the case of the two subtraction tasks in chapter 3, both the transformation and the action required by the child were very similar if not identical. Both tasks required children to not only transform the materials via removal, they both required the specific action of removing a piece of dowel from a stick. The identical transformations and actions required in these tasks means it is not possible to disentangle whether the similar success levels found in these two tasks was determined at the level of transformation or at the level of specific action.

The two reshaping tasks reported in chapter 3 present the opposite confound. These two tasks not only differed at the level of transformation (one required bending and the other unbending) they also differed at the level of specific action, straight pipecleaner to hook or bent pipecleaner to straight. In this case it is not possible to disentangle whether the differences seen between the two reshaping tasks were due to them differing at the level of transformation, or due to them differing at the action level.

In order to discover at which of these levels tool-making complexity might lay, in the current study we held mode of manufacture and transformation constant and varied the specific actions required. This design produced tasks that separated transformation and action allowing for direct comparison of these components that had previously covaried in chapter 3. Children were tested on four tasks. The tasks were split in to two sets that differed on the mode of manufacture required – subtracting or reshaping. Having two sets of tasks provided us with two opportunities to explore tool-making complexity. Each set of two tasks were identical at the level of mode and level of transformation (subtracting tasks required removal, reshaping tasks required unbending) but they required different specific actions. If children found one of the tasks in each set more difficult than the other, this would tell us that difficulty was due to the specific action that the task required. However, if children performed comparably on the two tasks in each set then this would suggest that task difficulty was due to either the tool-making mode or the type of transformation.

We also manipulated another dimension at the level of action which we termed directness. Within each pair, one of the tasks required children to make a tool that was needed to act directly on the target, whereas the other task required the tool to act indirectly by altering the structure of the apparatus. Evidence from studies with chimpanzees suggests that directness of action might alter task difficulty.

Chimpanzees were significantly hindered in their success on trap-tube problems when they had to use a tool and could not retrieve the target with their hands. A trap-tube task involves the subject retrieving a reward from a tube. The reward is placed in the middle of the tube and there are two traps, one at each end of the tube. During a trial only one of these traps is functional, subjects must therefore work out which is the functional trap and avoid it by pushing the reward out of the other end of the tube. Seed, Call, Emery & Clayton (2009) tested chimpanzees on two types of trap-tube problem. One version required subjects to retrieve the reward by pushing it out of the tube with a stick tool. The second version had finger holes along the length of the tube which enabled chimpanzees to move the reward with their fingers and as such did not require a tool. Chimpanzees were much more successful when they did not have to use a tool and they were able to act directly on the target reward. This finding suggests that difficulty on physical problem solving tasks could be moderated by the distance between the participant and their means for acting on the target. Acting more directly on the target reward is easier.

In the current study the distance between the subject and the means for acting on the target was moderated by the directness of the action. In two of the tasks children were required to make a tool which could act directly on the target object, i.e. children could physically move the target with the tool. In the other tasks children were required to make and enter something which altered the structure of the apparatus and made it possible to retrieve the target, but would not form a physical link between the subject and the target, i.e. children could not push or pull the reward towards them using the artefact they had made. Based on the findings of Seed et al. (2009) it is possible that children would be able to solve the physical problem we give them more readily if the solution enabled them to act directly on the target with the tool. Creating something which acts further away and isn't merely an extension of the hand that can contact the reward may be much more difficult for children. If children perform differently on the direct and indirect tasks, we will need to perform a second experiment to determine whether differences were due to directedness or due to the different specific actions.

4.2 Method

4.2.1 Participants

Twenty-one 4-to-5-year-olds (11 boys) mean age 5 years 4 months (5; 4) (range 4; 11 – 5; 10), and 21 5-to-6-year-olds (9 boys) mean age 6; 4 (range 5; 10 – 6; 10) participated in the study. Children were recruited from an infant school in South Birmingham that served a working and middle class population.

4.2.2 Materials

The apparatus and materials used for the hooks and tube tasks were identical to those used in the subtract version of the hooks task and the unbending version of the tube task from chapter 3.

The apparatus for the shelf task consisted of a clear plastic box (height = 22.5cm, width = 18.5cm, depth = 8cm), which had a hole (5.5cm x 4cm) cut into the middle of the front face. A rubber ball (4cm diameter) with sticker attached was placed on a shelf (5.5cm x 8cm) inside the top left hand corner of the apparatus. There was a hole (2cm x 3cm) on the left-hand side of the box to allow children access to push the ball. Inside the box were wooden sticks attached from the back to the front, on which the cardboard shelf (length 5.5cm, width = 5.5cm) could be placed. The reward could be retrieved by inserting the piece of cardboard through the hole and balancing it on the wooden sticks such that a shelf was created. Once children had pushed the ball through the top opening the cardboard shelf would ‘catch’ it and children could retrieve the ball through the front opening. The cardboard shelf was presented with a piece of dowel (length = 10cm, diameter = 0.5cm) through a hole in the middle. The distracter item was a piece of blue cloth (length = 5.5cm, width = 5.5cm). A plastic barrier (length = 27.5cm, height = 22.5cm) was attached to the left-hand edge of the box to prevent children from pushing the ball and catching it with their other hand.

The apparatus for the bridge task consisted of a clear plastic box (length = 20cm, height = 12cm, depth = 14cm), mounted in an open box (length = 31cm, width = 18cm, height = 5.5cm) on top of a piece of dowel (diameter = 1.5cm) to allow the plastic box to pivot. Inside the plastic box at both ends (left and right) were mounted two ledges (length = 8cm, width (4.5cm), on the left hand-side of which a rubber ball (diameter = 4cm) was placed. On the right-hand side of the box was a slot (height = 4.5cm, length = 7cm) from which the ball could be retrieved. There was also a slot along the front of the box (length = 16cm, height = 2cm), into which materials could be inserted. The materials presented to solve the task were a silver bendy strip (length = 15cm, width = 8cm) made from duct tape with lengths of wire inside, and a blue piece of cloth (length = 15cm, width = 8cm).

The materials for the warm-up exercise were a black bendy strip (length = 6cm, width = 6cm), black cloth (length = 9cm, width = 9cm), a 5cm piece of dowel (diameter = 1.5cm) with a hole drilled in one end, a 4cm piece of dowel (diameter = 0.5cm), a piece of cardboard (length = 4cm, width = 2cm) with a hole in one end, a green pipecleaner (length = 20cm) and a piece of black string (length = 44cm).

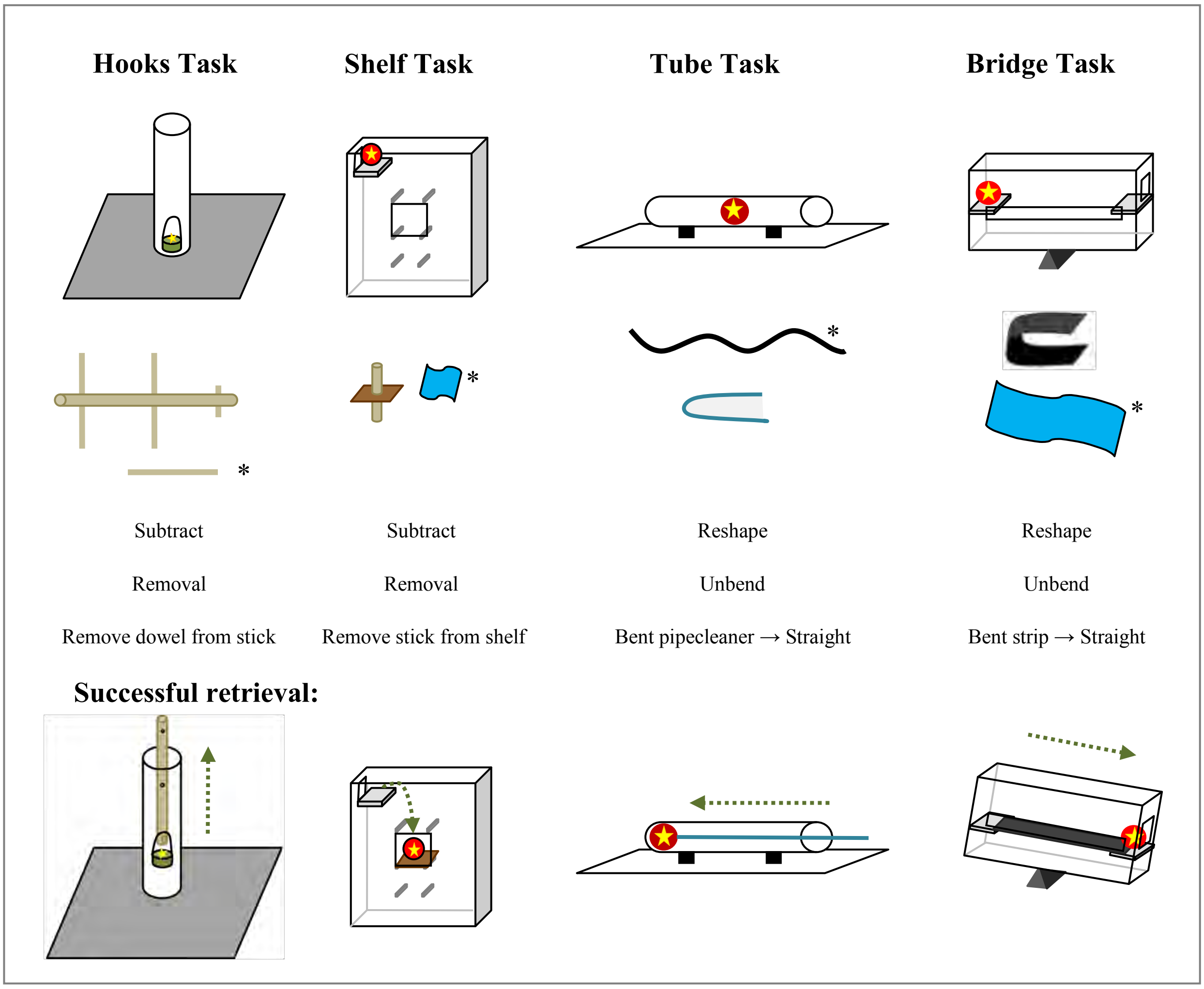

Figure 4. 2 *Apparatus used for the four tasks.*
*Note: * indicates distracter material, green arrow indicates direction of movement.*

4.2.3 Procedure

4.2.3.1 Warm-up

Children first received a warm-up exercise in which they experienced each of the materials. Children were shown versions of each material that were used in the main task one at a time and were given demonstrations of their properties before being able to handle them themselves. The materials used in this warm-up phase differed in colour, size, and in some cases shape from those used in the main task. The pipecleaner and bendy strip were introduced as being ‘bendy’ and bent in the middle to demonstrate. The string and cloth were introduced as being ‘wiggly’ and then shaken to demonstrate. The small dowel along with the card and the larger dowel each with a hole drilled through were introduced with the experimenter saying look what I can do with these, and the small dowel was passed through the holes in the larger dowel and the cardboard.

4.2.3.2Main Task

Children received all four innovation tasks. The tasks were presented so that children alternated between the types of transformation needed; the tasks were also grouped so that the two direct action tasks and the two indirect action tasks were performed together. The order of presentation was counterbalanced across participants.

For all tasks the experimenter drew the child's attention to the sticker in the bucket (hook task) or attached to the ball (shelf and bridge task) or pom-pom (tube task), she then told the child *'If you can get that out of there you can win that sticker'* and then traced with her finger the route required to get the sticker out. For the shelf and bridge tasks children were given additional information to ensure they were aware of how the apparatus worked. In the shelf task children were told *'you can poke the ball through this hole here* (point to hole at top of apparatus) *and it's your job to try and get it out of this hole here* (point to hole on front of apparatus)'. For the bridge task the additional information children received was *'You need to get the ball out of this hole here, and you can tip this like this* (demonstration of tipping)'. After these instructions children were told on all tasks *'here are some things that can help you'* and the experimenter brought out the relevant materials. Children were given one minute to interact with the apparatus to try and solve the task. If they were not successful during this time they were encouraged to put down the materials they were using and then taking her own materials the experimenter said *'look at this'* and showed children the endstate-tool they were trying to achieve. Children were given a further 30 seconds to try and complete the task. If they were still unsuccessful the experimenter again encouraged them to put down their materials and then after saying *'watch this'* demonstrated, with her own materials, the action that needed to be done with the materials to make the required tool. Children were then allowed to interact with the apparatus again.

4.3 Results

There were no gender differences (lowest p = .220) and no differences with age (lowest p = .102) on any of the four tasks, as such all data were collapsed into one sample. Order of presentation also made no difference to success levels (lowest p = .183).

At all stages of the tasks children were coded as being successful only if they made the required tool and then used it correctly to retrieve the target. A Cochran's Q test was used to compare children's performance across all tasks. McNemar tests were then used to make individual comparisons with a Bonferroni correction applied (p <.008). Comparisons were made between children's performance on tasks requiring different specific actions for each mode/transformation type individually. This enabled us to discover whether the specific action or direct vs. indirect actions varied performance. Second we compared performance across mode/transformation type to determine whether children performed comparably on all tool-innovation tasks.

Table 4. 1 *Numbers of successful children in experiment 4 for all tasks pre- and post-demonstrations.*

Mode/ Transformation	Task	N	Success pre-demonstration	Success after endstate demo	Success after action demo
Subtracting/ Removing	Hooks	42	13	14	12
	Shelf	40	15	11	14
Reshaping/ Unbending	Tube	42	32	8	2
	Bridge	41	27	10	4

Significant differences in children's success on the four tasks pre-demonstration were found, Cochran's Q test = 24.946, $p < .001$ (see table 4.1). McNemar tests were then used to compare performance for the two tasks requiring different actions within the same mode/transformation. There was no difference in success levels between the two subtracting/removing tasks which required different specific actions, McNemar test, $p = .424$. Similarly for the reshaping/unbending tasks, there was no difference in success levels between the two tasks requiring different specific actions, McNemar test, $p = .332$. As well as differing in terms of the action required, the two tasks within each mode/transformation also differed based on whether they acted directly or indirectly on the target object. A lack of difference in success levels between the two tasks within each set suggests no difference in complexity based on directness of action.

Next we compared performance across the two different tool-making modes/transformations in order to discover whether children performed comparably on all tool-innovation tasks. First we compared success across modes/transformations by grouping tasks based on directness of action. For the two tasks requiring children to act directly on the target, the reshaping/unbending task (tube unbend) was significantly easier than the subtracting/removing task (hook removal), McNemar test, $p < .001$. When comparing tasks requiring children to act indirectly, a trend was found that the reshaping/unbending task (bridge unbend) was easier for children than the subtracting/removing task (shelf removal) but this did not reach significance after applying a Bonferroni correction for the 6 multiple comparisons (the alpha level was set at $p < .008$), McNemar test, $p=.031$. Further evidence that the reshaping/unbending tasks were easier to achieve comes from comparisons between the hook subtracting task and the bridge reshaping task, McNemar test, $p = .001$, and between the shelf subtracting task and the tube reshaping task, McNemar test, $p = .001$. These results demonstrate that children did not simply perform comparably on all tool-making tasks, the reshaping tasks were found to be significantly easier.

Coding success as outlined above does not give a full picture of the complexity of behaviours observed during the tasks. Table 4.2 shows that many children transformed the materials in to the required tools but then failed to use the tool correctly to retrieve the target. In the analyses above these children would have been coded as unsuccessful which would put them in the same category as children that made no such attempt to transform materials. By coding behaviours in more detail we are given a richer insight into how children behave when confronted with tool-innovation tasks. In table 4.2 children were categorised as either succeeding on the task (making a tool and using it correctly), transforming the materials (transformed correctly but did not use to retrieve target) or being unsuccessful (did not transform correctly or retrieve target). The most noticeable pattern from table 4.2 is that, except for one child, all children who transformed the materials in to the required tool but then failed to use the tool correctly to retrieve the target did so in the indirect tasks (shelf and bridge). Examination of table 4.2 shows that in the innovation phase of the tasks (pre-demonstration) only a few children transformed materials but did not use them correctly. Of these children those in the subtraction tasks then went on to use the subtracted piece (the dowel) on the apparatus. In the bridge reshaping task children who correctly transformed the material either entered the tool in the wrong slot or entered the narrow end of the material and tried to carry the ball from one platform to the other.

The most interesting differences in transformation behavior were seen following the endstate-tool demonstration. For the two subtracting tasks similar levels of success are seen following this demonstration for the two tasks. However, whereas children in the direct task made no attempts to make the required tool, the majority of children in the indirect task transformed the material correctly but did not then use the tool they created to retrieve the target. A similar pattern may exist for the reshaping tasks, however due to lower numbers of children requiring demonstrations this is not clear. Possible reasons for the differences in behavior between the direct and indirect tasks will be addressed in the discussion.

Table 4. 2 *Number of children succeeding and transforming materials in experiment 4.*

		Success								
		Pre-demonstration			After end-state demo			After action demo		
Mode	Task	No	Trans-formed	Yes	No	Trans-formed	Yes	No	Trans-formed	Yes
Subtract										
	Hooks	28	1	13	15	-	14	3	-	12
	Shelf	21	4	15	1	13	11	1	-	13
Reshape										
	Tube	10	-	32	2	-	8	-	-	2
	Bridge	9	5	26	1	3	10	-	1	3

4.4 Discussion

The current chapter built on the findings from chapter 3 and queried whether Kacelnik et al.'s. (2006) suggestion that tool-making complexity is determined at the level of tool-making mode also applies to humans. In chapter 3 no differences in success levels were found between different modes of tool making within two different tasks. The current study investigated whether the study in chapter 3 was focusing at the incorrect level of tool making. Additional analyses comparing the tool-making modes across tasks found differences in success within the different modes, however, it could not be determined if these differences were determined at the level of transformation or the level of specific action. The current study manipulated action within two types of mode/transformation in order to discover whether complexity is determined at the level of action. No differences in success levels were found suggesting that tool-making complexity is not determined by the specific action a tool-making episode requires.

In the current work children were tested on four tasks, two subtracting and two reshaping. Within each pair, the tasks were identical at the level of tool-making mode and transformation but differed in the specific action required. No differences were found between the success levels within each pair. There were however differences in success between the different types of mode/transformation, with the reshaping/unbending tasks being significantly easier than the subtracting/removal tasks. This suggests that children did not merely find all tasks equally difficult.

At first glance the finding that the two unbending tasks were easier to achieve than the subtracting tasks appears to contradict the findings from chapter 3 that found no hierarchy of tool-making mode complexity. In chapter 3 the different tool-making modes were compared within the same task. In the current study different transformations, which were purposefully confounded with mode due to the findings from chapter 3, were compared across different task types. As the success levels for the different modes arise from different tasks, it is not possible to directly compare children's success rates for the different tool-making modes as there are many different factors that may be influencing the results.

The current work also tested how 'distance' between the participant and the target influenced tool-innovation difficulty. Children were given tasks in which the 'tool' to be made either acted directly on the target object or indirectly by altering the structure of the apparatus. As well as testing whether directness alters task difficulty this manipulation may have further-reaching consequences. Some researchers have recently suggested a narrower definition of what constitutes tool use. A new definition suggests that to be a tool an object must have a '*dynamic mechanical interaction*' (St. Amant & Horton, 2008, p. 1203) that is the tool must act directly on a target, i.e. it must be an extension of the hand,

and not merely something that can be placed somewhere and left to its own devices. The design of the current study had potential to provide supportive evidence for this narrow definition. Comparison of direct vs. indirect tasks could provide insight in to differences between children's ability to innovate tool and non-tool objects. The hook and tube tasks clearly required the making of a tool as they acted directly on the target, whereas the shelf and bridge tasks required the creation of a non-tool object, as they acted indirectly by altering the structure of the apparatus. No differences between indirect and direct actions for tasks requiring the same transformation were found. This could be taken as evidence against a narrower definition of what constitutes a tool or alternatively it could provide evidence that there is nothing particularly special about tool use. The complexity of innovation may be comparable for all tasks requiring physical cognition and manipulation of materials whether or not the end product turns out to be a tool. Whatever definition of a tool one decides to use, the clear finding from the current study is that there is no difference in difficulty between the innovation of artefacts that act directly or indirectly on a target.

The directness of the action did not affect children's success rates, however, differences in children's behaviours were observed. In the indirect action tasks many children transformed the materials correctly but did not then go on to use the tool they created to retrieve the target. This was most noticeable following the first endstate-tool demonstration. As no differences in success levels were seen in either the innovation phase or following the endstate-tool demonstration this finding appears odd. One potential explanation could be that the transformation required was clearer to see in the indirect tasks. For example in the indirect subtracting tasks children were shown the piece of cardboard (the shelf) they needed to achieve. It may have been easier for children to see how this had been transformed from the initial dowel and cardboard they were presented with than it was for children to notice that a piece of dowel was missing from the hooks tool. In the case of the shelf tool, the tool was made up from only two materials, cardboard and dowel, it would therefore be easy to recognize that one of these pieces was missing and work out the transformation required. Conversely, the hook tool was made up of four pieces and therefore it would have been more difficult for children to realize which piece was missing and how the materials would need to be transformed.

The suggestion that there was a clearer transformation in the indirect tasks may address why children more readily transformed the materials but it does not address why there is no difference in success rates. If children have managed to transform the materials correctly then why did they not then go on to use the created tool to solve the task? Previous tool-making studies have shown that once a child has created a tool they almost always go on to use that tool correctly and retrieve the target (Beck et al., 2011; Cutting et al., 2011). We suggest these children are unsuccessful in retrieving the target in the indirect tasks due to there being a greater number of potential actions that could take place on the apparatus. In the direct tasks once a tool has been made there are very few options as to what the child can do with the tool. In the hooks task there is only one opening with which to insert the tool and in the tube task it does not matter which opening one chooses as entering the tool will always contact the target. In contrast the shelf and bridge tasks each have two openings, only one of which can be utilized effectively. This provides the child with more options as to where to insert their tool and as such the correct course of action may be less obvious. We suggest that these greater degrees of freedom could explain why in the indirect tasks children did not use their correctly transformed materials to retrieve the target reward. We will return to this point in more detail in the general discussion chapter.

4.5 General discussion of chapters 3 and 4

Taken together the results from chapters 3 and 4 lead to the tentative suggestion that tool-making complexity might be determined at the level of transformation, i.e. removing or unbending, but is not dependent on the general level of tool-making mode or the specific action required. First, the findings from chapter 3 suggested that tool-making complexity is not determined by the mode of tool making. No differences were found between the subtract, add and reshape modes of tool making when compared within a task. Additional analyses from this chapter found differences in success levels within the different tool-making modes when compared across the two different tasks for some modes but not for others. This finding led to the suggestion that tool-making complexity might be determined at a lower level.

In chapter 4 two lower levels of tool making were suggested: transformation and action. As these two levels were confounded in chapter 3, experiment 4 separated them by holding transformation and mode constant and varying the actions that tasks required. Experiment 4 found no difference in success levels between tasks requiring the same mode/transformation and different specific actions. This suggested that tool-making complexity is not determined at the level of action; however from this study it was not possible to disentangle mode and transformation.

No single study could directly address where tool-making complexity might lay due to the three levels of tool making being determined by each other, for example it would be impossible to hold action constant but vary the transformation. However, if results from experiments 3 and 4 are taken together, along with previous research showing there to be a stable difference in success levels between the hooks and unbending reshaping tasks (Cutting et al., 2011, and additional analyses, chapter 3), these findings suggest that complexity is determined at the level of transformation. There is no difference in complexity between different modes of tool making on a singular task, but there are consistent differences between tasks using the same mode but different transformations. Additionally when holding the mode and transformation constant specific action made no difference to complexity.

The theory that complexity is determined by transformation supports the suggestion made at the beginning of this chapter that the previous work on complexity in chapter 3 and suggestions by Kacelnik et al. that focused attention on tool-making mode were too broad. Many different tool-making episodes may require the making of a tool through reshaping for example, but it would be surprising to find that all of these were equally difficult. For example reshaping material could not only involve bending as in the hook experiments (Beck et al., 2011) or unbending as demonstrated here and in previous work (Cutting et al.), but could include an infinite number of tools that could be created in an infinite number of ways. To think that innovating the solution to bend a hook is comparably difficult to innovating the solution of scrunching up leaves to use as a sponge just because they require the same tool-making mode (reshaping) not only appears implausible but is also improbable given the current findings.

Altogether chapters 3 and 4 advance our understanding of tool-making complexity. Complexity is not simply determined at the general level of tool-making mode as previously suggested (Kacelnik et al., 2006), but is more likely dependent on the transformation of materials that is required. This finding gives us an insight into which elements make tool innovation so difficult for young children. Knowing that the transformation required determines tool-making complexity would allow for future research to investigate why certain transformations may be more difficult for children to achieve than others. The new terminology we have used in this chapter could also be applied to the non-human animal literature to better define and compare the tool-making episodes that have been observed in other species.

Chapter 5

The Role of Executive Function in Tool Innovation

5.0 Overall chapter summary

The current chapter investigated the role that executive function may play in tool-innovation tasks. In experiment 5 we investigated whether children's poor performance in tool- innovation tasks is due to impulsive behaviours. In this study we attempted to decrease impulsive behaviours in children by introducing a delay before children encountered the innovation task. Additionally children were given the opportunity to explore the materials prior to engaging in the main task, as there is anecdotal evidence that children do not take time to do this independently.

In experiment 6 we investigated the role of executive functions in tool innovation in more detail by examining children's performance on a battery of executive function tasks in addition to the hook-innovation task. Children were tested on a series of tasks testing their abilities in inhibitory control, task switching, working memory and ill-structured problem solving to see whether any of these abilities correlated with performance on a tool-innovation task.

5.1 Experiment 5: Impulsivity and Exploration.

Being able to stop what one is doing is termed inhibitory control and is a key component of executive function (Miyake et al., 2000). In previous studies we have used perseveration as an indirect measure of children's inhibitory control. In chapter 3 we defined perseveration as children entering an unsuccessful tool/material in to the task apparatus and then failing to stop using that material for the whole time period. Overall we observed very little perseverative behaviour in young children during our tasks. As such we ruled out the possibility that children's poor tool-innovation performance was due to an inability to overcome perseveration of incorrect strategies. Many children in these studies did not display any perseverative behaviour, but were still unable to successfully innovate tools. Additionally, for many successful children they had no need to overcome perseveration because they immediately innovated a successful tool without implementing an unsuccessful initial solution. It was concluded that overcoming perseveration is not the limiting step for children to succeed in these tasks; however it is a necessary obstacle to overcome. In the current chapter we explored inhibitory control more directly by examining how impulsiveness can affect children's performance on tool-innovation tasks.

Impulsivity is a multi-dimensional construct that is characterised by an inability to wait, failure to inhibit inappropriate actions or behaviours, and the inclination to act without forethought or any consideration of the consequences (Reynolds, Ortengren, Richards & De Wit, 2006). Whereas the measure of perseveration in chapter 2 measured children's inhibitory control by assessing ability to stop implementing an unsuccessful action, in the current study we use impulsivity as a wider measure of inhibitory control which is most characterised by children's tendency to act quickly without any thought of the consequences.

It has been observed by the experimenter during tool-innovation tasks that children have a tendency to pick up materials and act on the apparatus immediately after receiving instructions. Very few children appeared to take time in considering their options. We suggest that children's difficulty on tool-innovation tasks may be due to their impulsive behaviour, more specifically their inability to inhibit using the presented materials immediately without any thought of the effects the materials will have or any consideration of how they might modify the materials.

Children may act impulsively on tool-innovation tasks because they allow their attention to be caught by irrelevant stimuli (Schachar & Logan, 1990), or they may act on naïve theories which prevent them from thinking through their actions more rationally. Children's poor performance due to naïve theory is demonstrated in Karmiloff-Smith & Inhelder's (1975) balance rod experiment in which children aged 6 who were in the grips of a naïve theory performed poorly in comparison to both older and younger children. In this task children were presented with a series of unevenly weighted rods that needed to be balanced on a pivot. Four-year-olds had no theory with which to approach the task, and as such adopted a flexible trial-and-error approach, successfully completing the task. Children aged six and eight had a theory that all rods must balance in the middle, a strategy that did not work with the weighted rods they were presented with. Eight-year-olds were able to overcome this theory and flexibly adapted their actions to succeed on the task, however, the six-year-olds who were in the grip of their naïve theory were unable to overcome their conviction that rods must balance in the middle and as such performed poorly, often giving up.

Further evidence for the detrimental effects of young children's naïve theories comes from studies reporting children's gravity errors (Hood, 1995; Hood, 1998; Hood, Wilson & Dyson, 2006). In tasks requiring children to search for a ball dropped down one of three opaque tubes which were interwoven and then each connected to three hiding boxes, children were shown to persistently look for a ball directly underneath the tube where it had been dropped. The 2-to-4-year-olds in these tasks were argued to be in the grips of a naïve theory surrounding gravity. They had a belief that all objects must fall directly downwards and were so entrenched in this theory they did not consider the effects of the tubes on the outcome. Support that these findings were due to a naïve theory surrounding gravity comes from evidence that children performed successfully on the tasks when the direction of the ball was reversed, i.e. the balls were sucked up the tubes (Hood, 1998). Naïve theories are not only limited to children, many adults possess naïve theories, and often answer physics problems incorrectly (McCloskey, 1983).

In our hook-innovation task it is possible that children have a naïve theory about how to solve the problem. All children in our tasks display behaviours that demonstrate their understanding about the need for contact, i.e. all children immediately use the materials to contact the bucket containing the reward; perhaps children act on a naïve theory surrounding contact. Children are aware of the need for contact from a very young age, with infants as young as 6-months demonstrating this awareness (Leslie & Keeble, 1987; Oakes & Cohen, 1990). Children may become so entrenched in their theory about contact that they are unable to inhibit their prepotent response of merely contacting the bucket containing the sticker with the materials they are given. Once children have acted impulsively on this naïve theory they may not be able to inhibit their actions and think about the problem in a different way. This may mean that they act perseveratively with one material; although as evidence from chapter 2 shows this is not usually the case, or may mean that they use multiple materials and strategies but ultimately are confined by their naïve theory of merely contacting the bucket.

In the current study we aimed to reduce children's impulsive behaviour by introducing a delay between the introduction to the task and their opportunity to act on the apparatus. Previous research has shown that a delay can significantly improve children's performance on tasks requiring inhibition. In a study using a typical inhibition task for children – the day-night stroop – Diamond, Kirkham and Amso (2002) found that introducing a short delay between showing children the card and asking for their response, significantly improved children's performance. Children were shown a set of cards that either had a sun or a moon on them and were required to respond day to the moon card and night to the sun card. This required inhibitory control, as children had to inhibit their initial reaction of what the cards depicted and respond with the opposite. Diamond et al. concluded that it takes children several seconds to compute the correct answer for the day-night stroop task, and making children take extra time to construct their answers significantly improved their success rates. This delay strategy has been used in other studies where inhibitory control is thought to aid performance. Beck, Carroll, Brunsden and Gryg (2011) introduced a delay by asking children to wait until a doll had gone down a slide before giving their responses to counterfactual questions. Children in the delay condition performed significantly better on counterfactual questions than children who did not experience the delay.

It has also been observed in previous tool-innovation studies that children rarely explore the materials prior to using them in the task. This may mean that children are unaware of the properties of the materials, or that the properties are not highlighted for them and so they do not consider them in the context of solving the task. Previous studies have provided structured exploration activities in the form of warm-up exercises in which children have wound the pipecleaner around a pen and placed the string on an s-shaped pattern for example (Beck et al., 2011; Cutting et al., 2011). However, research in to children's exploratory behaviours has shown that children learn more about materials when they are allowed to explore them for themselves rather than being shown ways they might work. Bonawitz et al. (2009) measured children's exploratory behaviours under different pedagogical conditions, and found that children who were shown one action that could be done with a toy in a pedagogical context, were less likely to explore and find other actions than children in the accidental demonstration and no demonstration conditions. It is possible that in studies where we have demonstrated actions with the materials that we have inadvertently reduced children's explorative behaviours. Children did not perseverate on the actions they were shown in demonstrations within the main task; however they may have thought that apart from using the materials as they were presented that this was the only thing that could be done with them. In the current study we allowed children to explore the materials for themselves to enable them to discover the materials' properties, which in turn may aid them in the innovation task.

In experiment 5 we gave half of children a delay before interacting with the task, during which they were given opportunity to explore materials. If children in the delay-explore condition were more successful at innovating the required tools than children in the control condition we planned to perform further studies to disentangle whether children were aided by the delay or the opportunity for exploration.

5.1.1 Method

5.1.1.1 Participants

The final sample consisted of 29 4- to 5- year olds (13 boys), mean age 4 years 8 months (4; 8), (range 4; 3 to 5; 2), and 24 6- to 7-year-olds (11 boys), mean age 6; 9 (range 6; 3 to 7; 2) from a Primary School in South Birmingham, UK. The ethnic composition of the sample was 91% Caucasian, 6% Black and 3% other/unknown.

5.1.1.2 Materials

The materials were identical to those used in Experiment 1, chapter 2.

5.1.1.3 Procedure

Children were alternately assigned to either the control group or to the experimental (delay/explore) group based on the teacher's class list. Participants were tested by a female experimenter in a quiet area just outside the main classroom. The child and experimenter sat facing each other across a table. All children received both the hook and unbending tasks. The order was counterbalanced across participants.

Children were shown the relevant transparent tube, either the vertical-tube with the bucket containing a sticker already in place in the bottom or the horizontal-tube containing the pom-pom with sticker attached. Children were then told that if they could get the sticker out they were allowed to keep it. The experimenter then brought out the string and pipecleaner (either straight or bent in half) and told the child *'Here are the things you're going to use. This one's bendy* (bend pipecleaner slightly and straighten) *and this one's* *wiggly* (wiggle string)'. The materials were placed on the base of the apparatus which was out of the child's reach. In the control condition, the apparatus was then put in front of the child and the experimenter said *'Can you work out how to get the sticker out?'* In the experimental condition after the demonstrated materials were placed on the base of the apparatus, the experimenter brought out identical new materials and said *'These are just the same as the ones you are going to use. But before you try to get the sticker for real, you can play with these and try to work out how you will get the sticker out'*. Children were given 10 seconds to explore the materials. The materials were then taken away and the apparatus and original materials were placed in front of the child and the experimenter said *'ok, can you get the sticker out?'* The children were then given one minute to try to retrieve the sticker. No feedback was given, but children were given neutral prompts if required. Examples of prompts used include *'Can you think how you might be able to get the sticker out?'* and *'Maybe you could use these things to help you.'* If, after one minute, the child had not retrieved the sticker, they were encouraged by the experimenter to put down the materials they were using. The experimenter then gave the child an endstate-tool demonstration followed by the action demonstration if required. See experiment 3 for full details of the demonstration procedure.

5.1.2 Results

There were no differences due to gender for either age on each task (Vertical-tube task: 4-to-5-year-olds, Fisher's Exact Test, $p = .123$, 6-to-7-year-olds, Fisher's Exact Test, $p = .192$; Horizontal-tube task: 4-to-5-year-olds, Fisher's Exact Test, $p = .688$, 6-to-7-year-olds, Fisher's Exact Test, $p = .408$) and so all data were collapsed across gender for subsequent analyses. We will first report descriptive data of the exploratory behaviours conducted by children during the delay. We will then report Chi-square analyses comparing performance between the experimental and control groups on both the vertical- and horizontal-tube tasks. After excluding successful children we then compared performance between the two groups following the two demonstrations.

Table 5. 1 *Behaviours displayed by children in the Delay/Explore condition during the delay.*

Age Group	N	Actions: Touch	Pick up	Combine	Bend (non-target tool)	Bend/unbend Target tool
Vertical-tube task						
4 to 5	14	2	8	2	1	1
6 to 7	11	1	2	1	4	3
Horizontal-tube task						
4 to 5	14	2	7	1	2	2
6 to 7	11	2	3	4	0	2

Table 5.1 displays the behaviours carried out by children during the delay/explore period. The younger children had a tendency to just pick up the materials, and explored them very little. Older children showed more exploratory behaviours such as combining and bending. These descriptive data show that children engaged with the materials in the delay/explore condition, meaning that they had a different experience to children in the control condition. Whether children were in the delay/explore condition or the control condition made no difference to children's first choice of material in either of the main innovation tasks. All except 3 children chose the pipecleaner first in the vertical-tube task, all children that chose the string were in the control group but there was no difference between conditions, Fisher's Exact Test, p = .238. Nine children in total chose the string

first in the horizontal-tube task, 4 children were in the delay/explore group and 5 were in the control group again demonstrating no difference between the two conditions, Fisher's Exact Test, p >.999.

For the vertical-tube task Chi-square analyses reveals no difference in success levels between children who were in the delay-explore condition and children in the control condition, χ^2 (1, N = 49) = .783, p = .376. This finding was seen in both age groups; 4-to-5-year-olds, Fisher's Exact Test, p =.596, 6-to-7-year-olds, Fisher's Exact Test, p > .999. There was also no difference in success levels between the two groups for the horizontal-tube task, χ^2 (1, N = 53) = .305, p = .581. Again this finding was the same across both age groups; 4-to-5-year-olds, Fisher's Exact Test, p > .999, 6-to-7-year-olds, Fisher's Exact Test, p > .697 (See table 5.2).

Table 5. 2 *Behaviours and number of children succeeding at different task stages.*

			Touched First		Used First		Success		
Age Group	Condition	N	Pipe-cleaner	String	Pipe-cleaner	String	Pre-demo	After Endstate Demo	After Action Demo
Vertical-tube task									
4 to 5	Delay	14	13	1	14	-	3	4	5
	Control	13	12	1	10	3	1	6	4
6 to 7	Delay	11	11	-	11	-	5	5	1
	Control	11	11	-	11	-	4	6	1
Horizontal-tube task									
4 to 5	Delay	14	9	5	10	4	4	5	3
	Control	15	12	3	10	5	5	3	3
6 to 7	Delay	11	7	4	11	-	4	5	1
	Control	13	13	-	13	-	6	3	4

Condition made no difference to success levels following the first endstate demonstration for either age group on both tasks (Vertical-tube task: 4-to-5-year-olds, Fisher's Exact Test, $p = .680$, 6-to-7-year-olds, Fisher's Exact Test, $p > .999$; Horizontal-tube task: 4-to-5-year-olds, Fisher's Exact Test, $p = .650$, 6-to-7-year-olds, Fisher's Exact Test, $p = .266$). For the vertical-tube task, the 6- to- 7- year- olds were more successful following the endstate demonstration than were the younger 4- to- 5- year- olds, χ^2 (1, N = 36) = 5.783, $p = .016$. No difference between the age groups was seen following the endstate demonstration for the horizontal-tube task, χ^2 (1, N = 3) = 1.463, $p = .226$.

Condition also made no difference to success levels following the second action demonstration for the 4-to-5-year-olds for either task (both tasks = Fisher's Exact Test, $p > .999$). No comparisons could be made for the 6-to-7-year-olds as all children in both conditions were successful following the action demonstration. There was also no difference in success levels between the two age groups on both tasks, (Vertical-tube task, Fisher's Exact Test, $p > .999$; Horizontal-tube task, Fisher's Exact Test, $p = .102$).

5.1.3 Discussion

In experiment 5 children between the ages of 4 and 7 were given a delay between being introduced to the tool-innovation problem and being allowed to interact with the apparatus. During the delay children were given the opportunity to explore the materials they would be using to solve the problem. The delay and opportunity for exploration did not help children on the subsequent tool-innovation task. We concluded that neither impulsive behaviour nor failure to explore materials was sufficient to explain children's difficulty with tool innovation.

The objective of this study was to investigate whether children's tool-innovation difficulty could be explained by either impulsive behaviour or the lack of exploration of the materials. Observations made in previous studies suggested that when confronted with a tool-innovation task young children spend very little time considering their options or planning a strategy before embarking on a solution. Children had a tendency to pick up the materials and insert them into the apparatus immediately after they were presented, sometimes even before the experimenter had fully finished giving the instructions. It was suggested that this impulsive behaviour may have been the cause of children's tool-innovation difficulty. Perhaps children were unable to inhibit their impulsive response of going directly for the reward and as such were prevented from thinking about the problem more rationally.

Research into children's inhibition has shown that children significantly improve on tasks requiring inhibition if there is a delay between the question or setting of the problem and their response. Diamond et al. (2002) suggested that the delay improves performance as it gives children time to compute their answer. Beck, Carroll, Brunsden and Gryg (2011) have argued that this explanation does not adequately explain the effects of a delay in their counterfactual tasks as children were capable of computing the correct answers for future hypotheticals even if there was no delay. Instead Beck et al. suggest that children have a tendency to act impulsively based on their current prepotent response. The addition of a delay allowed children to avoid errors caused by their impulsive behaviours rather than aid inhibition of a particular response. Children's difficulty in tool innovation might have been due to impulsive behaviours based on a default theory regarding the need for contact, if so a delay may have helped children avoid errors based on this impulsiveness. However, we found no improved performance in the current study when this manipulation was adopted. This suggests that behaving impulsively is not the limiting step in tool innovation.

The second identified possible constraint on children's tool-innovation performance was the failure to explore the task materials. In previous studies children engaged in a regimented warm-up exercise in which they manipulated the materials in certain ways as defined by the experimenter. Research into children's exploratory behaviour has shown that children are better able to find the properties of materials if they are allowed to explore for themselves rather than being shown one way of manipulating them (Bonawitz et al., 2009). Based on this finding, rather than showing children things that the materials could do, we let them explore the materials for themselves. If children's poor tool-innovation ability was due to them not exploring the materials they were given then we might expect improved performance in the group of children who got to explore the materials. This was not the case. Children who independently explored the materials were no more likely to innovate a tool than children who did not explore the materials. This suggests that a lack of exploration is not sufficient for explaining children's tool-innovation difficulties.

In the current study we gave half of children both a delay and the opportunity to explore the materials with the view that if children in this group were more successful we would investigate further to determine what was driving this effect. As there was no improvement in this group, we conclude that neither impulsivity nor lack of exploration was sufficient to explain tool-innovation difficulties, and as such we leave our investigations on these lines of research here.

5.2 Experiment 6: Executive Function Battery

In experiment 6 we investigated the role of executive function in tool innovation more fully by testing children on a series of executive function tasks, and correlating them with tool-innovation performance. In the current and previous chapters we have suggested various components of executive function as possible explanations for children's poor tool-innovation abilities. In chapter 2 we discussed the possibility of switching difficulties, and in experiment 5 we investigated impulsivity, an aspect of inhibitory control. In the current experiment we expand on these ideas and test children's executive abilities more fully through an extensive executive control battery. Children's performance on executive tasks tapping inhibition, working memory and task switching, as well as an ill-structured executive task that tests these constructs in combination were related to performance on the hook-innovation task.

5.2.1.1 Inhibitory Control

In previous experiments various aspects of inhibitory control were explored. In experiment 2 perseveration was used as an indirect measure of children's ability to inhibit their incorrect actions. Perseveration of an incorrect strategy for the whole task period was taken as an indicator of children's lack of inhibitory control. Very low levels of perseveration were observed in experiment 2 and consequently inability to overcome perseveration was ruled out as a reason for children's poor tool-innovation performance.

Experiment 5 directly investigated the role of impulsivity, the inability to inhibit the tendency to act without waiting, on children's poor innovation ability. This study concluded that children's poor performance was not due to them acting impulsively as preventing children from acting immediately on the apparatus did not improve task performance. Whilst both of these studies gain us some insight in to the role of inhibitory control in tool innovation they do not provide a general measure of children's inhibitory abilities. In the current study tasks specifically thought to tap inhibition were used to provide a baseline measure of children's general ability. Measures of simple and complex inhibition were gained and then compared to tool-innovation performance. Based on definitions put forward by Surtees, Burns, Beck, Riggs and Apperly (under review) we use simple inhibition to refer to measures obtained when tasks are thought to tap inhibitory control without the involvement of other executive functions. Complex inhibition is used to refer to measures that tap inhibitory control but also require additional components of executive control such as working memory.

5.2.1.2 Task Switching

In chapter 2 (Cutting et al., 2011) children's ability to switch between different strategies on two different tool-innovation tasks was examined. Performance on the hook-bending task and the unbending task were compared to see how well children could switch between the different solutions each of the tasks required. It was concluded that children's poor performance was not due to the need to switch away from previously correct strategies. On their second task children did not perseverate on techniques and strategies that had been successful in their first task. This finding provides evidence against task switching difficulty between tasks; however it does not address how well children can switch between alternative strategies *within* a tool-innovation task. In the studies in this thesis children have been observed to switch between different materials in varying amounts. A measure of children's ability to switch between different strategies would enable us to discover whether readiness for switching ability might aid children's performance. Although we have observed children's ability to switch between different strategies in tool-innovation tasks, as evidenced by a lack of perseveration, this switching may be cognitively demanding and as such limit children's success. Obtaining measures of children's general task-switching abilities in the current executive function tests will allow for a fuller understanding of how task-switching capacity may affect tool-innovation performance. The current study will provide measures of both local and global switching abilities and compare these to success on the tool-innovation task. Local switch costs occur when one must switch between different rules or strategies within a task. Global switch costs refer to the decrease in performance when one knows that there is a possibility that a switch may be needed (Burns, Riggs & Beck, 2011). The tool-innovation task is most likely to engage local switching ability, as children will need to switch between different strategies within a single task. However, if children are aware that there are many different strategies to try then there may also be global switch costs.

5.2.1.3 Working Memory

To date no studies have investigated the role of working memory in children's tool-innovation performance. fMRI studies have however indicated that working memory may be important for tool-using actions (Johnson-Frey, Newman-Norlund & Grafton, 2005). It is reasonable to assume that working memory ability may be important when confronted with an innovation task. Children must be able to remember the goals of the task, they must remember what strategies and techniques they have already tried and they must hold in mind what they plan to try next. By including a measure of working memory in our executive battery we will be able to discover the extent of the relationship between working memory and tool innovation and discover whether poor working memory may be a limiting factor on children's performance.

5.2.1.4 Ill-structured tasks

In chapter 2 it was suggested that tool innovation may be an ill-structured problem. An ill-structured problem is one that is missing information from either the start state, goal state or the transformation required to go between the two (Reitman,1965). The defining feature of an ill-structured problem is the need to generate this additional information for oneself (Goel, 2010). To generate this knowledge and use it to solve the task requires multiple executive functions in combination but does not simply reduce down to the sum of its parts.

Recent experiments with brain damaged patients (Goel, Pullara & Grafman, 2001) and children with autism (White, Burgess & Hill, 2009) have used ill-structured tasks to demonstrate the poor performance of clinical populations compared to normal controls.

Previously patients had performed comparatively to controls on traditional executive tasks measuring inhibitory control, task switching and working memory such as the ones used in the current study (Shallice & Burgess, 1991). This finding was unexpected as many of the patients had difficulty performing seemingly simple tasks in everyday life, for example visiting the shops or cooking a simple meal. Ill-structured tasks have provided a useful measure of executive function that more closely resemble the tasks faced by participants in everyday life. These tasks more readily captured the difficulties seen in clinical groups when compared to controls. Researchers concluded that the difficulty of these ill-structured tasks lies in the fact that they require multiple executive functions as described above and do not simply reduce down to the sum of their parts. Tool innovation fits the definition of an ill-structured problem. The start and goal states are well-defined but there is information missing about the transformation required to go between the two. As such tool innovation, like the studies cited above, is likely to require the use of executive functions, but may do so in conjunction with each other and may not simply be explained by individual executive components.

Based on this suggestion the current experiment tested for a relationship between tool innovation and another ill-structured task, namely the Six Part Test (Emslie, Wilson, Burden, Nimmo-Smith, Wilson, 2003). The Six Part Test requires children to engage with three different subtests each containing two parts. Children must achieve as much as possible whilst following rules laid out for them. The Six Part Test requires children to engage multiple executive functions. Children must to use their working memories to hold in mind the rules and what they have done and need to do next. They must also engage their inhibitory abilities to stop on a task and move on to a different one, and to prevent themselves from simply moving along the line of tasks that would mean a rule break. Finally children must be able to switch between different tasks.

5.2.2 Method

5.2.2.1 Participants

The final sample consisted of 43 participants aged 6 to 8 years (25 boys), mean age 7 years 6 months (7; 6), range 6; 8 to 8;5, from a Primary School in South Birmingham, UK.

5.2.2.2 Design

Children were tested individually in two sessions by the same experimenter. Each session lasted around 15 to 20 minutes. The two sessions were administered to children at least 3 days apart and a maximum of 14 days apart. In the first session, participants first completed the Six Part Test from the Behavioural Assessment for Dysexecutive Syndrome for Children (BADS-C) battery (Emslie et al., 2003) and then the British Picture Vocabulary Scale-II (BPVS-II, Dunn, Dunn, Whetton & Burley, 1997). In the second session children were given the hook-innovation task and the EF tasks in a fixed order: Hooks task (Beck et al., 2011), simple inhibition (‘the pictures task’, Burns et al., 2012, adapted from Davidson, Amso, Anderson & Diamond, 2006), working memory (a counting recall task based on Alloway, Gathercole, Willis & Adams, 2004; Burns et al.), and finally a task of complex inhibition and task switching (‘the eyes task’, Burns et al., an adaptation of the arrows task from Davidson et al.). The tasks were presented in a fixed order to children to ensure that they received as closely as possible the same experience.

5.2.2.3 Materials and Procedure

5.2.2.3.1 Six Part Test

The Six Part Test is a subtest from the BADS-C battery (Emslie et al., 2003). The test is a version of the Six Elements Test for adults (Burgess et al., 1996) that has been modified to be suitable for use with children. The task is laid out as per Figure 5. 1(a), and contains three types of task that each have two versions. The green ‘How Many?’ tasks required children to turn over cards to reveal a number of pictures (see Figure 5. 1 (b)). Children counted the number of pictures and wrote their answers down on the paper provided. The blue 'What is it?' cards required children to turn over cards to reveal pictures (see Figure 5. 1 (c)). Children were then required to write down what the picture was on the paper provided. All words were short in length consisting of 3-5 letters; children were aided with the spelling of the words if required. The red 'Sort me' tasks consisted of two boxes, one containing multiple types of beads, the other containing nuts and bolts. The lids of the boxes contained a picture, and children were required to find the relevant beads or nuts from the boxes that matched the picture and put them in the lids. Children were given instructions as to how to carry out each task and were told they had 5 minutes to complete as much as they could of each of the six tasks and it was emphasised that they would not be able to complete all of the tasks because they did not have enough time. Additionally children were given two simple rules to follow. First, they were told they must complete a little bit of every single task during the 5 minutes. Second, they could not do two types of the same task in a row, e.g. if children were working on the green *'how many? 1'* task they could not move on to work on the green *'how many? 2'* task next, they must switch to work on one of the blue *'what is it?'* or red *'sort me'* tasks. Children were given 5 minutes to engage with the task and a timer was in view so that they could check their progress.

Children received an overall score out of sixteen for the Six Part Test. The score was calculated in the following way: children were awarded 2 marks for each subtask they attempted (maximum of 12 marks). One mark was deducted for any rule breaks on each of the three types of tasks (up to a maximum of 3 marks), additional rule breaks within each type of task were not penalised. Marks were then added or deducted based on the strategies children used. If children used a clear pattern of responses to avoid breaking the order rule they were awarded 2 marks, for example, G1, B1, R1, G2, B2, R2. If children had a clear strategy for trying to attempt all 6 parts they were awarded an additional 2 marks, examples of strategies include undertaking a set number of items on each subtest before switching, or attempting a task for a set amount of time, or a combination of both of these. Children had one mark deducted if they returned to any part 3 or more times.

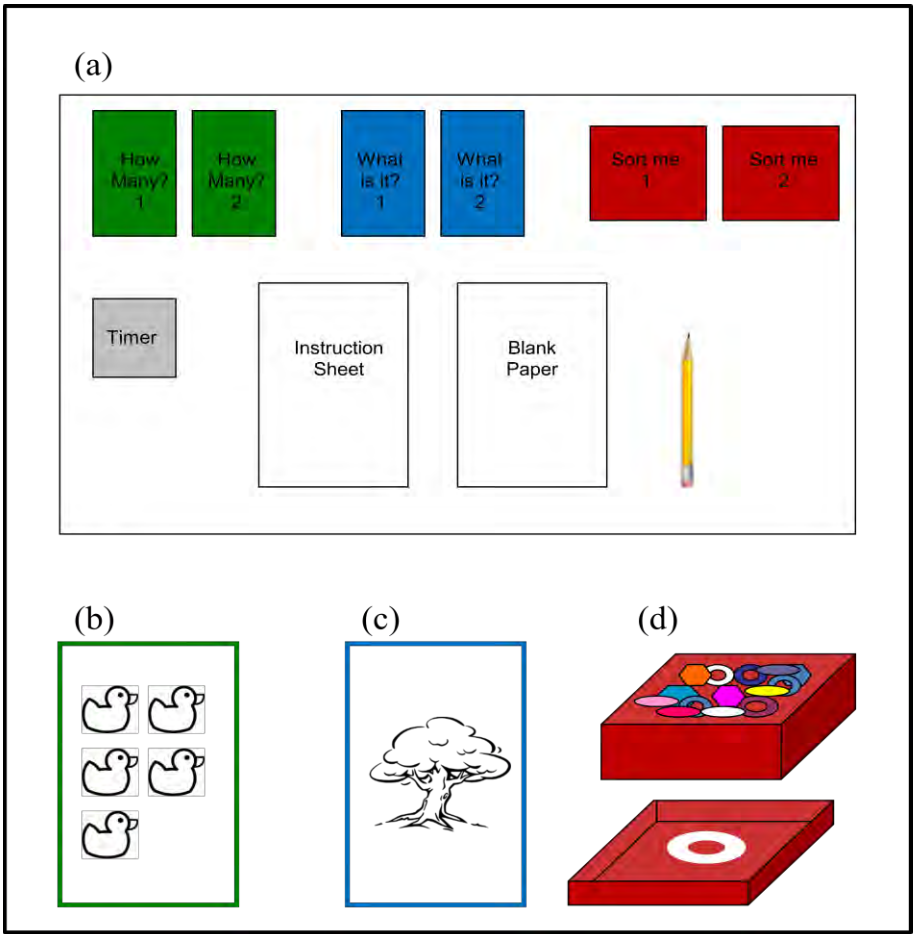

Figure 5. 1 *Materials for the Six Parts Test. (a) Layout of cards and materials for the test. (b) An example of a 'How Many?' Card. (c) An example of a 'What is it?' card. (d) The 'Sort Me' beads task.*

5.2.2.3.2 BPVS II

The British Picture Vocabulary Scale-II (BPVS-II, Dunn et al., 1997) is a measure of children's receptive vocabulary. On each trial children were presented with four outline drawings and were asked to point to the picture that corresponded to a target word spoken by the experimenter. Trials were administered in sets of 12 which increased in difficulty.

Children started with the set which was indicated as being appropriate for their age. The test was terminated if children only succeeded on four or fewer trials within a set. The dependent measure was the total number of correct responses.

5.2.2.3.3 Hooks Task

Children were presented with the hook-innovation task using the same instructions as in experiment 1. There was no familiarisation warm-up but children received the two-stage demonstration procedure as required (see experiment 3 for details of demonstrations).

5.2.2.3.4 Executive Function tasks

The executive function tasks were presented on a 17 inch screen laptop using E-Prime 2.0 (Psychology Software Tools Inc.) For the ‘Eyes’ and ‘Pictures’ tasks children made responses using two custom built button boxes. The top faces of the boxes were 12cm x 14cm and they had a depth of 3.5cm at the back sloping to 2.5cm at the front. A circular plastic button (diameter 2.5cm) was present on the top of each box. On the left-hand box this button was blue and on the right-hand box this button was green. Responses in the counting recall task were made verbally. All tasks had a pseudorandom trial order to ensure that all children had a very similar experience. It also ensured that there were equal numbers of congruent and incongruent trials in the pictures and eyes tasks, and equal numbers of switch and non-switch trials in the eyes task. The pictures and eyes tasks had similar training procedures, where after receiving instructions children received four practice trials with feedback. Children were required to succeed on three out of four practice trials to proceed to the main task. If children did not reach this threshold they received additional sets of four trials until the criterion was reached. The maximum iteration that any child required was two sets.

5.2.2.3.4.1 Pictures task

The pictures task (Burns et al., 2012, adapted from Davidson et al., 2006) gives a measure of inhibitory control and takes the form of a classic spatial Simon task. Children were first presented with two pictures (a monkey and a cat) which were paired with the two response buttons positioned in front of the participant. Children were instructed to press the left-hand (blue) button when they saw the cat and the right-hand (green) button when they saw the monkey. A small picture of each stimulus printed onto card was placed above the relevant response button so as to reduce memory demands. This was done to reduce extraneous variance allowing us to measure the variable of interest – inhibitory control. The task consisted of 20 trials in which the pictures were presented individually in a pseudorandom order on either the left-hand side or right-hand side of the computer screen. Half of the trials required a congruent response, such that the stimulus was presented at the same side as the response button, and half of trials were incongruent, meaning that the picture was presented at the other side to the response button (see Figure 5. 2). The incongruent trials were the main source of interest as these allowed measurements of children's ability to inhibit their prepotent response of pressing the response button on the same side as stimulus presentation. Accuracy and response times were recorded by the E-Prime software. To account for children merely pressing buttons or getting distracted, anticipatory responses, less than 200ms, were removed prior to analyses. Responses greater than 2.5 standard deviations from the mean were also removed, as per Davidson et al. (2006). A trade-off between accuracy and reaction time was calculated to give an overall processing cost for both the congruent and incongruent trial types. This was calculated by dividing each child's mean reaction time (ms) by the proportion of correct responses such that larger scores equalled greater processing costs. The measure of simple inhibitory control used in subsequent correlational analyses was determined by the processing costs for the incongruent trials in comparison to the congruent trials that did not require inhibition. We term this measure simple inhibition due to there being few other executive demands at play.

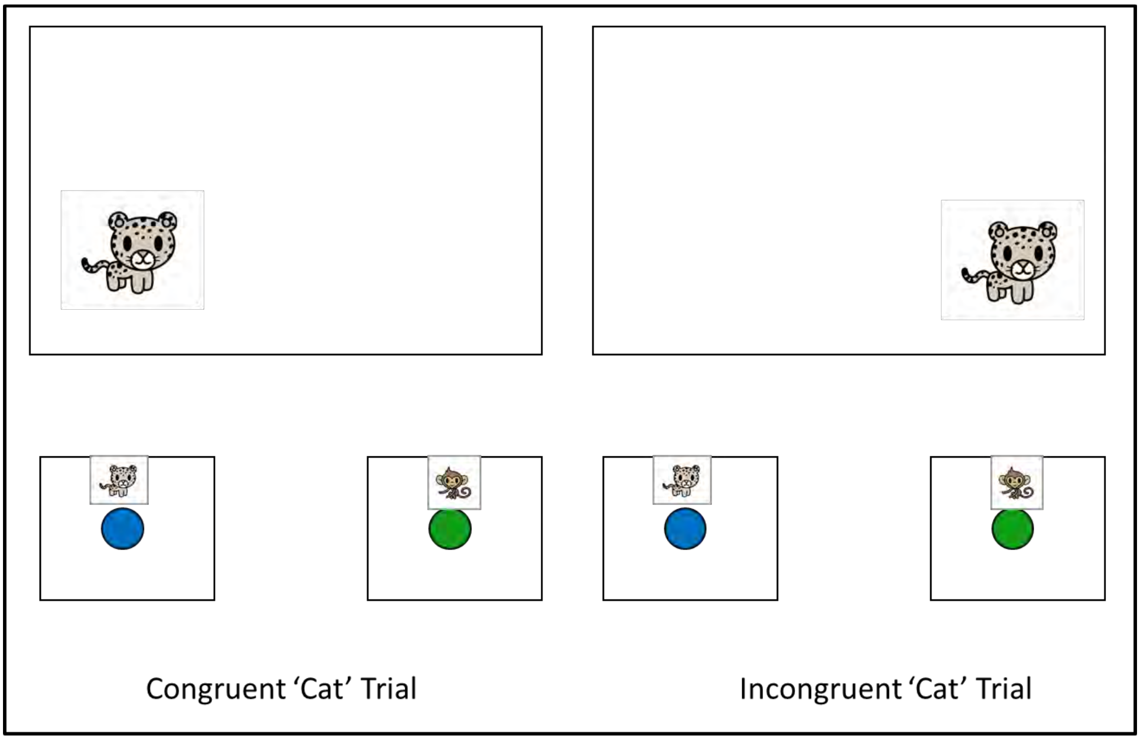

Figure 5. 2 *An example of a congruent trial (left) and incongruent trial (right) for the pictures task.*

5.2.2.3.4.2 Counting Recall

The counting recall task measured children's verbal working memory (Alloway, Gathercole, Gathercole & Pickering, 2000; Burns et al., 2012). On each trial children were presented with an array of red dots and blue squares on the laptop screen (see Figure 5. 3), and were instructed to only count the red dots. The arrays contained between 4 and 7 red dots. The array then disappeared and children were asked to verbally recall how many red dots they had counted. Children began by recalling one screen at a time, and succeeded in a block if they were correct on at least four out of six trials. If they reached this threshold they then proceeded to the next block in which each trial consisted of recalling the numbers of dots in two arrays, then three arrays and so on up to a maximum of five. Each block consisted of 6 trials, and children needed to achieve 4 trials correctly to proceed, if children got the first four trials correct they proceeded automatically and were credited as achieving all 6 correctly. The test was terminated when children were incorrect on three trials within a single block or they had completed all of the available trials. The total number of correct trials was calculated as the dependent measure of working memory. Children received four warm up trials with feedback prior to starting the task. Two of the warm-up trials had one array and two contained two arrays.

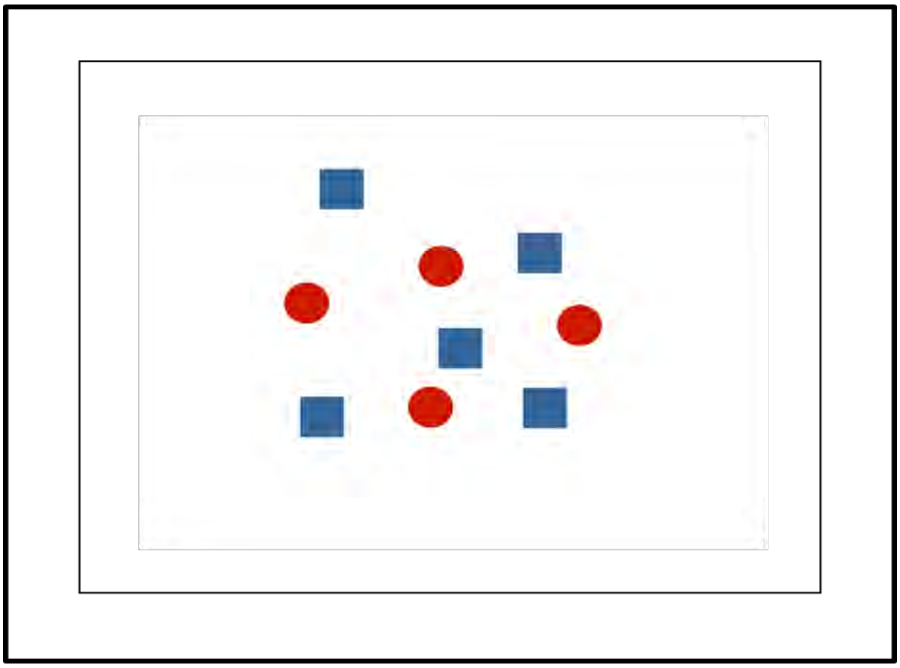

Figure 5. 3 *Example of an array from the Counting recall task.*

5.2.2.3.4.3 Eyes Task

The eyes task (Burns et al., 2012) is an adaptation of Davidson et al.'s (2006) arrows task. The task has both inhibitory and task switching demands. The stimuli in the task were faces presented on the laptop screen. Faces could be presented on either side of the screen and had eyes that looked either straight downwards or downwards and across at a 45° angle (see Figure 5. 4). Children were instructed to press the button the eyes were looking towards. When the eyes looked downwards the response children were required to give was congruent with the side of the screen where the face was presented. When the eyes looked across the screen the response was incongruent with the position the face was presented. To succeed on this task children must learn two rules, when the eyes are looking downwards they must press the button on the same side as the stimulus, and when the eyes are looking across they must press the button on the opposite side. Burns et al. found local switch costs indicating that children treated these as two separate rules and did not combine them in to one simpler rule, i.e. press where the eyes are looking, as one might expect. Children received 3 blocks of 20 trials; the first block contained all eyes looking downwards, the second block was all eyes looking downwards and across, and the third block was a mixture of the two.

Three measures were obtained from the eyes task; complex inhibition, local switch cost and global switch cost. The complex inhibition measure was obtained using the same method as the simple inhibition measure outlined in the pictures task, and provided a measure of the processing cost involved in responding to incongruent trials in comparison with congruent trials in the mixed block. In the eyes task this is termed complex inhibition due to the increased working memory demands, as unlike in the pictures simple inhibition task children were not provided with pictures of the relevant stimuli above the response buttons. Local switch cost refers to children's ability to switch between different rules within the mixed trial block, that is switching between eyes downwards and eyes downward and across trials. The local switch cost measure is calculated by comparing performance on switch versus non-switch trials within the mixed block. As per the inhibition measures a processing cost was produced by dividing accuracy on each of the trial types by the proportion of correct trials. The difference between the processing costs for switch and non-switch trials was used as a measure of local switch cost. Global switch cost refers to the cost to children's performance in the mixed block when they know that they might have to switch between rules relative to performance in the congruent or incongruent blocks where no switching is required. Global switch cost was calculated by comparing processing costs for congruent trials that follow congruent trials in the mixed block to congruent trials following congruent trials in the congruent block, and similar for incongruent trials. An average of these two differences was then used as the global switch cost measure.

.

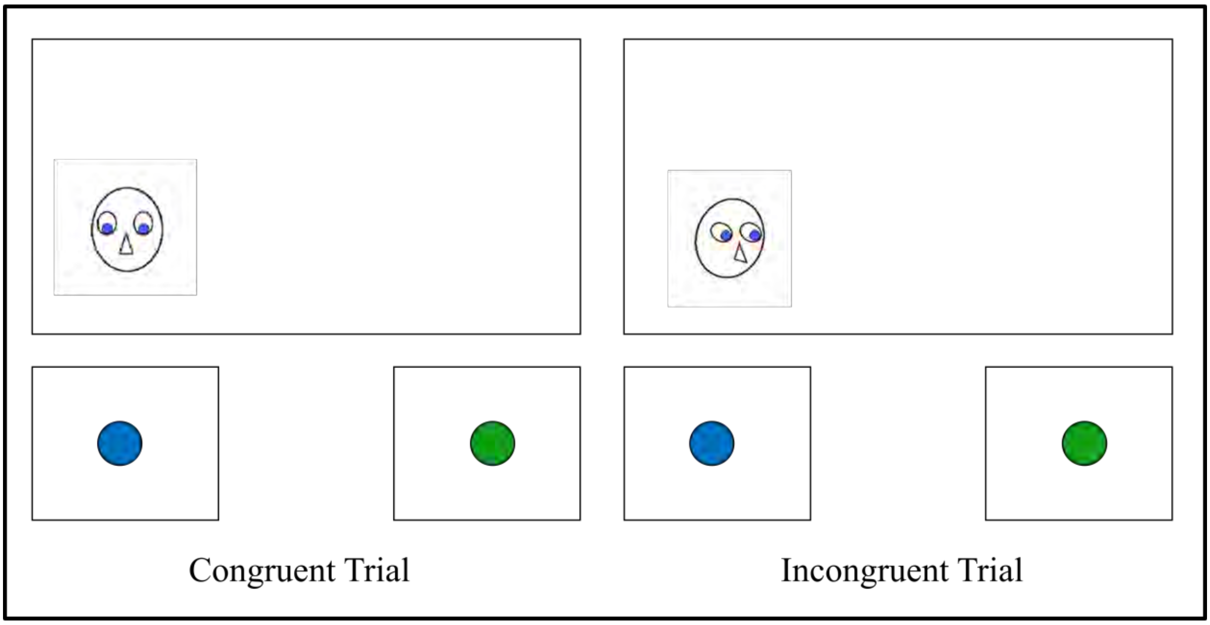

Figure 5. 4 *An example of a congruent trial (left) and incongruent trial (right) for the 'eyes' task.*

5.2.3 Results

Preliminary analyses revealed no gender differences for any of the tasks as such data were collapsed for all subsequent analyses. First we will report descriptive data for all tasks, before examining individual differences in task performance.

5.2.3.1 Six Parts Test

Children achieved a wide range of scores of the Six Parts tests (see table 5.3). Children were able to score a maximum of 16 points on the task, and so the mean score of

7.65 means that children were not performing at floor or ceiling and provided a good spread of results.

Table 5. 3 *Six parts test mean scores.*

	Range	Mean (standard deviation)
Overall Score	2 - 15	7.65 (3.91)
Number subtasks completed	1 - 6	3.95 (1.93)

5.2.3.2 Counting Recall Task

Children provided a wide range of scores on the counting recall task (7 to 21, out of a possible maximum score of 30). The mean score achieved was 14.09 (standard deviation 2.877), showing that there was a good spread of scores and that we were not seeing ceiling or floor performance.

5.2.3.3 Hook Innovation Task

Just under half of children (47%) were successful in innovating a hook tool to solve the task prior to receiving any demonstration from the experimenter (see table 5.4). Unsuccessful children then received the endstate demonstration; following this first demonstration the majority of the children (76%) successfully manufactured the required tool. All four of the remaining unsuccessful children successfully manufactured the required hook tool following the second action demonstration.

Table 5. 4 *Hook innovation task success levels*

	Success	
	No	Yes
Pre-demonstration	23	20
After endstate demonstration	4	19
After action demonstration	-	4

5.2.3.4 Executive Measures

Table 5.5 presents the mean score for the counting recall task, plus the mean accuracy and reaction times for the different trial types on the pictures and eyes tasks. I will first compare the accuracies and reaction time measures within different tasks before going on to use the calculated processing costs in the correlation analyses.

For the pictures task children were significantly more accurate, t (42) = 7.445, $p<$ .001, Cohen's d = 1.14, and faster at responding, t (42) = -6.205, p < .001, Cohen's d = 0.95, on the congruent trials than on the incongruent trials. For the mixed block in the eyes task children were more accurate for the congruent trials than they were on the incongruent trials, t (42) = 5.290, p <.001, Cohen's d = 0.81, but no significant difference was seen in the reaction times, t (42) = -1.722, p = .092. In the eyes mixed block children were more accurate, t (42) = -4.401, p < .001, Cohen's d = 0.67, and quicker at responding, t (42) = 6.774, p < .001, Cohen's d = 1.03, on the non-switch trials than they were on the switch trials. Children were more accurate, t (42) = 17.824, p < .001, Cohen's d = 2.72, and faster, t (42) = -5.027, p < .001, Cohen's d = 0.77, in the eyes down (congruent) block than they were in the eyes across (incongruent) block. This finding appears to be driven by surprisingly poor performance for the eyes across block during which many children appeared to persist with the rule they had been using in the eyes downwards block.

Table 5. 5 *Mean accuracies and reaction times for executive measures*.

	Mean (Standard Deviation)	
	Proportion Correct	Response Time (ms)
Pictures		
Congruent	.90 (.11)	689.08 (139.61)
Incongruent	.67 (.19)	810.358 (140.88)
Counting Recall	14.09 (2.88)	-
Eyes		
Mixed Block		
Congruent	.85 (.15)	900.89 (223.66)
Incongruent	.63 (.26)	956.52 (304.51)
Congruent following congruent	.87(.19)	818.65 (223.00)
Incongruent following incongruent	.72 (.30)	885.84 (299.70)
Switch	.69 (.17)	1016.47 (273.92)
Non-Switch	.78 (.21)	844.55 (215.26)
Eyes Downwards Block		
(matched trials)	.95 (.11)	620.34 (192.76)
Eyes Across Block (matched trials)	.40 (.12)	768.09 (272.23)

Table 5. 6. *Correlations between all measures (Partial correlations controlling for age and BPVS).*

	Six Parts	Simple Inhibition	Counting Recall	Complex inhibition	Local Switch Cost	Global Switch Costs	Hooks
Six Parts	-	-.058 (-.098)	-.014 (-.032)	.273 (.328)	.096 (.084)	.151 (.177)	-.073 (-.161)
Simple Inhibition		-	-.343* (-.371)*	.011 (.039)	.077 (.079)	.191 (.207)	.039 (-.014)
Counting Recall			-	-.233 (-.185)	.136 (.042)	-.190 (-.130)	.232 (.152)
Complex inhibition				-	-.064 (-.010)	.669** (.658)**	-.153 (-.064)
Local Switch Cost					-	-.121 (-.066)	.116 (.030)
Global Switch Cost						-	-.022 (.057)
Hooks							-

*p< .05 **p < .01

5.2.3.5 Individual Differences

Correlations between all measures were conducted, this included: Six Parts Score, simple inhibition (Pictures task), complex inhibition (Eyes task), working memory (Counting Recall), local and global switch costs (Eyes task), and the hook-innovation task (see table 5.6). The scores for the simple inhibition, complex inhibition, and local and global switch costs were calculated as processing costs meaning that larger scores signified poorer performance. For the Six Parts test and the working memory test larger scores indicated better performance. For the hook-innovation task there were two possible outcomes, pass or fail. Significant correlations were found between simple inhibition and working memory (p =.024), and complex inhibition and global switch costs (p < .001). These correlations were still significant after controlling for age and BPVS-II performance (simple inhibition and working memory, p = .017; complex inhibition and global switch cost, p < .001). No other correlations between the executive measures were found. Of most interest to the current study, none of the measures were found to significantly correlate with the tool-innovation measure.

5.2.4 Discussion

The main aim of the current study was to explore the relationships between executive functions and tool-innovation performance in children. Correlational analyses found no relationships between children's ability to innovate a hook tool and their performance on any of the executive measures. We suggest that the difficulties children display in the domain of tool innovation are not due to limited capacity for a singular executive function. Based on the findings of this chapter we will later consider whether we are justified in retaining the conclusion that hook innovation might be an ill-structured problem.

In the current study children undertook a series of tasks measuring their executive abilities; inhibitory control, working memory, task switching, and performance on an ill-structured task tapping all of these demands. None of these executive measures correlated with children's performance on the hook-innovation task. Based on this finding we concluded that tool-innovation difficulty is not driven by a deficit in executive function when measured individually. However, we still think it likely that executive functions play a significant role in tool-innovation tasks as it is difficult to envisage a task analysis that does not involve executive function. Children must have to use their working memory during the tasks, if not they would be unable to keep in mind their goal and the strategies they have used and wish to employ. Children must have to use some form of inhibitory control, they have to inhibit irrelevant actions and stop what they are doing if they are using an incorrect strategy or technique. Similarly, if children are using the wrong strategy they must be able to switch to an alternative one. Executive functions surely play a significant role in enabling children to successfully innovate tools; however, results from the executive function battery indicate that none of the factors have proven to be a limiting step on children's performance.

We suggest the lack of correlation between tool innovation and the executive measures to be due to the ill-structured nature of tool-innovation tasks. Ill-structured problems have been defined as tasks that have information missing from either their start state, goal state or the transformation required to go between the two (Reitman, 1965) . Using this definition tool innovation is an ill-structured problem that is missing information about the transformation required to achieve the goal. To solve ill-structured problems solvers need to engage multiple executive functions all in conjunction with each other and this does not simply reduce down to the sum of its parts. Tasks may involve inhibition, working memory and switching, but measures of these constructs individually do not capture the process by which they are engaged in ill-structured tasks where they are used in combination. Therefore, although none of the executive measures used in the current study correlated with tool-innovation performance, this does not necessarily mean that they are not engaged during tool-innovation tasks. Instead it means that we do not have the method to measure them accurately when they are used in combination.

One might question why, if tool innovation is an ill-structured task, it does not correlate with the ill-structured Six Part Test used in the current study. As stated above the general definition of an ill-structured problem is one that is missing information about the start state, goal state or transformations (Reitman, 1965). This wide definition of what constitutes an ill-structured problem means that ill-structuredness comes in many different shapes and forms. Whilst many ill-structured tasks may require some of the same executive functions to be deployed, these are likely to vary in amount and complexity. As such two tasks that require the same components of executive control may do so in varying amounts and therefore differ greatly in difficulty.

In the current study we have compared two very different types of ill-structured problem. The hook-innovation task has a clearly defined start state (the apparatus and materials) and a clearly defined goal (retrieve the bucket to win the sticker). The ill-structuredness of this task comes from the missing information regarding the transformation required to achieve the goal. The Six Part Test is much more open-ended. This task sets out the start state (3 different tasks, rules etc.) and sets constraints about legal transformations. The ill-structuredness of this task comes from the abstract goal that is set. Children are told that the goal of the task is to complete as much of the six tasks as they can within 5 minutes without breaking any of the rules. Although the goal is outlined for children it is not as concrete as the goal in the hook-innovation task. In the hook-innovation task it is very clear for children that they have accomplished the goal as they will have retrieved their prize. In the Six Part Test children have nothing to gauge their success against, and may be unaware of any failings. This can be seen more clearly if we compare errors between these two ill-structured tasks. Children may be unaware that by breaking the order rule in the Six Part Test they have not achieved their goal; in contrast it is very obvious if children have not achieved their goal when innovating a tool as they will not have retrieved the target. Due to these differences children did not perform comparatively on the two tasks despite them both being ill-structured problems. Although both may engage children's executive abilities of inhibitory control, working memory and switching, they do so in varying amounts and in very different manners.

5.3 Overall Chapter Summary

This chapter investigated the role of executive function in young children's tool-innovation difficulties. Experiment 5 concluded that difficulty could not be explained by children's impulsive behaviour. In addition this experiment also ruled out lack of self-directed exploration of the materials as a reason for children's difficulty. In experiment 6 executive function was investigated more fully with the use of an executive function battery. None of the executive measures were found to correlate with performance on the tool-innovation task. Despite a lack of correlation between hook innovation and the six part test we retain the suggestion that children's difficulty is due to the ill-structured nature of tool innovation. Tool innovation requires multiple executive functions all in conjunction with each other but does not simply reduce down to the sum of its parts. As such, measures of individual executive components did not provide an indication of how well children might perform on these types of task.

Chapter 6

Why Can't Children Piece Their Knowledge Together? The Puzzling Difficulty of Tool Innovation.

This chapter in its current form is currently under submission as:

Cutting, N., Beck, S. R., & Apperly, I. A. (under submission). Why Can't Children Piece Their Knowledge Together? The Puzzling Difficulty of Tool Innovation.

6.0 Abstract

Tool innovation – designing and making novel tools to solve tasks – is extremely difficult for young children. In the current studies we demonstrated different aspects of tool-making to children aged 4 to 6, to discover why this might be. In experiment 1 (N=59), older children successfully innovated the means to make a hook after seeing the pre-made target tool only if they had had chance to manipulate the materials in a warm-up. Older children who had not manipulated the materials, and all younger children performed at floor. In experiment 2 (N=50), younger children's poor performance was not explained by a failure to remember the manipulation warm-up. We conclude that children's difficulty is likely to be due to the ill-structured nature of tool-innovation problems, in which components of a solution must be both generated and coordinated. Older children struggled with generating components of the solution but could co-ordinate them, whereas younger children could not co-ordinate components, even when explicitly provided.

6.1 Introduction

Tools are an essential part of everyday life; it is hard to consider how we might get through the day without them. It has long been argued that tools played a crucial role in human evolution (Gibson & Ingold, 1993), and as such we are argued to be the ultimate possessors of tool cognition. It is therefore unsurprising that young children are prolific tool users. Their tool-using capacity is evident from a young age, with children as young as 2-years using simple tools such as spoons (Connolly & Dalgleish, 1989) and rakes (Brown, 1990). Despite being successful tool users and tool manufacturers following full instruction, children perform poorly on tasks requiring tool innovation, by which we mean creating a novel tool to solve a problem (Beck, Apperly, Chappell, Guthrie & Cutting, 2011). In the current experiments we sought to discover why tool innovation might be so difficult for children. Our strategy was to demonstrate different components of the task solution to see if this improved children's performance.

Children's tool-innovation difficulties have previously been demonstrated in a series of experiments requiring children to innovate the solution to a task in order to retrieve stickers (Beck et al., 2011; Cutting, Apperly & Beck, 2011; Cutting, Beck & Apperly, under submission). In these studies, children had great difficulty in generating the solution to bend a pipecleaner into a simple hook tool to retrieve a bucket from a narrow vertical tube. Children under the age of 5 rarely innovated the solution of a hook tool, and by the age of 8 only around half of children were successful on this task. This difficulty in tool innovation extends to making other tools using pipecleaners (Cutting et al., 2011), and to other materials and methods of tool making (Cutting et al., under submission).

Previous research investigating children's abilities has used the broad description that tool innovation is the making of novel tools to solve problems (Beck et al., 2011; Cutting et al., 2011). However, within this broad definition there are a number of potentially distinct components. In the current work, we distinguished between two types of tool innovation: *innovation of the solution* and *innovation of the means* required to succeed on tool-making tasks. Children were credited with *innovating the solution* to a task if they succeeded in making the required tool without having seen an example of that tool. This is what has been described as 'tool innovation' in previous research. In the current studies, children were alternatively credited with *innovating the means* to solve the task if they had seen an example of the endstate tool they were required to create but had to generate how to make the tool for themselves. After initial failure at innovating the task solution Cutting et al. (under submission) showed 4- to 7- year olds an example of the endstate tool needed (a readymade pipecleaner hook). Success levels following this demonstration were low suggesting children may also have difficulty with innovating the means to solve tool-making tasks. In this previous research success following an endstate-tool demonstration was more broadly termed ‘tool manufacture’, but this current work makes a clearer distinction. Following previous research (Cutting et al., 2011) *tool manufacture* is now more strictly defined as the making of a tool having seen a full demonstration (i.e. having seen both the means and the solution).

What is most surprising about the findings that children find both types of innovation so difficult is that children appear to possess all the relevant knowledge required to solve these tasks. Children are familiar with the properties of the materials, for example the pliant nature of pipecleaners. In previous studies children received manipulation exercises in which they bent pipecleaners prior to being given the tool-making task (Beck et al., 2011, experiment 3; Cutting et al., 2011, experiment 1). Manipulating the pipecleaner did not aid children on the subsequent tool-making tasks. This suggests that lacking knowledge about the properties of pipecleaners (or other materials) is not sufficient to explain children’s difficulty.

Another possibility is that a sense of a lack of permission is a potential constraint on children’s tool innovation. To ensure children knew they had permission to alter the pipecleaners in the context of the main task, Cutting et al. (2011, experiment 2) told children to *make something* with the materials they were given. Again, this did not help children to innovate the solution to the task.

As well as seemingly understanding the properties of pipecleaners and the fact that they are allowed to manipulate them, children also appeared to have the required knowledge about the physics of the problem they faced. In the hook-making task, children appeared to understand that a hook would be the most functional tool for the job: in a tool-selection task children as young as four years old chose the hooked tool over the straight tool first, when their task was to retrieve a bucket from a vertical tube using pre-made tools (Beck et al., 2011, experiment 1). Furthermore, children could also recognize a functional tool when shown how to make one: After initial failure on the hook-innovation task children readily manufactured a hook tool and used it correctly when shown a hook-making demonstration (Beck et al., 2011; Cutting et al., 2011). Note that children were only shown how to make the required tool; they were not given a demonstration as to how to use it.

Taken together this evidence suggests that children have the knowledge required to solve these tool-innovation problems. Children understand the properties of the materials they are given and are aware that they are allowed to manipulate them. Children understand the physics of the task and can recognize a hook as the most functional tool. So if children possess all of this knowledge why do they find tool innovation so difficult?

One possibility is that children's difficulty with tool innovation could be due to its ill-structured nature. Although there is no single agreed upon definition of what constitutes an ill-structured problem, a generally agreed upon framework is that an ill-structured problem is one that is missing information from its start state, goal state, or information regarding the transformation required to go between the two (Goel & Grafman, 2000). Following this definition tool innovation is an ill-structured problem: children are given the start state (the apparatus and the materials) and told that the goal is to retrieve the sticker, yet they are given no information regarding how they should go about this task. Compare this to Beck and colleagues' (2011, experiment 1) well-structured tool-selection task in which young children readily succeed. In this task children are given the start state (the apparatus and materials), the goal state (retrieve the sticker) and are given the choice between two possible means for effecting a transformation (use the straight pipecleaner or use the hooked pipecleaner). When information about the start state, goal and means were provided children found it trivially easy to retrieve the bucket.

The literature on ill-structured problems suggests that just having all of the individual items of domain knowledge is not sufficient to be successful in solving these types of problems (Chen & Bradshaw, 2007). Domain knowledge must be well-integrated into what is termed structural knowledge to enable people to utilize it effectively (Jonassen, Beissner & Yacci, 1993). Structural knowledge is the knowledge of different concepts and how they relate to each other (Jonassen et al.). When people have well-integrated structural knowledge they are able to represent problems more flexibly. This flexibility enables people to generate the required pieces of knowledge and successfully coordinate them. Some novices may possess all the relevant pieces of information, but only in experts is this knowledge integrated into structural knowledge that is flexible enough to solve the problem (Voss et al., 1986; Wineburg, 1998).

Applying this framework to tool innovation it is possible that although children undertaking these problems appear to possess all the knowledge required to solve the tasks, if this knowledge is not well integrated then they may still struggle to produce a solution. Children's difficulty in these tool-innovation studies may lie with generating all the required pieces of information from memory, coordinating all of these different pieces of knowledge, or a combination of both.

From previous studies we know that highlighting the properties of the materials was not sufficient to elicit tool innovation at any age tested. Children were not aided in innovating the solution to the task when they were given a warm-up manipulation exercise which highlighted information about the properties of the materials (Beck et al., 2011, experiment 3; Cutting et al., 2011, experiment 1).We also know that just seeing the endstate target tool that they were required to make, without any information regarding manipulation, was not sufficient to prompt children to innovate the means to make the tool for themselves (Cutting et al., under submission). This is particularly surprising given that children are able to see the utility of the end-state tool and select it to use themselves in the context of a tool-selection task (Beck et al., 2011, experiment 1).

In the current experiments we investigated whether children were able to coordinate information and successfully innovate the means to make a tool if we highlighted both the properties of the materials and the endstate tool required. By highlighting property information to half of the children before they attempted the task and then providing all children with an endstate demonstration after initial failure, we can begin to disentangle the minimum amount of information children required in order to successfully innovate a tool. Given previous findings we expected children who had information about material properties highlighted through the manipulation exercise to be no more successful in innovating the solution of a hook tool than children who did not receive the manipulation exercise. Second, if children failed to innovate the solution in this first stage we then compared the two groups on their ability to innovate the *means* to make the required tool following the endstate demonstration. Based on findings from Cutting et al. (under submission) we expected children in the no manipulation exercise condition to perform poorly following the endstate tool demonstration. This would demonstrate children's difficulty with generating additional information. For children in the manipulation exercise condition examination of performance following the endstate tool demonstration allowed us to discover whether children could successfully coordinate information when it is highlighted. If children's difficulty lies only with generating information we should expect to see successful coordination of information in children who experience having both information about properties and information about hooks highlighted for them (manipulation group). However, if children's difficulty lies with coordinating information we should find that even if children have both pieces of information highlighted they will still be unable to solve these tool-innovation tasks.

We tested children in reception (aged 4 to 5 years) and year one (aged 5 to 6 years) classes as these children perform near floor on previous tool-innovation tasks, and thus there was room for significant improvement.

6.2 Experiment 7

6.2.1 Method

6.2.1.1Participants

The participants were 29 children aged 4 to 5 years (13 boys), mean = 4 years 1 month (4;1), (range 4;1 – 5;1) and 30 children aged 5 to 6 years (17 boys), mean = 5;7 (range = 5;2 – 6;1) from a Primary school in Birmingham, UK. The ethnic composition of the sample was 96% Caucasian, 3% Black and 1 % Asian.

6.2.1.2 Materials

For the manipulation exercise, we used a pipecleaner (length = 29cm), a pen, a piece of string (length = 29cm), and a template of an S-shape printed onto card. For the main task the apparatus was a clear plastic tube (length = 22cm, width of opening = 4cm) attached vertically to a cardboard base (length = 35cm, width = 21cm), a bucket containing a sticker, a pipecleaner (length = 29cm) and a piece of string (length = 29cm) (see figure 1). The experimenter used an identical pipecleaner (length = 29cm) for the demonstrations.

6.2.1.3 Procedure

Before testing children were instructed by their class teacher not to tell other children how to play the games they would be playing with the experimenter to ensure they would be a nice surprise for everyone. All participants were tested by a female experimenter in a quiet area just outside the main classroom. The child and experimenter sat at right angles to each other at the corner of a table. Children were systematically allocated to either the manipulation or no manipulation group based on the teacher’s class list.

6.2.1.3.1 Manipulation exercise

Children in the manipulation group received the exercise prior to being given the main task. We based the exercise on the procedure from Cutting et al. (2011). Children watched as the experimenter demonstrated actions with the string and pipecleaner (order counterbalanced), which the child then copied. The pipecleaner was wound around a pen, and then removed to demonstrate that it kept its shape. The string was laid over the template to follow the S-shaped pattern.

6.2.1.3.1 Main Task

Children were shown the vertical transparent tube with the bucket containing a sticker already in place in the bottom. They were told that if they could get the sticker out of the tube they were allowed to keep it. The experimenter then brought out the string and pipecleaner and told the child that these were things that *'can help'* to get the sticker out. The children were then given one minute to try to retrieve the sticker. No feedback was given, but children were given neutral prompts if required. Examples of prompts used include *'Can you think how you might be able to get the sticker out?'* and *'Maybe you could use these things to help you.'* If, after one minute, the child had not retrieved the sticker, they were encouraged by the experimenter to put down the materials they were using. With the materials remaining on view in front of the participant, the experimenter then said '*look at this*,' and brought out a readymade pipecleaner hook for the child to view (endstate-tool demonstration). The children were again encouraged to retrieve the sticker using their own materials. If after 30 seconds the child had not retrieved the sticker, they were told to put down their materials. With their materials remaining in view as before, the experimenter said '*watch this'* and taking her own straight pipecleaner, held in the middle, bent one end to form a hook (action demonstration). Children were again encouraged to use their own materials to retrieve the sticker. If children were still not successful in making the required hook tool they were given verbal prompts such as '*Did you see what I did with mine?'* and then *'Can you do that?'*

6.2.2 Results and Discussion

There were no effects of gender on level of success pre-demonstration (Fisher's Exact Test, p > .999), or for success following the first endstate-tool, (χ^2 (1, N = 54) = 0.081, p = .776), or second action demonstrations (χ^2 (1, N = 32) = 0.125, p = .723). As such, data were combined across gender for subsequent analyses.

Children's success at innovating the solution during their first exposure to the apparatus was examined to see whether the manipulation exercise facilitated their performance. Overall children were very poor during their first exposure to the task, with only 5 out of 54 children successfully making a hook tool; all 5 children were in the older age group. Three of these children were in the manipulation group and 2 were in the no manipulation group, demonstrating no effect of condition, Fisher's Exact Test, p > .999. This is in line with previous findings (Beck et al., 2011, experiment 3; Cutting et al., 2011, experiment 1) showing that manipulation of materials prior to the main task does not aid children in innovating the solution.

Children who were successful on their first exposure to the task were excluded from subsequent analyses that compared success following the demonstrations.

To investigate whether children were more successful following the endstate demonstration if they had received the manipulation exercise Chi-square analyses were used to compare children's performance between the two conditions (manipulation vs. no manipulation). For the older age group, children were significantly more likely to make a hook tool following the endstate demonstration if they had received the manipulation exercise, Fisher's Exact Test, $p = .015$ (see table 6.1). This suggests that children are able to coordinate the information if both pieces of information have been highlighted for them. No such difference was seen for the younger children, Fisher's Exact Test, $p > .999$. One interpretation of this is that the younger children's difficulty is at least in part with coordinating the information to innovate the means to solve the task. However, the difference between the two age groups may be due to lower working memory capacity of the younger children. Perhaps younger children were unable to remember information that had been highlighted for them prior to them encountering the main task. We tested this possibility in experiment 2.

Table 6. 1 *Frequency of children's tool-making following different levels of instruction in Experiment 1.*

			Success		
Age Group (years)	Condition	N	Before Demonstrations	After endstate-tool demonstration	After action demonstration
4 to 5					
	Manipulation	15	-	5	6
	No manipulation	14	-	4	4
5 to 6					
	Manipulation	15	3	8	2
	No manipulation	15	2	2	5

For children requiring the second action demonstration, only around half were successful at innovating the tool needed (see table 6.1). Chi-square analysis revealed no difference in the levels of success for each group following the action demonstration for either the younger, χ^2 (1, N = 20) = 0.800, $p = .371$, or older children, Fisher's Exact Test, $p > .999$. This level of success was lower than previously observed following similar tool manufacturing demonstrations and we return to this in the General Discussion.

6.3 Experiment 8

In experiment 1 the 5-to-6-year-old children were successful in innovating the means to make the required hook tool, only if they had received both the manipulation exercise in which they manipulated a pipecleaner, *and* were shown an example of the pre-made tool required to solve the task. Individually these pieces of information were not sufficient to elicit either the innovation of the solution (following the manipulation) or the innovation of the means (following the endstate-tool demonstration) required to solve the task. However, even when given both these pieces of information, 4-to-5-year-olds did not solve the task. In experiment 2 we tested whether younger children's difficulty was due to them merely forgetting information that had been highlighted prior to them encountering the main task. We compared children's behavior on the same manipulation version of the task from experiment 1 in which the manipulation information was given to children before they attempted the task and the endstate information given after initial failure (high memory condition), to a low memory version where children were given both the manipulation and endstate information together after children's initial failure on the innovation task to minimize the chance of forgetting.

6.3.1 Method

6.3.1.1 Participants

The participants were 55 children aged 4 to 5 years (27 boys), mean = 4; 7 (range (4; 2 – 5; 2) from an Infant school in Birmingham, UK. The ethnic composition of the sample was 97% Caucasian, 2% Black and 1% Asian.

6.3.1.2 Materials

The materials were the same as in experiment 1, with the omission of the string and the S-shape printed onto card for the manipulation exercise.

6.3.1.3 Procedure

Participants were tested in a similar environment to that outlined in Experiment 1. Children were systematically allocated to either the low memory or the high memory condition alternately based on the teacher's class list. We changed the manipulation exercise to involve only the pipecleaner, so that children were not provided with irrelevant information which may add to their memory load unnecessarily. Children in the high memory condition received the task as in the manipulation condition in experiment 1. Children in the low memory condition were given the apparatus and told to try and retrieve the sticker. If unsuccessful, these children were then given the pipecleaner manipulation exercise immediately followed by the endstate-tool demonstration. As in experiment 1, children in both conditions received the action demonstration if they failed to make the required hook tool in the 30 seconds following the endstate demonstration.

6.3.2 Results & Discussion

There were no effects of gender on level of success pre-demonstration, Fisher's Exact Test, $p = .340$, or for success following the first endstate, χ^2 (1, N = 46) = 0.490, $p = .484$, or second action, demonstration, χ^2 (1, N = 32) = 3.802, $p = .051$. As such, data were combined across gender for subsequent analyses.

As in experiment 1 very few children successfully made a hook tool on their first exposure to the apparatus. Only 4 out of 50 children successfully made the required hook tool, 2 children in each condition. Using Chi-square analyses we next compared children's success at innovating the means to make a hook tool following the endstate demonstration for each of the conditions. We found no difference in success levels between the two groups, χ^2 (1, N = 46) = 0.107, p = .743 (see table 6.2), with only around a quarter of children successfully innovating the means to make a hook tool following the endstate demonstration. This suggests that children's poor performance in experiment 1 cannot be explained by underdeveloped memory.

Table 6. 2 *Frequency of children's tool-making following different levels of instruction in Experiment 8.*

		Success		
Condition	N	Before Demonstrations	After endstate-tool demonstration	After action demonstration
Low Memory	25	2	7	5
High Memory	25	2	6	9

As in experiment 1 success following the second action demonstration was lower than reported in previous studies. Just under half of children who failed to innovate after the first demonstration were successful following this second demonstration, with chi-square analysis revealing no difference in success between the two groups, χ^2 (1, N = 32) = 1.245, p = .265.

6.4 General Discussion

In the current work we distinguished between innovation of the solution and innovation of the means required to solve tool-making tasks. Innovation of the solution refers to when children were able to come up with and make the tool they needed by themselves without seeing an example of that tool. Innovation of the means refers to when children successfully made the tool needed to solve a problem after having seen an example of the tool required but not a demonstration of how to make that tool. The current findings suggest a series of limiting steps in innovation, with children getting stuck at different steps at different ages.

Children at both ages were poor at innovating the solution needed to succeed on the task, even if they had information about pipecleaner properties highlighted for them. This demonstrates difficulty with generating information in both age groups. Older children were successful at coordinating information to innovate the means only if they had information concerning both pipecleaner properties and the need for a hook highlighted for them. However, even if both pieces of information were highlighted younger children still failed to successfully coordinate information and innovate the means needed to make the required tool.

Overall, very few children spontaneously innovated the solution of a hook tool with either no additional information or just information about pipecleaner properties highlighted; all successful children were in the older age group. These results are in line with previous research demonstrating that young children have great difficulty in innovating the solutions to tool-making problems with either no additional information (Beck et al., 2011, experiment 2) or information highlighting pipecleaner properties (Beck et al., experiment 3; Cutting et al., 2011, experiment 1).

Children's ability to innovate the means to make a tool was tested following the endstate demonstration. Children were shown a pre-made pipecleaner hook, but were not shown how to make it. This enabled us to discover whether children could innovate the means to make the hook for themselves. Children in both age groups were extremely poor at innovating the means to make the hook tool following the endstate-tool demonstration if they had not had information regarding pipecleaner properties highlighted for them, i.e. children in the no manipulation group. Older children who had information regarding pipecleaner properties highlighted were significantly more successful in making the required hook tool following the endstate demonstration than children who had not had pipecleaner properties highlighted for them. This suggests that if both pieces of information were readily accessible to older children they were able to successfully coordinate it in to a solution. Conversely, children in the younger age group displayed great difficulty in innovating the means to make the required hook tool even if they had both pieces of information highlighted for them. This suggests that younger children face a limitation in the domain of tool making in that they are unable to coordinate information even when it is highlighted.

These findings from experiment 1 suggest that children's difficulty with tool innovation could be due to problems with generating information. Children were unable to generate additional information when given certain aspects of the task. For example, children who received the manipulation exercise which highlighted the pliable property of pipecleaners were unable to generate information about hooks needed to allow them to innovate the task solution. Similarly, following the endstate demonstration, children in the no manipulation condition, were unable to generate information regarding the properties of pipecleaners and successfully innovate the means of bending the pipecleaner into a hook.

Further evidence suggests progression in the way children between the ages of 4 and 6 can use information to solve problems. Older, 5 to 6 year old, children could successfully coordinate knowledge in to a useful solution if they had both information about pipecleaner properties and information about hooks highlighted for them. However, even when given all the pieces of knowledge required, younger children faced a limitation and could not coordinate this information in to a useful solution. In experiment 1 children in the manipulation condition received information about pipecleaner properties prior to the task, and then information regarding the endstate tool they needed to achieve after initial failure, however these children were unable to coordinate this information and combine it together to make the required tool. Only 5 out of 15 children were successful at making the hook tool following the endstate demonstration. In experiment 2 the possibility that children merely could not remember the information that had been highlighted was ruled out. Children in this experiment were not helped when all of the relevant pieces of information were highlighted just before they encountered the main task. Therefore, young children's difficulty seems unlikely to be explained by underdeveloped working memory.

Findings from both age groups fit with the suggestion that tool innovation is an ill-structured problem that requires solvers to both generate and coordinate knowledge in order to solve a task. As younger children have difficulty with both of these steps it is unsurprising that they perform so poorly on ill-structured tool innovation tasks. Performance appeared to improve with age with 5- and 6- year olds being able to coordinate information in to a useful solution if it had been highlighted, but these older children still displayed great difficulty with generating this information for themselves. Regardless of how good children's ability to coordinate knowledge is they can never succeed in solving the task if they are unable to generate the components of knowledge required. As such, we still see poor innovation ability in this older age group under some conditions (Beck et al., 2011; Cutting et al., 2011).

It is surprising that both age groups had difficulty in generating the required knowledge needed to innovate the solution as previous evidence suggests that children possess all of the individual pieces of knowledge required to solve this tool-innovation task. First, children recognized that a hook was a solution to the task. This is demonstrated in Beck and colleagues tool selection task (2011, experiment 1), in which children from the age of 4 understood the hook to be the most functional tool and chose it over a straight pipecleaner significantly more than chance. Furthermore there is a wealth of evidence showing that children can readily recognize a hook as a functional tool when shown how to make one. Without being shown how to use the tool children readily manufactured a hook tool and used it correctly when shown a hook-making demonstration (Beck et al., 2011; Cutting et al., 2011). Second, pipecleaners are readily accessible to young children in the UK and are used during craft sessions in both nurseries and infant schools. As such children are likely to possess knowledge about their pliant nature. Additional evidence for children's knowledge of pipecleaner properties comes from the manipulation exercises children encountered prior to being given the tool-making tasks (Beck et al., 2011, experiment 3; Cutting et al., 2011, experiment 1). So if children possess all of this information why can they not generate it in the context of a tool-innovation task?

Having domain knowledge may not be sufficient to solve ill-structured problems. To solve these problems knowledge must be well-integrated into what is termed structural knowledge. Only when we have structural knowledge can we represent problems effectively, and flexibly use this knowledge to solve them (Jonassen et al., 1993). The children in the current studies are novices. Although these children may possess all the independent pieces of knowledge the task requires, they do not have sufficient experience with the world and the materials to have integrated structural knowledge. We suggest that young children lack the flexibility needed to generate their knowledge from memory and then coordinate it in order to solve these tool-innovation tasks.

Despite showing children's difficulties in both innovating the solution and the means to solve tasks the current and previous work investigating children's tool making (Beck et al., 2011; Cutting et al., 2011) demonstrates children's aptitude for learning from others. Many children successfully solved the task and manufactured the required hook tool following the second action demonstration. However, these current results show success levels lower than those reported following the same demonstration in previous studies. Previously nearly all children were successful following an action demonstration (Beck et al.; Cutting et al.), but only around half the children in these current experiments were successful. This could of course be due to the sample of children used, but there may be an alternative explanation.

Children in these studies received two demonstrations from the same experimenter. For children requiring the second action demonstration the first endstate-tool demonstration proved to not be useful, as such these children may have subsequently lost faith in the experimenter as a source of useful information and therefore ignored information they were given in the second demonstration. Evidence from studies investigating children's trust in informants suggests that children do not use information provided by a previously inaccurate informant (Birch, Vauthier & Bloom, 2008; Nurmsoo & Robinson, 2009). This finding needs to be explored further to discover whether incremental demonstrations may be detrimental to children's learning about how to make tools. Discovering how children use sources of information is useful not only in this current line of work, but may have far reaching consequences in other domains where social learning occurs.

Altogether, these studies show that young children have difficulty both innovating the solutions and innovating the means to solve tool-making tasks. Findings from both age groups suggest children's main difficulty to be with generating their knowledge from memory. Younger children in these studies also displayed great difficulty with coordinating their knowledge. As children develop and integrate their knowledge they first improve in their capacity to coordinate information, and can do so readily if all the information needed is highlighted for them. We suggest that as children develop further their knowledge will become more integrated. This will allow them to access and generate their knowledge more flexibly, and along with their ability to coordinate knowledge, enable them to solve these ill-structured tool-innovation tasks.

Chapter 7

General Discussion

7.0 Overview

The research in this thesis aimed to discover the extent of young children's difficulties innovating novel tools and investigated reasons why tool innovation might be so difficult. In this concluding chapter I will first briefly summarise the findings from each of the chapters and then discuss them in relation to each other and other literatures. I will also discuss the wider implications and make suggestions for future research.

7.1 Chapter summaries

7.1.1 Chapters 2 to 4: How robust are children's tool-innovation difficulties?

Children's tool making is an exciting new area of research. In these early chapters broad definitions of different types of tool making were used. The term tool innovation was used to describe the making of a novel tool to solve a problem, whereas tool manufacture was used to describe tool making following instruction. The work in this thesis was predominantly interested in children's tool-innovation abilities.

The main focus of chapter 2 was to replicate and extend Beck et al.'s (2011) novel finding that tool innovation is difficult for young children. Children were tested on Beck et al.'s hook-innovation paradigm as well as a new tool-making task that required children to innovate a straight tool from a bent pipecleaner required to push a ball out of a transparent horizontal tube. Children between the ages of 4 and 7 were tested and similar success levels to the previous work were found. Children's difficulties were shown to extend to the new innovation task, although children did find this task comparatively easier.

Chapter 3 extended children's tool-innovation difficulties even further. Children were tested on the same two pieces of apparatus used in chapter 2, but this time the required tool could be made by one of four different tool-making modes as defined by B. Beck (1980). This not only allowed for the testing of children's innovation via different modes of tool making, but also tested children's ability to make tools from different materials. No differences in success levels were seen within each task for the different modes of tool making leading to the conclusion that contrary to suggestions from the animal literature (Kacelnik et al., 2006) there is no complexity hierarchy for different modes of tool making in human children. The main findings from chapters 2 and 3 are that tool-innovation difficulty is a replicable and robust finding that extends to various types of tool making.

Despite there being no differences in success levels between the different modes of tool making within tasks, differences were observed within the different tool-making modes across tasks. Using these additional analyses from chapter 3 as a basis, chapter 4 further investigated the physical aspects of the task which may alter task difficulty. Three levels of tool-making were proposed: mode, transformation and action. As transformation and action covaried in chapter 3, chapter 4 separated these components in order to discover where task complexity might lay. Children were tested on tasks that required the same mode and transformation i.e. reshaping/unbending or subtracting/removal, but different specific actions. Children performed comparably on tasks requiring the same mode/transformation, but differences were seen between the different types (i.e. subtract and reshape). The specific action required made no difference to success. Taking the results from chapters 3 and 4 together, it was suggested that research in chapter 3 was focusing at the wrong level of tool making. Rather than focusing at the broad level of tool-making mode, together these studies suggested that task difficulty was better categorised by the transformation that each tool-making episode requires.

7.1.2 Chapters 5 and 6: Why do children find tool innovation difficult?

Chapter 5 investigated whether executive function may play a role in children's tool-innovation ability. Experiment 5 found that children's task performance was not improved by a delay designed to prevent impulsive behaviour. Experiment 6 found that no singular executive function correlated with tool-innovation performance. Additionally no relationship was found between tool innovation and success on the ill-structured six parts test. Originally designed to test the proposal that tool innovation is an ill-structured problem, a lack of relationship does not necessarily dispute this proposal. It is possible that these two tasks could both be ill-structured yet show no relationship due to differences in their composition. Individual executive components may be required in differing amounts making it difficult to directly compare the two tasks. As such we retained the suggestion that the difficulty of tool innovation is due to it being an ill-structured problem. Ill-structured problems require the solver to generate solutions for themselves. This involves multiple executive functions in conjunction with each other and does not simply reduce down to the sum of its parts.

Building on this, chapter 6 aimed to discover which aspects of tool innovation make the task particularly difficult. Various aspects of the task solution were demonstrated to children in order to discover where difficulty may lay. Five- to six-year-olds were shown to be competent at coordinating knowledge when it was highlighted for them, but had difficulties in generating this knowledge for themselves. In contrast, younger children could not coordinate knowledge even when all aspects of the task solution were highlighted for them.

7.2 Discussion of General Findings

7.2.1 Are some tool-innovation tasks harder than others?

First, different elements of the task were manipulated to ensure that difficulty was not a product of the task design. Issues relating to permission and pragmatics were ruled out in experiments 1 and 2. Second, physical aspects of the tasks were manipulated, such as the apparatus and the type of tool required. Experiments 1 and 2 found low success levels in a new and seemingly simple task requiring children to unbend a pipecleaner to make a long straight tool. However, this task was found to be significantly easier than the hook bending task, a finding that has been replicated in all studies where both of these tasks were used. This difference in success levels led to the question of why some tool-innovation tasks may be easier for children to achieve than others.

Experiment 3 compared different types of tool making that used different materials. Using the vertical-tube apparatus from the original hook-making task and the horizontal-tube apparatus from the unbending task we tested children's ability to innovate tools using B. Beck's (1980) four different modes of tool making: detaching, subtracting, adding and reshaping. With the exception of detaching which we concluded was not a true test of tool making; no differences between the different tool-making modes were found on each of the tasks. Children found all tool-making modes equally difficult and performed poorly. Experiment 4 built on these findings and compared different tasks in more detail. Taken together, the results from chapters 3 and 4 led to the suggestion that rather than categorising difficulty of tool-making tasks based on B. Beck's modes, difficulty is better categorised by the transformation that is required. This means that rather than categorising tasks based on a mode such as reshaping which is broad in its definition, we should categorise tasks on the precise transformation within that mode, i.e. we should group all *unbending* reshaping tasks together and these differ in complexity to the *bending* reshaping tasks.

These findings provide a useful framework in which we can categorise and make predictions about the complexity of different tool-making tasks. This is useful not only in work such as this on human children's tool making but also in the comparative literature. I will return to this point more fully in section 7.4.

Using the finding that transformation is an important determinant of tool-making complexity as a basis, further work now needs to determine what exactly makes one transformation more difficult than another. This thesis presents findings that an unbending transformation is easier to achieve than a bending transformation (experiments 1-3), and that unbending is also easier than removal (experiment 4). In experiment 3 it was speculated that differences in success levels following an endstate-tool demonstration could be due children's ability to recognise the relationship between the endstate-tool they are shown and the raw materials they possess. The transformation required to achieve the endstate-tool from the raw materials may be clearer to see in some cases than others. Following experiment 3 it was suggested that higher success levels following demonstration in the horizontal-tube task for the add compared to the subtract mode of tool manufacture may have been due to the obviousness of the transformation. For the add mode the transformation was to join together three identical pieces of dowel to make a long stick, whereas in the subtract mode the transformation required was to remove a small piece of dowel from a larger more prominent piece. The transformation in the add mode would have been clearer for the child to see and therefore an easier transformation to execute. To fully understand tool-making complexity we need to untangle these differences further and investigate which elements of a transformation make it easier or harder for children to achieve.

7.2.2 What do children need to know in order to make a tool?

Throughout this thesis various studies have directly or indirectly addressed the question of what children need to know in order to be able to make a tool. Experiments 1 and 2 demonstrated that showing children an action demonstration in which they saw the tool they needed to make and how to make it, was sufficient to induce tool manufacture on the hooks and unbending tasks (see also Beck et al., 2011 for similar evidence for the hooks task). Experiment 1 replicated the findings of Beck et al. (Experiment 3) that demonstration of the means alone was not sufficient to elicit tool making in children. In these studies children received a warm-up in which they manipulated the pipecleaner. Highlighting the means by which the tool was required to be made did not aid children on the subsequent task. Chapters 3 to 5 used a two-stage demonstration procedure in which after initial failure children saw a demonstration of the endstate-tool they needed to achieve. For the majority of children seeing an example of the tool they needed to make without having information highlighted as to how to make that tool was not sufficient to enable them to make that tool for themselves. Taken together the research from chapters 2 to 5 suggests that in order to successfully innovate novel tools children must have knowledge regarding the tool they are required to make and knowledge of the means to make that tool. Neither piece of knowledge by itself was sufficient to elicit tool making, but when given together, as in the action demonstration, children were highly successful.

Evidence from experiment 4 suggests there is a third required factor for children's successful tool innovation. In order to successfully complete tool-innovation tasks children must also understand the means to use a tool they have created. Evidence from Beck et al.'s tool-selection task demonstrates that children as young as 4 were able to choose a hooked tool to complete a task and were successful in using it to achieve their goal. Therefore in our tool-innovation tasks requiring the making of a hook, we can safely presume that children had no difficulty in knowing the means to use the hook tool once they had made it. However, evidence from experiment 4 suggests that in some cases children may be able to successfully make a tool but then lack the knowledge to use that tool effectively. A previously unseen behaviour was encountered in experiment 4 where in the indirect tasks the majority of children correctly made the required tool following the end-state tool demonstration, but did not then go on to use the tool correctly to retrieve the target. Children needed to be prompted in how to use the tool to complete the task. This finding shows that as well as understanding what tool is needed to solve a task and the means to make that tool, children must also understand how to use the tool. It could be the case that in the hook tasks the means to use the tool was simple for children because they had good experience and knowledge of using hooks. Alternatively once children had made the tool following the demonstration it could have been that the number of options as to what they could do with their tool in relation to the task apparatus were minimal therefore making the correct means to use the tool very obvious. In contrast the indirect tasks in experiment 4 were more complex in terms of the apparatus design. There were more degrees of freedom in what one could do with a tool they had made as there were two potential entrances in to which the tool could be inserted. This greater complexity could explain why children did not know how to use their tool. A tool selection task using the indirect apparatus could help to confirm this.

In summary we suggest there to be three components of knowledge that children require in order to make a tool. First, they must know what tool is needed to solve the task. Second, they must know the means to make that tool. Finally, they must know how to use the tool once it has been made.

7.2.3 Tool Innovation is an ill-structured problem.

Throughout this thesis I have argued that tool innovation is an ill-structured problem. In this section I will bring together the evidence to support this. There are two strands of research in to ill-structured problem solving, both of which are relevant to this current work. The first strand of research provides a general framework for ill-structured problem solving and suggests the difficulty of these problems to be due to the need to both generate and coordinate knowledge. The second strand of research focuses on the executive control that is required for this generation and coordination of knowledge.

Based on Reitman's (1965) definition tool innovation is an ill-structured problem. Children are given information about the start and goal states but there is information missing about the transformation required to achieve the goal. Children must generate the information regarding the required transformation for themselves and then coordinate it in to a useful solution.

Experiment 7 concluded that the main difficulty faced by 4-to-6-year-olds was the requirement to generate knowledge needed to solve the task. That is, children must generate knowledge about the tool that is needed, the means to make that tool and the means to use the tool. Previous studies have demonstrated children's ability to easily generate knowledge about how to use hooks and so experiment 7 focused only on the generation of information about the tool required and the means to make the tool. Even if knowledge about either the required tool or the means to make the tool was highlighted, all children in this age range had great difficulty in generating the additional information required to solve the task.

Experiment 7 found development between the ages of 4 and 6 in the way children could use information to solve problems. When all of the relevant information was highlighted children aged 5 and above could successfully coordinate the information in to a useful solution. However, even when information regarding pipecleaner properties (means to make tool) and the need for a hook (tool required) were highlighted, the younger children failed to coordinate this information in to the required solution. Similar findings for children not being able to use information that we know they have has been reported by Brown (1990). Drawing on her own work and that of Bates, Carlson-Luden, and Bretherton (1980), Brown described the inconsistencies in infants' abilities based on their understanding of contact. Habituation paradigms show an understanding for the need of contact between 5 and 7 months, yet at 10 months infants were unable to use this knowledge to solve a tool-use problem unless the objects were already touching. Even by 13-to-18-months children could not succeed on such tasks unless they had received a demonstration, that is they could not envision the contact solution for themselves. These studies together with those reported in this thesis support suggestions in the ill-structured problem solving literature that having domain knowledge is not sufficient to be successful on problem solving tasks (Jonassen, Beissner & Yacci, 1993). Knowledge must be well-integrated into structural knowledge that can be flexibly generated and coordinated in order to solve problems. Although the infants in Brown's studies and the children in the current tool-innovation studies appeared to have all the required knowledge to solve the given tasks, these findings suggest that this knowledge was not established enough to enable them to successfully generate it in the context of the tasks. As infants or children gain more experience and knowledge, the information they possess becomes better integrated and children are able to access it more readily.

To generate and coordinate pieces of information requires multiple executive functions such as working memory – one must remember the goal, inhibition – one must inhibit irrelevant information, and task switching – one must switch between alternative strategies. Each of these executive functions is used in conjunction with each other and does not simply reduce down to the sum of its parts. Throughout this thesis various aspects of executive control have been investigated, and together the evidence supports this view.

In experiments 1 and 2 perseverative tendencies and the ability to switch between different task solutions were ruled out as possible limiting factors for children's tool-innovation difficulties. In experiment 5 inhibitory control was explored in more detail through a paradigm created to test children's impulsivity. Preventing children from acting impulsively did not aid tool-innovation performance. This provides further evidence that inhibitory control is not a limiting factor for children's abilities. Finally, in experiment 6 an extensive executive function battery was conducted alongside a tool-innovation task. None of the executive measures were found to correlate with tool-innovation performance. Taken together these studies suggest that tool-innovation difficulty cannot be explained by a deficit in a singular executive function. Tool innovation is not simply an inhibition task or a working memory task. Tool innovation requires multiple executive functions and none of those tested have been found to be a limiting factor on children's performance.

The main focus of research investigating ill-structured problem solving and executive function comes from the study of frontal lobe patients. Deficits in ill-structured problem solving have been found in patients who perform at an otherwise normal level on traditional executive function measures (Shallice & Burgess, 1991; Goel, Pullara & Grafman, 2001). Recent advances have discovered that patients showing this divergence in abilities have damage to an area of the frontal lobe known as Brodmann Area 10 (Burgess, Dumontheil & Gilbert, 2007). Brodmann Area 10 has been demonstrated to have protracted maturation between the ages of 5 and 11 (Dumontheil, Burgess & Blakemore, 2008). This taken together with evidence of improved tool-innovation performance between these ages further adds to support for the suggestion that tool innovation is an ill-structured problem.

To date there is very little other research on children's ill-structured problem solving. Studies that do investigate children's ill-structured problem solving abilities differ to tool-innovation tasks in the way they are ill-structured. Tool innovation is ill-structured because there is a lack of information regarding the transformation required to move from the start state to the goal state. In contrast tasks such as the six parts test (Emslie, Wilson, Burden, Nimmo-Smith, Wilson, 2003) are ill-structured due to the goal state being ill-defined. It is possible that no relationship was found between hook innovation and the six parts test in chapter 6 due to this difference in the way the tasks are ill-structured. To my knowledge tool innovation is the first task for children that is ill-structured due to a lack of information regarding transformation. As such results from tool-innovation tasks could be the first to show the development of this type of ill-structured problem solving ability in children and it would be interesting to see if difficulties extend to other ill-structured domains. In particular it would be interesting to see if children's abilities in other domains requiring innovation follow the same developmental trajectory demonstrated in this thesis. Additionally the categorisation of tool-innovation tasks as ill-structured problems opens up the possibility that these tasks could be useful in the testing of Dysexecutive syndromes such as patients with frontal lobe damage and children with autism. White, Burgess and Hill (2009) tested children with autism on traditional executive function tasks and more open-ended ill-structured tasks. Similar to the findings of frontal lobe patients (Shallice & Burgess, 1991), White et al. found that autistic children performed comparably to controls on traditional executive tasks but performed significantly worse on the ill-structured measures. Tool-innovation tasks have the potential to be a new measure for testing the abilities of autistic children.

7.3 A Model for Tool Innovation

Bringing together the components discussed above I would like to suggest a model outlining the process by which children are able to innovate the solutions for tool-making problems (see Figure 7. 1). I will take the innovation of a hook tool as an example to illustrate the model. To innovate the solution of a hook tool children need to generate the knowledge of what tool is needed (a hook), the means to make that tool (bending pliable pipecleaners) and how to use the tool (enter in to tube, and pull up bucket from handle). Once children have generated these pieces of knowledge they need to coordinate them in to the required solution. To generate and coordinate all of this knowledge requires multiple executive functions in conjunction with each other. Children must use their working memory to hold in mind the goal and the relevant pieces of information as they generate them, they must inhibit irrelevant information and naïve theories, and they must switch between different solutions when their initial ideas prove to be incorrect.

Experiment 7 suggested the main difficulty for children was to generate information. The results of this experiment do not enable us to tell whether there was a difference in difficulty for generating knowledge about the tool needed or knowledge about the means to make the tool. When either of these pieces of information were highlighted individually success levels were at floor, giving no indication of a difference between them. Coordinating information was found to be easier for children than generating information. Children in the older age group in experiment 7 successfully coordinated information if all the relevant pieces were highlighted. Coordination was difficult for the younger children even if the information was highlighted suggesting that capacity for coordination may develop between 5 and 6 years old. Further work needs to be done to establish which stages of this model are the most difficult to achieve. However, at this stage I would like to suggest that to successfully innovate a tool solution children must use their executive abilities to generate and coordinate their knowledge of what tool is needed, the means to make that tool, and how that tool should be used.

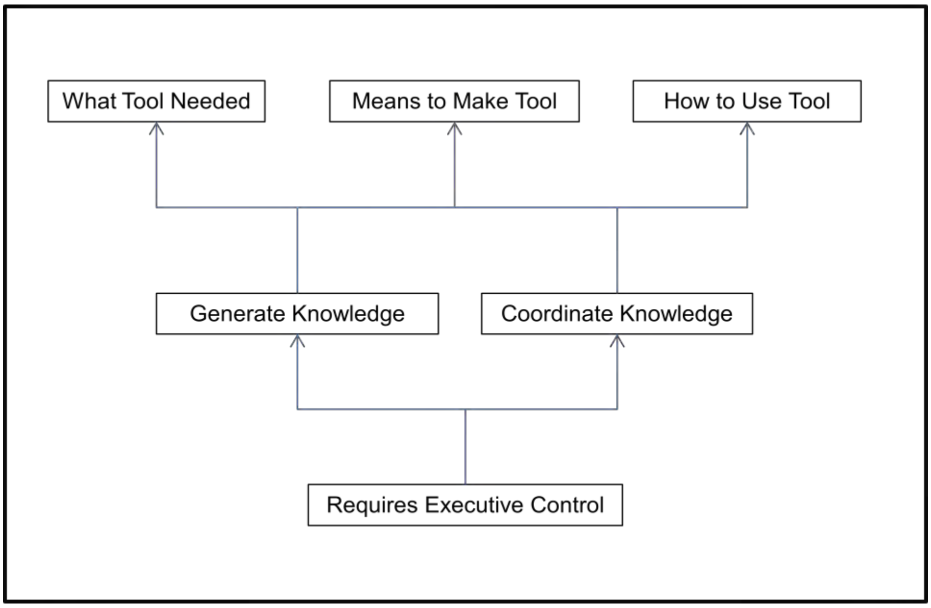

Figure 7. 1 *A model for tool innovation*

7.4 Contribution to Non-Human Animal Literature

This thesis provides the first series of studies investigating human children's tool making. Due to this fact the majority of the background literature that has informed the current work has come from the non-human animal literature. Since Goodall first reported chimpanzees using twigs to fish for termites in the 1960s there has been a huge interest in non-human animal tool use and tool making. The work in this thesis has suggested new ways of thinking about tool making and has provided new definitions and frameworks in which we can categorise and compare tool-making behaviours across species. The main contribution this thesis makes is to introduce new definitions for different levels of tool making. This is important as it enables us to directly compare different types of tool-making behaviour across a variety of species. By applying the framework of definitions set out in this thesis across different fields larger and more direct claims about different findings can be made.

This thesis proposes tool making to be split in to three different levels: innovation of the tool solution, innovation of the means to make a tool, and tool manufacture. Innovation of the tool solution refers to when one must make a tool without having seen an example of the tool they are required to make. This requires the maker to generate information about the type of tool that needs to be made, the means to make it, and how to use it to achieve their goal. Innovation of the means refers to when one has seen an example of the tool they need to make but have to generate the means to make that tool for themselves, and also know how to use that tool. Tool manufacture refers to where one makes a tool following instruction that demonstrates what tool needs to be made and how to make it. In the studies in this thesis children were required to work out the means of how to use the tool for themselves as no demonstrations of how to use the tools were given[2]. Under this current definition a demonstration showing how to use the tool would also be termed tool manufacture. Although not addressed in this current work, further consideration needs to take place to determine whether these two types of tool manufacture should be separated in to two separate terms to generate even finer definitions of tool-making behaviour.

Using these new definitions allows us to directly compare tool-making behaviours more precisely. If we take for example the hook-making behaviour of Betty the crow (Weir, Chappell & Kacelnik, 2002) and compare it to the hook-making abilities of young children, at first it would appear that Betty was outperforming the children. However, by more carefully defining different levels of tool making it is clear that these studies are not directly comparable. In the child studies no information was given about the tool required or the means to make the tool, as such these children were required to innovate the *solution* of a hook tool. In contrast in the corvid study Betty had had previous experience with a hook tool, and one was actually present at the time of testing (although her partner Abel had flown off with it!). This means that Betty had information about the tool she needed

[2] Exceptions to this can be found in chapter 4 where verbal prompts about how to use the made tools were given to children.

but not information about how to make one, as such Betty innovated the *means* to make a hook tool.

To my knowledge there are no studies investigating non-human animals' ability to innovate the solution in tool-making studies. Under the definitions laid out in this thesis tool-making studies with animals focus on the ability to innovate the means or investigate abilities for tool manufacture. In many studies animals first engage in a familiarisation phase where they get to interact with the apparatus and a suitable tool before being given unsuitable 'tools' they need to modify. For example, Visalberghi, Fragaszy and Savage-Rumbaugh (1995) first gave capuchin monkeys and great apes a straight stick tool with which to retrieve a reward from a tube before giving them the bundle and H-shape stick conditions which required modification. As the primates had seen an example of the tool required, this task tests ability for innovating the means to make a tool. In Bania et al.'s (2009) study investigating chimpanzees' abilities for subtraction and adding the experimenter gave the chimpanzee a demonstration of how to assemble or disassemble a tool before each trial. As such this study tested the ability of chimpanzees to manufacture tools.

As these examples demonstrate, clearer definitions of different tool making levels enable us to make more precise comparisons of tool-making behaviour. They also highlight gaps in the comparative literature, as there are no studies investigating abilities for innovating the solution to tool making studies.

The work in this thesis has also advanced the work in the non-human literature by making suggestions about the determinants of tool-making complexity. Studies of non-human primates have reported differing success levels within participants and species on different tool-making tasks. For example, Bania et al. (2009) found that chimpanzees were much more successful at subtracting than they were at adding, and Visalberghi et al. (1995) found great apes to be better on tool-making tasks than capuchin monkeys.

However, researchers do not address why these differences might be seen. At present researchers investigating non-human animal tool making have primarily focused on cataloguing tool-making behaviour (B. Beck, 1980; Schumaker, Walkup & Beck, 2011). This is undoubtedly a crucial first stage of investigation, but I would argue that as well as understanding who can do what, we also need to understand why they can do it and why some tool-making behaviours are easier than others. This thesis makes the first steps in understanding the complexity related to different aspects of tool-making behaviour. An understanding of tool-making complexity will allow us to predict different species' capacities in the domain of tools, which in turn can help to inform us about the evolution of our own tool-rich culture.

7.5 Social Learning

The primary aim of this thesis was to investigate children's tool innovation, however as this ability was demonstrated to be poor these studies have produced a rich data set concerning children's observational learning of tool making. Experiments in this thesis provide evidence for children's abilities to make tools via imitation and emulation. Imitation occurs when a model's goal is recognised and reproduced using the same specific actions used by the model (Tomasello, 1990). Children imitate the production of a tool in the current studies when they manufacture a tool following the action demonstration. Emulation occurs when an observer recognises and reproduces a model's goals but does not use the specific actions demonstrated by the model (Tomasello). In the current studies endstate emulation is observed when children make a tool following an endstate-tool demonstration. Children do not see the actions required to make the tool but can emulate the tool solution by innovating the means to make that tool for themselves.

7.5.1 Imitation of tool making

First, it is important to note that when children were given the opportunity for observational learning in the current studies they only received demonstrations of how to make the required tool. None of the studies in this thesis included a demonstration of how to use the tool. These are the first studies to investigate the social learning of tool *making* all other studies on children's tool behaviour have focused on the learning of tool *use.* In line with previous results from Beck et al. (2011) children in experiments 1 and 2 demonstrated great aptitude for learning about tool making from others. Following an action demonstration where children were shown what tool to make and how to make it success levels were near ceiling. These results add to evidence that children are excellent social learners for tool-related behaviours. Previously, children have been shown to imitate tool *use* from the age of two (Nielsen, 2006). This thesis shows that children readily imitate tool-making behaviours, with evidence from Beck et al. (2011) demonstrating this in children as young as 3 years old. It is possible that younger children may also be able to imitate tool-making behaviours but as yet we do not have evidence to support this. A potential difficulty in testing younger children's abilities is that although they may understand how to make a tool, young children may lack the dexterity to do so, making it difficult to test them. Together the findings from this thesis along with evidence from tool-using experiments support suggestions from the social learning literature that humans are adapted for learning from others and readily imitate a model's actions (Csibra & Gergely, 2009; Nielsen and Tomaselli, 2009; Whiten, McGuigan, Marshall-Pescini & Hopper 2009).

Success levels following the action demonstration from experiments 7 and 8 are at odds with the suggestion that children are excellent imitators. Children were much less successful following the action demonstration in these hook-making studies than they had been previously. One potential reason for this finding could be due to children's lack of trust in the model. Studies investigating children's trust in informants support this suggestion. Nurmsoo & Robinson (2009) found that when looking for a hidden toy in collaboration with a puppet, children were more likely to ignore the puppet's suggestion and guess for themselves if the puppet had previously been inaccurate. Similarly, 3-to-4-year-olds lost faith in an informant following inaccuracy in word and object-function learning tasks (Birch, Vauthier & Bloom, 2008). In experiments 7 and 8 children who reached the action demonstration phase had already received an endstate-tool demonstration from the experimenter. It is possible that because the information the experimenter gave them in this endstate-tool phase had proved insufficient to help them in the task, children then lost faith in the experimenter as a reliable source of information, and therefore did not to use the information the experimenter gave them in the action demonstration.

More research in to children's behaviour following different accuracies of demonstrations could provide us with more knowledge of how children learn from others and who children choose to learn from. In particular, this could guide us in understanding how tool-making methodology spreads within groups. Would children always go to the experts or group leaders? Or would they be willing to learn new tool-making skills from older children and those less skilled? Evidence to date suggests that children prefer models who are held in high esteem (Laland, 2004) and have a good track record of being accurate (Nurmsoo & Robinson, 2009). Children have also been shown to have preference for adult over child models (Jaswal & Neely, 2006; Rakoczy, Hamann, Warneken & Tomasello, 2010), although they will flexibly consider accuracy over age (Jaswal & Neely).

It would be interesting to see how children learn from their peers when there are no adults to guide their behaviour. It may be that a hierarchy of individuals based on their knowledge state emerges. These are important factors in the transmission of new skills that could help us to understand the evolution of human tool behaviour. Recent diffusion chain studies have shown that children faithfully transmit behaviours along a chain (Flynn & Whiten, 2008; Horner, Whiten, Flynn & de Waal, 2006). Diffusion chain methodology involves children being set a task, usually an artificial fruit (a puzzle box that can be opened by removing a number of defences), which they are then shown how to complete. Once the child has successfully retrieved their reward they then demonstrate the task to the next child and so on. Horner and colleagues seeded two chains of children with different methods to retrieve a reward on the same task. They found that children remained faithful to the method they observed, with fidelity at 100%.

Flynn and Whiten (2010) used open diffusion to see how behaviours spread more naturally. In this study children were first allowed to interact with an artificial fruit individually with no instructions or modelled demonstration, this was to give children time to acquire information about the apparatus. The apparatus was then placed in the classroom so that all children had access to it. Children discovered the two different methods for retrieving the reward during the individual phase, but quickly converged onto a single method during the open diffusion. This demonstrates children's ability to learn from others and how behaviours become normalised within groups. This study provides some insight into how children learn from each other and how behaviours spread. However, this task was comparably easier than the tool-innovation task, with over half of children succeeding in the initial phase. As such there were plenty of models to enable behaviours to spread quickly. It would be interesting to see how behaviours spread in a more difficult task such as tool innovation. Flynn and Whiten reported collaboration in their study and it is possible that this may be observed more when a task is harder to achieve.

7.5.2 Emulation of tool making

In experiments 3 to 8 we introduced a two-stage demonstration procedure where after initial failure on the tool-innovation phase of the task children first received a demonstration of the endstate-tool they needed to make. If children were not successful in making a tool following this demonstration they then proceeded on to the action demonstration which was conducted as in the previous experiments. As noted above the endstate-tool demonstration was designed to test whether children could innovate the means to make the shown tool for themselves. Under the social learning literature definitions this would be endstate emulation, i.e. could children emulate the goal of producing the required tool?

Results from experiments 3, 7 and 8 showed that young children (aged 4 to 6) were poor at emulating the goal and innovating the solution for themselves. Children get better at emulation with age, with 6-to-8-year-olds in experiment 6 performing well after the endstate-tool demonstration. This is in line with the suggestion of Want and Harris (2002) that children become able to emulate when they have better causal understanding of the task. Children are only able to emulate if they understand the causal mechanisms involved, if not children need to imitate the actions they have seen to achieve the goal. Looking at success levels across studies the results suggest young children to be more successful at making tools via imitation than they are by endstate emulation. This would support suggestions in the social learning literature that have shown two-and-a-half-year-olds to be more successful in completing a task following social conditions where they saw a model performing the correct action only or saw the correct action and the endstate, than when they saw the endstate only (Call, Carpenter & Tomasello, 2005). In the current tool-making studies there has not yet been direct comparison of these two types of demonstration. A study where half of the unsuccessful innovators received the endstate-tool demonstration and half received the action demonstration would help to confirm the suggestion that children are more successful at imitation than emulation and corroborate Call et al.'s results.

These studies provide the first evidence for children's aptitude for social learning in the domain of tool making. Together with the vast evidence for children's abilities to learn how to use tools from others these findings add to evidence to suggest that humans have adapted for social learning (Csibra & Gergely, 2009). This suggestion is more fully explored in the next section.

7.5.3 Evolution of Human Cumulative Culture

Human culture is unique. No other species has such a diverse culture that encompasses a huge range of behaviours including technology, language and social institutions (Boyd & Richerson, 1996; Whiten, 2011). Culture is defined as behaviour that is learned by all members of a group, is faithfully transmitted between individuals (Tomasello, Kruger and Ratner, 1993) and is not the product of genetic or environmental changes (Boyd & Richerson, 1996). Cultural variation refers to differences between groups due to differences in the behaviours that have spread within them. Cultural variation has been reported in a diverse range of species from chimpanzees (McGrew, 1992) to lizards and fish (Lefebvre & Palameta, 1988). Cultural variation can be seen between different chimpanzee populations in terms of behaviours such as tool use and grooming (see Whiten et al., 1999 for an extensive review). However, cumulative cultural evolution is very rare. Most culture produces behaviours that individuals could learn by themselves. In contrast cumulative culture evolution produces behaviours that individuals would be unable to invent individually. Boyd and Richerson argue that cumulative culture exists only in humans, songbirds and possibly chimpanzees. However in birds cumulative culture is limited to their songs (Boyd & Richerson) and Whiten (2011) suggests that examples of chimpanzee cumulative culture are rare and do not compare to behaviours seen in humans.

Recently there has been much speculation about what makes the cumulative evolution of human technological culture unique.

Tomasello et al. (1993) argued that human cumulative culture has been possible due to what they term the ‘ratchet-effect’. One generation starts off by doing something one way, which is transmitted by social learning to the next generation. This next generation faithfully copies these processes but modifies them to make them more efficient. This new process is then transmitted to the next generation who make their own modifications and so on. Over time this means significant changes are made that allow human culture to evolve. The ability of humans to achieve this ratchet effect is thought to be due to faithful transmission; as such the current focus of cumulative culture research is on differences in social learning ability between different species, most notably between humans and our nearest relatives, chimpanzees. High fidelity transmission ensures that all the required information is passed on. Humans have been demonstrated to be faithful imitators from infancy (Meltzoff, 1988) through to adulthood (McGuigan, Makinson & Whiten, 2011). The current work on children’s abilities to learn how to manufacture tools from others adds to the wealth of research in this area.

Differences between human and chimpanzee social learning are argued to be the reason for differences in our abilities for cumulative culture (Tomasello et al., 1993; Tomasello, 1999).When learning new behaviours humans focus on the body actions of the model, that is they copy the processes to achieve the goal as well as the goal itself (Tennie, Call & Tomasello, 2009). This fidelity means that rather than ‘reinventing the wheel’ each time a subject encounters a task they can use information they have gained from others as a starting point. In contrast chimpanzees focus on the physical effects or outcomes. Chimpanzees copy the product rather than the process. By failing to imitate the process needed to achieve a goal chimpanzees are at a disadvantage. Chimpanzees may be successful in creating the product but they have to innovate the means to create their goal from scratch each time, this requires greater cognitive capacity and is more time consuming.

Cumulative cultural evolution requires innovation as well as faithful transmission (Caldwell, Schillinger, Evans & Hopper, 2012; Tomasello, 1999). Successful transmission is of course important for the spread of behaviours, but behaviour is unable to evolve without innovation. There is a vast literature on the social learning and transmission of tool behaviours, but there is a gap in the literature when it comes to innovation. Innovations have been reported in other domains such as food preparation and threat displays. Japanese macaques have famously been reported to wash sweet potatoes (Kawai, 1965), and chimpanzees in from Gombe adopted a new aggressive gesture in the form of wrist shaking following an innovation from one individual (Kummer & Goodall, 1985). Innovations are difficult to study, especially ones that occur naturally as it is difficult to establish whether the behaviours observed are truly novel (Ramsey, Bastian & van Schaik, 2007).

This thesis provides the first evidence for the development of children's innovation abilities in the domain of tool making. These studies have clearly demonstrated that innovation is a difficult and late-developing ability. It has been speculated that this could be due to the ill-structured nature of tool innovation that places huge demands on executive functions. The difficulty of innovating a tool for oneself could explain why humans have a great aptitude for learning from others. New innovations are rare and as such it is vital that they are spread through social learning as individuals are unlikely to discover them for themselves. The faithful transmission of innovations ensures that new adaptions are kept and that culture is able to evolve. However, perhaps human adaptation for social learning has come at a cost to our ability to innovate. Young children are brought up in a culture where learning from others via both observational learning and teaching are the norm. It is possible that this over-reliance on others for knowledge prevents children from innovating for themselves. Until we have a greater understanding of the process by which innovations are made we will be unable to fully understand how human culture evolved. Social learning is undeniably important for the transmission of culture but without innovation evolution could not have occurred. This thesis makes the first steps in understanding the difficulties involved in innovation, and provides a solid basis on which future research can progress.

7.6 Future Research

Children's tool making is a new and exciting area of investigation. As such there are many possible avenues for future study. In the various sections above I have made some suggestions for future research. I will summarise the major suggestions in this section alongside some additional thoughts.

This thesis suggests the complexity of tool-making problems is due to the transformation of materials that is required rather than the broader level of tool-making mode as previously suggested (Kacelnik et al., 2006). I have speculated that some of the differences in success levels seen between different tasks may be due to how obvious the transformation required is. For example making a tool by removing a large piece of material will be much clearer to recognise than if only a small piece of material needed to be removed. At present this suggestion is merely based on observations of the different tasks, but this suggestion could be tested empirically.

Another possible reason for differences seen between different transformations could be due to 'undoing' being easier than 'doing'. For example in the unbending task a pipecleaner which would usually be presented straight is bent in half, therefore you must 'undo' what has been done to it to make the required tool. In the case of hook making the pipecleaner comes in its 'natural' form and you must 'do' something to it to make a tool. It seems plausible that 'undoing' something and putting it back to its natural state might be easier to achieve than ‘doing’ something to a raw material. There are of course more possibilities as to what one can ‘do’ with something, but limited ways in which to ‘undo’ it. All of this of course relies on the recognition that the material is not in its natural state in the case of ‘undoing’. The suggestion that ‘undoing’ might be easier than ‘doing’ is supported by anecdotal evidence that some children in the unbending task automatically straightened the pipecleaner, returning it to its ‘natural’ state, with little outward evidence that they knew they needed to do this for the task. Investigation of more tasks that involve ‘doing’ versus ‘undoing’ could test this suggestion.

Research investigating cumulative culture suggests that culture evolves due to collaboration (Tennie et al., 2009). Cumulative culture results in innovations and inventions that one individual could not arrive at by themselves (Boyd & Richerson, 1996). Based on this suggestion it would be interesting to discover whether children perform better on tool-innovation tasks if they have the opportunity for collaboration. This could take the form of a structured group activity where children are set the challenge of innovating a tool to win a reward either in pairs or small groups. Children have been shown to collaborate from a young age. Warneken, Lohse, Melis and Tomasello (2010) reported successful collaboration in problem solving tasks for children as young as 3 years. Furthermore, children appeared to recognise the contribution of their collaborator as they tended to share rewards equally. Supporting suggestions regarding cumulative culture, collaboration has been shown to create greater levels of success than participants could achieve individually in the domains of mathematics (Mullins, Rummel & Spada, 2011) and route planning (Gauvain & Rogoff, 1989). Mullins et al. suggest that when learning mathematics collaboration aids students as they are required to verbalise their thoughts and explanations which makes them more explicit. This elaboration helps to progress learning.

An alternative design would be to see how long and by what means children are able to solve a tool-innovation problem if it was left in a communal area over time so that all members of a group (e.g. a class) had full access to the problem. As well as showing us how innovations may develop this would also give opportunity to discover how new innovations spread amongst a group through social learning. Flynn and Whiten (2010) noted that the majority of studies in to the social learning of tool use use a dyadic design, where imitation of behaviours is assessed following a demonstration from a single model. Although some skills may be learnt in this way, they propose that the majority of behaviours are more likely to be learnt following multiple demonstrations from a number of models, with the learner having numerous attempts at the task themselves. Using a more naturalistic study design we will be able to discover how innovations are discovered and transmitted throughout a group. Children might imitate following direct observation, alternatively they may emulate the result if they come across someone else's premade tool. This study design would also allow us to see if children will teach each other once they have discovered a solution. Open diffusion such as this has been tested in children on tool-use experiments where the aim is to open an artificial fruit to retrieve a reward (Flynn & Whiten, 2010. See also Flynn & Whiten 2008; Whiten et al., 1996 for more structured diffusion chains using an artificial fruit). Flynn and Whiten found evidence of imitation, collaboration and teaching in their tool using study. It would be interesting to discover whether these would also be found in a more complex task requiring tool making. Learning how to use tools is important, but to discover more about human tool evolution more research needs to uncover the processes by which individuals learn how to make tools.

The current research has suggested that children should have all the required knowledge needed to solve tool-innovation problems but lack the ability to generate this knowledge in the context of the task. One way in which this suggestion could be tested is to investigate children's implicit knowledge. Research on gesture has found that before children are able to explicitly pass tasks they often demonstrate understanding of task components through their implicit gestures (Church & Goldin-Meadow, 1986). Using Piagetian conservation tasks Church and Goldin-Meadow found that many children answered the conservation questions incorrectly but made spontaneous gestures that corresponded with the correct answers. These gestures were thought to convey implicit knowledge that the child had but was unable to explicitly verbalise. On other tasks children who spontaneously produced gestures containing solution relevant information were found to be more receptive to later instruction than children that did not produce relevant solutions in their gestures (Church & Goldin-Meadow; Alibali & Goldin-Meadow, 1993; Perry, Church & Goldin-Meadow, 1988).

In more recent work Broaders, Wagner Cook, Mitchell and Goldin-Meadow (2007) actively encouraged children to gesture whilst completing maths problems. They found that children elicited implicit knowledge about the tasks through their gestures despite being unable to solve them explicitly. However the children whose gestures revealed implicit understanding were more likely to benefit from instruction in a follow-up session. This study concluded that gesture plays a causal role in learning. This argument is based on evidence that gesturing enables learners to express their implicit knowledge; in turn this implicit knowledge makes children more likely to learn from instruction.

Gestures are also argued to highlight and help structure information (Alibali, Spencer, Knox & Kita, 2011). As well as gaining insight in to which children are on the verge of solving tool-making tasks, gesture may also help children to succeed. It has been argued that gesturing whilst problem solving reduces cognitive load (Goldin-Meadow, Nusbaum, Kelly & Wagner, 2001). This has been documented in adult studies looking at gear movement problems (Alibali et al, 2011), and maths and word learning studies with children (Goldin-Meadow et al.). Gesture may be a useful medium with which to gain more knowledge about what children understand in tool-innovation problems, and also to gain understanding of methods that can be used to progress children’s abilities.

Anecdotally I have observed gestures during tool-innovation tasks that suggest children may have implicit knowledge about the tasks. In one particular example a 3-year-old girl gestured a hook shape by drawing it in the air, although she was unable to successfully make a pipecleaner hook to complete the task. Asking children to gesture whilst explaining how they intend or how they tried to solve the task following failure may help to identify children who are on the cusp of being able to solve tool-innovation problems. Gesture may also aid children in solving tasks following demonstrations. It would be interesting to see if children who implicitly portray correct solutions in their gestures are more successful following the endstate-tool demonstration than children who do not show implicit knowledge in their gestures.

7.7 Conclusions

This thesis contains the first series of studies to investigate children's tool making. In contrast to the vast literature demonstrating children's impressive abilities for tool use, this work concludes that there is a divergence in children's abilities in the domain of tool making. Children have demonstrated a great aptitude for learning how to make tools from others, a finding that fits well with social learning literature investigating tool use which suggests humans to be well adapted for learning from others. In contrast, children display great difficulty in innovating a novel tool for themselves. Children's difficulties in tool innovation have been shown to be a robust phenomenon spanning several tasks and tool-making methods. This thesis takes the first steps in discovering what determines the complexity of different tool-making episodes, suggesting that research should focus on the transformation needed. Throughout this thesis it has been argued that the difficulty of tool innovation is due to its ill-structured nature. This concluding chapter amalgamated the evidence for this suggestion drawing on different problem solving literatures to support this view. Finally, this thesis provides much needed structure to the human and non-human tool behaviour literatures by providing new definitions and frameworks with which to describe and test behaviours seen both in the wild and the laboratory.

Recently, there has been great interest in discovering why human tool behaviour is unique and how our cumulative culture has evolved. Innovation is a vital component of cumulative culture, yet it has been neglected in favour of the study of social learning. This thesis provides the first evidence for the development of innovation ability and demonstrates that innovation is a difficult and late-developing capacity. The fact that 'reinventing the wheel' for oneself is comparatively difficult could help to explain why humans have become such proficient social learners. Both innovation and social learning are clearly important for the evolution of tool culture and both have the potential to be the limiting step in this evolution. Current researchers have focused on examining the differences in social learning behaviours between human and non-human primates but more attention should focus on different species' abilities to innovate. Tool innovation is cognitively demanding and could be the key component that makes human tool culture uniquely complex.

References

Alibali, M. W., & Goldin-Meadow, S. (1993). Transitions in learning: What the hands reveal about a child's state of mind. *Cognitive Psychology*, 25, 468-523.

Alibali, M. W., Spencer, R. C., Knox, L., & Kita, S. (2011). Spontaneous Gestures Influence Strategy Choices in Problem Solving. *Psychological Science, 22*(9), 1138-1144.

Alloway, T. P., Gathercole, S. E., Willis, C. S., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology,* 87, 85-170.

Allport, A., & Wylie, G. (2000). Task-switching, stimulus–response bindings, and negative priming. In: Monsell, S., & Driver, J. S. (Eds.). *Control of Cognitive Processes: Attention and Performance XVIII*. MIT Press, Cambridge, MA.

Anderson, V. (1998). Assessing executive functions in children: Biological, psychological and developmental considerations. *Neuropsychological Rehabilitation*, 8, 319–349.

Bania, A. E., Harris, S., Kinsley, H. R., & Boysen, S. T. (2009). Constructive and deconstructive tool modification by chimpanzees (Pan Troglodytes). *Animal Cognition, 12,* 85-95.

Barr, R., & Wyss, N. (2008). Re-enactment of televised content by 2-year-olds: Toddlers use language learned from television to solve a difficult imitation problem. *Infant Behavior and Development,* 31, 696-703.

Barrett, T., Davis, E., & Needham, A. (2007). Learning about tools in infancy. *Developmental Psychology,* 43, 352-368.

Bates, E., Carlson-Luden, V., & Bretherton, I. (1980). Perceptual aspects of tool using in infancy. *Infant Behavior and Development*, 3, 127-140.

Beck, B. B. (1980). *Animal tool behaviour: the use and manufacture of tools by animals.* New York; London: Garland STPM.

Beck, S. R., Carroll, D. J., Brunsdon, V. E. A., & Gryg, C. K. (2011). Supporting children's counterfactual thinking with alternative modes of responding. *Journal of Experimental Child Psychology,* 108*,* 190-202.

Beck, S. R., Apperly, I. A., Chappell, J., Guthrie, C., & Cutting, N. (2011). Making tools isn't child's play. *Cognition*, 119, 301–306.

Beck, S. R., Chappell, J., Apperly, I. A., & Cutting, N. (2012). Tool innovation may be a critical limiting step for the establishment of a rich tool-using culture: a perspective from child development. *Behavioral and Brain Sciences,* 35 (4). Commentary on Vaesen (2012).

Birch, S. A. J., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*, 107, 1018-1034.

Bird, C. D., & Emery, N. J. (2009). Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proceedings of the National Academy of Science of the United States of America*, 106, 10370–10375.

Bloom, P. (1996). Intention, history, and artefact concepts. *Cognition*, 60, 1-29.

Boesch, C., & Boesch, H. (1990). Tool use and tool making in wild chimpanzees. *Folia Primatologica,* 54, 86-99.

Bonawitz, E. B., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L.E. (2011). The double-edged sword of pedagogy: Teaching limits children's spontaneous exploration and discovery. *Cognition,* 120(3)*,* 322-330.

Boyd, R., & Richerson, P. J. (1996). Why Culture is Common but Cultural Evolution is Rare. *Proceedings of the British Academy*, 88, 73–93.

Broaders, S., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General*, 136, 539–550.

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology*, 26(2), 571–593.

Brown, A. L. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science,* 14, 107-133.

Brugger, A., Lariviere, L. A., Mumme, D. L., & Bushnell, E. W. (2007). Doing the right thing: Infants' selection of actions to imitate from observed event sequences. *Child Development*, 78, 806-824.

Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Sciences*, 11 (7), 290-298.

Burgess, P. W., Alderman, N., Emslie, H., Evans, J. J., Wilson, B. A., & Shallice, T. (1996). The simplified six element test. In Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (Eds.) *Behavioural Assessment of the Dysexecutive Syndrome*. Bury St. Edmunds, UK: Thames Valley Test

Burns, P., Riggs, K. J., & Beck, S. R. (2012). Executive control and the experience of regret. *Journal of Experimental Child Psychology,* 111*,* 501-515.

Caldwell, C. A., Schillinger, K., Evans, C. L., & Hopper, L. M. (2012). End state copying by humans (*Homo sapiens*): Implications for a comparative perspective on cumulative culture. *Journal of Comparative Psychology*, 126 (2), 161-169.

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*,* 151-163.

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

Casler, K., & Kelemen, D. (2005). Young children's rapid learning about artifacts. *Developmental Science,* 8, 472-480.

Casler, K., Terziyan, T., & Greene, K. (2009). Toddlers view artifact function normatively. *Cognitive Development,* 24*,* 240-247.

Chaigneau, S. E., Castillo, R. D., & Martinez, L. (2008). Creators' intentions bias judgments of function independently from causal inferences. *Cognition,* 109*,* 123-132.

Chen, C., & Bradshaw, A. C., (2007). The Effect of Web-Based Question Prompts on Scaffolding Knowledge Integration and Ill-Structured Problem Solving. *Journal of Research on technology in Education*, 35 (4), 359-375.

Chevalier, N., & Blaye, A. (2009) Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. *Developmental Psychology,* 45, 782-797.

Church, R. B., & Goldin-Meadow, S. (1986). The Mismatch Between Gesture and Speech as an Index of Transitional Knowledge. *Cognition*, 23, 43-71.

Connolly, K., & Dalgleish, M. (1989). The emergence of a tool-using skill in infancy. *Developmental Psychology,* 25, 894–912.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences,* 13*,* 148-153.

Cutting, N., Apperly, I. A., & Beck, S. R. (2011). Why do children lack the flexibility to innovate tools? *Journal of Experimental Child Psychology*, 109, 497–511.

Cutting, N., Beck, S.R., & Apperly, I.A. (under submission). Is There a Complexity Hierarchy in Human Children’s Tool Making?

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44, 2037–2078.

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

Defeyter, M. A., Hearing, J., & German, T. C. (2009). A developmental dissociation between category and function judgments about novel artifacts. *Cognition,* 110*,* 260-264.

Dennett, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Diamond, A. (2006). The early development of executive functions. In Bialystok, E., & Craik, F. (Eds.), *Lifespan Cognition: Mechanisms of Change, pp. 70-95.* NY: Oxford University Press.

Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology*, 38, 352–362.

Dumontheil, I., Burgess, P.W., & Blakemore, S.J. (2008). Development of rostral prefrontal cortex and cognitive and behavioral disorders. *Developmental Medicine and Child Neurology***,** 50 (3), 168-181.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, 58, 270.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scale* (2nd Ed.). London: NFER-Nelson.

Eslinger, P. J., & Damasio, A. R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology,* 35, 1731-1741.

Emslie, H. C., Wilson, F. C., Burden, V., Nimmo-Smith, I., Wilson, B. A. (2003) *Behavioural Assessment of the Dysexecutive Syndrome in Children (BADS-C).* Thames Valley Test Company, Bury St Edmunds.

Espy, K. (1997). The shape school: Assessing executive function in preschool children. *Developmental Neuropsychology*, 13, 495–499.

Flynn, E. G., & Whiten, A. (2008). Imitation of hierarchical structure versus component details of complex actions by 3- and 5-year-olds. *Journal of Experimental Child Psychology,* 101(4), 228-240.

Flynn, E. G., & Whiten, A. (2010). Studying children's social learning experimentally "in the wild". *Learning and Behavior,* 38(3), 284-296.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development,* 10, 483-527.

Gardiner, A., Bjorklund, D., Greif, M. L., & Gray, S. (2012). Choosing and using tools: Prior experience and task difficulty influence preschoolers tool use strategies. *Cognitive Development*, 27 (3), 240-254.

Gathercole, S. E., & Pickering, S. J. (2000). Assessment of working memory in six- and seven-year old children. *Journal of Educational Psychology,* 92, 377-390

Gauvain, M., & Rogoff, B. (1989). Collaborative problem solving and the development of children's planning skills. *Developmental Psychology,* 25, 139-151.

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

German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review,* 7, 707–712.)

German, T. P., & Johnson, S. (2002). Function and the origins of the design stance. *Journal of Cognition & Development*, 3, 279–300.

Gibson, K. R., & Ingold, T. (1993). *Tools, language and cognition in human evolution.* Cambridge: Cambridge University Press.

Goel, V. (2010). Neural Basis of Thinking: Lab Problems vs. Real-World Problems. *Wiley Interdisciplinary Reviews: Cognitive Science*, Vol. 1, pp. 613-621.

Goel, V., Pullara, D., & Grafman, J. (2001). A Computational Model of Frontal Lobe Dysfunction: Working Memory and the Tower of Hanoi. *Cognitive Science*, 25 (2), 287-313.

Goel, V. (1995). *Sketches of thought*. MIT Press; Cambridge, MA.

Goel, V., & Grafman, J. (2000). Role of the right prefrontal cortex in ill-structured planning. *Cognitive Neuropsychology*, 17, 415-436.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science,* 12(6), 516-522.

Hart, B. L., & Hart, L. A. (1994). Fly switching by Asian elephants: tool use to control parasites. *Animal Behaviour*, 48, 35-45.

Hart, B. L., Hart, L. A., McCoy, M., & Sarath, C R. (2001). Cognitive behaviour in Asian elephants: use and modification of branches for fly switching. *Animal Behaviour*, 62, 839-847.

Hayne, H., Herbert, H., & Simcock, G. (2003). Imitation from television by 24- and 30-month-olds. *Developmental Science, 6,* 254-261.

Herbert, J., & Hayne, H. (2000). Memory retrieval by 18- to 30-month-olds: Age-related changes in representational flexibility. *Developmental Psychology*, 36(4), 473-484.

Hernik, M. & Csibra, G. (2009). Functional understanding facilitates learning about tools in human children. *Current Opinion in Neurobiology,* 19, 34-38.

Herrmann, E., Call, J., Hernández-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, 317, 1360–1366.

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

Hood, B. M. (1998). Gravity does rule for falling events. *Developmental Science, 1,* 59-63.

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

Horner V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children. *Animal Cognition*, 8,164–181

Horner, V., Whiten, A., Flynn, E. G., & de Waal, F. B. M. (2006). Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. *Proceedings of the National Academy of Science,* 103, 13878–13883.

Hunt G. R., & Gray R. D. (2003). Diversification and cumulative evolution in New Caledonian crow tool manufacture. *Proceedings of the Royal Society B*, 270, 867–874.

Hunt, G. R., & Gray, R. D. (2004). Direct observations of pandanus-tool manufacture and use by a New Caledonian crow (Corvus moneduloides). *Animal Cognition*, 7, 114–120.

Hunt, G. R. (1996). Manufacture and use of hook-tools by New Caledonian crows. *Nature,* 379, 249–251.

Hunt, G. R., & Gray, R. D. (2002). Species-wide manufacture of stick type tools by New Caledonian crows. *Emu,* 102, 349–353.

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science*, 17, 757-758.

Jaswal, V. K. (2006). Preschoolers favor the creator's label when reasoning about an artifact's function. *Cognition,* 99, B83-B92.

Johnson-Frey, S.H., Newman-Norlund, R.D., & Grafton, S.T. (2005). A distributed network in the left cerebral hemisphere for planning everyday tool use skills. *Cerebral Cortex*, 15, 681-695.

Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Jonides, J., and Smith, E. E. (1997). The architecture of working memory. In Rugg, M. D. (Ed.), *Cognitive Neuroscience*, Sussex, England: Psychology Press. pp. 243-276.

Kacelnik, A. (2009). Tools for thought or thoughts for tools? *Proceedings of the National Academy of Sciences USA,* 106, 10071-10072.

Kacelnik, A., Chappell, J., Weir, A.A.S., & Kenward, B. (2006). Cognitive adaptations for tool-related behaviour in New Caledonian Crows. In: *Comparative cognition: experimental explorations of animal intelligence* (eds. Wasserman, E.A., & Zentall, T.R.), pp. 515-528. Oxford: Oxford University Press.

Kaplan, G., &, Rogers, L. (1994). *Orangutans in Borneo*. Armidale, Australia: University of New England Press.

Karmiloff-Smith, A., & Inhelder, B. (1975). If you want to get ahead, get a theory. *Cognition*, 3 (3), 195-212.

Kawai, M. (1965). Newly-acquired pre-cultural behavior of the natural troop of Japanese monkeys on Koshima islet. *Primates*, 6 (1), 1-30.

Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition*, 70, 241–272.

Kemler Nelson, D. G., Egan, L. C., &, Holt, M. (2004). When children ask "What is it?" What do they want to know about artifacts? *Psychological Science*, 15, 384–389

Killen, M., & Uzgiris, I. (1981). Imitation of actions with objects: The role of social meaning. *Journal of Genetic Psychology*, 138, 219 – 229.

Kummer, H., & Goodall, J. (1985). Conditions of innovative behavior in primates. *Philosophical Transactions of the Royal Society London B,* 308, 203-214.

Laland K. N. (2004). Social learning strategies. *Learning Behaviour*, 32, 4–14

Lane, F. (1957). *Kingdom of the octopus*. London: Jarrold.

Lehto, J. (1996). Are Executive Function Tests Dependent on Working Memory Capacity? *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 49 (1), 29-50.

Leslie, A. M., &, Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition,* 25, 265- 288

Lefebvre, L., & Palameta, B. (1988). Mechanisms, ecology and diffusion of socially learned, food-finding behaviour in feral pigeons. In Zentall, T., & Galef, B. G. (Eds.). *Social Learning: Psychological and Biological Perspectives*, pp. 141–163. Hillsdale, New Jersey: L. Erlbaum.

Luna, B., Thulborn, K. R., Munoz, D. P., Merriam, E. P., Garver, K. E., Minshew, N. J., Keshavan, M. S., Genovese, C. R., Eddy, W. F., & Sweeney, J. A. (2001). Maturation of widely distributed brain function subserves cognitive development. *NeuroImage*, 13(5), 786-793.

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, 104, 19751-19756

McCloskey, M. (1983). Intuitive physics. *Scientific American,* 248(4), 122-130.

McGrew, W. C. (1992). *Chimpanzee material culture: implications for human evolution*. New York/Cambridge, UK: 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, 1-18.

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, 367-381.

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, 353-364.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, memory, and cognition,* 22, 1423-1442.

Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, 24 (4), 470-476.

Meltzoff, A. N. (1995). Understanding the Intentions of Others - Re-enactment of Intended Acts by 18-Month-Old Children. *Developmental Psychology,* 31, 838-850.

Merlen, G., &, Davis-Merlen, G. (2000). Whish: More than a tool-using finch. *Noticias de Galapagos*, 61, 2-9.

Mesulam, M. M. (1986). Frontal cortex and behavior. *Annals of Neurology, 19*(4), 320-325.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contribution to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology,* 41, 49–100.

Monsell, S. (1996) Control of mental processes. In Bruce, V. (Ed.). *Unsolved Mysteries of the Mind: Tutorial Essays in Cognition,* pp. 93–148, Erlbaum.

Morris, N., & Jones, D. M. (1990). Memory updating and working memory: The role of the central executive. *British Journal of Psychology*, 81, 111-121.

Mounoud, P. (1996). A recursive transformation of central cognitive mechanisms: the shift from partial to whole representation. In Sameroff, A. J., & Haith, M. M. (Eds.). *The five to seven year shift: the age of reason and responsibility*, pp. 85–110, Chicago: Chicago University Press.

Mullins, D., Rummel, N., & Spada, H. (2011) Are two heads always better than one? Differential effects of collaboration on students' computer-supported learning in mathematics. *International Journal of Computer-Supported Collaborative Learning*, 6 (3), 421-443.

Nagell, K., Olguin, R., & Tomasello, M. (1993) Processes of social learning in the tool use of chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Journal of Comparative Psychology*, 107,174–186.

Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology*, 42, 555–565.

Nielsen, M., & Tomaselli, K. (2009). Over-imitation in the Kalahari desert and the origins of human cultural cognition. *Nature Precedings,* 713.

Nurmsoo, E., & Robinson, E.J. (2009) Children's trust in previously inaccurate informants who were well or poorly-informed: When past errors can be excused. *Child Development, 80,* 23-27.

Oakes, L., & Cohen, L. B. (1990). Infant perception of causal events. *Cognitive Development,* 5,193–207.

Ottoni, E. B., & Mannu, M. (2001). Semifree-ranging tufted capuchins (*Cebus apella*) spontaneously use tools to crack open nuts. *International Journal of Primatology*, 22, 347-358.

Owen, J. (2005). Nut-cracking ape may boost gorillas' IQ rep. *National Geographic News*, October 25. http://news.nationalgeographic.com/news/.

Parish, A. R. (1994). Sex and food control in the 'uncommon chimpanzee': How bonobo females overcome a phylogenetic legacy of male dominance. *Ethology and Sociobiology*, 15 (3), 157-179.

Penfield, W., & Evans, J. (1935). The frontal lobe in man: a clinical study of maximum removals. *Brain*, 58, 115–33.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*(4), 359-400.

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

Prince, E., Lambeth, S., Schapiro, S., & 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.

Rakoczy, H., Hamann, K., Warneken, F., & Tomasello, M. (2010). Bigger knows better – young children selectively learn rule games from adults rather than from peers. *British Journal of Developmental Psychology,* (28), 785-798.

Ramsey, G., Bastian, M. L., &, van Schaik, C. (2007). Animal innovation defined and Operationalized. *Behavioral and Brain Sciences*, 30, 393-437.

Reitman, W. (1965). *Cognition and thought*. New York: Wiley

Reynolds, B., Ortengren, A., Richards, J. B., & de Wit, H. (2006). Dimensions of impulsive behaviour: Personality and behavioral measures. *Personality and Individual Differences,* 40, 305-315.

Ross, D. (1971). Protection of hermit crabs (*Dardanus* spp.) from octopus by commensal sea anemones (*calliactis* spp.). *Nature,* 230, 401-402.

Schachar, R. J., & Logan, G. D. (1990). Impulsivity and inhibitory control in normal development and childhood psychopathology. *Developmental Psychology,* 26, 710-720

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

Seed, A. M., 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, 23-34.

Shallice, T., & Burgess, P. W. (1991). Higher-order cognitive impairments and frontal lobe lesions in man. In Levin, H. S., Eisenberg, H. M., & Benton, A. L. (Eds.) *Frontal lobe function and dysfunction* (pp. 125–138). New York: Oxford University Press.

Silva, F. J., Page, D. M., & Silva, K. M. (2005). Methodological conceptual problems on the study of chimpanzees' folk physics: How studies with adult humans can help. *Learning and Behaviour, 33,* 47–58.

Simpson, A., & Riggs, K.J. (2011) Under what conditions do children have difficulty inhibiting imitation? Evidence for the importance of planning specific responses. *Journal of Experimental Child Psychology,* 109, 512-524.

Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A.W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience*, 24, 8223–8231.

Spence, K. W. (1937). Experimental studies of learning and higher mental processes in infrahuman primates. *Psychological Bulletin,* 34, 806–850.

St. Amant, R., & Horton, T. (2008). Revisiting the definition of animal tool use. *Animal Behaviour*, 75, 1199-1208.

Sugiyama, Y. (1985). The brush-stick of chimpanzees found in south-west Cameroon and their cultural characteristics. *Primates*, 26 (4), 361-374.

Sugiyama, Y. (1989). Description of some characteristic behaviours and discussion of their propagation process among chimpanzees of Bossou, Guinea. In Sugiyama, Y. (Ed.). *Behavioural studies of wild chimpanzees at Bossou, Guinea*. Inuyama, Japan: Kyoto University Primate Research Institute, 43-76.

Surtees, A. D. R., Burns, P., Beck, S. R., Riggs, K., & Apperly, I. A. (under review) The role of executive functions in avoiding egocentrism in perspective-taking.

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions of the Royal Society of London B Biological Sciences,* 364, 2405-2415.

Thorpe, W. H. (1956*). Learning and Instinct in Animals*. Methuen, London.

Tomasello, M. (1990). Cultural transmission in the tool use and communicatory signaling of chimpanzees. In Parker, S. T., & Gibson, K. R. (Eds.). *Language and intelligence in monkeys and apes: comparative developmental perspectives*, pp. 274–311. Cambridge, UK: Cambridge University Press.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press.

Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences,* 16, 495–552.

Uzgiris, I. C. (1981). Two functions of imitation during infancy. *International Journal of Behavioral Development*, 4, 1–12.

Vaesen, K. (2012). The cognitive bases of human tool use. *Behavioral and Brain Sciences,* 35(4), 203-218.

van Schaik, C. P., &, Fox, E. A. (1994). Tool use in wild Sumatran orang-utans (*pongo pygmaeus*). *XVth Congress of the International Primatological Society, August 3-8, Bali, Indonesia*. Indonesia: Indonesian Wildlife Society, 339.

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., Suzuki, A., Utami, S. S., & Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science*, 299, 102-105.

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, 52-60.

Voss, J. E., Blais, J., Means, M. L., & Greene, T. R. (1986). Informal reasoning and subject matter knowledge in the solving of economics problems by naive and novice individuals. *Cognition and Instruction,* 3, 269-302.

Walraven, V., van Elsacker, L., & Verheyen, R. (1995). Reactions of a group of pygmy chimpanzees (Pan paniscus) to their mirror-images: Evidence of self-recognition. *Primates*, 36, 145-150.

Want, S. C., & Harris, P. L. (2002). How do children ape? Applying concepts from the study of non-human primates to the developmental study of 'imitation' in children. *Developmental Science,* 5, 1-13.

Warneken, F., Lohse, K., Melis, A. P., &, Tomasello, M. (2010). Young Children Share the Spoils After Collaboration. *Psychological science*, 22, 2267-2273.

Weir, A.A.S., & Kacelnik, A. (2006). A New Caledonian crow (Corvus moneduloides) creatively re-designs tools by bending or unbending aluminium strips. *Animal Cognition,* 9, 317-334.

Weir, A. A. S., Chappell, J., & Kacelnik, A. (2002). Shaping of hooks in new Caledonian crows. *Science, 297,* 981

White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on "open-ended" executive function tests in autism. *Autism Research*, 2(3), 138-147.

Whiten, A., Goodall, J., McGrew, W.C., Nishida, T., Reynolds, V. , Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W., & Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399, 682-685.

Whiten, A., & Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: reappraisal of a century of research. *Advances in the Study of Behavior*, 21, 239–283.

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, 3-14.

Whiten, A. (2011). The scope of culture in chimpanzees, humans and ancestral apes. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 366, 997-1007.

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

Willatts, P. (1984). Stages in the development of intentional search by young infants. *Developmental Psychology*, 20, 389-396.

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life-span. *Developmental Psychology*, 35, 205–213.

Wineburg, S. (1998). Reading Abraham Lincoln: An expert-expert study in the interpretation of historical texts. *Cognitive Science,* 22, 319-34.

Zelazo, P. D., & Muller, U. (2002). Executive function in typical and atypical development. In Goswami, U. (Ed.) *Blackwell handbook of childhood cognitive development*.